

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/360963167>

Le choix des approches évaluatives

Chapter · May 2022

CITATIONS

0

READS

80

1 author:



Thomas Delahais

Quadrant Conseil

22 PUBLICATIONS 100 CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:



Evaluation of Legislation [View project](#)



Métiers de l'évaluation en France [View project](#)

Le choix des approches évaluatives

Thomas Delahais



Introduction

Le mot « approche » fait partie de ces termes évaluatifs employés à tout bout de champ, mais dont la définition est difficile. Après tout, qu'y a-t-il de commun entre l'expérimentation par assignation aléatoire et l'évaluation féministe, mis à part qu'on les qualifie toutes les deux d'approches évaluatives ? Cependant, les approches sont des éléments utiles pour appréhender la façon de faire des évaluations, ancrées dans la théorie comme dans la pratique, et qui permettent de mieux se comprendre. Il est donc utile de s'en servir comme « blocs de construction » d'une évaluation (section 1). Par où commencer ? Nous proposons ici de considérer l'éventail des approches possibles en partant de leur point d'entrée : le choix des méthodes; l'interrogation des valeurs; l'utilité des travaux; l'exigence de justice sociale (section 2). Nous terminerons enfin par quelques mots visant à faciliter aux lectrices et aux lecteurs le choix d'une approche appropriée (section 3).

Qu'est-ce qu'une approche évaluative ?

Une façon de définir les approches évaluatives est de les voir comme des modèles, comme un « ensemble de règles, de prescriptions et d'interventions, et de cadres qui guide [l'évaluateur(-trice)] et qui spécifie ce qu'est une bonne évaluation, ou une évaluation adéquate, et comment elle devrait être menée » (Alkin, 2013, p. 4). Ces approches peuvent être propres à l'évaluation ou bien être importées de la recherche, ou encore de pratiques voisines à l'évaluation.

Un entre-deux, entre théorie et pratique

Une approche est une sorte d'entre-deux entre la théorie et la pratique. D'un côté, une évaluation s'inscrit, implicitement ou explicitement, dans une certaine vision du monde, de la nature des choses, de l'éthique, c'est-à-dire dans des paradigmes évaluatifs (Dagenais et Ridde, 2012, p. 14), qui se déclinent eux-mêmes en théories évaluatives. De l'autre, elle consiste aussi à manipuler des méthodes ou des outils des sciences sociales et à mener un certain nombre d'opérations très pratiques, de collecte d'informations par exemple. Une approche ne doit être confondue ni avec les uns ni avec les autres.

Ainsi, on peut décrire une évaluation en disant qu'elle se place dans un paradigme positiviste, c'est-à-dire qu'elle s'appuie sur l'idée qu'il existe une réalité qui est indépendante des perceptions de chacun de nous; et dans une théorie de la causalité contrefactuelle, au titre de laquelle on pourra affirmer qu'une cause C produit un effet E uniquement si, en l'absence de C, E est absent.

Inversement, on pourrait présenter cette évaluation par les actes concrets qu'elle suppose : des outils ou méthodes tels que l'enquête, voire à un niveau plus micro comme un échantillonnage, la définition d'indicateurs, des analyses statistiques, des choix d'interprétation, etc.

Mais même les plus acharnés des théoriciens et théoriciennes ne se posent pas de questions relatives à la nature de la réalité à chaque évaluation, et les plus terre-à-terre des évaluateurs ou évaluatrices peuvent comprendre qu'une suite d'outils ou d'actes ne suffit pas à rendre compte de ce qu'ils ou elles veulent faire. Une approche fait le lien entre ces deux extrêmes en fournissant de nombreux éléments permettant de mener à bien une évaluation (Tableau 1)¹.

Tableau 1 — Éléments constitutifs d'une approche

Une approche fournit :	Afin de :
Un cadre paradigmatique, éthique et théorique.	Répondre à des questions relatives à la nature de la réalité et de la connaissance.
Une méthodologie (deMarrais et Lapan, 2004, p.5).	Décrire comment une évaluation devrait être menée en théorie pour qu'elle soit considérée comme étant de bonne qualité (chaque approche tendant à définir ses propres critères de qualité).
Des promesses et des situations d'usage préférentielles.	Choisir quand déployer la méthodologie, selon le contexte et les conséquences attendues de l'évaluation.
Un mode d'emploi (plus ou moins prescriptif), des éléments de posture et des concepts opératoires.	Mettre en œuvre cette méthodologie uniforme, indépendamment de la personne qui la déploie.
Une liste d'outils suggérés ou imposés, une façon de les articuler, parfois des indications particulières sur la manière de les utiliser.	Recueillir les informations nécessaires pour le bon fonctionnement de l'approche.

Les approches comme communautés de pratique

Toutes les approches n'intègrent pas la totalité de ces éléments, ou plutôt, celles et ceux qui formulent une approche ne peuvent préciser l'ensemble de ces caractéristiques en une seule fois. Le ferment initial d'une approche, c'est le plus souvent une théorie dont on envisage les applications pratiques, une méthodologie dont on explicite le cadre théorique, ou une montée en généralisation de pratiques personnelles.

Dans tous les cas, pour qu'une approche soit complète et qu'elle puisse guider les choix évaluatifs, elle doit pouvoir compter sur une « communauté de pratique » (Wenger-Trayner et Wenger-Trayner, 2015, p. 1), qui va progressivement l'étoffer. Certes, certaines personnalités en sont venues à incarner des approches – que l'on pense à Michael Q. Patton avec l'évaluation axée sur l'utilisation (2008). En réalité, cependant, l'élaboration des approches est un processus dynamique. Au sein de communautés de pratique se tiennent des débats parfois vifs sur leurs différents éléments constitutifs. Les approches sont effectivement testées sur le terrain dans une logique d'essai-erreur ou d'hybridation,

1. Cette liste s'appuie notamment sur les cinq critères proposés par Robin Lin Miller (2010) pour examiner les théories évaluatives : la spécificité opérationnelle, le champ d'application, la faisabilité en pratique, l'impact discernable et la reproductibilité.

des leçons sont tirées de contextes nouveaux, de nouvelles solutions à des problèmes repérés, etc. Il ne suffit donc pas de décrire une méthodologie ou de fournir un cadre théorique pour qu'une approche prenne vie².

Ainsi, lorsqu'une équipe d'évaluation souhaite s'engager dans une approche évaluative, par exemple une évaluation quasi expérimentale, elle peut s'appuyer à la fois sur des écrits qui dépeignent le déroulement d'une telle évaluation, mais aussi sur une série d'exemples auxquels se référer et de personnes-ressources auxquelles s'adresser. C'est le recours à ces exemples et à ces personnes qui permet à l'équipe d'utiliser concrètement une approche, en traitant de cas limites, de difficultés techniques ou éthiques, de biais repérés, etc.

Parfois, ces débats se font au sein d'une dénomination partagée. Mais très souvent, ils amènent à la constitution de familles d'approches. Ainsi, l'évaluation féministe regroupe sous son ombrelle plusieurs approches, qui ont pour point commun de s'inscrire dans une vision transformationnelle et qui se distinguent ainsi de l'approche attentive au genre (Podems, 2014). Inversement, on trouve autour de l'analyse de contribution de nombreuses propositions alternatives qui partagent des valeurs et des concepts similaires, et qui mentionnent toutes l'analyse de contribution dans leurs inspirations tout en s'en détachant³. Il peut être difficile de se retrouver dans ce foisonnement – d'où l'importance de communautés de pratiques et de personnes-ressources permettant d'y avoir accès – mais il participe à offrir la diversité de modalités qu'un évaluateur ou évaluatrice est en droit d'attendre d'une approche.

Les approches comme blocs de construction de l'évaluation

Enfin, on peut définir les approches par leurs usages, et les voir comme des éléments préfabriqués, des blocs mis à la disposition des évaluateurs et évaluatrices pour construire une évaluation. Chaque évaluation nécessite l'élaboration d'une démarche évaluative ad hoc. Par ce terme, nous désignons l'ensemble des choix faits par l'équipe d'évaluation et visant à atteindre des buts évaluatifs. Ces choix sont de natures multiples : ils sont méthodologiques bien sûr, mais relèvent aussi de l'organisation, de la posture, de la relation aux parties prenantes, etc. Or, plus l'évaluation est complexe, plus ces choix sont marqués par l'incertitude : que faudrait-il faire, que peut-on attendre de telle ou telle démarche ?

Toutes les évaluations ne se positionnent pas explicitement dans des approches spécifiques, et ce n'est pas forcément un gage de qualité d'une évaluation qu'elle se réclame d'une approche. Il est d'ailleurs à noter que les pratiques varient entre les pays et les cultures d'évaluation à ce sujet : en Europe, il est sans doute moins fréquent de se réclamer d'une approche en particulier qu'en Amérique du Nord.

Cependant, la mention d'une ou de plusieurs approches évaluatives a plusieurs avantages :

- Les approches servent à ouvrir les choix possibles pour une évaluation, tout en offrant des cadres robustes et prévisibles. Avoir accès à un large répertoire d'approches est une occasion de sortir de ses habitudes ou de sa zone de confort;

2. Egon Guba et Yvonne Lincoln (1989), par exemple, ont proposé une théorie constructiviste de l'évaluation et ont décrit dans le détail ce que pourrait être une évaluation « de 4^e génération ». Si la théorie s'est révélée très influente, en réalité elle n'a presque jamais été mise en pratique (Lay & Papadopoulos, 2007), et d'autres approches ou démarches s'inscrivant pour tout ou partie dans la théorie constructiviste ont prospéré.

3. On peut citer par exemple le Contribution tracing (Befani et Mayne, 2014), le Qualitative Impact Assessment Protocol (Copestake *et al.*, 2019), la méthode Impress (Barret *et al.*, 2017) ou l'Outcome Evaluation approach (Belcher, 2020).

