Raymond MASSÉ

Anthropologue, spécialiste en anthropologie de la santé
Professeur titulaire, département d'anthropologie, Université Laval.

(1993)

“Réflexions anthropologiques sur la fétichisation des méthodes en évaluation.”

Un document produit en version numérique par Jean-Marie Tremblay, bénévole,
professeur de sociologie au Cégep de Chicoutimi

Courriel: jean-marie_tremblay@uqac.ca

Site web pédagogique : <http://www.uqac.ca/jmt-sociologue/>

Dans le cadre de: "Les classiques des sciences sociales"

Une bibliothèque numérique fondée et dirigée par Jean-Marie Tremblay,
professeur de sociologie au Cégep de Chicoutimi

Site web: <http://classiques.uqac.ca/>

Une collection développée en collaboration avec la Bibliothèque
Paul-Émile-Boulet de l'Université du Québec à Chicoutimi

Site web: <http://bibliotheque.uqac.ca/>

Politique d'utilisation de la bibliothèque des Classiques

Toute reproduction et rediffusion de nos fichiers est interdite, même avec la mention de leur provenance, sans l'autorisation formelle, écrite, du fondateur des Classiques des sciences sociales, Jean-Marie Tremblay, sociologue.

Les fichiers des Classiques des sciences sociales ne peuvent sans autorisation formelle:

- être hébergés (en fichier ou page web, en totalité ou en partie) sur un serveur autre que celui des Classiques.
- servir de base de travail à un autre fichier modifié ensuite par tout autre moyen (couleur, police, mise en page, extraits, support, etc...),

Les fichiers (.html, .doc, .pdf., .rtf, .jpg, .gif) disponibles sur le site Les Classiques des sciences sociales sont la propriété des **Classiques des sciences sociales**, un organisme à but non lucratif composé exclusivement de bénévoles.

Ils sont disponibles pour une utilisation intellectuelle et personnelle et, en aucun cas, commerciale. Toute utilisation à des fins commerciales des fichiers sur ce site est strictement interdite et toute rediffusion est également strictement interdite.

L'accès à notre travail est libre et gratuit à tous les utilisateurs. C'est notre mission.

Jean-Marie Tremblay, sociologue
Fondateur et Président-directeur général,
[LES CLASSIQUES DES SCIENCES SOCIALES](#).

Cette édition électronique a été réalisée par Jean-Marie Tremblay, bénévole, professeur de sociologie au Cégep de Chicoutimi à partir de :

Raymond MASSÉ

Anthropologue, spécialiste en anthropologie de la santé, Professeur titulaire, département d'anthropologie, Université Laval.

“Réflexions anthropologiques sur la fétichisation des méthodes en évaluation.”

Un article publié dans l'ouvrage sous la direction de Judith Légaré et Andrée Demers, **L'évaluation sociale : savoirs, éthique, méthodes**, pp. 209-238. Actes du 59e congrès de l'ACSALF, 1991. Montréal, Les Éditions du Méridien, 1993, 372 pp.

[Autorisation formelle accordée par l'auteur le 14 novembre 2008 de diffuser cette œuvre dans Les Classiques des sciences sociales.]

Courriel : Raymond.Masse@ant.ulaval.ca

Polices de caractères utilisée :

Pour le texte: Times New Roman, 12 points.

Pour les citations : Times New Roman, 12 points.

Pour les notes de bas de page : Times New Roman, 12 points.

Édition électronique réalisée avec le traitement de textes Microsoft Word 2008 pour Macintosh.

Mise en page sur papier format : LETTRE (US letter), 8.5'' x 11''

Édition numérique réalisée le 2 mai 2009 à Chicoutimi, Ville de Saguenay, province de Québec, Canada.

Raymond MASSÉ

Anthropologue, spécialiste en anthropologie de la santé
Professeur titulaire, département d'anthropologie, Université Laval.

“Réflexions anthropologiques sur la fétichisation des méthodes en évaluation.”

Un article publié dans l'ouvrage sous la direction de Judith Légaré et Andrée Demers, **L'évaluation sociale : savoirs, éthique, méthodes**, pp. 209-238. Actes du 59e congrès de l'ACSALF, 1991. Montréal, Les Éditions du Méridien, 1993, 372 pp.

Table des matières

Introduction

Le fétichisme des outils et devis

L'évaluation des programmes de prévention des abus et négligences envers les enfants

LA FÉTICHISATION DES DEVIS D'ÉVALUATION.

Les conditions d'application des devis expérimentaux

LA FÉTICHISATION DES OUTILS DE MESURE : LE MYTHE DES ÉCHELLES ET OUTILS VALIDÉS.

Les défis posés par la mesure en évaluation

Conclusion

RÉFÉRENCES BIBLIOGRAPHIQUES

Raymond MASSÉ

**“Réflexions anthropologiques sur la fétichisation
des méthodes en évaluation”.**

Un article publié dans l'ouvrage sous la direction de Judith Légaré et Andrée Demers, **L'évaluation sociale : savoirs, éthique, méthodes**, pp. 209-238. Actes du 59e congrès de l'ACSALF, 1991. Montréal, Les Éditions du Méridien, 1993, 372 pp.

INTRODUCTION

[Retour à la table des matières](#)

Les évaluations de programmes de prévention ou de promotion de la santé sont actuellement dominées au Québec par des approches et des méthodes mal adaptées aux attentes de l'intervention. Sans tenir compte des besoins souvent peu ou mal exprimés des gestionnaires, les chercheurs ont tendance à privilégier les études d'impact de ces Programmes sur l'état de santé de la population-ciblée et ce, sans vraiment questionner l'adéquation de l'implantation, de la structure et du fonctionnement du programme. Le peu de préoccupation face à l'évaluation du fonctionnement des programmes a conduit à un désintérêt du vécu et des appréciations des bénéficiaires et des intervenants à l'intérieur de ces programmes. Or, bien que le rendement des programmes mérite d'être étudié, les finalités de la recherche évaluative ne peuvent se réduire à cette seule dimension (Rutman, Mowbray, 1983).

Nous proposons dans le texte qui suit que le recours abusif et inadéquat aux outils et données jugées « dures » (aux dépens d'une pluralité de méthodes qui comportent aussi des données « molles ») est conséquent à une forme de fétichisation des devis expérimentaux et des outils dits validés de cueillette de données. Il est temps de requestionner le recours abusif aux devis expérimentaux et aux outils « validés » dans le domaine des interventions psycho-sociales. Nous plaiderons ici

en faveur de la conjugaison d'une pluralité d'approches méthodologiques, qualitatives et quantitatives, en recherche évaluative.

Cet abus dans le recours aux méthodologies « objectives » s'exprime de deux façons. D'abord dans l'exclusion de toutes méthodes ou outils qualitatifs dans la mesure des résultats ou des conséquences d'un programme. Ensuite, dans la préférence donnée aux évaluations de résultats et d'impact aux dépens des évaluations centrées sur l'implantation, la mise en oeuvre et l'accueil du programme sur le terrain. On privilégie actuellement, à tort, une évaluation visant à porter un jugement sur le programme à travers ses résultats et impacts (évaluation sommative) aux dépens d'une évaluation visant à produire les informations permettant de l'améliorer, l'adapter et en faciliter la gestion (évaluation formative).

Notre réflexion sur cette fétichisation des méthodes quantitatives en évaluation aura pour cadre la notion de programme et s'inspirera plus particulièrement des expériences d'évaluation de programmes de prévention des abus et/ou négligences envers les enfants. Un premier postulat de base sera que tout programme est constitué de plusieurs composantes qui chacune mérite d'être évaluée. Les principales sont les objectifs, les ressources investies (*input*), les services mis en oeuvre (*output*), les résultats (*outcome*) et les conséquences (*impact*) (Champagne *et al.* 1985). Un second postulat veut que l'opposition des concepts d'évaluation quantitative et d'évaluation qualitative soit quant à elle vide de sens. S'engager dans ce genre de débat revient, croyons-nous, à consacrer la primauté de la méthode sur les finalités de l'évaluation. Or, les évaluations de structure, de processus, de résultats, de coûts-bénéfices, etc., auront, dans des proportions variables, tour à tour besoin de méthodes dites quantitatives et qualitatives de cueillette et d'analyse de données.

