

Approches qualitative et quantitative en évaluation de programmes

Qualitative and Quantitative Approaches in Program Evaluation

Normand PÉLADÉAU and Céline MERCIER

Volume 25, Number 2, Fall 1993

La construction des données

URI: <https://id.erudit.org/iderudit/001547ar>

DOI: <https://doi.org/10.7202/001547ar>

[See table of contents](#)

Publisher(s)

Les Presses de l'Université de Montréal

ISSN

0038-030X (print)

1492-1375 (digital)

[Explore this journal](#)

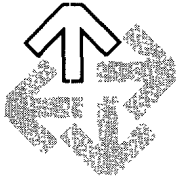
Cite this article

PÉLADÉAU, N. & MERCIER, C. (1993). Approches qualitative et quantitative en évaluation de programmes. *Sociologie et sociétés*, 25(2), 111–124.
<https://doi.org/10.7202/001547ar>

Article abstract

In program evaluation, it is common to oppose qualitative and quantitative evaluation. This paper questions the postulate which states that qualitative and quantitative approaches belong to different paradigms. Such a position appears to correspond little to real practices and is, in fact, very harmful to the development of methodologies for program evaluation. An examination of recent studies in program evaluation makes it possible to identify four different approaches liable, in our view, to bring about evolution in evaluation practices, based on criticisms which tenants of both quantitative and qualitative approaches have made of one another. The first of these strategies is the exclusive use of one approach, be it qualitative or quantitative. The second calls upon one or the other approach, depending on the context. The third and fourth imply simultaneous recourse to both approaches, through triangulation or integration by combination.

Approches qualitative et quantitative en évaluation de programmes



NORMAND PÉLADEAU et CÉLINE MERCIER

Historiquement, l'évaluation de programmes s'est constituée comme un domaine appliqué, relié à la prise de décision, à la gestion et à l'amélioration des programmes d'intervention. Elle consiste essentiellement en une collecte systématique d'information sur les composantes (la structure), le fonctionnement (le processus), les résultats (l'efficacité) ou l'impact d'un programme. Un programme peut lui-même être défini comme un ensemble d'activités, soutenues par des ressources humaines et matérielles, visant à intervenir dans un processus social avec l'intention de résoudre un problème ou de fournir un service.

Du fait que l'évaluation de programmes s'applique à un grand nombre de secteurs d'activités : développement international, éducation, environnement, justice, santé, services sociaux, etc., elle ne se trouve pas rattachée à une discipline particulière, mais emprunte ses méthodes à l'ensemble des sciences sociales. Avec l'expérience, l'évaluation a également modifié les pratiques méthodologiques existantes pour les adapter à son propre contexte et développé de nouvelles techniques de recherche. Située au confluent de plusieurs disciplines, les écrits sur l'évaluation de programmes constituent ainsi un point de vue privilégié pour appréhender la diversité des méthodologies ainsi que les enjeux des débats opposant différentes positions méthodologiques. Parmi ces débats, celui qui a suscité le plus de remous ces dernières années est sans nul doute celui qui oppose les approches qualitative et quantitative.

Bien que plusieurs auteurs s'interrogent sur le bien-fondé d'une telle opposition et refusent d'accorder à l'une ou l'autre des approches une quelconque supériorité, chacune a ses partisans, prêts à défendre l'idée que leur méthodologie est la plus appropriée pour répondre aux besoins de l'évaluation. La revue *Evaluation Practice* se faisait dernièrement l'écho de ce genre de débats, à partir des positions des deux derniers présidents de l'Association américaine d'évaluation. En réponse au discours de Yvonna Lincoln, la présidente sortante et l'une des chefs de file de l'évaluation de la « quatrième génération » (Lincoln, 1991), le président en titre, Lee Sechrest, déplorait le fait que l'on présente l'évaluation qualitative comme une méthodologie destinée à remplacer toutes les autres, et non comme un récent développement survenu dans le domaine de l'évaluation, à utiliser de façon conjointe avec les méthodes quantitatives (Sechrest, 1992). Le débat se

poursuivait dans le numéro suivant de la revue, avec la réponse de Lincoln et de Guba (1992) et l'appel à la conciliation du président élu (Fetterman, 1992).

L'existence d'un tel débat n'est pas problématique en soi et peut même comporter certains avantages. Il constitue en effet une source de stimulation intellectuelle où chacun se trouve amené à redéfinir sa position, à s'interroger sur ses méthodes à la lumière des critiques qui leur sont adressées et, éventuellement, à les perfectionner. Le problème surgit cependant lorsque la façon de poser les termes du débat a pour conséquence de faire obstacle au développement respectif des différentes méthodologies et, par conséquent, au domaine de l'évaluation de programmes. Il en est ainsi, à notre avis, lorsqu'on en vient à définir le débat en terme d'opposition épistémologique et non pas strictement méthodologique. Plusieurs auteurs, particulièrement certains tenants d'une approche qualitative de l'évaluation de programmes (Filstead, 1979; Guba et Lincoln, 1989; Lecompte, 1982; Patton, 1978, 1987), définissent, en effet, la nature de l'opposition entre les méthodologies quantitative et qualitative comme se situant fondamentalement au niveau des paradigmes. Nous sommes d'avis que cette tendance a pour effet de créer une polarisation entre les évaluateurs qualitatifs et quantitatifs et de faire obstacle à tout dialogue entre les tenants des deux approches. Cette vision réduit également la marge de manœuvre de chacune des approches en limitant les possibilités d'innovation et, plus particulièrement, celles impliquant une intégration de méthodes qualitatives et quantitatives.

On peut à cet égard soulever des doutes quant à l'applicabilité de la notion de paradigme à un domaine comme celui de l'évaluation de programmes, lequel, jusqu'à présent, s'est presque essentiellement défini par les méthodes utilisées ou par les idéologies (Scriven, 1989) et où on retrouve peu de réels fondements théoriques, sinon aucun (Chen, 1990). L'examen des textes portant sur l'évaluation de programmes ne permet pas de croire en l'existence d'une science unifiée, ni même à l'existence de problèmes méthodologiques importants qui demeureraient insolubles à l'intérieur de chacune des approches. En fait, si l'on devait caractériser l'état actuel du domaine de l'évaluation de programmes selon la perspective de Kuhn, il semblerait plus opportun d'opter plutôt pour un stade pré-paradigmatique. Sur cette base, Alexander (1986) a soulevé des doutes quant à l'applicabilité du concept de paradigme dans le domaine de l'évaluation de programmes. Par ailleurs, Reichart et Cook (1979) ont clairement fait la démonstration que le choix d'une méthodologie de recherche n'est pas indubitablement lié aux positions philosophiques, épistémologiques ou ontologiques du chercheur.

Dans le contexte de ce débat entre qualitatif et quantitatif, nous avons retenu des pratiques actuelles de recherche quatre démarches qui nous semblent susceptibles de faire évoluer la pratique de l'évaluation, à partir des critiques que s'adressent réciproquement les tenants des deux approches.

