

Sous la direction de  
**MARC ALAIN et DANNY DESSUREAULT**

# Élaborer et évaluer les programmes d'intervention psychosociale



Presses  
de l'Université  
du Québec

*Facebook : La culture ne s'hérite pas elle se conquiert*



**Élaborer  
et évaluer  
les programmes  
d'intervention  
psychosociale**

PRESSES DE L'UNIVERSITÉ DU QUÉBEC  
Le Delta I, 2875, boulevard Laurier, bureau 450  
Québec (Québec) G1V 2M2  
Téléphone: 418-657-4399 • Télécopieur: 418-657-2096  
Courriel: puq@puq.ca • Internet: www.puq.ca

Diffusion / Distribution :

**CANADA et autres pays**

PROLOGUE INC.  
1650, boulevard Lionel-Bertrand  
Boisbriand (Québec) J7H 1N7  
Téléphone: 450-434-0306 / 1 800 363-2864

**FRANCE**

AFPU-DIFFUSION  
SODIS

**BELGIQUE**

PATRIMOINE SPRL  
168, rue du Noyer  
1030 Bruxelles  
Belgique

**SUISSE**

SERVIDIS SA  
Chemin des Chalets  
1279 Chavannes-de-Bogis  
Suisse



La *Loi sur le droit d'auteur* interdit la reproduction des œuvres sans autorisation des titulaires de droits. Or, la photocopie non autorisée – le « photocopillage » – s'est généralisée, provoquant une baisse des ventes de livres et compromettant la rédaction et la production de nouveaux ouvrages par des professionnels. L'objet du logo apparaissant ci-contre est d'alerter le lecteur sur la menace que représente pour l'avenir de l'écrit le développement massif du « photocopillage ».

Sous la direction de  
**MARC ALAIN et DANNY DESSUREAULT**

# **Élaborer et évaluer les programmes d'intervention psychosociale**

**2009**



**Presses de l'Université du Québec**

Le Delta I, 2875, boul. Laurier, bur. 450  
Québec (Québec) Canada G1V 2M2

*Catalogage avant publication de Bibliothèque et Archives nationales du Québec  
et Bibliothèque et Archives Canada*

Vedette principale au titre :

Élaborer et évaluer les programmes d'intervention psychosociale

Comprend des réf. bibliogr.

ISBN 978-2-7605-2346-3

1. Service social – Évaluation. 2. Évaluation de programme. 3. Service social –  
Planification. 4. Analyse des besoins. I. Alain, Marc. II. Dessureault, Danny, 1957- .

HV41.E42 2009 361.0068 C2008-942434-4

Nous reconnaissons l'aide financière du gouvernement du Canada  
par l'entremise du Programme d'aide au développement de l'industrie  
de l'édition (PADIE) pour nos activités d'édition.

La publication de cet ouvrage a été rendue possible grâce à l'aide financière  
de la Société de développement des entreprises culturelles (SODEC).

**Mise en pages:** INFOSCAN COLLETTE-QUÉBEC

### **Couverture**

Conception et réalisation: RICHARD HODGSON

Illustration: *Timbrés et affranchis*, 2005, techniques mixtes sur toile

MIREILLE BOURQUE est une artiste audacieuse qui expérimente, de toile en toile, les médiums les plus diversifiés. Pour elle, le processus de création est semblable à une méditation. Elle est attentive aux émotions qui émergent en regard de ce qui apparaît sur la toile. Il s'agit, la plupart du temps, de personnages car ce thème favorise le rapprochement et lui permet d'instaurer une forme de dialogue. Aussi, depuis quelque temps, des visages apparaissent en premier plan et brisent l'anonymat des personnages: ils ont des choses à dire! Au-delà de sa démarche artistique, l'implication de Mireille Bourque au sein du programme *Vincent et moi* témoigne de sa volonté de jouer un rôle actif dans la démystification de la maladie mentale.



*Photo de l'illustration:* SIMON LECOMTE

1 2 3 4 5 6 7 8 9 PUQ 2009 9 8 7 6 5 4 3 2 1

*Tous droits de reproduction, de traduction et d'adaptation réservés*

© 2009 Presses de l'Université du Québec

Dépôt légal – 1<sup>er</sup> trimestre 2009

Bibliothèque nationale du Québec / Bibliothèque nationale du Canada

Imprimé au Canada

# TABLE DES MATIÈRES

---

|  |    |
|--|----|
| <b>Introduction et remerciements</b> . . . . .   | 1  |
| <i>Marc Alain et Danny Dessureault</i>   |    |
| <br>Partie I   |    |
| <b>Histoire, théories et modèles d'évaluation de programme.</b> . . . . .  | 7  |
| <br>Chapitre I   |    |
| <b>Des systèmes de classification des modèles d'évaluation de programmes<br/>d'intervention psychosociale à une proposition de modèle intégrateur.</b> |    |
| <b>Les trois phases du cube AVANT-PENDANT-APRÈS</b> . . . . .  | 9  |
| <i>Marc Alain</i>  |    |
| 1. Quelques classifications des modèles d'évaluation de programme . . . . .  | 12 |
| 1.1. Une perspective plus historique, celle de Rossi, Lipsey et Freeman (2004) . . . . .   | 13 |
| 1.2. Les quatre modèles d'évaluation de la perspective de Stufflebeam . . . . .  | 13 |
| 1.3. L'arbre des théories de Christie . . . . .  | 14 |
| 1.4. Les quatre générations de l'évaluation selon Guba et Lincoln . . . . .  | 16 |
| 1.5. L'évaluation guidée par la théorie du programme:<br>la contribution de Chen à la proposition d'un modèle intégrateur . . . . .                    | 17 |
| 2. La proposition d'un modèle intégrateur: le modèle du cube . . . . .   | 18 |
| 2.1. La phase AVANT . . . . .  | 20 |
| 2.2. La phase PENDANT . . . . .  | 21 |
| 2.3. La phase APRÈS . . . . .  | 24 |
| En guise de conclusion: comment les chapitres suivants s'intègrent au modèle . . . . .   | 25 |
| Bibliographie. . . . .   | 27 |

|  |           |
|--|-----------|
| Chapitre 2   |           |
| <b>Le développement de l'évaluation de programme . . . . .</b>   | <b>29</b> |
| <i>Martin Goyette</i>  |           |
| 1. Les sources de l'évaluation . . . . .   | 31        |
| 2. Les balbutiements de l'évaluation: le testing et la mesure . . . . .  | 32        |
| 3. La description . . . . .  | 33        |
| 4. Le jugement . . . . .   | 33        |
| 5. L'évaluation négociée . . . . .   | 36        |
| 6. Une complémentarité des paradigmes et des méthodologies . . . . .   | 40        |
| Bibliographie. . . . .   | 41        |
| Partie 2   |           |
| <b>De l'évaluation des besoins à la conceptualisation<br/>d'un programme d'intervention, la phase AVANT . . . . .</b>  | <b>43</b> |
| Chapitre 3   |           |
| <b>L'état subjectif des lieux: les zones de conflit et les zones d'entente<br/>dans la perception des besoins à combler et de la volonté<br/>des acteurs mobilisables à s'impliquer. . . . .</b> | <b>45</b> |
| <i>Sylvie Hamel, Marie-Marthe Cousineau et Martine Vézina</i>  |           |
| 1. Bref état de la situation . . . . .   | 46        |
| 1.1. Recommandations des experts pour la prévention<br>du phénomène des gangs . . . . .  | 48        |
| 1.2. Qu'est-ce qu'une innovation sociale? . . . . .  | 49        |
| 2. Le pouvoir aux communautés: une approche participative . . . . .  | 50        |
| 3. Une méthodologie principalement qualitative . . . . .   | 53        |
| 3.1. Plusieurs sources de données . . . . .  | 54        |
| 3.2. Une analyse s'appuyant sur la méthode de cas . . . . .  | 55        |
| 3.3. Deux modèles de départ . . . . .  | 56        |
| 4. Une évaluation sous plusieurs angles . . . . .  | 58        |
| 4.1. L'évaluation des besoins . . . . .  | 59        |
| 4.2. L'évaluation des processus . . . . .  | 61        |
| 4.3. L'évaluation de la qualité des plans d'action. . . . .  | 65        |
| Conclusion . . . . .   | 68        |
| Bibliographie . . . . .  | 69        |
| Chapitre 4   |           |
| <b>L'évaluation des besoins: un outil pour la planification . . . . .</b>  | <b>73</b> |
| <i>Line Massé</i>  |           |
| 1. Les objectifs de l'évaluation des besoins . . . . .   | 74        |
| 2. Les principaux concepts liés aux besoins. . . . .   | 75        |
| 3. Les approches d'analyse des besoins . . . . .   | 76        |
| 4. Les étapes de l'analyse de besoins . . . . .  | 77        |
| 4.1. Déterminer les fins de l'étude de besoins et les utilisateurs potentiels . . . . .  | 78        |
| 4.2. Définir la population à l'étude . . . . .   | 78        |
| 4.3. Déterminer les informations particulières à obtenir . . . . .   | 78        |

|  |            |
|--|------------|
| 4.4. Élaborer un protocole d'évaluation . . . . .  | 79         |
| 4.5. Recueillir l'information, analyser et interpréter les données . . . . .   | 80         |
| 4.6. Préparer le rapport . . . . .   | 81         |
| 4.7. Rédiger le rapport final et diffuser plus largement les résultats de l'étude . . . . .                                  | 81         |
| 5. Les techniques de collecte de données . . . . .   | 84         |
| 5.1. La revue des écrits . . . . .   | 84         |
| 5.2. L'approche par indicateurs sociaux . . . . .  | 85         |
| 5.3. L'enquête par questionnaires ou entrevues . . . . .   | 87         |
| 5.3.1. La formulation des questions . . . . .  | 88         |
| 5.3.2. Les choix des techniques propres à l'enquête . . . . .  | 88         |
| 5.3.2.1. Les questionnaires écrits . . . . .   | 89         |
| 5.3.2.2. Les entrevues face à face ou par téléphone . . . . .  | 89         |
| 5.3.2.3. Les enquêtes assistées par ordinateur . . . . .   | 90         |
| 5.4. Les informateurs clés . . . . .   | 91         |
| 5.5. Les entrevues de groupe et les techniques de travail en comité . . . . .  | 92         |
| 5.5.1. Le groupe nominal . . . . .   | 93         |
| 5.5.2. La technique Delphi . . . . .   | 94         |
| 5.6. Les forums communautaires ou audiences publiques . . . . .  | 95         |
| 6. L'utilisation de l'évaluation des besoins dans la planification de programme. . . . .                                     | 96         |
| Conclusion . . . . .   | 96         |
| Bibliographie. . . . .   | 98         |
| Chapitre 5   |            |
| <b>Éléments et critères d'évaluabilité d'un programme<br/>d'intervention psychosociale . . . . .</b>                         | <b>101</b> |
| <i>Pierre Potvin</i>   |            |
| 1. Le programme d'intervention . . . . .   | 102        |
| 2. L'évaluation de l'évaluabilité d'un programme (EEP):<br>ce que doit comprendre un programme pour être évaluable . . . . . | 105        |
| 2.1. La théorie du programme . . . . .   | 106        |
| 2.2. Le rôle important des objectifs dans un programme . . . . .   | 107        |
| 2.3. L'évaluation de la théorie du programme . . . . .   | 107        |
| 3. Outil pour évaluer l'évaluabilité d'un programme . . . . .  | 108        |
| 3.1. Analyse globale du programme . . . . .  | 109        |
| 3.2. Clientèle cible . . . . .   | 109        |
| 3.3. Problématique de la clientèle cible et analyse de besoins . . . . .   | 110        |
| 3.4. Objectifs et attente de résultats . . . . .   | 110        |
| 3.5. Théorie du programme . . . . .  | 110        |
| 3.6. Ressources humaines et matérielles . . . . .  | 111        |
| 3.7. Résultat de l'évaluation . . . . .  | 111        |
| Conclusion . . . . .   | 111        |
| Annexe – Évaluation d'un programme de réadaptation en centre jeunesse –<br>exemple d'un cas d'application de l'EEP . . . . . | 112        |
| Bibliographie. . . . .   | 113        |

## Partie 3

|   |            |
|---|------------|
| <b>De l'implantation d'un programme d'intervention à la mesure des premiers effets produits, la phase PENDANT . . . . .</b> | <b>115</b> |
|---|------------|

## Chapitre 6

|  |            |
|--|------------|
| <b>Les dimensions formative et sommative de l'évaluation d'implantation d'un programme: une combinaison des perspectives objective et subjective en lien avec les modèles d'évaluation basés sur la théorie des programmes . . . . .</b> | <b>117</b> |
| <i>Jacques Joly, Luc Touchette et Robert Pauzé</i>   |            |

|  |     |
|--|-----|
| 1. Pourquoi l'évaluation durant la phase PENDANT . . . . .   | 118 |
| 2. Les objets de l'évaluation d'implantation: dimensions à observer lors de la phase PENDANT . . . . . | 122 |
| 3. Les approches, stratégies et méthodes de collecte de données . . . . .                              | 128 |
| 3.1. Les approches . . . . .   | 128 |
| 3.2. Les méthodes . . . . .  | 131 |
| 3.3. Autres considérations . . . . .   | 134 |
| 4. Tout mettre ensemble avec un exemple: le programme CAFE . . . . .                                   | 136 |
| 4.1. Étape 1: préparer l'évaluation d'implantation. . . . .  | 137 |
| 4.2. Étape 2: l'implantation initiale . . . . .  | 138 |
| 4.3. Étape 3: l'implantation mature. . . . .   | 142 |
| Conclusion . . . . .   | 143 |
| Bibliographie. . . . .   | 144 |

## Chapitre 7

|  |            |
|--|------------|
| <b>Conjuguer l'évaluation aux temps de l'implantation et des impacts . . . . .</b> | <b>147</b> |
| <i>Diane Dubeau, Sylvain Coutu et Geneviève Turcotte</i>                           |            |

|  |     |
|--|-----|
| 1. ProsPère, un projet d'intervention communautaire visant à promouvoir l'engagement paternel. . . . . | 148 |
| 1.1. Origines du projet ProsPère . . . . .   | 148 |
| 1.2. Les démarches d'une action concertée . . . . .  | 149 |
| 2. Regards évaluatifs posés sur le projet . . . . .  | 152 |
| 2.1. Analyse du milieu . . . . .   | 152 |
| 2.2. Évaluation de l'implantation. . . . .   | 153 |
| 2.3. Évaluation des impacts . . . . .  | 153 |
| 2.4. Bilans de l'évaluation des impacts: résultats et considérations d'ordre méthodologique . . . . .  | 154 |
| 3. Vers une métaévaluation. . . . .  | 155 |
| 3.1. L'évaluation, une composante intégrante d'un programme . . . . .                                  | 155 |
| 3.2. L'évaluation: une question de contexte . . . . .  | 158 |
| 3.3. La convergence des subjectivités ne devient-elle pas objectivité? . . . . .                       | 159 |
| 4. Retour sur le projet ProsPère . . . . .   | 161 |
| 4.1. Facteurs facilitants . . . . .  | 162 |
| 4.1.1. La thématique du programme d'intervention . . . . .   | 162 |
| 4.1.2. Un attrait pour un modèle souple et pour la recherche action. . . . .                           | 162 |
| 4.1.3. Des personnes clés au départ . . . . .  | 162 |
| 4.1.4. Diversité de l'équipe des chercheurs . . . . .  | 162 |

|   |            |
|---|------------|
| 4.1.5. Stabilité des chercheurs et des membres<br>des comités de pilotage au fil des années . . . . .   | 163        |
| 4.1.6. Apports rapides et concrets pour les milieux<br>de leur partenariat établi avec les chercheurs . . . . .   | 163        |
| 4.1.7. Obtention de financement pour les volets évaluatifs . . . . .  | 163        |
| 4.2. Les défis rencontrés . . . . .   | 163        |
| 4.2.1. Peu de sources de financement pour les activités d'intervention . . . . .  | 163        |
| 4.2.2. Difficultés de concilier les impératifs parfois distincts<br>entre les milieux de la recherche et de la pratique<br>(échanciers, concentrations, priorités) . . . . .  | 164        |
| 4.2.3. Enjeux éthiques liés à la participation de plusieurs sites<br>quant à la diffusion des résultats (résultats individuels vs de groupe). . . . .   | 164        |
| 4.2.4. La fin du projet. . . . .  | 164        |
| Annexe 1 – Stratégies et cibles d'intervention pour un regroupement local<br>visant à promouvoir l'engagement paternel . . . . .  | 165        |
| Annexe 2 – Analyse des plans d'action annuels de CoopERE . . . . .  | 167        |
| Annexe 3 – Processus de production des effets . . . . .   | 168        |
| Annexe 4 – Démarches pour l'élaboration du devis d'évaluation d'impacts<br>de ProsPère et modèle d'évaluation adopté. . . . .   | 169        |
| Bibliographie. . . . .  | 172        |
| Partie 4  |            |
| <b>Du bilan postimplantation à la mesure des résultats<br/>et effets à court, moyen et long termes d'un programme<br/>d'intervention, la phase APRÈS . . . . .</b>  | <b>175</b> |
| Chapitre 8  |            |
| <b>La perspective plus classique de la mesure des résultats :<br/>éléments de méthode, le passage de la recherche scientifique<br/>à son adaptation aux réalités de la recherche terrain<br/>et de l'évaluation des effets d'un programme . . . . .</b> | <b>177</b> |
| <i>Danny Dessureault et Valérie Caron</i>   |            |
| 1. Les approches théoriques en évaluation de programmes . . . . .   | 179        |
| 1.1. L'approche classique ou centrée sur les effets . . . . .   | 180        |
| 1.2. L'approche centrée sur la théorie de programmes . . . . .  | 182        |
| 1.3. L'approche centrée sur les utilisateurs. . . . .   | 183        |
| 1.4. Le modèle logique . . . . .  | 184        |
| 2. Les méthodes pour conduire une évaluation de programme sommative . . . . .   | 185        |
| 2.1. Les devis de recherche . . . . .   | 186        |
| 2.1.1. Qu'est-ce qu'un résultat? . . . . .  | 187        |
| 2.1.2. Qu'est-ce qu'un indicateur? . . . . .  | 188        |
| 2.2. La collecte de données et analyses statistiques . . . . .  | 190        |
| 3. L'interprétation des résultats . . . . .   | 190        |
| 3.1. Dans le contexte d'un devis expérimental . . . . .   | 191        |
| 3.2. Dans le contexte d'un devis quasi expérimental . . . . .   | 192        |
| Bibliographie. . . . .  | 193        |

## Chapitre 9

**Les apports de la recherche qualitative en évaluation de programmes** . . . . . 195*Daniel Turcotte, Isabelle F.-Dufour et Marie-Christine Saint-Jacques*

|   |     |
|---|-----|
| 1. La recherche qualitative . . . . .   | 196 |
| 2. Le domaine de l'évaluation . . . . .   | 197 |
| 2.1. Les courants en évaluation de programmes . . . . .   | 198 |
| 2.2. La pertinence des méthodes qualitatives en évaluation de programmes . . . . .  | 200 |
| 3. La méthodologie de l'évaluation de programmes basée sur des méthodes qualitatives . . . . .                            | 201 |
| 3.1. Le contenu du devis de recherche . . . . .   | 202 |
| 3.1.1. Les éléments conceptuels . . . . .   | 202 |
| 3.1.2. Les éléments techniques . . . . .  | 203 |
| 3.2. La collecte de données . . . . .   | 204 |
| 3.3. L'analyse des données . . . . .  | 206 |
| 3.4. Le rapport . . . . .   | 208 |
| 4. La valeur scientifique de la recherche qualitative . . . . .   | 209 |
| 5. Illustrations d'une démarche de recherche évaluative combinant données quantitatives et données qualitatives . . . . . | 210 |
| Bibliographie . . . . .   | 216 |

## Chapitre 10

**La dimension formative du bilan d'implantation : des suites d'un programme à l'état des acquis subjectivement perçus par les intervenants au programme** . . . . . 221*Carl Lacharité*

|   |     |
|---|-----|
| 1. La démarche ethnographique du bilan des processus et des acquis d'un programme par les agents . . . . .  | 224 |
| 2. Le point de vue des acteurs dans l'évaluation de l'implantation d'un programme destiné aux familles vivant en situation de vulnérabilité . . . . . | 227 |
| Conclusion . . . . .  | 232 |
| Annexe – Bilan des processus d'implantation du PAPFC . . . . .  | 233 |
| Bibliographie . . . . .   | 239 |

## Chapitre 11

**Les ressources disponibles en évaluation de programme** . . . . . 241*Marc Alain et Pierre Potvin*

|  |     |
|--|-----|
| 1. Les organismes nationaux et de référence en évaluation accessibles par Internet . . . . .     | 242 |
| 2. Les revues et périodiques spécialisés . . . . .   | 246 |
| 2.1. La <i>Revue canadienne d'évaluation de programme</i> (<evaluationcanada.ca>) . . . . .      | 246 |
| 2.2. <i>L'American Journal of Evaluation</i> (<www.eval.org/Publications/AJE.asp>) . . . . .     | 247 |
| 2.3. <i>New Direction for Evaluation</i> (<www.eval.org/Publications/NDE.asp>) . . . . .         | 247 |
| 2.4. <i>Evaluation and Program Planning</i> . . . . .  | 249 |
| 2.5. Quelques autres publications accessibles par Internet ou par bibliothèque abonnée . . . . . | 251 |

|   |            |
|---|------------|
| 3. Quelques sites Internet consacrés à des questions de méthodologies appliquées à l'évaluation. . . . .                                      | 252        |
| 4. Quelques ouvrages de référence importants . . . . .  | 253        |
| 4.1. Les ouvrages théoriques importants . . . . .   | 253        |
| 4.2. Quelques ouvrages qui portent essentiellement sur la pratique de l'évaluation . . . . .  | 254        |
| 4.3. Les <i>textbooks</i> . . . . .   | 255        |
| Chapitre 12   |            |
| <b>Les programmes se cachent pour mourir :<br/>le bouclage de l'évaluation et de la programmation<br/>dans la modernité avancée . . . . .</b> | <b>257</b> |
| <i>Denis Allard</i>   |            |
| 1. Bouclage de l'évaluation par la programmation: un cas exemplaire . . . . .   | 260        |
| 2. Bouclage de la programmation par l'évaluation: un condensé d'histoire . . . . .  | 264        |
| 3. Bouclage des deux par la subjectivation: une hypothèse de travail . . . . .  | 270        |
| 4. Pour ne pas boucler trop vite . . . . .  | 275        |
| Bibliographie. . . . .  | 276        |
| <b>Notices biographiques. . . . .</b>   | <b>283</b> |



---

# **Introduction et remerciements**

Marc Alain  
Département de psychoéducation  
Université du Québec à Trois-Rivières

Danny Dessureault  
Département de psychoéducation  
Université du Québec à Trois-Rivières

L'évaluation de programmes est à la mode. En effet, pas une organisation gouvernementale ou paragouvernementale actuelle ne saurait mettre de l'avant un programme ou une politique sans proposer qu'une forme d'évaluation quelconque ne l'accompagne; de toute manière, peu importe que l'argent dépensé soit privé ou public, une forme ou une autre de reddition de comptes finira par opérer. De fait, aussi, à peu près tous les pays développés et en voie de l'être disposent d'une société nationale d'évaluation. Et, pourtant, l'évaluation fait encore peur. Si quelques pays peuvent se targuer d'avoir su intégrer presque totalement la démarche d'évaluation à tout projet d'intervention important – on pense ici plus particulièrement à la Grande-Bretagne et aux États-Unis –, c'est loin d'être le cas partout. Or du côté des évaluateurs, à tout le moins, on s'étonne de constater que certains semblent être contre la vertu, menacés qu'ils se sentent par le processus d'évaluation en lui-même. Car pour les évaluateurs, l'évaluation ne peut être que vertueuse, qualificatif obligé de la démarche (ce qui n'est malheureusement pas toujours le cas; on connaît trop d'exemple de mauvaise évaluation ou d'évaluation faite à mauvais escient). Mais si cette perception demeure, c'est peut-être essentiellement parce qu'on la connaît encore trop peu. L'évaluation, surtout lorsqu'elle est obligée, devient vite une bête noire à empêcher l'intervention. On la considère encore trop souvent comme une froide mécanique qui, dans un souci premier de vérification de l'efficacité et de l'efficience, n'accorderait que bien peu de considération à celles et ceux qui sont au cœur même de l'application de ces programmes et politiques. C'est réitérer à quel point, au bout du compte, on ne connaît encore que très peu les finalités et les modalités de l'évaluation de programme.

Or poser la question de ce qu'est l'évaluation de programmes dans le champ de l'intervention sociale devrait commencer par une réflexion sur ce qu'est, comme tel, un programme d'intervention, objet central de notre réflexion évaluative dans le cas de l'ouvrage que nous proposons ici. Commençons tout d'abord par la notion d'intervention. Les intervenants que nous sommes, ou que nous sommes appelés à devenir, ou avons été, savent en général en donner une assez bonne définition; elle consisterait, par exemple, à faire en sorte que la situation négative ou problématique vécue par un individu ou un petit groupe d'individus soit modifiée de manière à atténuer ou à éliminer les éléments négatifs et problématiques. Se pose ensuite la question de comment réaliser la modification en question. Il s'agit bien sûr d'abord d'un processus qui commence par un diagnostic des forces et faiblesses, des tenants et des aboutissants de la situation problématique; puis, de l'intervention proprement dite où, règle générale, on mobilisera les forces en présence; et, finalement, viendra le temps du bilan, histoire de jauger de la réussite ou de l'échec de l'intervention.

En un sens, la question de définir ce qu'est un programme ne sera pas très différente de ce que nous venons de réaliser en regard de l'intervention : un diagnostic, un traitement et une évaluation des résultats atteints. Mais il y a quelques différences, bien sûr : toute intervention planifiée et réfléchie doit elle-même s'inscrire dans un programme qui offre le soutien, l'encadrement matériel et humain, les ressources, les liens, le cadre théorique, le cadre juridique et tout autre considérant nécessaire à ce que notre intervention auprès d'un individu ou d'un groupe se fasse dans les conditions optimales de réussite et de cohérence. C'est là qu'intervient l'évaluation de programme : loin de n'être que ce moment ultime où se joue la question de l'atteinte ou non des résultats souhaités, elle peut être présente à tous les moments de la vie d'un programme, de sa conceptualisation initiale jusqu'à sa « mort », moment trop rare où un programme n'a plus de raison d'être si tant est que le problème initial qui avait justifié sa naissance est réglé.

L'évaluateur, dans une telle perspective, ne peut plus ou, mieux encore, ne doit plus être l'étranger qui soupèse et qui juge ; il est au cœur même de l'intervention, comme l'un des acteurs de la réussite de l'implantation d'un programme. C'est exactement dans ce sens que ce livre se présente, notre tout premier souci ayant toujours été que l'évaluation doit constituer l'un des outils au service de l'intervention. Oui, bien sûr, le processus lui-même de l'évaluation pourra servir des finalités pas toujours directement liées à l'intervention elle-même, des questions de rendements coûts/bénéfices par exemple, mais, pour que l'évaluation soit perçue comme un soutien à l'intervention, elle ne pourra considérer ces mêmes finalités que comme accessoires. Évidemment, le dire ou l'écrire est une chose, le faire réellement en est une autre. Ce sont justement les contributions des auteurs qui ont collaboré à cet ouvrage qui nous permettent le mieux de démontrer comment cette conjugaison intervention/évaluation est non seulement possible, mais carrément inévitable si tant est que notre premier souci est d'abord et avant tout centré sur l'être humain. Tous ceux qui ont participé à ce livre présentent leur réflexion sur l'évaluation sans jamais s'inscrire à contre-courant de cette idée de base.

Mais revenons un peu en arrière. La genèse de ce livre remonte à juin 2005. C'est dans la perspective d'élaborer le matériel de support au cours obligatoire pour les étudiants de deuxième cycle inscrits à la maîtrise en psychoéducation à l'Université du Québec à Trois-Rivières que la nécessité d'un livre récent et en français sur l'évaluation de programme se fit sentir. Très rapidement, les recherches pour dénicher un tel ouvrage se révélèrent à peu près vaine. Les bibliothèques universitaires proposaient, il est vrai, de très bons ouvrages, pour la plupart des *textbook* populaires dans les facultés des universités américaines et canadiennes, mais tous, bien évidemment, rédigés en langue anglaise. Il existait bien quelques livres en français, mais

il apparut assez vite que ces ouvrages étaient soit relativement anciens, soit encore beaucoup trop limités à une discipline précise et parfois bien éloignée de la psychoéducation et de l'intervention sociale en général (l'administration des petites et moyennes entreprises, l'éducation secondaire, pour ne donner que ces deux exemples). Une partie de ces mêmes ouvrages proposaient bel et bien une réflexion fouillée sur l'évaluation dans des domaines proches de ceux qui étaient visés par le cours d'évaluation mais, et c'était là un problème dont les pourtours commençaient alors à prendre forme, leur démarche reposait sur une approche centrée sur une série d'étapes prédéterminées dont il importait de suivre la séquence. Dès ce moment, en fait, il ne nous semblait pas nécessairement très logique « d'enfermer » la démarche d'évaluation dans un carcan plus ou moins rigide, alors que le sens commun dictait plutôt l'inverse, soit que la démarche évaluative devait être élaborée selon les paramètres mêmes du programme à évaluer. D'échanges de courriels en conversations téléphoniques avec les collègues d'autres universités et départements, cette impression d'absence se confirma peu à peu. Et, de fil en aiguille, est ainsi née l'idée de profiter de ce bagage d'expériences et de réflexions sur les tenants et les aboutissants de la démarche d'évaluation de programme qui existe au Québec pour proposer le projet d'un ouvrage collectif.

Mais, pour ce faire, encore fallait-il proposer à ces collaborateurs et collaboratrices un cadre conceptuel aussi inclusif que possible et susceptible d'intégrer en un tout à peu près cohérent des expériences et des façons de faire parfois très différentes les unes des autres. De plus, un tel cadre devait nous permettre de développer, pour les étudiants et évaluateurs intéressés, non pas une énième « recette » de l'évaluation, mais un exposé de l'ensemble des différentes facettes de la démarche évaluative, de l'analyse préliminaire des besoins jusqu'à l'évaluation des effets à moyen et long terme. C'est ainsi que le modèle intégrateur du « cube » est né ; le premier chapitre de notre livre en présente particulièrement la teneur. Tous les auteurs que l'on retrouve dans les pages suivantes se sont donc vu offrir la possibilité de réfléchir à l'une ou l'autre des étapes, voire pour certains à plusieurs étapes à la fois, d'une démarche de l'évaluation qui soit non pas considérée comme un processus à part de l'intervention, mais bel et bien comme une partie intégrante de cette même intervention. En ce sens, ce livre propose donc une suite intégrée de 12 chapitres, rédigés par des auteurs qui exploitent l'une ou l'autre des diverses phases de la vie des programmes d'intervention et qui font reposer leur démonstration sur des cas très concrets, maillant ainsi la réflexion plus théorique à la pratique de l'évaluation sur le terrain de l'intervention. Ce livre s'adresse à des professionnels en place et à de futurs professionnels de l'intervention qui ont ou auront la responsabilité de réfléchir à la pertinence d'élaborer des programmes, à les définir, à les implanter et à faire face aux imprévus incontournables et, bien sûr à en évaluer les retombées réelles,

prévues et imprévues. Ces professionnels, en retour, sauront encore mieux intervenir, maintenant conscients du fait que leur travail s'inscrit dans un processus plus large, un processus qu'ils seront à même d'améliorer tant ils y auront apporté leurs savoirs, leur passion pour le métier et leur amour pour les finalités de l'intervention.

Nous tenons, finalement, à remercier tous ceux et celles qui ont bien voulu accepté de mettre la main à la pâte pour ce qui est rapidement devenu un effort vraiment collectif, ce qui inclut, bien sûr, les étudiants et professionnels sans qui pareille tâche serait à toute fin utile impossible.

Marc Alain, Ph.D.

Danny Dessureault, Ph.D.

Octobre 2008



# PARTIE 1

---

## **HISTOIRE, THÉORIES ET MODÈLES D'ÉVALUATION DE PROGRAMME**

1. Des systèmes de classification des modèles d'évaluation de programmes d'intervention psychosociale à une proposition de modèle intégrateur
2. Le développement de l'évaluation de programme



# CHAPITRE 1

---

## **Des systèmes de classification des modèles d'évaluation de programmes d'intervention psychosociale à une proposition de modèle intégrateur**

Les trois phases du cube :  
**AVANT-PENDANT-APRÈS**

Marc Alain

Département de psychoéducation  
Université du Québec à Trois-Rivières

L'évaluation, dans le champ de l'intervention sociale et dans son enseignement se cantonne généralement à deux dimensions, soit, d'une part, tout ce qui est du domaine de l'évaluation de la situation qui fait l'objet de l'intervention et, d'autre part, toute la question de l'évaluation de l'intervention elle-même, un bilan qui cible les résultats d'une intervention donnée et qui fournit des indices de l'état d'évolution de la pratique de l'intervenant lui-même. Dans une perspective plus globale, l'évaluation du programme d'intervention dans lequel s'est inscrit l'intervention visée, est, quant à elle, trop souvent laissée qui aux gestionnaires soucieux d'une efficacité essentiellement économique et administrative, qui aux évaluateurs patentés et externes à qui l'on confie le dossier du programme qui « tourne mal » et dont il convient alors de démonter la mécanique en une opération où l'on se soucie de conserver la distance garante d'une neutralité objective (Stufflebeam, 2001 ; Shadish, Cook et Leviton, 1991). C'est d'autant plus le cas maintenant que les économies contemporaines placent l'objectivité et la rentabilité aux premiers rangs des qualités recherchées, et ce, trop souvent au détriment de la reconnaissance des savoirs de l'intervention sociale (Kazi et Rostila, 2002).

Or, ce faisant, c'est la compréhension des acteurs en situation qui se perd, l'expertise tacite que les intervenants sociaux finissent toujours par développer, non pas seulement de leur propre pratique, mais des finalités beaucoup plus larges des programmes dans lesquels leurs interventions s'inscrivent, tout comme celles des autres intervenants, qu'ils soient de leur milieu et des autres milieux concernés, de façon formelle ou informelle (Zuniga, 1994). Les intervenants sociaux sont pourtant mis en contact avec les théories qui sous-tendent l'évaluation de programme (Mercier et Perreault, 2001). C'est généralement lors de la formation initiale et continue que s'effectue l'enseignement de l'élaboration et de l'évaluation des programmes d'intervention. En effet, les cursus universitaires développés afin de former les intervenants sociaux ici, au Québec, et ailleurs en Amérique et en Europe tentent d'intégrer d'une manière plus ou moins extensive le champ de l'évaluation de programme. Le problème tient ici au fait que les éléments constituant les modalités et méthodologies de l'évaluation de programme sont, d'une part, souvent très complexes à maîtriser, sans compter que, d'autre part, ils sont souvent donnés par des spécialistes du domaine, ceux-là même qui deviennent, une fois leur enseignement terminé, les évaluateurs externes et neutres à qui on confie les dossiers et mandats d'évaluation. En d'autres termes, l'évaluation de programme constitue, pour la plupart des intervenants, qu'ils soient en cours de formation ou sur le terrain, une discipline plus ou moins obscure et loin de leurs premières préoccupations. Qu'il s'établisse, en conséquence, une méfiance plus ou moins ouverte à l'endroit de l'évaluation de programme se comprend d'autant mieux (Scriven, 1993).

Ce sont les protocoles d'évaluation participative et collaborative (O'Sullivan, 2004; Bernoux, 2004), connus aussi sous l'appellation de protocoles d'évaluation de quatrième génération (Guba et Lincoln, 1989) qui devaient pourtant modifier cet état de fait. Or, bien que les fondements théoriques de ces protocoles soient établis depuis maintenant plus de 15 ans (Guba, 1990), il est toujours surprenant de constater que l'on se situe encore beaucoup au plan des principes et très peu sur ceux de l'application concrète (Lay et Papadopoulos, 2007) et, surtout, de l'implication active des intervenants sociaux au cœur de l'évaluation des programmes dans lesquels ils œuvrent, alors qu'il n'est pourtant plus à démontrer la pertinence et l'importance de leur point de vue (O'Sullivan et D'Agostino, 2002).

Une partie de cette difficulté de faire passer l'évaluation collaborative de la théorie à la pratique tient peut-être encore aujourd'hui à la difficulté de cerner aussi clairement que possible le rôle de l'évaluateur et de son travail. Les théoriciens de l'évaluation plus classique, et l'on pense tout particulièrement à Stufflebeam (2001) et à Stufflebeam et Shinkfield (2007), ont assez clairement établi la distance qui existe entre ce que l'on attend des administrateurs et gestionnaires responsables et, surtout, imputables de la bonne marche d'un programme et de l'institution qui l'héberge, d'une part, et ce que l'on attend des évaluateurs, d'autre part. En effet, de ces derniers, on attend – même encore maintenant – un jugement éclairé, un avis sur ce qui va et sur ce qui ne va pas. Cette distinction étant établie, nous ne sommes malheureusement pas tellement plus avancés quant à savoir comment l'évaluateur va pleinement jouer son rôle. Car c'est ici précisément que les oppositions de principe sont encore les plus criantes, les débats occupant encore régulièrement la table des matières des revues spécialisées. Et il ne s'agit plus ici d'une question d'approches et de méthodes, le consensus étant à peu près établi maintenant que le quantitatif et le qualitatif ont tous deux leur place en fonction de leurs avantages et inconvénients respectifs. Non, le débat est plus profond et, jusqu'à un certain point, il revient à Guba et Lincoln de l'avoir lancé, en proposant une lecture plutôt radicale de ce que les évaluateurs des trois premières générations, pour reprendre leur terminologie, adoptaient comme position. À partir du moment où l'évaluateur, comme le proposent les deux théoriciens, dépend de sa relation avec les promoteurs et les gestionnaires du programme évalué pour émettre ses recommandations et son point de vue, force nous est de reconnaître la possibilité que l'évaluation soit teintée, l'adhésion aux principes éthiques des associations d'évaluation ne constituant en tout et pour tout que des garanties théoriques.

Bien sûr, Guba et Lincoln se sont alors positionnés, à la fois sur le plan de l'approche méthodologique, un débat qui, comme on l'a vu, est aujourd'hui quelque peu dépassé, et sur le plan du rôle de l'évaluateur. À notre sens, c'est cette dernière prise de position qui est au cœur du problème ; voici de quelle

manière. Pour Guba et Lincoln, à partir du moment où l'évaluateur externe est contraint à l'entretien d'une relation privilégiée avec les gestionnaires et promoteurs du programme, il ne serait plus alors en mesure de prendre en considération les points de vue et les perspectives opposées que l'on est beaucoup plus susceptible de rencontrer chez les intervenants, voire les clients du programme. Guba et Lincoln vont du même souffle reconnaître que si l'évaluateur ne peut être tout à fait neutre lorsqu'il entretient une relation privilégiée avec les gestionnaires, il ne peut pas l'être non plus s'il décide d'adopter la perspective des intervenants et des clients. C'est donc une perspective franchement phénoménologique que nos deux auteurs vont mettre de l'avant. Ici, l'évaluateur ne jouera globalement qu'un rôle de miroir des interprétations diverses et, éventuellement, celui de catalyseur d'un rapprochement des positions et des points de vue. En d'autres termes, il est très clair que ce n'est pas son jugement à lui, comme évaluateur, que l'on recherche, mais ce qu'il pourrait faire émerger des sens que tous les acteurs en situation vont faire ressortir. Pour remettre les choses dans une perspective plus pratique, deux défis se posent : celui, tout d'abord, d'en arriver à déterminer un tel consensus, en termes de temps, de patience, de capacité d'écoute et d'introspection, pour ne nommer que ces seuls éléments, et celui, ensuite, de réaliser la congruence entre les éléments subjectifs et toute la série des éléments très objectifs, qui illustrent des réalités concrètes et mesurables de ce que fait ou non un programme et qui ne sont pas, n'en déplaisent aux évaluateurs de la quatrième génération, des objets qui n'appartiennent qu'aux promoteurs et gestionnaires des programmes.

Si, donc, le modèle proposé par Guba et Lincoln achoppe à réaliser la congruence des éléments subjectifs et des éléments objectifs du programme à évaluer, se pourrait-il toutefois que, depuis la publication du *Fourth Generation Evaluation*, quelque théoricien du domaine soit parvenu à penser la quadrature du cercle de manière pratique et directement utilisable ? Mettons tout de suite la table en nous donnant quelques balises en passant rapidement en revue les systèmes de classification des modèles d'évaluation ; cet exercice nous permettra d'autant mieux d'en arriver alors avec une proposition concrète.

## 1. QUELQUES CLASSIFICATIONS DES MODÈLES D'ÉVALUATION DE PROGRAMME

La classification des différents modèles d'évaluation de programme semble être devenue, depuis un peu plus d'une décennie maintenant, un exercice à peu près obligatoire pour quiconque proposera un modèle différent. Trop souvent, d'ailleurs, une telle classification va permettre à son auteur de démontrer en quoi l'avenue proposée permet de combler les lacunes des

modèles précédents ; à ce titre, Guba et Lincoln (1989), mais aussi Stufflebeam et Shinkfield (2007) jusque dans une certaine mesure, demeurent des cas types de ces tendances. En tentant d'éviter de retomber nous-mêmes dans ces ornières, nous ferons ici un court survol des cinq classifications qui nous semblent bien résumer les diverses manières de réaliser cet exercice ; ce faisant, il n'en sera que plus facile alors de proposer non pas ici un énième modèle, mais ce qui va plutôt tenter de se présenter comme une intégration de ce que les grands travaux théoriques en évaluation de programme ont permis d'établir, sur les plans objectifs/subjectifs, et formatifs/sommatifs.

### 1.1. Une perspective plus historique, celle de Rossi, Lipsey et Freeman (2004)

Rossi, Lipsey et Freeman distinguent quatre périodes dans le développement des modèles d'évaluation de programme, chacune étant tributaire à la fois des lacunes de la précédente et de contingences plus historiques : 1) de la Révolution industrielle jusqu'aux années 1930 lorsque l'État contemporain prend véritablement forme, 2) les années suivant la crise de 1929 et les grands travaux publics menés aux États-Unis dans la foulée du *New Deal* de Roosevelt et de Hoover, 3) la décennie 1960-1970, une période marquée par une profonde remise en question de l'État providence et, 4) des années 1980 à nos jours, une période où l'on assiste à un éclatement des modalités d'évaluation, des modèles les plus stricts qui reposent sur des protocoles quasi expérimentaux aux adaptations plus proches du terrain de la pratique, telles que l'évaluation collaborative de O'Sullivan (2004) ou, encore, le courant de l'*empowerment evaluation* de Fetterman, Kaftarian et Wandersman (1996). Pour Rossi, Lipsey et Freeman, ce n'est essentiellement qu'à partir de la quatrième période que l'on assiste réellement à l'établissement de l'évaluation de programmes comme un véritable champ de connaissances, avec ses méthodes, ses *a priori* et ses assises théoriques.

### 1.2. Les quatre modèles d'évaluation de la perspective de Stufflebeam

Stufflebeam, dans un texte marquant publié dans *New Directions for Evaluation*<sup>1</sup> (2001) va proposer une classification hiérarchique, dirions-nous, des formes d'évaluation, et ce, en 22 modèles distincts. Il s'agit ici d'une hiérarchie en ce sens que l'auteur nous propose des modèles qui vont

1. C'est d'ailleurs à partir de ce même texte qu'il a récemment publié un ouvrage récapitulatif de ses idées : D.L. Stufflebeam et A.J. Shinkfield (2007). *Evaluation Theory, Models and Applications*, San Francisco, Jossey-Bass.

se complexifier au fur et à mesure qu'ils se déclinent ; pour l'entendement, l'auteur va proposer quatre grandes classes réunissant tous ces modèles. On passe ainsi, et c'est le premier stade, des pseudo-évaluations (qui le sont en vertu du fait qu'elles proposent des résultats partiels, voire carrément invalides), aux approches orientées en fonction d'intérêts et de visées plus méthodologiques (ce sont les *question/methods oriented approaches*), qui constituent le second stade de cette hiérarchie. On retrouvera ensuite, au troisième niveau, les approches évaluatives qui sont tributaires d'une reddition de comptes et de l'amélioration de la performance (c'est-à-dire *improvement/accountability approaches*). Ce n'est finalement qu'au quatrième et dernier niveau que vont se loger les modalités d'évaluation qui visent l'atteinte d'une perspective politique et sociale, des modalités, aux dires de Stufflebeam, qui tiennent compte de la globalité des points de vue et des finalités des programmes évalués. Cette approche en 22 modèles différents pourra laisser le lecteur quelque peu perplexe : en effet, pourquoi en arriver à un nombre aussi déterminé (pourquoi 22, et non 21 ou 23), et ce, sans compter que, pour atteindre le quatrième niveau hiérarchique, il conviendra de consentir à des efforts et des ressources passablement difficiles à réunir dans des économies de taille et de niveau de développement plus modestes qu'aux États-Unis. Car ce dernier pays constitue la seule économie, à notre connaissance, en mesure de faire vivre de très nombreuses firmes publiques, semi-publiques et privées dont l'unique finalité consiste à faire de l'évaluation de programme. C'est en grande partie pour tenter d'y voir plus clair que Christina Christie va proposer non pas une nomenclature hiérarchique ou chronologique des divers modèles d'évaluation, mais plutôt une forme de généalogie des théories qui sous-tendent ces modèles.

