

7

Les méthodes mixtes

Pierre Pluye

Les méthodes mixtes de recherche et d'évaluation combinent les méthodes de collecte et d'analyse de données qualitatives et quantitatives. Elles sont de plus en plus fréquemment utilisées et, au fil des ans, plusieurs termes ont été proposés pour les nommer (par exemple, multiméthodes). Johnson (2007) a effectué une synthèse critique de la littérature scientifique sur les méthodes mixtes et a proposé une définition de ces méthodes pour la recherche. Cette recension des écrits suggère de définir les méthodes mixtes pour l'évaluation de programmes de la manière suivante. Une évaluation mixte est un type d'évaluation dans laquelle un expert ou une équipe d'experts combinent les approches ou devis méthodologiques avec les techniques de collecte et d'analyse de données qualitatives et quantitatives pour approfondir la compréhension, ou interpréter, ou corroborer des résultats d'évaluation.

Les trois principales raisons de combiner méthodes qualitatives et quantitatives sont les suivantes. Les évaluateurs peuvent avoir besoin des méthodes qualitatives pour interpréter des résultats quantitatifs. Ils peuvent avoir besoin des méthodes quantitatives pour généraliser des résultats qualitatifs. Enfin, ils peuvent avoir besoin d'explorer ou de mieux comprendre un phénomène nouveau (méthodes qualitatives) et d'en mesurer l'ampleur, l'évolution, les causes et les effets (méthodes quantitatives).

Nous remercions Lucie Nadeau, Marie-Pierre Gagnon, Roland Grad, Janique Johnson-Lafleur et Frances Griffiths pour leur contribution à une version précédente de ce chapitre (2009).

Dans le tableau 7.1, nous présentons une liste non exhaustive de méthodes qualitatives et quantitatives susceptibles d'être combinées sous forme de méthodes mixtes. Les lecteurs qui ne les connaissent pas bien peuvent se référer aux dictionnaires de Muchielli (2004) pour les méthodes qualitatives et de Porta (2008) pour les méthodes quantitatives.

TABLEAU 7.1**Exemples de méthodes qualitatives et quantitatives**

Méthodes qualitatives	Méthodes quantitatives
Communication et information <ul style="list-style-type: none"> Analyse de discours Idéaux-types et « patterns » <ul style="list-style-type: none"> Analyse des événements Théorisation ancrée Étude de cas qualitative Actions sociales <ul style="list-style-type: none"> Herméneutique Histoires de vie Phénoménologie Société et interactions sociales <ul style="list-style-type: none"> Ethnographie Ethnométhodologie 	Études randomisées <ul style="list-style-type: none"> Essai contrôlé randomisé Études non randomisées <ul style="list-style-type: none"> Essai contrôlé Étude de cohorte Étude de comparaison de cas Étude pré-post / séries temporelles Enquête transversale (analytique) Études descriptives <ul style="list-style-type: none"> Étude de prévalence ou d'incidence (sans groupe de comparaison) Série de cas Rapport de cas

À titre d'illustration, une méthode mixte d'évaluation des services de santé mentale en milieu multiethnique pourrait combiner une enquête quantitative transversale (questionnaire structuré) avec une étude ethnographique (documents, entrevues et observations), permettant ainsi de mieux comprendre les résultats de l'enquête et d'élaborer un nouveau programme pour faciliter l'accès aux services (Groleau *et al.*, 2007). Dans cet exemple, l'enquête seule aurait suffi pour révéler les problèmes d'accessibilité, mais elle aurait été insuffisante pour permettre de créer un programme approprié. Comme l'indiquent Sharp et Frechtling (1997), « en se concentrant principalement sur des techniques quantitatives, les experts peuvent passer à côté de certaines parties importantes de l'histoire. Les évaluateurs expérimentés ont maintes fois constaté que de meilleurs résultats sont obtenus en utilisant des méthodes mixtes d'évaluation » (p. 2). De fait, on a depuis longtemps combiné ces méthodes en recherche et en évaluation, et cette combinaison peut même être vue comme inévitable dans le domaine de l'évaluation (Rallis et Rossman, 2003, p. 491). Mais ce n'est que récemment que ce type de combinaisons

a été conceptualisé en termes de méthodes mixtes (Creswell et Plano Clark, 2010 ; Greene, 2006 ; Johnson, 2007 ; Tashakkori et Teddlie, 2010).

En général, les méthodes mixtes sont utilisées en vue de combiner les forces et compenser les limites respectives des méthodes quantitatives et qualitatives. D'une part, les évaluations quantitatives examinent habituellement des associations de facteurs (par exemple, des causes et leurs effets) pouvant être généralisées à toute une population, notamment grâce à des inférences statistiques. D'autre part, les évaluations qualitatives fournissent des descriptions détaillées de phénomènes complexes fondées empiriquement sur un contexte spécifique pouvant être théoriquement applicables dans d'autres situations (par exemple, une étude ethnographique). En particulier, la partie qualitative d'une évaluation mixte peut fournir une meilleure compréhension du développement des programmes complexes (planification, implantation et pérennisation), et une compréhension approfondie des raisons pour lesquelles ces programmes fonctionnent bien dans certains contextes (ou ne fonctionnent pas dans d'autres situations).

Ce chapitre vise à présenter notre conception des méthodes mixtes, puis à décrire les approches, devis et techniques de collecte et d'analyse de données qualitatives et quantitatives qui sont utiles pour planifier ou implanter des méthodes mixtes pour l'évaluation des programmes. Une illustration sera tirée de notre travail en évaluation des technologies. Dans les sections finales, nous proposerons des critères de qualité pour planifier ou évaluer les méthodes mixtes, et pour effectuer une recension des écrits mixte pour l'évaluation des programmes.