La première s'inscrit dans un refus de toute intégration des approches et consiste à demeurer strictement dans l'ordre du qualitatif ou du quantitatif. Suivant ce scénario de l'exclusivité, les tenants de l'une ou l'autre école ne visent pas à intégrer ou à utiliser des éléments de l'autre approche, mais tentent néanmoins de répondre aux critiques qui leur sont faites par le développement de nouvelles méthodes ou la modification des pratiques existantes. La seconde, que l'on peut qualifier de position « contextualiste », consiste à choisir plutôt l'une ou l'autre des stratégies de recherche en fonction du contexte de l'évaluation. Ces deux dernières pratiques, qui semblent caractériser un nombre croissant d'évaluations, visent à combiner deux méthodologies au sein d'une même recherche. Dans un cas, cette combinaison se fait par l'adoption de démarches parallèles indépendantes, l'une quantitative et l'autre qualitative, dans une optique de triangulation. Dans l'autre, on procède à une démarche de recherche intégrant des éléments des deux grandes traditions de recherche.

L'EXCLUSIVITÉ DES APPROCHES

Malgré la critique faite précédemment de la polarisation suscitée par la définition du débat qualitatif/quantitatif, le choix exclusif d'une méthodologie, quelle qu'elle soit, comporte des avantages certains. Les chercheurs qui s'identifient à une approche donnée sont les plus susceptibles de réagir aux critiques et aux limites de leurs méthodes par le développement de nouveaux moyens et de nouvelles applications à l'intérieur même de leur modèle, et cela sans avoir recours à un autre type de méthodologie. Pour Reichardt et Cook (1979), la découverte d'un biais n'est pas nécessairement une raison pour rejeter la méthode, mais un défi pour l'améliorer. Certaines objections quant à l'utilisation en évaluation de programmes de démarches quantitatives ou qualitatives permettent d'illustrer comment chacune de ces approches a développé de nouvelles méthodes pour répondre à ces critiques.

DÉVELOPPEMENT DES MÉTHODES QUANTITATIVES

Deux critiques sont souvent faites quant au recours à des schémas expérimentaux en évaluation de programmes : le grand nombre de participants requis et la durée de l'enquête. Compte tenu de l'utilisation courante des statistiques inférencielles, il est souvent nécessaire de recruter un nombre important de participants pour être en mesure de tirer des conclusions valides. Ce problème est particulièrement marqué lorsque peu de clients reçoivent le même type d'interventions, soit qu'ils sont peu nombreux à présenter le même problème, soit que l'intervention est individualisée de façon à s'adapter aux besoins spécifiques du client. Par ailleurs, l'application d'un devis expérimental ou quasi expérimental implique généralement l'obtention de données de suivi sur une longue période. L'évolution des programmes étant ce qu'elle est, les résultats de l'étude arrivent souvent trop tard pour qu'ils puissent être d'une utilité quelconque. Pour remédier à ces problèmes, plusieurs auteurs ont proposé l'utilisation de devis expérimentaux à sujet unique (*single case experimental design*) dans le cadre de l'évaluation de programmes (Horn et Heerboth, 1982; Kratochwill et Levin, 1978; Nuehring, et Pascone, 1986). Cette méthode a l'avantage de pouvoir être appliquée à un très petit échantillon, voire à un seul sujet, et d'être en mesure de produire très rapidement de l'information quant à l'efficacité à court terme d'une intervention. Elle permet également d'obtenir des informations précises sur le processus d'intervention.

On reproche souvent aux schèmes expérimentaux ou quasi expérimentaux en évaluation d'être simplificateurs et de ne pas tenir compte de la complexité de la réalité étudiée. Cette critique, principalement formulée par des chercheurs qualitatifs, l'est également par les chercheurs quantitatifs. Ainsi, dans une recension de rapports d'évaluation, Lipsey (1988) constate que les évaluations visant à déterminer l'efficacité d'une intervention très souvent reposent sur un modèle de causalité simpliste qui sous-estime la complexité des processus en jeu. Pour cet auteur, le devis utilisé correspond rarement à la complexité des interactions et à la dynamique existant entre l'intervention, la clientèle, l'environnement et les changements observés. L'intervention évaluée comporte souvent plusieurs composantes, s'étend sur une longue période et vise des objectifs multidimensionnels et variables dans le temps. L'impact du programme implique la plupart du temps des effets directs et indirects, souvent diffus ou subtils, manifestes à plus ou moins long terme. Les effets du programme d'intervention se trouvent également grandement influencés par une série de facteurs extérieurs au programme.

Pour répondre à ces limites, Lipsey (1988) plaide en faveur du développement et de l'utilisation de méthodes quantitatives plus appropriées. Il suggère entre autres choses de traiter l'implantation d'un programme non pas comme une variable indépendante de nature dichotomique, mais plutôt comme une variable continue, ce qui serait plus conforme à la réalité de l'implantation des interventions (Cook et Poole, 1982; Mark, 1983). Il préconise également l'utilisation de certains types d'analyses multivariées, plus susceptibles de traduire la complexité des relations causales. À cet égard, on peut constater

qu'au cours de la dernière décennie, un progrès considérable a été réalisé par l'utilisation et le développement de techniques statistiques ou méthodologiques sophistiquées, lesquelles sont mieux en mesure de rendre compte de la complexité des phénomènes étudiés : cheminement de la causalité (*path analysis*) et procédure de *pattern matching* (Trochim, 1989) qui permettent de vérifier le niveau de correspondance entre une configuration d'hypothèses et un ensemble d'observations.

DÉVELOPPEMENT DES MÉTHODES QUALITATIVES

Cette disposition à modifier ses pratiques et à développer d'autres techniques de recherches pour répondre aux limites observées se retrouve également chez les évaluateurs de tradition qualitative. Ainsi, on reproche souvent aux études qualitatives d'être incapables de fournir des données valides et fiables susceptibles d'être généralisées à d'autres situations ou d'autres populations (Chen, 1990). Cette limite serait particulièrement grave dans les cas où la cueillette et l'interprétation des données ne sont faites que par une seule personne (Mechanic, 1989). En fait, les notions de validité et de fidélité ne sont pas étrangères aux chercheurs qualitatifs (Kirk et Miller, 1986). Ceux-ci font usage de concepts qui s'apparentent aux notions de fidélité et de validité interne et externe tels que la crédibilité, la transférabilité et la neutralité chez Guba et Lincoln (1989) ou les critères de précision et de consistance chez Cronbach et Suppes (1969). Un examen des études consacrées à la recherche qualitative laisse entrevoir d'ailleurs une préoccupation constante pour tout ce qui touche la qualité des informations recueillies. Ainsi, Silverman, Ricci et Gunter (1990) proposent une dizaine de stratégies différentes pour augmenter la validité et la fidélité des données des recherches qualitatives. À titre d'exemple, ils suggèrent une sélection et un entraînement rigoureux des travailleurs de terrain pour augmenter la portée (*comprehensiveness*) et la précision des informations recueillies. Ils recommandent également l'obtention de données d'informateurs clés par au moins deux interviewers ou à deux moments différents, pour identifier les biais d'interviewers ou pour corriger les inexactitudes ou les inconsistances dans les informations recueillies. Ils donnent également une série de recommandations touchant la sélection des lieux de cueillette, le choix des sources d'information, la composition de l'équipe de recherche et les modalités d'interprétation des données.

Cette capacité des approches de répondre aux critiques dans le cadre même de leurs orientations méthodologiques respectives amène à s'interroger sur leurs limites réelles : Dans quelle mesure une seule approche pourrait-elle éventuellement s'appliquer à tout objet d'étude ? Ou doit-on croire au contraire que chaque méthodologie a ses limites infranchissables et ne peut s'appliquer qu'à certains types d'objet ? Quoi qu'il en soit, la capacité d'innovation et de développement démontrée par chacune des approches semble peu correspondre au portrait statique et immuable que se renvoient leurs opposants respectifs.