- Les approches permettent de se placer dans une vision du monde, de ce qui est bien, de ce qui compte : expliciter ces valeurs permet de voir en quoi elles correspondent à ses valeurs personnelles ou à celles de l'organisation pour laquelle on travaille;
- Les approches facilitent la compréhension mutuelle entre les évaluateurs et évaluatrices, commanditaires, et autres parties prenantes de l'évaluation. Un choix commun d'approches est une assurance que certaines valeurs sont communes⁴ – du moins si la connaissance des approches est partagée. Parler d'approches est rassurant (il existe d'autres exemples de ce type d'évaluation, qui ont été menés à bien) et, in fine, contribue à créer un climat de confiance utile à l'évaluation.

L'éventail des approches

Note : dans cette partie, les références en français permettant d'en savoir plus sur les différentes approches présentées sont soulignées.

Dans tout l'éventail des possibles, nous avons privilégié les approches qui disposent aujourd'hui de communautés de pratiques actives, car ce sont les plus susceptibles de fournir aux équipes d'évaluation qui souhaiteraient les mettre en œuvre les ressources nécessaires pour le faire. Parmi ces approches, nous avons accordé une certaine importance à celles traitant des questions causales, car l'impact des interventions est au cœur de nombreux débats dans le monde de l'évaluation. Au-delà des choix méthodologiques⁵, l'impact est aussi une question d'approche, et c'est ce que nous illustrerons ici.

Toutes ces approches sont complexes et visent généralement à aborder différents aspects de la démarche évaluative. Il est donc difficile de les catégoriser sans tomber dans la superficialité. Nous avons choisi de recourir à quatre ingrédients de toute évaluation⁶ : les méthodes, l'utilisation, les valeurs et l'attention à la justice sociale. Nous considérons que toute évaluation doit, dans des proportions qui varient selon les cas, s'inscrire dans ces quatre dimensions pour mériter son nom. Néanmoins, ces quatre branches correspondent à des « portes d'entrée » dans l'évaluation, qui nous semblent pertinentes pour naviguer dans les approches.

Dans le cadre de ce chapitre, la description des approches est nécessairement trop brève pour rendre compte de toutes leurs facettes. Nous avons voulu mettre en exergue leurs apports, leurs valeurs sous-jacentes et les contextes dans lesquels elles brillent particulièrement. Lorsque c'est possible, des ressources en français sont suggérées pour aller plus loin.

Les approches orientées méthode

Les approches qui sont réunies dans cette partie accordent une importance particulière à la démarche d'investigation menée dans l'évaluation, ainsi qu'à la production de connaissances. C'est notamment ici que l'on retrouve les approches d'évaluation d'impact, qui visent à établir un lien de causalité entre

4. C'est d'autant plus important si, comme le soutient MQ Patton, « les gens recrutent probablement des évaluateurs parce qu'ils partagent des valeurs communes plutôt que pour les outils qu'ils proposent. » (cité dans Williams, 2018, p. 142).

5. Pour des listes de méthodes, voir Befani (2020) et Vaessen *et al.* (2020). Pour un aperçu des approches et méthodes en français, voir Quadrant Conseil (2017).

6. Nous reprenons ici la définition de l'évaluation en trois composantes proposées par Demarteau (2002, p. 457) en la complétant avec la préoccupation de justice sociale (Mertens et Wilson, 2012, p. 41).

une intervention et des changements observés. Nous distinguons dans un premier temps des approches qui se placent dans une logique expérimentale, où l'impact est la différence entre la situation avec et sans l'intervention, et qui sont d'abord préoccupées par la description de l'impact (Shadish *et al.*, 2002), et des approches dites basées sur la théorie, et en particulier les plus récentes, qui reconnaissent la complexité des contextes sociaux et qui cherchent en premier lieu à expliquer les conditions dans lesquelles les interventions ont des effets.

Encadré 1 – Approches orientées méthode

- Les évaluations d'impacts
 - Le devis expérimental (l'expérimentation par assignation aléatoire)
 - Les devis quasi expérimentaux

- Les approches basées sur la théorie
 - L'analyse de contribution
 - L'évaluation réaliste

Les évaluations d'impacts contrefactuels

Devis expérimental (ou l'expérimentation par assignation aléatoire)

L'expérimentation par assignation aléatoire (*randomized controlled trial*) est une approche « totale » qui dépasse l'évaluation pour se positionner comme outil de production de l'action publique par l'expérimentation. Par métonymie, cette approche est généralement appelée par le nom de sa méthode principale, l'essai contrôlé randomisé (ECR). L'ECR consiste à comparer, dans le cadre d'une expérimentation, les effets d'une intervention sur deux groupes, l'un des groupes y ayant accès et l'autre non, au regard d'un indicateur donné. Chaque groupe est tiré aléatoirement au sein de la population admissible, afin de prémunir l'expérimentation d'un biais de sélection. L'impact est alors la différence entre la valeur constatée pour le groupe avec l'intervention et le groupe sans intervention (White, 2014).

Au-delà de cet aspect, l'expérimentation par assignation aléatoire vise à créer les conditions pour juger « scientifiquement » du mérite intrinsèque d'une intervention hors influence de la mise en œuvre – le jugement étant positif si l'intervention marche dans ces conditions contrôlées. C'est le fait de fournir une preuve d'impact qui justifie, d'un point de vue éthique, le recours à un tirage aléatoire et les autres mesures mises en place – plutôt que d'aider ceux qui en auraient le plus besoin, par exemple.

Sous son apparente simplicité, la complexité de l'expérimentation par assignation aléatoire tient aux efforts extraordinaires nécessaires pour maintenir les conditions de la validité des résultats produits (Shadish *et al.*, 2012). Ces conditions limitent aussi son usage à des situations où un lien séquentiel direct entre une intervention et un effet peut être envisagé. L'expérimentation est conçue, le cas échéant, pour rendre cette situation observable.

Exemple 1 – Expérimentation par assignation aléatoire

Une expérimentation est engagée auprès de jeunes diplômés qui peinent à trouver du travail. Pour en estimer les effets sur l'accès à l'emploi, un ECR est organisé. Il vise à comparer trois groupes, l'un « traité », l'autre de contrôle, et enfin un groupe de « super-contrôle », comprenant des jeunes issus de territoires non affectés par l'expérimentation. Cette modalité cherche à vérifier si l'accès au travail des jeunes inscrits se fait au détriment des autres jeunes diplômés (effet de substitution). L'analyse réalisée présente à première vue un impact mesurable sur l'emploi à 12 mois : à la fin de l'expérimentation, 39,5 % des jeunes concernés sont en emploi, contre 37 % des jeunes du groupe de contrôle. Cependant, ces taux sont inférieurs à ceux identifiés dans le groupe de super-contrôle. De plus, la reproduction des enquêtes à 16 mois et à 20 mois montre qu'il n'y a plus, à ce stade, de différences entre le groupe test et le groupe de contrôle (cité dans Behaghel *et al.*, 2013).

Les devis quasi expérimentaux

Nous regroupons sous cette dénomination commune toutes les approches (*quasi-experimental designs*) qui visent à appliquer l'approche expérimentale, comparant une situation avec l'intervention, et une situation sans l'intervention, dans des contextes non expérimentaux. La comparabilité entre groupes ne pouvant être garantie à l'aide de l'assignation aléatoire, le groupe témoin doit être reconstitué, à partir de bases de données existantes ou produites pour l'occasion (par enquête, par exemple).

Des concepts et des méthodes statistiques sont utilisés pour garantir la comparabilité des groupes, en s'assurant de la proximité de la composition entre les deux groupes et/ou en prenant en compte leurs différences dans l'estimation de l'impact. Dans le premier cas, on utilise des méthodes d'appariement : pour chaque individu du groupe test, on trouve un « jumeau » (c'est-à-dire un individu aux caractéristiques observables similaires), en utilisant des techniques plus ou moins sophistiquées pour ce faire. Dans le second, on mesure les différences pour les retirer de l'opération : la méthode de double différence consiste ainsi à comparer deux groupes, l'un soumis à l'intervention, l'autre non, et ce, avant, à l'entrée, et à la sortie du dispositif, au regard d'un indicateur d'impact. Les niveaux mesurés sont différents, mais si les groupes sont suffisamment similaires, leurs évolutions le sont également. L'impact est estimé en défalquant de la différence entre les deux groupes à la sortie du dispositif, celle observée à l'entrée (White et Sabarwal, 2014).

Exemple 2 – Devis quasi expérimental

Un programme vise à lutter contre la malnutrition infantile au Bangladesh à travers des ateliers ciblant les mères de famille. Pour en vérifier les effets, des données relatives à l'âge, à la taille et au poids des enfants inclus dans le dispositif sont comparées à celles réunies dans le cadre d'une enquête nutritionnelle alimentaire. Un appariement par score de propension, « qui désigne la probabilité de participation à l'intervention au vu des caractéristiques observables de l'individu » (White et Sabarwal, 2014) permet de construire le groupe de contrôle. L'impact de l'intervention est ensuite calculé en comparant les chiffres obtenus en moyenne pour chaque groupe. L'évaluation montre que, contrairement à ce que semblaient montrer les données de suivi, les enfants passés par le programme ne se distinguent pas significativement des autres enfants bangladais à caractéristiques comparables (White, 2009).

Les approches basées sur la théorie

Nous regroupons sous cette ombrelle commune (*theory-based evaluation*) une série d'approches partageant des caractéristiques communes, et relevant d'une volonté d'expliquer la façon dont les interventions évaluées ont des effets (Brousselle et Buregeya, 2018). Elles sont appelées ainsi parce qu'elles organisent tout ce processus explicatif autour de la constitution puis de la vérification d'une théorie du changement (TdC) (Funnell et Rogers, 2011)⁷. Cette théorie « avec un petit t » (Chen et Rossi, 1987) est élaborée au début de l'évaluation à partir d'hypothèses implicites ou explicites portant sur l'intervention et sa capacité à avoir des conséquences, le pourquoi et le comment, et dans quels cas. Des questions d'évaluation sont ensuite posées pour interroger tout ou partie des hypothèses réunies dans la TdC, et des outils de collecte et d'analyse sont choisis et articulés pour ce faire. L'évaluation se termine par une mise à jour de la théorie, qui est mise à contribution pour améliorer la faculté de l'intervention à avoir des effets désirables.