Les méthodes mixtes et les débats entre méthodes quantitatives et qualitatives

La méthodologie des méthodes mixtes peut être conceptualisée à partir des débats méthodologiques entre les tenants des méthodes quantitatives et ceux des méthodes qualitatives. Ces débats suggèrent deux idéaux-types méthodologiques, souvent considérés comme deux paradigmes concurrents, qui sont respectivement associés aux méthodes quantitatives et qualitatives : l'empirisme logique (matérialisme, réalisme et arguments objectifs habituellement associés aux méthodes quantitatives) et le constructivisme (idéisme, relativisme et arguments subjectifs habituellement associés aux méthodes qualitatives). Un paradigme est défini comme étant une norme (ou un

standard) institutionnelle des sciences. Bien que selon Kuhn (1983), l'évolution des sciences soit caractérisée par des transitions successives d'un paradigme à l'autre, des révolutions, de nombreux historiens ou philosophes des sciences observent que plusieurs paradigmes ou « visions du monde » coexistent (Morgan, 2007). Par exemple, l'empirisme logique, le paradigme dominant des disciplines scientifiques au xx^e siècle, est contesté par les tenants du constructivisme (Denzin et Lincoln, 1994), et coexiste avec le constructivisme. Dans cette section, nous allons brièvement présenter ces deux idéaux-types, leurs principales forces et faiblesses en évaluation de programmes, et la façon dont ces idéaux-types permettent de conceptualiser les méthodes mixtes. Le terme « idéal », dans le concept (outil d'analyse) de l'idéal-type proposé par Max Weber, désigne l'idée, et un idéal-type est le groupement d'un certain nombre de caractéristiques typiques des phénomènes sociaux qui a pour but d'aider à la compréhension de ces derniers.

L'empirisme logique vise à étudier des phénomènes selon des lois empiriques, par exemple la force de l'association entre des facteurs importants de l'implantation des programmes (mesurés en valeurs de variables indépendantes) et les résultats observés en termes d'implantation (mesurés en valeurs de variables dépendantes). Les relations causales constituent un exemple de ces lois empiriques et sont présentées en détail au chapitre 6, intitulé « Attribution et causalité des effets d'un programme ». L'empirisme logique dérive du positivisme proposé au xix^e siècle, intègre les critiques postpositivistes du xx^e siècle et propose habituellement (mais pas nécessairement) que les lois empiriques soient informées par une théorie qui préexiste à la démarche de recherche (approche déductive ou confirmatoire). Par exemple, les empiristes logiques assument habituellement (mais encore là, pas nécessairement) qu'il n'y a pas de causalité sans organisation normative préalable du champ d'exploration empirique par une explication scientifique logique. L'étude de la causalité est alors fondée sur des hypothèses formulées à l'aide de relations conditionnelles (langage de la science selon l'empirisme logique)¹.

En évaluation de programmes, des biais peuvent influencer la validité des études quantitatives expérimentales ou quasi expérimentales. Les

1. Nous remercions François Béland, de l'Université de Montréal, pour ses commentaires sur l'empirisme logique. Ce paragraphe ne rend évidemment pas compte de tous les travaux des empiristes logiques, et notamment de ceux qui ont développé une logique « inductive ».

sources des biais sont plus ou moins contrôlées selon les devis utilisés, les études randomisées étant considérées comme les plus valides pour examiner des relations causales (plus valides que les études non randomisées, par exemple). Les études randomisées ont néanmoins leurs propres faiblesses (Mays *et al.*, 2005 ; Pawson, 2005). Ainsi, une évaluation avec essai randomisé contrôlé peut être inadéquate selon le contexte ou la question, et elle peut être contredite par des études non randomisées. Par exemple, les études randomisées sont moins appropriées que les études de cohorte pour évaluer les rares effets négatifs à long terme des programmes (la randomisation de milliers de personnes peut être nécessaire pour détecter ces effets, mais elle risque d'être irréalisable ou trop coûteuse). De plus, les programmes peuvent être conceptualisés en tant qu'interventions sociales complexes agissant sur des systèmes sociaux complexes. Or, les études randomisées fournissent habituellement peu ou pas de résultats empiriques expliquant pourquoi les effets des interventions varient selon différents contextes d'implantation.

Quant au constructivisme, il vise à découvrir et à explorer des phénomènes complexes, à décrire leurs variations selon leur contexte et à en proposer des interprétations différentes pour mieux les comprendre (approche inductive ou exploratoire). Le constructivisme dérive des approches dites compréhensives développées en Allemagne dans les années 1920 par Weber, et à Chicago, par les fondateurs d'une approche sociologique américaine appelée « interactionnisme symbolique ». Les chercheurs qualitatifs soulignent la nature socialement construite de la réalité, la relation intime entre le chercheur, le phénomène étudié et leur contexte (Denzin et Lincoln, 1994). Ils examinent les aspects historiques, interprétatifs et narratifs en fonction de leur contexte de production culturel, économique et social (Abbott, 1998).

Les méthodes qualitatives ont été développées de manière concomitante dans plusieurs disciplines et utilisent différentes procédures interprétatives (par exemple, l'ethnographie en anthropologie et les ethnométhodes en sociologie). Les méthodes qualitatives ont du succès, en évaluation de programmes, car elles satisfont les besoins des décideurs et des professionnels en termes de compréhension riche et détaillée des interventions sociales complexes (Patton, 2002), compréhension utile pour la planification, l'implantation et la pérennisation des programmes. Plus particulièrement, les études qualitatives fournissent des résultats empiriques expliquant pourquoi

les effets des interventions varient selon les différents contextes d'implantation. Cependant, les méthodes qualitatives sont critiquées parce que les résultats peuvent être idiosyncrasiques à un contexte spécifique (aucune connaissance transférable à un autre contexte), ou fondés sur des procédures interprétatives tacites, non reproductibles et non vérifiables (Huberman et Miles, 1991).