### 1.3. L'arbre des théories de Christie

Alkin et Christie (2005 ; voir aussi le texte que Christie publie avec Azzam<sup>2</sup> en 2005) vont proposer qu'à partir d'un tronc commun, les modèles d'évaluation vont ensuite prendre l'une des trois branches à émerger de ce tronc, à savoir, 1) la branche « *Values* », 2) la branche « *Use* » et 3) la branche « *Methods* ». Dans le premier cas, le modèle « *value-engaged evaluation* » propose des approches centrées essentiellement sur le contexte dans lequel l'évaluation se déroule, et ce, dans un respect aussi strict que possible des valeurs exprimées

- 
2. De manière tout à fait stimulante et constructive, ces deux articles proposaient à quatre théoriciens assez connus (Jennifer Greene, Gary Henry, Steward Donaldson et Jean King) de réaliser un devis d'évaluation à partir d'un seul et même cas de figure, celui de l'implantation d'une école de type alternatif en Californie. On voit alors émerger des approches et des soucis très différents d'un théoricien à l'autre, des différences que Christie tente ensuite d'inscrire à l'une ou l'autre des diverses branches de sa « généalogie » des modalités d'évaluation.

par l'ensemble des parties prenantes du programme (les « *stakeholders* »); en ce sens, l'approche, par souci de respect des valeurs démocratiques, pourra délaissier certains aspects à propos desquels le consensus est mal établi, quitte à ce que le bilan demeure volontairement partiel. On pourra comprendre que le courant de l'évaluation participative et collaborative trouvera ici sa niche. La seconde branche, « *use evaluation branch* », quant à elle, est essentiellement orientée en fonction de l'utilisation aussi concrète que possible non seulement des résultats de l'évaluation, mais du processus d'évaluation lui-même qui devra, à terme, être repris en main par les intervenants du programme comme une partie intégrante des modalités d'intervention; on visera ici le développement de la capacité d'un programme à générer lui-même ses propres évaluations d'impact. Finalement, on retrouvera à la troisième branche, nommée « *method oriented branch* », tous ceux et celles qui mettent de l'avant l'avancement des méthodes et des techniques propres aux sciences sociales. Ici, le souci de répondre aux plus hauts standards statistiques et méthodologiques est constant, et, en conséquence, il n'en est que plus logique de retrouver ici des chercheurs pour qui le quantitatif et le respect des protocoles quasi-expérimentaux sont les principales marques. Or c'est précisément à ce titre que Jane Davidson (2006) ne manque pas de critiquer l'approche de classification de Christie :

Un pattern fort intéressant émerge de l'approche préconisée par Alkin et Christie en regard de la branche sur laquelle ils positionnent chacun. La branche « *Methods* » est presque exclusivement occupée par des chercheurs quantitatifs, alors que tous ceux qui choisissent des approches qualitatives tout aussi sérieuses se retrouvent sur la branche « *Value* ». Alkin et Christie pensent-ils alors que les tenants des approches qualitatives sont gouvernés par un sens aigu des valeurs en jeu tandis que les tenants des approches quantitatives ne le seraient pas ? (Davidson, 2006, p. 275 – notre traduction)<sup>3</sup>.

Or ce fameux débat entre deux approches méthodologiques apparemment irréconciliables, c'est probablement à Guba et Lincoln (1989, 2005) qu'il convient de l'avoir exacerbé à un paroxysme rarement vu jusque-là; voyons brièvement ce qu'il en est.

---

3. Dans ce même texte, Davidson pousse plus loin cette idée qu'à partir d'une branche, on peut en créer de multiples autres, à un tel point qu'une certaine confusion risque d'entacher l'effort : « *There is so much of that in other disciplines, where research ends up being narrower and narrower until it is about trivia. Or, as they say, the specialist ends up knowing more and more about less and less until he or she knows everything about nothing* » (Davidson, 2006, p. 276).

## 1.4. Les quatre générations de l'évaluation selon Guba et Lincoln

Guba et Lincoln proposent une approche de l'évolution des modèles d'évaluation qui tient compte, à l'instar de la classification de Rossi, Lipsey et Freeman, des contingences plus historiques tout en incorporant un bilan critique des lacunes de ce qu'ils considèrent être la troisième génération de l'évaluation. En fait, pour ces auteurs, cette troisième génération aurait comme principaux défauts d'être centrée trop exclusivement sur des approches quantitatives qui excluraient de l'analyse toute forme de considération subjective, tout en étant, et ce n'est pas là le moindre des reproches adressés par Guba et Lincoln, axée sur un rapport malsain entre les administrateurs (les *sponsors* de la démarche évaluative) et les évaluateurs. Cette exclusivité à la fois de la méthode et des rapports privilégiés aurait donc comme conséquence qu'elle occulterait complètement tout point de vue un tant soit peu en dissonance, notamment celui des opposants au programme évalué. La charge de Guba et Lincoln à l'endroit des évaluateurs, dirions-nous, plus « classiques » converge vers la nécessité de repenser non seulement les méthodes d'approche en évaluation mais, mieux encore, toute la question des mandats dévolus aux évaluateurs. En préconisant le recours à des approches qualitatives et une participation de l'évaluateur au cœur même du programme évalué, Guba et Lincoln proposent en effet l'avènement d'une quatrième génération de l'évaluation, où l'évaluateur se fera le catalyseur de l'atteinte des consensus nécessaires à l'appropriation du programme par l'ensemble des parties prenantes. Il devient bien tentant alors de penser que c'est précisément à cette génération qu'appartiennent les tenants de l'évaluation collaborative (que l'on pense bien sûr à O'Sullivan, 2004, mais aussi à Bernoux, 2004, pour une rare et d'autant plus précieuse contribution en français à ce débat), mais également les tenants dit de l'« *empowerment evaluation* » chère à Fetterman et à ses collègues. De façon un peu étonnante, cependant, un très récent débat dans une édition de 2007 de l'*American Journal of Evaluation* (vol. 28, n° 2, juin 2007) opposait encore ces derniers auteurs (David Fetterman et Abraham Wandersman) à d'autres que l'on serait tenté d'identifier à la troisième génération d'évaluation (Michael Scriven et Steward Donaldson), un peu comme si rien n'était réglé, et qu'une sourde et stérile querelle opposait encore des approches marquées par des méthodologies apparemment incompatibles et à l'opposé l'une de l'autre. Or, bien que son approche fasse l'objet d'une sévère critique de la part de Stufflebeam<sup>4</sup> (2001), c'est tout de

4. « Overall, there really is not much to recommend theory-based program evaluation, since doing it right is usually not feasible and since failed or misrepresented attempts can be highly counterproductive. Nevertheless, modest attempts to model programs—labeled as such—can be useful for identifying measurement variables, so long as the evaluator does not spend too much time on this and so long as the model is not considered as fixed

même à Huey-Tsyh Chen (1990) qu'il revient d'avoir su aller au-delà de la querelle méthodologique pour proposer une logique de l'évaluation qui se veut pragmatique, en ce sens qu'il n'est plus question ici de méthodes ou de philosophie des rôles de l'évaluateur et de l'évaluation, mais bel et bien de son objet central : qu'est-ce que le programme entend changer et jusqu'à quel point ce qui doit être changé l'a oui ou non été en vertu de l'implantation de ce même programme ?

### 1.5. L'évaluation guidée par la théorie du programme : la contribution de Chen à la proposition d'un modèle intégrateur

Chen (1990) commence en fait un peu là où ses contemporains s'arrêtent : le choix de la méthode et les querelles que la justification de ces choix semblent générer. Pour Chen, c'est révéler ici une dichotomie où s'opposent les modèles d'évaluation orientés par la méthode aux modèles orientés en fonction de la théorie du programme. L'originalité de l'approche développée par Chen tient en ceci qu'il est l'un des premiers à proposer que les protocoles d'évaluation n'ont pas à respecter une logique dictée par une méthodologie donnée (voire une combinaison de méthodes<sup>5</sup>), mais bien de s'articuler en une structure guidée par le programme lui-même ! À ce titre, il n'existe donc plus un ou des modèles d'évaluation, mais autant de protocoles d'évaluation qu'il y a de programmes d'intervention. Il devient alors d'autant plus important de reconnaître que l'on ne pourra à peu près rien dire des résultats d'un programme (les « *outcomes* ») si l'on ne s'est pas arrêté, avec tout le souci nécessaire, aux processus d'implantation et aux contingences environnementales, organisationnelles, politiques, sociales et économiques du programme en question. Or, et outre la critique qu'oppose Stufflebeam à l'idée que le protocole d'évaluation devrait se coller à la théorie du programme<sup>6</sup>, on pourrait également reprocher à Chen que son idée risque de générer peut-être plus de confusion que de clarté dans l'esprit des parties prenantes du programme évalué : tout évaluer est-il possible,

---

*or as a validated theory. In the rare case where an appropriate theory already exists, the evaluator can make beneficial use of it to help structure and guide the evaluation and interpret the findings*» (Stufflebeam, 2001, p. 39).

5. «[...] an atheoretical approach to evaluation is characterized by adherence to a step-by-step cookbook of doing evaluation. Program evaluation in this situation can become a set of predetermined research steps that are uniformly and mechanically applied to various programs without concern for the theoretical implications of program content, setting, participants, implementing organizations, and so on» (Chen, 1990, p. 18).
6. Nous reconnaissons qu'il n'a probablement pas tort de penser que beaucoup de programmes n'ont tout simplement aucune théorie du changement dans ce qu'ils comptent apporter.

voire tout simplement souhaitable ? L'auteur reconnaît donc à cet égard la nécessité d'en arriver à un modèle intégrateur de l'ensemble des dimensions inhérentes aux protocoles d'évaluation :

À cette étape du développement de l'évaluation, les problèmes et les frustrations ne proviennent pas tant des lacunes méthodologiques que des difficultés à établir un cadre conceptuel qui lie ou intègre les activités de l'évaluation. [...] Il est important de développer un cadre conceptuel qui intègre les éléments de l'implantation au sein de l'ensemble de la démarche évaluative. En adoptant un tel cadre conceptuel, les évaluateurs, en plus de suivre pas à pas l'implantation du programme et de ses progrès, pourront aussi travailler conjointement avec les parties prenantes de manière à améliorer les processus d'implantation (Chen, 1990, p. 30-33 – notre traduction).

## 2. LA PROPOSITION D'UN MODÈLE INTÉGRATEUR : LE MODÈLE DU CUBE

Un modèle intégrateur au sens où l'entend Chen consisterait alors à construire un modèle de compréhension des dimensions de l'évaluation de programme qui puisse tenir compte de ce que certains théoriciens ont vu comme des contradictions entre « faire de l'évaluation classique » et les avantages d'une évaluation dite de « quatrième génération ». Ce modèle, qui plus est, devrait être en mesure de tenir compte autant que possible des éléments les plus constructifs de l'évaluation plus classique au sens où l'entendent Stufflebeam et Scriven, toutes les notions entourant la construction des listes de contrôle, par exemple, qui peuvent grandement soutenir les intervenants sur le terrain.

En d'autres termes, il n'y a ni rupture au sens de ce que Guba et Lincoln entrevoyaient (c'est-à-dire entre l'évaluation au service des administrateurs et l'évaluation au service de l'ensemble des *stakeholders*), ni non plus de « petite évaluation » pour alimenter la Grande évaluation, au sens où Fetterman semble le proposer. Mais, et c'est là peut-être le point essentiel d'une argumentation sur la nécessité de repenser la structure même de l'évaluation de programme, l'idée n'est plus de proposer une énième recette, un autre compendium des étapes obligées d'une démarche d'évaluation. C'est que l'un des problèmes actuels réside en ce que l'on semble vouloir multiplier à l'infini les études de cas en tentant d'en extraire des modalités génériques et transférables à toute sorte d'autres situations.

Le modèle que nous proposons ici se situe quelque part entre des considérants purement **théoriques** des finalités de l'évaluation, des considérants beaucoup plus **politiques** (l'évaluation dans l'action, chère à Zuniga, 1994) et des considérants purement **méthodologiques** (que l'on pense notamment à

Scriven). Nous proposons donc une image des dimensions de l'ensemble des processus impliqués dans l'évaluation de programme, de leur enchaînement logique dans le temps et de leurs interactions. Comprenons ici que l'idée ne consiste pas à dire que sans tenir compte de tous ces éléments, on ne fait pas d'évaluation, loin de là ; en fait, nous nous situons dans un contexte de relative précarité de moyens et de connaissances. Compte tenu de ces limites, il est assez rare, et c'est tout à fait compréhensible, de pouvoir procéder constamment à chacune des étapes d'un processus complet d'évaluation. Ça n'empêche cependant pas les évaluateurs internes ou externes, et c'est un peu là le défi que nous proposons par ce modèle intégrateur, de comprendre où se situe leur travail et de voir comment ils peuvent être en mesure de justifier quelles devraient être les dimensions à évaluer en priorité.

De plus, un tel modèle intégrateur a l'avantage de permettre de comprendre, pour quelqu'un qui aura à faire de l'évaluation a posteriori de l'implantation du programme, quels sont les divers aspects qu'il ne connaît pas et comment ces mêmes aspects l'empêchent d'expliquer certains des résultats qu'il dévoile. On retrouvera ces situations lorsqu'un programme atteint une tout autre clientèle que celle qui avait été prévue au départ, par exemple ; et c'est là un constat fort difficile à expliquer si on ne sait rien de la stratégie d'évaluation des besoins qui avaient menés à la justification de ce programme. En connaissant théoriquement l'importance de cette dimension, l'évaluateur sera en meilleure position pour poser les questions qu'il faut aux promoteurs du programme.

On a aussi très souvent tenté d'établir une dichotomie franche entre les éléments qui relèvent de l'implantation d'un programme d'intervention de ceux qui relèvent des effets et des résultats attribuables au même programme. Et c'est sans compter aussi que toute la question de l'évaluation des besoins préalable à la conceptualisation d'un programme d'intervention est très souvent traitée comme un objet à part de l'évaluation elle-même. De plus, nous avons vu plus haut à quel point une autre distinction existe encore entre les théoriciens de l'évaluation, selon qu'ils privilégient les aspects franchement objectifs du programme (entendus ici généralement comme des éléments mesurables et chiffrés) ou ceux plus subjectifs (souvent entendus dans le sens de la défense du point de vue des utilisateurs du programme, voire aussi des intervenants eux-mêmes, par opposition, bien souvent, à leurs patrons et autres responsables des organismes et des milieux où se mettent en place les programmes). Et il est, finalement, une dernière opposition potentielle, soit celle qui existerait entre les éléments formatifs d'un programme (généralement confinés à la seule phase de l'implantation du programme) et les éléments sommatifs (que l'on associe à la mesure des résultats et effets entraînés).

Pourtant, sans connaître de manière précise quels ont été les besoins évalués qui ont justifiés la mise en place du programme, il est pratiquement impossible pour l'évaluateur de savoir jusqu'à quel point ce sont bel et bien ces besoins auxquels le programme a répondu. De même, à quoi servirait-il de démontrer que le programme a répondu à des mesures objectives de changement, mais que l'on oublie, en réalisant ce portrait, qu'il a entraîné dans son sillage un niveau inexcusable d'insatisfaction parmi la clientèle, qui se serait bien plus volontiers contentée du statu quo ? Et, finalement, serait-il possible que des aspects sommatifs existent au moment même de l'évaluation des besoins, donc dans la phase « avant » de la mise en place du programme ? Car produire un répertoire des éléments sommatifs (peu importe la phase) exige de réaliser un bilan, un état des faits (tout autant objectifs que subjectifs). Dans la même logique, les aspects formatifs ne prennent pas fin lorsque le programme est implanté ; ainsi, par exemple, la trajectoire de service d'un client ne s'arrête pas nécessairement lorsque celui-ci est parvenu au terme du programme évalué : qu'a-t-on pensé pour assurer une continuité de service ?

En bref, le modèle que nous proposons ici se veut une façon d'intégrer en un seul schéma trois niveaux différents de préoccupations et d'éléments interreliés :

- premier niveau : les trois phases de développement d'un programme (c'est-à-dire les phases « AVANT », « PENDANT » et « APRÈS ») ;
- second niveau : la dualité des aspects sommatifs et formatifs ;
- troisième niveau : la dualité des aspects subjectifs et objectifs.

La figure 1.1 prendra alors la forme d'un cube composé de trois tranches qui symbolisent l'interaction des trois phases de développement (au premier plan le AVANT, au second le PENDANT, et au fond le APRÈS), cube qui croise les aspects subjectif et objectif, ainsi que formatif et sommatif.

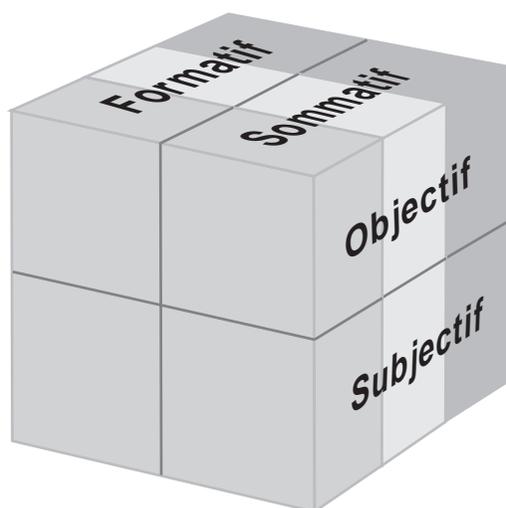
Voyons maintenant un peu plus en détails comment se définiraient les croisements potentiels de ces dimensions pour chacune des trois phases.

## 2.1. La phase AVANT

En toute logique, c'est véritablement sur la question de l'évaluation des besoins et des conditions préalables à l'implantation d'un programme d'intervention que l'essentiel de l'attention sera ici porté. Ce sont donc les considérants suivants qui seront abordés :

- évaluation des besoins ;
- état des ressources présentes ;
- état des ressources manquantes ;
- estimation des coûts d'un programme ;

Figure 1.1

**Le cube**

- évaluation de l'existence de programmes et initiatives déjà en place ;
- élaboration des éléments de justification de la pertinence d'un nouveau programme ;
- démonstration des éléments de continuité de ce nouveau programme avec les programmes et initiatives déjà en place ;
- évaluation de la capacité de mobilisation et de l'intérêt à se mobiliser des parties prenantes.

Comme le montre la figure suivante, on pourra réaliser cette évaluation de manière très exhaustive ou, encore, n'en réaliser qu'une partie mais en sachant aussi précisément que possible quels ont été les éléments qui, à cause de contraintes budgétaires ou un manque de ressources, par exemple, nous ont échappé.

## 2.2. La phase PENDANT

Nous nous situons ici dans le moment de l'implantation du programme d'intervention. Or on a trop souvent tenté de confiner cette phase à des notions qui relèvent d'une forme ou une autre de reddition de comptes relative à l'adéquation entre les besoins établis lors de la première phase – donc la planification des ressources et moyens à consacrer au programme pour que celui-ci soit susceptible de répondre à ces besoins – et les ressources effectivement consacrées. Ce serait oublier un peu vite que des éléments subjectifs

importants viennent ici jouer un rôle ; pensons seulement à la manière dont le programme a été présenté aux intervenants (un outil qui leur appartient, ou seulement une modalité supplémentaire de contrôle de leurs activités professionnelles ?) et à la réception que ceux-ci lui ont fait. Si, donc, on pourra considérer que l'adéquation entre les ressources investies et les objectifs visés demeure de l'ordre des éléments sommatifs, en revanche, tout ce qui a trait à la mesure de la capacité d'adaptation du programme aux circonstances et aux imprévus tiendra du formatif.

Figure 1.2

## Les éléments qui constituent la phase AVANT



On considérera donc, et sans prétendre ici à l'exhaustivité, les questions suivantes :

- Les ressources, humaines, financière et autres sont-elles en place ?
- A-t-on installé les infrastructures techniques nécessaires ?
- Quel est le degré de mobilisation des acteurs ?
- Le programme montre-t-il une bonne capacité à s'adapter aux changements et aux imprévus ?
- Le programme livre-t-il les extrants (*output*) prévus<sup>7</sup> ?
- Jusqu'à quel point les ressources, humaines, financières et autres sont-elles à la hauteur de ce qui avait été prévu au départ ?

Figure 1.3

### Les éléments de la phase PENDANT



7. Il importe ici de ne pas confondre la notion des extrants et celle des résultats, auquel cas on réfère aux *outcomes*, qui constituent la troisième et dernière phase du modèle.

Il va de soi que le recours à des listes de contrôle<sup>8</sup> pourra grandement soutenir le travail de terrain, particulièrement dans les situations où la cueillette de données serait effectuée par les intervenants au programme eux-mêmes<sup>9</sup>.

### 2.3. La phase APRÈS

Si tant est que la première phase mettait la priorité sur l'évaluation des besoins, et que la seconde, elle, le faisait sur les éléments de fonctionnement et d'implantation, la troisième sera fortement teintée par la question de l'évaluation des impacts et des résultats du programme d'intervention. Mais si, donc, on touche d'un peu plus près aux aspects sommatifs, c'est en considérant qu'un programme d'intervention ne se situe pas dans le vide, et qu'il fait généralement partie d'une chaîne d'intervention balisée d'éléments formatifs. On pourra aussi comprendre la question des aspects formatifs sur le plan de la continuité de l'amélioration des conditions telle que souhaitées au départ par les parties prenantes du programme.

Se retrouvent ici tous les éléments visant à s'assurer que les personnes et les groupes de personnes ayant bénéficié du programme ne se retrouvent pas laissées en plan s'il appert qu'un suivi du programme leur est nécessaire. Ce suivi devrait inclure les éléments du programme permettant d'en générer d'autres plus ou moins directement liés. Sans prétendre, pour une dernière fois, à l'exhaustivité, on se devrait d'aborder les éléments suivants :

- Quels sont les effets à court, moyen et long terme ?
- Quelle est l'importance de ces effets (c'est-à-dire la mesure de l'écart entre l'évolution des normales des conditions et ce que le programme a réellement produit) ?
- Jusqu'à quel point les personnes et groupes de personnes ciblées ont vu leurs conditions de base modifiées positivement par le programme ?
- Comment les intervenants ont-ils été en mesure d'accroître leurs connaissances et compétences d'intervenant ?
- Les effets induits par le programme sont-ils persistants ? Jusqu'à quel point ?
- Le programme a-t-il produit des conséquences et des effets non recherchés (tout autant positifs que négatifs) ?

---

8. Il s'agit ici bien sûr des fameuses *check-list* dont quelques auteurs, on pense particulièrement à Daniel Stufflebeam et Michael Scriven, se sont fait une quasi spécialité ; voir à cet effet Scriven (2007).

9. Les dix composantes du modèle de Gendreau (2001) pourraient très bien trouver ici une application tout à fait pertinente dans un souci d'évaluation.

C'est toujours en intercroisant nos axes dimensionnels sommatif et formatif, d'une part, ainsi qu'objectif et subjectif, d'autre part, que nous devrions obtenir les éléments illustrés à la figure suivante.

Figure 1.4

## Les éléments de la phase APRÈS



## EN GUISE DE CONCLUSION : COMMENT LES CHAPITRES SUIVANTS S'INTÈGRENT AU MODÈLE

Chacun des auteurs qui ont été approchés pour la rédaction de cet ouvrage collectif a contribué, au Québec et ailleurs, au développement de la discipline de l'évaluation de programme essentiellement en évaluant sur le terrain des pratiques de l'intervention sociale (en l'enseignant aussi, d'ailleurs, dans

la plupart des cas). À l'exception de quelques-uns dont nous évoquerons bientôt la contribution, toutes et tous ont été abordés avec le défi de montrer en quoi leur expérience pouvait ou non être liée à l'une ou l'autre – voire à plusieurs – des trois phases de notre modèle, en en faisant parfois aussi éclater les cadres. Voici plus précisément ce qu'il en est. Ce livre comprend quatre parties : si la première est consacrée aux questions plus théoriques et historiques, les deuxième, troisième et quatrième parties reprennent, quant à elles, les trois phases du modèle intégrateur que nous venons de présenter.

On trouvera au chapitre 3 une analyse fouillée des considérants plus subjectifs à traiter lors de la première phase, à savoir celle liée à l'élaboration d'un programme d'intervention ; ce chapitre rédigé par mesdames Hamel, Cousineau et Vézina illustre, par des exemples très concrets, toute l'importance du rôle de l'évaluateur en tant que révélateur des divers regards que les acteurs portent sur la situation, et le sens qu'ils lui attribuent, abordant en cela les considérations subjectives de la première et, en partie à tout le moins, de la seconde phase de développement d'un programme. Le quatrième chapitre, rédigé par madame Massé, se veut en quelque sorte le pendant plus objectif de ces considérations à envisager lors de la première phase de développement. L'auteure touche plus précisément la question de l'utilisation des tests et des outils plus cliniques dans leur apport à une planification réfléchie de l'intervention et de la mise sur pied de programmes ; ce chapitre passe en revue les précautions d'usage et la séquence des étapes d'une analyse aussi objective que possible des besoins, corrélât obligatoire mais trop souvent négligé du constat de l'existence d'un problème.

La troisième partie de l'ouvrage, quant à elle, tourne autour des modalités d'évaluation de l'implantation des programmes. Bien que composée de seulement deux chapitres, le premier étant le résultat du travail de messieurs Joly, Touchette et Pauzé et l'autre, celui de Mesdames Dubeau, Turcotte et de Monsieur Coutu, elle comporte des apports incontournables dans la conceptualisation des dimensions à aborder au moment de la phase dite du PENDANT, en conjuguant l'aspect tout autant formatif que sommatif et en incluant, autant que faire se peut, les dimensions objective et subjective. Tout autant au sixième qu'au septième chapitre, le recours à des exemples concrets va permettre au lecteur de se faire une idée précise de la mise en application des grands principes.

Les trois premiers chapitres de la quatrième partie sont centrés sur ce que d'aucuns considéraient il n'y a pas si longtemps comme le nerf de la guerre en évaluation, à savoir la question de la mesure des résultats d'un programme. Le chapitre de monsieur Dessureault et de madame Caron touche aux éléments et méthodes de mesure quantitative, tandis que le chapitre rédigé par monsieur Turcotte et mesdames Dufour et Saint-Jacques aborde la question sous l'angle de la contribution des approches qualitatives à cette

même préoccupation. Le chapitre 10, pour sa part, constitue la démonstration de l'importance de considérer la question de la continuité et des suites à donner à un programme d'intervention. Nous prenions en quelque sorte le pari de soulever l'importance qu'il fallait accorder au fait qu'un programme, une fois son existence à peu près complétée, devait tout de même être assuré d'une certaine pérennité, notamment par ce qu'il apportait de plus comme expérience pour les intervenants; Lacharité, auteur de ce chapitre, illustre précisément cette dimension et les mécanismes sous-jacents.

Quatre des douze chapitres que compte cet ouvrage ne s'inscrivent pas directement dans la logique du modèle intégrateur que nous avons présenté plus haut. Mis à part le cas du chapitre 11, entièrement dévolu à la question des ressources disponibles en évaluation, ce qui le rend, par le fait même, un peu à part, les trois autres chapitres constituent des compléments à la connaissance non pas propre au modèle intégrateur proposé ici, mais bien à la connaissance générale des réalités de l'évaluation de programme. Ainsi, le chapitre 2, rédigé par monsieur Goyette, propose un panégyrique du développement de la discipline, notamment au sein des institutions québécoises; le chapitre 5, rédigé par monsieur Potvin, aborde la difficile question du niveau de maturité nécessaire à un programme pour qu'il puisse faire l'objet d'une évaluation de ses résultats, une question que touche aussi un peu plus indirectement les auteurs du chapitre 6.

C'est finalement en guise de conclusion au présent ouvrage que nous présentons la contribution de Denis Allard. Ce texte, tout à fait original et très riche, présente la place de l'élaboration et de l'évaluation des programmes, la vie même d'un programme, comme s'inscrivant dans la trame de notre ère contemporaine; ce faisant, l'auteur en vient à proposer une structure de ce qu'est un programme que l'on pourra considérer comme complémentaire à celle du cube.

## BIBLIOGRAPHIE

- BERNOUX, J.-F. (2004). *L'évaluation participative au service du développement social*, Paris, Éditions Dunod.
- CHRISTIE, C.A. et T. AZZAM (2005). « What theorists say they do: A brief description of theorists' approaches », *New Directions for Evaluation*, vol. 106, p. 15-26.
- DAVIDSON, J. (2006) « Book review: Evaluation roots – Tracing theorists' views and influences », *American Journal of Evaluation*, vol. 27, p. 273-276.
- FETTERMAN, D.M., S.J. KAFTARIAN et A. WANDERSMAN (dir.) (1996). *Empowerment Evaluation*, Thousand Oaks, Sage publications.
- GENDREAU, G. (2001). *Jeunes en difficulté et intervention psychoéducative*, Montréal, Éditions Sciences et Culture.

- GUBA, E.G. (1990). *The Paradigm Dialog*, Newbury Park, Sage Publications.
- GUBA, E.G et Y.S. LINCOLN (1989). *Fourth Generation Evaluation*, Newbury Park, Sage Publications.
- GUBA, E.G. et Y.S. LINCOLN (2005). «Paradigmatic controversies, contradictions, and emerging confluences», dans N.K. Denzin et Y.S. Lincoln (dir.), *Handbook of Qualitative Research*, 3<sup>e</sup> éd., Thousand Oaks, Sage Publications, p. 191-215.
- KAZI, M.A.F. et I. ROSTILA (2002). *The Practice of Realist Evaluation in Two Countries*, Conférence présentée en octobre 2002, à l'occasion de l'European Evaluation Society, Séville, Espagne.
- LAY, M et I. PAPAPOULOS (2007). «An exploration of fourth generation evaluation in practice», *Evaluation*, vol. 13, n° 4, p. 495-504.
- MERCIER, C. et M. PERREAULT (2001). «L'évaluation de programme», dans H. Dorvil et R. Mayer (dir.), *Problèmes sociaux, théories et méthodologies*, Québec, Presses de l'Université du Québec, p. 413-430.
- O'SULLIVAN, R.G. (2004). *Practicing Evaluation, A Collaborative Approach*, Thousand Oaks, Sage Publications.
- O'SULLIVAN, R.G. et A. D'AGOSTINO (2002). «Promoting evaluation through collaboration: Findings from community-based programs for young children and their families», *Evaluation*, vol. 8, n° 3, p. 372-387.
- ROSSI, P.H., M.H. LIPSEY et FREEMAN, H.E. (2004). *Evaluation: A Systematic Approach*, 7<sup>e</sup> éd., Thousand Oaks, Sage Publications.
- SCRIVEN, M. (1993). «Hard-won lessons in program evaluation», *New Directions for Evaluation*, vol. 58, p. 8-34.
- SHADISH, W.R., T.D. COOK et L.C. LEVITON (1991). *Foundations of Program Evaluation*, Newbury Park, Sage Publications.
- STUFFLEBEAM, D.L. (2001). «Evaluation models», *New Directions for Evaluation*, vol. 89, p. 7-98.
- STUFFLEBEAM, D.L. et A.J. SHINKFIELD (2007). *Evaluation Theory, Models, and Applications*, San Francisco, Jossey-Bass.
- ZUNIGA, R. (1994). *L'évaluation dans l'action*, Montréal, Presses de l'Université de Montréal.

# CHAPITRE 2

---

## **Le développement de l'évaluation de programme**

Martin Goyette  
École nationale d'administration publique  
Université du Québec

L'affirmation selon laquelle l'évaluation de programme est d'abord et avant tout un outil d'aide à la prise de décision pour les différents détenteurs d'enjeux peut paraître courante. En effet, qui peut être contre la vertu, alors que l'intention est de comprendre les effets des interventions mises en œuvre pour mieux intervenir auprès des populations ou des groupes démunis? Ainsi, au Québec comme ailleurs, il est possible d'observer un foisonnement de travaux autour de l'évaluation de programme. Toutefois, depuis quelques décennies au Québec, un «double mouvement» favorise le développement de l'évaluation dans les secteurs de la santé et des services sociaux, de l'éducation et de l'emploi (Midy, 2001). D'une part, dans la continuité d'un courant anglo-saxon d'imputabilité, les bailleurs de fonds, ministères, agences et organismes veulent rendre des comptes à la population dans une perspective renouvelée de la démocratie et d'allocation optimale de ressources rares (Marceau, 2007). D'autre part, les organisations sociales veulent se donner les moyens d'atteindre les objectifs qu'ils se sont fixés, afin de favoriser l'adaptation, l'insertion, la citoyenneté ou la démocratie (Midy, 2001). Issus de la tradition évaluative, pour certains complémentaires pour d'autres contradictoires, ces courants amènent plusieurs manières de concevoir les rôles de l'évaluation et de l'effectuer.

C'est dans l'objectif, d'une part, de mieux comprendre ces complémentarités et ces contradictions entre différentes écoles de pensée en évaluation et, d'autre part, de mieux saisir comment celles-ci influencent notre pratique d'évaluation que nous avons voulu mettre en évidence les principales étapes du développement de l'évaluation. Nous proposons d'abord un historique de l'évaluation de programme. Ensuite, nous envisageons comment il est possible de réaliser des liens entre les éléments de cette histoire et certaines grandes perspectives théoriques et méthodologiques en évaluation.

Si le terme «évaluation» est polysémique, il faut immédiatement, pour les fins de ce texte, exclure certaines des conceptions qu'on peut s'en faire. Dans le domaine de l'intervention psychosociale, qui s'inscrit en continuité du *case work*, l'évaluation psychosociale renvoie assez souvent à l'évaluation de la situation d'un individu, de sa famille et de son environnement. Or l'évaluation de programme s'éloigne de l'évaluation psychothérapeutique, l'unité d'analyse n'étant pas l'individu ou la famille mais bien le système organisationnel qui veut mettre en place l'intervention (c'est-à-dire le projet ou le programme). En fait, l'évaluation de programme donne maintenant lieu à un vocabulaire scientifique propre qui veut la distinguer d'autres «disciplines» en sciences sociales (Dubois et Marceau, 2005).

Dans cette perspective, il y a quasiment autant d'histoires de l'évaluation que de gens qui la racontent. Ainsi, les représentations de l'évaluation varient en fonction des avancées des différentes disciplines (économie, science politique, sociologie, administration de la santé, etc.) ou des

impératifs de certains secteurs. Par exemple, comme le constatent Dubois et Marceau (2005), les éducateurs et les psychologues se regroupent autour de l'American Evaluation Association, et les évaluateurs en science politique autour de l'American Association of Public Policy. Les évaluateurs dans le domaine de la santé et des services sociaux se regroupent davantage autour d'associations professionnelles pertinentes à leur secteur, par exemple en santé publique, en épidémiologie ou en sociologie (Dubois et Marceau, 2005). Il s'en suit une fragmentation des connaissances en évaluation. En plus des périodiques spécialisés en évaluation (par exemple, *Evaluation Review*, *Evaluation Practice*, *New Directions for Evaluation*<sup>1</sup>), des textes significatifs sont publiés dans des revues sectorielles ou disciplinaires. Malgré cet éclatement, plusieurs situent les sources de l'évaluation de programme aux États-Unis et dans les pays anglo-saxons.

## 1. LES SOURCES DE L'ÉVALUATION

L'évaluation s'est développée aux États-Unis et dans les pays anglo-saxons dans le contexte de la révolution industrielle et des changements technologiques et sociaux qui l'ont accompagnée. En effet, au cours du XIX<sup>e</sup> siècle, l'industrialisation a amené une urbanisation plus importante de la société en modifiant sa structure. Pour l'analyse, il faut se souvenir que jusqu'à cette industrialisation, la responsabilité des soins aux malades et de l'aide aux démunis demeure une responsabilité individuelle et familiale. Cette responsabilité relève ainsi du domaine privé et l'État n'intervient qu'exceptionnellement en cas de crise (Mayer et Goyette, 2000). Au cours du XIX<sup>e</sup> siècle, plusieurs sociétés de secours mutuel naissent et développent leur action en complémentarité avec celle de l'Église. Toutefois, la structure d'aide en matière de services de santé et d'assistance aux démunis demeure toujours du domaine privé et communautaire. Elle relève essentiellement de familles, des congrégations religieuses et des associations de charité, et ce, avec une aide financière très particulière de la part de l'État. Cette structure traditionnelle d'aide va commencer à craquer de toutes parts avec les débuts de l'industrialisation et de l'urbanisation, qui ont pour conséquence d'augmenter sensiblement le nombre de pauvres et les problèmes sociaux, les solidarités familiales et communautaires ne pouvant plus jouer leur rôle (Mayer et Goyette, 2000). C'est dans ce contexte général que se sont développés dans les pays anglo-saxons différents programmes sociaux et de santé,

---

1. À ce sujet, la *Revue canadienne d'évaluation de programme*, publiée par la Société canadienne d'évaluation de programme, est une revue de référence au Canada. Nous reviendrons sur les activités de la Société canadienne d'évaluation.

afin de pallier aux effets négatifs de l'industrialisation et de fournir une structure sociale et des employés en santé, permettant ainsi la poursuite de cette industrialisation (Merrien *et al.*, 2005).

## 2. LES BALBUTIEMENTS DE L'ÉVALUATION : LE TESTING ET LA MESURE

Dans cette perspective, aux États-Unis comme en Grande-Bretagne, les différents gouvernements ont mis en place des réformes dans le domaine de la santé et de la santé publique (par la création de ce qui sera les hôpitaux), de la gestion des jeunes mineurs en danger (l'instauration des orphelinats) et de l'éducation (Madaus et Stufflebeam, 2000). C'est dans ce dernier domaine que l'évaluation a pris son envol, tellement qu'aux États-Unis jusqu'au milieu des années 1950 on associe l'histoire de l'évaluation de programmes sociaux avec celle de l'éducation (Madaus et Stufflebeam, 2000; Guba et Lincoln, 1989). C'est dans cette optique que certains travaux appréhendent le développement de la fonction d'évaluation dans les administrations publiques (Marceau, 2007; Jacob, 2006; Turgeon et Dubois, 2000; Madaus et Stufflebeam, 2000), tandis que d'autres mettent l'accent sur son développement paradigmatique (Guba et Lincoln, 1989; Madaus et Stufflebeam, 2000; Fontan et Lachance, 2005).

De leur côté, Madaus et Stufflebeam (2000) décrivent cette période de naissance de l'évaluation (jusqu'au début du xx<sup>e</sup> siècle) comme « l'âge des réformes ». Les balbutiements de l'évaluation prennent la forme de différentes commissions d'enquête gouvernementales qui, à partir de systèmes d'inspection externes, mettent l'accent sur l'évaluation des différentes ressources investies sur la base de normes d'experts. Ces différentes commissions d'enquête correspondent aux premiers efforts de standardisation, afin de favoriser les comparaisons et la prise de décisions. Ces commissions d'enquête sont innovatrices, notamment parce qu'elles permettent le développement de bureaucraties professionnelles et de programmes plus systématisés, ce qui facilite la mise en place d'évaluations. Elle donne lieu à l'instauration d'un « management scientifique » associé à « l'âge de l'efficacité et des tests » (Madaus et Stufflebeam, 2000) de 1900 à 1930. Pendant cette période, les efforts de standardisation et de systématisation se poursuivent afin de mieux allouer les ressources dans une optique de rendement optimal. Ce mouvement dans les administrations publiques n'est pas étranger à l'esprit du taylorisme de l'époque, qui prend son envol dans l'ensemble des sphères de la production (Madaus et Stufflebeam, 2000). Par exemple, en éducation, les approches retenues visent l'examen et le « *testing* » des apprentissages, afin de faciliter la comparaison entre les différentes écoles, et ce, au travers des districts et des États. Au-delà de l'avis des experts des commissions

d'enquête, cette approche a l'avantage de permettre de sortir d'une vision « subjective », notamment en s'intéressant aux résultats. Mais Madaus et Stufflebeam (2000) font remarquer qu'il ne s'agit pas encore d'évaluation. Guba et Lincoln (1989, p. 26) parlent de la période allant jusqu'en 1930 comme étant celle de la « mesure ». Cette appellation désigne une approche selon laquelle l'évaluateur, au moyen de tests, observe les écarts entre les objectifs visés et les résultats atteints. Dans cette perspective, le rôle de l'évaluateur est technique : il doit utiliser les instruments existants ou en créer de nouveaux. Si, du point de vue des paradigmes, l'évaluation s'est développée après cette période de « mesure technique », Guba et Lincoln (1989) insistent pour dire que ce courant est encore d'actualité, puisqu'il demeure au cœur de plusieurs travaux en éducation (et dans le social en général) visant à hiérarchiser la performance, par exemple, des écoles.

### 3. LA DESCRIPTION

Pour Madaus et Stufflebeam (2000), la période allant de 1930 jusqu'à la fin de la Deuxième Guerre mondiale donne lieu à un raffinement des techniques et des approches développées jusque-là. Dans la continuité d'une approche expérimentale et centrée sur la « mesure » et le « *testing* », on assiste également au développement d'un paradigme évaluatif centré sur la « description » (Guba et Lincoln, 1989). En effet, pendant cette période et pendant les deux décennies qui vont suivre, les évaluateurs vont tenter de comprendre les écarts observés entre les différents résultats du programme. Dépassant le rôle de technicien, il doit décrire les forces et les faiblesses du programme pendant son expérimentation, ce qui peut parfois permettre de bonifier *in vivo* le programme, dans une perspective d'évaluation formative. Toutefois, la plupart du temps, cette méthode, qui se situe dans une épistémologie positiviste et dans une perspective sommative, est utilisée pour tester le potentiel d'attribution des résultats d'un programme à un autre ou à certaines de ses composantes. Pour Guba et Lincoln (1989, p. 28), il s'agit de la naissance de l'évaluation de programme en tant que telle.

### 4. LE JUGEMENT

Les années d'après-guerre correspondent au développement accru des programmes sociaux et d'éducation dans un contexte d'apparente abondance des ressources. Pour l'évaluation, il s'agit d'une période de consolidation de ses dimensions techniques. Dans cette perspective, Madaus et Stufflebeam (2000) décrivent cette période, qui se poursuit jusqu'au début des années 1960, comme étant celle de « l'innocence », dans la mesure où les besoins

en évaluation ne sont pas criants. Toutefois, elle laisse place à un contexte moins favorable au développement des programmes sociaux, leur efficacité étant remise en cause. Cette désillusion permet la reconnaissance de l'utilité de l'évaluation au-delà de ses dimensions descriptives. Elle permet ainsi de porter un jugement en donnant des informations utiles à la prise de décision des gestionnaires publics. De cette manière, plusieurs méthodes d'évaluation sont développées dans la décennie de 1960, en reconnaissance des buts variés et de la complexité des programmes. Ce développement constitue la condition préalable à la professionnalisation de l'évaluation (Madaus et Stufflebeam, 2000). À partir du milieu des années 1970, avec la remise en question des États providence et le souci, dans les pays anglo-saxons, de contrôler les dépenses publiques, les besoins en évaluation s'accroissent. Après avoir développé de manière plus ou moins ordonnée des interventions publiques, il faut maintenant gérer la décroissance. Ainsi, l'évaluation permet de porter un jugement qui s'inscrit dans un acte social, donc nécessairement « subjectif », même s'il se veut neutre. Pour Guba et Lincoln (1989), il s'agit d'un développement important où, au-delà de la mesure et de la description, les évaluations visent à porter un « jugement » sur l'efficacité relative de l'objet évalué en fonction de critères de mérite. C'est à partir de ce paradigme central du « jugement » que la professionnalisation de l'évaluation s'amorce. Elle renvoie bien sûr à l'idée d'une profession avec son éthique, ses pratiques, ses formations, ses méthodes et ses standards (Madaus et Stufflebeam, 2000).