Malgré les avantages pour le développement des méthodes d'une position d'exclusivité, dans les conditions actuelles, cette dernière comporte cependant des risques importants. Au niveau pratique, elle peut faire en sorte que les tenants de l'une ou l'autre approche ne répondent pas adéquatement aux besoins des commanditaires d'une évaluation, par incapacité de voir certaines dimensions du problème posé, ou en raison des limites actuelles des méthodes propres à leur approche.

LE CHOIX DES APPROCHES SUIVANT LE CONTEXTE

Une deuxième stratégie qui semble recueillir la faveur d'un grand nombre d'auteurs est liée à l'idée que chaque forme d'évaluation est appropriée dans un contexte particulier. Il existerait en quelque sorte une espèce de division du travail, où chaque méthodologie se verrait attribuer un privilège de pratique, dans un champ précis de spécialisation. Ainsi, Patton (1987) identifie plus de vingt conditions où la cueillette d'information qualitative peut s'avérer utile. Pour cet auteur, l'évaluation qualitative serait particulièrement appropriée

dans les programmes dont les objectifs demeurent vagues ou très généraux ou dans ceux où les objectifs d'intervention sont individualisés. Elle s'avèrerait également propice à fournir des informations détaillées sur la dynamique d'un programme et son implantation, à documenter en profondeur certains aspects plus particuliers de l'intervention ou de la clientèle, ou à décrire la diversité des interventions. L'évaluation qualitative serait de même plus adéquate lorsque l'évaluation est réalisée à des fins formatives ou exploratoires, ou dans le but de développer une théorie de l'intervention. Enfin, lorsque le contexte exige de l'évaluateur un mode de cueillette personnalisé, lorsque les procédures de cueillette de données doivent interférer le moins possible avec les activités du programme ou lorsqu'il n'existe pas de mesures quantitatives appropriées, les méthodes qualitatives sont tout indiquées.

Rossi et Berk (1981) discutent pour leur part comment les approches qualitative et quantitative peuvent apporter une contribution spécifique à différents types d'évaluation. Ainsi, dans le cadre d'une analyse de besoins, les études qualitatives seraient plus susceptibles d'être efficaces lorsque l'étude vise à déterminer la nature des besoins, tandis que l'application de méthodes quantitatives serait plus en mesure de documenter l'étendue de ces besoins. Dans le contexte d'une évaluation d'impact, Rossi et Berk recommandent l'utilisation d'un devis expérimental comme étant le meilleur moyen de s'assurer que l'effet observé est réellement attribuable à l'intervention, en permettant l'élimination des autres hypothèses explicatives et en s'assurant que ce changement n'est pas le fruit du hasard. Ils considèrent néanmoins l'évaluation qualitative comme une meilleure option lorsqu'il s'agit d'identifier pourquoi une intervention est efficace ou non.

Pour Heilman (1980), l'application de l'une ou l'autre des approches pourrait en fait s'inscrire dans une séquence d'activités complémentaires, correspondant aux phases de développement d'un programme d'intervention. Aux premiers stades d'implantation d'un programme, le besoin de feedbacks continus pour permettre les ajustements nécessaires aux circonstances de l'intervention serait plus important. Dans un tel contexte, une évaluation formative de nature qualitative serait la plus appropriée. Une fois le programme bien implanté et la distribution des services devenue plus routinière, les besoins d'information sommative s'accroîtraient, nécessitant l'utilisation de méthodes quantitatives. Pour Heilman, on devrait également prendre en considération, dans le choix d'une méthode d'évaluation, les caractéristiques des utilisateurs actuels et potentiels, et leurs préférences pour un type ou un format particulier d'information.

Le courant de l'évaluation orientée vers la théorie (*Theory-driven program evaluation*, Chen, 1990) représente sans doute l'effort le plus systématique pour lier le choix d'une méthode au contexte d'un programme. Chen (1990) propose, en effet, un cadre d'analyse permettant : 1) de décrire un programme, son développement et son contexte, 2) d'établir à la suite de cette définition des objectifs d'évaluation, 3) pour finalement choisir une méthode d'évaluation appropriée.

L'idée primordiale sous-jacente à l'utilisation judicieuse d'une évaluation de type qualitatif ou quantitatif est la reconnaissance que chacune des traditions de recherche a permis le développement d'un champ d'expertise qui lui est propre et de méthodes appropriées pour répondre à certains types de questions. L'avantage qu'il y a, à court terme, de choisir la méthodologie en fonction du contexte est d'offrir à l'évaluateur une plus grande flexibilité pour répondre aux diverses demandes qui lui sont adressées. Le risque potentiel d'une telle perspective réside cependant dans une application trop rigide, susceptible d'empêcher l'innovation qui pourrait découler de l'application d'une méthodologie dans un contexte inhabituel. On pourrait ainsi envisager une évaluation d'impact qui serait de nature exclusivement qualitative. Il n'est pas non plus impossible de procéder à une évaluation de processus au moyen de méthodes quantitatives. En fait, malgré les avantages certains que peut comporter l'adoption d'une forme d'éclectisme méthodologique, ce type de stratégie a comme conséquence de ne pas encourager le développement des méthodologies, puisqu'il tend à répondre aux limites d'une approche par l'adoption d'une autre approche momentanément plus appropriée.

L'INTÉGRATION PAR TRIANGULATION

En soi, l'idée de faire usage de plusieurs méthodes dans une même recherche n'est pas nouvelle. En évaluation de programmes, comme dans bien d'autres domaines des sciences humaines et sociales, on valorise depuis de nombreuses années l'utilisation de plusieurs méthodes. Le concept de méthodes multiples se réfère à l'utilisation de mesures différentes d'un même construit (Campbell et Fiske, 1959), à la combinaison de divers devis quasi expérimentaux dans un devis synthétique (Chen, 1990), à la réalisation d'analyses multiples de données ou à la combinaison d'informations de différentes sources ou de différents observateurs (Denzin, 1978). La triangulation peut être conçue comme une modalité particulière d'utilisation de plusieurs méthodes où l'objectif recherché est d'accroître la vraisemblance des conclusions d'une étude par l'obtention de résultats convergents obtenus par des méthodes différentes. Cette stratégie de triangulation a souvent été proposée comme une formule propre à réconcilier les approches qualitative et quantitative.

Une des prémisses de l'idée d'utiliser plusieurs méthodes est que chacune de ces méthodes comporte des faiblesses qui se traduisent par des biais et des erreurs de mesure. L'utilisation de méthodes multiples permet de compenser les faiblesses de l'une par les avantages de l'autre. Par exemple, Smith (1986) souligne que si les études qualitatives fournissent des données écologiquement plus valides, elles peuvent néanmoins souffrir d'un biais introduit par l'observateur ou d'une trop grande identification du chercheur avec le programme d'intervention ou avec les participants à la recherche en raison de contacts personnels prolongés. Par ailleurs, si les questionnaires permettent d'obtenir rapidement des données représentatives et standardisées, ils suscitent d'avantage de réponses biaisées et peuvent impliquer une trop grande simplification des divers points de vue. On postule qu'en combinant dans une même étude plusieurs méthodes qui ne partagent pas les mêmes faiblesses, les biais de l'une feront contrepoids aux biais de l'autre, ce qui permet de s'assurer que les résultats obtenus ne sont pas un artefact de la méthode utilisée. Dans cette perspective, la triangulation consiste à procéder à l'intérieur d'une même étude évaluative à deux recherches indépendantes dont l'une est associée à une démarche qualitative et l'autre est de nature plus quantitative et d'en croiser les résultats. L'objectif de cet exercice de triangulation est l'obtention de données qui ne soient pas complémentaires mais concurrentes, dans le but d'établir le degré de convergence des conclusions.