Le fétichisme des outils et devis

[Retour à la table des matières](#)

Le paradigme hypothético-déductif a servi de cadre de référence quasi-exclusif à des dizaines de recherches visant l'évaluation de diverses interventions préventives dans le domaine de la santé et des services sociaux au Québec au cours des dix dernières années. Malgré l'importance des sommes consacrées, le

bilan général est catastrophique. Non pas que ces évaluations aient été faites par des chercheurs incompetents ou que les techniques utilisées aient manqué de rigueur scientifique. Le problème de base est que l'on n'a pas su viser les objectifs les plus pertinents et que l'on n'a pas su utiliser les outils appropriés aux objets évalués. Dans certains cas, les rapports postulés entre les interventions et les effets attendus ne reposent pas sur des modèles théoriques solides. Ailleurs les évaluateurs sont, lors du dépôt du rapport final, après investissement d'importantes ressources humaines et matérielles, dans l'incapacité de décrire ce qui s'est effectivement passé lors des interventions. Ils ne peuvent donc identifier les composantes du programme qui ont conduit aux échecs ou aux succès mesurés. On ne trouve pas réponse aux questions suivantes : Comment l'intervention a-t-elle fonctionné sur le terrain ? Qu'ont fait concrètement les intervenants sur le terrain ? Comment ont-ils appliqué l'intervention ? Comment cette intervention a-t-elle été reçue par les bénéficiaires ? Les services ont-ils été donnés avec la même diligence, la même intensité, le même dévouement, le même synchronisme tout au long du programme ? Bref, les apprentis sorciers de l'évaluation sont trop souvent incapables de fournir aux gestionnaires ou commanditaires un cahier de recommandations sur les modifications à apporter 1) au choix et à la formation des intervenants 2) à la nature ou l'intensité des interventions 3) aux critères optimaux de définition de la clientèle-cible ou 4) à la procédure d'implantation du programme.

Nous proposons que le principal responsable de ce bilan décevant est une sorte de culte de la quantification et sa vénération aveugle de deux instruments « fétiches » que sont les devis expérimentaux et les outils validés de mesure des variables à l'étude. Rappelons qu'au sens minimal un « fétiche » est un objet de culte auquel on attribue un pouvoir magique. Le fétichisme est le culte fondé sur la vénération de fétiches. La fétichisation devient donc le processus par lequel l'homme croit pouvoir contrôler un environnement donné, naturel ou surnaturel, à l'aide d'outils aux pouvoirs non contestables (sous peine d'hérésie). Je fais ici l'hypothèse que les devis d'évaluation expérimentaux et les outils de mesure validés font trop souvent office de fétiches entre les mains des chercheurs. Ces derniers les utilisent pour asseoir leur contrôle sur l'évaluation d'un programme donné. Ils jouent le rôle d'amulettes que l'on agite pour faire fuir les biais qui menacent la validité des résultats de l'évaluation. Parés de ces fétiches, les chercheurs s'imposent aux intervenants et bénéficiaires comme les sorciers du culte, seuls dépositai-

res légitimes du savoir. La manipulation de ces fétiches sans initiation étant dangereuse, on maintient ainsi une distance respectable entre initiés de l'évaluation et destinataires des résultats de l'évaluation.

Même si l'importance des approches qualitatives en évaluation est reconnue depuis longtemps (Patton, 1980 ; Preissle Goetz et LeCompte, 1984) et que ces méthodes qualitatives peuvent trouver des applications pertinentes pour l'évaluation des programmes de santé mentale (Rocheleau, 1982) ou dans le large champ de la prévention primaire (Willms *et al.*, 1990), force est de constater qu'en santé communautaire, on ne prend pas toujours la distance nécessaire face à une évaluation de programme emprisonnée dans un culte de la quantification qui, comme tous les cultes, a ses prêtres, ses credos, ses mythes et ses fétiches.

L'évaluation des programmes de prévention des abus et négligences envers les enfants

[Retour à la table des matières](#)

La prévention des abus et négligences appelle la mise en oeuvre d'un large spectre de programmes. On retrouve à ce titre des maisons communautaires de soutien aux parents, des programmes visant le dépistage précoce à l'école, des programmes de sensibilisation des intervenants, éducateurs, enfants et du grand public à cette problématique, des programmes visant le développement de liens parent-enfant, etc. (Massé, 1991).

Toutefois, les programmes qui ont fait l'objet d'une évaluation structurée sont essentiellement ceux visant à donner un soutien au rôle parental via un suivi intensif à domicile des nouveaux parents, suivi expérimenté avec des intervenants tant professionnels que non professionnels.

À partir des revues de littérature américaines portant sur la prévention des abus et négligence faites par Helfer (1982) ou plus récemment par Daro (1988) et après considération des expériences québécoises récentes, nous avons identifié neuf recherches évaluatives majeures portant sur des programmes de soutien au rôle parental dans le cadre de la prévention de la maltraitance (Ashem et Kurtz, 1987 ; Barth et coll., 1988 ; Durand et coll., 1988 ; Gray et coll., 1979 ; Larson,

1980 ; Lyons-Ruth et coll., 1990 ; Olds, 1982 et 1986 ; Séguin et coll., 1986 ; Siegel et coll., 1980). Toutes sont basées sur des devis expérimentaux ou quasi-expérimentaux. Il s'agit invariablement d'évaluations de projets pilotes qui consistent en des visites à domicile faites par des infirmières, médecins ou psychologues ou par des aidants naturels dans les mois suivants la naissance de l'enfant. Certains programmes incluent aussi des visites pré-natales. Les services offerts par ces programmes sont tantôt précisément définis à l'avance, tantôt assez souples dans leur application. Si le but ultime des programmes est la prévention de la négligence, les objectifs intermédiaires poursuivis vont de l'amélioration des connaissances sur les soins de santé à l'enfant, les techniques de stimulation, les étapes du développement de l'enfant, à la mobilisation du réseau de soutien social des parents en passant par le soutien psychologique aux mères, l'aide matérielle ou le soutien émotif aux parents.

Une analyse de ces recherches évaluatives met en évidence l'absence quasi-totale de considération pour une évaluation de la structure et du processus des interventions ou pour les méthodes qualitatives. L'évaluation des résultats prime sur la compréhension des interventions elles-mêmes. Bref, ces évaluations illustrent on ne peut plus clairement un certain aveuglement paradigmatique (Lecomte, 1982a). Dans les pages qui suivent, nous mettrons l'accent sur les abus ou les situations de non-pertinence du recours aux devis expérimentaux et aux outils de mesure standardisés, afin de sensibiliser les chercheurs à l'impératif de mettre les moyens au service des finalités de la recherche évaluative.

LA FÉTICHISATION DES DEVIS D'ÉVALUATION.