Les méthodes mixtes questionnent ces idéaux-types, et suggèrent des types méthodologiques différents. Par exemple, les études quantitatives exploratoires ne correspondent pas à l'idéal-type méthodologique de l'empirisme logique. Les résultats statistiques ou épidémiologiques peuvent être considérés comme des constructions sociales utiles pour explorer et comprendre des phénomènes sociaux (constructionnisme ou réalisme constructif). Ainsi, Kallerud et Ramberg (2002) favorisent l'interprétation constructiviste des enquêtes. Autre exemple, les études qualitatives inductives-déductives ne correspondent pas à l'idéal-type méthodologique du constructivisme. Les résultats qualitatifs sont souvent intégrés dans des recherches examinant des relations causales, et sont considérés comme des arguments scientifiques valides pour générer des théories ou des hypothèses (post-positivisme). Ainsi, selon Campbell (1988), une considération qualitative des différents contextes de mesure permet de proposer des hypothèses rivales plausibles.

Le débat entre méthodes quantitatives (ci-après QUANT) et méthodes qualitatives (ci-après QUAL) permet de concevoir les méthodes mixtes (ci-après MM) par rapport aux deux idéaux-types méthodologiques résumés ci-dessus. Tel que précisé en introduction de ce chapitre, en évaluation de programmes, les méthodes quantitatives et qualitatives coexistent, et cette coexistence contribue au développement des MM et, par ces dernières, au développement de combinaisons d'approches, de devis et de techniques.

Les méthodes mixtes : approches, devis et techniques de collecte et d'analyse de données

La présente section vise à présenter toutes les possibilités de combinaisons et d'intégration des approches, des devis et des techniques pour faciliter la planification et l'évaluation des MM. Tout comme Johnson (2007), nous sommes d'avis qu'il est légitime de considérer comme MM toute combi-

naison d'approches méthodologiques, de devis et de techniques de collecte ou d'analyse de données, mais uniquement lorsque ces combinaisons satisfont trois conditions: 1) au moins une méthode QUAL et une méthode QUANT sont combinées (par exemple, l'ethnographie et l'expérimentation); 2) chaque méthode est utilisée de façon rigoureuse par rapport aux critères généralement admis (voir la section suivante « Planification et qualité des MM en évaluation de programmes »); et 3) la combinaison des méthodes se traduit par une intégration des approches méthodologiques, des devis ou des techniques (par exemple, des observations et des mesures), et des données et (ou) des résultats. Inversement, selon nous, ne sont pas des MM: 1) une méthode QUANT avec collecte ou analyse de données qualitatives qui ne réfère pas à une méthode QUAL (par exemple, une étude de prévalence utilisant un questionnaire structuré auto-administré avec principalement des questions à réponse fermée, et quelques questions requérant des commentaires libres), ou vice-versa; 2) les combinaisons de méthodes QUANT; 3) les combinaisons de méthodes QUAL; et 4) les juxtapositions de méthodes QUAL et QUANT sans aucune intégration d'approches, de devis, de techniques, de données et (ou) résultats QUAL et QUANT.

Les approches

Les experts des MM comme Bryman (sociologie) et Cresswell (psychologie), ainsi que Greene, Morgan, Tashakkori et Teddlie (éducation) proposent différentes typologies des approches en MM. Nous nous inspirons ici des propositions de Teddlie et Tashakkori (2003). Évidemment, les MM sont impossibles lorsque les méthodes QUANT et QUAL sont considérées comme incompatibles ou *incommensurables* (positivisme radical ou constructivisme radical). Trois approches principales soutiennent les MM en évaluation de programmes.

L'approche *complémentaire* indique que les résultats QUAL et QUANT sont présentés séparément, mais qu'un élément QUAL contribue à un élément QUANT, ou vice-versa. Par exemple, les résultats QUAL et QUANT sont clairement distincts dans la section « Résultats » des rapports d'évaluation ou des publications, mais au moins une phrase suggère que les résultats QUAL et QUANT sont complémentaires.

L'approche par *tension dialectique* souligne les divergences toujours possibles entre les résultats QUAL et QUANT. Par exemple, des divergences

entre les résultats QUANT et QUAL sont clairement énoncées dans la section « Résultats » des rapports d'évaluation ou des publications, et au moins une phrase suggère comment résoudre ces divergences. L'approche MM par *tension dialectique* peut coexister avec la précédente. Elle postule l'existence de différentes « visions partielles du monde » (Teddlie et Tashakkori, 2003, p. 18). Trois principales stratégies sont proposées pour résoudre les divergences entre données ou résultats QUAL et QUANT : exclusion, réconciliation ou initiation (nouveau projet de recherche ou nouvelle collecte/analyse de données). Ces stratégies sont présentées ailleurs (Pluye, Grad, Levine et Nicolau, 2009).

L'approche MM par *assimilation* suggère que les résultats QUAL et QUANT soient assimilés ou transformés sous une seule forme QUAL ou QUANT. Par exemple, les fréquences d'occurrence des thèmes QUAL sont présentées, et il n'y a aucune distinction claire entre les résultats QUAL et QUANT dans la section « Résultats » des rapports d'évaluation ou des publications.

Les devis

Une classification usuelle des méthodes mixtes est fondée sur trois types de devis de recherche qui correspondent chacun à un type d'intégration des méthodes quantitatives et qualitatives (Creswell et Plano Clark, 2010). Premièrement, les devis de recherche séquentiels exploratoires qui utilisent une méthode qualitative suivie d'une méthode quantitative (QUAL puis QUANT). En termes d'intégration, les résultats quantitatifs sont mobilisés pour confirmer ou généraliser les résultats qualitatifs. Deuxièmement, les devis séquentiels explicatifs qui utilisent une méthode quantitative suivie d'une méthode qualitative (QUANT puis QUAL). Concernant l'intégration, les résultats qualitatifs sont mobilisés pour interpréter ou expliquer des résultats quantitatifs. Troisièmement, les devis concomitants combinent les méthodes quantitatives et qualitatives dès la planification de la recherche, ou au moment de la collecte et de l'analyse des données (QUANT et QUAL). Par exemple, ces méthodes peuvent étudier les mêmes phénomènes de manière complémentaire. L'intégration s'effectue au moment de la collecte et de l'analyse des données.