Dans sa forme la plus pure, la triangulation de méthodologies qualitative et quantitative implique une indépendance complète des processus de collecte de données qualitatives et quantitatives et des activités d'analyse, pour ainsi éviter toute contamination mutuelle. Elle implique également une absence d'interactions et de communications entre les différents chercheurs ou équipes de recherche tout au long du processus de recherche. L'intégration des méthodes ne survient en fait qu'à l'étape finale où une comparaison est faite des résultats et des conclusions des deux recherches indépendantes.

Dans la pratique cependant, ce type de stratégie d'intégration des méthodologies qualitative et quantitative, en plus d'introduire d'autres difficultés plus importantes, pose les mêmes problèmes que toute étude faisant usage de méthodes multiples. Le principal problème dans la conduite d'une démarche de triangulation survient lorsque les démarches donnent lieu à des conclusions divergentes ou qui entrent en contradiction les unes avec les autres (Shotland et Mark, 1987). Dans une telle situation, les chercheurs ne disposent souvent pas de règles précises et impartiales pour trancher en faveur de l'une ou l'autre des interprétations. La seule façon de tirer des conclusions est de privilégier l'une des approches, au risque du même coup de discréditer l'autre. C'est alors qu'apparaît l'arbitraire de la décision, prise tantôt en fonction des préférences méthodologiques des chercheurs, tantôt en raison de considérations politiques ou autres qui n'ont souvent rien à voir avec la rigueur méthodologique de la démarche. Plusieurs auteurs constatent dans ces situations une tendance marquée à reléguer un type de méthodologie à un rôle secondaire

et à utiliser les données qui en sont issues pour compléter les résultats obtenus par l'autre approche (Jick, 1983 ; Trend, 1979 ; Greene et McClintock, 1985).

La triangulation de méthodes qualitatives et quantitatives peut aussi faire en sorte que les différentes démarches de recherche produisent des résultats non comparables, du fait qu'elles concernent des questions différentes (Jick, 1983). Les conclusions sont alors non pas contradictoires ou rivales mais relativement indépendantes : elles relèvent de perspectives différentes d'un même problème ou traduisent des intérêts divergents. Pour reprendre l'analogie de la géométrie à l'origine de l'idée de triangulation, on ne peut établir la localisation d'un objet par triangulation si les deux observateurs ne pointent pas le même objet. Pour Kidder et Fine (1987), la probabilité d'un tel risque augmente lorsqu'on combine des méthodes qualitatives et quantitatives, dans la mesure où le processus inductif de la démarche qualitative encourage le chercheur à émettre ou à réviser certaines hypothèses en cours d'évaluation, lui permettant ainsi de s'écarter des objectifs fixés au départ. Pour d'autres auteurs, la nature même des données résultant d'études qualitatives et quantitatives est à ce point différente qu'elle empêche d'emblée toute possibilité de comparaison.

Un autre problème causé par l'utilisation de plusieurs méthodes tient au postulat selon lequel les biais de chacune des stratégies de recherche devraient s'annuler. Or, il n'est pas rare de constater que des méthodes différentes comportent des biais qui, loin de s'annuler, vont en fait dans la même direction. Cette situation a pour fâcheuse conséquence d'amener le chercheur à accorder une plus grande confiance à des conclusions fautives (Shotland et Mark, 1987). Pour ces auteurs, le problème est d'autant plus difficile à surmonter qu'il est souvent impossible de déterminer *a priori* la nature du biais d'une méthode.

À ces critiques on peut également ajouter un ensemble de problèmes pratiques tels que le dédoublement des activités de recherche, lequel entraîne des coûts considérables et nécessite une durée de réalisation plus longue, ainsi que les difficultés de collaboration entre chercheurs de tradition différente dans un contexte de polarisation des approches.

L'INTÉGRATION PAR COMBINAISON

Les scénarios que nous avons vus jusqu'à présent tentent de maintenir une indépendance des méthodologies de recherche, de telle sorte que chacune conserve l'intégrité de sa démarche. Le quatrième scénario implique, quant à lui, une combinaison délibérée au sein d'une même étude d'éléments associés à chacune des deux grandes approches par un même évaluateur, dans le but d'augmenter la qualité de l'évaluation. Contrairement à l'approche de triangulation qui exigeait un déroulement indépendant des recherches de type quantitatif et qualitatif, de façon à éviter toute contamination des démarches, ce quatrième scénario encourage au contraire une influence réciproque des méthodes. En fait, comme le soulignent Mark et Shotland (1987), l'indépendance des méthodes que sous-tend la logique de la triangulation devient souvent dysfonctionnelle lorsque l'on vise à assurer la complémentarité des données ou à ajuster les activités de recherche en cours d'exécution.

Les modalités d'agencement des approches qualitative et quantitative dans un même devis de recherche sont innombrables. Bien souvent, une des méthodologies s'avère nettement prédominante tandis que la seconde assume un rôle plus accessoire pour clarifier ou enrichir les résultats obtenus par la première (Lipse, Cordray et Berger, 1981 ; Trend, 1979). D'autres modalités d'agencement accordent plutôt une importance comparable à chacune des approches. Tel est le cas lorsque l'intégration des méthodes se fait de façon successive et qu'à différentes étapes de l'évaluation correspond une démarche de recherche différente ou, encore, lorsque l'étude implique la cueillette parallèle de données qualitatives et quantitatives dans le but de documenter des aspects différents d'un programme d'intervention. Ainsi, les évaluateurs d'orientation, autant qualitative que quantitative, peuvent tirer avantage d'une utilisation combinée de différentes méthodes.

Un des aspects particuliers du travail de l'évaluateur est qu'il est souvent amené à évaluer un programme dont il ne connaît préalablement pas la nature ou dont la problématique lui est peu familière. Une des premières tâches qu'il doit accomplir est de se familiariser avec le programme et sa problématique. En matière d'évaluation de programmes, on reconnaît qu'une bonne connaissance de l'organisation est importante dans la planification d'une évaluation. Elle permet de découvrir les composantes du programme, les objectifs réels d'une intervention, d'identifier les variables cibles de l'intervention et les indicateurs possibles. Cette familiarité avec le milieu et la population cible facilite aussi l'accès au terrain lors de la cueillette de données. Une bonne connaissance du processus d'intervention et de son contexte d'application s'avère également un atout important lors de l'analyse des données : elle guide l'identification des éléments de l'intervention aptes à expliquer les résultats obtenus, suggère des hypothèses alternatives et permet, dans certains cas, de juger de leur vraisemblance. En fait, le recours informel à des données qualitatives a souvent été évoqué par les évaluateurs de tradition plus quantitative. Pour Reichardt et Cook (1979), la démarche quantitative doit nécessairement reposer sur une connaissance préalable de nature qualitative. Cette connaissance qualitative entre en jeu dans le choix d'une stratégie d'analyse statistique, dans l'interprétation des résultats autant que dans leur généralisation. En fait, à chaque étape du processus de recherche, l'évaluateur aura recours à sa compréhension personnelle du programme, compréhension fondée sur ses contacts, ses échanges avec les intervenants ou les observations qu'il a pu faire dans le milieu de l'intervention. Certains auteurs voient même dans ce processus informel d'utilisation d'informations qualitatives pour le travail quantitatif une preuve que les deux approches sont, non seulement loin d'être opposées, mais souvent même indissociables (Campbell, 1979; Filstead, 1979).