Un aspect fondamental de ces approches est de considérer que les interventions ne « marchent pas » toutes choses égales par ailleurs : elles reconnaissent au contraire une diversité de cas, et considèrent le « contexte » comme un élément affectant les changements observés. C'est par ce biais qu'elles abordent les interventions comme des objets complexes dans des environnements complexes (Robert et Ridde, 2013).

Différentes approches distinctes prospèrent aujourd'hui au sein de cette famille. Nous en explorons deux ci-dessous.

L'analyse de contribution

L'analyse de contribution (*contribution analysis*) se distingue dès ses premiers pas par une démarche humble et pragmatique : son objectif n'est pas de prouver les effets des interventions, mais bien de réduire l'incertitude quant à leur contribution, de façon à éclairer la décision. C'est ainsi qu'elle est progressivement devenue une approche particulièrement adaptée aux situations complexes, lorsqu'une contribution est plausible, mais n'est certainement ni séquentielle ni linéaire (Mayne, 2012), ou lorsqu'il est probable, dès le départ, que l'intervention ne soit pas le principal vecteur du ou des changements recherchés (Delahais et Toulemonde, 2017).

Une originalité de l'analyse de contribution est que sa recherche de liens causaux ne se limite pas à l'intervention. Au contraire, elle vise spécifiquement à repérer des « paquets de causes », associant l'intervention et d'autres facteurs, et qui peuvent être des conditions nécessaires ou suffisantes aux changements escomptés.

La démarche d'analyse de contribution peut être assimilée à une sorte d'enquête : à partir de la théorie du changement initiale, des hypothèses de contribution sont bâties, et leur plausibilité testée de diverses façons. Ces hypothèses portent à la fois sur l'intervention et sur des explications alternatives. Elles sont ensuite testées, à charge et à décharge, lors de la collecte proprement dite. Dans ce processus, mesure des changements et estimation des contributions sont strictement séparées, ce qui demande à la fois un travail de documentation des changements et d'élaboration de tests empiriques relatifs à

7. Il existe de très nombreux termes qui peuvent être utilisés pour parler de chaînes causales représentant les interventions et leurs conséquences de façon séquentielle, nous utilisons ici ce terme de façon générique. Pour une vue générale, voir Funnell et Rogers (2011).

la contribution. L'analyse de contribution fournit un effort particulièrement important d'identification des impacts en contexte; en contrepartie, elle ne prétend pas à établir des modèles généraux, mais plutôt les ingrédients et les circonstances pour leur utilisation (Delahais *et al.*, 2020).

Exemple 3 – Analyse de contribution

Un centre de recherche souhaite évaluer sa contribution à la gestion durable des forêts dans le bassin du Congo sur les 20 dernières années. Il est évident qu'il n'en est pas le principal vecteur, mais les parties prenantes font valoir des contributions plausibles de la recherche qui peuvent être approfondies. Un système représentant les principaux acteurs (États, organisations internationales, acteurs de la recherche, ONG, entreprises, habitant(e)s...) ayant contribué à cette gestion durable est élaboré à partir de la littérature, et permet de documenter les principaux changements intervenus sur la période. Des hypothèses relatives à la contribution du centre de recherche aux différentes étapes de ce système sont émises, puis testées systématiquement au travers d'études de cas nationales et thématiques : in fine, l'évaluation fait apparaître sept contributions majeures de la recherche à la gestion durable des forêts, à plusieurs points du système, articulées avec d'autres facteurs expliquant ces changements (Delahais et Toulemonde, 2017). Par exemple, le centre de recherche fournit les connaissances nécessaires pour pouvoir développer une certification durable adaptée aux forêts de la région, elle-même nécessaire pour que les entreprises puissent mieux valoriser le bois produit de cette façon.

L'évaluation réaliste

L'approche réaliste (*realist evaluation*) tire son nom du paradigme réaliste critique dont elle est l'adaptation à l'évaluation (Pawson et Tilley, 1997). Elle se distingue ainsi par sa conceptualisation des effets, son objet et sa volonté de généralisation.

Dans l'approche réaliste, en effet, les interventions ne « marchent pas » d'elles-mêmes : elles fournissent à des acteurs une opportunité d'agir d'une certaine façon, en raison de leurs caractéristiques propres, des interactions qu'ils ont entre eux et avec les acteurs qui portent l'intervention, et selon le moment ou les circonstances. Les raisonnements multiples qui les amènent à se saisir ou non de l'intervention sont appelés « mécanismes. » De ceux-ci dépendent in fine les effets de l'intervention, et l'évaluation s'attache à les repérer et à comprendre les contextes dans lesquels ils sont activés ou non.

Là où l'analyse de contribution cherche à établir des systèmes, l'évaluation réaliste cherche plutôt à épuiser les explications relatives à un moment donné de la théorie du changement. Elle est notamment utilisée pour comprendre la façon dont sont mises en œuvre les interventions. Ces explications sont présentées sous la forme de configurations CME : dans tel Contexte, une intervention déclenche tel Mécanisme qui génère tel Effet (en anglais CMO, pour *Context-Mechanism-Outcome*). L'analyse méthodique de ces configurations doit aboutir à des théories de moyenne portée, qui permettent d'expliquer les régularités observées dans les configurations CME (Robert et Ridde, 2013), et ainsi, in fine, d'améliorer les interventions.

Exemple 4 – Évaluation réaliste

Le programme *Dance with Parkinson's* consiste en des cours de danse pour les personnes atteintes de cette maladie dégénérative. Cette intervention est un succès : les participants et participantes reviennent régulièrement. Ses promoteurs y ont apporté des améliorations continues, mais rencontrent des difficultés à maintenir leurs financements, et cherchent alors à prouver de façon systématique les effets obtenus. L'approche évaluative mise en place est participative et réunit les acteurs du programme pour construire une TdC, dans laquelle sont précisés les changements attendus sur l'état des participants et participantes, et les chemins par lesquels l'intervention peut exercer une influence sur cet état. Des observations et des entretiens avec les différentes personnes impliquées (enseignant(e)s, participant(e)s, leur famille) permettent de tester cette théorie. Plusieurs mécanismes sont identifiés, liés notamment à un sentiment de liberté retrouvée, à la confiance en soi, au plaisir également (les classes sont des moments amusants), qui génèrent des effets à court terme (meilleure humeur, moins de chutes...) et à long terme (meilleure qualité de vie pour les participants et participantes et pour ceux qui leur prodiguent des soins) (Gibson et Robichaud, 2020).

Les approches concernées par les valeurs

Nous proposons ici une série d'approches dans lesquelles ce sont les valeurs des parties prenantes qui sont au cœur de l'évaluation. Quelles que soient leurs différences par ailleurs, ces approches prévoient d'identifier les parties prenantes et leurs valeurs, et organisent l'évaluation de façon à examiner l'intervention au regard de celles-ci. Aussi ces approches comportent-elles toujours une dimension participative ainsi que des outils de collecte et d'analyse permettant de rendre compte finement des contextes, des attitudes et des comportements des acteurs.

Encadré 2 – Approches concernées par les valeurs

- La famille des approches participatives
- L'évaluation répondante
- L'évaluation démocratique délibérative

La famille des approches participatives

Les approches présentées dans les pages suivantes, qu'elles soient concernées par les valeurs, par l'utilisation ou par la justice sociale, accordent une place importante aux parties prenantes. Elles peuvent généralement être menées de façon participative (mais pas toujours).

Une évaluation participative donne à un groupe de parties prenantes un rôle décisionnel dans la conduite de l'évaluation, ce contrôle pouvant être total ou partagé; ces parties prenantes représentent une diversité de perspectives sur l'intervention, et à tout le moins d'autres points de vue que ceux des décideurs ou des financeurs; enfin, les parties prenantes doivent contribuer à la production de l'évaluation et de ses cadres, au minimum en influant sur la façon dont l'intervention va être jugée et sur ses usages. Le processus évaluatif organise la délibération entre les parties prenantes, permettant de s'informer, mais aussi d'évoluer dans ses préférences et d'aboutir à un jugement plus fondé (Delahais, 2021).

On peut considérer deux visées principales des évaluations participatives (Cousins et Whitmore, 1998). D'un côté, produire de meilleures évaluations, plus pertinentes, plus crédibles et plus utiles. C'est ainsi que les différentes approches concernées par l'utilité de l'évaluation voient en priorité la participation. De l'autre, donner à ceux qui n'ont pas la parole habituellement les moyens de se faire entendre et de changer les choses. C'est la perspective des approches orientées vers la justice sociale. Les deux approches présentées ci-dessous se placent entre ces deux visées, avec un souci d'être utile, mais sans méconnaître les enjeux de pouvoir et de domination qui affectent la société.

L'évaluation répondante

La prémisse essentielle de l'évaluation répondante (*responsive evaluation*) est de considérer que les parties prenantes ont des valeurs différentes; que ces valeurs jouent un rôle dans la façon dont les uns et les autres jugent une intervention; que l'évaluateur ou l'évaluatrice ne devrait pas imposer de valeurs particulières dans le jugement apporté. Cette ascendance constructiviste étant établie, tout le rôle de l'équipe d'évaluation est de construire le processus permettant de déterminer les valeurs qui comptent, de faire émerger les enjeux qui intéressent les différents acteurs impliqués pour pouvoir y répondre et de favoriser, in fine, une pluralité de jugements. En ce sens, l'évaluation répondante cherche à aider ces derniers à se faire leur avis sur l'intervention, de façon à faire les choix qui leur semblent adéquats.

Un des défis essentiels de l'évaluation répondante est de se donner les marges de manœuvre suffisantes pour identifier les sujets qui feront effectivement l'objet de l'évaluation. Au début de l'évaluation, les parties prenantes ont des questions relativement basiques, du type « est-ce que les objectifs sont atteints? ». Or, l'évaluation répondante ne donne pas une place prépondérante aux objectifs (ou à la logique d'intervention) dans la définition de ce qui compte. Il faut du temps et de la flexibilité, des interactions avec l'équipe d'évaluation, des premiers éléments de collecte pour que des questionnements plus essentiels émergent et guident la suite de l'évaluation. Les études de cas, qui explorent en profondeur une situation, son contexte, les effets attendus et inattendus d'une intervention, les différents points de vue à son sujet, sont l'outil privilégié de l'approche répondante, non seulement pour la richesse d'informations qu'elles apportent, mais aussi parce qu'elles permettent aux parties prenantes d'appuyer leur jugement.