La professionnalisation de l'évaluation se confirme aussi par son institutionnalisation grandissante en tant que fonction centrale de l'administration publique. Par exemple, après s'être doté d'une structure de contrôle budgétaire dans les années 1930, le Conseil du Trésor émet, en 1977, « l'acte de naissance de l'évaluation au niveau fédéral » canadien (Jacob, 2006), qui consacre la décentralisation de l'évaluation au sein des différents ministères et organismes fédéraux, avec la création souhaitée de leur propre service d'évaluation soutenu par la Bureau du contrôleur général du Canada (Jacob, 2006; Turgeon et Dubois, 2000). Bien que plusieurs jugent plutôt sévèrement les avancées du gouvernement du Canada (Jacob, 2006; Turgeon et Dubois, 2000) et du gouvernement du Québec (Marceau, 2007) en matière d'évaluation, notamment parce que les évaluations sont surtout axées autour de suivis de gestion à partir de données administratives, il n'en demeure pas moins que les arrangements institutionnels développés depuis la fin des années 1970 contribuent à la reconnaissance, en partie au moins, de l'évaluation en tant que profession<sup>2</sup>.

2. Sur ce registre, Jacob (2006) mentionne que la Société canadienne d'évaluation de programme a joué un rôle important en réalisant la recension des différentes formations universitaires en évaluation et en mettant en ligne une revue savante bilingue.

Dans ce contexte général, le développement de l'évaluation, jusqu'à la fin des années 1970, témoigne d'une prédominance de l'épistémologie positiviste<sup>3</sup>, n'accordant que peu de place aux aspects qualitatifs. Dans ces différents aspects de son développement, « de grands espoirs sont placés dans la science pour aider à porter un jugement objectif sur l'efficacité et le bien-fondé des mesures publiques » (Perret, 2007, p. 31). Dans cette perspective, « le plus souvent l'acteur principal de l'évaluation est l'évaluateur, spécialiste des méthodes de recherche quantitative (modélisation, plans d'expérience, statistique) [...]. Une fois l'évaluation décidée par une autorité politique ou administrative légitime, c'est à l'évaluateur qu'il revient de concevoir et de mettre en œuvre le protocole de recherche, d'en produire des résultats et d'en proposer des conclusions » (Perret, 2007 p. 31).

Dubois et Marceau (2005, p. 14-21) schématisent les principales dimensions analytiques de l'épistémologie positiviste en évaluation<sup>4</sup>. Pour les positivistes, « l'évaluation possède une essence propre à elle-même » et est « constituée de faits et d'expériences empiriques ». Sa finalité est d'expliquer, surtout en mesurant l'efficacité et l'efficience. Les évaluations réalisées le sont pour les « décideurs publics de haut niveau » et pour les contribuables payeurs de taxes, dans une perspective où « l'évaluateur est indépendant de l'objet évalué car la connaissance est objective et « acontextuelle ». En

---

Avec la proposition de standards de pratique et d'un code d'éthique, la Société canadienne d'évaluation participe à la qualité des évaluations et à sa professionnalisation (Jacob, 2006). Il faut également noter le travail de la Société québécoise d'évaluation de programme.

3. Au sujet de l'opposition épistémologique, il faut dire que certains auteurs ne croient pas que l'évaluation constitue, au sens de Khun, une discipline, puisque les oppositions ne sont pas surtout théoriques mais méthodologiques (voir notamment Péladeau et Mercier, 1993).
4. Plusieurs théories en évaluation peuvent être associées en tout ou en partie à la perspective positiviste (Dubois et Marceau, 2005, p. 15-17). Certaines approches sont ainsi orientées « vers une plus grande amélioration des programmes et une responsabilisation plus grande de la part des décideurs publics ». Ainsi, pour la théorie centrée sur les besoins des citoyens (*consumer-oriented studies*), l'évaluation « est un processus qui permet de déterminer le mérite et la valeur du programme où le bien-être du citoyen demeure la justification première ». L'évaluation permet de **juger** « les bénéfices [...] retirés et les coûts assumés par les citoyens ». Pour la théorie orientée vers la prise de décision et la responsabilisation des actions publiques (*decision/accountability-oriented evaluation approach*), c'est l'amélioration continue du programme et le jugement sommatif du mérite et de la valeur qui sont centraux. D'autres approches s'appuient davantage sur des méthodologies ou des questions précises. La première théorie s'appuie sur les « études expérimentales » (*experimental studies*) pour déterminer les rapports de cause à effet entre variables indépendantes et dépendantes « pour répondre à des questions précises ». Une autre approche veut évaluer les programmes à partir « d'une conception théorique » (*program theory-based evaluation*) « en schématisant les liens unissant le programme aux autres variables de la problématique d'intervention, afin de mieux comprendre les raisons du succès ou de l'échec d'un programme ». Pour sa part, l'analyse coût-avantages (*cost-benefit analysis approach*) permet de comprendre, du point de vue économique, comment les ressources investies sont liées aux effets.

fait, comme la connaissance est objective, l'évaluateur doit se garder d'influencer le programme évalué. Du point de vue méthodologique, les protocoles d'évaluation sont très structurés et développés à partir de théories et d'hypothèses qui portent sur des variables dépendantes et indépendantes. De ce point de vue, la recherche évaluative permet de confirmer ou d'infirmer les hypothèses dans une perspective de généralisation éventuelle des résultats, avec des tests statistiques basées sur des données principalement quantitatives. Les « devis expériment[aux] ou quasi expériment[aux] avec données avant et après intervention et groupe(s) contrôle(s) » sont généralement privilégiés. Par ailleurs, un courant très fort dans les théories au cœur du noyau positiviste promeut l'utilisation de devis qui analysent les effets du programme en lien avec les coûts qu'il engendre pour la société. Du point de vue méthodologique, on parlera ici d'analyses coût-avantages, coût-utilité ou coût-efficacité.

## 5. L'ÉVALUATION NÉGOCIÉE

Or ce modèle, consacré dans le paradigme évaluatif du jugement (développé durant les décennies 1960 et 1970), est remis en cause par plusieurs. En fait, cette critique du modèle dominant positiviste en évaluation trouve sa source dans une critique plus générale de la recherche en sciences sociales.

Si les approches positivistes visent souvent à rendre compte des facteurs causaux « liant une intervention donnée et les effets observés » en écartant les différents obstacles à la validité interne, ces travaux ne s'attardent pas, la plupart du temps, à comprendre les mécanismes sous-jacents à l'efficacité des programmes (Desrosiers *et al.*, 1998, p. 41). C'est pourquoi ces approches ont été fortement critiquées. D'abord, on souligne « [leur] incapacité à déterminer dans quelle mesure une absence d'effet est liée ou non à des difficultés d'implantation du programme » (Desrosiers *et al.*, 1998, p. 41) et, ensuite, « le temps, l'argent et les compétences » techniques importantes qu'elles requièrent (par exemple pour des programmes locaux menés par des organismes communautaires). Ainsi, il arrive souvent que les résultats et conclusions d'une évaluation soient rendus publics après que les gestionnaires ou les politiques aient fait part de leur décision. Par ailleurs, les approches positivistes centrées sur les effets parviennent difficilement à évaluer des programmes où les interventions sont individualisées, de manière à s'adapter aux difficultés d'un client (Péladeau et Mercier, 1993), comme c'est régulièrement le cas dans le domaine social. Enfin, il faut considérer « un autre obstacle majeur associé aux contraintes liées à la comparaison de sujets exposés à l'intervention et aux sujets du groupe contrôle : en sciences sociales, il est

difficile, d'un point de vue éthique, d'assigner des sujets de manière aléatoire à une intervention et d'en refuser l'accès à d'autres pour les mettre dans le groupe contrôle» (Chagnon et Mishara, 2004, p. 12).

Avec bien d'autres<sup>5</sup>, Guba et Lincoln (1989), en proposant une «quatrième génération d'évaluation», mettent en forme les principales critiques du modèle dominant positiviste. Guba et Lincoln (1989) développent leurs critiques autour de trois arguments liés. Premièrement, ils questionnent la relation de proximité trop forte entre les gestionnaires et les évaluateurs dans la détermination des mandats et des questions d'évaluation. En fait, les évaluations seraient surtout guidées par les besoins des gestionnaires au détriment des besoins et intérêts d'autres acteurs. Dans cette perspective, il est possible de croire que l'évaluateur perd la possibilité de contribuer à la formulation des questions. Ainsi, les modes de gestion peuvent ne pas être pris en compte comme variable explicative de la variation des résultats. Enfin, dans cette grande proximité, les gestionnaires auraient un rôle trop important dans l'acceptation des résultats. La place centrale des gestionnaires dans le processus amène une deuxième critique de Guba et Lincoln (1989). En effet, les paradigmes évaluatifs développés jusqu'au début des années 1970 mettent la question de l'objectivité au centre des préoccupations. Cependant, pour Guba et Lincoln (1989), cette présomption d'objectivité des jugements évaluatifs ne tient pas. La réalité décrite par les positivistes n'existe pas, dans la mesure où toute personne qui l'observe en aura une représentation mentale particulière (Zuniga, 2001, 1994). De plus, les parties prenantes auront elles aussi un point de vue. C'est ce que Zuniga (1994) a nommé la double subjectivité.

L'évaluation est donc un acte social et politique et il faut tenir compte de la multiplicité des intérêts et des valeurs. Ainsi, les clients de l'évaluation ne sont plus seulement le gestionnaire ou le bailleur de fonds de l'évaluation, mais aussi les intervenants au cœur du programme, les bénéficiaires, etc. Or considérer l'évaluation comme un acte négocié entre les parties prenantes pour que les intérêts de chacun soient respectés, c'est considérer que l'évaluateur n'est plus dans une position d'expert omnipotent. En effet, «l'évaluation ainsi que ses conclusions ne doivent pas représenter uniquement le jugement objectif de l'évaluateur, mais elles doivent plutôt inclure et correspondre aux valeurs dominantes des acteurs engagés dans l'évaluation (Weiss, 1972)» (Dubois et Marceau, 2005, p. 18). En ce sens, Guba et Lincoln (1989) critiquent la dominance de l'approche scientifique positiviste en évaluation. À l'opposé du positivisme, les évaluateurs constructivistes considèrent l'évaluation comme une «réalité pluraliste», «priviliégiant nettement les valeurs

---

5. Voir notamment le texte de Dubois et Marceau (2005) pour une mise en perspective des enjeux épistémologiques en évaluation.

et l'intérêt des individus et des groupes touchés par les problématiques ciblées par le programme» (Chagnon et Mishara, 2004, p. 9). Partant de cette conception, la participation directe dans le programme et son environnement leur permet de « négocier et construire un sens commun relatif au programme » (Dubois et Marceau, 2005, p. 18). En effet, au contraire des tenants de la conception « acontextuelle », pour les constructivistes, il est primordial de cerner l'influence du contexte et de l'environnement sur les variations dans l'implantation et les résultats d'un programme. Dans cette perspective, la finalité de l'évaluation est de comprendre (et non pas d'expliquer), à partir des expériences personnelles des parties prenantes. Le rôle de l'évaluateur consiste à dégager un jugement collectif sur l'objet évalué dans une perspective de médiation (Fontan et Lachance, 2005).

Cette compréhension vise l'amélioration continue du programme. Le modèle de raisonnement prôné par Guba et Lincoln (1989) est donc inductif et le processus de recherche mise principalement sur des données qualitatives (mais aussi quantitatives), à partir d'un devis flexible qui s'adapte aux conditions naturelles de l'objet étudié<sup>6</sup>. Ici, les données n'ont pas à être standardisées, puisque la triangulation des informations vise surtout à établir le degré de convergence des conclusions (Péladeau et Mercier, 1993).

Cette dernière évolution des paradigmes en évaluation a coïncidé avec la reconnaissance de nouvelles pratiques évaluatives au Québec. Ces dernières se sont développées dans un contexte bien particulier. Dans les années 1990, le Québec a connu d'importantes réformes administratives de ses systèmes de soins et de services sociaux. Dans le contexte de coupures financières, le domaine de la santé mentale, avec le rapport Harnois, a été un précurseur. Les réformes misent sur une plus grande décentralisation des structures, sur une désinstitutionnalisation des services vers la communauté

---

6. Guba et Lincoln (1989) ne sont pas les seuls à s'inscrire dans une perspective constructiviste. Dubois et Marceau (2005) identifient cinq théories qui s'inscrivent au moins partiellement dans le paradigme constructiviste. Les études orientées vers le client du programme (*client-centered studies, responsive evaluation*) visent la réalisation des évaluations en partenariat avec les acteurs au cœur du programme, au contraire de certains modèles positivistes « qui se concentrent sur les besoins des citoyens (*consumer-oriented studies*) » (Dubois et Marceau, 2005, p. 21). L'évaluation constructiviste, telle qu'on la présente plus haut, s'inscrit aussi au nombre des théories identifiées par Dubois et Marceau (2005). « La troisième théorie concerne l'évaluation démocratique et délibérative (*deliberative democratic evaluation*) proposée par House et House (1998) » (Dubois et Marceau, 2005, p. 21), qui proposent un processus de validation des résultats pour consolider les conclusions. » La quatrième correspond aux travaux sur l'utilisation des résultats (*utilization-focused evaluation*), et la cinquième concerne l'évaluation par étude de cas (*case study evaluations*), qui permet l'étude du programme dans sa complexité.

(Mayer et Goyette, 2000). Comme les autres, elles veulent mettre le citoyen au cœur des préoccupations. Du point de vue des services, on ne parle plus que de la dispensation en partenariat.

Or ce contexte de réformes ne doit pas faire oublier celui des coupures budgétaires qui, comme ailleurs au Canada et aux États-Unis, a exigé un meilleur regard évaluatif sur les programmes mis en place. Ainsi, depuis lors, la clé du financement des programmes gouvernementaux est la démonstration par l'évaluation de l'utilité des interventions. Cette demande pour l'évaluation a ainsi provoqué un rapprochement entre les praticiens du social et les évaluateurs. Elle a aussi orienté la dispensation de fonds de recherche universitaire d'au moins deux manières. Premièrement, la plupart des programmes de financement exigent maintenant des partenariats préalables avec des milieux de pratique (Chagnon et Mishara, 2004). Ensuite, les appels d'offres conjoints entre les grands organismes et ministères demandeurs d'évaluation et, par exemple, le Fonds québécois de recherche sur la société et la culture (FQRSC) sont de plus en plus nombreux et très souvent orientés autour de dimensions évaluatives. En bref, partenariat et évaluation sont deux termes au cœur de la recherche universitaire québécoise.

Le contexte de changement évoqué dans les systèmes de services sociaux depuis les années 1990 a également apporté la reconnaissance des organismes communautaires en tant que nouveaux partenaires. Forts des réformes décentralisatrices et des coupures dans les services publics, les organismes communautaires participent activement à la dispensation de services aux citoyens (Mayer et Goyette, 2000). La plupart du temps, ils ne veulent pas être assimilés à des organismes intermédiaires intégrés aux réseaux de services publics. De même, les regroupements d'organismes communautaires ont toujours été critiques envers l'évaluation de leur pratique et encore davantage maintenant que les résultats de ces évaluations orientent les décisions en termes de financement. Craignant pour l'autonomie de leur mission déjà bien malmenée par l'orientation des programmes gouvernementaux de services aux citoyens, ils militent pour des pratiques évaluatives participatives, d'*empowerment*, etc.<sup>7</sup>, en reprenant les principales critiques énoncés par Guba et Lincoln (1989). En effet, la diversité d'intérêts des acteurs impliqués dans ces programmes (et leur rapport de force) façonne la mise en œuvre et leur évaluation. Certains auteurs affirment que l'évaluation en partenariat implique une asymétrie de pouvoirs (Hébert *et al.*, 2005; MSSS, 1997; McHardy, 2002). Si les relations entre l'État et le milieu communautaire se situent sur un continuum allant de la coproduction à la

---

7. Les *Cahiers de recherche sociologique* ont présenté un numéro thématique intitulé «L'évaluation sociale : un enjeu politique» (vol. 35, 2001), dans lequel plusieurs contributions défendent une vision participative de l'évaluation sociale.

sous-traitance (Hébert *et al.*, 2005), McHardy (2002) est toutefois d'avis que l'évaluation participative contribue à rééquilibrer ces rapports de force. Si certains acteurs du milieu communautaire ont d'abord craint de voir leur autonomie diluée au travers du processus d'évaluation (MSSS, 1997), ils l'utilisent maintenant pour s'approprier leur action et définir leur stratégie. Le Comité ministériel sur l'évaluation (CME) a donc proposé aux organismes communautaires une approche participative et négociatrice pour l'évaluation de leurs interventions (MSSS, 1997). Cette approche se prétend plus qu'inclusive, car elle préconise l'atteinte d'un consensus entre les partenaires. Présentée de cette manière, l'évaluation n'est plus qu'un simple outil de gestion et de contrôle des dépenses de l'État. L'évaluation devient un outil pour les organismes communautaires favorisant le développement de l'esprit associatif et des revendications du mouvement. Il devient un outil politique. Et, dans cette perspective, il faut prendre en compte que toute évaluation est chargée de politique même quand elle n'y aspire point.

Si l'évaluation répond à une demande d'imputabilité, elle témoigne également de dynamiques où les enjeux politiques sont omniprésents. Il faut donc développer une théorie de la méthode évaluative qui lui donnera à la fois de la rigueur, mais aussi de « l'interaction politique » (Zuniga, 2001, p. 24).

## 6. UNE COMPLÉMENTARITÉ DES PARADIGMES ET DES MÉTHODOLOGIES

La présentation des différentes oppositions entre le positivisme et constructivisme, et l'organisation de l'argumentation historique autour des générations de l'évaluation de Guba et Lincoln (1989), ne doivent pas faire oublier qu'en évaluation, comme dans la recherche en sciences sociales, on assiste assez souvent à des bricolages en fonction de plusieurs traditions de recherches. Ainsi, par exemple, des travaux davantage associés à la mesure en côtoient d'autres associés à la description et au jugement. Il y a donc « coexistence ». Également, on peut remarquer régulièrement des travaux à l'intérieur desquels il est possible de retrouver plusieurs paradigmes évaluatifs. Ainsi, comme le font remarquer Guba et Lincoln (1989), l'histoire de l'évaluation de programme n'est pas linéaire. Elle « ne se résume pas au remplacement d'une génération par une autre » (Fontan et Lachance, 2005, p. 6). Toutefois, on peut observer l'émergence des approches alternatives, qui se soucient des processus liés aux programmes et à l'atteinte de leurs objectifs, plutôt qu'à la simple démonstration de leurs effets. En effet, « l'utilisation maximale des résultats aux différentes phases du programme, la théorisation du programme, la participation des différents acteurs touchés par le programme

à ses différentes phases de développement et l'appropriation du programme par ces groupes et leur communauté» fait partie de la boîte à outils des évaluateurs (Chagnon, 2004, p. 10)

Il n'y a donc pas de paradigme unique : si la puissance du paradigme positiviste avec les méthodes quantitatives a été pendant longtemps à l'avant-scène, nous assistons maintenant à une coexistence des approches. Le choix qui préside aux différentes approches et paradigmes d'intervention dépend de plusieurs facteurs (Desrosiers *et al.*, 1998, p. 43-44), dont du type d'évaluation que nous voulons mettre en œuvre. En effet, une évaluation d'implantation, des effets ou du rendement peut être opérationnalisée de différentes manières. De plus, la perspective qu'on adopte au regard des finalités formative ou sommative a également une influence sur le choix des approches. Bien sûr, les contraintes organisationnelles (éthiques, politiques, etc.) et opérationnelles (temps, ressources disponibles, etc.) ont une influence cruciale sur le développement des évaluations.

Toutefois, nous sommes tenté de dire que la « position idéologique ou les modes de pensée » (Desrosiers *et al.*, 1998, p. 44) qui oriente la manière de définir les problèmes oriente également les solutions à y apporter. Ainsi, la définition des problèmes est au cœur du processus évaluatif. C'est d'ailleurs pour cette raison que nous avons opté, tout au long de notre texte, pour le choix du terme « paradigme évaluatif », mettant l'accent sur le rapport de l'évaluateur avec le monde qui l'entoure. De cette manière également, il faut saisir que, sous ces aspects d'objectivité et de rationalité, l'évaluation est toujours au cœur d'enjeux politiques qu'il faut absolument prendre en compte.

L'histoire de l'évaluation témoigne ainsi d'une prise en compte des éléments objectifs et subjectifs de ses processus. Il s'agit ici de reconnaître leur juste place afin de les comprendre en complémentarité et non pas en opposition perpétuelle.

## BIBLIOGRAPHIE

- CHAGNON, F. et B.L. MISHARA (dir.) (2004). *Évaluation de programmes en prévention du suicide*, Québec, Presses de l'Université du Québec.
- DESROSIERS, H. *et al.* (1998). *Cadre pratique pour l'évaluation de programmes : applications en promotion de la santé et en toxicomanie*, Québec, Ministère de la Santé et des Services sociaux.
- DUBOIS, N. et R. MARCEAU (2005). « Un état des lieux théoriques de l'évaluation : une discipline à la remorque d'une révolution scientifique qui n'en finit pas », *The Canadian Journal of Program Evaluation*, vol. 20, n° 1, p. 1-36.
- FONTAN, J.M. et É. LACHANCE (2005). « Pour une évaluation de cinquième génération », *Cahiers de l'ARUC-ÉS*, vol. 18, p. 1-25.

- GUBA, E.G. et Y.S. LINCOLN (1989). *Fourth Generation Evaluation*, Newbury Park, Sage Publications.
- HÉBERT, J. et al. (2005). *Analyse des pratiques d'évaluation dans les organismes communautaires. Rapport de recherche*, Montréal, Service aux collectivités de l'Université du Québec à Montréal.
- JACOB, S. (2006). « Trente ans d'évaluation de programme au Canada : l'institutionnalisation interne en quête de qualité », *Revue française d'administration publique*, vol. 3, n° 119, p. 515-531.
- MADAUS, G.F. et D.L. STUFFLEBEAM (2000). « Program evaluation: A historical overview », dans D.L. Stufflebeam, G.F. Madaus et T. Kellaghan (dir.), *Evaluation Models*, Boston, Kluwer Academic Publishers, p. 4-18.
- MARCEAU, R. (2007). « L'expérience québécoise des trente dernières années en évaluation », dans S. Jacob, F. Varone et J.-L. Genard (dir.), *L'évaluation des politiques au niveau régional*, Bruxelles, P.I.E. Peter Lang, p. 143-155.
- MAYER, R. et M. GOYETTE (2000). « Politiques sociales et pratiques sociales au Québec depuis 1960 », dans J.-P. Deslauriers et Y. Hurtubise (dir.), *Introduction au travail social*, Québec, Les Presses de l'Université Laval, p. 35-65.
- MCHARDY, J. (2002). « New partnerships require new approaches to participatory program evaluations: Planning for the future », *The Canadian Journal of Evaluation*, vol. 17, n° 2, p. 89-102.
- MERRIEN, F.-X., R. PARCHET et A. KERNEN (2005). « L'édification de l'État social », dans F.-X. Merrien, R. Parchet et A. Kernen, *L'État social. Une perspective internationale*, Paris, Armand Collin, p. 63-94.
- MIDY, F. (2001). « L'évaluation aujourd'hui au Québec : déplacement des enjeux et redéfinition du champ », *Cahiers de recherche sociologique*, vol. 35, p. 55-79.
- MINISTÈRE DE LA SANTÉ et DES SERVICES SOCIAUX (1997). *Les organismes communautaires et bénévoles. Une évaluation respectueuse des organismes communautaires et bénévoles implique un processus de négociation*, Québec, Gouvernement du Québec, Comité ministériel.
- PÉLADEAU, N. et C. MERCIER. (1993). « Approches qualitative et quantitative en évaluation de programmes », *Sociologie et Sociétés*, vol. XXV, n° 2, p. 111-123.
- PERRET, B. (2007). « L'évaluation, un nouveau mode de construction de l'action publique ? », dans B. Bouquet, M. Jaeger et I. Sainsaulieu (dir.), *Les défis de l'évaluation en action sociale et médico-légale*, Paris, Dunod, p. 29-39.
- TURGEON, J. et N. DUBOIS. (2000). « L'évaluation des politiques et des programmes au niveau du gouvernement fédéral canadien », dans S. Trosa et Y. Cousquer (dir.), *L'évaluation au service de l'avenir: Rapport annuel 1999*, Paris, La Documentation française, p. 183-204.
- ZUNIGA, R.B. (1994). « Chapitre 2 », dans R. Zuniga, *L'évaluation dans l'action*, Montréal, Presses de l'Université de Montréal, p. 51-73.
- ZUNIGA, R.B. (2001). « L'évaluation sur la place publique : science, éthique et politique », *L'évaluation sociale: un enjeu politique*, vol. 35, p. 15-29.

# PARTIE 2

---

## DE L'ÉVALUATION DES BESOINS À LA CONCEPTUALISATION D'UN PROGRAMME D'INTERVENTION, LA PHASE AVANT

3. L'état subjectif des lieux
4. L'évaluation des besoins
5. Éléments et critères d'évaluabilité  
d'un programme d'intervention  
psychosociale



# CHAPITRE 3

---

## **L'état subjectif des lieux** Les zones de conflit et les zones d'entente dans la perception des besoins à combler et de la volonté des acteurs mobilisables à s'impliquer

Sylvie Hamel  
Département de psychoéducation  
Université du Québec à Trois-Rivières

Marie-Marthe Cousineau  
École de criminologie  
Université de Montréal

Martine Vézina  
Hôpital Rivière-des-Prairies

En 1995, le Service de police de la Ville de Montréal (SPVM) confiait au Centre jeunesse de Montréal – Institut universitaire (CJM-IU) et à l'Institut de recherche pour le développement social des jeunes (IRDS) le mandat de conduire une recension des écrits sur les gangs de rue. Puis, en 1997, on leur demandait de mener une recherche terrain dans la région de Montréal, auprès d'intervenants clés et de jeunes ayant fait l'expérience des gangs, afin de compléter le portrait de la situation. Les résultats de ces deux recherches allaient servir à l'élaboration d'un Plan stratégique quinquennal, déposé en 1998 (Hamel, Fredette, Blais, Bertot et Cousineau, 1998). Ce Plan stratégique quinquennal est mieux connu aujourd'hui sous le nom de *Jeunesse et gangs de rue*. Plus précisément, la réalisation du plan stratégique visant la prévention du phénomène des gangs de rue s'amorcera en 2000 grâce à un financement fourni pour une période de cinq ans par les ministères de la Justice et du Solliciteur général du Canada, dans le cadre de la Stratégie nationale pour la sécurité communautaire et la prévention du crime. Dans trois agglomérations, des intervenants et des citoyens ont uni leurs compétences pour élaborer et mettre en œuvre des plans d'action locaux novateurs, visant à prévenir le phénomène des gangs en s'appuyant sur une approche de développement social communautaire.

Une évaluation de ce projet fut menée entre 2000 et 2005. Celle-ci visait notamment le développement d'une compréhension en profondeur des mécanismes soutenant le développement des plans d'action par les communautés. Cet aspect de la démarche est celui que l'équipe de recherche chargée de l'évaluation voulait le mieux documenter, considérant le peu de connaissances existant sur *comment développer et faire ensemble un plan d'action pour prévenir le phénomène des gangs*.

L'objectif de ce chapitre consiste à présenter les stratégies qui ont été employées par l'équipe de recherche pour conduire cette évaluation, dont les approches et les cibles étaient multiples et complémentaires. Mais, d'abord, des éléments de contexte seront présentés, de même que les principaux cadres de référence, afin de mieux comprendre les choix qui ont été faits pour conduire cette évaluation.

## 1. BREF ÉTAT DE LA SITUATION

Bien que le phénomène des gangs de rue existe depuis longtemps en Amérique, tout porte à croire qu'il progresse et se transforme. Une tournée provinciale conduite à l'automne 2002 auprès d'informateurs clés provenant de 14 régions du Québec (Cousineau, Hamel et Desmarais, 2004) confirme cette impression. La raison d'être des gangs, qui jadis se limitait à des questions d'honneur et de protection du territoire, comporte aujourd'hui de

nouveaux enjeux. Ces groupes font désormais d'importantes affaires, fort lucratives, en s'adonnant, entre autres choses, au trafic de stupéfiants et au proxénétisme.

À plus large échelle, l'Enquête policière canadienne sur les gangs de jeunes réalisée en 2002 révèle que, sur l'ensemble du territoire canadien, sept provinces reconnaissent avoir un problème de gangs (Chettleburgh, 2003). Le Québec fait partie de ces provinces et Montréal apparaît comme l'une des villes les plus affectées. Cette enquête révèle également que le Canada compte 434 gangs de jeunes sur son territoire, regroupant quelque 7 000 membres. Plus récemment, le Service canadien de renseignements criminels (2006) estimait la présence de 300 gangs de rue dans le pays, pour un total d'environ 11 000 membres.

À Montréal, la Division du renseignement du SPVM évalue la présence d'une vingtaine de gangs majeurs (groupes bien établis avec des activités criminelles organisées) et encore davantage de gangs émergents (nouveaux regroupements violents qui tentent de s'imposer sur son territoire (SPVM, 2005). Le SPVM estime aussi que, depuis 1989, sept homicides par année seraient attribuables aux gangs. Cela est sans compter que, depuis 2002, 397 tentatives de meurtre ont également été reliées à ces groupes (Barbeau, 2005).

Évidemment, le phénomène des gangs englobe une diversité de groupes qui ne s'adonnent pas tous à la violence et à la criminalité avec la même intensité (Hébert, Hamel et Savoie, 1997). Certains d'entre eux sont composés de très jeunes adolescents qui cherchent à imiter les gangs et à leur ressembler, tandis que d'autres rivalisent avec les organisations criminelles (Hamel, Cousineau et Fredette, 2004). Quoi qu'il en soit, devant l'apparente flambée du phénomène, le premier réflexe des services policiers fut de tenter d'éradiquer ces groupes, par une action essentiellement répressive. Les forces de l'ordre ont toutefois vite fait de constater que la répression n'offre que peu de résultats si elle est utilisée de manière unilatérale (Klein, 1995). Non seulement cette stratégie est insuffisante pour faire face au phénomène des gangs, mais elle pourrait même avoir pour effet d'augmenter leur violence, de renforcer leur noyau et de rallier plus sérieusement certains membres. Il aura finalement été convenu que la répression ne peut fonctionner que si elle porte un message qui est véritablement dissuasif et que, pour ce faire, celle-ci demande à être jumelée à d'autres stratégies (soutien, accompagnement, formation, réinsertion) venant compléter l'intervention initiée par l'action des policiers. Autrement dit, cette stratégie n'offre son plein rendement qu'à l'intérieur d'une approche globale et intégrée.

## 1.1. Recommandations des experts pour la prévention du phénomène des gangs

Une approche globale et intégrée est celle que recommandent les plus grands experts dans le domaine de la prévention du phénomène des gangs (Short et Hugues, 2006; Howell, 2000; Spergel et Grossman, 1998; Spergel, 1995). Ces derniers sont d'avis qu'une action multistratégique constitue la solution la plus adéquate pour répondre à la complexité du phénomène.

Spergel (1995) fut parmi les premiers à proposer un modèle d'intervention relativement structuré dont les composantes sont: la mobilisation communautaire, l'accès aux possibilités, l'intervention sociale, la répression et le développement organisationnel. Ce modèle demande que des intervenants (représentants des forces de l'ordre, agents de probation, intervenants sociaux œuvrant dans les milieux communautaires et scolaires) s'entendent par consensus sur le choix des cibles (certains jeunes, certains gangs ou certains contextes en particulier). Puis, ils doivent développer des liens de confiance et d'interdépendance pour que les stratégies qu'ils mettront en opération soient bien organisées et intégrées, s'adressant à un problème dont ils connaissent bien la nature sur un territoire local et circonscrit. Non seulement ce modèle peut donner lieu au développement de nouveaux programmes servant à couvrir l'ensemble des stratégies d'un continuum d'intervention (prévention, intervention, répression, réinsertion), mais il peut également augmenter le potentiel des services et des stratégies d'action qui sont déjà en place, en les coordonnant et en développant des procédés ou des façons d'opérer plus systématiques (Coolbaugh et Hansel, 2000).

Le modèle proposé par Irvin Spergel a grandement influencé les premières recommandations que nous avons faites au SPVM pour l'élaboration du Plan stratégique quinquennal, ainsi que les principes sur lesquels *Jeunesse et gangs de rue* s'est finalement fondé. Spergel fit l'expérimentation de son modèle dans le cadre d'un projet intitulé Little Village Gang Violence Reduction Project, mené entre 1987 et 1991 à Chicago (Spergel et Grossman, 1998). Les résultats de l'évaluation de ce projet, rendus publics récemment (Spergel, Wa et Sosa, 2006), rapportent que l'expérience eut des retombées intéressantes, se traduisant notamment par le développement de liens novateurs entre les policiers et les intervenants communautaires. Les officiers de police auraient développé une meilleure approche avec les jeunes membres de gangs et, dans certains cas, les travailleurs de rue les auraient guidés afin d'éviter qu'ils n'arrêtent les mauvaises personnes, sans pour autant perdre le respect des jeunes.

L'application du modèle est également associée à une certaine réduction de la criminalité des gangs, particulièrement chez les jeunes qui présentent de lourdes histoires accompagnées de nombreuses arrestations pour crimes violents durant la période ayant précédé le programme. Il s'avère en

fait que les tendances indiquant une baisse de la criminalité n'auraient été maintenues que pendant les trois premières années du projet, qui s'étala sur cinq ans. Ce résultat s'expliquerait par un certain effritement de la cohésion du programme causé par l'incapacité des acteurs à créer de véritables alliances avec la communauté elle-même. Un comité conseil composé de citoyens et de divers représentants du milieu fut formé pour faire face à cette difficulté, mais il se serait rapidement dissolu. Apparemment, ses membres n'auraient pas tous adhéré de la même façon au principe d'une coordination entre les services, sans compter que de nombreux antagonismes entre groupes locaux, agences sociales ainsi que représentants de l'Église et du système de justice auraient eu pour effet d'embrouiller la communication entre ces diverses instances. Après tant d'efforts, Spergel et ses collaborateurs en arrivent à la conclusion qu'encore trop peu d'attention a été consacrée au défi des relations interorganisationnelles.

Ces résultats mettent en évidence que les processus des approches globales et intégrées sont peu connus. Il faut dire que ces approches sont peu développées, et encore moins évaluées. Il est facile d'imaginer que leur développement puisse engendrer diverses difficultés techniques reliées à l'implantation de nouvelles structures et de nouveaux mécanismes d'intervention au sein d'environnements complexes. Mais, plus fondamentalement encore, une telle démarche est reliée à d'importants enjeux sociaux. Elle force en effet la rencontre et la négociation entre différentes cultures d'intervention pour une nouvelle distribution de leurs pouvoirs d'action. Cela étant, il est apparu que le développement d'approches globales et intégrées, notamment pour la prévention du phénomène des gangs, devrait être examiné de la même manière que le développement d'une innovation sociale.

## 1.2. Qu'est-ce qu'une innovation sociale ?

L'innovation sociale est un concept dont on parle de plus en plus puisque, devant l'aggravation et la complexification de certains problèmes sociaux, dont les gangs de rue, on souhaite que de nouvelles solutions se mettent en place. Cette idée se rapproche assez bien d'ailleurs de la définition du Gouvernement du Québec (cité dans Lévesque, 2005), qui associe l'innovation sociale à « toute nouvelle approche, pratique, ou intervention, ou encore tout nouveau produit mis au point pour améliorer une situation ou solutionner un problème social et ayant trouvé preneur au niveau des institutions, des organisations et des communautés ». L'innovation sociale comporte donc, d'une part, une dimension technique (nouvelle approche, pratique ou intervention), et, d'autre part, une dimension sociale représentée par les acteurs qui s'approprient et qui portent cet objet technique. En fait, l'innovation sociale s'appuie d'abord sur la capacité des acteurs à coordonner leurs

relations entre eux de telle sorte qu'ils puissent améliorer leurs performances pour, en définitive, parvenir à régler des problèmes importants (Harrison et Vézina, 2006; Nadeau, 2006).

La création d'associations inédites entre les acteurs ou le réaménagement de systèmes existants comporte cependant des difficultés qui sont plus faciles à surmonter lorsque les acteurs s'approprient le nouvel objet qu'ils portent et parviennent à lui donner un sens commun (Alter, 2002; Edquist et Hommen, 1999; Nadeau, 2006). Pour qu'une transformation s'opère, une nouvelle vision doit d'abord émerger (Lévesque, 2005), ce qui peut être ardu lorsque les acteurs impliqués défendent des systèmes de croyance divergents (Gauthier, 2006). Il s'avère en effet que, bien souvent, la diversité et l'hétérogénéité des acteurs, qui constituent la richesse du réseau qu'ils sont en train de former, posent au départ plusieurs problèmes de traduction. La solution doit alors passer par la négociation et le compromis (Callon, 1986; Callon, Lascoumes et Barthe, 2001; Klein et Harrisson, 2007).

Le processus conduisant au développement d'une innovation sociale serait donc, par définition, lié à une certaine lutte de pouvoir entre les acteurs qui, forcément, sont appelés à revoir leur position sur l'échiquier de l'action et de l'intervention. Par contre, les théories les plus récentes dans le domaine de la sociologie des organisations prétendent que les controverses et les rapports de force ne sont pas sensés perdurer (Klein et Harrisson, 2007). Ces théories avancent plutôt que les organisations s'appuient sur un certain nombre d'accords et de conventions implicites et que, par économie pour leur fonctionnement, celles-ci se mobilisent normalement pour que les protagonistes trouvent des moyens de coopérer en dépit de leurs intérêts divergents. Dans le processus conduisant au développement d'une innovation, il y aurait donc des espaces pour le dialogue et l'apprentissage (Gauthier, 2006), faisant qu'il est possible que des acteurs différents, défendant autrefois des positions divergentes, se retrouvent d'accord à propos d'un nouvel objet constituant une innovation sociale. C'est pourquoi Michel Callon et Bruno Latour (1986) sont d'avis qu'étudier un objet innovant revient à analyser le réseau qui le soutient. Comprendre une nouvelle technique, une nouvelle intervention, c'est comprendre quels acteurs se trouvent intéressés, liés, attachés ensemble et par quels procédés, dans la réalisation de cette innovation.

## **2. LE POUVOIR AUX COMMUNAUTÉS : UNE APPROCHE PARTICIPATIVE**

Les méthodes et les stratégies que nous avons utilisées pour l'évaluation du projet Jeunesse et gangs de rue tiennent compte des considérations qui viennent d'être énoncées. Elles s'adressent aux deux dimensions, technique et sociale, qui normalement interviennent dans le développement d'une innovation sociale.

La dimension sociale est en fait celle que nous avons considérée en premier lieu lorsque nous avons choisi d'embaucher un agent de liaison dans chacune des localités ciblées, tous choisis par les membres de leur communauté. Les tâches premières de ces acteurs consistaient à organiser et à coordonner les rencontres et les activités dans les milieux. Mais, plus tard, une autre fonction, plus importante, leur fut attribuée : celle d'intermédiaire entre les milieux et l'équipe de recherche. À ce titre, ils devaient faciliter la communication entre les deux protagonistes en les aidant, l'un et l'autre, à mieux comprendre leur réalité et leur culture respective. Leur implication constituait du même coup un moyen de surmonter les enjeux liés à l'appropriation du projet.

Au début de cette démarche, il était difficile pour les communautés de concevoir que nous voulions *construire avec elles* des plans d'action dont elles deviendraient éventuellement maîtres d'œuvre, plutôt que d'*expérimenter* et de *valider* auprès d'elles des techniques déjà élaborées. Mais, de notre point de vue, il était impossible que nous puissions seuls élaborer pour les collectivités des techniques capables de les rendre meilleures, plus fortes et plus attrayantes pour les jeunes et, par le fait même, plus résistantes au phénomène des gangs. Pour cela, il nous fallait recourir aux communautés elles-mêmes, à leurs expertises, leurs forces, leur connaissance des jeunes, leurs moyens et leurs ressources. Cette position rejoint l'idée de plus en plus répandue selon laquelle les évidences scientifiques ne sont pas suffisantes à la conception des programmes et, en conséquence, que la participation des milieux s'impose comme une dimension incontournable et essentielle à l'élaboration de l'action des pouvoirs publics. En fait, les programmes les plus prometteurs s'inscrivent dans une logique d'intérêt, d'effets d'entraînement et de compromis qui s'appuie elle-même sur des principes de légitimité et de crédibilité (Bilodeau, 2000). Dans cette perspective, les programmes capables de viser les multiples facettes d'une problématique complexe, comme le phénomène des gangs, sont ceux qui, à toutes les étapes, s'allient des acteurs influents, œuvrant dans des sphères d'activités pertinentes et complémentaires.

Jeunesse et gangs de rue mène donc à une *évaluation participative* et, à ce titre, veut contribuer à produire des connaissances sur les processus associés au développement de nouveaux plans d'action pour la prévention du phénomène des gangs. Elle veut, en plus, favoriser la participation et l'éducation des acteurs afin qu'ils deviennent plus habiles à contrôler leur vie et leur environnement (Fine, Torre, Boudin, Bowen, Clark, Hylton, Martinez, Roberts, Smart et Upegui, 2003; Jason, Keys, Suarez-Balcazar, Taylor et Davis, 2002). L'éducation conduirait à une conscientisation de ces acteurs quant à leur réalité sociale et aux processus qui y sont associés, ce qui, en retour, viendrait augmenter leur pouvoir et leur capacité de transformer leur milieu et leur réalité.

Certains préfèrent utiliser le terme *évaluation collaborative* (O’Sullivan, 2004), la collaboration étant conçue comme une formule qui demande aux acteurs un plus grand engagement que leur simple participation. Bien souvent, la participation des acteurs les limite à se soumettre au devis des chercheurs. La collaboration, elle, induirait une véritable intégration des acteurs dans un certain nombre de décisions en lien non seulement avec l’action, mais aussi avec la recherche et l’évaluation. Le pouvoir et la capacité des acteurs à prendre le contrôle de leur environnement ne pourraient d’ailleurs être augmentés que par l’entremise d’une réelle collaboration (implication).

Cela dit, le pouvoir que l’on souhaite ainsi donner aux acteurs nous renvoie sans contredit au créneau de *l’empowerment evaluation* (Fetterman et Wandersan, 2007). Dans ce cadre, la participation est encouragée dans le but explicite d’augmenter le pouvoir des participants et des communautés. En fait, cette approche est celle qui s’avance le plus sur le terrain de la collaboration, faisant qu’elle n’est pas seulement favorisée dans le dessein d’enrichir ou d’augmenter la validité de la recherche, mais bien pour augmenter la capacité des acteurs à diriger eux-mêmes les opérations entourant la planification du programme, son implantation et son évaluation<sup>1</sup>. Cette façon de faire aurait pour effet, en retour, d’augmenter les probabilités que les programmes donnent les résultats escomptés (Jason, Keys, Suarez-Balcazar, Taylor et Davis, 2002).

Une valeur fondamentale de ce type d’évaluation consiste donc à influencer la qualité des programmes, en augmentant le pouvoir et la responsabilité des acteurs. En participant au processus de l’évaluation, les acteurs démystifient le fonctionnement du programme et son utilisation, et s’approprient la recherche comme un outil pouvant les aider à progresser. Dans ce cadre, l’efficacité du chercheur tient à sa capacité de faire participer les gens et de leur rendre accessibles les outils de la recherche. Fetterman (1994) précise que les rôles du chercheur relèvent à la fois de l’éducation, de la facilitation, de la médiation, de l’illumination et de la libération. L’évaluateur est donc un expert qui, parmi d’autres experts, partage ses connaissances, indique ce qu’il pense être le meilleur, mais qui, dans une perspective de démocratisation du savoir, veut que les acteurs participent aux processus décisionnels et deviennent éventuellement capables de les assumer seuls. Il va sans dire que ce type d’évaluation exige du chercheur qu’il soit plongé dans l’action. Et si celui-ci se montre véritablement fidèle aux principes de cette approche, les participants ne seront pas ses subordonnés, mais bien des collaborateurs, des associés, en quelque sorte, préoccupés par la production d’un savoir

1. Dans cette perspective, Micheal Patton (1998, p. 226) écrivait : « *Findings have a very short “half-life” – to use a physical science metaphor. They deteriorate very quickly as the world change rapidly. Specific findings typically have a small window of relevance. In contrast, learning to think and act evaluatively can have an ongoing impact.* »

qui prend tout son sens pour eux. Dans ce cadre toujours, le chercheur, de son côté, ne peut pas non plus se considérer comme un observateur neutre cherchant à comprendre la situation faisant l'objet de la recherche. Il se lie plutôt aux autres pour mieux saisir la nature des faits qui émergent de cette expérience (Dolbec, 2003).