[Retour à la table des matières](#)

Les devis expérimentaux reposent sur un paradigme hypothético-déductif fondé sur la vérification d'hypothèses prédéfinies, à l'aide de données « objectives » considérées comme traduisant fidèlement les faits. L'essai contrôlé avec groupes expérimental et contrôle randomisés sert d'étalon de mesure à partir duquel sera évaluée la qualité des autres devis d'évaluation (jugés de pré ou quasi-expérimentaux). Seul ce type de devis permettrait de contrôler la plus large part

des biais qui menacent la validité interne (degré de certitude avec laquelle les variations observées au niveau des variables dépendantes sont dues à l'intervention) et la validité externe (capacité de généraliser la relation intervention-résultats à d'autres contextes d'intervention) de la recherche. Le recours aux devis expérimentaux repose sur une condition préalable qu'est l'identification des objectifs « attendus » du programme évalué et la définition de variables qui en traduisent l'essence. Ces variables dépendantes devront être suffisamment sensibles pour refléter adéquatement les modifications obtenues dans les facteurs qu'elles mesurent (Rossi et coll., 1979). Elles devront traduire les effets spécifiques du programme évalué en fonction des relations qui semblent, en pré-expérimentation, refléter les liens les plus plausibles et vérifiables entre les inputs et les *outputs* (Chen et Rossi, 1983).

Les conditions d'application des devis expérimentaux

[Retour à la table des matières](#)

Une fois considérés ces paramètres d'application des devis expérimentaux et quasi-expérimentaux, on peut identifier une série de problèmes majeurs qui en invalident le recours pour l'évaluation des programmes de prévention de problèmes psycho-sociaux. Ces problèmes découlent du fait que ce type de programme, et en particulier les programmes de prévention des abus et négligences, ne rencontre généralement pas les conditions préalables à une application judicieuse des méthodes expérimentales. Parmi ces conditions, les cinq principales sont l'homogénéité ou la stabilité des interventions ; un programme offrant des services avec une intensité suffisante ; la possibilité d'identifier en pré-expérimentation les effets traduits en un nombre limité d'objectifs immuables ; la répliquabilité du programme d'un environnement à un autre et la possibilité de contrôler parfaitement les caractéristiques des sujets visés.

1) La structure même des devis expérimentaux présuppose que *l'intervention devrait être homogène* pour toutes les personnes qui en bénéficieront. Le devis expérimental requiert implicitement que tous les bénéficiaires recevront des servi-

ces de même nature, avec le même dosage, par des intervenants semblables, dans des contextes comparables, etc. Bref, il requiert une rigidité dans le programme.

Consciemment ou non, on considère le programme d'intervention comme un kit monolithique appliqué intégralement, sans adaptations notables, d'un individu à l'autre. Même si parfois on est conscient de l'existence de multiples variantes individuelles du programme, le résultat est le même : on en repousse l'étude du revers de la main. On est alors plus en mesure de certifier que le programme évalué fut bien le programme planifié. Les devis expérimentaux fondés sur des variables dépendantes non flexibles, qui sont par surcroît les mêmes pour chacun des bénéficiaires, ne peuvent composer avec le « risque » qu'il y ait eu en réalité autant de programmes que de bénéficiaires. Face à l'impossibilité de décomposer le programme en une multitude de variables couvrant l'ensemble des variantes de l'intervention et face à l'exigence de flexibilité posée par les administrateurs pour modifier le programme en réponses à des circonstances changeantes (Hudson, 1982), on considérera « le programme » comme une seule variable indépendante.

En fait, les problématiques et faisceaux de problèmes (isolement, toxicomanies, conflits conjugaux, violence conjugale, pauvreté, etc.) rencontrées dans les familles abusives ou négligentes sont tellement spécifiques à chacune d'elles qu'il risque d'y avoir autant de formes d'intervention, de nature, de durée et d'intensité variables que de bénéficiaires. Cette diversité fut reconnue et mise en évidence par certaines recherches évaluatives tel le projet « De la Visite » (Massé, 1988). Dans ce dernier exemple, même si les deux grands objectifs initiaux étaient la mobilisation du soutien social et l'amélioration des compétences parentales et environnementales, les intervenantes non-professionnelles impliquées dans le suivi des nouveaux parents avaient comme consigne d'adapter la nature, l'intensité et l'étendue de leurs interventions aux problématiques particulières de chaque famille. D'ailleurs comme c'est fréquemment le cas dans ce genre de programme, ce sont les bénéficiaires des services qui déterminent la fréquence, l'intensité voire la nature des services. Cette flexibilité dans l'investissement de ressources (variables d'expérimentation) et dans le choix des problématiques visées par l'intervention (variables dépendantes) constituant la principale force de ce programme, a toutefois eu comme effet d'invalider plusieurs volets du devis expérimental. L'intervention a porté sur les variables dépendantes pressenties dans moins de la moitié des cas et encore là, seules certaines des composantes de l'isolement social et

de la compétence parentale ou environnementale ont été concernées. L'intensité des services fournis ne corrèle aucunement avec les résultats obtenus (si définis quantitativement) dans la mesure où les familles contactées plusieurs fois par semaine et sur plusieurs mois étaient celles qui, au départ, présentaient le plus de problèmes et le moins de potentiel de changement. Sans être vraiment en mesure de le quantifier, l'équipe d'évaluation concluait que les effets de l'intervention étaient plus indirects : « sécurisation » des parents, réduction du stress conjugal, amélioration de l'estime de soi chez la mère, prévention d'épisodes de violence par la simple présence sur place d'une intervenante, etc.

Le problème n'est pas tant que l'on ne puisse viser à une homogénéisation des services (en en contrôlant au maximum tous les aspects). La réalité est que, dans la plupart des problématiques psycho-sociales, une telle homogénéisation des services n'est pas souhaitable. Dans le domaine des programmes de soutien au rôle parental, on s'entend aujourd'hui, au contraire, (Bouchard, 1989) pour considérer comme une composante essentielle de toute intervention préventive ou thérapeutique, qu'elle soit souple, adaptée aux besoins spécifiques et aux attentes des parents.

Or, d'évidence, certains programmes ont encadré les interventions dans un moule qui ne présente pas cette flexibilité. À titre d'exemple, on peut référer au projet d'éducation et de soutien aux parents expérimenté par Séguin et coll. (1986) qui établissait précisément les thèmes à aborder avec les parents lors de chacune des six rencontres pré-natales. De même, le programme de visites pré et post-natales évalué par Larson (1980) était bâti sur un protocole prédéfini de conseils et d'informations à donner aux parents par des étudiants en psychologie de l'enfance.

Cette homogénéisation des interventions se trouve aussi limitée par l'impératif d'adapter les programmes aux besoins et attentes des divers sous-groupes culturels à l'intérieur d'un même groupe ethnique. L'un des apports de l'anthropologie de la santé publique (Massé, 1991) serait, en ce sens, de sensibiliser les administrateurs de programmes à une approche « émiqne » des interventions (c'est-à-dire respectueuse du point de vue de la population-cible elle-même). L'évaluation devrait alors compter avec une analyse des « savoirs populaires » relatifs aux variables d'impact (par exemple la compétence parentale, la discipline physique ou les conditions propices au développement de l'enfant). Une fois conscientisés face à

l'existence de variantes importantes de ce savoir d'un sous-groupe social (et d'une sous-culture à une autre), les intervenants ne pourront que reconnaître l'obligation d'élaborer des programmes adaptés à chaque sous-groupe et sous-culture. Pas plus que de savoir populaire global, il n'existe de population-cible globale. L'un des mandats de la recherche évaluative consiste donc, non plus à mesurer la validité externe d'un programme (définie comme la possibilité de reproduire en bloc un programme dans des populations-cibles comparables) mais d'en mesurer la « validité écologique ». Cette dernière serait la capacité des diverses composantes d'un programme à s'adapter aux conditions environnementales (sociales, économiques, culturelles, etc.) caractéristiques d'un sous-groupe social défini. Étant donné qu'il n'existe jamais deux terrains d'expérimentation tout à fait semblables aux plans des clientèles, des intervenants, etc., l'évaluateur devra plaider en faveur d'une évaluation sensible aux variations dans les contextes d'expérimentation.