Dans cette approche, les interactions entre les parties prenantes et l'équipe d'évaluation sont nombreuses, mais les rôles sont clairement répartis : c'est l'équipe d'évaluation qui, à partir des interrogations des acteurs, mène ses investigations et en rapporte ensuite les résultats. Cependant, les acteurs de l'intervention sont considérés comme des participants actifs à l'enquête, plutôt que des sujets : on leur demande comment ils perçoivent les changements, plutôt que de leur faire passer des tests pour les mesurer (Stake, 2003).

Exemple 5 – Évaluation répondante

Un programme aide les anciens patients et patientes d'un établissement psychiatrique à retrouver un emploi. Ce programme se construit chemin faisant, et l'évaluation vise à permettre à ses acteurs, et en particulier aux personnels, à réfléchir à leurs pratiques et à les améliorer. Très rapidement, il apparaît à l'évaluatrice que les pratiques des thérapeutes sont probablement un obstacle à la réinsertion des patients et patientes. Dans leur vision des choses, ces derniers ont besoin de l'expertise, des soins et de la protection des personnels; ils ne savent pas ce qui est bon pour eux. Pour faire évoluer leur pratique, il faut faire changer ce récit. L'évaluatrice va alors collecter d'autres histoires, qui ne cadrent pas avec ce récit dominant, auprès de thérapeutes, de patients et patientes, et de spécialistes, et les rapporte aux parties prenantes. Elle montre comment les patients et patientes vivent ces pratiques et comment elles les maintiennent dans la dépendance. Ce processus amène certains personnels à se rendre compte que leur pratique pose problème, et à engager un processus de transformation. Le rapport ne comprend ni conclusions (à chacun d'en tirer) ni recommandations (Abma, 1998).

L'évaluation démocratique délibérative

Cette approche (*democratic deliberative evaluation*) a l'ambition de contribuer à la bonne marche de la démocratie, en faisant la démonstration de ce qu'est un processus démocratique, mais aussi en permettant que les décisions prises dans le cadre des institutions démocratiques le soient à partir de constats robustes et de conclusions étayées. Pour cela, trois principes majeurs sont mis de l'avant : l'inclusion des parties prenantes; le dialogue; la délibération.

L'évaluation démocratique partage l'idée que les interventions sont l'objet de nombreuses perspectives, mais elles ne sont pas toutes égales : pour être équitable, l'équipe d'évaluation doit rechercher activement la participation de certaines parties prenantes et doit éviter que les plus puissants n'imposent leurs vues. Ces parties prenantes doivent pouvoir avoir l'occasion de dialoguer sur les enjeux identifiés et les analyses apportées par l'évaluation. Ce dialogue permet à l'équipe d'évaluation de comprendre pleinement les positions et les points de vue des groupes en présence, et à ceux-ci de clarifier leurs propres intérêts vis-à-vis de l'intervention. C'est grâce à ce processus que sont définis les principaux sujets d'investigation. Le processus de délibération vise enfin à mettre en perspective ces différents points de vue avec les informations collectées lors de l'évaluation : il ne s'agit pas de les prendre pour argent comptant, et l'objectif de l'évaluation est bien d'aboutir à un jugement de la valeur de l'intervention qui peut être acceptable pour le plus grand nombre. C'est là où l'équipe d'évaluation doit prendre ses responsabilités et proposer, in fine, les conclusions qui lui semblent les plus adaptées, les débattre avec les parties prenantes (désormais averties de la diversité des perspectives et des éléments les soutenant ou pas) et les consolider (House et Howe, 2013).

Exemple 6 – Évaluation démocratique délibérative

Un programme fournit une éducation bilingue à des élèves qui ne maîtrisent pas suffisamment la langue officielle d'enseignement. Le contexte, dans cette ville américaine, est difficile, avec une animosité croissante entre communautés, qui se cristallise autour de ce programme. L'évaluateur commence par rencontrer les parties prenantes, y compris les principaux soutiens et adversaires du programme. Des mesures sont ensuite prises pour s'assurer que tous les groupes en présence acceptent la démarche évaluative et ses résultats : ainsi, un comité composé des différentes parties est réuni; les indicateurs de performance sont discutés jusqu'à ce qu'ils soient unanimement considérés comme des indicateurs de succès; la collecte est menée par deux anciens directeurs d'établissement, respectés de tous, etc. L'évaluateur accorde une attention particulière aux affirmations des parties. Par exemple, certaines familles se plaignent que les élèves sont envoyés trop tôt en classe d'anglais : le niveau des élèves lors du passage d'une classe à l'autre fait donc l'objet d'investigations spécifiques. Se basant sur cinq ans d'évaluation, l'évaluateur met de l'avant que le programme fonctionne désormais correctement, et que les parties prenantes peuvent échanger entre elles, et considère que le processus mis en place, transparent, équitable, juste, y est pour quelque chose (Mertens et Wilson, 2012)⁸.

Les approches concernées par l'utilisation

Nous regroupons ici des approches foncièrement concernées par l'utilisation de l'évaluation, et qui sont organisées en conséquence. Michael Patton a une importance singulière, tant pour son investissement de longue date sur ce sujet, que pour sa capacité à formaliser ses pratiques évaluatives sous la forme de modèles complets. Nous commencerons donc par deux d'entre elles, et terminerons par une troisième approche, la récolte des incidences, qui se place dans la filiation de Patton tout en s'en démarquant par une méthodologie originale.

Encadré 3 – Approches concernées par l'utilisation

- L'évaluation axée sur l'utilisation (ÉAU)
- L'évaluation évolutive
- La cartographie et la récolte des incidences

8. Cette évaluation a été menée par E. House lui-même. Elle est résumée dans Mertens et Wilson (2012, p. 176-178).

L'évaluation axée sur l'utilisation (ÉAU)

Lorsqu'elle apparaît sur la scène évaluative, cette approche (*utilization-focused evaluation*) suscite la controverse par une proposition essentielle : « les évaluations devraient être jugées en fonction de leur utilité » (Patton, 2008; Patton et Labossière, 2012), et non en fonction de la méthodologie déployée, par exemple. L'utilisation n'est pas quelque chose que l'on attend vaguement à la fin de l'évaluation, mais l'objet même de toute la démarche. L'ÉAU n'est ainsi pas prescriptive en matière d'outils de collecte ou d'analyse, l'important est que ceux-ci permettent de répondre aux besoins d'information affirmés par les usagers.

Un des éléments essentiels de l'ÉAU est qu'elle est entièrement organisée autour des personnes qui vont être impliquées dans le processus évaluatif. Elle démarre donc par l'identification des principaux utilisateurs attendus d'une évaluation, qu'il s'agisse des commanditaires ou d'autres acteurs, et la détermination de leur niveau de participation à la démarche évaluative. Les parties prenantes s'accordent ensuite sur les utilisations prévues de l'évaluation (quels besoins d'évaluation, pour quoi faire, à quel moment) et précisent le questionnement évaluatif en conséquence. Les utilisateurs sont amenés à s'impliquer à la fois dans le choix des outils mis en œuvre pour répondre aux questions, puis dans l'interprétation des résultats obtenus, toutes ces étapes visant à renforcer leur engagement et in fine l'utilisation de l'évaluation. Ainsi, le déroulement de l'évaluation est, tout autant que les résultats, vecteur de meilleure utilisation.

L'ÉAU fait de l'évaluation un processus dynamique, susceptible d'évoluer en cours de route pour préserver, renforcer ou élargir le potentiel d'utilisation. Elle donne un rôle crucial d'animation et de négociation à l'évaluateur ou à l'évaluatrice, qui doit en plus intégrer des enjeux éthiques plus importants qu'à l'accoutumée, relatifs aux usages qui seront faits de l'évaluation.

Exemple 7 – Évaluation axée sur l'utilisation

Deux expérimentations visent à accueillir des jeunes en grande précarité en résidence (sous forme d'internat ou de séjour pendant les vacances scolaires) pendant leur scolarité. Dans un premier temps, l'équipe d'évaluation échange sans démarche évaluative prédéfinie avec des utilisateurs potentiels. La construction du référentiel est ensuite engagée de façon itérative, elle permet d'identifier une problématique et des questions d'évaluation intéressant toutes les parties ainsi que les principaux éléments de méthode pour y répondre. Les échanges font apparaître le besoin de mettre en place un système de suivi, pour aider les opérateurs à ajuster et à structurer l'action au fur et à mesure : le choix des indicateurs fait l'objet de négociations avec les utilisateurs potentiels, dans l'optique de trouver un bon équilibre entre ce qui serait le plus utile et ce qui est possible de faire pour suivre sans nuire au travail des équipes. Le travail sur les écarts entre le cadre de vie des jeunes et celui proposé dans les structures d'accueil (culture, discipline, modes d'interaction, confort) s'avère assez rapidement un enjeu clé de la réussite du projet sur le long terme, et les porteurs de projet demandent un cadre leur permettant d'y réfléchir : l'équipe d'évaluation identifie et propose alors des concepts théoriques qui serviront à structurer l'évaluation ainsi qu'à outiller les opérateurs. Les échanges lors de l'évaluation, y compris sur la collecte effectuée et l'interprétation des données, favorisent l'appropriation de l'évaluation par tous et toutes. In fine, des résultats distincts sont fournis aux différentes catégories d'utilisateurs (Devaux-Spatarakis *et al.*, 2021).

L'évaluation évolutive

Contrairement à l'ÉAU, qui s'apparente presque à une posture pouvant s'appliquer à tout contexte évaluatif, l'évaluation évolutive (*developmental evaluation*) est conçue pour maximiser l'utilisation de l'évaluation dans le cadre très précis de l'innovation sociale, ou plus largement de changements radicaux qui induisent une très forte incertitude quant à leur mise en œuvre et à leurs impacts. Ces innovations fonctionnent par itérations successives, qui peuvent être très fréquentes. Or, une évaluation classique prend du temps pour poser ses critères et collecter des informations pour renseigner ces derniers; elle risque d'être systématiquement trop tardive pour être utile – ou pire, inadéquate – par exemple parce qu'elle travaille sur des hypothèses qui sont déjà obsolètes.