Ces trois types de devis MM sont proposés pour planifier des études simples plutôt que des programmes de recherche composés de plusieurs

études, mais ils peuvent certainement s'appliquer aux deux (Creswell et Plano Clark, 2010). On peut transposer cette nuance dans le domaine des évaluations de la manière suivante. Les MM en évaluation de programmes peuvent combiner des méthodes QUAL et QUANT à l'intérieur d'une étude d'évaluation unique ou dans un ensemble d'évaluations étroitement liées. À titre d'illustration, nous présentons ici un devis d'évaluation des technologies. Cet exemple montre comment combiner une évaluation QUANT longitudinale (étude de cohorte) avec une évaluation QUAL (étude de cas multiple) dans un devis MM concomitant.

Il existe aussi des variantes de ces trois devis de base, par exemple les devis multiniveaux. Le chapitre de Dagenais et Dupont (chapitre 22) qui présente l'évaluation du programme appelé « Équipe d'intervention jeunesse » en fournit une illustration. Ce programme pilote (2003-2006) comportait l'analyse de la mise en œuvre de 17 équipes de coordination pour améliorer l'accessibilité et la qualité des services pour des jeunes en difficulté (0-18 ans). Trois aspects du programme ont été évalués à deux niveaux, (a) les jeunes (1. implantation et 2. résultats), et (b) les équipes (3. processus de coordination). *Au premier niveau*, l'évaluation a utilisé un devis concomitant (triangulation). La composante QUANT consistait en analyses statistiques descriptives des données collectées pour chaque jeune ($N = 441$), dans une base de données sur internet, par les coordinateurs du programme (étude de prévalence). La composante QUAL consistait en études de cas ($N = 29$), un cas étant défini par un jeune, son entourage et les professionnels chargés de leur fournir des services. Les cas ont été sélectionnés de manière aléatoire de façon à ce qu'ils soient représentatifs de la population des 441 jeunes. Les évaluateurs ont mené 114 entrevues semi-structurées pour collecter des données, et une analyse thématique des transcriptions de ces entrevues. Les résultats QUANT suggèrent que le programme est bien implanté, atteint la clientèle appropriée et fournit comme prévu des services provenant de différentes organisations. Les résultats QUAL suggèrent que les résultats du programme sont positifs en termes de satisfaction des jeunes, de leur famille et des professionnels, par exemple, les parents ont fait part d'une amélioration de l'accès et de la qualité des services offerts. *Au deuxième niveau*, un échantillon de quatre équipes a été sélectionné de manière à refléter les différents types d'équipes en termes de composition, de fonctionnement, et de clientèle. Pour chaque équipe, des données additionnelles ont été

collectées en effectuant des entrevues individuelles avec les coordinateurs, les agents de liaison et les responsables des services, également en menant un *focus group* avec eux. Les évaluateurs ont effectué une analyse thématique des transcriptions des entrevues et des *focus groups*, tout en tenant compte des données de « premier niveau ». Les résultats suggèrent par exemple que le coordinateur avait un rôle de facilitateur central pour le bon fonctionnement des équipes.

Un exemple de devis concomitant

- *Évaluation quantitative* (Grad *et al.*, 2005). Nous avons élaboré une méthode pour évaluer systématiquement l'impact des recherches d'informations dans les ressources électroniques. Nous avons évalué une ressource, InfoRetriever, avec une cohorte de 26 résidents. InfoRetriever est un engin de recherche qui combine sept bases de données, incluant un manuel de référence, des lignes directrices, des systèmes experts, et des synopsis d'articles de recherche sélectionnés pour leur pertinence et leur validité en médecine familiale. Un questionnaire électronique d'évaluation était rempli par ces professionnels pour rapporter l'impact perçu de chaque item d'information trouvé dans InfoRetriever. Le questionnaire a permis d'évaluer 2495 items d'information. Les rapports d'impact positif les plus fréquents étaient « Ma pratique est améliorée » et « J'ai appris quelque chose de nouveau ».
- *Évaluation qualitative concomitante* (Pluye *et al.*, 2007). Notre méthode a permis d'évaluer systématiquement la pertinence situationnelle des ressources électroniques. Nous avons d'abord identifié sept raisons pour lesquelles les professionnels cherchent des informations : quatre objectifs cognitifs (par exemple, répondre à une question clinique) et trois objectifs organisationnels (par exemple, échanger des informations avec d'autres professionnels). Puis, nous avons conduit une étude de cas multiple. Les cas consistaient en des recherches effectuées dans InfoRetriever par 17 résidents pendant deux mois. Les entrevues avec ces professionnels (collecte de données qualitatives) étaient guidées par les données du questionnaire mentionné ci-dessus (collecte concomitante de données quantitatives). Notre analyse concomitante des données qualitatives et quantitatives a porté sur 156 recherches critiques (contexte et conséquences clairement décrits) liées à 877 items d'information. Pour chaque cas, les données qualitatives étaient assignées à un objectif cognitif ou organisationnel (analyse thématique). Les résidents ont atteint leurs objectifs dans 85,9 % des cas (pertinence situationnelle). Les résultats soutiennent les objectifs proposés et suggèrent des niveaux de pertinence.
- *Évaluation quantitative modifiée*. Grâce aux résultats du devis concomitant, notre nouveau questionnaire intègre les sept objectifs ci-dessus pour évaluer systématiquement l'impact et la pertinence situationnelle des informations.