2) Le recours aux devis expérimentaux implique *une identification a priori des variables dépendantes (nature des résultats attendus)*. L'approche dominante dans l'évaluation des effets d'un programme consiste à identifier, à anticiper, avant même son implantation, la nature des résultats attendus du programme d'activités et le nombre de variables qui seront considérées. Bien sûr, l'ABC de toute bonne évaluation consiste à identifier le plus clairement possible les objectifs qu'elle vise, quel que soit le paradigme d'évaluation retenu. Or, il est évident que, même après avoir fait une bonne revue de littérature sur la problématique concernée, il restera toujours une large part d'arbitraire dans le choix des variables susceptibles de mesurer le degré d'atteinte des objectifs. Les chercheurs joueront alors aux apprentis-sorciers et retiendront en général celles qui apparaissent les plus « logiques ». La marge d'erreur s'accroît encore dans le cas de programmes novateurs qui n'ont jamais été expérimentés. On ne peut alors que spéculer sur la nature des effets attendus d'un programme novateur.

Les devis expérimentaux exigent que les retombées d'un programme soient réduites à un nombre limité de variables. Cette limitation est justifiée par la longueur maximale de l'entrevue avec le bénéficiaire (et du questionnaire), l'absence d'outils de mesure « validés », etc. Or il est clair que des programmes visant des problématiques aussi larges que la modification de comportements à risque pour la santé, la promotion de saines habitudes de vie, la prévention des abus et de la

négligence ou la promotion de la santé mentale devront miser sur une multiplicité d'approches, d'activités, de messages qui auront des effets sur un large spectre de facteurs. Les impératifs des méthodes quantitatives ne laissent aucune place à des outils de cueillette de données sensibles aux écarts entre variables d'impact prédéterminées et les composantes de la problématique effectivement influencées par le programme. L'impératif de conserver dans son intégralité les mêmes outils de mesure pour tous les bénéficiaires, du début à la fin de l'expérimentation, coupe court à tout ajustement en cours de route. C'est tout le problème de la mesure des effets non-attendus qui se trouve ici posé. Malgré son importance, la solution à ce problème est clairement hors de portée des devis expérimentaux.

Les auteurs du projet « De la Visite » ont clairement démontré que le suivi fait à long terme avec des familles à risque d'abus ou de négligence n'a eu aucun impact mesurable (avec les outils validés utilisés) sur les deux variables dépendantes choisies après une saine revue de littérature, soit l'adéquation du comportement maternel et l'isolement social. Pourtant, grâce au couplage d'un devis qualitatif poussé au devis expérimental, ils ont pu mesurer des effets de l'intervention non-professionnelle sur l'estime de soi de la mère, la qualité de la relation conjugale ou diverses conditions environnementales favorables à une saine relation parent-enfant. Les outils de mesure qualitatifs ont également mis en évidence des effets sur des composantes de l'isolement social inaccessibles à des grilles fermées (timing de l'aide reçue, revalorisation de la mère dans son réseau social, etc.) (Durand, Massé, Ouellet, 1988).

La divergence importante dans les effets mesurés par les deux devis (qualitatif et quantitatif) dans ce projet illustre clairement les limites que Trend (1979) a soulignées à propos de la « complémentarité » des approches. Ce dernier rappelle qu'à moins de réduire les méthodes qualitatives à une logique expérimentale, on risque le plus souvent de se retrouver face à des résultats carrément non convergents. La complémentarité devient conflit. L'unanimité dans les résultats provenant d'approches diverses devrait être considérée avec suspicion. L'engouement pour la « triangulation », depuis la proposition de Denzin (1978) comme approche par accouplement de plusieurs sources d'informations et méthodes, ne doit pas occulter le problème réel de réconciliation de résultats contradictoires.

3) *Les programmes d'intervention doivent avoir une intensité et une spécificité minimale pour avoir un effet visible* sur les facteurs de risque visés. Or une analyse rapide des programmes de prévention des abus ou négligence montre qu'ils n'ont généralement pas l'intensité nécessaire pour prétendre influencer de façon notable les variables dépendantes retenues. Certains vont jusqu'à considérer que, dans le cas des programmes sociaux, il est utopique d'espérer des résultats marqués à partir d'interventions faibles et dispersées. « De telles attentes témoignent autant de la naïveté des personnes chargées de définir les objectifs des programmes que de la position des chercheurs qui acceptent sans discussion la tâche d'évaluer de tels programmes » (Hudson, 1982, p. 84). La question ici est moins de s'assurer d'une puissance statistique des tests (rejet fondé de l'hypothèse nulle) par un devis adéquat que d'une puissance suffisante des interventions.

On veut, par exemple, réduire l'isolement social des nouveaux parents, améliorer la compétence parentale ou directement réduire l'incidence des abus et négligences en faisant intervenir une ressource pendant quelques heures réparties sur quelques semaines alors que, d'évidence, seule une intervention intensive échelonnée sur plusieurs mois peut laisser de l'espoir. Ainsi, Siegel (1980) rapporte les résultats d'un programme consistant 1) à favoriser le contact mère-enfant après l'accouchement à l'hôpital et 2) à réaliser neuf visites à domicile effectuées par des non-professionnelles durant les trois premiers mois du post-partum. Les effets mesurés concernent les comportements d'attachement mère-enfant, les attitudes d'acceptation-rejet et d'implication-détachement, mais aussi l'incidence de signalement d'épisode d'abus ou de négligence mettant en cause les parents visités et leur fréquentation des urgences hospitalières. On ne peut être surpris d'apprendre que le programme en question n'a eu aucun effet sur l'intensité de l'attachement maternel, l'incidence de signalement d'abus ou négligence chez les familles suivies ou l'utilisation des services de santé. Aucune surprise non plus dans le fait qu'ailleurs six sessions d'information de 20 minutes intégrées aux cours prénatals habituels plus une seule visite à domicile de 60 minutes n'ait permis de réduire l'isolement social, d'améliorer l'environnement familial, de favoriser le développement de l'enfant ou de réduire l'incidence de signalement d'abus ou négligence (Séguin et coll., 1986). Dans ce dernier cas, on ne peut qu'être frappé par la démesure des moyens utilisés. Des interventions aussi légères paraissent écrasées sous le poids de l'outillage de mesure : deux groupes témoins et un groupe expérimen-

tal et le recours à six mesures répétées (prénatal, deux semaines, un, trois, six et douze mois) pour une batterie de cinq outils de mesure standardisés. On peut douter ici que, face à des interventions aussi légères, même un appareillage qualitatif hyper-sensible ait réussi à mettre des effets en évidence. C'est donc moins le devis expérimental qui est en cause que l'intensité du programme lui-même. On est ici témoin d'une manifestation d'« aveuglement quantitatif » qui amène les évaluateurs à prioriser les moyens sur les fins. Les efforts mis dans la préparation d'un outillage sophistiqué de mesure semblent avoir masqué aux chercheurs eux-mêmes la nature et les limites de l'intervention qui prend inévitablement des allures de « prétexte » à évaluation.

4) La focalisation des devis expérimentaux sur la mesure des résultats repose sur un autre postulat soit *la répliquabilité du programme*. On peut en fait se questionner sur la pertinence d'évaluer les résultats d'un programme que l'on sait non-répliquable. L'expérimentation ne peut fournir les bases d'une généralisation des résultats sans la possibilité de reprendre le programme dans des conditions similaires (Crane, 1986). Même Cronbach, pourtant reconnu pour son penchant pour les méthodologies quantitatives, reconnaît que « la répliquabilité des programmes n'est qu'une figure de style » (cité par Crane, 1986, p. 61). Or d'évidence, dans le domaine psycho-social en général et dans le cadre des programmes de soutien au rôle parental en particulier, un même programme ne peut jamais être expérimenté deux fois dans un contexte parfaitement comparable. Tel est le cas de la plupart des projets pilotes. À titre d'illustration, référons encore une fois à l'évaluation du projet « De la Visite » (Durand, Massé, Ouellet, 1988). Le rendement mesuré était largement tributaire des attributs des intervenantes non-professionnelles impliquées, des dynamiques de quartiers particuliers, des qualités d'une coordonnatrice du suivi des familles, du défi posé à l'équipe par un projet pilote novateur, etc. Tous ces facteurs ne pourraient être reproduits sans modifications. De plus, les interventions évaluées se sont étendues sur 18 mois. Or, on peut se demander ce qu'il adviendrait de ces interventions après 10 ans d'implantation, de rodage et de pénétration dans les mœurs de la population-cible.