### 3. UNE MÉTHODOLOGIE PRINCIPALEMENT QUALITATIVE

Une petite partie seulement de la recherche associée à l'évaluation participative menée par l'équipe de Jeunesse et gangs de rue utilise une méthodologie quantitative, prenant la forme de questionnaires dont l'analyse est essentiellement descriptive. Ces questionnaires constituent, comme nous le verrons plus loin, l'essentiel d'une enquête sur le climat social conduite dans les localités pour évaluer dans quelles conditions de départ ces dernières amorcent la démarche. L'évaluation que nous voulions mener, par ailleurs, répondait davantage à une posture constructiviste et commandait plutôt que nous utilisions une méthode qualitative. Nous voulions en effet saisir la réalité des acteurs à partir de leurs propres cadres de référence, ceux avec lesquels ils interprètent les faits, de même que leurs actions et les pensées et sentiments qui s'y rattachent. Cette approche s'inspire de la phénoménologie qui défend l'idée selon laquelle il est essentiel de prendre en considération le sens qu'accordent les individus au monde qui les entoure étant donné qu'ils sont en rapport avec lui et qu'ils se projettent en lui, ce qui les amène à développer une conscience à son égard (Boutin, 2006), et à poser des actions en conséquence. En clair, la méthode qualitative ou *interprétative* que nous avons utilisée pour faire notre recherche allait nous permettre de voir comment les acteurs auxquels nous étions associées se représentent et construisent le monde, leur monde.

Une approche qualitative en évaluation est également utile quand la recherche ne poursuit pas d'objectifs particuliers sinon que de savoir comment se vit le programme (Patton, 1987). Autrement dit, une telle approche est utile à un stade exploratoire, alors que les connaissances sur le programme ne sont pas encore assez avancées pour pouvoir formuler des questions d'évaluation précises. Adoptant une approche inductive, on souhaite faire émerger la théorie des données recueillies, plutôt que de vérifier si celles-ci concordent avec des hypothèses posées au départ, comme le commande une approche déductive. En fait, l'approche inductive est associée à un contexte de l'activité scientifique dite de découverte (Lessard-Hébert, Goyette et Boutin, 1996).

### 3.1. Plusieurs sources de données

Les données de l'évaluation réalisée dans le cadre de Jeunesse et gangs de rue proviennent de sources variées. Les informations issues d'entretiens semi-dirigés réalisés auprès d'acteurs clés y occupent une place importante. Plusieurs grilles d'entrevue sont conçues dans le but d'obtenir des témoignages diversifiés sur les expériences de partenariat vécues par certaines personnes entourant le projet Jeunesse et gangs de rue. À partir de canevas proposés par l'équipe de recherche, les grilles d'entretien sont ajustées pour les comités d'action locale, possédant chacun ses propres caractéristiques et cheminement. Les modifications sont apportées sous l'œil avisé des agents de liaison qui, comme toujours, portent une attention particulière à la réalité et aux enjeux de leur communauté et veulent s'assurer que les questions posées à leurs partenaires soient pertinentes pour qu'ils en comprennent le sens et appréhendent déjà les retombées que pourrait avoir leur témoignage. À partir de la deuxième année du projet, les agents de liaison deviennent eux-mêmes des intervieweurs auprès des partenaires.

Ces entretiens ont été réalisés à chacune des quatre années du projet avec des acteurs clés, pas nécessairement toujours les mêmes, ayant contribué aux travaux de Jeunesse et gangs de rue. Au total auront été interviewés : 33 intervenants des milieux institutionnels et communautaires à la fin de l'an 1 du projet (printemps 2001), 51 à la fin de l'an 2 (printemps 2002), 28 à la fin de l'an 3 (printemps 2003), incluant cette fois les agents de liaison, et 29 à la fin de l'an 4 (printemps 2004), incluant, encore une fois, les agents de liaison.

La méthode d'échantillonnage des intervenants à interviewer est fondée sur une technique non probabiliste qui vise à sélectionner des personnes susceptibles de présenter une diversité de points de vue et de révéler l'essence de la réalité observée (Poisson, 1991). Ainsi, les participants sont choisis en fonction de leur sexe, de leur degré et de leur durée d'implication dans les comités de travail, du type de milieu qu'ils représentent (public, communautaire, citoyen) et du territoire.

Une première approche a consisté dans une évaluation des besoins du milieu. Pour ce faire, des entretiens téléphoniques (35 en tout) ont été réalisés auprès d'informateurs clés, et quatre groupes de discussion auprès de jeunes, dont un certain nombre provenait des centres jeunesse, ont été tenus.

Parmi les informations recueillies, nous comptons également tous les procès-verbaux et documents de travail issus des réunions des comités (notamment les plans d'action présentés dans différents formats, intermédiaires et finaux), les rapports d'activités de l'équipe de recherche, les rapports annuels des organismes partenaires, divers documents d'information sur les

partenaires et leurs services, de même que des données issues d'observations participantes. Les membres de l'équipe de recherche (chercheures, coordonnatrice, professionnelle) se sont relayées tout au long du processus pour pouvoir participer à toutes les rencontres des comités de travail. Elles discutaient régulièrement de leurs observations portant notamment sur les défis de la concertation, la capacité des acteurs à travailler ensemble, les meilleurs moyens pour faciliter leur travail et leur organisation. Des notes étaient colligées durant ces rencontres d'équipe. Les principales idées qui en émergeaient étaient ensuite portées à la connaissance des agents de liaison qui, en retour, pouvaient ou non corroborer les conclusions qui en étaient tirées.

### 3.2. Une analyse s'appuyant sur la méthode de cas

Le matériel recueilli provenant des diverses sources fut traité de façon convergente en suivant un mode de triangulation. Ce principe se traduit en sciences sociales de la façon suivante : si un minimum de trois sources d'information convergent vers le même résultat, celui-ci est considéré comme un fait solide (Yin, 1989). Il fut codifié à partir d'une grille développée à cette fin, utilisée avec une intention de conceptualisation en tentant de relier les différents éléments émanant des données, tout en les confrontant aux éléments du modèle théorique de départ. Tout au long du processus d'analyse, la corroboration, qui se base essentiellement sur l'échange d'interprétations entre les membres de l'équipe de recherche (Poupart, Deslauriers, Groulx, Laperrière, Mayer et Pires, 1997), fut privilégiée pour le contrôle de la validité interne des analyses.

Cette façon de faire s'allie à la méthode de cas (de Bruyne, 1974; Grawitz, 1993). Yin (1989, 1993) est l'un des rares qui l'envisage en tant que procédé, qu'il définit comme «une enquête empirique qui étudie un phénomène contemporain dans son contexte de vie réelle, où les limites entre le phénomène et le contexte ne sont pas nettement évidentes, et dans lequel des sources d'information multiples sont utilisées» (Yin, 1989, p. 6). Ainsi considérée, la méthode des cas consiste à rapporter un événement à son contexte et à le considérer sous cet aspect pour voir comment il s'y manifeste et se développe. En d'autres mots, il s'agit, par la méthode des cas, de saisir comment un contexte produit l'événement que l'on veut aborder. Cette voie est privilégiée pour l'analyse de la démarche partenariale de Jeunesse et gangs de rue.

Trois cas sont examinés dans le cadre de cette recherche, soit trois localités ayant servi de sites de construction du modèle d'action. Les comités qui soutiennent l'action en construction siègent tous dans la région montréalaise. Ces localités sont choisies en fonction de deux critères: 1) elles se

disent préoccupées par le phénomène des gangs et 2) elles sont reconnues pour leur expérience de concertation. Il s'agit de Villeray-La Petite-Patrie, Montréal-Nord et Vieux-Longueuil.

Les trois cas sont d'abord analysés séparément (analyse intrasite), puis comparés (analyse intersite). Cette approche permet de généraliser avec plus de confiance les conclusions de ce type de recherche parce qu'elle permet d'observer, le cas échéant, s'il y a reproduction, ou convergence à tout le moins, des processus étudiés ou de certains éléments des processus étudiés dans différents milieux (Champagne, Contandriopoulos et Denis, 1990).

L'analyse est de nature explicative puisqu'elle cherche non seulement à déterminer quelles sont les conditions favorisant la réussite des changements de pratiques d'intervention en matière de gangs de rue, mais elle s'efforce aussi de comprendre comment ces conditions agissent sur cette réussite. Cet examen s'exécute en développant progressivement l'explication (des succès ou des revers des changements des pratiques d'intervention) à partir d'une série d'opérations interreliées auxquelles se greffent progressivement de nouvelles propositions théoriques de plus en plus précises, jusqu'à l'obtention d'un modèle théorique d'évaluation ajusté à la réalité observée (Chen, 1990). Cette technique itérative correspond à ce que Yin nomme la « construction de l'explication » (*explanation-building*; Yin, 1989), laquelle se compose des étapes suivantes :

1. choisir un modèle théorique préalable à la collecte de données ;
2. comprendre les résultats d'un premier cas à partir de ce modèle ;
3. réviser le modèle théorique ;
4. comparer d'autres éléments du cas à cette révision ;
5. réviser de nouveau le modèle théorique ;
6. comparer les révisions avec les données de deux ou trois cas ;
7. répéter ce processus autant de fois que nécessaire.

### 3.3. Deux modèles de départ

Deux modèles de départ sont utilisés ici pour guider l'évaluation. D'une part, il y a celui de la sociologie de la traduction (Callon, 1986; Callon et Latour, 1986). Il s'agit là d'une théorie relativement récente dans le champ de la sociologie des organisations qui est de plus en plus utilisée pour faire l'étude du développement des innovations sociales. Celle-ci prescrit quatre critères servant à évaluer la qualité de la solution ou de l'innovation qui prend forme (Bilodeau, 2000) : 1) la solidité du modèle d'action ; 2) l'ancrage de la solution dans son milieu ; 3) la solidité et la force d'entraînement du réseau d'acteurs et 4) la viabilité de la solution.

Le plan d'action dont il faut pouvoir évaluer la solidité, selon le premier critère proposé par le modèle de la sociologie de la traduction, est celui que nous avons nous-mêmes développé à partir de nos premières recherches sur les gangs. Le tableau qui suit résume ses principales composantes.

Tableau 3.1

### Composantes des théories du problème et de l'action promues

| Théorie du problème   | Théorie de l'action   |  |
|---|---|--|
| Philosophie   | Philosophie   | Principes directeurs de l'action   |
| Les jeunes découvrent dans les gangs le moyen de combler leurs besoins les plus fondamentaux (protection, appartenance, reconnaissance et valorisation), moyen qu'ils ne trouvent nulle part ailleurs dans leur environnement (famille, école et communauté). | La prévention des gangs tient à la mise en place des conditions nécessaires pour que <i>les jeunes s'attachent et s'intègrent à leur communauté comme ils s'attachent et s'intègrent aux gangs.</i> | <ul style="list-style-type: none"> <li>• Perspective multistratégique (sensibilisation, possibilités, intervention, mobilisation, répression; sur un continuum d'intervention).</li> <li>• Approche globale ciblant jeunes, écoles, familles, communauté.</li> <li>• Reconnaissance des forces et de l'autonomie des individus et des collectivités.</li> <li>• Développement social.</li> <li>• Partenariat local soutenu par une collaboration régionale.</li> </ul> |

Les modèles utilisés au départ, visant à guider l'évaluation, servent aussi à nous préserver d'un piège important que peut renfermer l'évaluation formative. Cette façon de faire permet en effet d'éviter de formuler des recommandations qui viennent uniquement des chercheurs, à propos de la démarche qu'ils étudient et qui se met en forme. Il est donc plus prudent de s'appuyer sur une théorie du changement et de rapporter des résultats qui rendent une image plus neutre, plus descriptive, de la progression à court ou à moyen terme de l'action (Reichardt, 1994; Weiss, 1997). Cette approche, qui s'inspire de la *theory driven evaluation* (Chen, 1990), permet de comparer ce qui est en train de se passer et de se développer à un cadre normatif ou idéal qui « anticipe » en quelque sorte les étapes qui devraient être franchies pour atteindre le résultat souhaité (Christie et Alkin, 2003). Les tenants de cette approche déplorent que, parce qu'on n'arrive à attribuer aucun effet aux programmes mis en œuvre, ces programmes ne sont pas efficaces. Ils considèrent plutôt que les chercheurs n'arrivent pas, bien souvent, à mettre à profit les dispositifs méthodologiques nécessaires pour mesurer les résultats qui méritent d'être mesurés (Chen et Rossi, 1980; Chen

et Rossi, 1989). Les projets de longue haleine, s'appuyant sur des approches participatives, souffriraient tout particulièrement de ce genre de problème (Hughes et Traynor, 2000).

Pour élaborer avec les communautés des plans d'action destinés à la prévention du phénomène des gangs, et pour évaluer les processus se rattachant à leur construction, nos modèles se sont avérés particulièrement utiles. Ils ont guidé l'évaluation, bien entendu, mais ils nous ont aussi permis de structurer les messages que nous avons la responsabilité de livrer au moment de la promotion de la démarche.

## 4. UNE ÉVALUATION SOUS PLUSIEURS ANGLES

La complexité de l'objet à évaluer nous a incitées non seulement à combiner plusieurs approches (*empowerment evaluation* et *theory driven evaluation*) mais aussi à faire une évaluation qui cible plusieurs dimensions à la fois. Cela permet notamment d'ouvrir sur un plus grand nombre de questions et de recommandations, et de développer une meilleure conscience des effets secondaires (non anticipés) ayant une influence sur l'objet qu'on évalue (Bledsoe et Graham, 2005). À vrai dire, il était bien difficile de ne choisir qu'une seule dimension à évaluer sans courir le risque de développer une connaissance incomplète à propos de l'innovation qui était en train de se mettre en forme. Et, malgré que nous ayons fait une évaluation sous plusieurs angles, nous ne considérons pas qu'elle soit exhaustive. En réalité, elle s'inscrit dans le créneau de la quasi-évaluation qui qualifie les recherches évaluatives qui ne couvrent qu'une partie des dimensions possibles (Stufflebeam, 2001). Les quatre principales dimensions que nous avons évaluées sont : les besoins, les processus, la qualité des plans d'action mis en forme et les retombées de la démarche. Pour les fins de ce chapitre, nous présentons les trois premières dimensions, soit celles qui, dans le cadre de notre démarche, ont généré les connaissances pouvant, à notre avis, se révéler les plus utiles à ceux ou celles qui pourraient avoir envie de se lancer dans une aventure s'inspirant de la nôtre.

### 4.1. L'évaluation des besoins

Au début de notre démarche, notre intention était de faire état de la situation des gangs dans chacune des localités visées. Les quatre stratégies que nous avons employées pour ce faire sont résumées dans le tableau 3.2. Celui-ci contient aussi quelques mots sur les retombées que ces stratégies ont eut sur les milieux.

Tableau 3.2

### Synthèse des quatre stratégies employées pour faire état de la situation des gangs au sein des trois localités ayant participé à Jeunesse et gangs de rue<sup>2</sup>

| Les stratégies                                      | Leurs modalités   | Les retombées pour les milieux  |
|---|---|---|
| Portrait de quartier                                | Cette collecte de données officielles permet de connaître les caractéristiques sociodémographiques des milieux, leurs principales composantes ainsi que des renseignements se rapportant à la criminalité.  | Les données des portraits de quartier ne sont pas utilisées comme telles par les acteurs durant la démarche. Elles sont néanmoins mises à la disposition de la communauté, étant jugées potentiellement utiles aux organismes désireux de faire des demandes de financement.  |
| Entretiens téléphoniques auprès d'informateurs clés | Vingt informateurs clés provenant de 18 organismes sont rejoints au début de la démarche dans Villeray-La Petite-Patrie. Dix-sept informateurs clés provenant de 17 organismes sont rejoints au début de la démarche dans Montréal-Nord. Lors de ces entretiens, les répondants indiquent ce qu'ils savent à propos de l'étendue et des manifestations du phénomène dans leur localité, leurs définitions de ce problème social, les actions qu'ils souhaitent voir se réaliser pour intervenir sur le phénomène. | Ces entretiens permettent de connaître l'existence de dynamiques qui pourraient être difficilement décelées autrement. Ces entretiens révèlent également les principales préoccupations des intervenants, de même que leurs priorités. Nous constatons aujourd'hui que les plans d'action mis en œuvre sont fidèles aux résultats de ces entretiens. Les propos des intervenants annonçaient les choix qui allaient être faits quant aux orientations et aux axes des plans d'action. |

2. Tiré de S. Hamel, M.-M. Cousineau et M. Vézina, avec la collaboration de S. Léveillé, *Guide d'action intersectorielle pour une prévention du phénomène des gangs*, Montréal, Institut de recherche pour le développement social des jeunes et Centre international de criminologie comparée, pour le Centre national pour la prévention du crime, 2007.

Tableau 3.2 (suite)

| Les stratégies               | Leurs modalités   | Les retombées pour les milieux   |
|------------------------------|---|--|
| Enquête sur le climat social | <p>Cette enquête rejoint 154 répondants de Villeray-La Petite-Patrie, 142 répondants de Montréal-Nord et 133 répondants du Vieux-Longueuil. Les groupes ciblés par cette enquête sont des résidents adultes, des jeunes et des commerçants. La collecte de données par questionnaire est laborieuse et demande des ajustements aux modes de recrutement des répondants. Les membres des comités s'impliquent néanmoins activement pour la réalisation de la collecte de données, y voyant l'occasion de connaître l'opinion et la réalité vécue par les citoyens.</p>   | <p>Les résultats de l'enquête arrivent tardivement dans le processus. Il est donc difficile de les utiliser pour qu'ils servent directement à orienter les plans d'action. Par contre, les résultats de cette enquête intéressent certaines instances et permettent aux comités d'action locale de gagner une certaine crédibilité dans leur milieu, notamment en pouvant fournir des données empiriques sur la perception du climat social dans la localité, et de faire connaître leurs travaux sur les gangs.</p>   |
| Groupes de discussion        | <p>Dans la localité du Vieux-Longueuil, quatre groupes de discussion sont tenus avec des jeunes, dont un certain nombre proviennent des centres jeunesse. Les groupes sont constitués par le comité d'action locale, qui contribue également à l'organisation et à l'animation. Les jeunes sont appelés à indiquer notamment ce qu'ils savent à propos de l'étendue et des manifestations du phénomène dans leur localité, à préciser leurs définitions de ce problème et quelles sont, selon eux, les meilleures pistes de solution à envisager. De manière plus générale, les jeunes s'expriment également sur leurs besoins et sur les activités qu'ils souhaitent voir se mettre en place dans leur localité.</p> | <p>Le portrait des gangs que dressent les jeunes illustre bien la complexité du phénomène. Au regard des constats et des conclusions tirés des discussions de groupe, les membres du comité d'action locale ajustent les cibles et les objectifs de leur plan d'action. Ils consacreront la plus grande partie de leur énergie aux jeunes à risque qui, apparemment, sont ceux qui dérangent le plus et qui sont aussi les plus susceptibles d'adhérer aux gangs de rue. Pour eux et à leur demande, les membres du comité tenteront de mettre un réseau de parrainage qui pourra mettre les jeunes en lien avec des personnes significatives qui partagent les mêmes passions qu'eux.</p> |

Par ces diverses collectes de données, nous voulions répondre au besoin tout naturel que les acteurs des milieux avaient de connaître l'ampleur du problème, et de savoir ce qu'il en était. Mais, rapidement, nous avons constaté que les données probantes issues des états de situation réalisés dans chacune des localités n'étaient pas suffisantes pour que les communautés puissent véritablement se positionner en regard du phénomène des gangs. Il

faut dire que celui-ci tend à évoluer rapidement et que les états de situation devraient idéalement être mis à jour régulièrement. À cet égard d'ailleurs, des experts dans le domaine, tel Howell (1998), recommandent de considérer avec prudence les informations visant à faire état de la situation des gangs dans un milieu donné qui sont généralement ponctuelles, statiques et peu représentatives d'un ensemble de points de vue. De telles données pourraient conduire à sous-estimer ou, à l'inverse, à surestimer la gravité du problème en question. On recommande plutôt d'établir des réseaux de communication et d'échanges d'information, constitués d'acteurs différents qui, jetant des regards variés sur le phénomène, conduisent à produire une vision nuancée de la situation à l'étude, et à constamment mettre à jour l'information sur son évolution.

Par ailleurs, le portait des manifestations du phénomène sur lequel on veut agir n'est qu'une partie de ce qui doit être mis au jour pour aider les communautés à se mobiliser. Il faut aussi leur donner une image précise des moyens et des ressources dont elles disposent pour agir. C'est pourquoi nous avons aussi fait, à cette étape de l'évaluation des besoins, des inventaires des organismes et des ressources existant déjà dans chacun des milieux, fournissant des informations mises à jour à propos non seulement de leur localisation mais aussi de leur mission, des clientèles desservies et des services offerts.

Par ailleurs, il faut laisser place aux échanges, aux discussions et aux argumentations servant à cheminer vers une entente qui servira d'assise à l'orientation qui sera donnée à l'action. Pour ce faire, les éléments de connaissance constituant l'état de la situation doivent être soumis à un processus de reconnaissance et d'appropriation mettant à profit les diverses sources de savoir issues à la fois des faits, des croyances et des expériences des intervenants. Dans la foulée des discussions qui ont été entamées dans le cadre du projet Jeunesse et gangs de rue, divers sujets ont été abordés qui tous ont eu leur utilité, même si certains semblaient passablement s'éloigner de l'objet principal. Les plus importants sont, à nos yeux, ceux ayant soulevé la controverse. Il faut comprendre que les acteurs concernés tentaient de comprendre comment ils pouvaient réaménager les systèmes d'intervention existants pour mieux faire face au phénomène des gangs de rue. Dès lors, nous avons aussi compris que l'évaluation des processus allait être d'une grande utilité dans le cadre de ce projet.

## 4.2. L'évaluation des processus

Une part importante de l'évaluation des processus s'appuie sur l'analyse des entretiens individuels conduits après des participants sur une base annuelle. Au début, les acteurs se soumettaient à l'exercice sans beaucoup

de conviction, le considérant comme un mal nécessaire puisqu'ils avaient accepté de travailler avec une équipe de recherche *qui devait accumuler toutes sortes d'informations pour pouvoir faire son travail*. Avec le temps, cependant, nous avons constaté que les acteurs se montraient de plus en plus intéressés par cette approche. Malgré leurs horaires chargés, les « habitués » ont toujours trouvé le temps nécessaire pour se prêter à ces entretiens printaniers et, à l'automne suivant, ils étaient curieux d'en connaître les résultats livrés par des rapports intérimaires et des bilans produits par l'équipe de recherche. Il est arrivé d'ailleurs que certains comités se servent des bilans des entretiens comme instruments de démarrage pour la nouvelle année qui s'amorçait.

Entre 10 et 17 personnes, selon les années et les localités, ont ainsi été rencontrées sur une base annuelle. Il s'agissait de recueillir leur point de vue et de sonder leur satisfaction sur divers sujets tels le fonctionnement de leur comité d'action locale, la qualité des relations entre les membres, le rôle de l'agent de liaison, l'avancement des travaux, la mise en réseau, le partenariat et la concertation locale, de même que la démarche de recherche participative. Les entretiens en personne étaient d'une durée moyenne d'une heure ; elles étaient généralement menées par les agents de liaison qui, connaissant bien les participants, pouvaient parfois, par ces rencontres, faire le point sur certains aspects de la démarche pour le bénéfice du fonctionnement des comités.

Pour illustrer les résultats possibles de l'évaluation des processus, nous donnons en exemple quelques citations d'informateurs rencontrés à la fin de la première et de la deuxième année.

La première citation choisie est représentative de l'opinion de plusieurs acteurs dans l'une des localités, concluant, à la fin de la première année, qu'ils sont satisfaits des échanges entre eux lors des réunions. Cette opinion se manifeste à différents égards. Près du tiers des membres estiment que le comité poursuit clairement un but commun :

*On a un peu tous le même but, c'est-à-dire faire de quoi avec nos ados, nos jeunes et notre population.*

De plus, la grande majorité témoigne du progrès de l'avancement des travaux réalisés durant l'année :

*Il y a plein d'autres personnes que j'ai rencontrées ou avec qui j'ai eu l'occasion d'échanger qui font partie d'autres organismes. Je trouve que cela facilite tellement le contact. Quand, à un moment donné, on a des questions, on ne se demande plus à qui on devrait les poser et on sait davantage vers qui se tourner, car on a un visage, on a un nom. C'est plus facile comme cela.*

Il est certain que ces renseignements étaient précieux à l'équipe de recherche qui, au début du projet, apprenait à connaître les milieux et leurs façons de faire. Il lui devenait en contrepartie possible d'en savoir un peu plus sur ce qui était attendu d'elle et de s'ajuster en conséquence. Les entretiens réalisés avec des informateurs de cette localité ont d'ailleurs été très encourageants à la fin de la première année, certains ayant exprimé que la recherche contribuait à consolider les actions déjà entreprises dans les milieux. Les notions de permanence et de continuité, dans le sens d'un continuum d'actions, ont aussi été évoquées à quelques reprises :

*Le rôle qu'elle [la recherche] est en train de jouer et qui devrait se poursuivre, ce que je pense que cela va donner à long terme, selon moi, c'est la permanence de certains projets et la continuité aussi d'un projet à un autre.*

En plus du fait que ces entrevues aient été une excellente tribune pour les participants désireux de faire savoir des choses importantes à leurs yeux à l'équipe de recherche, il s'est révélé que les résultats en ressortant, étant rendus accessibles à tous, permettaient aussi aux membres des comités eux-mêmes de vérifier si leurs opinions ou même leurs inquiétudes étaient partagées par d'autres, et si l'équipe de recherche les avait bien perçues. Ce faisant, les participants pouvaient, dans une certaine mesure, estimer les chances que la démarche d'ensemble soit influencée par leur point de vue. Certains diront qu'ils avaient tout le loisir de s'exprimer en comité. Nous savons toutefois que toute vérité n'est pas toujours bonne à dire. À tout le moins, cet adage convient bien aux groupes de travail qui apprennent à se connaître et qui ont du pain sur la planche. Certains propos se livrent mieux en coulisse, sur une base anonyme, et ne risquent pas, de cette façon, de ralentir la cadence de travail du comité. Ils donnent néanmoins quelques repères, dessinent la ligne à suivre ainsi que les règles à ne pas enfreindre. Bref, ces *conversations de couloirs* constituent en soi des lieux d'échange entre les membres des comités qui, à certaines étapes charnières d'une démarche collective, notamment au début, ne pourraient pas, sans doute, résister au face à face. Ces dernières sont d'autant plus productives si elles se produisent en présence de l'agent de liaison qui peut prendre action en regard du message qui est passé.

À titre d'exemple, il aurait sans doute été démotivant, malgré le bilan positif qu'ils faisaient de la première année, que des membres de cette localité expriment ouvertement au comité leur inquiétude persistante à ne pas voir un jour leurs travaux de réflexion prendre la forme d'actions concrètes. Par contre, sur une base individuelle et en dehors des rencontres, une telle affirmation peut avoir un effet constructif. Dans un tel contexte, cette position invite à la vigilance plutôt qu'à la démobilisation des troupes :

*Il va être important qu'il y ait des résultats tangibles tout au long du processus. Les gens n'ont pas le temps d'attendre quatre ans.*

En contrepartie, les entrevues que nous avons réalisées dans une autre localité, durant cette même année, laissent entrevoir d'autres dynamiques :

*Je ressens encore des fois des tensions comme si le milieu institutionnel enlevait quelque chose au communautaire.*

Même si les dynamiques que révèlent ces propos étaient palpables durant les réunions du comité, jamais elles n'ont été aussi clairement énoncées. Les confidences livrées à l'agent de liaison sur une base individuelle et anonyme auront permis de considérer cette controverse, sans que le comité ne soit dissolu, malgré qu'il ait été ralenti pendant un certain temps en raison de ce litige latent.

Ces entretiens, tenus en dehors des comités, ont donné des indications précieuses aux membres de l'équipe de recherche, certains membres se faisant très explicites quant à la marche à suivre pour assurer l'avenir et l'autonomie du comité ; pour eux, de nouveaux membres ne pourront se joindre à l'initiative que lorsque le comité sera véritablement prêt.

*Actuellement, je dirais non [à la pertinence d'aller chercher d'autres partenaires]. [...] Avant d'aller chercher de nouveaux partenaires, il faut définir les orientations.*

En revanche, d'autres réflexions exprimées en dehors du comité ont fait savoir à l'équipe de recherche qu'il fallait discuter de certains sujets en grand groupe, sinon le fonctionnement du comité risquait de s'en trouver dérangé, voire paralysé. Il en est ainsi de cette proposition concernant la présentation des résultats produits par l'équipe de recherche :

*L'important serait peut-être d'avoir du feedback, ça pourrait être un rapport intérimaire. Je pense qu'il devrait y avoir deux ou trois fois dans l'année des mises au point de faites sur le travail qui a été fait par le comité [...] Techniquement, on pourrait faire un point statutaire à l'ordre du jour : rapport verbal de l'état d'évolution de la recherche.*

Dès le retour, à l'automne, des ententes ont été établies avec le comité à cet égard.

En somme, les entretiens individuels s'adressaient à la dimension sociale de l'innovation qui était en train de se construire, en aidant en quelque sorte les acteurs à traverser les périodes difficiles, en leur cédant la parole et en leur permettant ainsi d'avoir une certaine emprise sur les enjeux se rattachant à la démarche. Mais, en revanche, les discussions survenues en grand groupe ont aussi contribué à la dimension sociale de l'innovation, d'une autre façon cependant. Pour leur part, les réunions étaient indispensables à l'unité et à la solidarité des groupes. Par exemple, les membres pouvaient y découvrir que d'autres partageaient leur point de vue.

En jumelant les résultats obtenus dans le cadre des entretiens individuels aux résultats provenant de l'analyse des autres sources de données constitués des procès-verbaux des réunions de comité, des rapports produits par l'équipe de recherche et de divers documents synthèses produits à différentes occasions pour rendre compte de l'action en cours, nous avons pu conduire l'analyse des processus à un autre niveau. En combinant l'ensemble de ces informations pour chacun des cas étudiés, nous avons pu, en effet, bonifier le modèle promu au départ, en ajoutant à ses composantes des éléments s'attachant plus particulièrement aux enjeux et à l'enchaînement des étapes qui devaient d'abord être franchies avant de *pouvoir mettre en œuvre un plan d'action global et intégré pour la prévention du phénomène des gangs*.

Au final, nous avons tiré de l'expérience conduite dans chacune des localités y participant un *Guide d'action intersectorielle pour la prévention du phénomène des gangs* (Hamel, Cousineau, Vézina et Léveillé, 2007). Ce guide offre quelques repères utiles reliés aux grandes étapes de la traversée, en particulier la formation des comités d'action, la production de l'état de la situation et la problématisation qui en découle, l'élaboration du plan d'action et sa mise en application.

Avec les connaissances émanant de l'évaluation fine que nous avons faite des processus en œuvre tout au long de l'expérience, nous pouvons considérer l'influence des contextes et de la composition des réseaux, traiter de leur évolution au fil des étapes qui ont été franchies, des controverses qui ont été soulevées dans le cours de la démarche, de même que des méthodes qui ont été appliquées pour les résoudre. Nous pouvons même donner une certaine image des formes successives que prenaient les innovations qui se développaient, lesquelles ont d'abord pris la forme de cadres de référence et de philosophies d'intervention, avant de devenir de véritables plans d'action dont découlent aujourd'hui quelques projets novateurs.

### **4.3. L'évaluation de la qualité des plans d'action**

Si nous pouvons prétendre aujourd'hui avoir été témoins du développement de projets novateurs, c'est que nous avons aussi fait l'évaluation de la qualité des plans d'action qui ont été examinés à lumière des critères et des principes directeurs émanant de la sociologie de la traduction (Callon et Latour, 1986), ainsi que des modèles du problème et de l'action que nous avons nous-mêmes développés. Cette analyse estime la qualité des plans d'action, ce qui en soi ne donne qu'une valeur théorique aux actions qui seront mises en œuvre. Elle permet néanmoins de jalonner le processus de construction de l'innovation sociale. Dans le cadre d'une démarche de longue durée où l'actualisation des plans d'action et, en conséquence, l'évaluation de leurs

impacts peuvent tarder à venir, il peut être utile en effet de se donner quelques repères en cours de route, afin de pouvoir juger si, par exemple, la démarche mérite d'être poursuivie ou, encore, si des correctifs doivent être apportés afin de ne pas perdre essentiellement les orientations qui avaient été convenues au départ. En un mot, il s'agit de suivre (évaluer) l'implantation de la démarche afin d'assurer sa réussite.

Les sources documentaires (formes intermédiaires et finalisées des plans d'action, synthèses des travaux des comités, rapports de recherche, procès-verbaux des réunions de comité) sont celles qui ont été le plus utilisées pour conduire cette analyse. Celle-ci s'adresse en quelque sorte à la dimension technique de l'innovation sociale, et s'organise à partir d'une grille renfermant des questions dont quelques-unes sont données en exemple ici.

- Le plan d'action s'appuie-t-il sur une action multistratégique et multisectorielle ?
- Le plan d'action s'appuie-t-il sur une approche globale ?
- Le plan d'action s'appuie-t-il sur les forces des jeunes et des collectivités ?
- Le plan d'action témoigne-t-il d'une certaine transformation du réseau ?

Pour illustration, nous présentons ici certains résultats ressortis de cette analyse pour une des localités en particulier : Montréal-Nord.

En 2003, nous remarquons que le plan d'action de Montréal-Nord s'inscrivait principalement dans une approche de prévention primaire, faisant qu'une première partie du continuum d'intervention était développée de manière importante (activités de sensibilisation, conférences), tandis qu'à l'autre bout du continuum, les actions s'adressant aux jeunes faisant partie de gangs de rue étaient moins développées, malgré que quelques avancées importantes aient été faites dans le cadre de cet axe depuis les débuts de la démarche. Jadis, ce travail était entièrement confié aux policiers. Pour la plupart des intervenants, l'idée de se rapprocher de ces jeunes évoquait le risque de devoir leur consacrer toutes leurs énergies. Mais depuis que le travail de rue était revenu en force dans la localité, les policiers n'étaient plus les seuls à intervenir auprès de cette catégorie de jeunes.

Nous pouvions donc constater que le Service de police, de même que le Centre jeunesse, n'étaient plus, comme avant, les seules institutions à vouloir agir pour résoudre la problématique des gangs. Nous notions que d'autres acteurs, des intervenants du milieu communautaire, du milieu scolaire et même des membres de l'administration municipale, s'ajoutaient peu à peu. Il y avait là une amélioration importante, une source vitale pour le comité d'action de la localité qui, depuis longtemps, déplorait l'absence des deux

dernières institutions. De nouveaux partenaires s'alliaient donc aux membres du comité pour les aider à mettre en œuvre un plan d'action qui prenait de plus en plus forme et demandait de s'allier de nouveaux effectifs.

À ce moment, nous soulignons que les jeunes devaient constituer la plus importante des cibles de l'intervention, sans que la communauté ne soit évacuée pour autant. Celle-ci était mobilisée par la promotion du Cadre de référence qu'avait développé le comité d'action locale, visant à établir et, dans une certaine mesure, à rétablir les faits sur la présence et le phénomène des gangs dans le quartier. On voulait ainsi modifier les perceptions des résidants et de plusieurs intervenants qui, avec le temps, avaient développé l'habitude de se protéger des gangs, de leurs membres et même des jeunes en général.

La philosophie émanant du Cadre de référence qui était promu par les membres du comité d'action locale mettait en valeur la force, le talent et la vitalité des jeunes, et déjà quelques projets inclus dans le plan d'action adoptaient cette approche. La promotion du cadre de référence constituait en fait le premier axe du plan d'action. Le deuxième axe, quant à lui, visait à rétablir les faits concernant la présence et l'action des gangs de rue dans la localité, ternissant son image et imprégnant les perceptions des membres de la communauté quant à leur sécurité et quant à la dangerosité des jeunes. Le troisième axe du plan d'action se vouait pour sa part essentiellement à la prévention par la mise en place et la consolidation de projets répondant aux besoins des jeunes.

Cette dernière facette du plan d'action s'est ensuite beaucoup développée, par l'entremise de trois projets intitulés l'Animation du milieu urbain, l'Aventure urbaine et un Itinéraire pour tous. En 2003, ce sont les forces de la communauté qui ont été mobilisées en premier lieu pour démarrer ces projets, dont certains demandaient la participation de plusieurs organismes. Nous pensons notamment à l'Animation du milieu urbain, auquel se greffent des dizaines d'organismes. Ce projet veut répondre au fait que de plus en plus de jeunes du quartier fréquentent désormais la rue, les parcs et autres endroits publics, n'ayant nulle part ailleurs où aller pour socialiser et se divertir, deux éléments essentiels à leur développement. Pour éviter que la situation ne se détériore davantage, et parce que celle-ci est propice à la formation de gangs, il est apparu capital pour la communauté que des intervenants animent ces lieux, y assurent une présence, recréent d'une certaine façon ce que les jeunes peuvent retrouver dans les maisons de jeunes (activités sportives, culturelles et ludiques), mais à ciel ouvert. Pour son fonctionnement, le projet Animation du milieu urbain s'appuie sur la mobilisation de l'ensemble des ressources du milieu pour aller vers les jeunes, là où ils se trouvent, et pour qu'ils disposent gratuitement d'un maximum de lieux et d'activités en dehors des heures de classe.

En 2004, le projet mettait en scène les intervenants de trois maisons de jeunes, du service de police, de la municipalité, de deux écoles secondaires, de deux centres pour jeunes raccrocheurs et de deux entreprises d'économie sociale qui, ensemble, ont réussi à rejoindre 7 922 jeunes qui ont participé à des activités régulières et à d'autres plus ponctuelles (événements spéciaux). En 2005, 9 631 jeunes se sont inscrits à des activités régulières et 1 404 à des activités ponctuelles (Hamel, Cousineau et Vézina, 2006). En 2006, le projet était mis en candidature pour recevoir le prix d'excellence du Réseau québécois de Villes et Villages en Santé, qui se consacre aux projets de développement durable. Cela signifie que le projet initial est resté fidèle aux principes de son modèle de départ et que le réseau qui le porte se consolide et s'allonge constamment. À cela s'ajoute également le fait que les porteurs du projet de l'Animation du milieu urbain sont aujourd'hui responsables du développement d'un cursus de formation provinciale destinée à tous les acteurs pouvant être appelés à faire face au phénomène des gangs.

Ainsi, les développements dont nous avons été témoins dans cette localité répondent bien, à notre avis, à la définition d'une innovation sociale et se rapprochent également du modèle théorique du problème et de l'action promu au début de cette démarche.

Nous pouvons conclure, à l'issue de notre analyse s'inscrivant dans un processus d'évaluation participative, qu'il est possible de bâtir des plans d'action servant à prévenir le phénomène des gangs en s'inspirant des principes du développement social.

## CONCLUSION

Si Jeunesse et gangs de rue n'est pas encore rendu à l'étape de faire la démonstration de ses effets, l'évaluation que nous en avons faite montre à tout le moins que le projet a certainement franchi des étapes importantes. Cependant, notre attention n'était pas portée sur un « programme » en particulier ni sur ses effets possibles. Elle se concentrait plutôt sur les processus conduisant à son développement. Cette approche fut significative pour les acteurs participant à la démarche, si bien qu'ils se sont appuyés sur les résultats qui émanaient de l'analyse pour dénouer des enjeux ou encore pour franchir des passages plus difficiles. Au final, certains quartiers, Montréal-Nord notamment, ont réussi à mettre en place les conditions nécessaires à l'établissement de changements durables au sein de leur communauté. Ces changements s'adressent entre autres choses à la vision des intervenants sur le pouvoir qu'ils ont face au phénomène des gangs et s'installent également au sein de nouveaux systèmes d'action, incluant un plus grand nombre d'acteurs. En soi, ce résultat n'est pas négligeable. Nous croyons bien entendu

qu'il est directement relié à la très grande compétence des gens des milieux avec lesquels nous avons travaillé. Mais nous croyons aussi que ce résultat est relié au fait que l'approche que nous avons choisie dans le cadre de cette évaluation était appropriée au contexte, de même qu'au stade de développement de l'objet d'étude (programme, innovation, modèle d'intervention) qui était en train de se mettre en forme. Cet aspect est crucial à notre avis dans le créneau de l'évaluation.

## BIBLIOGRAPHIE

- ALTER, N. (2002). *Les logiques de l'innovation*, Paris, La Découverte.
- BARBEAU, D. (2005). *Actualité Gang De Rue: Bilan des projets du SPVM (2005)*, Montréal, Publication du Service de police de la Ville de Montréal.
- BILODEAU, A. (2000). *Les conditions de qualité de la planification participative et de l'action en partenariat dans le champ de la prévention – promotion de la santé et du bien-être, le cas des priorités régionales 1995-1998 de Montréal-Centre*, Thèse de doctorat, Université de Montréal.
- BLEDSON, K.L. et J.A. GRAHAM (2005). «The use of multiple evaluation approaches in program evaluation», *American Journal of Evaluation*, vol. 26, n° 3, p. 302-319.
- BOUTIN, G. (2006). *L'entretien qualitatif*, Québec, Presses de l'Université du Québec.
- CALLON, M. (1986). «Éléments pour une sociologie de la traduction: la domestication des coquilles Saint-Jacques et des marins-pêcheurs dans la baie de Saint-Brieuc», *L'Année sociologique*, n° 36, p. 169-208.
- CALLON, M., P. LASCUMES et Y. BARTHE (2001). *Agir dans un monde incertain. Essai sur la démocratie technique*, Paris, Seuil.
- CALLON, M. et B. LATOUR (1986). «Les paradoxes de la modernité: comment concevoir les innovations?», *Prospective et santé*, n° 36, p. 13-25.
- CHAMPAGNE, F., A.P. CONTANDRIOPOULOS et J.L. DENIS (1990). *La recherche évaluative en santé*, Montréal, Groupe de recherche interdisciplinaire en santé de l'Université de Montréal.
- CHEN, H-T. (1990). *Theory-driven Evaluations*, Beverly Hills, Sage Publications.
- CHEN, H-T. et P.H. ROSSI (1980). «The multi-goal, theory-driven approach to evaluation. A model linking basic and applied social science», *Social Forces*, vol. 59, n° 1, p. 106-122.
- CHEN, H-T. et P.H. ROSSI (1989). «Issues in the theory-driven perspective», *Evaluation and Program Planning*, vol. 12, p. 299-306.
- CHETTLEBURGH, M.C. (2003). *Résultats de l'enquête policière canadienne sur les gangs de jeune de 2002*, Ottawa, Rapport présenté au Solliciteur général du Canada, Atwood Strategy Corporation.

- CHRISTIE, C.A. et M.C. ALKIN (2003). «The user-oriented evaluator's role in formulating a program theory: Using a theory-driven approach», *American Journal of Evaluation*, vol. 24, n° 3, p. 373-385.
- COOLBAUGH, K. et C.J. HANSEL (2000). *The Comprehensive Strategy: Lessons Learned From Pilot Sites*, Washington, Department of Justice, Office of Justice Programs, Office of Juvenile Justice and Delinquency Prevention.
- COUSINEAU, M.-M., S. HAMEL et A. DESMARAIS (2004). *Tournée provinciale pour l'établissement d'un portrait du phénomène des gangs au Québec. Huit synthèses régionales et une synthèse globale*, Montréal, Comité organisateur du Colloque pour la création d'un réseau québécois d'échange sur les jeunes et les gangs de rue.
- DE BRUYNE, P., AVEC LA PARTICIPATION DE J. HERMAN et H. SCHOUTHEETE (1974). *Dynamique de la recherche en sciences sociales: les pôles de la pratique méthodologique*, Paris, Presses universitaires de France.
- DOLBEC, A. (2003). «La recherche-action», dans B. Gauthier (dir.), *Recherche sociale, de la problématique à la collecte de données*, Québec, Presses de l'Université du Québec, p. 505-540.
- EDQUIST, C. et L. HOMMEN (1999). «Systems of innovation», *Technology in Society*, vol. 21, n° 1, p. 63-79.
- FETTERMAN, D.M. (1994). «Empowerment evaluation», *Evaluation Practice*, vol. 15, n° 1, p. 1-15.
- FETTERMAN, D. et A. WANDERSMAN (2007). «Empowerment evaluation. Yesterday, today, and tomorrow», *American Journal of Evaluation*, vol. 28, n° 2, p. 179-198.
- FINE, M. et al. (2003). «Participatory action research: From within and beyond prison bars», dans P.M. Camic, J.E. Rhodes et L. Yardley (dir.), *Qualitative Research in Psychology, Expanding Perspectives in Methodology and Design*, Washington, American Psychology Association, p. 173-198.
- GAUTHIER J. (2006). «L'apprentissage: source d'innovation dans les politiques d'aide sociale?», dans Y. Comeau et L. Thuy-Diep (dir.), *Les innovations sociales: émergence, processus de construction et diffusion*, Actes du 7<sup>e</sup> colloque des étudiants-es du CRISES, Montréal, Cahiers du CRISES – Collection Hors-Série, p. 92-98.
- GRAWITZ, M. (1993). *Méthodes des sciences sociales*, Paris, Dalloz.
- HAMEL S. et al. (2007). *Guide d'action intersectorielle pour la prévention du phénomène des gangs*, Montréal, Institut de recherche pour le développement social des jeunes, pour le Centre national pour la prévention du crime.
- HAMEL S., M.-M. COUSINEAU et M. VÉZINA (2006). *Jeunesse et gangs de rue: évaluation d'un projet d'action intersectorielle pour la prévention du phénomène des gangs*, Montréal, Institut de recherche pour le développement social des jeunes, pour le Centre national pour la prévention du crime.
- HAMEL, S. et al. (1998). *Jeunesse et gangs de rue, Phase II. Résultats de la recherche-terrain et proposition d'un plan stratégique quinquennal*, Montréal, Institut de recherche pour le développement social des jeunes.