Sources : Grad, R. M. *et al.* (2005), « Assessing the impact of clinical information-retrieval technology in a family practice residency », *Journal of Evaluation in Clinical Practice*, vol. 11, n° 6.

Pluye, P. *et al.* (2007), « Seven reasons why family physicians search clinical information-retrieval technology: Toward an organizational model », *Journal of Evaluation in Clinical Practice*, vol. 13, n° 1.

Les techniques

Les techniques mobilisées en MM pour évaluer des programmes peuvent s'inspirer de tous les types de combinaisons des techniques pour la collecte et l'analyse de données QUAL et QUANT. Bryman (2006) a effectué une recension des écrits couvrant 232 articles en sciences sociales qui rapportaient l'emploi de MM. Il en a conclu que l'entrevue structurée, ou questionnaire, tend à prédominer du côté des techniques QUANT (enquêtes transversales), et que l'entrevue ouverte ou semi-structurée tend à prédominer du côté des techniques QUAL. Selon leur synthèse des écrits, les principales techniques de collecte de données associées aux MM sont, par ordre alphabétique : les entrevues individuelles (non structurées, semi-structurées, structurées), les groupes de discussion, les questionnaires (avec ou sans réponses ouvertes), l'observation participante (de type ethnographique) et la recension de documents. Selon cette synthèse des écrits, les principales techniques d'analyse de données associées aux MM sont, par ordre alphabétique : l'analyse thématique qualitative, les analyses statistiques, et l'analyse de contenu ou la quantification de données qualitatives (par exemple, la quantification des réponses fournies par les interviewés).

De plus, une proportion substantielle des articles mentionne que la collecte de données QUANT et QUAL n'était pas fondée, comme on pourrait s'y attendre, sur des instrumentations distinctes. Une telle combinaison des techniques correspond au *Devis concomitant* présenté plus haut. Cette combinaison suppose par exemple que des entrevues ethnographiques peuvent être analysées en utilisant des statistiques textuelles (MM lexico-sémantiques). Cette combinaison de techniques illustre encore une fois comment les MM suggèrent des types méthodologiques différents par rapport aux idéaux-types *empirisme logique* et *constructivisme*. Cet aperçu des approches, devis et techniques permet de planifier les MM en évaluation de programmes, et d'aborder la question de la qualité des MM.

La planification et la qualité des méthodes mixtes en évaluation de programmes

Il existe peu de recommandations spécifiques pour planifier les MM en évaluation de programmes (Greene, Caracelli et Graham, 1989 ; Greene,

2006), mais les évaluateurs peuvent s'inspirer des recommandations existantes pour planifier des recherches MM (Creswell et Plano Clark, 2010). Au moment où nous écrivons ce chapitre, il n'existe toujours pas de norme généralement admise pour planifier une MM de « bonne qualité » ou pour évaluer la qualité méthodologique des MM. Selon Creswell *et al.* (2004) et O'Cathain (2010), des travaux sont nécessaires pour préciser les critères de qualité des MM.

À partir de notre recension des écrits (Pluye, Gagnon, Griffiths et Johnson-Lafleur, 2009), d'une étude pilote (Pace *et al.*, 2012), d'ateliers et de consultations avec des experts (Pluye *et al.*, 2011), nous proposons de concevoir la qualité méthodologique des MM de la manière suivante (tableau 7.2). Les MM doivent réconcilier les idéaux-types *empirisme logique* et *constructivisme* résumés plus haut. La théorie constructionniste de Hacking (1999) permet, à notre avis, cette réconciliation. Les MM créent un « effet de boucle » entre les données probantes provenant des études QUAL et QUANT, et construisent ainsi un nouveau « type mixte » de données probantes. Cette conception des MM soutient l'usage de différents critères pour différentes méthodes. Ainsi, nous proposons : 1) d'évaluer la qualité méthodologique des composantes QUAL des MM à partir de critères utilisés pour évaluer les méthodes QUAL ; 2) d'évaluer les composantes QUANT des MM à partir de critères utilisés en épidémiologie pour évaluer les méthodes QUANT ; et 3) d'évaluer la qualité générale des MM en fonction de ce que nous avons présenté dans la section précédente sur les approches, devis et techniques des MM. Fondé sur cette conception, un outil d'évaluation de la qualité des méthodes mixtes est disponible gratuitement sur internet (<http://mixedmethodsappraisaltool-public.pbworks.com>). Cet outil tient sur une seule page, il est fourni avec un manuel qui en facilite l'utilisation.

Il faut souligner que l'évaluation de la qualité des méthodes QUAL est controversée, et donc, par conséquent, celle des composantes QUAL des MM l'est aussi pour au moins trois raisons : quelle que soit leur qualité, certaines études QUAL peuvent *illuminer*, c'est-à-dire clarifier ou améliorer la compréhension des phénomènes observés (Sandelowski, 1993) ; peu de méthodes QUAL correspondent à des critères de qualité explicites généralement admis par une communauté de chercheurs (Sandelowski *et al.*, 2007) ; enfin, les méthodes QUAL ne sont pas des entités fixes telles que présentées dans les outils d'évaluation de la qualité, et nombreux sont