Il apparaît clairement que, dans le domaine du psycho-social en général, la « généralisabilité » des résultats à d'autres expérimentations d'un même programme, est un argument qui ne peut plus justifier le recours aux devis expérimentaux.

Il devient un prétexte justifiant a priori le recours à des méthodes lourdes non pertinentes.

5) Enfin, *soulignons qu'en intervention sociale, il est à peu près impossible de contrôler les caractéristiques des sujets à l'étude* et d'obtenir des groupes expérimentaux et témoins qui soient parfaitement comparables. On ne pourra appairer les cas et les témoins que sur un petit nombre de variables (généralement socio-démographiques, telles le sexe, l'âge, le niveau de scolarité, la parité, etc.), mais non sur la multitude de facteurs (toxicomanies, antécédents de négligence, tensions conjugales, etc.) qui, dans un cadre systémique, influent sur le problème visé (ex : violence).

De plus, on doit généralement compter avec des taux de refus importants chez les répondants échantillonnés tant du côté des témoins que des bénéficiaires du programme, de même qu'avec un taux de déperdition important dans le groupe expérimental. Ce problème de la « mortalité expérimentale » est encore plus critique dans le cas des seules interventions qui valent la peine d'être évaluées, soit celles qui offrent un suivi intensif s'étendant sur plusieurs mois. Enfin, les conditions requises pour le recours aux statistiques inférentielles ne sont généralement pas remplies (caractère aléatoire des échantillons, normalité dans la distribution des valeurs des variables analysées, etc.).

LA FÉTICHISATION DES OUTILS DE MESURE : LE MYTHE DES ÉCHELLES ET OUTILS VALIDÉS.

[Retour à la table des matières](#)

Les grands prêtres des méthodes quantitatives établissent la distinction entre les devis expérimentaux et quasi-expérimentaux par le fait que l'allocation aléatoire des cas entre groupes contrôle et expérimental est, dans le second cas, rendue impossible par divers facteurs (ex : Cook, Cook et Mark, 1982). La marge d'erreur dans les mesures d'évaluation est alors attribuable à l'impossibilité de contrôler des biais tels l'historicité, la maturation, la régression statistique, etc. Le discours

sur les limites de la recherche est centré sur la comparabilité des groupes à l'étude ou sur les accrocs au déroulement du programme (mortalité expérimentale, etc.).

La question des limites des outils de mesure est généralement escamotée. Pourtant, les outils de mesure « validés » sont au cœur du processus d'évaluation quantitative et on leur fait une confiance sans borne pour la mesure des effets du programme. On attribue en fait à ces outils des vertus semblables à celles qu'attribuent les sorciers à leurs fétiches. Trop souvent les évaluateurs se contentent de puiser à même une boîte à outils ceux qui traditionnellement ont été utilisés pour des fins semblables et ce, sans se questionner plus à fond sur leur adéquation. Ce comportement rappelle celui des sorciers qui déduisent l'efficacité du fétiche aux vertus magiques de la fréquence de son utilisation avec « succès ». On est en fait confronté au cercle vicieux « plus le monde en mange, plus on les aime ; plus on les aime, plus le monde en mange ».

Deux exemples frappants de tels outils méritent d'être soulignés dans le domaine de l'évaluation des programmes de soutien au rôle parental : le « Arizona Social Support Interview Schedule » (Barrera, 1980) pour la mesure de l'isolement social et le « Home Observation for the Measurement of the Environment » (Caldwell et Bradley, 1982) pour la mesure de l'adéquation de l'environnement physique, social et parental de l'enfant. Ces deux outils ont été utilisés dans la très grande majorité des projets de recherche évaluative (ou épidémiologiques) portant sur les abus et négligences au Québec depuis 10 ans (ex : Durand, Massé, Ouellet, 1988 ; Larson, 1980 ; Séguin, 1986 ; Lepage et Guay, 1990 ; Laurendeau, M.C. et coll., 1990 ; D.S.C. Laval, 1986 ; etc.). On les retrouve de même dans sept des neuf recherches évaluatives majeures déjà identifiées. De plus, malgré les réserves importantes formulées sur la validité de ces outils par les chercheurs qui y ont déjà eu recours, plusieurs recherches maintiennent l'idée de les intégrer à des protocoles de recherches en élaboration. Deux raisons majeures sont alors invoquées. La première tient au fait que les construits mesurés par ces outils se rapprochent le plus des construits sur lesquels on croit que l'intervention aura des effets. La seconde, plus terre-à-terre, est qu'il est pratiquement impossible d'obtenir le financement d'une recherche qui ne repose pas sur des outils validés et fréquemment utilisés par d'autres chercheurs. Sur un plan épistémologique, la première est la plus grave. On subordonne la nature postulée des effets d'un programme à l'existence d'outils de mesure. On ajuste les effets attendus aux problèmes de mesure.

Bref, on subordonne les objectifs du programme aux méthodes. Or, dans la plupart des cas, la valeur de ces outils et leur inadéquation face aux facteurs à mesurer n'en justifient pas l'utilisation. L'élaboration de nouveaux outils adaptés à un programme spécifique devient, la plupart du temps, une avenue incontournable.

Les défis posés par la mesure en évaluation

[Retour à la table des matières](#)

Forts de ces constats, nous pouvons identifier trois défis qui attendent les évaluateurs et méritent de faire l'objet de débats plus approfondis. Il en va de l'évolution de la recherche évaluative dans le sens d'une meilleure adaptation des moyens aux fins d'évaluation. D'abord, il faut dénoncer les prétentions d'universalité des outils de mesure. Puis il faut redonner une crédibilité aux perceptions des individus, intervenants ou bénéficiaires, face aux effets d'un programme. Ensuite, il faut plaider pour l'élaboration d'outils de mesure adaptés à des objectifs d'évaluation et non à des objectifs épidémiologiques.

1) Lutter contre les prétentions d'universalité des outils de mesure.

Un simple regard critique sur les évaluations de programme dans le domaine social au Québec montre à l'évidence que les mêmes outils de mesure sont utilisés dans des contextes d'évaluation très divers. Les outils vedettes de mesure de l'isolement social, de la compétence parentale, de la dépression, etc., se voient attribuer des aptitudes universelles et ce, en deux sens différents :

A) D'abord malgré des différences importantes dans la nature et l'intensité des interventions proposées par les divers programmes de prévention, on retiendra les mêmes composantes de l'isolement social ou de la compétence parentale comme indicateurs des effets, soit celles mesurées par les outils populaires. Ce qui est en cause ici c'est la prétention d'universalité des variables dépendantes retenues pour toutes les sauces. On mesure des effets universels attribués à des programmes différents. Les outils de mesure doivent avoir la souplesse nécessaire pour couvrir

les diverses formes des phénomènes mesurés. Or ce n'est pas possible de mesurer exactement la compétence parentale, l'amour, la qualité de la communication parent-enfant... On se rabat alors sur des variables dépendantes mesurables (poids de l'enfant, niveau de développement psycho-moteur, cognitif, etc.) qui ne traduisent que les composantes les plus manifestes des résultats.