Sieber (1973) souligne à cet égard que la familiarité du chercheur avec la situation ou la problématique à l'étude se trouve souvent à l'origine des hypothèses ou des théories qui feront l'objet d'une vérification empirique, bien qu'il en soit fait rarement mention dans le rapport final de l'enquête. On constate également l'absence, dans les rapports d'évaluation, de toute référence à la démarche de familiarisation entreprise par l'évaluateur préalablement à l'établissement du devis d'évaluation ou en cours d'évaluation. Les manuels de méthodologies quantitatives demeurent également muets quant il s'agit de traiter de cet aspect du travail du chercheur. Par contre, en matière d'évaluation, cette familiarisation se trouve explicitement reconnue comme une étape préalable à la recherche, dans la mesure où l'évaluateur doit d'abord s'assurer de l'évaluabilité du programme (Wholey, 1982).

Sans doute, le principal reproche adressé aux évaluateurs de programmes adeptes des méthodes quantitatives et des devis expérimentaux ou quasi expérimentaux est-il de mettre un accent démesuré sur la rigueur méthodologique de l'étude avec, comme conséquence, de ne pas refléter adéquatement la réalité quotidienne du programme d'intervention. Autrement dit, ces chercheurs ont trop souvent fait preuve d'une préoccupation disproportionnée quant à la validité interne de l'étude, au détriment de sa validité externe. Une telle critique suggère donc que cet effort de familiarisation demeure encore aujourd'hui trop peu répandu dans les pratiques de recherche des évaluateurs. Elle justifie également que la démarche qualitative informelle soit systématisée par ceux qui la pratiquent et soit documentée dans le rapport d'évaluation. Campbell (1979), sans doute l'auteur le plus fortement identifié à l'orthodoxie quantitative, formule d'ailleurs une telle recommandation. Il considère en effet important d'obtenir et de rapporter les réactions et les interprétations des résultats de tous les participants intéressés à l'évaluation : « L'histoire narrative est une partie indispensable du rapport final, et les meilleures méthodes qualitatives devraient être utilisées dans sa préparation » (p. 52, c'est nous qui traduisons). De façon générale, une utilisation plus structurée des techniques développées en recherche qualitative permet d'augmenter la quantité de données et de maximiser la validité des informations quantitatives recueillies.

Une des stratégies que semblent privilégier plusieurs auteurs pour assurer la validité autant externe qu'interne de l'évaluation est de faire précéder l'étude quantitative d'une démarche de recherche qualitative au cours de laquelle l'évaluateur est amené à se familiariser avec les différents aspects spécifiques du programme qu'il doit évaluer. Ainsi, Ianni et Orr (1979) suggèrent une première étape de recherche où les résultats des méthodes ethnographiques et cliniques serviront à identifier les variables importantes, à construire les instruments d'évaluation et à établir la stratégie d'analyse. Smith (1988) explique de façon détaillée comment il est possible d'extraire de textes tels que des articles de journaux, des rapports ou des transcriptions d'entrevues non structurées des énoncés pertinents pour les questions qui seront insérées dans une entrevue ou un questionnaire de nature plus quantitative.

La mise en œuvre d'une démarche qualitative préalable à l'évaluation quantitative comporte cependant certains inconvénients. En plus de nécessiter un temps plus long pour mener à bien l'évaluation, l'évaluateur formé dans une tradition quantitative est souvent peu familier avec les techniques ethnographiques ou d'observation participante. Une des solutions possibles pour contrer les inconvénients d'un manque de familiarité avec le milieu, et qui pourrait s'avérer particulièrement appropriée dans un contexte de ressources limitées, est d'adopter une démarche de consultation intensive et extensive auprès de l'ensemble des personnes impliquées dans l'évaluation, avant même d'établir le devis ou les objectifs d'évaluation. On retrouve dans la littérature en évaluation qualitative de nombreuses méthodes de travail de groupe destinées à systématiser un tel processus de consultation, que ce soit la technique des groupes nominaux (Delbecq, Van de Ven et Gutafson, 1975), la technique du groupe Delphi (Barrington et Barrington, 1986) ou la technique des *focus group* (Morgan, 1988). Dans cette optique, Chen (1990) démontre la contribution capitale d'un travail avec les groupes d'intérêt (*stakeholders*) préalable à toute démarche d'évaluation dans le but de définir la théorie du programme d'intervention et de cerner à partir d'un cadre d'analyse exhaustif le contexte global de l'intervention, contexte autant physique, social que politique. C'est à partir de ce travail de consultation que seront établies les priorités d'information de l'organisme, les objectifs de l'évaluation et, par conséquent, la méthode à utiliser.

Une autre étape où l'apport de données qualitatives peut s'avérer bénéfique, voire essentiel, est la phase de l'analyse et de l'interprétation des résultats. Un chercheur qui a une connaissance en profondeur de la réalité de l'intervention et du travail des intervenants parce qu'il a passé de longues heures sur le terrain sera plus en mesure de proposer différentes interprétations des résultats. Par ailleurs, de par leur connaissance de la gestion quotidienne du programme, les intervenants peuvent proposer des explications des résultats obtenus ou identifier la logique qui se cache derrière un profil de résultats et qui demeure obscure à l'évaluateur moins familier avec le processus de l'intervention. D'où l'intérêt de présenter les résultats préliminaires aux différents groupes impliqués (administrateurs, promoteurs, intervenants, clients) pour leur permettre de réagir et de soulever des hypothèses que l'évaluateur pourra par la suite tenter de vérifier s'il dispose des données nécessaires ou, à tout le moins, d'inclure dans sa présentation finale des résultats. Omettre une telle consultation expose l'évaluateur au risque de présenter des interprétations facilement récusables ou des conclusions peu réalistes en raison de l'omission d'éléments importants. Campbell (1979) va même jusqu'à suggérer que les résultats d'analyses quantitatives devraient être considérés avec suspicion lorsqu'ils contredisent l'évaluation qualitative du personnel ou de la clientèle, et cela tant qu'on n'a pas été en mesure d'identifier la source de cette divergence.

Une autre utilisation des données qualitatives dans le cadre d'évaluations quantitatives est celle où, à l'étape de la présentation des données, l'évaluateur se réfère à des données qualitatives pour illustrer et enrichir les données d'analyses quantitatives. On confère volontiers à ce type d'information la capacité de donner de la « profondeur » et du « sens » aux données statistiques. Plusieurs raisons justifient une telle utilisation de données qualitatives. Un des problèmes souvent évoqués dans le domaine de l'évaluation

de programmes est la faible utilisation des résultats. Pour plusieurs auteurs, une partie importante du problème pourrait résider dans le niveau de langage souvent trop technique des rapports d'évaluation ou de la présentation des résultats d'analyses dont la lecture est considérée comme trop aride (Mowbray, 1988; Patton, 1987). La présentation d'histoires de cas, la retranscription d'extraits d'entretiens seraient des moyens efficaces d'alléger la lecture, de susciter un intérêt plus grand et, également, de faciliter la compréhension des points soulevés par les évaluateurs. Comme le souligne Smith (1986), cet argument peut s'appliquer également aux recherches qualitatives dont le format peut déplaire à certains groupes ou à certains commanditaires, plus à l'aise avec une présentation formelle de données quantitatives, particulièrement dans un contexte où ceux-ci doivent justifier leur prise de décision.