TABLEAU 7.2

Qualité des méthodes mixtes en évaluation de programmes

Examiner :	Critères de qualité des méthodes mixtes
1. Composante qualitative	<ul style="list-style-type: none"> • Sources de données pertinentes pour répondre à la question • Analyse des données pertinentes pour répondre à la question • Prise en compte du contexte de collecte des données • Réflexivité des évaluateurs sur leurs méthode et résultats
2. Composante quantitative randomisée	<ul style="list-style-type: none"> • Randomisation (ou génération de séquence) appropriée • Allocation appropriée (par ex., à l'aveugle) • Résultats complets • Faible taux d'abandon
3. Composante quantitative non randomisée	<ul style="list-style-type: none"> • Recrutement minimisant les biais • Mesures appropriées (validées ou standards) • Participants similaires dans les deux groupes (ou différences prises en compte dans l'analyse) • Résultats complets, taux de réponse et de suivi acceptables
4. Composante quantitative descriptive	<ul style="list-style-type: none"> • Échantillonnage approprié pour répondre à la question • Échantillon représentatif de la population étudiée • Mesures appropriées (validées ou standards) • Taux de réponse acceptable
5. Composante mixte	<ul style="list-style-type: none"> • Devis MM pertinent pour répondre aux questions • Intégration des données et (ou) résultats QUAL et QUAN • Discussion des limites associées à cette intégration (par ex., divergences QUAL vs QUANT)

ceux qui critiquent la prémisse voulant qu'il existe clairement une bonne et une mauvaise manière d'appliquer les méthodes QUAL.

Par conséquent, alors même que se développent les MM, il faut tenter d'en promouvoir la qualité en s'appuyant sur des critères de rigueur et ne pas tuer dans l'œuf la créativité et les découvertes associées aux composantes qualitatives des MM en imposant des critères stricts qui, justement, empêcheraient de bénéficier de la richesse que peut offrir la combinaison

des méthodes QUAL et QUANT. Pour cela, il serait possible d'utiliser les critères proposés au tableau 7.2 pour analyser séparément, en évaluation de programmes, les études et les composantes QUAL ou QUANT de bonne vs de faible qualité méthodologique.

Pour illustrer l'utilisation de cet outil, nous résumons ici la description et la discussion des méthodes utilisées au chapitre 21 pour évaluer un programme. Nous appliquons ensuite les critères présentés ci-dessus pour en évaluer la qualité méthodologique. Ridde *et al.* (chapitre 21) présentent l'utilisation d'un devis séquentiel explicatif pour (a) mesurer les dépenses effectuées par les femmes lors d'un accouchement eutocique dans une maternité de première ligne d'un district du Burkina Faso au moyen d'une enquête de prévalence, puis (b) comprendre ces résultats au moyen d'une étude de cas qualitative. Ils ont présenté les résultats quantitatifs sur l'écart constaté entre le forfait officiel (norme) et les paiements déclarés par les femmes (pratique), puis les résultats qualitatifs pour comprendre cet écart. En plus des leçons apprises de ces résultats pour améliorer la politique nationale en soins de santé primaire, les auteurs suggèrent des recommandations pour limiter l'écart de mise en œuvre de politiques publiques similaires (censées favoriser l'équité).

Composante quantitative descriptive: En conduisant une enquête auprès des ménages, l'échantillonnage était approprié pour répondre aux questions sur les dépenses. L'échantillon était représentatif de la population étudiée puisqu'il incluait quasiment toute la population (98,7 % des femmes ayant accouché et enregistrées dans les maternités). Les auteurs mentionnent que le questionnaire était adapté d'une précédente recherche. Pour être systématique et vérifier que les mesures étaient appropriées, nous pourrions contacter les auteurs et demander des précisions sur la validité et l'adaptation du questionnaire.

Composante qualitative: Les sources de données étaient pertinentes pour répondre à la question (*focus groups* et entrevues individuelles). Les auteurs justifient de manière convaincante les raisons du choix des informateurs clés parmi les agents de première ligne dans les centres de santé, les responsables des maternités ou des centres de santé et les membres de l'équipe cadre du district. Les résultats quantitatifs étaient présentés par les chercheurs et une discussion était menée pour examiner les pistes d'explications. L'analyse thématique inductive des données était pertinente pour répondre à la question. Les auteurs expliquent comment ils ont pris

en compte le contexte de la collecte des données, et ils fournissent une description détaillée du site, de la population, de l'organisation pyramidale (hiérarchique) de la société et du système de santé (pouvoir important des professionnels), et de la mise en œuvre du programme évalué. La réflexivité des évaluateurs sur leur méthode et leurs résultats est largement discutée. Par exemple, comment les biais inhérents aux discussions de groupe (contrôle social) ont été diminués en ayant également recours à des entrevues individuelles en profondeur qui ont été conduites dans des maternités identifiées comme « cas extrêmes » grâce à l'analyse des données quantitatives (ayant la moyenne la plus élevée en termes de paiement déclaré par les femmes).

Composante mixte: Le devis mixte était pertinent pour répondre aux questions. Le devis séquentiel explicatif est justifié de la manière suivante : après l'analyse des données quantitatives sur les dépenses effectuées par les femmes, quatre pistes qualitatives d'explications à l'écart constaté entre le forfait officiel et le paiement déclaré par les femmes ont été formulées. L'intégration des données s'est faite entre les étapes quantitative et qualitative (sélection de cas extrêmes), puis les résultats qualitatifs ont été intégrés aux résultats quantitatifs au moment de l'interprétation des résultats par les chercheurs. Finalement, les auteurs ont discuté les limites associées à cette intégration en termes de forces et faiblesses respectives des composantes quantitatives et qualitatives, par exemple sur l'absence de mesure de l'importance et de la signification du lien de causalité entre les écarts observés et les explications de ces écarts.

Une recension mixte des écrits pour les évaluations de programme

Dans toute démarche d'évaluation, la recension des écrits constitue un élément essentiel qui permet de préciser les objectifs de l'évaluation en fonction de ce qui est déjà connu, ou de discuter de ses résultats. On connaît les métasynthèses d'études QUAL (par exemple, les méta-ethnographies) et les revues systématiques d'études QUANT (par exemple, les revues des essais contrôlés randomisés avec méta-analyse). On connaît cependant moins ce que nous appelons les *recensions mixtes des écrits*, c'est-à-dire les revues concomitantes d'études QUAL, QUANT et MM (Pluye, Gagnon *et al.*, 2009). Par exemple, effectuer une recension mixte permet de faire la synthèse de tout type de connaissances empiriques entourant les

programmes à évaluer ou en cours d'évaluation. Des exemples peuvent être trouvés en éducation, en santé, en gestion et en sciences sociales. Nous résumons en encadré une recension mixte que nous avons effectuée pour améliorer notre méthode d'évaluation des technologies.