B) Ensuite, on ne remet pas suffisamment en question la validité de construit et de contenu de ces outils. Faisons pour l'instant l'hypothèse que leur validité métrique est adéquate. Il est beaucoup moins évident que ces outils sont aptes à saisir toute l'essence des « construits » théoriques (ex : compétence, isolement) que l'on prétend mesurer et qu'ils soient, de même, en mesure de cerner toutes les facettes ou dimensions de ces concepts complexes. En fait, ces limites sont de deux ordres : même les meilleurs outils ne mesureront jamais qu'un nombre limité de dimensions du « processus environnemental » ; ensuite, et ce point est fondamental, le nombre, la nature et l'importance relative des dimensions de concepts tels la compétence ou l'isolement variera profondément d'un groupe ethnique, d'une classe sociale et d'une génération à l'autre. Soyons clairs ici. Notre propos n'est pas de faire la promotion d'outils idéaux qui n'existeront jamais. Il s'agit plutôt de remettre en question le recours aveugle et sans recul critique à ces outils qui caractérise trop souvent les évaluations de programmes psycho-sociaux. Il ne s'agit pas non plus de dénoncer une certaine forme de « colonialisme psychométrique » qui caractérise la recherche psycho-sociale au Québec, ni de mousser un nationalisme futile qui mettra l'accent sur des outils québécois aux dépens d'outils américains. Nous faisons un plaidoyer en faveur d'un souci constant d'adapter les outils de mesure aux formes mouvantes que prennent certains problèmes de santé d'un groupe ethnique, d'une classe sociale ou d'une génération à une autre. Bref il s'agit d'une dénonciation sous tous azimuts des prétentions d'universalité et des exportations abusives des outils, des concepts et des modèles.

L'anthropologie épidémiologique (ex : Janes, Stall et Gifford, 1986) n'a pu que constater les limites inhérentes à la définition d'une foule de problèmes de santé ou de facteurs de risque, en particulier dans le domaine psycho-social. Sa proposition d'opérer un virage à 180 degrés par rapport à la démarche épidémiologique est d'un fort intérêt pour la recherche évaluative. Plutôt que de partir d'une définition fixe, à prétention universelle, d'un problème de santé, on devrait privilégier une consultation directe de la population concernée dans le but de faire émerger la

ou les définitions qu'elle donne à ce problème. Seule une compréhension de la définition populaire de ces phénomènes et des « savoirs populaires » permettra d'ajuster le contenu des programmes de prévention aux réalités vécues et interprétées par les populations-cibles. Un rappel des postulats de base du relativisme culturel en évaluation ne pourrait être que bénéfique pour resituer la validité externe des outils de mesure dans des limites plus justes.

*2) Lutter contre l'occultation de la sensibilité
dans la mesure des effets d'un programme*

Ce qui constitue l'essence du débat ici, c'est la crédibilité que l'on devrait accorder aux opinions, jugements et appréciations des intervenants et des bénéficiaires quant aux effets des services investis. C'est la considération des données qualitatives comme valides pour l'évaluation des résultats. Nous considérons impensable d'envisager l'évaluation d'un programme sans prendre en considération les résultats ressentis par les premiers concernés. Les familles bénéficiaires des interventions ne peuvent être confinées au rôle d'acteurs passifs qui ne présentent d'autre intérêt que de justifier l'existence du programme. L'évaluateur doit redonner le droit de parole aux participants aux programmes et rendre crédible leur discours. À l'image des anthropologues, il doit servir d'intermédiaire, de *culture broker* entre la population et les gestionnaires des programmes. Il doit défendre la place d'une approche « émique » des effets d'un programme.

Un enjeu majeur ici en recherche évaluative est la lutte contre l'occultation de la sensibilité dans la mesure des effets. Il est difficilement acceptable que des programmes soient évalués sans qu'aucun des dispensateurs et bénéficiaires de services ne soient questionnés sur leur appréciation des résultats et retombées du programme. On constate malgré tout fréquemment que certains programmes étalés sur plusieurs mois sont évalués sans que le chercheur n'ait eu plus de quelques heures de discussion avec les premiers concernés et sans qu'il n'ait été présent sur le terrain d'expérimentation. L'un des facteurs « d'insensibilisation » des chercheurs face aux effets perçus de l'intervention est l'absence de contact direct de ces derniers avec les répondants. Parce qu'ils opèrent la cueillette des données à l'aide d'outils dont on ne requestionne jamais la pertinence en cours de projet et

parce qu'ils délèguent les entrevues à des interviewers, les responsables des évaluations de type expérimental n'ont souvent aucun contact direct avec les bénéficiaires et intervenants. Le fait de ne vivre ces contacts qu'à travers l'écran cathodique d'un micro-ordinateur et de ne discuter avec la population que par le biais d'indicateurs quantifiés rend le chercheur souvent aussi insensible à la complexité des variables étudiées que les outils de mesure qu'il utilise.

Il devrait pourtant devenir non pas un traducteur servile des opinions des bénéficiaires, mais un décodeur du discours populaire et un médiateur entre les administrateurs de programmes et la population-cible. Il est clair que l'on ne pourra jamais remplacer les données « solides » par les opinions populaires. Toutefois la propension actuelle à n'attribuer de crédibilité qu'aux données recueillies par des outils standardisés peu sensibles entraîne des lacunes telles en évaluation que l'adjonction de mesures plus qualitatives s'impose. La promotion d'une plus grande sensibilité dans les mesures en évaluation passe par l'assouplissement d'un paradigme du retrait sécurisant du chercheur en faveur d'un paradigme de l'engagement, la promotion d'un paradigme du terrain aux dépens d'un paradigme de la recherche de bureau.

*3) Plaider en faveur d'outils de mesure spécifiques
à des fins d'évaluation et non à des fins de mesure épidémiologique.*

Dans un numéro récent de la revue *Evaluation and Program Planning*, Hans et Bernstein (1991) font une réflexion intéressante sur l'adaptation des outils de mesure pour l'évaluation des programmes destinés aux mères adolescentes. Quoique tel ne soit pas leur propos initial, leurs réflexions conduisent à la constatation que les outils existants de mesure de l'adéquation de la relation parent-enfant furent élaborés pour des fins de recherche dans un cadre clinique, mais sont mal adaptés aux exigences de la recherche évaluative. Ces outils ne seraient pas adéquats pour la mesure des effets de l'intervention. Ils ne seraient aucunement conçus pour répondre aux besoins des autres composantes de l'évaluation que sont la planification et la gestion des processus et de la structure de l'intervention.

Dans le cadre des programmes de prévention destinés aux mères adolescentes et leurs enfants (Ounce of Prevention Fund Programs), ils ont mis deux ans à éla-

borer un guide d'observation des interactions parent-enfant. Cet outil permet de comprendre les rapports entre parent et enfant et d'identifier les forces et problèmes qui devront guider l'intervention. De leur expérience, ils tirent neuf leçons. J'extrapolerai ici pour en faire neuf critères de base qui devraient guider l'élaboration des outils de mesure dans tout projet d'évaluation d'intervention psychosociale. Ces critères nient totalement la possibilité d'existence d'outils universels utilisables dans des contextes d'évaluation différents.