Un autre avantage de la diversification des modes de présentation des résultats a été souligné par Patton (1987). Selon lui, la présentation routinière de données, sous un même format, aux responsables d'un programme, peut les amener, par habitude, à ne pas réagir aux changements qui surviennent. Par conséquent, modifier la façon de présenter les résultats peut stimuler leur attention et susciter de nouvelles questions. Certains auteurs suggèrent également que cette façon de briser la routine pourrait s'avérer bénéfique au chercheur lui-même. L'utilisation de nouvelles sources de données pourrait être en effet un excellent moyen de lancer de nouvelles pistes de recherche, de susciter de nouveaux questionnements, d'innover dans la façon de mesurer un concept, ou même de renouveler l'intérêt du chercheur pour une problématique ou ses résultats (Patton, 1987; Rossman et Wilson, 1985).

Il n'y pas que les évaluations quantitatives qui profitent de l'intégration de procédures de recherche associées à l'autre approche. En fait, malgré la polarisation entre les tenants d'approches qualitative et quantitative, on constate que la plupart des évaluateurs qualitatifs font usage, ou, à tout le moins, admettent la possibilité d'utiliser des techniques d'enquête ou de sondage, de questionnaires fermés ou de toutes autres techniques de cueillette quantitatives. Pour Guba et Lincoln (1989), « il y aura des circonstances où les méthodes quantitatives — test ou autres instruments de mesure, ou présentations numériques — seront, et devraient être utilisées » (p. 259, c'est nous qui traduisons).

L'utilisation qui semble la plus courante consiste à se référer aux résultats d'enquêtes ou de sondages au moment du processus de sélection des personnes ou des groupes qui feront l'objet d'une étude qualitative approfondie (McClintock, Braunon et Maynard-Moody, 1983). Une critique fréquente des évaluations qualitatives concerne précisément la représentativité des échantillons sélectionnés. Sieber (1973) souligne la présence d'un biais élitiste des chercheurs qualitatifs dans la sélection des informateurs clés et des informations recueillies, biais qui entraîne une surreprésentation des personnes d'un statut social plus élevé, occupant des postes plus importants ou dont la facilité d'élocution donne l'impression d'une plus grande connaissance. De façon générale, le recours aux données démographiques sur la population cible ou l'exécution d'une enquête préalable à l'étude sur le terrain pour documenter la répartition de certaines caractéristiques pertinentes peuvent guider le processus de recrutement et éviter ainsi ou corriger en cours d'étude les biais de sélection. Ce genre d'informations peut également s'avérer grandement utile lorsque la stratégie d'échantillonnage vise à identifier des cas extrêmes ou déviants, ou des exemples qui semblent confirmer ou infirmer les hypothèses de l'évaluateur. La possibilité d'établir la représentativité des échantillons peut également légitimer l'extrapolation des conclusions de l'étude qualitative. Cet avantage est d'autant plus primordial que, dans le cas des études qualitatives utilisant les récits de vie ou les histoires de cas, le faible nombre de répondants rend ces études particulièrement vulnérables à une critique de non-représentativité.

Smith (1986) souligne également que dans un contexte de conflit d'intérêts majeur entre groupes divergents, l'utilisation exclusive de méthodes qualitatives pourrait ne pas s'avérer suffisamment convainquante. Chen (1990) précise que, dans un tel contexte, les stratégies qualitatives traditionnelles s'avèrent souvent inaptes à établir un consensus ou à

entraîner une vision commune de la réalité. L'utilisation d'une méthode de sondage auprès des différentes parties pourrait alors permettre de les rassurer quant à la représentativité des points de vue considérés.

L'utilisation de méthodes quantitatives peut également contribuer au travail de cueillette de données qualitatives. Un exemple particulièrement intéressant est fourni par l'étude de Friedrich et Reams (1987) sur l'évolution des symptômes chez des enfants victimes d'abus sexuels. Dans cette recherche évaluative, les auteurs font usage de profils de changements de comportements établis à partir d'un questionnaire standardisé et procèdent ensuite à l'histoire de cas de chacun des enfants participant à l'étude.

Rossmann et Wilson (1985) démontrent également comment l'obtention de données quantitatives peut permettre de documenter la répartition des catégories d'une typologie créée à partir d'analyses qualitatives, de mieux comprendre ainsi ces données et suggérer de nouvelles interprétations. Ils font de la sorte la démonstration que les informations quantitatives peuvent ajouter de la profondeur aux résultats qualitatifs et en faciliter l'interprétation. Sieber (1973) démontre également comment une technique d'enquête peut permettre d'atténuer le problème de l'illusion holistique (*holistic fallacy*) souvent présent dans les études qualitatives. Ce problème résulte d'une tendance du chercheur à percevoir les différents aspects de la situation comme plus congruents qu'ils ne le sont en réalité et à minimiser l'importance d'éléments inconsistants avec la vision d'ensemble. L'obtention de données d'enquête sur les différents aspects étudiés permet d'identifier et de corriger un tel biais dans la perception.

Cette énumération des circonstances où l'on peut avantageusement intégrer les méthodes qualitatives et quantitatives au sein d'une même étude ne peut être exhaustive. Elle démontre cependant l'intérêt pour l'évaluateur de diversifier ses techniques de recherche, quel que soit le courant qu'il privilégie.

CONCLUSION

Dans cet article, nous avons tenté de faire ressortir les efforts en évaluation de programmes visant à dépasser l'opposition encore trop fréquente entre le qualitatif et le quantitatif. Quatre stratégies différentes ont ainsi été identifiées, toutes susceptibles, à leur manière, de contribuer au développement des méthodologies de recherche et à mieux répondre ainsi aux problèmes pratiques de l'évaluation.

Il n'en demeure pas moins qu'il existe souvent un écart considérable entre les possibilités offertes par les innovations méthodologiques et leur application dans l'évaluation d'un programme donné. On ne peut que constater, à l'instar de Lipsey (1988), que, malgré la panoplie des méthodes disponibles et les possibilités offertes par leur intégration, les évaluateurs se cantonnent souvent dans des pratiques routinières d'évaluation, et se satisfont de la mise en œuvre de modèles simplistes. Bien des méthodes demeurent trop peu utilisées par manque de familiarité, ou pour des raisons d'économie de temps ou d'argent. Une partie du problème réside sans doute dans la formation académique des évaluateurs jugée par plusieurs auteurs comme nettement déficiente (Sechrest, 1992). À cet égard, on ne peut que déplorer la pénurie de programmes de formation universitaire en évaluation propres à favoriser l'apprentissage des multiples méthodologies et à sensibiliser les étudiants au contexte particulier du travail d'évaluation.