Nous avons également examiné un échantillon de 59 recensions mixtes des écrits en sciences de la santé. Ces recensions constituent un type de MM dont les données sont documentaires (documents rapportant des évaluations ou des études tant QUANT, QUAL que MM). Chaque recension mixte examine en moyenne 26 études QUANT, 17,9 études QUAL et 3,7 études MM. Les trois approches mentionnées plus haut sont utilisées par les recensions mixtes. Les résultats QUAL et QUANT sont complémentaires dans 54 % des cas, ils sont en tension dialectique dans 7 % des cas, et sont assimilés à ou transformés en résultats QUANT ou QUAL dans 36 % des cas.

Bénéficiant du développement des MM, les recensions mixtes des écrits ont récemment émergé comme une nouvelle forme de recension des écrits. Des 59 recensions examinées dans le domaine de la santé, 56 (95 %) ont été publiées depuis 2000. Selon notre échantillon, les recensions mixtes sont utiles pour explorer de nouvelles connaissances (par exemple, découvrir les tendances actuelles concernant le programme évalué) et (ou) pour confirmer des savoirs (par exemple, justifier les prémisses du programme évalué). Les recensions mixtes sont utilisées dans des buts exploratoires (62 %) ou de validation (19 %), ou encore dans ces deux buts à la fois (19 %).

Nous avons identifié trois types de recensions mixtes (*systématiques*, *reproductibles* ou *de convenance*) selon qu'elles satisfont ou non aux caractéristiques des recensions systématiques des écrits (Moynihan, 2002) : formulation d'une question ; identification des études potentiellement pertinentes ; sélection des études pertinentes ; évaluation de la qualité des études sélectionnées ; synthèse des résultats des études retenues pour leur pertinence et leur qualité. Nous avons identifié 17 *recensions systématiques* (29 %) qui satisfont aux cinq caractéristiques mentionnées, et 23 (39 %) *recensions reproductibles* qui satisfont quatre caractéristiques (question, identification, sélection et synthèse) et n'évaluent pas la qualité des études sélectionnées pour leur pertinence. Quant aux 19 *recensions de convenance* (32 %), elles satisfont uniquement à deux caractéristiques (question et synthèse), car l'identification et la sélection ne peuvent pas être reproduites

à partir des informations disponibles et parce qu'il n'y a pas d'évaluation de la qualité des études retenues. Si les évaluateurs décident de conduire une recension systématique, ils peuvent utiliser le tableau 7.2 pour évaluer la qualité des études QUAL, QUANT et MM sélectionnées pour leur pertinence.

Une recension mixte des écrits en évaluation des technologies

- L'évaluation des technologies est complexe et ouverte au pluralisme méthodologique. Nous avons identifié, sélectionné, évalué et fait la synthèse de tous les types d'études empiriques examinant l'impact des ressources électroniques sur les professionnels.
- Méthodes. Nous avons recensé les écrits scientifiques internationaux des années 1960 à 2004. Deux réviseurs ont identifié indépendamment les études potentiellement pertinentes en scrutant respectivement 3 368 et 3 249 références (titre, auteurs, revue, résumé) en utilisant deux stratégies de recherche dans des bases de données bibliographiques. Des études additionnelles ont été identifiées à la main (recherche dans des dossiers personnels et des revues spécialisées) et en effectuant une recherche des articles citant les articles sélectionnés par l'entremise de la base de données bibliographique ISI Web of Sciences. Les textes entiers de 605 articles ont été examinés pour leur pertinence. Parmi ceux-ci, 40 (6,6 %) ont été évalués indépendamment par deux réviseurs pour leur pertinence et leur qualité méthodologique. Ces articles rapportaient des recherches évaluatives quantitatives, qualitatives ou mixtes. De ces derniers, 26 (4,3 %) ont été retenus comme pertinents et de qualité acceptable pour faire une synthèse. Pour chaque article retenu, deux équipes ont effectué une analyse thématique qualitative pour faire la synthèse des résultats quantitatifs et qualitatifs.
- Résultats. Les études quantitatives observationnelles suggèrent qu'environ un tiers des recherches d'informations dans les ressources électroniques ont un impact positif sur les médecins. Les études quantitatives expérimentales ou effectuées en laboratoires informatiques (simulations) sont contradictoires quant à l'impact des ressources électroniques par rapport aux autres sources d'information comme le matériel éducatif imprimé. Les résultats qualitatifs et la qualification des résultats quantitatifs soutiennent et complètent les types d'impact que nous avons proposés dans une étude pilote.

Source: Pluye, P. *et al.* (2005), « The Impact of Clinical Information-Retrieval Technology on Physicians: A Literature Review of Quantitative, Qualitative and Mixed-method Studies », *International Journal of Medical Informatics*, vol. 74, n° 9.

Conclusion

Les chercheurs et les évaluateurs qui utilisent les méthodes mixtes rejettent la thèse d'incommensurabilité qui sépare le constructivisme et l'empirisme logique en deux idéaux-types méthodologiques incompatibles (Johnson, 2007). Le présent chapitre montre comment les propositions de Hacking (constructionnisme) peuvent aider à conceptualiser les MM et les données probantes mixtes, et à partir de cette conceptualisation, il décrit trois approches méthodologiques MM pour l'évaluation des programmes, trois types de devis MM et diverses combinaisons de techniques de collecte et d'analyse de données pour les MM. De plus, il propose des

critères pour évaluer la qualité des MM. Selon notre expérience, les MM devraient être planifiées, implantées et évaluées par un expert en méthodes QUAL et QUANT, ou par au moins deux experts ayant des expertises complémentaires en méthodes QUAL et QUANT. Par exemple, notre évaluation des techniques et notre recension des écrits ont été conduites par un chercheur ayant une formation en épidémiologie avec une expertise en MM et en méthodes QUAL, et par des chercheurs détenant une expertise en méthodes QUANT.