- HARRISSON D. et M. VÉZINA (2006). «L'innovation sociale: une introduction», *Annals of Public and Cooperative Economics*, vol. 77, n° 2, p. 129-138.
- HÉBERT, J., S. HAMEL et G.J. SAVOIE (1997). *Jeunesse et gangs de rue. Phase I. Revue de littérature* – Rapport présenté au Service de police de la communauté urbaine de Montréal, Montréal, Institut de recherche pour le développement social des jeunes.
- HOWELL, J.C. (1998). *Youth Gangs: An Overview*, Washington, Department of Justice, Office of Justice Programs, Office of Juvenile Justice and Delinquency Prevention.
- HOWELL, J.C. (2000). *Youth Gang Program and Strategies*. Washington, Department of Justice, Office of Justice Programs, Office of Juvenile Justice and Delinquency Prevention.
- HUGHES, M. et T. TRAYNOR (2000). «Reconciling process and outcome in evaluating community initiatives», *Evaluation*, vol. 6, n° 1, p. 37-49.
- JASON, L.A. et al. (dir.) (2002). *Participatory Community Research. Theories and Methods in Action*, Washington, American Psychological Association.
- KLEIN, J-L. et D. HARRISSON (dir.) (2007). *L'innovation sociale. Émergence et effets sur la transformation des sociétés*, Québec, Presses de l'Université du Québec.
- KLEIN, M.W. (1995). *The American Street Gang: It's Nature, Prevalence, and Control*, New York, Oxford University Press.
- LESSARD-HÉBERT, M., G. GOYETTE et G. BOUTIN (1996). *La recherche qualitative: fondements et pratiques*, Montréal, Éditions Nouvelles.
- LÉVESQUES, B. (2005). *Innovations et transformations sociales dans le développement économique et le développement social*, Montréal, CRISES, Études théoriques, 0507.
- NADEAU J. (2006). «Le partenariat: une innovation pour le développement des communautés?», dans Y. Comeau et L. Thuy-Diep (dir.), *Les innovations sociales: émergence, processus de construction et diffusion*, Actes du 7<sup>e</sup> colloque des étudiants-es du CRISES, Montréal, Cahiers du CRISES – Collection Hors-Série, p. 85-91,
- O'SULLIVAN, R.G. (2004). *Practicing Evaluation. A Collaborative Approach*, Beverly Hills, Sage Publications.
- PATTON, M.Q. (1987). *How To Use Qualitative Methods in Evaluation*, Newbury Park, Sage Publications.
- PATTON, M.Q. (1998). «Discovering process use», *Evaluation*, vol. 4, n° 2, p. 225-233.
- POISSON, Y. (1991). *La recherche qualitative en éducation*, Québec, Presses de l'Université du Québec.
- POUPART, J. et al. (1997). *La recherche qualitative: enjeux épistémologiques et méthodologiques*, Montréal, Gaëtan Morin.
- REICHARDT, C.S. (1994). «Summative evaluation, formative evaluation, and tactical research», *Evaluation Practice*, vol. 15, n° 3, p. 275-281.

- SERVICE CANADIEN DE RENSEIGNEMENTS CRIMINELS (2006). *Rapport annuel sur le crime organisé au Canada*, Ottawa, Document de consultation du gouvernement du Canada.
- SERVICE DE POLICE DE LA VILLE DE MONTRÉAL (2005). *Plan d'action corporatif 2005 du SPVM*. 3 décembre 2004, Montréal, Document interne.
- SHORT J.F. et L.A. HUGUES (2006). *Studying Youth Gngs*, Toronto, AltaMira Press.
- SPERGEL, I.A. (1995). *The Youth Gang Problem: A Community Approach*, New York, Oxford University Press.
- SPERGEL I.A. et S.F. GROSSMAN (1998). «The little village project: A community approach to the gang problem», dans P.L. Ewalt, E.M. Freeman et D.L. Poole (dir.), *Community Building*, Washington, National Association of Social Workers, p. 130-148.
- SPERGEL, I.A., K.M. WA et R.V. SOSA (2006). «The comprehensive, community-wide gang problem program model: Success and failure», dans J.F. Short et L.A. Hugues (dir.), *Studying Youth Gangs*, Toronto, AltaMira Press, p. 203-224
- STUFFLEBEAM, B.M. (2001). «Evaluations models», *New Directions for Evaluation*, n° 89, p. 7-98.
- WEISS, C.H. (1997). «How cans theory-based evaluation make greater headway?», *Evaluation Review*, n° 21, p. 501-524.
- YIN, R.K. (1989). *Case Study Research: Design and Methods*, 2<sup>e</sup> éd., Newbury Park, Sage Publications.
- YIN, R.K. (1993). *Applications of Case Study Research*, Newbury Park, Sage Publications.

# CHAPITRE 4

---

## **L'évaluation des besoins** Un outil pour la planification

Line Massé  
Département de psychoéducation  
Université du Québec à Trois-Rivières

La raison d'être première de mettre en place ou de maintenir un programme psychosocial est la naissance ou la présence d'un problème social. S'il n'y a pas de problèmes significatifs ou de besoins perçus pour une intervention, il n'y a généralement pas de base pour affirmer la nécessité de mettre en place un programme. Ainsi, une forme importante d'évaluation consiste à évaluer la nature, l'importance et la distribution d'un problème psychosocial ainsi que ses répercussions, dans quelle mesure il y a besoin d'intervenir sur cette situation, la population qui doit être visée et les implications de ces différentes circonstances pour la conception et l'implantation de l'intervention. Cet état des lieux réfère communément à l'évaluation des besoins<sup>1</sup>. Cette évaluation s'effectue habituellement avant de procéder à la mise sur pied d'un service, d'un programme ou d'une intervention. Elle peut également être utile pendant son déroulement pour adapter l'intervention aux besoins changeants de la population. Dans ce chapitre, nous présentons les objectifs de cette démarche ainsi que les principaux concepts qui y sont reliés. Nous exposons par la suite les étapes qui sont généralement suivies pour sa réalisation ainsi que les principales techniques utilisées pour la collecte de données. Pour terminer, nous donnons un exemple d'évaluation des besoins qui illustre certains des éléments présentés dans ce chapitre.

## 1. LES OBJECTIFS DE L'ÉVALUATION DES BESOINS

L'évaluation de besoin a pour but de fournir aux intervenants, aux gestionnaires ou aux politiciens des informations valides et fiables sur les principaux problèmes de la communauté et sur ses besoins en matière de services ou d'intervention. Ce processus peut servir à circonscrire les principaux problèmes et besoins d'une population ou d'un groupe cible afin de mieux orienter l'intervention qui sera faite auprès de cette population par la suite. Les problèmes et les besoins à documenter dépendront naturellement du domaine d'intervention privilégié : promotion de la santé, prévention des difficultés psychosociales auprès des populations à risque ou intervention auprès des personnes présentant des difficultés d'adaptation psychosociale. L'identification des problèmes porte sur l'ampleur des problèmes existants, sur leurs causes et leurs conséquences. On pourra aussi s'intéresser aux facteurs qui mettent la population à risque de développer des problèmes ou au contraire à ceux qui en diminuent le risque. Cette exploration fournit

---

1. On utilise aussi les termes « étude » ou « analyse » de besoins ou « évaluation de la pertinence ». Le terme « étude de faisabilité » renvoie à une démarche similaire mais, dans ce cas, on s'intéresse seulement à vérifier si une intervention ou un programme est réalisable dans les conditions techniques, financières, calendaires, politiques ou sociales actuelles ou à venir.

également un inventaire des ressources existantes dans la communauté et l'état de leur utilisation et, à partir de ce portrait, permet de dresser le tableau des besoins qui sont comblés et de ceux qui sont non satisfaits.

L'évaluation des besoins est souvent considérée comme la première étape dans le développement et la planification d'un nouveau programme afin de fournir des informations sur les services requis et la façon dont ces services devraient être offerts à la population. Si on ne comprend pas bien la nature et l'ampleur des problèmes de la population visée et les besoins qui en découlent, les objectifs du programme ne peuvent pas être bien spécifiés, ni ses modalités bien définies. En ce qui concerne les programmes ou les services existants et bien établis, l'évaluation des besoins peut être fort utile afin de vérifier si des ajustements sont nécessaires, si le programme est encore pertinent compte tenu des autres services offerts ou des besoins changeants de la population, ou encore pour identifier si d'autres groupes pourraient en bénéficier. À cet égard, on pourra examiner s'il y a un écart ou non entre les besoins actuels des participants ciblés et les activités ou interventions en cours. On pourra également considérer les points suivants : 1) la disponibilité des services ou des programmes, 2) leur accessibilité, 3) la connaissance de ces derniers par la population, 4) le niveau d'intégration ou de continuité des services ou des programmes, 5) leurs particularités et leur complémentarité avec les autres services ou programmes, et 6) le niveau des ressources disponibles et leur distribution.

Les études de besoins ne sont néanmoins pas toujours faites en référence à un programme particulier. Elles peuvent être un outil de planification et de décision fort utile pour les décideurs et ceux qui élaborent les politiques sociales afin d'établir des priorités par rapport aux demandes de la population et aux nombreux problèmes présentés. De même, pour les organismes et leurs gestionnaires, ces études peuvent orienter la distribution des ressources entre les différentes unités d'un service.

## 2. LES PRINCIPAUX CONCEPTS LIÉS AUX BESOINS

L'évaluation des besoins s'intéresse tant aux problèmes sociaux d'une communauté qu'aux besoins qui en découlent. Quoique apparentés, on distingue habituellement les concepts de problèmes et de besoins. Selon Desrosiers *et al.* (1998), « la notion de problème renvoie à l'existence dans la communauté, d'un état ou d'un phénomène jugé indésirable » (p. 26). On reconnaît habituellement que les problèmes sociaux ne sont pas des phénomènes objectifs, mais plutôt des phénomènes subjectifs basés sur des jugements de valeur ; ils sont des constructions sociales et politiques qui émergent des intérêts des parties impliquées et des conditions qu'elles

observent (Rossi *et al.*, 2004, p. 125). Ce qui explique que la définition des problèmes varie grandement selon le contexte sociopolitique dans lequel on se trouve. Par exemple, les problèmes de pauvreté sont définis différemment en Amérique du Nord qu'en Afrique. Pour bien définir les problèmes, deux autres concepts seront utiles, soit l'incidence et la prévalence. L'incidence réfère au nombre de nouveaux cas d'un problème particulier qui sont identifiés dans une région géographique pour une période particulière de temps (p. ex., le nombre de personnes qui ont été diagnostiquées comme ayant une dépression au cours de l'année). La prévalence correspond au nombre de cas existants dans une région géographique particulière à un moment précis (p. ex., le nombre de personnes qui souffrent actuellement d'une dépression au Québec).

Quant au concept de besoin, il se définit comme un écart entre une situation existante et une situation jugée idéale, normale, minimale ou souhaitable par une population dans un contexte social donné (Posavac et Carey, 2007, p. 114). Même si les besoins d'une population découlent des problèmes qu'elle rencontre, l'existence d'un problème ne veut pas nécessairement dire qu'il y a un besoin insatisfait, notamment s'il existe des ressources disponibles, accessibles et appropriées pour résoudre ce problème dans la communauté. Dans le même ordre d'idées, il peut y avoir besoin d'un service ou d'un programme dans une communauté sans qu'un problème particulier soit actuellement présent, comme c'est le cas, par exemple, d'une population à risque qui nécessite une intervention en matière de promotion de la santé (Desrosiers *et al.*, 1998, p. 26).

### 3. LES APPROCHES D'ANALYSE DES BESOINS

On distingue habituellement trois façons d'aborder l'analyse des besoins (Kaufman, 1972; Mayer et Ouellet, 1991; Nadeau, 1988): l'approche déductive, l'approche inductive et l'approche mixte ou la triangulation. L'approche déductive consiste à identifier les besoins d'une population à partir d'une recension des recherches menées sur le sujet ou de l'expérience sur le terrain afin de dresser une première liste de besoins, qui sera par la suite validée auprès de la population cible. Cette approche est souvent utilisée lorsque la situation problématique est bien documentée. L'approche inductive, quant à elle, «s'appuie sur le postulat que la population dans le besoin est la mieux placée pour définir ses besoins et la satisfaction de ces derniers, c'est-à-dire la situation désirée» (Mayer et Ouellet, 1991, p. 72). Selon cette approche, c'est la population qui définit les problèmes qu'elle rencontre et qui établit les priorités à propos des besoins à combler. Cette approche est fréquemment utilisée lors que les problèmes sont sensibles dans la population, lorsque les phénomènes à l'étude sont très peu documentés ou lorsque l'on veut que

la population soit partie prenante des décisions qui seront prises. La troisième approche, dite mixte, consiste à allier des techniques propres aux deux approches et à trianguler les informations issues des différentes sources de données afin de vérifier les convergences et les divergences. L'approche mixte est celle qui permet la détermination optimale des besoins, mais elle n'est pas toujours possible en raison du temps et des ressources disponibles.

## 4. LES ÉTAPES DE L'ANALYSE DE BESOINS

L'évaluation des besoins peut s'avérer une activité complexe et coûteuse qui requiert beaucoup de temps et d'énergie. Souvent, le personnel de l'organisme promoteur du programme ou les bailleurs de fonds croient qu'ils ont une connaissance suffisante des problèmes sociaux ou des besoins de la population ciblée et que cette opération est une perte de temps. Il faudra faire attention de ne pas sauter trop vite à cette conclusion, car comme le notent Rossi *et al.* (2004, p. 122), il est très facile pour un programme d'être basé sur des postulats erronés, soit en raison d'un diagnostic incomplet des problèmes, d'un changement dans les problèmes rencontrés ou dans la population cible depuis que le programme a été mis en place, ou encore en raison d'informations biaisées ou de stéréotypes qui mènent à des visions déformées de la réalité. Cela ne veut pas dire qu'il sera toujours nécessaire de mettre en branle un processus complexe de collecte de données. Dans certains cas, le processus suivi pourrait être limité. C'est le cas lorsque l'organisme a déjà une bonne idée des besoins prioritaires à combler et de la façon de le faire, ou lorsqu'il y a beaucoup d'informations disponibles sur le sujet. Dans d'autres cas, le budget et les ressources disponibles, de même que le temps alloué au projet, limiteront l'envergure que pourra prendre l'étude. Si on décide d'entreprendre une collecte de données, Desrosiers *et al.* (1998, p. 76) recommandent de constituer un groupe responsable de l'étude dont les membres représentent les différentes fonctions de soutien à un tel projet (planification, recherche et gestion), de même que les intérêts des divers groupes (usagers, population, etc.). Beadel de Palomo et Luna (2000) suggèrent de confier les responsabilités suivantes à ce comité : déterminer l'envergure de la collecte et les résultats désirés ; approuver le design de l'étude ; surveiller le processus ; choisir les consultants ; s'assurer que toutes les populations affectées ou concernées sont incluses ; faciliter l'organisation du processus, en particulier le dépistage des participants ; participer à l'interprétation des résultats et à l'élaboration des recommandations.

Quel que soit le contexte de l'étude de besoin, le processus suivi est sensiblement le même et il est similaire à celui des autres formes d'évaluation. Dans ce qui suit, nous décrivons brièvement chacune des étapes généralement suivies (Nardi, 2003ab ; Posavac et Carey, 2007 ; Royse *et al.*, 2006 ;

Rossi, Freeman et Lipsey, 2004; Witkin et Altschuld, 1995). L'encadré 1 complète ces informations en présentant différentes questions qui pourront guider l'évaluateur tout au long de ce processus.

#### **4.1. Déterminer les fins de l'étude de besoins et les utilisateurs potentiels**

Les utilisateurs de l'étude de besoins sont les personnes qui vont agir sur la base des résultats recueillis (personnel du programme, administrateurs, bailleurs de fonds, politiciens, etc.) et les citoyens qui vont être affectés par les décisions qui seront prises (clients passés, actuels ou potentiels; autres membres de la communauté, autres organismes touchés, syndicats, entreprises, etc.). L'implication des deux groupes tout au long du processus va habituellement faciliter l'analyse des résultats et l'implantation des recommandations qui seront faites par la suite. Connaître les utilisateurs potentiels et les fins de l'étude aide l'évaluateur à mieux cerner les problèmes et les solutions à considérer, et peut aussi limiter l'étendue des problèmes touchés.

#### **4.2. Définir la population à l'étude**

Il s'agit ici de circonscrire le groupe sur lequel portera l'analyse. L'étude peut porter sur l'ensemble d'une population, sur une partie de celle-ci ou sur une clientèle plus précise considérée à risque ou en besoin. Une clientèle à risque est définie comme un ensemble d'individus ou d'unités dans une région particulière qui présentent des caractéristiques ou des facteurs qui augmentent la probabilité de présenter ou de développer des problèmes d'adaptation psychosociale. Une clientèle en besoin réfère à des individus ou à des unités dans une région particulière qui manifestent actuellement une condition problématique particulière. Il faudra aussi déterminer le niveau de représentativité de la population visée. S'intéresse-t-on à la population locale, régionale, provinciale ou nationale? Désire-t-on comparer avec un échantillon normatif ou avec une autre population (d'une autre région par exemple)?

#### **4.3. Déterminer les informations particulières à obtenir**

À partir des décisions prises aux deux premières étapes, il faudra déterminer les informations qui seront utiles pour orienter les prises de décisions. Ces informations pourront porter tant sur la population visée, sur les problèmes qu'elle présente, sur les solutions possibles, sur les ressources disponibles pour répondre à ces besoins que sur l'utilisation de ces services.

On distingue habituellement deux types d'informations ou de données : primaires et secondaires. Les données primaires sont des données originales que l'évaluateur collecte et analyse, par enquête ou entrevues de groupe par exemple. Les données secondaires sont des données que quelqu'un d'autre a collectées, mais qui peuvent être utilisées pour répondre aux questions de l'étude. Dans certains cas, l'évaluateur pourra même réanalyser ces données pour les fins précises de son étude, par exemple dans le cas des banques de données de Statistique Canada ou de grandes enquêtes nationales ou internationales.

#### 4.4. Élaborer un protocole d'évaluation

Afin de ne pas réinventer la roue, un travail exploratoire sera nécessaire afin de déterminer si les informations désirées existent déjà ou s'il faut mettre en place des moyens spéciaux pour les obtenir. Avant de déployer toute une batterie de moyens pour recueillir les besoins, il est souvent préférable de commencer par une analyse secondaire des données disponibles, en faisant une revue des écrits ou en utilisant l'approche des indicateurs sociaux (voir plus loin au point 5 – Les techniques de collecte de données). On peut aussi procéder à l'inventaire des ressources, des programmes et des services existants pour résoudre la problématique identifiée, tout autant pour s'assurer qu'il n'y a pas de dédoublement que pour vérifier les ressources qui sont manquantes ou les populations qui ne sont pas desservies. Ces premières démarches permettront de mieux cibler les informations manquantes. Le choix des méthodes à utiliser dépendra de différents critères, que l'on pense aux objectifs, au type d'information recherchée, au temps alloué, aux ressources et budget disponibles ou toute autre contrainte de la situation (son caractère d'urgence, par exemple). Plusieurs techniques peuvent être utilisées (voir au point 5) et il est souvent préférable de les utiliser en combinaison, car certaines se complètent bien. Par exemple, les groupes de discussion peuvent illustrer, valider et enrichir les données fournies par les informateurs clés sur un problème ou une population. En fonction des méthodes choisies, il faudra localiser les sources d'informations nécessaires. Dans la mesure du possible, il sera préférable de recourir à plusieurs sources d'informations afin de garantir la validité de la démarche. La consultation d'experts, de chercheurs, de praticiens, de gestionnaires, de politiciens, de groupes d'intérêts ou de membres actifs de la communauté pourra guider l'identification des problèmes les plus saillants et des besoins de la clientèle, la détermination des ressources manquantes, l'établissement des priorités d'action et la planification de l'intervention ou des services. Leurs opinions pourront être recueillies lors d'entrevues individuelles (approche par informateurs clés), lors d'entrevue de groupe (groupes de discussion, groupe nominal) ou lors de travail en comités (technique Delphi, comité d'évaluation).

Selon les techniques choisies et la complexité des démarches qui seront entreprises, il pourra être nécessaire d'élaborer un protocole d'évaluation qui précisera la méthode privilégiée. Ce protocole comprend habituellement une description des objectifs de l'étude, des variables étudiées, des participants ou de l'échantillon, des instruments utilisés, de la procédure suivie pour colliger les informations et l'échéancier de travail.

## 4.5. Recueillir l'information, analyser et interpréter les données

Les analyses choisies dépendront de la nature des informations recueillies. Habituellement, les informations recueillies par utilisation de questionnaires ou d'entrevues structurées appellent un traitement quantitatif, alors que les informations obtenues dans les écrits ou recueillies à l'aide d'entrevues individuelles semi-structurées ou d'entrevues de groupe requièrent une analyse qualitative<sup>2</sup>. L'analyse des résultats doit se faire avec la collaboration de différents utilisateurs de l'étude de besoins, tant pour s'assurer de la justesse des interprétations, que pour orienter éventuellement les analyses à faire. En effet, par leurs questions, les utilisateurs peuvent contribuer à l'identification d'analyses auxquelles les évaluateurs n'avaient pas pensé et dont les résultats pourront leur être fort utiles. Cette étape exige l'intégration des informations sur la population cible, sur les problèmes et les besoins reconnus, sur les solutions possibles pour répondre à ces besoins, sur l'inventaire des ressources et des services existants ainsi que les différents points de vue qui ont été émis à ce sujet. L'analyse des résultats doit faire ressortir les problèmes les plus saillants ainsi que ceux qui peuvent se prêter plus facilement à une intervention. Elle doit permettre d'identifier les solutions possibles en tenant compte de l'environnement actuel de services. L'analyse doit aussi préciser les groupes cibles du programme ou de l'intervention. Enfin, tout ce processus mènera à la détermination des priorités d'intervention et à la formulation des recommandations. Dans le cas d'un nouveau programme, le processus pourra mener à la formulation de ses objectifs. Pour déterminer les priorités, l'évaluateur pourra avoir recours à des méthodes de prise de décision de groupe ou à des approches de consensus telle la technique du groupe nominal ou la technique Delphi (voir plus loin aux points 5.5.1 et 5.5.2). Morneau (1992, cité dans Desrosiers *et al.*, 1998) identifie différents critères pour déterminer les priorités parmi les besoins identifiés: «l'ur-

---

2. Pour plus de détails sur les différentes méthodes d'analyses quantitatives pouvant être utilisées, le lecteur pourra se référer au chapitre 9 de ce livre ou à l'ouvrage de Howel (2008) portant sur la question. Pour ce qui est des analyses qualitatives, le chapitre 10 aborde sommairement cette question. Pour aller plus loin, le lecteur pourra consulter les ouvrages de Miles et Huberman (2003) ou Paillé et Mucchielli (2003).

gence de la situation, les conséquences pour les personnes et leur fonctionnement, les effets sur l'environnement immédiat des personnes (soutien), le volume de la population concernée, la faisabilité (solutions possibles), la disponibilité des ressources, la concordance avec la mission et le mandat de l'organisation» (p. 76).

#### **4.6. Préparer le rapport**

Habituellement, le rapport doit présenter les fins de l'étude, la méthode utilisée ainsi que ses limites, une synthèse des problèmes identifiés et des solutions possibles, l'inventaire des ressources et des services disponibles ainsi que les recommandations de l'évaluation quant aux besoins prioritaires à combler et aux actions qui devront être prises en ce sens. Il y a plusieurs façons de présenter l'information. Beadle de Palomo et Luna (2003) font les suggestions suivantes pour améliorer la présentation des résultats : considérer les utilisations qui seront faites des résultats, présenter les résultats par population cible ou sujet et non selon les sources d'information ou les méthodes utilisées, choisir un format clair et facile à comprendre, utiliser une présentation qui facilite la prise de décision (par exemple, des graphiques). Dans le cas de problèmes particulièrement controversés dans la population ou de situations où d'autres personnes ou organismes pourraient être affectés par les résultats de l'étude, il faudra faire attention à la façon de présenter les informations afin de ne pas faire de bévues diplomatiques ou politiques. Pour éviter ces bévues, mais surtout pour s'assurer que le rapport fait une bonne synthèse des informations recueillies et que les recommandations sont pertinentes, il est recommandé de communiquer les résultats préliminaires aux parties prenantes pour obtenir leur rétroaction (Royse *et al.*, 2006, p. 62).

#### **4.7. Rédiger le rapport final et diffuser plus largement les résultats de l'étude**

Une fois le rapport final rédigé, il faudra le diffuser de façon adéquate aux décideurs, aux utilisateurs éventuels et à tout autre auditoire susceptible d'être affecté par les résultats. Pour s'assurer d'une bonne diffusion, il faudra aussi prévoir des communications orales (conférence publique, conférence dans des congrès, etc.) ou l'utilisation des nouvelles technologies (site Internet, courriel, etc.).

Encadré I

## Questions pouvant guider le processus d'évaluation des besoins

---

### Fins de l'étude

- Est-il nécessaire de documenter davantage la nature et l'ampleur des problèmes de la communauté ou d'un groupe cible particulier?
- L'étude doit-elle aider les décideurs à établir les besoins prioritaires à combler afin d'orienter l'offre de services?
- L'étude a-t-elle pour but de réaménager les ressources ou de modifier un programme existant afin de mieux répondre aux besoins de la population?
- L'information recueillie doit-elle justifier la pertinence d'un nouveau programme?
  - Quelle est la nécessité d'agir sur la situation problématique?
  - Doit-on déterminer s'il y a suffisamment de clients potentiels avec un problème particulier pour justifier la création d'un nouveau programme?
  - Doit-on déterminer s'il y a d'autres interventions ou ressources dans la communauté pour répondre aux besoins identifiés?
- L'information recueillie doit-elle permettre d'orienter l'élaboration d'un nouveau programme afin que celui-ci soit taillé aux besoins du groupe cible?

### Utilisateurs potentiels

- Quels sont les principaux intéressés par les résultats de l'étude? Les clients ou les clients potentiels? Le personnel du service, du programme ou de l'organisme? Les décideurs (fonctionnaires, gestionnaires, politiciens, responsables)?
- Est-ce que les résultats de l'étude de besoins seront vraiment utilisés ou serviront-ils seulement à des fins politiques ou bureaucratiques?
- Est-ce que les résultats de l'étude serviront à d'autres membres de la communauté ou à d'autres organismes?

### Contraintes de l'étude de besoin

- Quel est le budget disponible?
- Quel est le temps alloué au projet? Est-ce une situation d'urgence sociale?
- Quel est le cadre social, juridique, politique et économique dans lequel se situe le projet? La question est-elle controversée dans la population?
- Quelles sont les ressources matérielles (ordinateurs, enregistreur numérique, centre d'appels, logiciels de traitement de données, salles d'entrevue, etc.) et humaines disponibles?

### Population ciblée

- S'intéresse-t-on aux besoins de la population générale, d'une population à risque ou d'une population en besoin?
- L'étude doit-elle s'effectuer au plan local, régional, provincial ou national?
- Désire-t-on comparer avec un échantillon normatif ou avec une autre population?

**Informations à recueillir**

- Quel est le portrait actuel de la situation problématique visée ?
  - Quelle est l'ampleur des problèmes existants ?
  - Quelles sont leurs causes et leurs conséquences ?
  - Quelle est la population touchée directement ou indirectement par ces problèmes ?
  - Quels sont les facteurs de risque et de protection associés aux problèmes sur les plans individuel, familial, social, environnemental ou politique ?  
Quels sont les facteurs présents dans la population visée ?
  - Quels sont les autres problèmes sociaux pressants ?
  - Quelles sont les attitudes, les normes et le point de vue de la collectivité par rapport à ces problèmes ?
  - Quelles sont les différences entre certains groupes de la population concernant les besoins à combler ou les ressources disponibles ?
- Quelles sont les solutions possibles ou les possibilités d'actions pour satisfaire le besoin reconnu (activités, interventions, programmes, ressources ou services) ?
  - Quels sont leurs clientèles, leurs objectifs et leurs particularités ?  
Quelle est leur efficacité ? Quels sont leurs coûts ? Quelle est leur faisabilité dans le contexte actuel ?
  - Sont-ils facilement disponibles et accessibles ? Sont-ils connus par la clientèle visée ou par les intervenants ? Quels sont les obstacles à l'accessibilité ou à l'utilisation des services perçus ou vécus par les membres de la population cible ?
  - Quel est le niveau d'intégration ou de continuité des services ?  
Quelles sont leurs particularités et leurs complémentarités ?
  - Quels sont le niveau et la distribution des ressources ?
  - Quelles sont les solutions qui sont perçues comme acceptables par la population, les clients potentiels et les décideurs ?
- Quels sont les problèmes futurs ?

**Sources d'information**

- Quels sont les écrits et les banques de données qui seront consultées ?
- Quelles sont les personnes qui seront sollicitées ?
  - Clients ou client potentiels : membres de la communauté desservie, population à risque ou population en besoin, entourage des personnes en difficulté ?
  - Informateurs clés : experts, chercheurs, praticiens, gestionnaires, politiciens, groupes d'intérêts ou membres actifs de la communauté ?

**Protocole d'évaluation**

- Quelles sont les informations (secondaires) désirées qui existent déjà ?
- Quelles sont les nouvelles informations (données primaires) qui sont nécessaires et quelles sont les sources pour les obtenir ?
- Quelles sont les techniques les plus appropriées selon les fins et les contraintes de l'étude ?
- Est-il nécessaire d'élaborer un protocole détaillé de l'évaluation ?
- Quelle sera la méthode utilisée : échantillon, instruments utilisés, procédure de collecte, traitement des informations (analyses qualitatives ou quantitatives, logiciels utilisés, etc.) ?
- Comment le processus sera-t-il coordonné et contrôlé ?

## Encadré I (suite)

**Analyses des informations**

- L'écart entre la situation actuelle et la situation désirée est-il assez grand pour qu'on reconnaisse qu'il y a un véritable besoin?
- Quels sont les besoins ou les problèmes jugés prioritaires?
- Les besoins peuvent-ils être satisfaits sans ressources additionnelles? Quelles sont les ressources manquantes? Quelles sont les ressources à mobiliser?
- Les ressources et programmes existants sont-ils connus, disponibles, accessibles et acceptables pour les clients potentiels?
- Le programme est-il pertinent par rapport aux besoins identifiés dans la population ciblée? Quels sont les éléments qui justifient sa pertinence?
- Quels sont les éléments de continuité et de spécificité de ce nouveau programme avec les programmes et les initiatives déjà en place?
- À quelle population doit s'adresser le programme ou le service?
- Dans le cas d'un programme existant, est-ce que les conditions qui ont suscité la mise en place du programme existent toujours?

**Rapport et communication des résultats**

- Le rapport présente-t-il les fins de l'étude, la méthode suivie et ses limites, les problèmes et les solutions identifiés, les ressources et les services disponibles, les besoins prioritaires à combler ainsi que les recommandations pour l'action?
- Les résultats préliminaires de l'étude ont-ils été présentés aux parties prenantes afin d'obtenir leur rétroaction?
- Quand, comment et dans quelle forme les résultats seront-ils présentés?
- Quels sont les efforts nécessaires pour promouvoir l'acceptation des résultats de l'évaluation par la communauté?

## 5. LES TECHNIQUES DE COLLECTE DE DONNÉES

Plusieurs techniques de collecte de données peuvent être utilisées lors d'une évaluation des besoins. Certaines sont plus particulières à l'approche déductive (la revue des écrits, l'approche par indicateurs sociaux, l'enquête par questionnaire ou entrevue structurée), d'autres sont plus particulières à l'approche inductive (les entrevues individuelles auprès des informateurs clés, les entrevues de groupe [le *focus group*, la technique Delphi, le groupe nominal], les forums communautaires ou les audiences publiques). Dans ce qui suit, nous traçons les grandes lignes de ces techniques en identifiant leur utilité ainsi que leurs avantages et désavantages.

### 5.1. La revue des écrits

La revue des écrits consiste à examiner des rapports de recherche, des avis, des articles scientifiques ou professionnels, des livres ou d'autres publications portant sur la problématique visée (prévalence, facteurs de risque et de protection, causes, conséquences, problèmes associés, etc.) et sur les actions possibles pour résoudre cette problématique. Dans la perspective où

l'étude de besoins doit mener au développement d'un programme, on peut également inventorier les programmes existants sur le sujet. Cette recension peut s'effectuer par repérage avec des moteurs de recherche Internet (Google, Alta Vista, Lycos, Yahoo, etc.), par des recherches informatisées dans les banques usuelles (CBCA Education, CUBIQ, Current Contents, Education Research Complete, Dissertation Abstract International, ERIC, Francis, NJCR, Medline, PsycINFO, PsycArticles, Repère, Sage, Sociological Abstracts, Web of Science, etc.), par des recherches manuelles dans certains livres de base ou revues importantes et, enfin, par les contacts avec des personnes ou des organismes qui pourraient avoir produit ou connaître des recherches. Cette recherche peut également être facilitée par la consultation préalable de centres de veille et d'information ou d'observatoires dans le domaine psychosocial. Ces observatoires ont généralement pour mandat d'inventorier des documents disponibles sur les principaux résultats de recherche, sur l'analyse des données probantes existantes et sur les programmes d'intervention disponibles pour une problématique donnée. Le CLIPP (Centre de liaison sur l'intervention et la prévention psychosociale – <[www.clipp.ca](http://www.clipp.ca)>) et l'Observatoire canadien pour la prévention de la violence à l'école (<[www.preventionviolence.ca](http://www.preventionviolence.ca)>) sont de bons exemples de centres de veille et d'information. La revue des écrits s'avère une approche facilement accessible et peu coûteuse. En revanche, elle exige une bonne capacité de synthèse et les données disponibles ne sont pas toujours suffisamment à jour ou précises pour le groupe ciblé (Desrosiers *et al.*, 1998, p. 65).

## 5.2. L'approche par indicateurs sociaux

Cette technique consiste à inférer les besoins d'une population ou d'un groupe cible à partir d'indicateurs sociaux disponibles pour une région géographique particulière (Strupeck, 2003). Les indicateurs sociaux sont des mesures statistiques globales qui décrivent des aspects clés d'une situation sociale ou de problèmes sociaux ou de santé. Ces indicateurs peuvent prendre différentes formes : 1) démographiques (description de la population par âge, par sexe, par catégorie socioprofessionnelle, par secteur d'activité, par origine ethnique, etc.); 2) mesures du bien-être collectif (p. ex., taux de criminalité ou taux de chômage); 3) épidémiologiques (données sur les divers facteurs conditionnant l'apparition, la fréquence, le mode de diffusion et l'évolution des maladies physiques ou mentales affectant des groupes d'individus); 4) sociosanitaires (description des services ou des ressources disponibles : type, clientèle, accessibilité, utilisation, etc.); et 5) budgétaires (p. ex., les dépenses gouvernementales pour l'éducation, la santé, les services sociaux ou la sécurité publique). Cette approche suppose que les indicateurs sociaux sont reliés aux problèmes du groupe cible ou à leurs besoins. Ainsi, pour bien interpréter les données recueillies, les évaluateurs doivent se référer à un modèle explicatif concernant le problème social à l'étude.

Les informations sur les indicateurs sociaux peuvent provenir de plusieurs sources : archives publiques, rapports d'organismes publics ou parapublics, bases de données informatisées, revues, observatoires, ministères, agences, organismes communautaires, etc. Il existe plusieurs revues scientifiques qui se spécialisent dans la publication de telles données, par exemple *Annual Demographic Statistics*, *Canadian Social Trends*, *Demography*, *European Journal of Population*, *Population*. Pour ce qui est des indicateurs canadiens, plusieurs banques de données accessibles sur Internet constituent d'excellentes sources d'informations ; notons entre autres celles de Statistique Canada, de l'Institut de la statistique du Québec, de l'Enquête nationale sur la santé de la population ou de l'Enquête sur la santé dans les collectivités canadiennes.

Cette approche s'avère particulièrement utile pour décrire une population, pour estimer la prévalence et la distribution de problèmes sociaux (Rossy, Freeman et Lipsey, 2004), pour définir les besoins de services pour une région particulière ou pour une population cible (Favaro et Ferris, 1991 ; Strupeck, 2003) ou pour estimer la population qui pourrait bénéficier du programme qu'un organisme veut mettre en place (Royse *et al.*, 2006, p. 63). Les indicateurs sociaux sont accessibles, faciles à utiliser et permettent des comparaisons normatives. Ils peuvent également fournir des informations utiles pour constituer des bases d'échantillonnage pour des enquêtes ou des sondages. Toutefois, les indicateurs couramment disponibles sont parfois limités, la couverture usuelle des banques de données étant souvent circonscrite à des problèmes courants comme la pauvreté, le chômage, la criminalité ou la composition des ménages. En outre, la fiabilité de certaines banques est instable, les changements de définitions des concepts dans le temps rendent parfois les comparaisons et les interprétations difficiles (Desrosiers *et al.*, 1998, p. 64). Les données sont parfois périmées, incomplètes, peu fiables (en raison de la méthode utilisée ou de pressions politiques) ou proviennent d'organismes qui ne sont pas vraiment similaires (région géographique différente, critères d'éligibilité différents, etc.). Il faudra faire particulièrement attention lorsque les données ont été obtenues par Internet, car les sources sont souvent peu fiables (Kremer, 2003). Par ailleurs, lorsqu'il s'agit d'enquêtes nationales, il est souvent difficile d'obtenir des informations valables au plan local<sup>3</sup>. Enfin, l'utilisation d'indicateurs sociaux peut parfois mener à des erreurs d'interprétation, en particulier à des sophismes écologiques (Royse *et al.*, 2006, p. 66). Il y a un sophisme écologique lorsqu'on applique les

---

3. Par ailleurs, il peut être intéressant de comparer des données obtenues localement avec celles obtenues au plan régional, provincial ou national afin de justifier la pertinence d'un programme particulier auprès de la population visée, par exemple lorsque l'on peut démontrer que le taux de prévalence d'un problème est plus élevé au plan local qu'au plan provincial.

résultats issus de mesures prises au niveau du groupe (ou du milieu de vie) au niveau individuel, c'est-à-dire lorsque l'on fait des inférences sur la nature des individus composant une population en se basant seulement sur des ensembles de statistiques sur le groupe auquel ils appartiennent, comme si chaque membre du groupe manifestait les mêmes caractéristiques du groupe dans son ensemble. Le fait, par exemple, que les quartiers à forte proportion de personnes sans emploi enregistrent un fort taux de criminalité ne signifie pas nécessairement que les personnes sans emploi sont plus à risque d'être des criminels! Les stéréotypes sont une forme de sophisme écologique. Toutes ces limitations ne veulent pas dire que les indicateurs sociaux ne peuvent être utilisés comme sources d'information valables pour l'étude de besoins, mais seulement qu'ils doivent être utilisés pour ce qu'ils sont, des symptômes du problème et des indices sur son importance et son étendue.

### 5.3. L'enquête par questionnaires ou entretiens

Les enquêtes (ou sondage) sont souvent utilisées pour estimer l'ampleur des besoins et l'acceptabilité des solutions possibles pour répondre à ceux-ci, quand il n'y a pas suffisamment d'informations disponibles à ce sujet. Elles sont habituellement réalisées auprès d'un échantillon représentatif de la population visée ou auprès d'un ensemble de personnes touchées directement ou indirectement par le problème (p. ex., individus présentant le problème, famille, proches, etc.). Comme l'élaboration d'un échantillon représentatif de la population est extrêmement difficile et peut représenter des coûts élevés, Posovac et Carey (2007, p. 119) suggèrent un compromis, soit l'utilisation d'un groupe intact qui pourrait être représentatif des besoins de la population. Prenons l'exemple d'un gestionnaire d'une maison de la famille qui aimerait connaître les besoins des parents d'enfants d'âge scolaire. Il pourrait réaliser une enquête auprès des parents des enfants fréquentant une école publique dans le quartier où se situe la maison de la famille. Les enquêtes et les sondages sont tellement fréquents et semblent si faciles à répondre que les gens ne se rendent pas compte de l'ampleur de la tâche demandée et des efforts requis pour bien mener une enquête et obtenir une information valide et fiable. Cette approche requiert une connaissance approfondie des méthodes de recherche, en particulier concernant les techniques de construction de questionnaires ou de grilles d'entretien, la conduite d'entretien, les procédures d'échantillonnage ainsi que les analyses (statistiques ou qualitatives) des résultats<sup>4</sup>. Le cadre limité de ce chapitre ne nous permettant pas de présenter de façon détaillée la façon

---

4. Pour avoir des informations détaillées sur les techniques d'enquête et d'échantillonnage, voir entre autres Berthier (2006), Ardilly (2006) ou le manuel très exhaustif de Gubrium et Holstein (2002).

de mener une enquête, nous nous limiterons à discuter de certaines considérations importantes concernant la formulation des questions et le choix des techniques de collecte de données.

### **5.3.1. La formulation des questions**

La formulation des énoncés est une question très délicate lors d'une enquête de besoins. Posavac et Carey (2007, p. 119) identifient différentes erreurs à éviter. La première est d'aller trop vite aux solutions possibles avant de laisser les répondants s'exprimer sur les besoins non comblés, tant sur le plan de la communauté que sur le plan individuel. La deuxième erreur est de formuler des questions tendancieuses ou suggestives qui vont orienter dans un sens les réponses des participants. Les auteurs citent le cas du biais de consentement, où il a été démontré que les personnes tendent plus facilement à être en accord qu'à être en désaccord avec une question. Pour éviter ce biais, les auteurs suggèrent d'éviter les questions demandant l'accord ou le désaccord avec un énoncé en utilisant par exemple des questions où on demande de comparer entre eux des énoncés. La troisième erreur est de ne pas aborder les coûts d'un service potentiel ou les effets de sa mise en place (par exemple, le retrait d'un autre service), ce qui aurait pour effet de gonfler le taux d'accord envers ce service.

Une fois les thèmes choisis, l'enquêteur devra choisir entre la formulation de questions fermées ou à choix de réponse ou la formulation de questions ouvertes. Les questions fermées sont les plus fréquemment utilisées lors des enquêtes de besoin, car elles sont plus rapides à coder et à analyser et elles permettent de quantifier plus facilement l'étendue ou l'ampleur des problèmes ou des besoins. Elles ne laissent pas place néanmoins à l'expression de problèmes ou de besoins qui n'auraient pas été perçus par les concepteurs du questionnaire ou à l'expression de réactions spontanées, ce qu'offrent les questions ouvertes. Les questions ouvertes permettent également d'aller plus en profondeur dans la compréhension des situations problématiques vécues ou dans les réactions des personnes aux solutions proposées. Elles demandent toutefois des efforts plus grands pour l'analyse des données.

### **5.3.2. Les choix des techniques propres à l'enquête**

Différentes techniques de collecte de données peuvent être utilisées: questionnaires envoyés par la poste, entrevues face à face ou par téléphone, ou enquêtes assistées par ordinateur. Chacune comporte ses avantages et désavantages et doivent être utilisées selon les situations.

### ***5.3.2.1. Les questionnaires écrits***

Les questionnaires écrits sont habituellement envoyés par la poste ou distribués en personne auprès d'individus regroupés dans un organisme (par exemple, dans une école). Lorsqu'ils ne sont pas remplis sur place, les enquêteurs fournissent habituellement des enveloppes de retour affranchies. Les enquêtes postales obtiennent généralement des taux de non-réponses plus élevées que toutes les autres méthodes d'enquête, en particulier lorsqu'elles sont réalisées auprès de populations défavorisées (Gubrium et Holstein, 2002). Répondre à un questionnaire écrit est plus ardu pour les populations qui ont des difficultés en lecture ou les minorités ethniques qui ne maîtrisent pas suffisamment la langue du questionnaire. Il est souvent recommandé que l'enquêteur soit présent lorsque les participants remplissent les questionnaires afin de répondre à leurs questions au besoin. Le grand avantage de ce type d'enquête est de rejoindre un grand nombre de personnes en peu de temps et pour des coûts modérés (Mayer et Ouellet, 1991, p. 74).