Avec l'évaluation évolutive, l'évaluateur ou l'évaluatrice intègre les équipes chargées de l'intervention évaluée, et s'attache à apporter des rétroactions rapides et pertinentes pour permettre aux innovateurs d'apprendre en temps réel et de mener à bien leurs ambitions. La posture évaluative en est considérablement modifiée : l'évaluatrice ou l'évaluateur est aussi intéressé que les autres à la réussite de l'intervention et agit en conséquence. Elle ou il met à profit sa pensée critique et sa posture évaluative, notamment pour cadrer les enjeux émergents, identifier des changements en cours, construire des typologies, poser des hypothèses et mettre en place des moyens pour les vérifier. L'évaluation doit aussi être en position de passer d'un rôle à un autre dans sa relation avec les parties prenantes : membre de l'équipe, ami critique, accompagnateur de changements, communicant, le tout demandant une forte flexibilité et des qualités dépassant largement celles qui sont généralement associées à l'évaluation (Dozois *et al.*, 2011; Turcotte *et al.*, 2021).

Exemple 8 – Évaluation évolutive

De plus en plus, les multinationales du numérique délocalisent certaines de leurs activités dans des pays du sud. Un appel de projets d'une Fondation états-unienne est engagé au Kenya, en Afrique du Sud, au Ghana et en Inde pour vérifier si des actions de formation des jeunes leur permettraient de mieux bénéficier de ces occasions d'emploi. L'équipe d'évaluation, engagée en amont, s'attache à partager ses objectifs de créer de la connaissance partagée et de contribuer à l'amélioration collective, de façon à éviter la défiance qu'une démarche plus classique aurait pu faire naître. Elle multiplie les échanges, et répond aux besoins des porteurs et porteuses quand ils se manifestent. Rapidement, l'équipe d'évaluation est considérée comme faisant partie de l'équipe projet, ce qui facilite la transmission d'informations; inversement, les porteurs et porteuses intègrent les questions d'évaluation dans leurs processus et s'attachent à contribuer à l'acquisition de connaissances à travers leurs actions (Harji et Jackson, 2016).

Les approches de cartographie et récolte des incidences

Ces deux approches, distinctes, mais partageant de nombreuses caractéristiques, représentent un continuum allant du suivi à l'évaluation, jusqu'à la mise en œuvre des interventions. Leur originalité est qu'elles naissent d'une définition spécifique des incidences (*outcomes*) : « des changements dans le comportement, les relations, les actions, les activités, les politiques, ou les pratiques d'un individu, d'un groupe, d'une communauté, d'une organisation ou d'une institution » (Earl *et al.*, 2001). Cette définition met les acteurs au centre des réflexions (et des processus, souvent participatifs); tient à

distance les objectifs ou autres résultats prédéterminés; et s'ouvre au contraire à l'ensemble des conséquences, attendues ou non, désirables ou non, des interventions. Pour cela, elle est particulièrement intéressante dans des contextes où ce sont les conséquences dans leur globalité qui importent (par exemple dans l'action sociale ou culturelle), ou encore lorsque l'évaluation porte sur des grappes de projets aux effets très différents, qui ne pourraient pas être captés par une théorie du changement.

La cartographie des incidences (*outcome mapping*) vise, en amont, à identifier les incidences et à construire des instruments de planification et de mesure autour d'elles. C'est un processus itératif, qui peut tout autant être mobilisé pour le suivi qu'à la mise en œuvre de l'intervention. La récolte des incidences (*outcome harvesting*) part, elle, en aval, de changements concrètement observés et cherche, à partir de là, à repérer les contributions de l'intervention évaluée. Elle peut se faire en temps réel, et contribuer ainsi à faire évoluer les interventions, ou bien servir à réaliser une évaluation rétrospective.

L'utilisation est au cœur de l'ensemble de ces démarches. La 1^{re} étape de la récolte des incidences consiste ainsi à identifier avec les utilisateurs des questions qui répondent à leurs besoins, et la dernière à organiser une discussion entre « récolteurs » d'information et utilisateurs. Les processus sont conçus pour être simples d'accès. L'accent sur les personnes permet d'impliquer plus facilement toutes les parties prenantes, en particulier celles qui ne se retrouveraient pas dans des objectifs jugés artificiels (Wilson-Grau et Britt, 2012).

Exemple 9 – Récolte des incidences

Une organisation à but non lucratif accompagne des communautés locales dans la mise en place d'actions en faveur du développement des tout-petits et du bien-être de leurs familles. Elle met en place une récolte des incidences portant sur ses soutiens apportés aux groupements de partenaires locaux pour juger rétrospectivement de son intervention. Une partie de la démarche est réalisée par entretiens individuels auprès de 20 groupements, permettant de recueillir 223 incidences, dont 24 négatives. Elles sont ensuite regroupées et analysées selon plusieurs variables (axes stratégiques, temporalité, durabilité, thématiques d'interventions, etc.) pour rendre compte des apports du programme. Parallèlement, des ateliers participatifs de récolte sont organisés dans les regroupements qui le souhaitent et ils visent à consolider leurs apprentissages dans une perspective de pérennisation des acquis après dix ans de soutien. Les participants et participantes présentent des changements qu'ils ont observés, donnent un exemple concret, exposent les conditions de succès nécessaires pour produire les changements et le faire perdurer, et réfléchissent au lien avec les contributions de l'organisation qui les a soutenus⁹.

Les approches concernées par la justice sociale

Les approches présentées ici mettent l'évaluation au service d'une société plus juste et plus équitable. Nous parlerons dans un premier temps des approches transformatives, avant de terminer par une approche très différente, l'évaluation émancipatrice.

9. Cet exemple est tiré d'une expérience réelle, menée au Québec entre 2017 et 2019 par l'organisation Avenir d'Enfants. Pour en savoir plus : <https://agirtot.org/thematiques/recolte-des-effets-1-de-2/>

Encadré 4 – Approches concernées par la justice sociale

- Les approches transformatives
- L'évaluation féministe
- L'évaluation attentive aux différences culturelles et l'évaluation autochtone
- L'évaluation émancipatrice

Les approches transformatives

Ces approches sont inspirées de courants critiques tels que la recherche-action participative, les luttes anti-discriminations ou le féminisme. Ces approches ont des proximités avec les approches centrées sur les valeurs, mais s'en distinguent en cela qu'elles nomment clairement les maux qu'elles combattent (le racisme, le sexisme, et toute forme de discrimination et d'oppression) et cherchent à donner aux groupes n'ayant pas habituellement voix au chapitre les moyens de se faire entendre en prenant le pouvoir sur les processus évaluatifs.

L'évaluation est donc une activité politique. Cela veut dire qu'elle est explicitement mise au service d'une lutte pour la justice sociale (et doit donc être utile si elle veut faire évoluer le statu quo). Mais cela signifie aussi qu'elle est elle-même inscrite dans des relations de pouvoir et des inégalités structurelles. L'évaluatrice ou l'évaluateur doit donc être conscient de la place occupée, de son pouvoir, de la nature de ses savoirs (légitimés par le modèle dominant) et de ses lunettes culturelles, pour pouvoir les mettre à distance; chercher constamment à organiser les conditions de la meilleure prise en compte de l'expertise des groupes vivant les inégalités; avoir un sens intense de l'éthique; tout cela en gardant en permanence à l'esprit le souci de la transformation sociale, et les chemins par lesquels l'évaluation peut y contribuer.

Les approches transformatives ne sont pas prescriptives en termes d'étapes, mais Donna Mertens propose le déroulement suivant : une phase initiale qui doit permettre de construire une évaluation avec un vrai potentiel de transformation, et dans laquelle il faut donc connaître les publics, créer avec eux des relations de confiance et trouver les façons de faire bon usage de leur expertise; une phase de collecte d'informations qui peut utiliser une large variété d'outils, selon les besoins; une phase d'interprétation collective des données collectées, avec un objectif de construction des usages de l'évaluation (Mertens, 2017).

L'évaluation féministe et l'évaluation attentive aux différences culturelles, présentées ci-dessous, s'inscrivent dans la famille des approches transformatives.

L'évaluation féministe

Cette approche (*feminist evaluation*) prend comme point de départ les rapports de genre et avance l'idée que ceux-ci génèrent des injustices sociales qui sont systématiques et structurelles. En « chaussant des lunettes de genre », en analysant les rapports de genre et leurs conséquences, comment ils affectent l'accès aux interventions évaluées, leur mise en œuvre ou leurs effets, l'équipe d'évaluation effectue un pas de côté qui lui permet de voir l'intervention différemment. À noter que malgré son choix

d'appellation, l'évaluation féministe est aussi sensible aux autres sources d'injustice et cherche à comprendre comment ces facteurs s'additionnent et se renforcent mutuellement (intersectionnalité) (Podems, 2018).

Exemple 10 – Évaluation féministe

Un programme s'adressant à des personnes toxicomanes et présentant des troubles psychiatriques leur propose un module optionnel visant à leur faire prendre conscience des traumatismes qu'elles ont vécus. Les évaluatrices ont repéré que les femmes étaient très peu nombreuses à s'inscrire à ce module et ont donc intégré un cadre féministe à leur évaluation. Cela leur a permis de mettre en regard le discours des thérapeutes, qui expliquaient initialement le faible taux de recours par le caractère facultatif du module et ses exigences en matière de temps, de réflexion et de travail personnel, avec l'expérience des femmes. Celles-ci, en effet, s'inscrivent moins, car elles déclarent moins que les hommes les traumatismes qu'elles ont vécus, et ce pour plusieurs raisons : 1) les femmes ne font pas toujours le lien entre leur santé mentale et leur expérience traumatique; 2) pour les femmes, la culture rurale rend plus difficile la reconnaissance qu'elles puissent avoir besoin d'aide; 3) les thérapeutes ont tendance à ne considérer comme traumatiques que les abus physiques et sexuels. C'est la réflexivité sur la manière dont les valeurs personnelles des évaluateurs et évaluatrices « s'alignent, s'entremêlent et/ou entrent en conflit avec les valeurs des principales parties prenantes lorsqu'elles définissent les problèmes, identifient les populations cibles, conçoivent les programmes, mesurent les résultats et rendent compte des conclusions » qui permet de produire des connaissances capables de générer du changement (Sielbeck-Mathes et Selove, 2014).