Un autre obstacle à l'innovation méthodologique tient au conservatisme inhérent aux processus de révision des projets de recherche et des programmes d'intervention. Qu'il s'agisse de juger de la valeur d'un devis d'évaluation ou de l'opportunité de poursuivre, de modifier ou d'interrompre un programme, la référence à des procédures classiques et éprouvées, référence souvent peu réaliste si on tient compte du contexte, tend généralement à l'emporter sur des approches éventuellement mieux adaptées, mais moins reconnues dans la communauté scientifique ou par les organismes gestionnaires.

Cependant, le caractère essentiellement appliqué du travail en évaluation de programmes le rend propre à favoriser le développement et l'innovation méthodologiques.

Les conditions d'opération souvent difficiles, mais aussi bien les questions d'éthique, empêchent l'application intégrale de techniques et de devis de recherches standards. L'évaluateur doit s'intégrer au milieu et adapter les procédures de recherche aux contraintes liées au déroulement du programme et à son environnement organisationnel. Le contexte politique dans lequel s'inscrit généralement la demande d'évaluation et la nécessité de tenir compte des exigences et des préférences de multiples groupes d'intérêt l'obligent à s'interroger sur les valeurs sous-jacentes aux différents choix méthodologiques et à négocier avec les groupes d'intérêt les éléments du processus de recherche. Enfin, la nécessité pour l'évaluateur de fournir des informations utiles en temps requis et sous une forme compréhensible pour différents auditoires l'amène également à modifier ses pratiques habituelles. Autant d'éléments qui favorisent l'ouverture et le rapprochement des différentes pratiques de recherche, y compris entre les approches qualitative et quantitative.

Normand PÉLADEAU et Céline MERCIER
Unité de recherche psychosociale
Centre de recherche de l'hôpital Douglas
6875, boul. LaSalle
Verdun (Québec)
Canada H4H 1R3

RÉSUMÉ

Dans le domaine de l'évaluation de programmes, il est courant d'opposer évaluation qualitative et quantitative. Cet article remet en question le postulat suivant lequel approche qualitative et approche quantitative relèveraient de paradigmes différents. Il nous apparaît qu'une telle position correspond peu aux pratiques réelles et est de fait néfaste au développement des méthodologies en matière d'évaluation de programmes. Un examen des travaux récents en évaluation de programmes nous a permis d'identifier quatre démarches différentes qui nous semble susceptibles de faire évoluer la pratique de l'évaluation, à partir des critiques que s'adressent réciproquement les tenants des approches quantitative ou qualitative. La première de ces stratégies est celle de l'utilisation exclusive d'une approche, qu'elle soit qualitative ou quantitative. La seconde se réfère à l'utilisation de l'une ou l'autre des approches, suivant le contexte. Les troisième et quatrième combinent les deux approches, sous forme de triangulation ou d'intégration par combinaison.

SUMMARY

In program evaluation, it is common to oppose qualitative and quantitative evaluation. This paper questions the postulate which states that qualitative and quantitative approaches belong to different paradigms. Such a position appears to correspond little to real practices and is, in fact, very harmful to the development of methodologies for program evaluation. An examination of recent studies in program evaluation makes it possible to identify four different approaches liable, in our view, to bring about evolution in evaluation practices, based on criticisms which tenants of both quantitative and qualitative approaches have made of one another. The first of these strategies is the exclusive use of one approach, be it qualitative or quantitative. The second calls upon one or the other approach, depending on the context. The third and fourth imply simultaneous recourse to both approaches, through triangulation or integration by combination.

RESUMEN

Dentro del dominio de la evaluación de programas, es corriente oponer evaluación cualitativa y cuantitativa. Este artículo cuestiona el postulado según el cual los métodos cualitativos y cuantitativos dependen de paradigmas diferentes. Nos parece que una tal posición corresponde poco a las prácticas reales y es de hecho nefasta al desarrollo de metodologías en evaluación de programas. Un examen de los trabajos recientes en evaluación de programas nos ha permitido identificar cuatro tipos de gestión diferentes que nos parecen susceptibles de hacer evolucionar la práctica de la evaluación, a partir de las críticas que se hacen recíprocamente los partidarios de los métodos cualitativos y cuantitativos. La primera de esas estrategias es aquella de la utilización exclusiva de una metodología, ya sea cualitativa o cuantitativa. La segunda se refiere a la utilización de una u otra de estas metodologías, según sea el contexto. La tercera y la cuarta implican el recurso simultáneo a las dos metodologías, bajo la forma de triangulación o de integración por combinación.

BIBLIOGRAPHIE

- ALEXANDER, H. A. (1986), « Cognitive relativism in evaluation », *Evaluation Review*, vol. 10, n° 3, pp. 259-280.
- BARRINGTON, G. V., BARRINGTON, G. V. (1986), « The Delphi as a naturalistic evaluation tool », *The Canadian Journal of Program Evaluation*, vol. 1, n° 2, pp. 81-88.
- CAMPBELL, D. T., FISKE, D. W. (1959), « Convergent and discriminant validation by the multitrait-multimethod matrix », *Psychological Bulletin*, vol. 56, pp. 81-105.
- CAMPBELL, D. T. (1979), « Degrees of freedom » and the case study », in T. D. Cook, C. S. Reichardt (eds.), *Qualitative and Quantitative Methods in Evaluation Research*. Beverly Hills, CA, Sage Publications, pp. 49-67.
- CHEN, H.-T. (1990), *Theory-Driven Evaluations*. Beverly Hills, CA, Sage Publications.
- COOK, T. J., POOLE, W. K. (1982), « Treatment implementation and statistical power ». *Evaluation Review*, vol. 6, n° 3, pp. 425-430.
- CRONBACK, L. J., SUPPES, P. (1969), *Research for tomorrow's schools : Disciplined inquiry in education*, New York, Macmillan.
- DELBECQ, A. L., VAN DE VEN, A. H., GUSTAFSON, D. H. (1975), *Group Techniques for Program Planning*, Scott, Foresman.
- DENZIN, N. K. (1970), *The Research Act : A Theoretical Introduction to Sociological Methods*, New York, NY, Aldine.
- FETTERMAN, D. M. (1992), « In response to Lee Sechrest's 1991 AEA presidential address : « Roots : Back to our first generations », *Evaluation Practice*, vol. 13, n° 3, pp. 171-172.
- FILSTEAD, W. J. (1979), « Qualitative methods : A needed perspective in evaluation research », in T. D. Cook, C. S. Reichardt (eds.), *Qualitative and Quantitative Methods in Evaluation Research*, Beverly Hills, CA, Sage Publications, pp. 33-48.
- FRIEDRICH, W. N., REAMS, R. A. (1987), « Course of psychological symptoms in sexually abused young children », *Psychotherapy*, vol. 24, n° 2, pp. 160-170.
- GREENE, J., McCLINTOCK, C. (1985), « Triangulation in evaluation. Design and analysis issues », *Evaluation Review*, vol. 9, n° 5, pp. 523-545.
- GUBA, E. G., LINCOLN, Y. S. (1989), *Fourth Generation Evaluation*, Beverly Hills, CA, Sage Publications.
- HEILMAN, J. G. (1980), « Paradigmatic choices in evaluation methodology », *Evaluation Review*, vol. 4, n° 5, pp. 693-712.
- HORN, W. F., HEERBOTH, J. (1982), « Single-case experimental designs and program evaluation », *Evaluation Review*, vol. 6, n° 3, pp. 403-424.
- IANNI, F. A. J., ORR, M. T. (1979), « Toward a rapprochement of quantitative and qualitative methodologies », in T. D. Cook, C. S. Reichardt (eds.), *Qualitative and Quantitative Methods in Evaluation Research*. Beverly Hills, CA, Sage Publications, pp. 87-98.
- JICK, T. D. (1983), « Mixing qualitative and quantitative methods : Triangulation in action », in J. V. Maanen (eds.), *Qualitative Methodology*. Beverly Hills, CA, Sage Publications, pp. 135-148.
- KIDDER, L. H., FINE, M. (1987), « Qualitative and quantitative methods : When stories converge », in M. M. Mark, R. L. Shotland (eds.), *Multiple Methods in Program Evaluation. New Directions for Program Evaluation* (n° 35). San Francisco, Jossey-Bass, pp. 57-76.
- KIRK, J., MILLER, M. L. (1986), *Reliability and Validity in Qualitative Research*. Beverly Hills, CA, Sage Publications.
- KRATOCHWILL, T. R., LEVIN, J. R. (1978), « What time-series designs may have to offer educational researchers », *Contemporary Educational Psychology*, vol. 3, n° 4, pp. 273-329.
- KUHN, T. S. (1970), *The Structure of Scientific Revolutions* (2^e édition). Chicago, Ill., University of Chicago Press.
- LINCOLN, Y. S. (1991), « The arts and sciences of program evaluation », *Evaluation Practice*, vol. 12, n° 1, pp. 1-8.
- LINCOLN, Y. S., GUBA, E. G. (1992), « In response to Lee Sechrest's 1991 AEA Presidential Address : « Roots : Back to our first generations », *Evaluation Practice*, vol. 13, n° 3, pp. 165-170.
- LIPSEY, M. W., CORDRAY, D. S., BERGER, D. E. (1981), « Evaluation of a juvenile diversion program : Using multiple lines of evidence », *Evaluation Review*, vol. 5, n° 3, pp. 283-306.
- LIPSEY, M. W. (1988), « Practice and malpractice in evaluation research », *Evaluation Practice*, vol. 9, n° 4, pp. 5-24.
- MARK, M. M. (1983), « Treatment implementation, statistical power, and internal validity », *Evaluation Review*, vol. 7, n° 4, pp. 543-549.
- MARK, M. M., SHOTLAND, R. L. (1987), « Alternative models for the use of multiple methods », in M. M. Mark, R. L. Shotland (eds.), *Multiple Methods in Program Evaluation. New Directions for Program Evaluation* (n° 35), San Francisco, Jossey-Bass Inc., pp. 95-100.
- McCLINTOCK, C. C., BRANNON, D., MAYNARD-MOODY, S. (1983), « Applying the logic of sample surveys to qualitative case studies : The case cluster method », in J. V. Maanen (eds.), *Qualitative Methodology*, Beverly Hills, CA, Sage Publications, pp. 149-178.