### ***5.3.2.2. Les entrevues face à face ou par téléphone***

Le choix entre les entrevues téléphoniques ou face à face dépendra de différents critères, dont le type d'information recherchée (démographiques, personnelles, délicates ou non, etc.), la complexité du sujet, les ressources disponibles, le temps alloué, les interviewers disponibles et le besoin d'uniformité entre les interviewers. Shuy (2002) résume bien les avantages et désavantages associés à chacune des techniques. Bien que plusieurs personnes soient peu enthousiastes à répondre à des entrevues téléphoniques, celles-ci sont souvent moins coûteuses et permettent un meilleur contrôle des variables situationnelles, en particulier une plus grande standardisation de la façon de mener l'entrevue lorsque celles-ci sont réalisées dans un centre téléphonique en présence d'autres interviewers et sous la supervision occasionnelle d'un responsable. Le rapport coûts/bénéfices est plus grand pour les entretiens téléphoniques. Le temps d'exécution est jusqu'à 10 à 20 % lorsque le questionnaire est administré par téléphone plutôt qu'en personne et, en plus, la réalisation d'entrevues face à face nécessite du temps additionnel pour les arrangements nécessaires et pour les déplacements. Par contre, les entrevues face à face génèrent un taux de réponses moins élevé que les entrevues téléphoniques et des réponses généralement plus détaillées. Les entretiens face à face semblent aussi plus appropriés lorsqu'il s'agit de questions délicates ou controversées ou pour les problématiques plus complexes. Ils semblent également plus à propos avec les répondants marginalisés, les personnes âgées, les malentendants, les répondants allophones et les individus ayant des capacités limitées.

### **5.3.2.3. Les enquêtes assistées par ordinateur**

Avec la popularité sans cesse grandissante des moyens de communication électroniques, de plus en plus d'enquêtes de besoin se réalisent par Internet. Les enquêtes assistées par ordinateur peuvent prendre différentes formes : questionnaires envoyés par courriel grâce à des listes d'envoi, questionnaires en ligne ou discussions un à un par des programmes de clavardage. Dans les enquêtes par courriel, les questions sont soit envoyées dans le texte du message et le répondant doit taper « répondre » et insérer ses réponses dans le texte, soit le questionnaire est en pièce jointe, soit enfin un lien figure dans le texte et mène à un questionnaire en ligne. La première formule est plus appropriée lorsque les personnes ont des connaissances très limitées de l'Internet et lorsque le questionnaire est court. La deuxième formule, soit la pièce jointe, est plus hasardeuse, car plusieurs personnes ignorent comment avoir accès à une pièce jointe, le système antivirus de certains bloque l'accès au document, ou encore les répondants trouvent que la démarche leur demande trop de temps. Les courriels avec des liens dans le texte renvoyant à un questionnaire en ligne obtiennent des taux de réponse plus élevés (Mann et Stewart, 2002). Les questionnaires en ligne sont soit accessibles sur un site Internet public, soit sur un site privé dont l'accès est limité aux personnes qui ont été désignées. Ces questionnaires sont interactifs, permettent d'ajuster le déroulement des questions selon les réponses reçues et d'enregistrer les réponses dans le format voulu. Certains problèmes techniques peuvent toutefois être rencontrés, entre autres des incompatibilités avec le système d'exploitation de certains répondants, ou le trafic Internet qui peut rendre momentanément le questionnaire non disponible ou provoquer des blocages. Les discussions un à un par des programmes de clavardage sont plus rarement utilisés, mais lorsqu'elles le sont, elles remplacent les entrevues face à face ou par téléphone, en particulier auprès des jeunes.

Dans une revue des écrits à ce sujet, Mann et Stewart (2002) établissent les coûts et bénéfices associés à ces techniques. Le plus gros problème réside dans la sous-représentativité de la population touchée (plus de riches, plus de jeunes, plus d'urbains, plus d'hommes) et de la difficulté à atteindre certaines populations. Toutefois, lorsque l'évaluateur cherche à rejoindre des groupes de personnes qui utilisent régulièrement l'Internet dans le cadre de leur travail ou de leur loisir, c'est un médium de premier choix. Aussi, les listes d'envoi par courriel, les questionnaires en ligne et les forums de discussion peuvent être particulièrement utiles pour rejoindre rapidement des personnes partageant intérêts, expériences ou expertises sur un thème particulier et qui sont dispersés géographiquement. Elles sont judicieuses également pour rejoindre les jeunes ou avoir accès à des endroits pour lesquels l'accès est restreint (comme les prisons ou certaines communautés religieuses).

Le recrutement se fait souvent par les listes d'envoi ou en mettant des messages sur des sites Internet d'intérêts particuliers. Les participants aux enquêtes assistées par ordinateur, surtout lorsque les formulaires sont en ligne, apprécient la rapidité d'exécution de cette méthode (plus rapide que par écrit, système qui permet d'éviter certaines questions non pertinentes, etc.) et le fait qu'ils puissent remplir le questionnaire au moment qui leur convient, au travail ou à la maison. Le faible coût de ces enquêtes est sûrement l'avantage majeur. Une fois le questionnaire informatisé placé en ligne, on évite les coûts d'envoi, de déplacement et plusieurs frais liés à l'organisation de la collecte. De plus, lorsque les questionnaires sont en ligne ou numérisés, les données peuvent être facilement transférées à des logiciels de traitement de données. Par contre, ces enquêtes n'éliminent pas les problèmes d'éthique liés au consentement libre et éclairé et à la confidentialité des données. Aussi, il faut prévoir un dispositif pour l'authentification des répondants afin de s'assurer de leur origine et qu'ils ne répondent pas plus d'une fois à l'enquête. Dans le cas des questionnaires en ligne, il y a aussi le danger que certains groupes d'intérêts noyautent l'enquête, en demandant par exemple à tous leurs membres d'y répondre.

#### 5.4. Les informateurs clés

Les informateurs clés sont des personnes dont l'expérience ou l'emploi les place en bonne position de connaître les problèmes d'une communauté ou les besoins qui ne sont pas comblés. Il peut s'agir de spécialistes du sujet traité (universitaires, chargés de dossiers, etc.), de représentants de groupes sociaux, d'organismes ou d'institution, de membres actifs dans la communauté (membres du clergé, politiciens, responsables d'organismes bénévoles, etc.) ou d'intervenants et de décideurs qui agissent auprès de la même population ou sur le même problème. Des entrevues semi-structurées sont généralement utilisées et on encourage les informateurs à donner leurs points de vue personnels sur la question. Pour favoriser une bonne représentativité des informateurs consultés, il peut être utile d'avoir recours à la technique de l'échantillon *boule de neige*, qui consiste à demander aux premiers informateurs clés consultés de recommander d'autres personnes qui pourraient fournir des informations utiles. Il est très facile, rapide et peu coûteux de recourir à la méthode des informateurs clés. Cette méthode s'avère particulièrement utile lorsqu'il y a peu de données objectives disponibles, que le problème visé est relativement récent ou que les fonds disponibles sont insuffisants pour soutenir une meilleure approche. Cette méthode permet de nuancer l'information recueillie par indicateurs et par enquête, d'explorer des pistes inusitées et de dégager des éléments d'une stratégie d'amélioration de la situation (Desrosiers *et al.*, 1998, p. 65). Par contre, elle requiert

une bonne maîtrise des techniques d'entrevue<sup>5</sup> et il est difficile d'assurer une bonne représentativité de l'échantillon. De plus, la qualité et l'utilité des informations dépendent grandement de la crédibilité des informateurs consultés. Aussi, bien que les informateurs clés puissent fournir des informations très utiles concernant les caractéristiques de certaines populations cibles ou les services requis, peu disposent des informations nécessaires pour bien estimer le nombre de personnes affectées par une condition sociale ou pour identifier la distribution démographique ou géographique de ces personnes, et ils ont souvent tendance à surestimer la population touchée (Rossi *et al.*, 2004, p. 135). Un autre problème relatif à cette technique est relié au domaine d'expertise des informateurs. Le fait d'avoir une expertise dans un domaine particulier peut amener les informateurs à penser que la façon de répondre aux besoins non comblés est reliée à leur champ d'expertise (Posavac et Carey, 2007, p. 122). Par exemple, pour les problèmes d'apprentissage, les enseignants peuvent avancer que le meilleur moyen pour améliorer la situation est de réduire le nombre d'élèves par classe, alors que les orthopédagogues peuvent proposer d'augmenter l'accessibilité à leurs services. Enfin, lorsque vient le temps d'interpréter les résultats, l'évaluateur doit tenir compte des biais possibles des informateurs (leurs valeurs, leurs croyances, ce qu'ils peuvent perdre ou gagner de la surestimation ou de la sous-estimation du problème ou des solutions proposées, etc.).

## 5.5. Les entrevues de groupe et les techniques de travail en comité

Les entrevues de groupes et les techniques de travail en comité rassemblent des techniques comme le groupe de discussion (*focus group*, groupes à interactions réciproques), le groupe nominal et la technique Delphi. Alors que les deux premières techniques regroupent face à face un maximum de 10 à 12 personnes pour une durée d'environ deux heures, les participants à la technique Delphi sont isolés et leur participation anonyme se fait par écrit. Lors des entrevues de groupe, on présente habituellement aux participants des questions ouvertes auxquelles ils sont invités à répondre. Ces techniques permettent de recueillir dans un délai relativement court une somme intéressante d'informations, de clarifier et de regrouper divers énoncés, puis d'établir éventuellement un ordre des priorités pour l'intervention (Boussougou *et al.*, 2003). Tout comme pour les informateurs clés, la qualité et la représentativité des membres constituent une variable importante à considérer. Pour ce qui est des entrevues de groupe face à face, Duchesne et Haegel

---

5. Pour avoir des informations détaillées sur la conduite des entrevues, voir entre autres Boutin (1997) ou Blanchet et Gotman (2007).

(2005, p. 47) recommandent l'homogénéité sociale entre les participants d'un même groupe (âge, éducation, statut socioéconomique, etc.) afin de favoriser la prise de parole. Comme le but premier est le libre échange d'opinions, de grandes différences entre les membres peuvent intimider certains de prendre la parole et limiter leur participation. Pour s'assurer de la représentativité de la population, les auteurs suggèrent de tenir plusieurs groupes.

Ces techniques présentent plusieurs avantages : nuancer les informations recueillies par la revue des écrits, les indicateurs ou les enquêtes ; favoriser le partage des idées ; permettre de rejoindre plus de personnes et de recueillir plus d'informations en même temps que lors d'entrevues individuelles. Néanmoins, elles arborent les mêmes limites que celles qui touchent les informateurs clés et nécessitent de solides habiletés d'animation de groupe. Le chapitre 10 fournissant de plus amples informations sur la façon de mener des groupes de discussion, nous nous attarderons ici aux méthodes axées sur la recherche de consensus telles la technique Delphi et la technique du groupe nominal, qui s'avèrent particulièrement utiles pour l'établissement des priorités d'action. Les lecteurs intéressés à approfondir leurs connaissances sur la conduite des entrevues de groupe pourront consulter les ouvrages de Boutin (2007), Duchesne et Haegel (2005) ou Fern (2001).

### **5.5.1. Le groupe nominal**

Développée initialement par Delbecq, Van de Ven et Gustafson (1975) comme technique d'identification des problèmes, le groupe nominal est un regroupement de six à dix personnes (experts, informateurs clés ou utilisateurs potentiels d'un service) qui, en réponse à un problème commun ou à une question, travaillent d'abord indépendamment, puis partagent leurs idées selon un processus structuré dirigé par un animateur afin d'en arriver à un consensus sur les énoncés ou les propositions prioritaires. Cette technique est particulièrement utile pour les problèmes qui requièrent une intervention rapide (Mayer et Ouellet, 1991). Royse *et al.* (2006, p. 68) résument bien les étapes du processus :

1. L'animateur pose une question ou un problème à résoudre à l'ensemble des participants.
2. Chacun des participants doit y proposer par écrit le plus grand nombre d'idées dans un laps de temps déterminé.
3. L'animateur demande à tour de rôle à chacun des participants de communiquer une idée. Les idées sont inscrites au fur et à mesure sur un tableau ou une grande feuille de papier. Le processus continue ainsi de suite jusqu'à ce qu'il n'y ait plus d'idées à exprimer.
4. Une période de discussion suit où les participants peuvent développer, éliminer, combiner ou ajouter de nouvelles idées.

5. Par la suite, les participants votent individuellement de façon anonyme sur les cinq idées les plus importantes en les plaçant par rang de priorité.
6. Pour diminuer la dispersion des votes, resserrer le consensus ou éviter les incompréhensions, le groupe discute des votes anonymes.
7. Après la discussion, il y a un vote final.

Cette technique présente plusieurs avantages sur les groupes de discussion moins structurés. Elle permet d'avancer plus rapidement et de maintenir la motivation des membres du groupe tout au long du processus, vu que tous ont une chance égale de s'exprimer (Mayer et Ouellet, 1991, p. 90). Elle permet également d'éviter les problèmes communs liés à la dynamique des groupes, en particulier la monopolisation du temps de parole par certains individus ou les problèmes d'influence (Favaro et Ferris, 1991, p. 15). Par contre, le laps de temps restreint alloué pour les discussions peut en frustrer certains et mener à un exposé superficiel des problèmes ou des propositions.

### 5.5.2. La technique Delphi

La technique Delphi (Delbecq *et al.*, 1975) consiste à sélectionner des informateurs clés et à leur demander de répondre par écrit à une série de questionnaires successifs sur une problématique. Ces informateurs ne se rencontrent jamais en personnes et leurs réponses demeurent anonymes. Le processus se déroule habituellement selon les étapes suivantes :

1. Sélection des informateurs clés (de 5 à 15 personnes).
2. Élaboration du premier questionnaire avec des questions ouvertes et générales et envoi aux participants.
3. Compilation des questionnaires retournés ainsi qu'analyse et synthèse des réponses; pour chacune des questions, des analyses statistiques descriptives des réponses sont faites (p. ex., moyennes, fréquences, etc.).
4. Élaboration du deuxième questionnaire: ce questionnaire pose les mêmes questions que le premier, mais présente cette fois les réponses originales du répondant ainsi que la distribution des réponses des autres enquêtés. Ce questionnaire peut également inclure une échelle de Likert pour chacune des réponses afin de solliciter le degré d'accord des participants avec l'énoncé.
5. Distribution du second questionnaire au même groupe. À cette étape, les participants peuvent changer leurs réponses originales, mais lorsque celles-ci s'éloignent trop de la majorité des autres membres, les participants doivent justifier leur point de vue.

Le processus continue ainsi jusqu'à ce que les membres du groupe aient atteint un certain consensus. Habituellement, après trois questionnaires, on arrive à un consensus.

D'après Mayer et Ouellet (1991, p. 84), cette technique se prête bien à l'analyse des problèmes complexes. Elle permet de travailler avec des individus qui sont dispersés géographiquement et les répondants peuvent prendre le temps nécessaire pour bien étayer leurs points de vue tout en conservant l'anonymat. Toutefois, même si on demande aux gens de répondre de façon indépendante, on ne peut savoir si tel est le cas, surtout lorsque des individus d'un même organisme participent à l'enquête (Favaro et Ferris, 1991, p. 15). Par ailleurs, ce processus peut s'avérer relativement long (jusqu'à trois mois) et on ne peut savoir à l'avance combien d'interactions (ou de tours de questionnaires) seront nécessaires pour atteindre le consensus. Cette technique demande que les participants possèdent de bonnes habiletés d'écriture et de synthèse et soient suffisamment motivés pour aller au bout du processus qui nécessite un certain investissement en temps de leur part.

## 5.6. Les forums communautaires ou audiences publiques

Les forums communautaires ou audiences publiques sont des assemblées ouvertes à tous les membres d'une communauté afin de recueillir leurs points de vue sur les besoins particuliers à combler et la façon de le faire (Mayer et Ouellet, 1991, p. 94). Ainsi, contrairement aux entrevues de groupe, les participants ne sont pas sélectionnés et on vise l'expression du plus grand nombre de personnes possible, peu importe leur origine, leur statut ou l'emploi occupé. Pour attirer un bon auditoire, l'événement doit être connu par la population et doit être facilement accessible. Habituellement, le forum se déroule selon les règles et le fonctionnement des assemblées délibérantes. McKillip (1998) recommande que les commentaires des participants soient limités à trois minutes et que les détails procéduraux soient annoncés au début de l'événement et rappelés régulièrement. Cette technique a l'avantage d'être accessible à tous ceux qui veulent participer, nécessite un temps de préparation et de réalisation minime, et entraîne peu de frais. De surcroît, elle requiert peu d'expertise en recherche pour interpréter ou synthétiser les résultats. Si le forum attire un grand nombre de personnes représentatives du milieu, il permet l'expression d'un grand nombre d'opinions et la confrontation des idées, favorise la compréhension de la dynamique du milieu et permet de jeter les bases d'une collaboration ultérieure (Desrosiers *et al.*, 1998, p. 65). Plusieurs problèmes sont néanmoins liés à la conduite des forums communautaires. Un premier problème est lié à la faible participation à ce type d'événement, à moins que le sujet soit controversé. Souvent, les participants se limitent au personnel des organismes qui œuvrent auprès

de la population concernée. Un deuxième problème concerne la représentativité des participants : certains points de vue peuvent être sous-représentés ou surreprésentés. Des groupes d'intérêt peuvent noyauter l'assemblée en y plaçant tous leurs partisans ou des groupes relativement peu nombreux, mais bien représentés, peuvent dominer oralement l'assemblée. Par ailleurs, ces audiences attirent souvent des personnes qui ont des griefs personnels à communiquer et dont les commentaires fournissent peu d'informations utiles à propos des besoins de la communauté. Un autre problème est lié au fait que ces forums tournent souvent autour d'une problématique très délimitée, laissant en plan d'autres problèmes qui pourraient être plus majeurs pour la communauté visée. Finalement, il y a danger de créer des attentes dans la population qu'on ne pourra combler par la suite.

## **6. L'UTILISATION DE L'ÉVALUATION DES BESOINS DANS LA PLANIFICATION DE PROGRAMME**

Une fois qu'on a bien identifié les besoins d'une population cible, la planification du programme peut commencer. Le bilan de la mesure des besoins devrait permettre d'établir la conjoncture idéale entre les besoins prioritaires à combler, les ressources à mobiliser, les ressources mobilisables et la complémentarité avec les services et les initiatives déjà en place. Les chances de succès du nouveau programme augmenteront si on intègre les parties prenantes dans la planification et la sélection des services à offrir ainsi que dans la définition des modalités d'implantation et d'accès. La définition des objectifs du programme devrait découler des besoins prioritaires identifiés et des caractéristiques de la population ciblée. Tout en étant propre aux besoins identifiés, le nouveau programme devrait tenir compte de l'inventaire des ressources afin d'éviter les dédoublements et de s'assurer de sa complémentarité avec les autres programmes existants. Le bilan de l'étude de besoins devrait également permettre au planificateur d'avoir une bonne idée du contexte dans lequel sera implanté le programme et de l'acceptabilité de celui-ci par la communauté afin d'en tenir compte, tant pour la conception du programme que pour la promotion qui en sera faite par la suite.

## **CONCLUSION**

Étape souvent escamotée ou faite de façon informelle, lorsqu'elle est faite de façon systématique, l'évaluation de besoins peut devenir un outil de planification très puissant pour établir les priorités quant aux besoins à desservir et pour développer des stratégies pour y répondre. Cette évaluation n'a pas toujours à devenir un processus complexe et onéreux. Dans bien des cas, une

bonne revue des écrits, une analyse des indicateurs sociaux et quelques entrevues individuelles ou de groupe avec des informateurs clés et des membres de la population cible pourront permettre de dresser un bon portrait de la situation actuelle qui justifie que des actions soient envisagées ainsi que des solutions possibles. En guide de conclusion, l'encadré 2 présente un exemple d'évaluation de besoins qui illustre différentes techniques ou démarches qui sont proposées dans ce chapitre<sup>6</sup>.

#### Encadré 2

### **Une étude de besoins en promotion de la santé de personnes fréquentant des services pour personnes handicapées en Irlande**

#### **Fins de l'étude**

- Le premier objectif était d'évaluer les habitudes de vie de clients avec une déficience intellectuelle, physique ou sensorielle fréquentant différents services offerts par le Health Services Executive East Coast Area, en Irlande. Le deuxième objectif était d'évaluer les besoins en promotion de la santé perçus par ces clients de façon à identifier les domaines à cibler pour l'éducation à la santé et la santé préventive.

#### **Population cible**

- Douze organisations offrent des services de jour, de répit ou d'hébergement aux personnes avec une déficience intellectuelle, physique ou sensorielle dans la région de la côte est d'Irlande (environ 350 000 personnes). Il y a 60 centres qui fournissent des services à environ 900 clients.

#### **Méthode**

- L'approche utilisée pour l'étude de besoins est mixte avec triangulation des données. Elle comprend trois parties : une enquête par entrevue structurée, des groupes de discussion et une enquête postale.

#### **Enquête par entrevues structurées**

- Participants : un échantillon de 425 clients répartis par âge, sexe et type de déficiences (247 clients avec une déficience intellectuelle et 180 clients avec une déficience physique ou sensorielle) ; les clients de moins de 15 ans ont été exclus de l'étude (taux de réponse 98,8%).
- Procédure : cinq interviewers ont reçu un entraînement sur la façon de faire une entrevue, en particulier avec une personne présentant une déficience. Deux questionnaires d'entrevues semi-structurées ont été développés à partir de quatre questionnaires existants, un pour les personnes avec une déficience physique ou sensorielle et un autre pour les personnes avec une déficience intellectuelle.
- Analyses : les résultats ont été analysés avec les logiciels Epi-Info et SPSS.

6. Le chapitre 3 présente également l'utilisation de différentes démarches ou techniques proposées dans ce chapitre, dont l'enquête par questionnaires et par entrevues, les groupes de discussion, les indicateurs sociaux et les entretiens téléphoniques auprès d'informateurs clés.

## Encadré 2 (suite)

**Groupes de discussion**

- Participants : deux groupes de clients avec une déficience intellectuelle ( $n = 12$ ) ; deux groupes d'aidants naturels (8 mères, 2 pères, 1 membre de la fratrie) ; trois groupes de dispensateurs de services ( $n = 12$ ).
- Instrument : un guide de discussion a été élaboré à partir de différents thèmes ayant émergé lors de l'enquête ou de l'analyse des résultats de l'enquête.
- Analyses : analyse de contenus (méthode non spécifiée).

**Enquête postale**

- Participants : 25 gestionnaires de centres de service (taux de réponse = 88%)
- Instrument : questionnaire structuré élaboré à partir de deux questionnaires existants sur les initiatives en promotion de la santé qui seraient profitables pour les clients.
- Analyses : les résultats ont été analysés avec les logiciels Epi-Info et SPSS.

**Résultats**

- Pour la plupart des réponses ouvertes, les clients répondaient des réponses très brèves ou pas de réponse du tout. Les clients présentant une déficience intellectuelle ont éprouvé beaucoup de difficulté à répondre aux questions relatives au temps ou à la fréquence de certains comportements liés à la santé. Les méthodes pour améliorer les services et les obstacles furent peu abordées. Les principaux résultats de cette étude sont que les clients veulent recevoir plus de services liés à des thérapies par les arts, plus d'activités physiques et plus de thérapies de relaxation. Un autre besoin est le désir de participer à plus d'activités sociales et de loisirs. Les dispensateurs de soins ont mis plus d'importance sur la physiothérapie, l'orthophonie, la thérapie occupationnelle et le soutien social ou psychologique. Ils ont également recommandé qu'il y ait plus de programmes relatifs aux habitudes alimentaires et à l'activité physique. Les aidants naturels ont, quant à eux, fait ressortir les besoins pour les services thérapeutiques et de réadaptation, et ont souligné les bénéfices de la thérapie par les arts ainsi que des activités sociales et de loisirs. Ainsi, les clients ont fait une contribution importante à l'étude en faisant ressortir des priorités différentes de celles établies par les dispensateurs de services ou les aidants naturels.

Source : Boland, Daly et Staines, 2008.

**BIBLIOGRAPHIE**

- ARDILY, P. (2006). *Les techniques de sondage*, Paris, Éditions Technip.
- BEADLE DE PALOMO, F. et E. LUNA (2003). *The Needs Assessment: Tools for Long-term Planning*, Washington, AED Center for Community-Based Health Strategies, <coach.aed.org/pubs/factsheets/NeedsAssessment.pdf>, consulté le 2 août 2008.
- BERTHIER, N. (2006). *Les techniques d'enquêtes en sciences sociales : méthodes et exercices corrigés*, 2<sup>e</sup> éd., Paris, Colin.
- BLANCHET, A. et A. GOTMAN (2007). *Enquête et ses méthodes : l'entretien*, Paris, Colin.

- BOLAND, M., L. DALY et A. STAINES (2008). «Methodological issues in inclusive intellectual disability research: A health promotion needs assessment of people attending Irish disability services», *Journal of Applied Research in Intellectual Disabilities*, vol. 21, p. 199-209.
- BOUSSOUGOU, J.-P. et al. (2003). *Évaluation de programmes éducatifs: guide pratique*, Eastman, Éditions Behaviora Inc.
- BOUTIN, G. (1997). *L'entretien de recherche qualitatif*, Québec, Presses de l'Université du Québec.
- BOUTIN, G. (2007). *L'entretien de groupe en recherche et formation*, Montréal, Éditions nouvelles.
- BRADSHAW, J. (1977). «The concept of social need», dans N. Gilbert et H. Specht (dir.), *Planning for Social Welfare: Issues, Models, and Tasks*, Englewood Cliffs, Prentice-Hall, p. 290-296.
- CHAMBERS, D.E., K.R. WEDEL et M.K. RODWELL (1992). *Evaluating Social Programs*, Boston, Allyn and Bacon.
- DELBECK, A., A.H. VAN DE VEN et H. GUSTAFSON (1975). *Groups Techniques for Program Planning: A Guide to Nominal Group and Delphi Processes*, Glenview, Scott, Foresman.
- DESROSIERS, H. et al. (1998). *Cadre de pratique pour l'évaluation de programmes: applications en promotion de la santé et en toxicomanie*, Québec, Gouvernement du Québec, Ministère de la Santé et des Services sociaux.
- DUCHESNE, S. et F. HAEGEL (2005). *Enquête et ses méthodes: l'entretien collectif*, Paris, Colin.
- FAVARO, P. et L.E. FERRIS (1991). «Program evaluation with limited fiscal and human resources», dans A.J. Love (dir.), *Evaluation Methods Sourcebook*, Ottawa, Canadian Evaluation Society, p. 6-25.
- FERN, E.F. (2001). *Advanced Focus Group Research*, Thousand Oaks, Sage Publications.
- GATES, B.L. (1980). *Social Program Administration: The Implementation of Social Policy*, Englewood Cliffs, Prentice-Hall.
- GUBRIUM, J.F., et J.A. HOLSTEIN (dir.) (2002). *Handbook of Interview Research: Context and Method*, Thousand Oaks, Sage Publications.
- HOWEL, D.C. (2008). *Méthodes statistiques en sciences humaines* (Traduction de *Statistical Methods for Psychology* de M. Rogier, V. Yzerbyt et Y. Bestgen), Bruxelles, De Boeck.
- KAUFMAN, R.A. (1972). *Educational System Planning*, Englewood Cliffs, Prentice Hall.
- KREMER, M.A. (2003). «Resources for data gathering and comparison», dans D.A. Nardi et J.M. Petr (dir.), *Community Health and Wellness Needs Assessment: A Step-by-step Guide*, New York, Thomson, Delmar Learning, p. 243-257.
- LAPOINTE, J.J. (1992). *La conduite d'une étude de besoins en éducation et en formation: une approche systémique*, Québec, Presses de l'Université du Québec.

- MANN, C. et F. STEWART (2002). « Internet interviewing », dans J.F. Gubrium et J.A. Holstein (dir.), *Handbook of Interview Research: Context and Method*, Thousand Oaks, Sage Publications, p. 59-82.
- MASLOW, A.H. (1954). *Motivation and Personality*, New York, Harper and Row.
- MAYER, R. et F. OUELLET (1991). *Méthodologie de recherche pour les intervenants sociaux*, Montréal, Gaëtan Morin Éditeur.
- MCKILLIP, J. (1998). « Need analysis: Process and techniques », dans L. Brickman et D.J. Rog (dir.), *Handbook of Applied Social Research Methods*, Thousand Oaks, Sage Publications, p. 261-284.
- MILES, M.B. et A.M. HUBERMAN (2003). *Analyses des données qualitatives* (Traduction de la 2<sup>e</sup> édition américaine par M.H. Rispal), Paris, De Boeck.
- NADEAU, M.-A. (1988). *L'évaluation de programme: Théorie et pratique*, Québec, Presses de l'Université Laval.
- NARDI, D.A. (2003a). « Introduction to health and wellness needs assessment », dans D.A. Nardi et J.M. Petr (dir.), *Community Health and Wellness Needs Assessment: A Step-by-step Guide*, New York, Thomson, Delmar Learning, p. 1-22.
- NARDI, D.A. (2003b). « Steps of the community health and wellness needs assessment process », dans D.A. Nardi et J.M. Petr (dir.), *Community Health and Wellness Needs Assessment: A Step-by-step Guide*, New York, Thomson, Delmar Learning, p. 23-45.
- PAILLÉ, P. et A. MUCCHIELLI (2003). *L'analyse qualitative en sciences humaines et sociales*, Paris, Colin.
- POSAVAC, E.J. et R.G. CAREY (2007). *Program Evaluation: Methods and Case Studies*, 7<sup>e</sup> éd., Upper Saddle River, Pearson Practice Hall.
- ROSSI, P.H., H.E. FREEMAN et M.W. LIPSEY (2004). *Evaluation: A Systematic Approach*, 6<sup>e</sup> éd., Thousand Oaks, Sage Publications.
- ROYSE, D. et al. (2006). *Program Evaluation: An Introduction*, 4<sup>e</sup> éd., Toronto, Thomson Brooks/Cole.
- SHUY, R.W. (2002). « In-person versus telephone interviewing », dans J.F. Gubrium et J.A. Holstein (dir.), *Handbook of Interview Research: Context and Method*, Thousand Oaks, Sage Publications, p. 59-82.
- SINGLETON, R.A. et B.C. STRAITS (2002). « Survey interviewing », dans J.F. Gubrium et J.A. Holstein (dir.), *Handbook of Interview Research: Context and Method*, Thousand Oaks, Sage Publications, p. 59-82.
- STRUPECK., C.D. (2003). « Analyzing indicators and using the analysis », dans D.A. Nardi et J.M. Petr (dir.), *Community Health and Wellness Needs Assessment: A Step-by-step Guide*, New York, Thomson, Delmar Learning, p. 219-241.
- WITKIN, B.R. et J.W. ALTSCHULD (1995). *Planning and Conducting Needs Assessments: A Practical Guide*, Thousands Oaks, Sage.

# CHAPITRE 5

---

## **Éléments et critères d'évaluabilité d'un programme d'intervention psychosociale**

Pierre Potvin  
Département de psychoéducation  
Université du Québec à Trois-Rivières

# 1. LE PROGRAMME D'INTERVENTION

Le présent chapitre a pour objectif d'aborder l'évaluabilité d'un programme. Les principales questions que se pose l'évaluateur avant même de planifier l'évaluation des effets d'un programme sont les suivantes: est-ce que le programme est suffisamment développé pour faire l'objet d'une évaluation? Jusqu'à quel point le programme, tel que représenté par le modèle descriptif, est suffisamment clair pour rendre l'évaluation réalisable et éventuellement utile? En d'autres mots, est-il évaluable?

Comme cet ouvrage sur l'évaluation de programmes le démontre, il existe plusieurs types d'évaluation. Ce chapitre sur l'évaluabilité d'un programme concerne plus particulièrement l'évaluation des effets d'un programme d'intervention et les conditions préalables à ce type d'évaluation.

Un programme d'intervention psychosociale vise à répondre à des besoins d'une clientèle qui éprouve un problème. Le développement d'un programme est donc étroitement lié à une problématique et à l'analyse des besoins qui a été réalisée sur le sujet. Les objectifs d'intervention d'un programme, quant à eux, découlent des besoins prioritaires identifiés au sein de la clientèle. Les besoins retenus serviront donc de base pour l'élaboration des objectifs du programme d'intervention (Monette, Charrette et Jobin, 1998).

Ces besoins indiquent les compétences à acquérir par la clientèle cible du programme. Le point commun entre ces compétences rejoint le but global du programme d'intervention, qui est le pivot entre les besoins et les objectifs généraux et particuliers qui en découlent (Monette, Charrette et Jobin, 1998).

Rappelons ici l'importance de bien distinguer la notion de problème, qui réfère à l'existence chez la clientèle cible d'un état ou d'un phénomène jugé indésirable (Beaudry, 1986, dans Paquette et Chagnon, 2000), alors que la notion de besoin réfère à l'importance plus ou moins grande d'éléments ou de solutions propres à contrecarrer ou à prévenir un quelconque problème (Beaudry, 1992, dans Paquette et Chagnon, 2000).

Avant d'aborder directement la notion d'évaluabilité de programme, il importe de se rappeler ce qu'on entend par programme d'intervention, et quelles en sont les différentes phases d'élaboration. Il existe de nombreuses définitions de programmes; en voici quelques-unes.

Selon Pineault et Daveluy (1995), un programme est un ensemble organisé, cohérent et intégré d'activités et de services réalisés simultanément ou successivement, avec les ressources nécessaires, dans le but d'atteindre des objectifs déterminés, en rapport avec des problèmes de santé précis, et ce, pour une population définie.

Plante (1994), pour sa part, mentionne qu'un programme connote l'idée d'une transformation. C'est un assemblage réfléchi d'idées, de stratégies et de personnes mis en place pour transformer des choses ou l'état des choses avec une perspective de permanence.

Nadeau (1988), quant à lui, définit de façon opérationnelle le programme comme un ensemble organisé de buts, d'objectifs particuliers, de contenus organisés de façon séquentielle, de moyens didactiques, d'activités d'apprentissage et de procédés d'évaluation pour mesurer l'atteinte de ces objectifs.

Toujours pour mieux comprendre ce qu'est un programme d'intervention, il devient également nécessaire d'en connaître la méthode de développement. Il existe différentes phases de développement d'un programme. Pour cet aspect, nous nous référons aux travaux de Loisel (2001) et à ceux de Tard, Beaudoin, Turcotte et Ouellet (1997).

En fait, la méthode de développement de programme peut se référer à la recherche développement telle que la présente Loisel (2001) lorsqu'il mentionne que ce type de recherche réfère au processus utilisé pour développer et valider des produits éducatifs (programmes, logiciels, méthodes, stratégies...). Ce type de recherche vise, par l'utilisation de connaissances scientifiques et de données de recherche, à produire des objets ou des procédés nouveaux (Legendre, dans Loisel, 2001). En principe, un programme d'intervention devrait être développé selon les cinq phases suivantes.

1. Phase d'analyse préalable: Elle permet de mettre en évidence les caractéristiques du problème (p. ex.: l'outil va servir à résoudre quel problème?), à bien identifier les besoins et les caractéristiques de la clientèle cible.
2. Phase de recension: Elle comprend la recension des écrits concernant l'état de connaissances (pratiques ou théoriques) sur les programmes pertinents pour résoudre un problème particulier auprès d'une clientèle particulière. On recherche également des modèles « programmes exemplaires » qui vont aider à l'orientation de la production.
3. Phase de production: C'est la phase du développement du programme comme tel, où les utilisateurs le mettent sur pied concrètement afin de s'assurer d'une validité « terrain » (connaissance expérientielle).
4. Phase de mise à l'essai (expérimentation, prétest): Lors de cette phase, les utilisateurs du programme sont formés à son utilisation; vient en deuxième lieu l'expérimentation (prétest) du programme par un groupe restreint d'utilisateurs représentatifs.
5. Phase d'implantation (mise en œuvre) d'évaluation et de révision: C'est une mise en œuvre du programme sur une plus grande échelle que lors du prétest. Durant cette phase, on procède à une évaluation d'implantation en apportant les corrections et les ajustements nécessaires. Une

vérification du niveau de satisfaction des utilisateurs et de l'atteinte des objectifs est également effectuée. Enfin, les corrections ainsi que la révision et la mise au point finale du programme complètent le processus de développement de programme.

Ces différentes phases dans le développement d'un programme font également appel à différents types d'évaluation d'un programme. Ainsi, la première et la deuxième phase (Phase d'analyse préalable et Phase de recension) s'appuient sur l'identification d'une problématique et sur l'évaluation des besoins de la clientèle cible visée par le futur programme, alors que la Phase de mise à l'essai et la Phase d'implantation peuvent faire l'objet d'évaluation formative, d'évaluation d'implantation ou de mise en œuvre. Précisons ici que l'évaluation formative maintient une rétroaction constante tout au long du processus d'élaboration et d'implantation du programme, permettant à celui-ci d'être continuellement modifié à mesure que des données nouvelles sont connues. L'objectif même de l'évaluation du processus est d'améliorer une action en cours, ou un programme en développement ou en activité (Tard *et al.*, 1997). On cherche également à accroître l'évaluabilité du programme (Trevisan, 2007). Ce type d'évaluation peut par exemple révéler des fautes ou des problèmes dans la livraison du programme, dans son application, ou que la clientèle cible n'est pas bien définie ou, encore, que l'intervention elle-même a besoin d'être reconceptualisée (Rossi et Lipsey, 1999).

Selon les différents moments de vie d'un programme, certains types d'évaluation seront appropriés et d'autres impraticables. Dans certains cas, le programme doit avoir atteint son plein développement pour que l'évaluation soit réalisable. C'est le cas si l'évaluation porte sur les effets, l'impact, le rendement ou l'efficacité du programme. Il sera très difficile, sinon impossible, de procéder à une évaluation des effets d'un programme qui n'a pas atteint un certain niveau de développement, une certaine stabilité, une certaine maturité: la stabilité est l'accueil de la même clientèle cible avec les mêmes modes d'intervention et les mêmes activités depuis une période relativement longue (Paquette et Chagnon, 2000).

Comme le mentionnent Lecomte et Rutman (1982), à moins de définir un programme de façon raisonnablement cohérente et exacte, il est impossible d'attribuer les résultats évalués à une intervention clairement reconnue. De plus, pour qu'un programme soit évaluable, il est important qu'il ait atteint un certain niveau de maturité, qu'il ait fait l'objet d'un développement, d'une application, d'un ajustement. Ce n'est qu'après ces ajustements et une relative stabilité qu'il sera possible d'évaluer les effets d'un programme.

## 2. L'ÉVALUATION DE L'ÉVALUABILITÉ D'UN PROGRAMME (EEP)<sup>1</sup> : CE QUE DOIT COMPRENDRE UN PROGRAMME POUR ÊTRE ÉVALUABLE

Trevisan (2007) a réalisé une imposante étude sur l'évaluation de l'évaluabilité intitulée « Evaluability Assesment from 1986 to 2006 ». Quelque 23 publications scientifiques sur l'évaluation de l'évaluabilité publiées durant ces 20 ans ont été analysées et les résultats permettent de mieux comprendre le but de ce type d'évaluation ou de préévaluation, ainsi que les méthodologies utilisées. Cet auteur mentionne, entre autres, que ce type d'évaluation était particulièrement utilisé durant les années 1970 à 1980, puis a subi une certaine baisse durant les années 1980 jusque vers 1995, pour par la suite redevenir une pratique bien utilisée jusqu'à présent. Ce type d'évaluation est donc reconnu comme pertinent dans la pratique d'évaluation de programmes.

Selon Trevisan (2007), l'évaluation de l'évaluabilité d'un programme (EEP) a été développée initialement durant les années 1970 surtout par les travaux de Joseph Wholey et ses collègues du Urban Institute, aux États-Unis. Le principal problème à l'origine de cette pratique de l'EEP était la difficulté d'évaluer les programmes peu développés ou mal formulés. C'était une façon d'améliorer l'évaluation d'un programme que de réaliser une activité de préévaluation. À ce sujet, une recherche de Rog (1985, cité par Trevisan, 2007) qui examine 57 études démontre que la moitié de ces études examinées amène au constat que ces programmes ne peuvent qu'être partiellement évalués.

Au départ, la finalité de l'EEP était centrée sur l'évaluation de la maturité du programme pour faire l'objet d'une évaluation de ses effets (évaluation sommative). L'étude de Trevisan (2007) démontre que l'utilisation de l'EEP maintient cette finalité principale de vérifier si le programme est prêt pour une évaluation de ses effets (sommative). Cependant, avec les années, s'est ajoutée l'utilisation de l'EEP dans le but de réaliser une évaluation formative afin de contribuer au développement du programme, à son changement et à son amélioration. On utilise ce type de préévaluation pour revoir les buts et objectifs du programme, modifier certaines composantes du programmes, augmenter la compréhension du programme auprès des responsables et, enfin, pour développer la logique du programme.

---

1. L'évaluation de l'évaluabilité se traduit en anglais par *evaluability assesment* (EA); voir Trevisan (2007).

Comme mentionné plus haut, l'évaluabilité d'un programme est particulièrement concernée lorsqu'on veut réaliser une évaluation des effets d'un programme d'intervention (évaluation sommative).

L'EEP consistera donc à vérifier si le programme est suffisamment développé, stable et mature pour faire l'objet d'une évaluation de ses effets. C'est une analyse de la qualité du programme, un processus qui vise à clarifier le design du programme, à explorer sa réalité et, si nécessaire, à aider à le reformuler afin qu'il rencontre les critères lui permettant d'être évaluable.

Pour être évaluable, un programme doit être clairement articulé et explicite, les objectifs et les effets attendus, précisés et une présentation d'un lien logique entre le programme, les objectifs et les effets, attendus (Tard, Beaudoin, Turcotte et Ouellet, 1997). De plus, il doit comporter une description de sa clientèle cible, de la problématique associée, de la théorie sur le problème, de la théorie d'actions préconisées, des ressources humaines et financières utilisées, de la nature des services offerts et des activités réalisées ainsi que des résultats escomptés (Paquette et Chagnon, 2000).

L'exemple d'un cas d'application de l'EEP est donné en annexe.

## 2.1. La théorie du programme

La théorie du programme ou la logique du programme concerne la conception de la structure, du fonctionnement et des procédures du programme pour atteindre ses objectifs. Elle doit décrire clairement comment sont intégrés dans le programme les composantes du programme, les objectifs, les activités et la clientèle cible. Toutes les ressources reconnues comme pouvant contribuer à l'atteinte des objectifs doivent faire partie de cette description (Rossi et Lipsey, 1999).

La théorie du programme comporte une logique entre les actions à réaliser et le problème identifié. Elle présente un rationnel entre les actions posées, les buts, les objectifs et les effets recherchés. Ce rationnel est issu d'une hypothèse d'explication du problème ou d'une théorie sur les causes du problème ainsi que d'une hypothèse sur la façon de le résoudre (Paquette et Chagnon, 2000).

La théorie du programme c'est en quelque sorte la description détaillée de la relation entre les ressources du programme, les activités et les résultats du programme. Elle montre comment le programme fonctionne et, surtout, comment il permet de contribuer à produire les résultats attendus. Cette théorie du programme explique pourquoi le programme fait ce qu'il fait et présente un rationnel qui explique pourquoi, en faisant les choses de cette façon, on obtiendra les résultats désirés. De plus, cette théorie du programme précise sur quelle base on s'appuie pour dire que le programme

peut éventuellement atteindre les objectifs établis. Cette théorie a été construite aussi bien à partir des théories et de la recherche qu'à partir des idées tirées de l'expérience acquise en ce domaine (Rossi et Lipsey, 1999).

Le principe de la théorie du programme se base sur le fait que si les buts et les objectifs du programme ne sont pas raisonnablement reliés à ce que le programme veut changer ou améliorer, ou si les hypothèses ou présupposés et attentes intégrés dans le programme ne représentent pas une approche crédible, il y a peu de chance que le programme soit efficace et produise les effets attendus (Rossi et Lipsey, 1999).

## **2.2. Le rôle important des objectifs dans un programme**

Les objectifs sont intégrés dans la théorie du programme. Ces objectifs servent à préciser les effets attendus, les changements souhaités chez la clientèle cible (Paquette et Chagnon, 2000). Ils servent également de critères à l'évaluation de la réussite du programme; ils doivent donc être clairement articulés pour favoriser le choix ou le développement d'instruments de mesure. Les objectifs garantissent que les variables clés ont été identifiées et qu'elles feront partie de l'évaluation (Lecomte et Rutman, 1982).