L'évaluation féministe, approche transformationnelle, ne doit pas être confondue avec les approches attentives au genre (*Gender-Responsive Evaluation*), qui visent plutôt à évaluer les interventions au regard d'un objectif d'égalité des sexes et d'autonomisation des femmes, sans remettre en question leur dimension structurelle (ONU Femmes, 2015).

L'évaluation attentive aux différences culturelles et l'évaluation autochtone

Cette approche transformationnelle (*culturally responsive evaluation*) part du principe que « des valeurs et des croyances liées à la culture sont au cœur de tout effort évaluatif » (Hood *et al.*, 2015, p. 283), la culture étant définie comme « une accumulation de comportements, valeurs, coutumes ou croyances, apprises et partagées, qui sont communs à un groupe ou une société en particulier » (idem). L'évaluation attentive aux différences culturelles porte une attention particulière aux groupes culturels marginalisés ou discriminés en raison de leur origine, de leur apparence physique, de leur capacité à s'exprimer dans la langue officielle ou usuelle d'un pays, ou de leur situation économique ou ethnique, parmi tant d'autres.

Les postulats de cette approche évaluative ont plusieurs implications majeures. La première, c'est la nécessité de repérer les groupes culturels qui sont concernés ou touchés par les interventions évaluées – sachant que chaque individu a des identités multiples, il y a donc un travail important à faire pour comprendre lesquels forment une typologie pertinente dans un contexte donné. La deuxième, c'est la prise en compte des valeurs et des croyances liées à la culture de ces groupes dans l'ensemble de la démarche évaluative : questionnement évaluatif autour d'enjeux qui préoccupent et affectent ces

groupes; conception et mise en œuvre d'outils permettant de faire ressortir les modes de connaissance propres à ces groupes, et favorisant leur investissement dans l'évaluation; interprétation des résultats à l'aune des modes de connaissance de ces groupes. La troisième et dernière implication est la nécessité pour l'évaluateur ou l'évaluatrice d'avoir ou de développer des compétences culturelles qui sont appropriées pour le groupe culturel concerné. Outre des savoirs, savoir-faire et savoir-être que tout évaluateur ou évaluatrice peut acquérir, il s'agit aussi, à chaque nouvelle évaluation, de se placer dans une posture d'échange avec les groupes culturels concernés, pour recommencer le processus d'apprentissage culturel, mais aussi pour construire les relations de confiance obligatoires à la bonne tenue de l'évaluation.

L'évaluation indigène et la volonté de décoloniser l'évaluation prolongent et élargissent l'évaluation attentive aux différences culturelles. Dans ce cas, c'est plus largement la « connaissance occidentale » qui est mise à distance; par exemple, les hypothèses relatives à l'individu autonome et rationnel; le dualisme cartésien entre corps et esprit; l'importance donnée à la maîtrise technique au détriment des savoirs d'usage ou de la sagesse pratique, etc. (Wehipeihana et McKegg, 2018). Un enjeu est alors d'imaginer des évaluations s'inscrivant beaucoup plus largement dans des systèmes de connaissance indigènes conçues et mises en œuvre par et avec des évaluateurs et évaluatrices indigènes. C'est ce que recherche notamment le mouvement *Made in Africa* (Frehiwot, 2019). Nan Wehipeihana décrit quant à elle l'évaluation autochtone comme « une évaluation par des personnes autochtones, pour les personnes autochtones, avec des autochtones et en tant qu'autochtones; et où il n'y a pas de rôle tenu pour les non-autochtones, sauf sur invitation » (Wehipeihana, 2019).

Exemple II – Évaluation autochtone

Il s'agit d'étudier la participation des Maoris à un programme sportif. Un critère classique consisterait à compter le nombre de Maoris dans ce programme. Mais l'équipe d'évaluation cherche à développer des critères prenant en compte la participation « en tant qu'autochtones » : Y avait-il une gouvernance, une gestion, un personnel maori? Parlait-on maori? La mise en œuvre intégrait-elle des pratiques maories? Les activités sportives avaient-elles lieu dans des endroits ayant une signification culturelle pour les Maoris? Les activités sportives elles-mêmes étaient-elles maories? Ces critères sont établis par un processus de dialogue et de délibération avec les opérateurs maoris du programme, et servent à s'accorder sur ce que signifie participer « en tant que Maori » à des activités sportives et récréatives (Wehipeihana et McKegg, 2018).

Une évaluation est engagée au Sénégal avec de jeunes évaluatrices et évaluateurs. Ceux-ci engagent deux approches de collecte de données adaptées à la culture et de nature autochtone. La première, dite « méthodologie Ataya », est un groupe de discussion qui tourne autour de la préparation et de la consommation du thé, un moment très important dans la vie quotidienne des Sénégalaises et des Sénégalais. La seconde est une approche de collecte de données dans laquelle les jeunes évaluatrices et évaluateurs recueillent des données qualitatives auprès d'autres jeunes en utilisant les maximes de Kocc Barma, un philosophe cayorien (avant qu'il ne soit « le Sénégal »).¹⁰

10. Cette expérience nous a été relatée par Thomas G. Archibald, que nous remercions.

L'évaluation émancipatrice

Cette approche (*Empowerment Evaluation*) peut être définie comme « le recours à des concepts, techniques et constats issus de l'évaluation pour améliorer [les interventions] et favoriser l'autodétermination » (Fetterman et Wandersman, 2007). L'évaluation émancipatrice est nécessairement une démarche participative. Pourtant, malgré son nom, elle n'est pas nécessairement transformationnelle : elle peut être employée dans un but très pratique de renforcement des capacités d'acteurs impliqués dans un projet, ou pour intégrer l'évaluation dans la routine organisationnelle d'une institution, par exemple. Nous choisissons cependant ici de l'aborder sous l'angle de la justice sociale et de sa promesse : redonner du pouvoir d'agir aux parties prenantes. L'évaluation émancipatrice cherche alors à créer les conditions pour qu'un groupe de personnes, qui peuvent être des travailleurs ou travailleuses de terrain, des destinataires d'une intervention, des groupes marginalisés, engagent une transformation psychologique qui va leur permettre de s'approprier les interventions et de se donner les moyens d'en prendre les rênes.

Il existe toute une famille d'évaluations « de type *empowerment* » (Cousins et Whitmore, 1998; Ridde *et al.*, 2005). Dans la version proposée par David Fetterman, l'évaluation émancipatrice démarre par une étape lors de laquelle le groupe impliqué va s'accorder sur les finalités et les missions de l'intervention évaluée. Les participants et participantes font ensuite un état des lieux de la situation actuelle : où en est-on dans la mise en œuvre et les effets de l'intervention? L'approche recommande le recours à une notation de 1 à 10, qui servira de niveau de référence pour le futur; les participants et participantes doivent engager une négociation pour s'accorder sur cette note, et utiliser des données (ou faire appel à des faits vérifiables) pour appuyer leur opinion. Cette étape peut se répéter à différents stades de l'avancement du programme, permettant ainsi de s'entendre sur l'évolution du programme. La troisième étape donne aux participants et aux participantes la possibilité de se projeter dans le futur, en établissant de nouveaux objectifs, des stratégies pour les atteindre, et les informations à collecter pour les vérifier (Bélanger, 2011).

Dans l'évaluation émancipatrice, l'évaluatrice ou l'évaluateur est un « ami critique ». Elle ou il accompagne, soutient, forme selon les besoins, mais ne détermine pas le processus. Surtout, l'équipe d'évaluation n'organise ni ne mène les investigations, les participants et participantes étant amenés à décider totalement ce dont elles ou ils ont besoin pour leur initiative, ce qui constitue des indices acceptables de la réussite ou de l'échec des interventions.

Exemple 12 – Évaluation de type émancipatrice

Une expérimentation est engagée en France auprès de l'Inspection du travail. Celle-ci vise d'une part à changer les pratiques des inspectrices et des inspecteurs, qui d'ordinaire travaillent seuls, pour passer à une pratique collective; et d'autre part à les amener à évaluer leurs actions. Pour ce faire, le ministère étudie initialement la mise en place d'indicateurs quantitatifs. Cependant, pas question d'envisager une évaluation extérieure de l'Inspection du travail, dont l'autonomie est protégée par une convention internationale, et qui voit dans l'évaluation une tentative de leur hiérarchie de s'immiscer dans leurs pratiques. La solution est de concevoir une démarche mettant les membres de l'Inspection en situation de construire et de mettre en œuvre leurs propres évaluations. Ils sont accompagnés pour cela par un évaluateur qui les forme, les accompagne collectivement dans l'élaboration de leur théorie du changement, de leurs questions d'évaluation et de leurs outils de collecte. Les évaluations proprement dites, cependant, sont menées de bout en bout par les membres de l'Inspection. Outre les apports de la démarche en matière de réflexion collective sur les actions engagées, l'évaluation offre la possibilité aux personnes impliquées de se pencher sur leurs pratiques professionnelles; elle leur donne aussi des compétences et une expérience leur permettant de s'affirmer face aux directives du ministère en matière d'évaluation (Delahais, 2021).

Déterminer l'approche qui convient : quels critères de choix ?

Nous avons dit précédemment que, pour toute évaluation, il fallait déterminer une démarche évaluative ad hoc, et que celle-ci pouvait être élaborée en s'appuyant sur une ou plusieurs approches, elles-mêmes adaptées au contexte d'intervention. Une bonne démarche évaluative dépend ainsi tout autant de l'intervention évaluée, de ce qui est attendu de l'évaluation ainsi que du contexte et des conditions de l'évaluation, mais aussi de ce que les parties prenantes impliquées jugent comme étant une bonne stratégie. Faut-il dès lors redire qu'il n'y a pas une hiérarchie des approches gravée dans le marbre? Plutôt, nous considérerons que toute approche évaluative a des caractéristiques qui lui sont propres, qui lui permettent de briller dans certains contextes ou d'apparaître foncièrement inadaptée dans d'autres.