En conclusion, nous proposons un plan de cours pour enseigner les MM qui peut être utile aux évaluateurs, aux étudiants et aux professeurs dans le domaine de l'évaluation de programmes. Ce cours a été implanté en 2008. Bien que celui-ci soit destiné aux étudiants dans le domaine de la santé, il a été construit sur la base d'autres programmes nord-américains en éducation, en psychologie et en sociologie. Le cours est conduit par un chercheur QUAL (sociologue), un chercheur QUANT (épidémiologiste) et un chercheur détenant une formation en épidémiologie avec une expertise en MM et en méthodes QUAL.

Enseigner les méthodes mixtes : plan de cours

- Cours 1. Introduction : Afin de préparer les étudiants pour les études mixtes, une brève introduction à chaque type de recherche est fournie aux étudiants.
- Cours 2. Projet de recherche : Les étudiants ébauchent un projet de recherche utilisant les méthodes mixtes.
- Cours 3. Recension mixte des écrits : Les étudiants apprennent les techniques pour conduire une analyse documentaire incluant des études qualitatives, quantitatives et mixtes.
- Cours 4 et 5. Devis mixtes : Les devis mixtes sont présentés, et chaque étudiant applique un devis approprié pour son projet.
- Cours 6. Échantillonnage et instrumentation : Les étudiants construisent l'échantillonnage et l'instrumentation appropriés, avec des composantes qualitatives et quantitatives, pour appliquer leur devis de recherche.
- Cours 7. Collecte de données et réflexion sur l'instrumentation : Les étudiants appliquent et présentent leur instrumentation à leurs pairs et en discutent avec eux.
- Cours 8. Divergences : Les étudiants utilisent trois stratégies pour analyser les divergences entre données qualitatives et quantitatives.
- Cours 9. Analyser des mots et des nombres : Les étudiants organisent et analysent leurs données qualitatives et quantitatives.
- Cours 10. Évaluation de la qualité méthodologique des études utilisant les méthodes mixtes : Les différentes méthodes pour évaluer la qualité des études qualitatives, quantitatives et mixtes sont présentées.
- Cours 11. Session de lecture critique : Les étudiants évaluent de façon critique un article qui utilise un devis mixte.
- Cours 12. Projets individuels : Les étudiants présentent leurs projets (communication orale) et reçoivent les commentaires et suggestions des instructeurs et de leurs pairs.

Source : Levine, A., B. Nicolau et P. Pluye, « Applied Mixed Methods in Health Research », plan de cours, Université McGill.

Pour aller plus loin :

CRESWELL, J. W. et V. L. PLANO CLARK (2010), *Designing and conducting Mixed Methods Research*, Thousand Oaks, Sage Publications.

Ce livre contient les plus récentes propositions de John Creswell (auteur du livre à succès *Research Design*) et de Vicki Plano Clark sur les types de devis de recherche et d'évaluation utilisant des méthodes mixtes. Il comprend des recommandations pratiques pour élaborer, étape par étape, un devis d'évaluation utilisant des méthodes mixtes. Ce livre constitue un manuel de base suffisant pour mener à bien une évaluation mixte.

GREENE, J. C., V. J. CARACELLI et W. F. GRAHAM (1989), « Toward a conceptual framework for mixed-method evaluation designs », *Educational Evaluation and Policy Analysis*, vol. 11, n° 3, p. 255-274.

Dans cet article pionnier, un cadre conceptuel des méthodes mixtes en évaluation est proposé à partir de la littérature et de l'analyse de 57 recherches évaluatives utilisant des méthodes mixtes. Cinq stratégies d'évaluation utilisant des méthodes mixtes sont identifiées : la triangulation, la complémentarité, le développement, l'initiation et l'expansion.

TASHAKKORI, A. et C. TEDDLIE (2010), *Handbook of Mixed Methods in Social and Behavioral Research*, Thousand Oaks, Sage Publications.

Ce livre présente les fondements de l'utilisation des méthodes mixtes. Les défis théoriques sont très bien développés et illustrés à l'aide d'exemples concrets. Tous les aspects des méthodes mixtes sont abordés par les auteurs clés dans le domaine. L'ouvrage offre une synthèse des méthodes mixtes très complète. Par rapport au manuel de Creswell et Plano Clark (2010), il constitue une référence incontournable pour les évaluateurs, les étudiants et les chercheurs qui veulent utiliser les méthodes mixtes de manière plus avancée.

Internet

Mixed Methods for Health Research : Overview of mixed methods research design & review.

Cet ouvrage, libre d'accès en ligne, contient une liste de ressources mises à jour par Frances Griffiths et Felicity Boardman de l'Université de Warwick (glossaire, bibliographie, experts, cours et formation, logiciels, sites internet, conférences et réseaux) :

www2.warwick.ac.uk/fac/med/research/hsri/primary_care/research_/centrepatexp/complexityhealth/mixedmethods

Journal of Mixed Methods Research.

La publication internationale de référence pour les articles qui portent sur les méthodes mixtes dans toutes les disciplines. Soutenu par les chefs de file de ce domaine, chaque numéro explore des recherches et des évaluations originales utilisant des méthodes mixtes, et présente des articles méthodologiques et théoriques qui contribuent aux connaissances sur ces méthodes. Le site de la revue : www.sagepub.com/journalsProdDesc.nav?prodId=Journal 201775