Monette, Charrette et Jobin (1998) mentionnent que les objectifs d'un programme doivent être formulés en établissant quels sont les résultats désirés et quels indices permettront de savoir si ces résultats sont effectivement atteints. Les objectifs d'intervention doivent préciser le changement que le programme souhaite voir s'effectuer chez la clientèle cible. Il faut donc s'assurer que les indicateurs d'effets ou de résultats soient clairement définis. Ces indicateurs servent de critères objectifs pour mesurer le degré de réussite du programme.

## **2.3. L'évaluation de la théorie du programme**

Certains programmes s'appuient sur une théorie articulée; ils sont en conséquence construits sur la base d'une théorie. Cependant, il est rare que les programmes offrent à l'évaluateur une théorie du programme suffisamment explicite et détaillée qui permet d'entreprendre une évaluation significative. Dans plusieurs cas, le programme s'appuie plutôt sur une théorie implicite, qui demande à être découverte ou dévoilée. Dans ce cas, l'évaluateur doit travailler à extraire et décrire, par différents moyens, cette théorie implicite (Rossi et Lipsey, 1999). Ce travail visant à préciser et développer la théorie du programme fait partie des activités de l'EEP (Trevisan, 2007).

La première étape de l'évaluation de la théorie du programme consiste donc à produire une description articulée, explicite des conceptions, des hypothèses et des attentes qui constituent le rationnel de la façon dont le programme doit opérer. Cette description comprend l'ensemble des hypothèses sur lesquelles s'appuie le programme pour orienter les stratégies utilisées afin d'atteindre les buts et les objectifs en relation avec les résultats attendus (Rossi et Lipsey, 1999).

Une deuxième étape de l'évaluation de la théorie du programme est de chercher à décrire et à comprendre la réalité du programme telle que vue par les responsables du programme et le personnel qui l'applique. Il s'agit ici de comprendre ce qui est important dans ce programme pour les responsables et ceux qui y prennent part (Rossi et Lipsey, 1999; Trevisan, 2007).

L'aboutissement de l'évaluation de l'évaluabilité du programme, dont fait partie l'évaluation de la théorie du programme, conduit souvent à ce que les responsables du programme reconnaissent le besoin de modifier le programme, de le bonifier, de le préciser (Rossi et Lipsey, 1999).

### **3. OUTIL POUR ÉVALUER L'ÉVALUABILITÉ D'UN PROGRAMME**

Les méthodes habituellement les plus utilisées pour l'EEP sont, en ordre décroissant, *a*) les entrevues avec les responsables du programme, avec le personnel qui applique le programme, *b*) la revue et l'analyse de documents, *c*) la visite de sites où est appliqué le programme, *d*) le groupe de discussion, *e*) le sondage par questionnaire (Trevisan, 2007).

Cette étape de l'évaluation de l'évaluabilité du programme comporte des négociations et des investigations réalisées conjointement par l'évaluateur et les responsables du programme afin de déterminer si ce dernier rencontre les préconditions pour une évaluation de ses effets (Rossi et Lipsey, 1999).

Lecomte et Rutman (1982) mentionnent que, dans le cas où les conditions préalables à l'évaluation des effets d'un programme n'existent pas ou sont insuffisamment élaborées, les chercheurs évaluateurs doivent collaborer avec les planificateurs et les gestionnaires au processus d'articulation du programme, préciser les objectifs et autres effets, et analyser les hypothèses émises au départ. En d'autres termes, si la conclusion de l'évaluateur en arrive au résultat que le programme n'est pas prêt pour l'évaluation des effets, il faut alors entreprendre de développer, ajuster, adapter les aspects du programme qui demandent correction avant de reprendre son évaluation.

Ces auteurs vont plus loin en recommandant que les chercheurs évaluateurs refusent d'effectuer l'évaluation de l'efficacité et de l'efficacité d'un programme qui n'est pas évaluable. Le rationnel sous-jacent à cette affirmation s'appuie sur le fait qu'il est probable que l'évaluation prématurée d'un programme non évaluable produise des études inappropriées, peu utiles pour la planification, la gestion, ou comme base d'imputabilité.

En s'appuyant sur ce qui a été présenté précédemment et sur les divers auteurs qui soutiennent cette question de l'évaluabilité d'un programme, voici une série d'items qui permettent de vérifier si les effets d'un programme sont évaluables. Cette liste permet également d'identifier les aspects du programme qui demanderont à être développés avant d'entreprendre son évaluation. La liste d'items de vérification de la qualité du programme à évaluer comporte sept grands thèmes : analyse globale du programme, clientèle cible, problématique de la clientèle cible et analyse de besoins, objectifs et attente de résultats, théorie du programme, ressources humaines et matérielles et résultat de l'évaluation.

### **3.1. Analyse globale du programme**

- L'accès à des documents qui décrivent clairement le programme est possible.
- La description du programme est bien articulée.
- Le programme constitue une solution appropriée pour résoudre le problème visé.
- Les composantes du programme, les activités et le fonctionnement du programme sont suffisamment définis.
- Les modalités d'intervention sont bien opérationnalisées (consignes, procédures, activités, déroulement, etc.).
- Le programme offre un cheminement à travers une série d'activités prédéterminées en vue de changer l'état de la clientèle cible.
- Le programme a fait l'objet d'évaluations lors de la mise à l'essai, et d'ajustements si nécessaire.
- Le programme a été appliqué à quelques reprises et présente un niveau de maturité (une certaine stabilité) permettant son évaluation.

### **3.2. Clientèle cible**

- La clientèle cible qui peut bénéficier du programme est bien identifiée et ses caractéristiques, bien définies.
- Les modalités de sélection ainsi que les critères de sélection et d'exclusion sont précisés.

### 3.3. Problématique de la clientèle cible et analyse de besoins

- La problématique associée à la clientèle cible est bien documentée et soutenue par une théorie sur le problème.
- Cette théorie explicative du problème est valide.
- Le programme est le résultat d'une analyse de besoins de la clientèle et l'information sur les besoins prioritaires de la clientèle cible est bien définie.
- Le niveau de correspondance entre les caractéristiques de la clientèle cible (problème, capacités) et les activités du programme est bien dosé et adéquat.

### 3.4. Objectifs et attente de résultats

- Le programme répond à des objectifs généraux liés à la mission du milieu.
- Le but et les objectifs du programme sont bien définis, clairement énoncés et mesurables.
- Les objectifs disposent d'une définition opérationnelle des résultats attendus avec des indicateurs et des critères de succès.
- Une précision est apportée sur les changements attendus de comportements ou d'attitude chez la clientèle cible.
- Les facteurs qui peuvent faciliter ou entraver la réalisation des objectifs sont présentés.
- Les différents degrés de réussite sont précisés.
- La façon de mesurer les progrès est présentée.
- Le lien articulé entre le programme et les objectifs ou les résultats attendus est suffisamment solide pour justifier une évaluation.

### 3.5. Théorie du programme

- La « logique » du programme ou « la théorie du programme » est clairement présentée par la conception de la structure, du fonctionnement et des procédures du programme pour atteindre ses objectifs.
- Cette théorie du programme explique pourquoi le programme fait ce qu'il fait et présente un rationnel qui explique pourquoi, en faisant les choses de cette façon, on obtiendra les résultats désirés.

- On retrouve un lien logique entre les activités du programme et ses résultats, et les postulats de causalité sont plausibles en liant le programme aux résultats.
- En résumé, le programme repose sur une théorie explicite justifiant le choix des interventions, une théorie sur le problème et sur l'action ainsi qu'un rationnel sur lequel s'appuient les interventions. Cette théorie du programme est assez solide pour justifier une évaluation.

### **3.6. Ressources humaines et matérielles**

- Les caractéristiques et compétences nécessaires du personnel qui applique le programme sont présentées clairement.
- Les ressources humaines, matérielles et financières sont bien décrites, adéquates et suffisantes pour la réalisation du programme.

### **3.7. Résultat de l'évaluation**

- L'utilisation de l'information issue des résultats de l'évaluation des effets du programme est adéquatement précisée.

## **CONCLUSION**

Il existe de nombreux programmes en intervention psychosociale. L'une des difficultés liées à l'évaluation de ces programmes est le fait que peu de programmes rencontrent les critères d'évaluabilité. En fait, peu de programmes sont évaluables quant à leurs effets et demandent un développement.

L'évaluateur doit se poser la question si le programme est prêt pour être évalué quant à ses effets. Est-il suffisamment décrit, articulé? La théorie du programme est-elle présente, explicite? Le programme a-t-il atteint un niveau acceptable de stabilité, de maturité pour pouvoir être évalué?

L'évaluateur sera confronté à ces questions et, plus souvent qu'autrement, à la suite de l'évaluation de l'évaluabilité du programme, il devra proposer aux responsables du programme un projet de développement et d'amélioration du programme afin de le rendre évaluable.

Annexe

## **Évaluation d'un programme de réadaptation en centre jeunesse – exemple d'un cas d'application de l'EEP**

---

Les autorités d'un centre jeunesse du Québec demandent à une équipe d'évaluateurs de l'Université du Québec à Trois-Rivières de réaliser l'évaluation des effets de leur programme intitulé Programme d'intervention en milieu scolaire et familial.

Le premier constat des évaluateurs est qu'il n'y a pas d'écrit précis et développé sur le programme à évaluer. Ce programme existait dans les faits depuis plusieurs années mais, comme c'est le cas dans de nombreux milieux d'intervention, il n'y avait pas de documents qui en décrivaient les principales composantes. Ce programme n'était donc pas évaluable, du moins quant à ses effets.

Les évaluateurs ont donc procédé à une activité de préévaluation qui consistait principalement à décrire l'ensemble du programme d'intervention. Cette opération s'est réalisée par des entretiens avec le responsable du programme, les principaux intervenants ainsi qu'avec le directeur du service clinique du centre jeunesse de qui relevait le programme. De plus, une analyse de documents divers en rapport avec le programme fut effectuée.

Le modèle de référence qui a servi de support à la description du programme de réadaptation en milieu scolaire et familial est le modèle de Samson-Saulnier (1990, 1992). Ce modèle propose des composantes pouvant servir à l'élaboration d'un programme ou à l'évaluation d'un programme existant. Ces composantes sont : les intrants qui font référence à la mission et à la structure du programme ; la transformation qui concerne le processus, soit les activités cliniques et les services prévus par le programme ; les extrants, soit la production représentée par les services offerts par le programme ; et les conséquences qui font référence aux résultats du programme.

La problématique de la clientèle pour qui était fait le programme a d'abord été décrite. Cette présentation s'appuyait sur une recension des écrits et sur les connaissances expérientielles des responsables du programme. La mission du centre jeunesse responsable du programme, les valeurs prônées dans le milieu, de même que l'ensemble des autres services du centre ont également été présentés.

Le programme comme tel a été décrit et opérationnalisé avec le modèle de Samson-Saulnier (1990, 1992). Cela a donné lieu à préciser les objectifs du programme, l'approche utilisée, les principes d'intervention, la clientèle cible participante au programme, la structure du programme qui concerne

son organisation et les ressources humaines, physiques et matérielles, les activités cliniques et les outils de soutien à l'intervention (comme le Plan individualisé d'intervention).

Les objectifs et les attentes de résultats ont fait l'objet d'une analyse plus particulière afin de les rendre mesurables et d'identifier les outils de collecte de données et de mesure qui permettraient de vérifier les effets du programme.

Durant cette période de description du programme, le plan d'évaluation des effets (sommatif) du programme était développé en s'assurant de prendre en compte le programme, sa clientèle cible, ses objectifs, ses attentes de résultats et le protocole d'évaluation.

En conclusion, cette opération de préévaluation et d'évaluation de l'évaluabilité du programme nous a mené à devoir développer à fond la description du programme et à bien préciser les attentes de résultats qu'on devait mesurer.

Le rapport d'évaluation de ce programme est accessible sur le site Internet de Pierre Potvin, <[www.uqtr.ca/~potvin](http://www.uqtr.ca/~potvin)>, dans le menu Publications communications – Rapports de recherche VJSD.

## BIBLIOGRAPHIE

- LECOMTE, R. et L. RUTMAN (1982). *Introduction aux méthodes de recherche évaluative*, Québec, Presses de l'Université Laval.
- LOISELLE, J. (2001). «La recherche-développement en éducation: sa nature et ses caractéristiques», dans M. Anadón, *Nouvelles dynamiques de recherche en éducation*, Québec, Presses de l'Université du Québec, p. 77-97.
- MONETTE, M., M. CHARRETTE et I. JOBIN (1998). *L'élaboration du programme*, Cap-Rouge, Presses Inter Universitaires.
- NADEAU, M.-A. (1988). *L'évaluation de programme. Théorie et pratique*, Québec, Les Presses de l'Université Laval.
- PAQUETTE, F et F. CHAGNON (2000). *Cadre de référence pour le développement et l'évaluation des programmes aux centres jeunesse de Montréal*, Montréal, Les centres jeunesse de Montréal, coll. «Développement et évaluation de programmes».
- PINEAULT, R. et C. DAVELUY (1995). *La planification de la santé. Concepts – méthodes – stratégies*, Montréal, Éditions nouvelles.
- PLANTE, J. (1994). *Évaluation de programme*, Notes de cours, Québec, Université Laval.
- ROSSI, P.-H. et M.W. LIPSEY (1999). *Evaluation. A Systematic Approach*, 6<sup>e</sup> éd., Newbury Park, Sage Publications.

- SAMSON-SAULNIER, G. (1990). *L'évaluation de programme*, Cahier du participant lors de sessions de formation données à l'Association des centres d'accueil du Québec.
- SAMSON-SAULNIER, G. (1992). *Évaluation de programme: sessions de formation sur l'évaluation de programme, manuel du participant*, Montréal, Association des centres d'accueil du Québec.
- TARD, C. et al. (1997). *Outils d'évaluation pour l'amélioration de la qualité des programmes*, Centre de recherche sur les services communautaires, Québec, Université Laval, Association des centres jeunesse du Québec.
- TREVISAN, M.-S. (2007). «Evaluability assessment from 1986 to 2006», *American Journal of Evaluation*, vol. 28, n° 3, p. 290-303.

# PARTIE 3

---

## **DE L'IMPLANTATION D'UN PROGRAMME D'INTERVENTION À LA MESURE DES PREMIERS EFFETS PRODUITS, LA PHASE PENDANT**

6. Les dimensions formative et sommative de l'évaluation d'implantation d'un programme
7. Conjuguer l'évaluation aux temps de l'implantation et des impacts



# CHAPITRE 6

---

## **Les dimensions formative et sommative de l'évaluation d'implantation d'un programme**

Une combinaison des perspectives  
objective et subjective en lien  
avec les modèles d'évaluation  
basés sur la théorie  
des programmes

Jacques Joly  
Département de psychoéducation  
Université de Sherbrooke

Luc Touchette  
Département de psychoéducation  
Université de Sherbrooke

Robert Pauzé  
Département de psychoéducation  
Université de Sherbrooke

La phase Pendant, du modèle intégrateur proposé dans cet ouvrage, qui correspond à la période d'implantation d'un programme, est certainement une phase fascinante du point de vue de l'évaluateur, mais aussi une phase exigeante et parfois difficile pour les personnes qui implantent le programme et pour l'ensemble des parties prenantes (*stakeholders*) confrontées aux dures réalités de la programmation au quotidien. Chose certaine, cette période fait de plus en plus l'objet de travaux de recherche qui visent à proposer des stratégies permettant d'améliorer les programmes et de mieux comprendre comment se déroule leur implantation. De plus, ces études cherchent de plus en plus à identifier des conditions ou des facteurs qui favorisent ou nuisent à l'implantation de programmes de qualité.

On s'intéresse aussi depuis déjà un certain temps aux liens entre cette phase et les autres phases, soit l'Avant (conception et planification des programmes) et l'Après (les résultats des programmes). Il existe en effet plusieurs études qui suggèrent que les résultats d'un programme sont directement reliés à ce qui se passe lors de l'implantation du programme, notamment sa fidélité d'implantation (Dumas *et al.*, 2001; Dusenbury *et al.*, 2003; Kam *et al.*, 2003; Noell, 2008, dans Burns *et al.*, 2008; Ringwalt *et al.*, 2003). Par contre, le lien entre la phase Avant et la phase Pendant demeure peu étudié, même si plusieurs auteurs mentionnent que les caractéristiques mêmes d'un programme, comme sa complexité, ont un effet sur son implantation.

De plus, nos expériences passées et actuelles nous permettent de croire que l'évaluateur peut jouer un rôle clé pendant l'implantation d'un programme, et peut devenir un collaborateur essentiel pour résoudre avec les parties prenantes une foule de contraintes. Nous verrons d'ailleurs comment plusieurs stratégies s'offrent aux évaluateurs et aux parties prenantes afin de travailler conjointement au succès de l'implantation d'un programme.

Dans ce chapitre, nous abordons d'abord les raisons qui motivent les évaluations d'implantation, les dimensions abordées lors des évaluations d'implantation, les approches et méthodes utilisées pour ce type d'évaluation, et nous proposerons un exemple appliqué au contexte d'un programme particulier, comme le programme Crise-Ado-Famille-Enfance (CAFE).

## 1. POURQUOI L'ÉVALUATION DURANT LA PHASE PENDANT

Selon le modèle proposé dans le cadre de cet ouvrage, la phase Pendant suit l'Avant et précède l'Après. Trivial, pensez-vous? Sachez que, pendant très longtemps, l'évaluation de programme s'est surtout concentrée à évaluer les résultats des programmes en utilisant des méthodologies suivant le modèle de la boîte noire (*black box evaluation*). S'inspirant de la méthodologie

expérimentale, à l'aide d'une méthodologie la plus rigoureuse possible, expérimentale ou quasi expérimentale, les évaluateurs comparaient alors les résultats en termes d'atteinte des objectifs de tous les participants au programme (groupe expérimental) à tous les non-participants ou participants à un autre programme (groupe témoin ou contrôle) sans tenir compte de l'implantation du programme. Des différences significatives dans la direction souhaitée étaient alors interprétées comme indicateurs de l'efficacité ou de l'incidence du programme. Par contre, si les résultats ne correspondaient pas à ceux escomptés (ou étaient non significatifs), on concluait plutôt que le programme n'était pas efficace. Les modèles d'évaluation ont bien évolué depuis ce temps : plus d'importance est accordée aux phases Avant et Pendant, le rôle de l'évaluateur a été remis en question, la place des parties prenantes dans le processus d'évaluation a été discuté de long en large et les méthodologies appropriées (quantitatives, mixtes ou qualitatives) ont fait l'objet de longs débats (Alkin et Christie, 2004).

Si nous nous plaçons dans une perspective plus moderne de la théorie des programmes, nous constatons que la pratique traditionnelle de l'évaluation amenait les évaluateurs d'alors à conclure à l'échec de la théorie du programme (*theory failure*) à la suite de l'absence d'effets significatifs attendus, alors qu'il était possible que l'échec en soit une d'implantation (*implementation failure*), l'absence d'effet étant possiblement reliée à une implantation lacunaire. Cette distinction entre échec de la théorie des programmes et échec d'implantation est très clairement exposée par Rossi *et al.* (2004, p. 79). Aujourd'hui, les principaux modèles d'évaluation (Chen, 2005; Rossi, Lipsey et Freeman, 2004, et Patton, 1997, pour ne citer que ceux-là) proposent d'étudier l'implantation des programmes et d'en tenir compte à la phase d'évaluation des résultats.

Deux termes sont souvent utilisés pour définir l'évaluation de programme pendant leur mise en œuvre : l'évaluation d'implantation ou l'évaluation de processus. Nous utiliserons dans ce chapitre le terme « évaluation d'implantation » pour faire référence à l'ensemble des approches, stratégies ou méthodes utilisées pour évaluer un programme pendant sa phase d'implantation. À l'occasion, nous ferons même une distinction entre implantation initiale et implantation mature. La première se situe, comme vous vous en doutez, au tout début de l'implantation, alors que la seconde se rapporte à une période plus longue qui s'étalera sur un certain temps selon la durée des interventions dans le programme, et qui pourra même chevaucher la phase des résultats, ou l'Après, selon le modèle présenté ici. Vous comprenez bien qu'un programme a une durée de vie, comme une personne, naît, vieillit et mourra éventuellement (voir, à cet égard, la contribution d'Allard, dans ce même ouvrage au chapitre 12). Et, comme chez une personne, les périodes d'enfance, d'adolescence et de maturité ne se terminent pas et

ne commencent pas une journée bien précise. Quoique, dans le cas d'un programme, il existe bien un jour où le premier intervenant rencontre son premier client, ce qui peut correspondre au début de l'implantation.

L'évaluation d'implantation joue deux rôles majeurs, selon Rossi *et al.* (2004, p. 175) : donner un feedback aux décideurs afin d'améliorer le plus rapidement possible le programme, ce qui correspond à sa dimension formative, et jouer un rôle complémentaire à l'évaluation des résultats ou de l'incidence du programme en permettant d'interpréter les résultats obtenus en fonction du programme réellement implanté, ce qui correspond davantage à sa mission sommative. L'évaluation d'implantation peut alors prendre deux formes : une activité réalisée en soi en dehors des activités courantes du programme, ou un monitoring continu incluant un système de gestion de l'information. De son côté, Chen (2005, p. 129) distingue entre l'évaluation d'implantation initiale et l'évaluation d'implantation mature. Il propose pas moins de huit approches possibles pour réaliser l'une ou l'autre. Lors de l'évaluation initiale, l'objectif est de fournir à temps de l'information sur les problèmes d'implantation pour aider les parties prenantes à les résoudre, ce qui rejoint le point de vue de Rossi *et al.* (2004) et la dimension formative du modèle proposé dans le présent livre. L'évaluation pendant la période plus mature du programme permet, entre autres choses, de porter un jugement sur la qualité de l'implantation du programme à des fins de reddition de comptes (*accountability*) ou à des fins d'analyse et de compréhension des processus afin de mieux interpréter les résultats, voire l'absence de résultats. Ainsi, l'évaluation de l'implantation mature démontre aussi son caractère sommatif. Mais il ne faut pas oublier que, même à une étape plus avancée, l'évaluation d'implantation sert à informer les parties prenantes des problèmes rencontrés afin de continuer à améliorer le programme.

Avant de passer aux aspects plus techniques de l'évaluation d'implantation, nous aimerions aussi distinguer programmes expérimentaux, développés et appliqués dans un contexte de recherche (*demonstration project*) et programmes diffusés à plus grandes échelles (*dissemination project*). Les premiers sont des programmes souvent conçus par des experts et mis en œuvre dans des milieux particuliers avec ou sans la contribution directe de ces milieux à la conception du programme. La tendance actuelle est toutefois de mettre à contribution des représentants de ces milieux non seulement à la conception, mais surtout à la planification de ces programmes. Très souvent, il s'agit de personnel cadre ou professionnel. Ces programmes font l'objet d'une grande attention, notamment au moment de l'implantation, de manière à s'assurer qu'ils sont implantés conformément au projet accepté. Le terme anglais *efficacy evaluation* est utilisé pour identifier les évaluations des résultats (*outcomes*) de ces programmes et en déduire leur efficacité. Un exemple de ce type de programme est CAFE, qui a été conçu en

collaboration avec une équipe de chercheurs de l'Université de Sherbrooke et plusieurs CLSC de la région de la Montérégie, et qui a fait l'objet autant d'une évaluation d'implantation que d'une évaluation des résultats (Joly *et al.*, 2005; Pauzé *et al.*, 2007)<sup>1</sup>.

Les seconds sont des programmes déjà existants que des personnes décident d'implanter dans un milieu donné. Par exemple, c'est le cas encore du programme CAFE, qui est implanté dans d'autres régions. C'est aussi le cas de multiples programmes dans tous les milieux d'intervention et de prévention, comme en milieu scolaire, où on plante régulièrement une panoplie de programmes de prévention ou d'intervention touchant des problématiques comme la consommation de drogues, la violence directe ou indirecte, l'abandon scolaire, le suicide et bien d'autres. Dans ces cas, ce sont des personnes du milieu, souvent des professionnels comme des psychoéducateurs, qui décident d'implanter ces programmes plus ou moins bien documentés, conçus par d'autres équipes ou experts plus ou moins disponibles pour les aider à entreprendre correctement de tels programmes. À la suite d'une recension des écrits portant sur des programmes de prévention de la violence en milieu scolaire, Dusenbury *et al.* (2003) ont observé que ces programmes diffusés à plus grande échelle sont souvent implantés dans des conditions imparfaites et s'avèrent très différents des programmes expérimentaux originaux jugés efficaces. Or on ne peut assumer que, dans de telles conditions, les programmes auront la même efficacité puisque les auteurs des grands modèles d'évaluation (Chen, 2005; Rossi *et al.*, 2004) et la littérature scientifique (Dumas *et al.*, 2001; Botvin *et al.*, 1990; Ringwalt *et al.*, 2003; Rohrbach *et al.*, 1993) soutiennent l'idée selon laquelle les programmes moins bien implantés sont moins efficaces. Le terme anglais *effectiveness evaluation* est plutôt utilisé pour identifier ces évaluations en milieu d'intervention.

La distinction entre les deux contextes d'implantation est importante car elle motive encore plus le recours à des évaluations d'implantation. En effet, comme le précisent si bien Rossi *et al.* (2004, p. 177), il n'est généralement pas sage de prendre pour acquis l'implantation du programme. Il est plutôt fréquent, dans un contexte de diffusion d'un programme, que celui-ci soit modifié, sans égard à la théorie d'action sous-jacente au programme, ou réduit, n'atteignant plus ainsi la quantité d'intervention nécessaire pour produire un effet. De plus, dans un contexte de diffusion, les ressources disponibles (financières, matérielles ou autres) sont souvent moindres que dans un contexte expérimental.

---

1. Ce programme existe toujours, mais a connu des modifications à la suite de son évaluation.

En bref, précisons que les avantages de l'évaluation d'implantation sont, selon Werner (2004, p. 10): 1) donner une rétroaction rapide aux gestionnaires, 2) fournir les informations nécessaires pendant la période formative du programme, 3) fournir de l'information contextuelle et ethnographique riche, 4) fournir de l'information sur le programme tel qu'il est et 5) fournir de l'information limitée mais stratégique où et lorsque c'est nécessaire. Mais, toujours selon Werner, il existe aussi des limites aux évaluations d'implantation: 1) elles ne fournissent pas d'information directe et précise sur les résultats ou le rapport coûts-bénéfices, et 2) elles proposent des jugements sur la base de données qualitatives et subjectives. Ce propos n'est pas nécessairement partagé par tous car, comme on l'a mentionné plus tôt, certains pensent plutôt que l'évaluation d'implantation peut être sommative et servir à des fins de reddition de comptes.

Nonobstant les limites possibles de l'évaluation d'implantation, il est dorénavant reconnu comme essentiel de procéder systématiquement à de telles évaluations pour tous les programmes. Nous verrons d'ailleurs que l'on peut s'acquitter de cette tâche même avec des ressources très limitées. Il faut juste savoir sur quoi porter son regard et profiter d'approches ou de méthodes efficaces et adaptées à ces ressources.

## **2. LES OBJETS DE L'ÉVALUATION D'IMPLANTATION : DIMENSIONS À OBSERVER LORS DE LA PHASE PENDANT**

Si l'évaluation de l'implantation vise à améliorer le programme (dimension formative) et à porter un jugement sur ce qui a été réalisé par rapport à ce qui était prévu (dimension sommative), alors quels sont les aspects qui doivent faire l'objet de l'évaluation? Pour répondre à cette question, nous nous plaçons encore dans la perspective de la théorie des programmes, qui est un modèle plutôt bien soutenu dans le domaine de l'évaluation des programmes (Donaldson, 2003; Chen, 2005; Rossi *et al.*, 2004). Cette perspective d'évaluation vise essentiellement à comprendre comment les programmes atteignent leurs objectifs ou pourquoi ils ne les atteignent pas, à identifier les composantes les plus efficaces, les mécanismes médiateurs et les caractéristiques des clients, des intervenants ou de l'environnement qui modèrent les mécanismes d'action. D'un côté, selon cette perspective, le plan d'action du programme, né de l'étape de la conception et de la planification, guide l'évaluation de l'implantation du programme. D'un autre côté, la théorie d'action du programme, ou le modèle de changement qui établit un lien entre les interventions, les déterminants (ou variables proximales) et les résultats

(variables distales), guide l'évaluation des résultats du programme ou son incidence. Bien évidemment, nous nous concentrons ici sur l'évaluation du plan d'action.

Comme vous l'aurez remarqué dans des chapitres précédents, il est souvent d'usage en évaluation de programme d'établir les objets d'évaluation par la formulation de questions d'évaluation. Dans le cas d'une évaluation d'implantation, Rossi *et al.* (2004, p. 171) proposent une liste de 17 questions générales d'évaluation, sans compter les sous-questions, qui permettent de répondre aux deux questions principales de l'évaluation d'implantation : 1) est-ce que le programme atteint sa clientèle cible ? et 2) est-ce que la prestation de service et les fonctions de soutien sont cohérentes avec le plan du programme (ou son devis) ou avec tout autre standard approprié ? Voici quelques exemples de questions d'évaluations proposées :

- Combien de personnes reçoivent les services ?
- Est-ce que celles qui reçoivent les services sont celles ciblées par le programme ?
- Est-ce qu'elles reçoivent la quantité, le type et la qualité appropriés de services ?
- Est-ce qu'il y a une partie de la clientèle cible qui ne reçoit pas les services ?
- Est-ce que le personnel est en nombre suffisant et compétent pour offrir le service ?
- Est-ce que le programme est bien organisé ? Est-ce que le personnel travaille bien l'un avec l'autre ?
- Est-ce que le programme est coordonné efficacement avec les autres programmes ou agences avec lequel il interagit ?
- Etc.

Werner propose aussi une série de questions et de sous-questions d'évaluation assez précises qui peuvent guider l'évaluation d'implantation (Werner, 2004, p. 16). Il propose tout d'abord une série de questions qui concernent l'implantation qui doivent être posées et répondues avant même le début de l'implantation. Par exemple, quels sont les types et les niveaux de ressources nécessaires pour implanter le programme tel que planifié ? Est-ce que le programme est adapté à son environnement ? Cette dernière question peut d'ailleurs être posée à la fin de la phase Avant, toujours à l'étape de la planification, mais elle peut aussi être reprise à la phase Pendant, lors de l'implantation du programme et surtout lors de son implantation initiale. Voici quelques autres exemples de questions et sous-questions proposées par Werner :

- Est-ce que les ressources pour opérer le programme sont en place et, si non, comment et pourquoi ?
  - Est-ce que toutes les ressources sont disponibles ?
    - Le nombre, la localisation et les infrastructures des bureaux, des points de contacts. Le nombre et le type de personnel de gestion et d'intervention pour soutenir la demande (*caseload*).
    - Les systèmes d'information nécessaires pour la gestion des cas et du programme.
  - Quelles sont les conséquences du niveau de ressources disponible sur les opérations du programme ?
- Est-ce que les processus et systèmes fonctionnent tels que planifiés et, si non, comment et pourquoi ?
  - Est-ce que le processus de référence fonctionne comme prévu ?
    - Est-ce que le personnel applique les critères d'éligibilité comme prévu ?
  - Est-ce que le personnel utilise les nouvelles politiques et procédures ?
    - Est-ce que les personnes qui demandent ou reçoivent des services sont référées aux bons services ?
  - Est-ce que les personnes qui demandent ou reçoivent des services participent aux différentes activités ou services ?
    - Est-ce qu'elles ont le soutien nécessaire pour le faire ?
  - Comment et pourquoi les différents processus ou systèmes ne fonctionnent pas comme prévu ?
- Est-ce que le programme atteint la population cible, avec les services appropriés, en quantité prévue et, si non, comment et pourquoi ?
  - Quel est le taux de participation aux activités ?
  - Quelles activités sont proposées aux clients ou réalisées ?
    - Combien ?
    - Quelle durée ?
    - Qualité de la prestation ou du contenu ?
    - Délais d'attente ?
    - Abandons ?
  - Quels sont les services connexes qui peuvent aider le client à participer ?

Évidemment, toutes ces questions doivent être adaptées au contexte particulier du programme qui fait l'objet de l'évaluation. De plus, ce ne sont que des exemples de question, mais on comprend que la clientèle rejointe,

les interventions réalisées, les conditions pour les réaliser, la quantité et la qualité de celles qui sont réalisées, les difficultés rencontrées sont autant de dimensions au cœur de l'évaluation d'implantation. Enfin, le nombre de questions est aussi un élément à considérer car, comme on le verra plus loin, il faudra s'assurer que l'on répond aux questions posées avec des méthodes adéquates mais, plus il y aura de questions, plus la collecte de données sera exigeante et coûteuse. Par conséquent ce n'est pas tant le nombre de questions qui est important dans une démarche d'évaluation d'implantation que la pertinence des questions posées pour atteindre nos objectifs d'améliorer le programme et de se prononcer sur la qualité de son implantation.

De son côté, Chen (2005) ne propose pas de listes de questions. Il suggère plutôt plusieurs approches et stratégies qui permettent d'évaluer un programme dans sa phase d'implantation initiale et dans sa phase d'implantation mature. Nous y reviendrons à la section suivante. À cette étape c'est le plan d'action élaboré à l'étape de conception et de planification qui sert de grille à l'évaluation. Peu importe l'approche particulière choisie, le regard se portera toujours sur l'une ou l'autre, ou sur l'ensemble des composantes du plan d'action du programme. Selon Chen, il y a six composantes dans un plan d'action :

1. La clientèle ciblée : incluant les critères d'inclusion bien définis de manière à éviter les problèmes de surcouverture (rejoindre des personnes non ciblées) ou des problèmes *de sous-couverture* (ne pas rejoindre les personnes ciblées) (voir Rossi *et al.*, 2004, p. 185, à ce sujet).
2. Les protocoles d'intervention et de services : description détaillée des activités, des méthodes d'intervention, de l'organisation du service, etc. Idéalement expliqués dans des manuels ou de la documentation adéquate.
3. Les personnes qui implantent (qui offrent le service) : leurs compétences techniques et culturelles (au sens de culture du milieu), leur implication et leurs qualifications. Sans oublier la formation et la supervision nécessaires.
4. L'organisation qui implante le programme : capacité à offrir l'expertise technique et culturelle et les ressources nécessaires pour planter le programme.
5. Les partenaires : on peut distinguer ces partenaires en fonction de leurs liens avec le programme. Certains jouent un rôle direct sans lesquels on ne peut offrir le programme, d'autres un rôle indirect (comme un autre service dans la même organisation), et d'autres un rôle éloigné (comme une autre organisation ou institution qui peut offrir d'autres services à la même clientèle).

6. Le contexte écologique: tout autre élément, dans l'organisation ou à l'extérieur, qui ne fait pas partie du programme comme tel mais qui peut influencer sur le bon déroulement du programme.

Ces éléments sont utiles pour planifier une évaluation d'implantation dans la mesure où ils nous aident à avoir un regard d'ensemble sur le programme. Nous avons donc tenté de mettre en relation les différentes questions d'évaluation avec les composantes du plan d'action, mais en établissant du même coup la relation avec le modèle du cube. Ainsi, le tableau 6.1 présente une série de questions d'évaluation à titre d'exemples, par composante du plan d'action que l'on devrait retrouver pour un programme, en distinguant entre les dimensions formative et sommative, subjective et objective. Nous nous sommes permis de suggérer que la dimension formative est surtout dominante lors de l'implantation initiale, alors que la dimension sommative l'est davantage lors de l'implantation mature. Nous sommes toutefois conscients que ces deux périodes se chevauchent, et nous ne limitons pas nécessairement la dimension formative à la seule période d'implantation initiale. Enfin, précisons que les questions ont été libellées de manière à pouvoir s'appliquer au programme CAFE que nous utilisons comme exemple à la section 4.

Tableau 6.1

### Exemples de questions d'évaluation illustrant les dimensions évaluées lors d'une évaluation d'implantation en fonction des composantes d'un programme

| Composantes du plan d'action    | Formative (Surtout en implantation initiale)  | Sommative (Surtout en implantation mature)  |
|---------------------------------|---|---|
| <b>Clientèle</b>                |   |   |
| Objective                       | Quelle est la nature de la crise vécue par les jeunes et leur famille?  | Quelles sont les caractéristiques des familles desservies par le programme?                         |
| Subjective                      | Que pensent les intervenants de la participation au programme des jeunes et des familles?   | Est-ce que la clientèle référée au programme correspond à la clientèle visée par le programme?      |
| <b>Processus d'intervention</b> |   |   |
| Objective                       | Quelles sont la rapidité, la durée et l'intensité des interventions réalisées auprès des premiers clients?<br>Est-ce que les intervenants appliquent fidèlement le programme? | Quel pourcentage de clients a reçu une intervention rapide, brève, intensive et dans la communauté? |

Tableau 6.1 (suite)

| <b>Composantes du plan d'action</b> | <b>Formative (Surtout en implantation initiale)</b>   | <b>Sommative (Surtout en implantation mature)</b>  |
|-------------------------------------|---|--|
| Subjective                          | Est-ce que certaines composantes du processus d'intervention posent problème? Lesquelles?         | Est-ce que le protocole d'évaluation a été utile pour la planification des interventions? Comment?                     |
| <b>Personnes qui implantent</b>     |   |  |
| Objective                           | Comment participent les intervenants aux supervisions?  | Quel est l'indice de fidélité de l'implantation de chaque intervenant?   |
| Subjective                          | Quel est le sentiment de compétence des intervenants dans ce programme?                           | Est-ce que les attitudes des intervenants favorisent l'implantation du programme? En quoi?                             |
| <b>Organisation</b>                 |   |  |
| Objective                           | Quels aménagements aux horaires et aux charges de travail ont été réalisés par l'organisation?    | Combien d'heures de formation et de supervision ont été offertes aux intervenants?                                     |
| Subjective                          | Quelle est l'opinion des intervenants à l'égard des manuels proposés pour appliquer le programme? | Quelle est l'opinion des intervenants à l'égard du soutien de la direction?  |
| <b>Partenaires</b>                  |   |  |
| Objective                           | Quels sont les protocoles d'arrimage avec les partenaires?  | Combien de dossiers ont été référés par ou vers les partenaires?   |
| Subjective                          | Comment perçoit-on la collaboration entre partenaires?  | Est-ce que les partenaires ont facilité (ou nui) à l'implantation du programme?  |
| <b>Contexte écologique</b>          |   |  |
| Objective                           | Est-ce que des éléments hors du programme nuisent à l'implantation de celui-ci? Lesquels?         | Est-ce que des contraintes, à l'extérieur du programme, ont empêché d'implanter correctement le programme? Lesquelles? |
| Subjective                          | Quelles sont les difficultés rencontrées dans l'implantation du programme?                        | Est-ce que le contexte organisationnel et le contexte social favorise l'implantation du programme? Comment?            |

S'il fallait tenter de distinguer entre implantation initiale et implantation mature, nous devrions tenir compte de la durée des interventions. Dans un programme où l'intervention dure environ un mois, la période d'implantation initiale n'a forcément pas la même durée que dans un programme dont

l'intervention s'étend sur un an. Nous pourrions suggérer, à titre de règle du pouce (*rule of thumb*), que l'implantation initiale dure au moins tout un cycle d'intervention pour les premiers clients. Après quelques corrections, commencerait alors la période d'implantation mature.

Armés d'une série de questions potentielles, il nous reste à déterminer une approche et des méthodes de collecte de données qui permettent d'atteindre nos objectifs d'évaluation. C'est ce que nous proposons dans la section suivante. Car c'est bien ce qui est au cœur de l'évaluation basée sur la théorie des programmes : formuler, prioriser et répondre à d'importantes questions d'évaluation (Donaldson, 2003, p. 115)

### **3. LES APPROCHES, STRATÉGIES ET MÉTHODES DE COLLECTE DE DONNÉES**

Une bonne compréhension des principes à la base de l'évaluation d'implantation ne se traduit pas nécessairement par une habileté à mettre en place et à réaliser une telle évaluation. N'est-ce pas là d'ailleurs la distinction entre connaissances et compétences ? Il faut aussi avoir recours à des approches ou des méthodes qui nous permettent de se doter d'un plan d'attaque pratique et réaliste. Dans cette section, nous présentons quelques approches possibles qui, utilisées conjointement, permettront de répondre à une gamme variée de questions, subjectives et objectives, formatives et sommatives. Nous verrons aussi qu'il faut faire appels à diverses méthodes de collecte de données afin de produire des rapports d'évaluation utiles. Toutefois, la diversité des méthodes a ses avantages mais aussi ses inconvénients. Enfin, nous aborderons quelques autres considérations importantes lors d'une évaluation d'implantation.

#### **3.1. Les approches**

Rossi *et al.* (2004, p. 175) reconnaissent qu'il existe une diversité d'approches, qui portent différents noms. Elles peuvent être réalisées sporadiquement ou en mode continu, conduites par un évaluateur interne ou externe ou même sans l'aide d'un expert, et visant des objectifs tout aussi variés, comme on l'a vu précédemment. Ils classent toutefois ces approches en deux grandes catégories : 1) les évaluations d'implantation ou de processus typiquement conduites par des spécialistes, avec ou sans la participation du personnel et ne faisant pas partie des opérations courantes du programme, et 2) le monitoring continu intégré aux activités régulières, souvent sur une base client par client, portant sur des indicateurs particuliers du programme.

Werner (2004, p. 81) propose plutôt des méthodes de collectes de données qui visent à raconter « l'histoire du programme » et à répondre à la grande question « Que se passe-t-il dans ce programme ? ». Cette question a deux composantes majeures : une non quantitative (le mot est bien de Werner), et une quantitative. Nous reviendrons aux méthodes à la section suivante.

Comme nous l'avons mentionné plus tôt, Chen (2005) propose huit approches différentes pour l'évaluation d'implantation initiale ou mature<sup>2</sup>. Nous présentons ici, brièvement, cinq de ses approches. Les autres nous paraissent moins intéressantes et peu utilisées. La première porte le nom d'évaluation formative. Elle peut être utilisée autant en implantation initiale que mature. Il s'agit d'une approche qui utilise des méthodes de recherche flexibles (souvent en relâchant même les critères de rigueur) pour identifier les barrières ou les éléments facilitateurs de l'implantation afin de corriger les problèmes rapidement. Elle est réalisée en six étapes : 1) révision de la documentation, 2) identification des éléments cruciaux au succès de l'implantation, 3) sélection des méthodes de collecte de données, 4) identification des problèmes, 5) recherche des sources des problèmes et aide à l'identification des solutions, 6) soumission des résultats aux parties prenantes et documentation des changements apportés au programme (pour évaluation future). Chen propose quatre types de recherches formatives : 1) l'observation sur le site et la vérification (l'évaluateur participe au programme), 2) les groupes de discussion, 3) les entrevues intensives, et 4) le balayage complet (*comprehensive scanning*), qui fait appel à une panoplie de méthodes de recherche (analyse de dossiers, entrevues, conférences téléphoniques, courriels, visites de site, enquêtes et autres) afin de vérifier plusieurs composantes du plan d'action, mais toujours dans une perspective de regard rapide. Chen précise bien que les résultats de ces recherches formatives doivent être utilisés avec prudence et qu'ils n'attestent pas de la qualité du programme. Ils ne servent qu'à fournir un feedback rapide pour améliorer le programme le plus tôt possible.

La seconde est la « revue de programme/réunion de développement » (*program review/developpement meeting*). Selon cette approche, l'évaluateur anime des réunions auxquelles assistent gestionnaires, intervenants et autres parties prenantes dans le but de discuter des défis auxquels fait face

---

2. De fait, chez Chen, ces approches sont imbriquées dans des stratégies. Alors qu'une stratégie est une orientation générale que se donnent les évaluateurs et les parties prenantes en lien avec les objectifs de l'évaluation, les approches sont des ensembles systématiques de procédures et de principes qui guident la planification et la réalisation de l'évaluation (Chen, 2005, p. 51). Malheureusement, nous avons observé dans nos cours que ce système de classification des approches et stratégies, selon les phases du programme, crée passablement de confusion chez les lecteurs moins avertis. De là vient notre décision de simplifier ici la présentation en n'abordant que ces approches.

le programme. Idéalement, l'évaluateur est externe (assurant sa neutralité) et favorise une expression libre des opinions sur le programme. Les thèmes abordés concernent évidemment les composantes du plan d'action : clientèle cible, protocoles d'intervention, organisation, partenaires, personnes qui implantent et contexte écologique. Les grands avantages de cette approche sont les coûts très faibles, un sens plus grand de l'appropriation par les intervenants, et un feedback très rapide. Par contre, les inconvénients sont nombreux : les intrants sont essentiellement basés sur des impressions, l'approche favorise la création d'un consensus mais pas la précision, et les personnes plus articulées peuvent dominer les réunions. Il en découle que cette approche ne peut être utilisée à des fins de reddition de comptes ou afin de se prononcer sur la qualité d'implantation.