Pour élaborer une démarche évaluative adaptée aux enjeux, faisable et utile, on pourra ainsi se poser les questions rassemblées dans le tableau 2.

Tableau 2 — Questions clés pour établir une bonne démarche évaluative

L'approche est-elle adaptée à l'intervention évaluée?	Quelle est la nature des instruments évalués? L'intervention est-elle innovante, expérimentale, ou au contraire éprouvée? Quels sont les publics concernés, leur nombre, le contexte d'intervention, le type de conséquences attendues et les chemins par lesquels ces impacts peuvent être obtenus, etc.
L'approche répond-elle adéquatement aux questions d'évaluation?	Les questionnements relèvent-ils de l'impact, ou d'autres registres (efficacité, pertinence, cohérence...)? L'approche doit-elle permettre l'émergence des questionnements?
L'approche fournit-elle le type de processus ou de résultats qui rendra l'évaluation utile?	Les attentes portent-elles sur des éléments de description ou d'explication causale? Des conclusions ou des recommandations sont-elles attendues? Y a-t-il des attentes en termes d'inclusion, de participation, d'expertise à prendre en compte?
L'approche peut-elle être crédible dans le contexte où elle sera mise en œuvre? (Hurteau <i>et al.</i> , 2012) ¹¹	Les commanditaires et les parties prenantes partagent-elles les valeurs exprimées dans l'approche envisagée? S'accordent-elles sur la façon dont l'évaluation devrait être menée, ce à quoi elle devrait servir? Qu'est-ce qui apparaît dans ce cadre comme une preuve crédible?
L'approche est-elle faisable en contexte?	Peut-elle être menée dans les contraintes de temps, de budget et de données disponibles qui sont propres à chaque évaluation (Bamberger <i>et al.</i> , 2021)?
L'équipe d'évaluation est-elle en mesure de mener cette approche?	L'équipe d'évaluation se sent-elle en mesure de la mener à bien? Fait-elle partie de sa zone de confort (en termes de valeurs, de rapport à la connaissance, d'éthique...)? A-t-elle les compétences techniques, sociales et culturelles nécessaires?

On ajoutera que tout cela définit une bonne démarche évaluative a priori, car en réalité, les équipes d'évaluation, leurs commanditaires, voire les autres parties prenantes font des choix tout au long d'une évaluation pour la mener à bien, pour la rendre plus utile, pour s'adapter à des circonstances existantes... sa qualité effective ne peut ainsi être jugée qu'a posteriori. En effet, il est difficile d'imaginer exécuter *fidèlement* une approche dans toutes ses composantes, dans la mesure où celles-ci sont mouvantes. Certaines approches comportent une dimension méthodologique très prescriptive, dont le suivi méticuleux est essentiel pour la validité des résultats obtenus. C'est par exemple le cas pour l'essai contrôlé randomisé, qui fournit des règles strictes sur la composition d'un échantillon ou la tenue d'une expérimentation. Même là, cependant, les marges de manœuvre sont importantes (Jatteau, 2016), et il est utile de conserver une attitude réflexive sur les choix effectués pour s'assurer que l'on fait les meilleurs choix pour l'évaluation. Ainsi, selon le contexte, la capacité d'adaptation d'une stratégie évaluative sera également un critère de qualité à prendre en compte. Dans tous les cas, les équipes d'évaluation seront amenées à faire des choix et pourront s'appuyer sur les approches, surtout lorsque celles-ci expriment clairement des principes, une éthique ou des valeurs pour les guider.

Parce qu'une bonne démarche évaluative dépend d'autant de critères, les choix d'approches qui peuvent être effectués pour l'alimenter sont en réalité limités. De nombreuses personnes préfèrent s'abstraire de ces questions en ne s'appuyant que sur une seule approche, ou sur un petit nombre de méthodes, qui coïncident avec leur croyance en ce qu'une évaluation devrait être, et plus globalement avec leur zone de confort. Des équipes d'évaluation vont ainsi cadrer une évaluation d'une façon qui

11. Toutes les approches que nous proposons ici s'appuient sur des méthodologies éprouvées, mais celles-ci ne sont qu'un élément de la crédibilité (Hurteau *et al.*, 2012).

s'inscrit dans l'approche qu'elles préfèrent, ou ne répondre qu'à des commandes qui s'y inscrivent. Des commanditaires posent leurs questions d'évaluation en fonction de ce à quoi pourra répondre l'approche qui leur est la plus familière, ou qui leur semble la plus crédible.

Dans l'ensemble, gageons que cela ne prêle pas réellement à conséquence. Néanmoins, il est utile de se donner, en début d'évaluation, la possibilité de sortir de la routine en étudiant comment plusieurs approches évaluatives pourraient s'appliquer à une évaluation, et définir ainsi un « champ des possibles ». Cet éventail d'approches pouvant être déployées permet à la fois de mieux répondre à une commande évaluative; mais aussi de réinterroger cette commande, en offrant d'autres façons de mener l'évaluation que celles que l'on envisageait initialement. L'objectif n'est pas forcément d'aboutir à l'approche idéale, mais au meilleur compromis, fait consciemment et dans un cadre précis, pour réaliser une évaluation.

Conclusion : élargir sa zone de confort

Nous avons, dans les pages précédentes, passé en revue 14 approches et familles d'approches. Certaines sont très semblables à une méthodologie; d'autres touchent à la posture et à la relation au monde; toutes, car c'était un élément essentiel de choix ici, sont utilisées pratiquement par des équipes d'évaluation pour élaborer leurs stratégies évaluatives et guider leurs efforts au quotidien. Cette liste est loin d'être définitive : certaines tomberont peut-être en désuétude dans les prochaines années; d'autres propositions sont peut-être en voie de devenir des approches à part entière.

Chacun est-il amené, dans sa carrière, à utiliser toutes ces approches? Sans doute non : nous avons toutes et tous une « zone de confort », délimitée par nos valeurs, nos habitudes, ce que nous percevons de nos compétences et du champ des possibles en matière d'évaluation. Mais il est possible d'élargir cette zone de confort, en repérant d'autres approches que celles que nous utilisons habituellement; de se redonner une certaine liberté de choix, en considérant plus systématiquement plusieurs stratégies évaluatives; d'hybrider et de sophistiquer nos pratiques, s'inspirer d'autres réflexions et d'autres pratiques que les siennes. C'est ce à quoi ce chapitre aspire.

Messages clés

- Les approches sont des ensembles de propositions relatives à ce qui constitue une bonne évaluation. Elles combinent de nombreux éléments, et en particulier un cadre paradigmatique, des promesses en termes de résultats et d'usages potentiels, des prescriptions méthodologiques ainsi que des préconisations pratiques permettant de les mettre en œuvre;
- Les approches sont comme des blocs de construction d'une évaluation. S'appuyer sur des approches lors de l'élaboration d'une démarche évaluative ouvre le champ des possibles en matière d'évaluation, permet d'apporter de la crédibilité méthodologique et de mieux se comprendre entre parties prenantes de l'évaluation;
- Il existe de nombreuses approches, qui correspondent à des contextes et des attentes différentes, mais aussi à la vision du monde et aux valeurs des différentes parties prenantes, évaluateurs et évaluatrices y compris. Il faut donc choisir en conséquence celles qui conviennent.

Références bibliographiques

- Abma, T. A. (1998). Storytelling as Inquiry in a Mental Hospital. *Qualitative Health Research*, 8(6), 821-838. <https://doi.org/10.1177/104973239800800608>
- Alkin, M. C. (2013). *Evaluation roots : A wider perspective of theorists' views and influences* (2^e éd.). SAGE Publications.
- Bamberger, M., Rugh, J., Church, M. et Fort, L. (2021). L'évaluation en situation réelle : concevoir des évaluations d'impact sous contraintes de budget, de temps et de données (trad., publication originale en 2004). Dans T. Delahais, A. Devaux-Spatarakis, A. Revillard, et V. Ridde (dir.), *Évaluation. Fondements, Controverses et Perspectives*. Éditions science et bien commun. <https://scienceetbiencommun.pressbooks.pub/evaluationanthologie/>
- Barret, D., Blundo Canto, G., Dabat, M.-H., Devaux-Spatarakis, A., Faure, G., Hainzelin, E., Mathé, S., Temple, L., Toillier, A. et Triomphe, B. (2017). *Guide méthodologique ImpresS. Évaluation ex post des impacts de la recherche agronomique dans les pays du Sud*. Cirad.
- Behaghel, L., Crépon, B., Gurgand, M., Kamionka, T., Lequien, L., Rathelot, R. et Zamora, P. (2013). L'accompagnement personnalisé des demandeurs d'emploi. *Revue française d'économie*, 28(1), 123-158. <https://doi.org/10.3917/rfe.131.0123>
- Befani, B. et Mayne, J. (2014). Process Tracing and Contribution Analysis : A Combined Approach to Generative Causal Inference for Impact Evaluation. *IDS Bulletin*, 45(6), 17-36. <https://doi.org/10.1111/1759-5436.12110>
- Befani, B. (2020). *Choosing Appropriate Evaluation Methods Tool* (version 2). CECAN.
- Belcher, B. M., Davel, R. et Claus, R. (2020). A refined method for theory-based evaluation of the societal impacts of research. *MethodsX*, 7. <http://dx.doi.org/10.1016/j.mex.2020.100788>
- Bélanger, J.-F. (2011). L'évaluation habilitative. Dans P. M. Daigneault (dir.), *Les approches théoriques en évaluation*. Perfeval.
- Brousselle, A. et Buregeya, J.-M. (2018). Theory-based evaluations : Framing the existence of a new theory in evaluation and the rise of the 5th generation. *Evaluation*, 24(2), 153-168. <https://doi.org/10.1177/1356389018765487>
- Chen, H. T. et Rossi, P. H. (1987). The theory-driven approach to validity. *Evaluation and Program Planning*, 10(1), 95-103. [https://doi.org/10.1016/0149-7189\(87\)90025-5](https://doi.org/10.1016/0149-7189(87)90025-5)
- Copestake, J. G., Morsink, M. et Remnant, F. (2019). *Attributing development impact : The Qualitative Impact Protocol (QuIP) case book*. Practical Action Publishing.
- Cousins, J. B. et Whitmore, E. (1998). Framing participatory evaluation. *New Directions for Evaluation*, 1998(80), 5-23. <https://doi.org/10.1002/ev.1114>
- Dagenais, C. et Ridde, V. (2012). *Approches et pratiques en évaluation de programmes*. Les Presses de l'Université de Montréal.
- Delahais, T. et Toulemonde, J. (2017). Making rigorous causal claims in a real-life context : Has research contributed to sustainable forest management? *Evaluation*, 23(4), 370-388. <https://doi.org/10.1177/1356389017733211>