- MECHANIC, D. (1989), « Medical sociology : Some tensions among theory, method, and substance », *Journal of Health and Social Behavior*, vol. 30, pp. 147-160.
- MORGAN, D. L. (1988), *Focus Groups as Qualitative Research*, Beverly Hills, CA, Sage Publications.
- MOWBRAY, C. T. (1988), « Getting the system to respond to evaluation findings » in J. A. McLaughlin, L. J. Weber, R. W. Covert, R. B. Ingle (eds.), *Evaluation Utilization. New Directions for Program Evaluation* (n° 39), San Francisco, Jossey-Bass Inc., p. 47-68.
- NUEHRING, E. M., PASCONE, A. B. (1986), « Single-subject evaluation : A tool for quality assurance », *Social Work*, vol. 31, n° 5, pp. 359-365.
- PATTON, M. Q. (1978), *Utilization-Focused Evaluation*, Beverly Hills, CA, Sage Publications.
- PATTON, M. Q. (1982), *Practical Evaluation*. Beverly Hills, CA, Sage Publications.
- PATTON, M. Q. (1987), *How to Use Qualitative Methods in Evaluation*, Beverly Hills, CA, Sage Publications.
- REICHARDT, C. S., COOK, T. D. (1979), « Beyond qualitative versus quantitative methods », in C. S. Reichardt, T. D. Cook (eds.), *Qualitative and Quantitative Methods in Evaluation Research*, Beverly Hills, CA, Sage Publications, pp. 7-32.
- ROSSI, P. H., BERK, R. A. (1981), « An overview of evaluation strategies and procedures », *Human Organization*, vol. 40, n° 4, pp. 287-299.
- ROSSMAN, G. B., WILSON, B. L. (1985), « Numbers and words. Combining quantitative and qualitative methods in a single large-scale evaluation study », *Evaluation Review*, vol. 9, n° 5, pp. 627-643.
- SCRIVEN, M. C. (1989), « Evaluation ideologies, in G. F. Madaus, M. S. Scriven, D. L. Stufflebeam (eds.), *Evaluation Models. Viewpoints in Educational and Human Services Evaluation*, Boston, MA, Kluwer-Nijhoff.
- SECHREST, L. (1992), « Roots : Back to our first generations », *Evaluation Practice*, vol. 13, n° 1, pp. 1-8.
- SHADISH, W. R., COOK, T. D., LEVITON, L. C. (1991), *Foundations of Program Evaluation. Theories of Practice*. Beverly Hills, CA, Sage Publications.
- SHOTLAND, R. L., MARK, M. M. (1987), « Improving inferences from multiple methods », in M. M. Mark, R. L. Shotland (eds.), *Multiple Methods in Program Evaluation. New Directions for Program Evaluation* (n° 35), San Francisco, Jossey-Bass Inc., pp. 77-94.
- SIEBER, S. D. (1973), « The integration of fieldwork and survey methods », *American Journal of Sociology*, vol. 73, pp. 1335-1359.
- SILVERMAN, M., RICCI, E. M., GUNTER, M. J. (1990), « Strategies for increasing the rigor of qualitative methods in evaluation of health care programs », *Evaluation Review*, vol. 14, n° 1, pp. 57-74.
- SMITH, M. L. (1986), « The whole is greater : Combining qualitative and quantitative approaches in evaluation studies », in D. D. Williams (eds.), *Naturalistic Evaluation. New Directions for Program Evaluation* (n° 30), San Francisco, Jossey-Bass, pp. 37-54.
- SMITH, R. B. (1988), « Linking quality et quantity. Part II. Surveys as formalizations », *Quality and Quantity*, vol. 22, n° 1, pp. 3-30.
- TREND, M. G. (1979), « On the reconciliation of qualitative and quantitative analyses : A case study », in T. D. Cook, C. S. Reichardt (eds.), *Qualitative and Quantitative Methods in Evaluation Research*, Beverly Hills, CA, Sage Publications, pp. 68-86.
- TROCHIM, W. M. K. (1989), « Outcome pattern matching and program theory », *Evaluation and Program Planning*, vol. 12, n° 4, pp. 355-366.
- WEBB, E. T., CAMPBELL, D. T., SCHWARTZ, R. D., SECHREST, L., GROVE, J. B. (eds.) (1981), *Non-reactive Measures in the Social Sciences*, Boston, Houghton Mifflin Company.
- WHOLEY, J. S. (1982), « L'appréciation de l'évaluabilité d'un programme », in R. Lecomte, L. Rutman (eds.), *Introduction aux méthodes de recherche évaluative*. Québec, Presses de l'Université Laval, pp. 47-63.