- 1) Le modèle théorique sous-jacent aux choix des items ou facteurs exprimant l'adéquation des relations parent-enfant devra être conforme aux consensus établis parmi les chercheurs. Mais il devra aussi être partagé par le personnel impliqué dans la gestion et la production des services. Les intervenants non-professionnels requièrent généralement une formation minimale pour être fonctionnels. Toutefois, si la théorie ne colle pas à leurs observations et expériences quotidiennes sur le terrain, cette formation risque de ne pas être intégrée dans l'intervention auprès des familles. Ils ont donc décidé de privilégier les facteurs suivants comme marqueurs d'un risque dans la relation parent-enfant : l'adéquation de la réponse des parents aux besoins de l'enfant, leur sensibilité aux intérêts de l'enfant, l'affection pour l'enfant et l'aide donnée à l'enfant pour qu'il apprenne.
- 2) L'outil doit être fondé sur l'observation de comportements spécifiques de la dyade sur le terrain, dans le contexte concret où ils se déroulent et prennent sens.
- 3) Dans la mesure où les programmes de soutien aux familles devraient mettre l'emphasis sur les forces et non les faiblesses des parents, l'outil de mesure doit couvrir les comportements positifs et non les seuls marqueurs du risque.
- 4) L'outil doit minimiser les biais culturels. Pour ce faire, il ne devrait pas mesurer les éléments susceptibles de prendre une signification différente d'un groupe ethnique à un autre.
- 5) L'outil doit avoir suffisamment de flexibilité pour pouvoir être utilisé avec fiabilité dans plusieurs contextes d'observation.

- 6) L'outil doit être valide pour couvrir l'étendue d'âge de l'enfant du début à la fin du programme d'intervention. Si le programme s'adresse aux enfants allant de deux semaines à deux ans, comme c'est le cas pour plusieurs programmes de prévention de la violence ou soutien au rôle parental, l'outil devra n'inclure que des items pertinents tout au long de ces deux ans d'observation.
- 7) L'outil doit fournir des informations tant sur le parent que sur l'enfant et l'interaction dans la dyade.
- 8) L'outil doit être facile d'utilisation et maniable. Il doit pouvoir être administré par des non-professionnels sans scolarité poussée, moyennant une légère formation.
- 9) Enfin, l'outil doit répondre aux critères habituels de fiabilité et de validité.

Nous ne pouvons que souscrire à ces constatations. Tous les outils de mesure de la relation parent-enfant, et de façon plus large, tous les outils de mesure utilisés dans les évaluations de programme de prévention devraient posséder ces qualités. Là où nous pouvons avoir des doutes, c'est lorsque Hans et Bernstein prétendent avoir réussi à faire de leur Parent-Infant Observation Guide (PIOG) cet outil miracle. En fait, les auteurs énumèrent tout un ensemble de critères qui démontrent l'impossibilité d'élaborer des outils à valeur universelle. Ils ont peut être réussi à élaborer un instrument valide dans les cadres de leur projet. Mais la liste des critères énumérés démontre qu'il faudrait refaire le travail dans tout autre contexte. Poser comme conditions que l'outil soit adapté :

- a) à l'âge de l'enfant,
- b) au contexte dans lequel se déroule l'interaction mère-enfant,
- c) aux croyances, valeurs et idéologies prévalentes dans un groupe ethnique (et ajoutons une sous-culture intra-ethnique) particulière,
- d) et qu'il tienne compte de la diversité des théories sur le développement de l'enfant (et nous ajouterions de la conception populaire de la compétence parentale)...

... c'est reconnaître l'impossibilité d'élaborer un instrument qui puisse être valide à l'extérieur d'un cadre bien particulier. C'est aussi et surtout reconnaître l'obligation d'élaborer autant d'outils qu'il existe de contextes d'évaluation.

CONCLUSION

[Retour à la table des matières](#)

Les premières générations de recherche évaluative dans le domaine psychosocial ont été caractérisées par une sorte d'invasion par les méthodes scientifiques. La recherche évaluative est devenue simplement un nouveau champ d'application des méthodes, procédures et outils propres à la recherche scientifique. Il s'agissait d'abord de se donner les moyens de mesurer scientifiquement l'habileté d'un programme à atteindre des objectifs fixés à l'avance. Pré-test, groupes contrôles et outils de mesure standardisés ont trouvé un nouveau débouché. La seule raison d'être de la recherche évaluative était la mesure des résultats anticipés. Or, il est progressivement devenu évident que les personnes concernées par le programme, soit les gestionnaires, les intervenants et les bénéficiaires des services ne sont pas préoccupés au premier abord par la seule mesure des résultats. On souhaite mieux connaître les besoins de la clientèle, l'adéquation des ressources investies et des services produits, les difficultés de fonctionnement, la barrière à l'implantation des services, la réceptivité des clients, etc. Or, comme le notaient récemment Hurlé et Nadeau (1987) dans le domaine de la psycho-éducation, la recherche évaluative n'est pas toujours en mesure de fournir ce type d'information étant donné ses objectifs et ses orientations méthodologiques.

Les finalités de l'évaluation devraient être constructives et non de poser un jugement « guillotine » sur l'existence du programme.

Trop d'évaluations arrivent à conclure que le projet n'a pas d'effet sans que l'on puisse dire quoi que ce soit d'intelligent sur ce qui a ou n'a pas marché. Plusieurs programmes présentent des résultats surprenants (impacts positifs sur les mauvaises variables dépendantes, impacts négatifs sur ce que le programme devait améliorer) et ce, sans que l'on puisse conclure à leur inutilité. Les organismes subventionnaires québécois ont dépensé plusieurs millions de dollars pour finan-

cer des évaluations dans les dix dernières années, projets qui n'auraient jamais dû être financés dans la forme qu'ils prenaient, c'est-à-dire avec des devis exclusivement expérimentaux. Dans la très grande majorité des cas, aucune évaluation de processus ou de structure, ou formative, ne fut faite. Les conclusions en arrivent à dire des banalités sur les variations dans les scores aux variables dépendantes, variables qui ne couvrent qu'une partie des effets attendus. Rien n'est dit sur les effets non attendus. Aucune information n'a pu être produite qui permet de réorienter, de bonifier, d'ajuster les interventions. La plupart du temps, les chercheurs sont conscients de ces lacunes. Mais après coup.

Nous croyons que la recherche est mûre pour dépasser le stade des *mea culpa*. Les chercheurs en évaluation doivent être habilités à choisir et à adapter leurs méthodes à la complexité et à l'unicité de l'intervention sociale (Lecomte, 1982a). Ils doivent Prendre conscience du fait que leurs choix méthodologiques sont souvent fondés sur des présupposés méthodologiques. Les limites des devis expérimentaux et des outils de mesure, les apports des méthodes qualitatives, l'importance des évaluations de processus et de structure, l'importance d'une évaluation constructive sont aujourd'hui tellement bien reconnus et démontrés qu'il n'y a plus d'excuses valables pour reproduire les mêmes erreurs. Les organismes subventionnaires devraient les premiers ouvrir leurs oeillères. Ils devraient arrêter de jouer aux grands prêtres du fétichisme des méthodes « dures » et de chercher à convertir les fidèles-chercheurs au culte du quantitatif sous peine d'excommunication. Au lieu de simplement questionner l'adéquation des méthodes à l'orthodoxie quantitative, ils devraient plutôt se demander quels seront les apports de cette recherche évaluative, quelles en seront les retombées, non pas sur la réputation du chercheur et du bailleur de fonds, mais sur l'intervention évaluée et sur les personnes qui en sont bénéficiaires.

RÉFÉRENCES BIBLIOGRAPHIQUES

[Retour à la table des matières](#)

ASHEM, B. et S. KURTZ (1987) *Parents Helping Parents Project : Evaluation Report*, Department of Public Health, City of Toronto.

BARRERA, M. (1980) « A method for the assesment of social support networks in community survey research », *Connections*, 3 (3) :8-13.

BARTH, R.P., HACKING, S. et J.R. ASH (1988) « Preventing Child Abuse : an experimental evaluation of the child parent enrichment project », *Journal of Primary Prevention*, 8 (4) : 201-217.

BERTHELET, D. (1987) *Les méthodes de recherche appliquées à l'Évaluation d'impact*, Rapport de recherche, Groupe de recherche interdisciplinaire en santé, Université de Montréal.

BOUCHARD, C. (1989) « Maladaptation parentale et interventions préventives », in : L. Rocheleau, L. Séguin, M. Cournoyer et C. Chamberland, *Vivre avec un nourrisson, La périnatalité au Québec*, Ministère de la santé et des services sociaux, p. 135-162.