- Delahais, T., Sage, K. et Honoré, V. (2020). Evaluators in Transition. *Zeitschrift für Evaluation (ZfEv)*, 2020(2), 239-260. DOI :10.31244/zfe.2020.02.03
- Delahais, T. (2021). Les évaluations participatives : Plus facile à dire qu'à faire? *LeGes*, 32(1). <https://doi.org/10.38023/265714d6-463d-4ae3-80ef-5588489d0ec3>
- deMarrais, K. B et Lapan, S. D. (dir.). (2004). *Foundations for research : Methods of inquiry in education and the social sciences*. Routledge –Taylor & Francis Group.
- Demarteau, M. (2002). A theoretical framework and grid for analysis of programme-evaluation practices. *Evaluation*, 8(4), 454-473. <https://doi.org/10.1177/13563890260620649>
- Devaux-Spatarakis, A., Harmach, S., Alouis, J. et Sage, K. (2021). Concilier les différents intérêts d'apprentissage d'une expérimentation : retour d'expérience sur le rôle de l'évaluateur pour favoriser les usages de l'évaluation. Dans A. Kerivel et S. James (dir.), *Comment évaluer des projets innovants encore en construction? INJEP Notes & Rapports 2021/14*.
- Dozois, E., Langlois, M. et Blanchet-Cohen, N. (2011). *DE 201 : Guide du praticien de l'évaluation évolutive*. Fondation de la famille J.W. McConnell et l'Institut international des droits de l'enfant et du développement.
- Earl, S., Carden, F. et Smutylo, T. (2001). *Outcome Mapping : Building Learning and Reflection into Development Programs*. IDRC.
- Fetterman, D. et Wandersman, A. (2007). Empowerment Evaluation : Yesterday, Today, and Tomorrow. *American Journal of Evaluation*, 28(2), 179-198. <https://doi.org/10.1177/1098214007301350>
- Frehiwot, M. (2019). *Made in Africa Evaluation : Decolonizing evaluation in Africa* (third quarter 2019). EvalMatters. <http://idev.afdb.org/sites/default/files/Evaluations/2020-03/article%203%20Decolonizing%20evaluation%20in%20Africa.pdf>
- Funnell, S. C. et Rogers, P. J. (2011). *Purposeful program theory : effective use of theories of change and logic models*. Jossey-Bass.
- Gibson, R. et Robichaud, S. (2020). Evaluating Dancing with Parkinson's : Reflections from the perspective of a community organization. *Evaluation and Program Planning*, 80(101449). <https://doi.org/10.1016/j.evalprogplan.2017.05.010>
- Guba, E. G. et Lincoln, Y. S. (1989). *Fourth Generation Evaluation*. SAGE Publications.
- Harji, K. et Jackson, E. T. (2016). *Developmental evaluation in practice : Lessons from evaluating a market-based employment initiative*. The Rockefeller Foundation. <https://www.issuelab.org/resources/27188/27188.pdf>
- Hood, S., Hopson, R. K. et Kirkhart, K. E. (2015). Culturally responsive evaluation. Dans E. K. Newcomer, P. H. Hatry et S. J. Wholey (dir.), *Handbook of practical program evaluation*. Wiley. <https://doi.org/10.1002/9781119171386.ch12>
- House, E. et Howe, K. (2003). Deliberative Democratic Evaluation. Dans T. Kellaghan et D. L. Stufflebeam (dir.), *International Handbook of Educational Evaluation* (p. 79-100). Springer.
- Hurteau, M., Houle, S. et Guillemette, F. (dir.). (2012). *L'évaluation de programme axée sur le jugement crédible*. Presses de l'Université du Québec.

- Jatteau, A. (2020). *Faire preuve par le chiffre? Le cas des expérimentations aléatoires en économie*. Institut de la gestion publique et du développement économique.
- Lay, M. et Papadopoulos, I. (2007). An Exploration of Fourth Generation Evaluation in Practice. *Evaluation*, 13(4), 495-504. <https://doi.org/10.1177/1356389007082135>
- Mayne, J. (2012). Contribution analysis : Coming of age? *Evaluation*, 18(3), 270-280. <https://doi.org/10.1177/1356389012451663>
- Mertens, D.M. et Wilson, A.T. (2012). *Program Evaluation Theory and Practice : A Comprehensive Guide*. Guilford Press.
- Mertens, D. (2021). La recherche transformationnelle : personnelle et sociétale (trad., publication originale en 2017). Dans T. Delahais, A. Devaux-Spatarakis, A. Revillard, et V. Ridde (dir.), *Évaluation. Fondements, Controverses et Perspectives*. Éditions science et bien commun. <https://scienceetbiencommun.pressbooks.pub/evaluationanthologie/>
- Miller, R.L. (2010). Developing Standards for Empirical Examinations of Evaluation Theory. *American Journal of Evaluation*, 31(3), 390-399. <https://doi.org/10.1177/1098214010371819>
- ONU Femmes (2015). *Gestion des évaluations sensibles au genre*. <https://www.unwomen.org/sites/default/files/Headquarters/Attachments/Sections/Library/Publications/2015/UN-Women-Evaluation-Handbook-fr.pdf>
- Patton, M.Q. (2008). *Utilization-focused evaluation* (4^e éd.). SAGE Publications.
- Patton, M.Q. et LaBossière, F. (2012). L'évaluation axée sur l'utilisation. Dans C. Dagenais et V. Ridde (dir.), *Approches et pratiques en évaluation de programmes* (p. 145-160). Presses de l'Université de Montréal.
- Pawson, R. et Tilley, N. (1997). *Realistic evaluation*. SAGE Publications.
- Podems, D. (2014). Feminist Evaluation for non-feminists. Dans S. Brisolaro, D.M. Seigart et S. SenGupta (dir.), *Feminist evaluation and research : Theory and practice*. The Guilford Press.
- Podems, D. (2018). *Rendre l'évaluation féministe praticable* (4^e trimestre, 2018). Evaluation Matters. <https://idev.afdb.org/sites/default/files/Evaluations/2020-03/Rendre%20l%27%20evaluation%20feministe%20praticable.pdf>
- Quadrant Conseil. (2017). *Comment évaluer l'impact?* <https://www.quadrant-conseil.fr/ressources/ArbreImpact.html>
- Ridde, V., Baillargeon, J., Ouellet, P. et Roy, S. (2005). L'évaluation participative de type empowerment : Une stratégie pour le travail de rue. *Service social*, 50(1), 263-279. <https://doi.org/10.7202/009972ar>
- Robert, É. et Ridde, V. (2013). L'approche réaliste pour l'évaluation de programmes et la revue systématique : De la théorie à la pratique. *Mesure et évaluation en éducation*, 36(3), 79-108. <https://doi.org/10.7202/1025741ar>
- Shadish, W. R., Cook, T. D. et Campbell, D.T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton Mifflin Company.

- Sielbeck-Mathes, K. et Selove, R. (2014). An Explication of Evaluator Values : Framing Matters. Dans S. Brisolaro, D. Seigart et S. SenGupta (dir.), *Feminist evaluation and research : Theory and practice*. Guilford Publications.
- Stake, R. (2003). Responsive Evaluation. Dans T. Kellaghan et D.L. Stufflebeam (dir.), *International Handbook of Educational Evaluation* (p. 63-68). Springer.
- Stufflebeam, D. L. et Coryn, C.L.S. (2012). *Evaluation Theory, Models, and Applications* (2^e éd.). Jossey-Bass.
- Turcotte, P.-L., Rey, L. et Brousselle, A. (2021). L'évaluation évolutive, de la théorie à la pratique : Perspectives de praticiens québécois. *Canadian Journal of Program Evaluation*, 36(1). <https://doi.org/10.3138/cjpe.70502>
- Vaessen, J., Lemire, S. et Befani, B. (2020). *Evaluation of International Development Interventions : An Overview of Approaches and Methods*. World Bank.
- Wehipeihana, N. et McKegg, K. (2018). Values and Culture in Evaluative Thinking : Insights from Aotearoa New Zealand : Values and Culture in Evaluative Thinking. *New Directions for Evaluation*, 2018(158), 93-107. <https://doi.org/10.1002/ev.20320>
- Wehipeihana, N. (2019). Accroître la compétence culturelle, une étape nécessaire en appui à une évaluation menée par les personnes autochtones. Dans T. Delahais, A. Devaux-Spatarakis, A. Revillard et V. Ridde (dir.), *Évaluation. Fondements, Controverses et Perspectives*. Éditions Sciences du Bien Commun.
- Wenger-Trayner, E. et Wenger-Trayner, B. (2015, 15 avril). *Communities of practice : A brief introduction*. Wenger-Trayner. <https://wenger-trayner.com/wp-content/uploads/2015/04/07-Brief-introduction-to-communities-of-practice.pdf>
- White, H. (2009). *L'évaluation d'impact basée sur la théorie : Principes et pratique*. Wageningen University & Research.
- White, H. (2014). *Essais contrôlés randomisés (ECR)* (Note méthodologique n° 7). UNICEF – Centre de recherche Innocenti.
- White, H. et Sabarwal, S. (2014). *Méthodes et modèles quasi expérimentaux* (Note méthodologique n° 8). UNICEF – Centre de recherche Innocenti.
- Wilson-Grau, R. et Britt, H. (2012). *Outcome harvesting*. Ford Foundation.