BRADLEY, R.H. (1981) The HOME inventory : A review of findings from the Little Rock Longitudinal study. *Infant Mental Health Journal*, 2 (3) : 198-205.

CALDWELL, B. et R. BRADLEY (1982) « *Home Observation for Measurement of the environment* ». New York, Dorsey. »

CHAMPAGNE, F., CONTANDRIOPOULOS, A.P. et R. PINEAULT (1985) *Health care program evaluation : a proposed framework*. Groupe de recherche interdisciplinaire en santé, Université de Montréal.

CHEN, H.T. et P.H. Rossi (1983) « Evaluating with sense. The theory-driven approach ». *Evaluation review*, 7 (3) : 283-302.

CHRISMAN, N. (1977) « The health seeking process : An approach to the natural history of illness », *Culture, Medecine and Psychiatry*, 1 (4).

COOK, T.D., LOMAX COOK, F. et M.M. MARK (1982) « Modèles expérimentaux et quasi expérimentaux en recherche évaluative » in : Roland Lecomte et Léonard Rutman (éds) : *Introduction aux méthodes de recherche évaluative*, Université de Carleton, Ottawa.

CRANE, J.A. (1986) « A model for single system evaluation of treatments », *Canadian Journal of Program Evaluation*, 1 (1) : 59-74.

DARO, D. (1988) *Confronting child Abuse. Research for effective program design*. N.Y, The Free Press.

DENZIN, N. (1978) *The research act*, N.Y. : McGraw-Hill.

Département de Santé Communautaire de Laval. (1986) *La prévention de la négligence des enfants de 0-5 ans en milieux à risque*, Protocole de recherche du projet « La Parentelle ».

DURAND, D., MASSÉ, R. et F. OUELLET (1988) *Intervenantes non professionnelles et soutien au rôle parental : « De la Visite » : Expérimentation et évaluation d'une intervention novatrice en prévention de l'enfance maltraitée*, D.S.C. Verdun, Lakeshore, Sacré-Cœur.

HANS, S.L. et V.J. BERNSTEIN (1991) « Adolescent parenting programs : Assessing parent-infant interaction », *Evaluation and Program Planning*, 14 :87-95.

HELPER, R.E. (1982) « A review of the litterature on the prevention of child abuse and neglect », *Child Abuse and Neglect*, 6 : 251-261.

GRAY, J.D., CUTLER, A., DEAN, G.D., et C.H. KEMPE (1979) « Prediction and prevention of Child Abuse and Neglect », *Journal of Social Issues*, 35 (2) : 127-139.

HUDSON, J. (1982) « Les problèmes de mesure en recherche évaluative », in : Roland Lecomte et Leonard Rutman (éds) : *Introduction aux méthodes de recherche évaluative*, Université de Carleton, Ottawa.

HURTEAU, M. et M.A. NADEAU (1987) « Évaluation de programme, recherche évaluative et recherche », *Revue Canadienne de psycho-éducation*, 16 (2) : 91-96.

IGUN, U.A. (1979) « Stages in health-seeking : A descriptive model ». *Social Sciences and Medecine*, 13A : 445-456.

JANES, C.R., STALL, R. et S.M. GIFFORD (1986) *Anthropology and Epidemiology*, Reidel Publishing Co.

LAURENDEAU, M.C. et coll. (1990) *Parents Magazine : un véhicule d'information sur l'éducation des jeunes enfants et les ressources de soutien aux familles à l'usage des parents et des CLSC*, Apprentissage et Socialisation, 13 (3) : 175-186.

LECOMTE, R. (1982) « Les paradigmes méthodologiques de la recherche évaluative : leurs fondements et leurs répercussions », in : Roland Lecomte et Léonard Rutman (éds) : *Introduction aux méthodes de recherche évaluative*, Université de Carleton, Ottawa.

MASSÉ, R. (1988) *De la Visite : expérimentation et évaluation d'une intervention novatrice en prévention de l'enfance maltraitée*, Vol 4 : « Analyse des composantes quantitatives et qualitatives des ressources investies et des services produits », D.S.C. Verdun, Sacré-Cœur et Lakeshore.

MASSÉ, R. (1990) « Évaluation critique de la recherche sur l'étiologie de la violence envers les enfants », *Santé Mentale au Québec*, XV (2) : 107-127. [Texte disponible dans [Les Classiques des sciences sociales](#). JMT.]

MASSÉ, R. (1991) *L'enfance négligée et abusée physiquement : facteurs de risque et prévention*. Document de recherche préparé pour le « Groupe de travail sur les problèmes de l'enfance et la jeunesse », Ministère de la Santé et des Services Sociaux, D.S.C. Verdun.

OLDS, D.L. (1982) « The prenatal/early infancy project. An ecological approach to prevention of developmental disabilities », in : Jay Belsky (éd.), *The Beginning*, N.Y., Columbia University Press.

OLDS, D.L., HENDERSON, C.R., CHAMBERLAIN, R. et R. TATELBAUM (1986) « Preventing child Abuse and Neglect : A randomized trial of nurse home visitation », *Pediatrics*, 78 (1) : 65-77.

LARSON, C.P. (1980) « Efficacy of prenatal and postpartum home visits on child health and development », *Pediatrics*, 66 (2) : 191-197.

LECOMTE, R. (1982a) « Les paradigmes méthodologiques de la recherche évaluative : leurs fondements et leurs répercussions », in : Roland Lecomte et Léonard Rutman (éds) : *Introduction aux méthodes de recherche évaluative*, p. 1-22, Université de Carleton, Ottawa, 1982a.

LECOMTE, R. (1982b) « Les apports de l'évaluation qualitative et critique en recherche évaluative », in Roland Lecomte et Leonard Rutman *Introduction aux méthodes de recherche évaluative*, « Université de Carleton, Ottawa.

LYONS-RUTH, K., CONNELL, D.B. et H.U. GRUNEBAUM (1990) « Infants at social risk : Maternal depression and family support services as mediators of infant development and security of attachment », *Child Development*, 61 : 85-98.

PATTON, M.Q. (1980) *Qualitative Evaluation Methods*. Sage Publications, Beverly Hills.

PRESSLE GOETZ, J. et M.D. LECOMPTE (1984) *Ethnography and Qualitative Design in Educational Research*, Academic Press.

ROCHELEAU, B. (1982) « Qualitative methods evaluation research : Major applications for mental health program evaluation », in : Gerald J. Stahler., et William R. Tash (éds) : *Innovative approaches to Mental Health Evaluation*, Academic Press, N.Y.

Rossi, P.H., FREEMAN, H.E. et S.R. WRIGHT (1979) *Evaluation : A systematic approach*, Sage, CA, Beverly Hills.

SÉGUIN, L., FERLAND, F., LAMBERT, J. et D. OUELLET (1986) *Évaluation de l'efficacité d'interventions éducatives et de support auprès des parents à la période périnatale*, Département de Médecine Sociale et Préventive, Université de Montréal.

SIEGEL, E., BAUMAN, K.E., SCHAEFER, E.S., SAUNDERS, M.M. et D.D. INGRAM (1980) « Hospital and home support during infancy : impact on maternal attachment, child abuse and neglect and health care utilization », *Pediatrics*, 66 : 183-190.

TREND, M.G. (1979) « On the reconciliation of qualitative and quantitative analyses », in : Thomas D. Cook et Charles S. Reichard, (éds), *Qualitative and quantitative methods in evaluation research*, Sage, Beverly Hills.

WILLMS, D., BEST, J.A., TAYLOR, D.W., J.R. GILBERT *et al.* (1990) « A systematic approach for using qualitative methods in primary prevention research », *Medical Anthropology Quarterly*, 4 (4), 391-409.

Fin du texte