Une troisième approche, parmi celles préconisées par Chen (2005), est l'approche d'évaluation de la fidélité. Elle consiste essentiellement à évaluer dans quelle mesure le programme est implanté comme prévu. Selon Chen, elle vise surtout à évaluer si certaines composantes du plan d'action ont été implantées de manière conforme au plan : clientèle rejointe, processus de référence, protocole d'intervention, prestation de services et autres aspects. C'est une approche nettement plus sommative que les deux précédentes, et elle fait appel à des évaluations plus rigoureuses afin de porter un jugement sur la qualité d'implantation. C'est aussi une approche adéquate dans une optique de reddition de comptes, et elle permet de prendre en compte des indicateurs de l'implantation lors de l'évaluation des effets du programme. En effet, il est alors possible d'introduire les indicateurs de fidélité à titre de variables modératrices dans les analyses des résultats du programme. La fidélité d'implantation des programmes est en soi un sujet de recherche actuel qui attire l'attention de nombreux chercheurs (Burns *et al.*, 2008 ; Crea *et al.*, 2008 ; Dumas *et al.*, 2001 ; Dusenbury *et al.*, 2003 ; Joly *et al.*, 2005 ; Ringwalt *et al.*, 2003 ; McGrew et Griss, 2005).

La quatrième approche, tirée de Chen (2005), est aussi présentée de manière très détaillée par Rossi *et al.* (2004). Il s'agit du monitoring. Au moins trois sortes d'informations devraient faire partie d'un système de monitoring : 1) les caractéristiques des clients, 2) les données cliniques sur les clients, et 3) les interventions effectuées auprès des clients. Le monitoring se distingue des autres approches par le fait que l'éventail des données recueillies dans le monitoring est beaucoup plus étroit, la profondeur des données, moindre, mais l'étendue dans le temps, plus grande. Cette approche peut s'avérer particulièrement appropriée pour documenter les processus cliniques à implanter dans certains programmes d'intervention.

La cinquième et dernière approche présentée est celle identifiée par Chen (2005) comme étant l'évaluation d'implantation basée sur la théorie des programmes (*theory-driven process evaluation*). Même si toutes les autres

sont aussi orientées par la théorie des programmes, car elles s'inspirent du plan d'action, cette dernière approche est par nature plus holistique. Elle vise en effet à évaluer systématiquement toutes les composantes du plan d'action à l'aide d'une variété de méthodes. Le produit final doit être un jugement basé sur une comparaison systématique du programme implanté et du programme prévu dans le plan d'action, concernant toutes ses composantes. Cette approche est très près de l'approche de la fidélité. On pourrait même affirmer qu'elles sont identiques si l'approche de fidélité était appliquée à toutes les composantes du programme. Par contre, cette dernière approche permet en plus, à cause de son caractère holistique, de détecter tous les effets inattendus du programme, positifs ou négatifs.

L'évaluateur est donc placé devant une diversité d'approches. Elles peuvent être utilisées seules ou en combinaison selon les besoins d'information et les ressources disponibles. Dans le cas où on dispose de peu de ressources, l'approche « revue de programme/réunion de développement » peut être très justifiée. En tout cas bien meilleure que l'absence d'évaluation d'implantation. Dans d'autres cas, où on dispose de plus de ressources, la dernière approche, plus holistique, est nettement préférable dans une perspective sommative. Jointe à une évaluation d'implantation formative (la première présentée), elles offrent une assurance qualité très élevée si les collectes de données sont bien réalisées et les rapports, produits à temps.

En définitive, ce sont des négociations entre évaluateur et parties prenantes, idéalement avant que ne commence l'implantation, qui guidera le choix de l'une ou l'autre des approches. Un bon moyen est de procéder d'abord à une discussion franche et ouverte à propos des questions d'évaluation. Cette discussion devrait permettre d'en prioriser un certain nombre, comme le démontre le tableau 6.1.

### 3.2. Les méthodes

Nous nous sommes beaucoup inspirés de Chen (2005) dans la section précédente. Un évaluateur qui voudrait s'inspirer strictement de Chen pour proposer un plan d'évaluation d'implantation aurait vite quelques difficultés, car celui-ci aborde très peu les questions méthodologiques. Il présuppose que le lecteur est familier avec les différentes méthodes de collecte de données. D'ailleurs, plusieurs livres de référence en évaluation de programmes abordent peu les aspects méthodologiques. Il faut se référer à des ouvrages de références en méthodologie scientifique.

Le livre de Werner (2004) est déjà plus intéressant à cet égard, et beaucoup plus utile. Son chapitre 2 (p. 27-80) est entièrement dédié aux méthodes de collecte de données dans un contexte d'évaluation d'implantation. Il y propose diverses méthodes, comme :

- L'analyse documentaire du matériel écrit: minutes des réunions de développement, lois régissant le programme, budgets, manuels du programme, matériel d'information (brochures), données statistiques ou administratives, rapports de recherche, etc.
- L'analyse des données administratives: caractéristiques des clients et utilisation des services contenues dans les dossiers papiers ou informatisés, banques de données sur la clientèle.
- Les entrevues: opinions des personnes directement concernées par le programme, incluant les clients. Cette méthode peut être rapide et peu coûteuse. Elle peut être réalisée en groupe. Permet d'obtenir des opinions sur comment les choses se passent, pourquoi elles se passent ainsi et comment améliorer le programme.
- L'observation: l'évaluateur observe directement (ou indirectement à l'aide d'un enregistrement sonore ou vidéo) le déroulement d'une entrevue, l'animation d'un groupe ou le fonctionnement de réunions de travail, selon les activités prévues par le programme.
- Les enquêtes ou questionnaires: données systématiques en provenance d'un grand nombre de répondants, incluant les intervenants ou les clientèles. Peuvent comprendre des opinions sur le programme (une ou plusieurs composantes), des connaissances sur les politiques ou les services, des questions de compréhension sur la théorie du programme (ou sa philosophie), des caractéristiques des clientèles, des indicateurs d'utilisation des services, des indicateurs de résultats (certains questionnaires peuvent être des tests standardisés et validés), des listes de vérification et même des indicateurs de satisfaction.

Nous remarquons que toute une panoplie de méthodes de collecte de données est disponible. Desrosiers *et al.* (1998, p. 65) proposent d'ailleurs un tableau synthèse des diverses méthodes de collecte de données utiles en évaluation de programme ainsi que leurs avantages et inconvénients. Ils suggèrent (p. 63) que l'étude de dossiers, les analyses de bases de données (analogue au monitoring), le journal de bord, l'observation sur le terrain ou l'observation participante, les entrevues individuelles auprès d'informateurs clés ainsi que des questionnaires d'évaluation et de satisfaction par les participants sont les méthodes les plus adaptées à l'évaluation d'implantation.

Le choix des méthodes sera donc fonction, dans l'ordre, des questions d'évaluation prioritaires, d'une ou des approches choisies, des ressources disponibles pour réaliser l'évaluation et, forcément, de l'ampleur du programme. Un programme national à la grandeur du pays ne commande pas les mêmes approches et les mêmes méthodes de collecte de données qu'un programme implanté à plus petite échelle. On remarque tout de même une grande diversité de méthodes qui sont proposées. Rossi *et al.* (2004,

p. 16) définissent d'ailleurs l'évaluation de programme comme étant « l'utilisation des méthodes de recherche sociale afin d'étudier systématiquement l'efficacité des programmes d'intervention sociale ». Dans le domaine de l'évaluation, et plus particulièrement des évaluations d'implantation, il y a un assez large consensus pour combiner plusieurs méthodes afin de répondre aux diverses questions d'évaluation. On peut donc affirmer que le paradigme des recherches mixtes est plutôt bien accepté dans ce domaine.

Il y a forcément des avantages à cette multiplicité des approches et méthodes. On y trouve son compte en fonction des ressources disponibles. Il est aussi possible de choisir ses méthodes dans une optique de triangulation afin d'augmenter la validité interne de notre évaluation, tout en cherchant à maintenir optimale la validité externe. Certaines méthodes couplées à certaines approches, comme des entrevues de groupe dans une approche de recherche formative, peuvent être plus utiles pour identifier rapidement certains problèmes de fonctionnement, mais ne permettent pas de porter un jugement valide sur la qualité du programme. De fait, la dimension formative est mieux servie par ces méthodes que la dimension sommative. Mais c'est aussi ce que nous désirons en début d'implantation.

Mais il y a aussi plusieurs inconvénients. On a fait état plus tôt des limites aux interprétations de résultats en raison de la rigueur scientifique moindre de certaines méthodes. Il est aussi très important de souligner que la multiplicité des méthodes peut aussi poser plusieurs problèmes logistiques. Plus il y a de méthodes de collectes de données, plus la pression est grande sur les intervenants. Le temps consacré à répondre aux questionnaires, passer les entrevues, être observé, sans compter l'impression d'être constamment évalué, peut sembler trop exigeant pour eux. Il faut donc savoir doser. Enfin, Bryman (2007) a fait état des difficultés inhérentes aux approches mixtes, par exemple, le fait qu'un chercheur va maintenir une préférence pour une méthode sur une autre et y donner plus d'importance, le fait qu'il est difficile de coordonner les diverses méthodes dans le temps, ou encore le fait que les différentes méthodes demandent différentes habiletés de la part des chercheurs. Ceux-ci ne peuvent être aussi habiles avec toutes les méthodes.

Nous donnerons un exemple à la section suivante de la combinaison de méthodes et approches pour évaluer l'implantation d'un programme de crise.

### 3.3. Autres considérations

On ne peut terminer cette section sans aborder quelques autres considérations importantes dans la planification et la réalisation d'une évaluation d'implantation. Parmi celles-ci, mentionnons le rôle de l'évaluateur, la relation entre l'évaluateur et les parties prenantes, l'utilisation des standards et la production des rapports.

La question du rôle de l'évaluateur pourrait faire l'objet non pas d'un chapitre mais d'un livre en soi. En effet, le rôle de l'évaluateur est un des éléments au cœur des distinctions et des débats entre les différents modèles ou théories d'évaluation<sup>3</sup>. C'est même un des aspects des théories d'évaluation que Alkin et Christie (2004) ont utilisé pour classer ces modèles ou théories dans un arbre qui les distingue (les branches) à partir de leurs racines communes. Résumer tous les différents rôles possibles c'est faire l'histoire des théories d'évaluation. Mais, de manière très résumée, ce rôle part d'un extrême, où l'évaluateur est un expert externe qui se prononce sur le programme et affirme que celui-ci est bon ou mauvais, à un autre extrême, où l'évaluateur joue un rôle de facilitateur qui aide les parties prenantes à se prononcer sur leur programme mais sans se prononcer lui-même.

Aborder le rôle de l'évaluateur, qui peut être externe ou interne au programme, c'est aussi aborder la question des relations et des rôles des parties prenantes. Si l'évaluateur est un expert externe qui portera seul un jugement sur la qualité du programme, le rôle des parties prenantes à ce niveau sera forcément limité. À l'inverse, un évaluateur facilitateur laissera plus de place aux parties prenantes.

Les modèles d'évaluation basés sur la théorie des programmes ne confinent pas l'évaluateur à un seul rôle. Pour Chen (2005, p. 268), le rôle de l'évaluateur est pluriel. C'est un évaluateur au sens strict (*assessor*), mais aussi un fournisseur d'information, un expert (*troubleshooter*), un facilitateur, un consultant, un partenaire de développement et même un éclairer. Selon notre expérience, ce rôle doit être précisé au départ et réévalué par la suite. La tendance des milieux est de voir le rôle de l'évaluateur surtout comme un chercheur et non pas comme un partenaire. Mais la participation du chercheur aux réunions de développement permet de modifier ce rôle en cours de route. Encore ici, la seule solution consiste à négocier ce rôle dès le départ, tout en prévoyant une redéfinition en cours de route si nécessaire.

---

3. En anglais, le terme *evaluation theory* fait référence à un modèle d'évaluation comprenant des stratégies ou des approches. À ne pas confondre avec *theory-driven evaluation*, qui est en fait un modèle d'évaluation préconisé par certains auteurs. La *theory-driven evaluation* est une *evaluation theory* parmi d'autres.

En plus du rôle de l'évaluateur il faut aborder la question des standards, aspect plus technique de l'évaluation d'implantation. Nous l'avons passée sous silence jusqu'à maintenant, mais nous devons l'aborder ici. Un standard est un critère auquel sera comparé le programme implanté. Très souvent, les standards sont quantitatifs. Par exemple, on pourrait fixer un standard pour chacune des questions objectives du tableau 6.1. Ainsi, pour la question « Combien de clients ont été référés au programme? », on devrait fixer un standard avant même de commencer l'implantation du programme. Une bonne planification du programme devrait comprendre une telle information. Combien de clients les gestionnaires veulent-ils desservir la première année du programme? L'année suivante? Un monitoring ou une analyse des dossiers de la clientèle permettra de compter le nombre de clients, et le résultat sera alors comparé au standard. Plus on sera près des standards pour les différentes composantes du programme, plus on aura une implantation réussie.

Définir les standards est un exercice exigeant qu'il n'est pas toujours possible de réaliser. Dans certains cas, le programme sera défini de manière trop large ou pas assez précise, et il sera impossible de déterminer des standards. Il faut être conscient qu'il sera alors plus difficile de réaliser une évaluation d'implantation qui vise à se prononcer sur la qualité de l'implantation. Elle pourra être essentiellement formative. Mais, lorsque le programme est relativement bien planifié, l'évaluateur peut aussi proposer des standards aux parties prenantes. Cet exercice permet notamment de clarifier les attentes de chacun. Nous avons vécu une situation où le fait de définir des standards a permis de constater que la direction de l'organisation avait des attentes très différentes de celles des intervenants à l'égard du programme. L'évaluateur devient alors un facilitateur et un négociateur. Ce ne sont pas tous les modèles d'évaluation qui proposent de travailler avec des standards pour les évaluations d'implantation. À notre avis, le recours à des standards est très utile, mais nous ne pouvons généraliser cette pratique à toutes les évaluations d'implantation.

Enfin, nous aimerions terminer cette section en abordant la question des rapports d'évaluation. La première qualité d'un rapport est sans contredit son « *timing* ». Apprendre qu'on a un problème d'implantation un an en retard n'est utile pour personne et est incompatible avec les visées de l'évaluation d'implantation, surtout dans un cadre formatif. L'évaluation d'implantation signifie nécessairement la rédaction de plus d'un rapport pendant l'implantation. Nous proposons de négocier la fréquence et la nature des rapports d'évaluation dès le départ, au même moment où on négocie les questions d'évaluation.

Le contenu des rapports devra être adapté à la clientèle. Il n'y a pas de recettes miracles à cet égard. Un bon dosage de précision et de vulgarisation est nécessaire. Les analyses plus techniques peuvent être placées en annexe alors que le texte peut rapporter les faits saillants. Il est important de rappeler les objectifs de l'évaluation, les questions, les approches et méthodes utilisées. On peut structurer le texte par composantes du programme (voir le tableau 6.1) et résumer les différents résultats obtenus avec chaque méthode de collecte de données, ou encore structurer le texte par méthode de collecte de données (plus facile à produire ainsi). Si on procède de cette manière, il faudra écrire une synthèse par composante du programme et répondre aux questions posées.

Composantes du programme, questions d'évaluation, approches, méthodes de collecte, rôle des évaluateurs et des parties prenantes ainsi que rapports d'évaluation sont autant d'éléments qui composent un puzzle avec lequel on devra jouer. Plus le programme sera bien planifié, plus facile sera l'évaluation d'implantation. La tâche peut parfois sembler insurmontable. Ce sont les choix parmi les possibilités énumérées plus tôt qui la rendent possible.

## **4. TOUT METTRE ENSEMBLE AVEC UN EXEMPLE : LE PROGRAMME CAFE**

Il est maintenant temps d'intégrer les parties en un tout. Pour ce faire, nous proposons un exemple amélioré inspiré de nos travaux des dernières années dans le cadre du programme Crise-Ado-Famille-Enfance (CAFE) auprès des CLSC de la région de la Montérégie (Joly *et al.*, 2005 ; Pauzé *et al.*, 2007). Nous débutons par une brève description du programme, puis nous présentons par étape l'intégration des approches, incluant la détermination des rôles des évaluateurs, les questions d'évaluation, les méthodes de collectes de données et la production des rapports.

Le programme CAFE vise à offrir aux jeunes âgés de 6 à 17 ans et aux familles en crise, n'ayant pas de dossier ouvert en protection de la jeunesse, une intervention de première ligne (dans la communauté), rapide (dans les deux heures suivant la demande), brève (durée de 10 semaines) et intensive (2 ou 3 interventions par semaine). Chaque intervenant (plus d'une trentaine dans une région comme la Montérégie) assume une charge moyenne de 8 dossiers à la fois. Le principal objectif visé par ce programme est de maintenir le jeune dans son milieu familial par la réduction des conflits parents-jeune, des dysfonctionnements familiaux et des problèmes de comportement du jeune qui sont à l'origine de la crise familiale. Le programme CAFE prévoit, tel que recommandé par Roberts (2000), une séquence d'interventions en trois étapes.

Étape 1: apaiser l'urgence. Il s'agit à cette étape d'assurer la sécurité de chacun, de répondre aux besoins immédiats de la famille et de mettre en place un contexte favorable à l'intervention.

Étape 2: évaluer. À la fin de la première étape, généralement après deux semaines, l'intervenant procède à l'évaluation des caractéristiques du jeune et de la famille à l'aide des questionnaires standardisés qui composent le protocole d'évaluation multidimensionnel élaboré par Toupin et Pauzé (2003). Il élabore par la suite un plan d'intervention.

Étape 3: résolution des problèmes. Cette étape s'appuie sur les résultats obtenus à l'aide des questionnaires d'évaluation. L'objectif visé est d'identifier les cibles d'intervention clé, de trouver des réponses appropriées à la situation de crise et de réduire l'incidence des facteurs de risque auxquels le jeune et la famille sont confrontés pour prévenir l'émergence d'autres crises.

Il s'agit bien évidemment d'une description très sommaire et certainement insuffisante pour construire un véritable plan d'évaluation. Il faut plus de détails sur la théorie du programme (modèle qui relie les interventions aux résultats attendus) mais surtout sur le plan d'action. Si un tel plan n'existe pas, l'évaluateur doit suggérer aux responsables du programme d'en élaborer un en expliquant qu'il est impossible de vérifier si on est sur le bon chemin quand on ne connaît pas la destination.

#### **4.1. Étape 1 : préparer l'évaluation d'implantation**

Il va de soi que la première étape consiste à clarifier les rôles de chacun, à proposer une démarche d'évaluation, à choisir et à prioriser les questions d'évaluation, à choisir les méthodes de collecte de données appropriées et à s'entendre sur des modalités de rapports.

En ce qui concerne les rôles, nous proposons un partage de ceux-ci où nous, les évaluateurs, agissons d'abord à titre de consultants mais aussi comme de véritables évaluateurs qui auront à poser un jugement sur le programme réellement implanté, dans un esprit de collaboration avec les parties prenantes. Les parties prenantes auront pour tâches de discuter nos propositions, de les adapter, de proposer des solutions alternatives qui leur conviennent et de mettre en place les décisions prises. La création d'un comité d'évaluation est alors essentielle. Nous préférons que ce comité soit intégré, si possible, aux activités du comité responsable du développement du programme. L'évaluation sera alors moins perçue comme une recherche en soi, réalisée à ses propres fins, mais plutôt comme une activité nécessaire visant à améliorer le programme. Cette proposition s'apparente d'ailleurs à l'approche d'évaluation que Chen (2005) nomme « revue de programme/ réunion de développement ».

Le choix d'une ou des approches est aussi important à cette étape car chacune détermine les questions d'évaluation et surtout les méthodes de collecte de données. Lors d'une première rencontre, les évaluateurs discutent avec les parties prenantes de leurs attentes à l'égard de l'implantation de leur programme. Les évaluateurs proposent ensuite, pour discussion, une ou des démarches avec une liste de questions à prioriser. Nous proposons ici de distinguer l'implantation initiale de l'implantation mature afin de déterminer d'avance un moment où un jugement devra être posé sur ce qui a été réalisé au début et afin de discuter des modifications à apporter pour la suite de l'implantation. Aux fins de cet exemple, nous supposons que les parties prenantes ont entériné les questions qui apparaissent au tableau 6.1 présenté à la section 2.

## 4.2. Étape 2 : l'implantation initiale

La première décision est d'établir la durée approximative de l'implantation initiale. Puisque la durée prévue d'intervention est d'environ 10 semaines, puisqu'il faut permettre à un intervenant d'expérimenter le programme sur au moins deux cas, sinon davantage, en supposant que chaque intervenant a un nouveau cas par semaine (avec une charge de 8 dossiers), il faudra deux semaines pour avoir au moins deux cas et 10 semaines supplémentaires pour compléter l'intervention du second cas. On peut donc proposer une évaluation d'implantation initiale d'environ trois à quatre mois. Elle sera plus longue si on ajoute plus de cas, mais ce n'est pas vraiment nécessaire. Selon notre expérience, les difficultés apparaissent généralement tôt dans l'implantation d'un programme.

Compte tenu de l'aspect formatif important de l'évaluation de l'implantation initiale, nous proposons ici une approche de recherche formative basée sur la théorie des programmes (*theory-driven process evaluation*) en raison de son caractère holistique (Chen, 2005, p. 173). Nous y joignons un monitoring de certaines activités et des caractéristiques de la clientèle, et la « revue de programme/réunion de développement » pour organiser un lieu d'échange consacré à l'implantation. Ce choix d'approches exige que la liste de questions touche toutes les composantes du programme, qu'une diversité de méthodes de collecte d'information soit utilisée et que les rapports soient présentés rapidement en réunion de développement. Il n'est pas question de prendre trois mois pour analyser les données. Il faut agir rapidement. C'est le but de l'opération.

Afin de dresser la liste des méthodes de collecte de données, nous proposons au tableau 6.2 les méthodes de collecte de données privilégiées pour répondre à chaque question. Une question peut appeler plus d'une méthode et, à l'inverse, une méthode peut permettre de répondre à plus

d'une question. On y constate une diversité de méthodes, dont des journaux de bord, des entrevues, des questionnaires d'opinion, des questionnaires sous forme de protocoles d'évaluation des jeunes et des familles, des observations (sous forme d'enregistrements audio) et des analyses documentaires. Il faut nécessairement disposer de certaines ressources pour effectuer ce type d'évaluation, mais pas autant qu'on pourrait le penser. Il est possible de réduire le nombre de méthodes différentes (comme tout faire par entrevues) ou d'en ajouter. C'est une question d'équilibre. Mais il ne faut pas oublier que les coûts directs et indirects d'implantation lacunaire peuvent être nettement plus élevés que les coûts mêmes de l'évaluation. Sans entrer dans tous les détails méthodologiques, voici quelques précisions sur les méthodes de collecte des données.

Le monitoring comprend deux méthodes de collecte des données :

- Les questionnaires du protocole d'évaluation multidimensionnel. Ces questionnaires sont administrés par l'intervenant environ deux semaines après le premier contact. Il s'agit d'une batterie de tests et questionnaires qui mesurent les caractéristiques de la famille, des parents, de la relation parents-enfants et des jeunes selon une perspective écosystémique (Toupin et Pauzé, 2003). Chaque intervenant doit faire remplir ces questionnaires par tous les enfants référés. Ils sont nécessaires à la planification de l'évaluation, et il en résulte une banque de données sur les caractéristiques des clientèles.
- Les journaux de bord (jdb). Remplis par l'intervenant, les jdb ont la forme d'une feuille de format lettre, recto verso, précodifiée, qui permet à l'intervenant de noter pour chaque intervention effectuée auprès d'un cas les personnes visées, le contexte d'intervention, la nature de l'intervention et toute autre information pertinente, incluant les difficultés rencontrées dans le dossier. À remplir au moins pour toutes les interventions effectuées pour les deux premiers cas. Ils permettent de vérifier si les interventions sont rapides, brèves, intensives et effectuées dans le milieu du jeune.

Des entrevues individuelles peuvent être réalisées auprès des intervenants et des gestionnaires (incluant les conseillers). Ici, des entrevues au début du programme, après la formation, puis à un mois et à trois mois, pourraient être fort à propos. Il s'agit de courtes entrevues qui comprennent essentiellement les questions qui apparaissent au tableau 6.2 vis-à-vis les entrevues comme méthode de collecte. La grille d'entrevue est pratiquement élaborée lorsque les questions ont été ainsi identifiées. Il n'est pas nécessaire de trop en sélectionner. Ces entrevues peuvent même être faites au téléphone à un moment opportun pour les intervenants. Les gestionnaires étant moins nombreux, les entrevues peuvent être autant faites au téléphone qu'en personne.

Tableau 6.2

## Exemples de questions d'évaluation et de méthodes de collecte des données proposées pour l'évaluation de l'implantation initiale

| Composantes du plan d'action    | Évaluation formative – Implantation initiale  | Méthodes de collecte de données : sources d'information  |
|---------------------------------|---|--|
| <b>Clientèle</b>                |   |  |
| Objective                       | Quelle est la nature de la crise vécue par les jeunes et leur famille?  | Monitoring : Questionnaires du protocole d'évaluation multidimensionnel : jeunes et familles   |
| Subjective                      | Que pensent les intervenants de la participation au programme des jeunes et des familles?   | Entrevues individuelles : intervenants   |
| <b>Processus d'intervention</b> |   |  |
| Objective                       | Quelles sont la rapidité, la durée et l'intensité des interventions réalisées auprès des premiers clients?<br>Est-ce que les intervenants appliquent fidèlement le programme? | Monitoring : journaux de bord : intervenants<br>Questionnaire d'autoévaluation de la fidélité : intervenants<br>Observation : enregistrement de séances d'intervention |
| Subjective                      | Est-ce que certaines composantes du processus d'intervention posent problème? Lesquelles?   | Entrevues : intervenants   |
| <b>Personnes qui implantent</b> |   |  |
| Objective                       | Comment participent les intervenants aux supervisions?  | Journal de bord : superviseur  |
| Subjective                      | Quel est le sentiment de compétence des intervenants dans ce programme?   | Entrevues : intervenants   |
| <b>Organisation</b>             |   |  |
| Objective                       | Quels aménagements aux horaires et aux charges de travail ont été réalisés par l'organisation?  | Analyse documentaire : dossiers administratifs<br>Entrevues : gestionnaires  |
| Subjective                      | Quelle est l'opinion des intervenants à l'égard des manuels proposés pour appliquer le programme?   | Entrevues individuelles : intervenants   |

Tableau 6.2 (suite)

| <b>Composantes du plan d'action</b> | <b>Évaluation formative – Implantation initiale</b>                                       | <b>Méthodes de collecte de données: sources d'information</b> |
|-------------------------------------|---|---|
| <b>Partenaires</b>                  |   |   |
| Objective                           | Quels sont les protocoles d'arrimage avec les partenaires?                                | Entrevues: gestionnaires                                      |
| Subjective                          | Comment perçoit-on la collaboration entre partenaires?                                    | Entrevues: gestionnaires<br>Entrevues: intervenants           |
| <b>Contexte écologique</b>          |   |   |
| Objective                           | Est-ce que des éléments hors du programme nuisent à l'implantation de celui-ci? Lesquels? | Entrevues: gestionnaires<br>Entrevues: intervenants           |
| Subjective                          | Quelles sont les difficultés rencontrées dans l'implantation du programme?                | Entrevues: gestionnaires<br>Entrevues: intervenants           |

Des questionnaires d'autoévaluation de la fidélité peuvent être soumis après une certaine période d'application. Nous développons présentement un tel instrument mais il en existe de très simples qui ne reposent que sur quelques questions (Ringwalt *et al.*, 2003). Ces instruments donnent un score de fidélité. Dans le présent projet, ils constituent un complément intéressant aux journaux de bord et aux observations. Ces données permettent de vérifier si le programme est vraiment implanté comme prévu.

Des observations sous forme d'enregistrement sonore de quelques entrevues (ciblées à l'avance) effectuées par les intervenants permettent aussi de vérifier la qualité des interventions. Elles peuvent être réalisées pour un seul cas par intervenant. Dumas *et al.* (2001) recommandent d'utiliser ces enregistrements pour vérifier la fidélité d'implantation, et de les utiliser sous forme de rétroaction lors des séances de supervision. Le superviseur peut aussi remplir un journal de bord où il note ces observations sur les agissements et les attitudes des intervenants. Évidemment, ces données ne pourront être dévoilées sans protéger l'anonymat des intervenants et des familles participantes.

Enfin, quelques analyses documentaires permettent de vérifier les aménagements du programme, ainsi que les outils dont il dispose. C'est le cas des manuels, mais aussi des brochures ou encore des données administratives. Les évaluateurs procèdent à une analyse de toute la documentation pertinente et peuvent alors être en mesure de suggérer des améliorations.

Les méthodes d'analyses doivent être simples de manière à ne pas retarder la production de rapports tout aussi simples : tableaux synthèses de données quantitatives (statistiques descriptives), résumés des entrevues par thème, court texte sous forme de questions et réponses sont autant d'outils à la disposition des évaluateurs. Rappelons que le but premier est de fournir une rétroaction rapide au comité de développement du programme. Avec de telles méthodes, nous sommes assurés de couvrir toutes les composantes du programme avec un minimum de rigueur. Les résultats sont discutés en comité de développement, de même que la recherche de solutions. L'implantation des modifications est la responsabilité des parties prenantes.

### 4.3. Étape 3 : l'implantation mature

La démarche pour l'implantation mature est assez semblable, mais il est essentiel d'apporter des ajustements. On pourra conserver les mêmes approches, mais adapter les méthodes de collecte de données à un nombre plus grand d'intervenants et de cas et, éventuellement, prévoir un roulement de personnel. Une stratégie possible est de faire en sorte que les collectes de données et la présentation des résultats soient étalées dans le temps. Un mois, on présente au comité de développement les résultats des jdb, un autre mois, les entrevues, un autre mois, les questionnaires de fidélité, et on recommence ainsi de suite pendant toute la période d'implantation mature, qui pourra durer plus d'un an, s'il le faut.

Voici brièvement quelques adaptations possibles. Au chapitre du monitoring, on pourra procéder à une réévaluation à l'aide du protocole d'évaluation à la fin de la période d'intervention afin de se doter d'un posttest. Les premiers résultats prétest et posttest de ces indicateurs seront des premiers indicateurs de performance qui pourront éventuellement faire partie de l'évaluation des résultats. Il sera bon d'analyser les évolutions observées en fonction des autres indicateurs d'implantation. Une analyse de cas est une méthode appropriée ici. Mais on ne pourra utiliser à cette fin les résultats des évaluations qu'à condition d'atteindre un niveau satisfaisant d'implantation du programme. Pour ce qui est des jdb, on pourra sélectionner des cas types pour lesquels les jdb sont complétés plutôt que de les faire remplir pour chaque intervention dans tous les cas. La pression sur les intervenants est alors moins grande.

Les entrevues seront plus espacées dans le temps, aux six mois, par exemple. Les questionnaires d'autoévaluation de la fidélité pourront aussi être soumis aux six mois, mais on les décalera dans le temps afin d'éviter que ces collectes aient toutes lieu en même temps.

Les observations pourront être délaissées lorsqu'un intervenant atteint un bon niveau de fidélité. Elles seront utilisées pour tout nouvel intervenant. N'oublions d'ailleurs pas de questionner les intervenants qui quitteront le programme sur les raisons de leur départ, car il s'agit d'une donnée importante.

Les résultats seront encore une fois présentés, dans la mesure du possible, rapidement lors des réunions de développement. La dimension formative est toujours présente. Mais la dimension sommative prend maintenant plus d'importance et l'accumulation des données dans le temps permettra de dresser un portrait d'ensemble de l'implantation du programme. Les évaluateurs auront l'obligation de se prononcer sur cette implantation et de suggérer des solutions aux différents problèmes. Les décisions appartiendront aux parties prenantes. La cessation du programme peut même faire partie des recommandations possibles si l'implantation est trop lacunaire. Enfin, advenant que le programme soit maintenu, les données sur l'implantation, comme la fidélité par chacun des intervenants et les caractéristiques des clientèles, pourront être utilisées lors de l'analyse des résultats du programme ou de son incidence.

Enfin, pourquoi ne pas terminer sur un mot au sujet de l'évaluation de l'évaluation que l'on nomme maintenant méta-évaluation (Scriven, 2004, p. 183). En effet, pourquoi ne pas profiter des différentes réunions du comité de développement pour vérifier en cours de route si la démarche d'évaluation est satisfaisante pour les parties prenantes et les évaluateurs ? Il peut arriver que des ajustements à ce chapitre soient autant nécessaires.

## CONCLUSION

Il est clair qu'il est primordial d'évaluer l'implantation des programmes, pour s'assurer qu'ils sont implantés correctement, pour les améliorer ou pour en tenir compte à la phase Après. Il existe une grande variété d'approches pour y arriver. Comme nous l'avons vu, la littérature regorge de suggestions, allant des plus simples aux plus complexes, en dépit du fait que nous n'ayons pas fait appel à toutes les théories possibles d'évaluation des programmes. Nous avons d'entrée de jeu situé notre approche strictement dans une perspective des théories des programmes. Les méthodes de collecte des données suggérées sont tout aussi variées. C'est leur utilisation qui fait la particularité de l'évaluation de programme. Nonobstant le rôle attribué aux évaluateurs, les parties prenantes sont considérées de plus en plus comme des parties participantes aux fins d'évaluation. C'est le retour des résultats aux parties prenantes qui rend d'ailleurs fascinante cette opération d'évaluation. C'est même un des endroits privilégiés où la science rejoint le monde réel. C'est pour cela qu'on s'y intéresse en tant que chercheurs, mais aussi en tant que citoyens.

## BIBLIOGRAPHIE

- ALKIN, M.C. et C.A. CHRISTIE (2004). « An evaluation theory tree », dans M.C. Alkin (dir.), *Evaluation Roots: Tracing Theorists Views and Influences*, Thousand Oaks, Sage Publications, p. 12-65.
- BOTVIN, G.J. *et al.* (1990). « Preventing adolescent drug abuse through a multi-modal cognitive-behavioral approach: Results of a 3-year study », *Journal of Consulting and Clinical Psychology*, vol. 58, p. 437-446.
- BRYMAN, A. (2007). « Barriers to integrating quantitative and qualitative research », *Journal of Mixed Methods Research*, vol. 1, n° 1, p. 8-22.
- BURNS, M.K., R. PETERS et G.H. NOELL (2008). « Using performance feedback to enhance implementation fidelity of the problem solving team process », *Journal of School Psychology* (sous presse).
- CHEN, H.T. (2005). *Practical Program Evaluation*, Thousand Oaks, Sage Publications.
- CREA, T.M., C.L. USHER et J.B. WILDFIRE (2008). « Implementation fidelity of team decisionmaking », *Children and Youth Services Review* (sous-presse).
- DESROSIERS, H. *et al.* (1998). *Cadre de pratique pour l'évaluation de programmes: applications en promotion de la santé et en toxicomanie*, Québec, Gouvernement du Québec, Ministère de la Santé et des Services sociaux.
- DONALDSON, S.I. (2003). « Theory-driven program evaluation in the new millennium », dans S.I. Donaldson et M. Scriven (dir.), *Evaluating Social Programs and Problems: Visions for the New Millennium*, Thousand Oaks, Sage Publications, p. 109-141.
- DUMAS, J.E. *et al.* (2001). « Promoting intervention fidelity: Conceptual issues, methods, and preliminary results from the early alliance prevention trial », *American Journal of Preventive Medicine*, vol. 20, n° 1, p. 38-77.
- DUSENBURY, L. *et al.* (2003). « A review of research on fidelity of implementation: Implications for drug abuse prevention in school settings », *Health Education Research*, vol. 18, n° 2, p. 237-256.
- JOLY, J. *et al.* (2005). *Évaluation de l'implantation du programme Crise-Ado-Famille*, Rapport de recherche, Sherbrooke, Université de Sherbrooke.
- KAM, C.M., M.T. GREENBERG et C.T. WALLS (2003). « Examining the role of implementation quality in school-based prevention using the PATHS curriculum », *Prevention Science*, vol. 4, n° 1, p. 55-63.
- MCGREW, J.H. et M.E. GRISS (2005). « Concurrent and predictive validity of two scales to assess the fidelity of implementation of supported employment », *Psychiatric Rehabilitation Journal*, vol. 29, n° 1, p. 41-47.
- PATTON, M.Q. (1999). *Utilization-focused Evaluation*, 3<sup>e</sup> éd., Thousand Oaks, Sage Publications.
- PAUZÉ, R., L. TOUCHETTE et J. JOLY (2007). « Évolution des familles impliquées dans le programme Crise-Ado-Famille », *Thérapie familiale*, vol. 28, n° 1, p. 45-62.

- RINGWALT, C.L. *et al.* (2003). «Factors associated with fidelity to substance use prevention curriculum guides in the nation's middle schools», *Health Education & Behavior*, vol. 30, p. 375-391.
- ROBERTS, R. (dir.) (2000). *Crisis Intervention Handbook. Assessment, Treatment, and Research*, New York, Oxford University Press.
- ROHRBACH, L.A. et J.W. GRAHAM et W.B. HANSEN (1993). «Diffusion of a school-based substance abuse prevention program : predictors of program implementation», *Preventive Medicine: An International Journal Devoted to Practice and Theory*, vol. 22, n° 2, p. 237-260.
- ROSSI, P.H., M.W. LIPSEY et H.E. FREEMAN (2004). *Evaluation: A Systematic Approach*, 7<sup>e</sup> éd., Thousand Oaks, Sage Publications.
- SCRIVEN, M. (2004). «Reflections», dans M.C. Alkin, (dir.), *Evaluation Roots: Tracing Theorists Views and Influences*, Thousand Oaks, Sage Publications, p. 183-195.
- TOUPIN, J. et R. PAUZÉ (2003). *Guide du portrait synthèse du jeune et de sa famille*, Groupe de recherche sur les inadaptations sociales de l'enfance, Sherbrooke, Université de Sherbrooke.
- WERNER, A. (2004). *A Guide to Implementation Research*, Washington, The Urban Institute Press.



# CHAPITRE 7

---

## **Conjuguer l'évaluation aux temps de l'implantation et des impacts**

Diane Dubeau

Département de psychoéducation et psychologie  
Université du Québec en Outaouais

Sylvain Coutu

Département de psychoéducation et psychologie  
Université du Québec en Outaouais

Geneviève Turcotte

Institut de recherche pour le développement  
social des jeunes (IRDS)

Qu'ils soient de nature conceptuelle, méthodologique ou interpersonnelle, les défis de l'évaluation d'un programme d'intervention psychosociale sont en effet nombreux pour quiconque se lance dans cette entreprise. Évaluer est une opération qui revêt une importance toute particulière du fait qu'elle permet, ultimement, de critiquer ou de documenter la valeur intrinsèque d'un programme. L'erreur la plus courante est cependant de penser que l'évaluation répondra à toutes les questions, et ce, pour toutes les personnes concernées par le programme. Bien souvent, ces questions ne peuvent être résolues simplement au moyen d'une seule approche car elles réfèrent à des formes d'évaluation distinctes, chacune utilisant sa propre logique pour aborder, sous autant d'angles différents, les réalités d'un même objet d'étude. Pour obtenir un portrait aussi complet et valide que possible de ces réalités, on se doit de multiplier, pour un même programme ou projet d'intervention, les angles d'observation, les outils d'évaluation et les méthodologies. Cette approche a été adoptée pour évaluer le projet d'intervention communautaire ProSPère, implanté dans deux territoires du Québec. L'objectif principal de ce programme consiste à valoriser l'engagement des pères auprès de leurs jeunes enfants.

Pour commencer ce chapitre, nous décrirons brièvement le projet d'intervention et retracerons les différentes étapes qui ont marqué son développement. Nous aborderons, dans un second temps, les différents regards évaluatifs portés sur le projet au cours des dix années de sa réalisation. Enfin, nous discuterons de trois principaux enjeux de l'évaluation qui émanent de notre expérience soit : 1) l'évaluation comme composante intégrante du développement d'un programme ; 2) la nécessaire prise en compte du contexte ; et 3) la convergence des subjectivités vers une objectivité.

## **1. PROSPÈRE, UN PROJET D'INTERVENTION COMMUNAUTAIRE VISANT À PROMOUVOIR L'ENGAGEMENT PATERNEL**

### **1.1. Origines du projet ProSPère**

Parmi les diverses recommandations émises en 1991 dans le rapport du groupe de travail sur le développement des jeunes «Un Québec fou de ses enfants», figurait celle de revaloriser le rôle paternel. En 1993, le principal auteur de ce rapport, Camil Bouchard, a entrepris de donner suite à ces conclusions en initiant le projet de recherche-action ProSPère, inscrit à la

programmation du GRAVE-ARDEC<sup>1</sup>. Avec l'étroite collaboration d'un comité promoteur intéressé par la thématique paternelle<sup>2</sup>, il fut décidé d'implanter le projet dans deux communautés québécoises considérées comme vulnérables sur les plans de leurs caractéristiques socioéconomiques et du taux relativement élevé de signalements pour mauvais traitements<sup>3</sup>. Trois entités principales soutiennent et suivent ce projet, soit deux comités de pilotage (une pour chacune des communautés: Initiative Place-O-Père de la municipalité de Pointe-Calumet et CooPère du quartier Rosemont à Montréal) et le comité de la recherche, qui regroupe plusieurs chercheurs issus des milieux institutionnels et universitaires.

## 1.2. Les démarches d'une action concertée

Il apparaît pertinent de préciser dès maintenant deux particularités du projet ProsPère: la présence de chercheurs dès le début du projet et l'adoption d'une perspective épistémologique de recherche-action qui mise sur la reconnaissance d'une expertise partagée entre les milieux de la recherche et de la pratique. Comme dans toutes les démarches d'actions concertées, le projet s'est développé selon un processus en trois étapes, chacune comportant des enjeux et des défis particuliers: 1) la mobilisation des partenaires autour d'un comité de pilotage local chargé de développer le projet dans leur communauté; 2) la recherche de consensus autour d'une logique d'action; et 3) l'élaboration d'un plan annuel d'actions (Ouellet, Paiement et Tremblay, 1995; Goodman, Wandersman, Chinman, Imm et Morissey, 1996).

Près d'une année d'échanges entre les différents acteurs du projet a été allouée afin de dégager une vision commune de la thématique (l'engagement paternel) et des pistes d'action (modèle d'intervention). Cette période initiale

- 
1. Groupe de recherche et d'action sur la victimisation des enfants – Alliance de recherche pour le développement des enfants dans leur communauté (GRAVE-ARDEC). Cette équipe de recherche est subventionnée par le Fonds québécois de recherche sur la société et la culture (FQRSC) et les Instituts de recherche pour la santé (IRSC). Les différents axes de recherche qui structurent la programmation de cette équipe touchent l'épidémiologie et l'étiologie des mauvais traitements, ainsi que le développement et l'évaluation de programmes permettant de couvrir le continuum de l'intervention, en aval et en amont des problèmes de victimisation.
  2. Le comité promoteur est formé de chercheurs et de gestionnaires en provenance de l'Université du Québec à Montréal, des Centres jeunesse de Montréal et des Laurentides, du CLSC Jean-Olivier-Chénier ainsi que de la Direction de la santé publique de Montréal-Centre.
  3. Cinq critères ont présidé au choix des sites d'expérimentation du modèle d'intervention. Les sites à retenir devaient: 1) comporter un volume suffisant de familles ayant au moins un enfant âgé entre 0 et 5 ans; 2) compter au moins 25 % de familles sous le seuil de faible revenu; 3) avoir un taux élevé de signalements pour mauvais traitements; 4) être homogènes quant à la composition ethnoculturelles de la population (plus de 70 % de la population ayant le français comme langue maternelle); et 5) ne pas avoir fait l'objet d'une expérimentation ou d'un travail d'enquête intensif dans le passé.

du projet a également permis de documenter les déterminants de l'engagement paternel (Turcotte, Dubeau, Bolte et Paquette, 2001) et de proposer une définition de ce concept. Cette recherche documentaire, alimentée par les échanges entre les partenaires, s'est traduite par un modèle d'intervention visant à guider l'action en prenant assise sur : 1) un cadre théorique susceptible de tenir compte de la complexité des besoins et des rôles des pères : le modèle écologique; et 2) une approche locale inspirée des principales stratégies de la promotion de la santé et de l'action intersectorielle (voir l'annexe 1). Le modèle d'intervention présenté aux partenaires des deux communautés se voulait relativement ouvert, proposait un cadre général et des lignes directrices, sans toutefois clairement expliciter les activités à mettre en place pour atteindre les objectifs visés. Il revenait ainsi à chacune des communautés d'identifier et d'implanter des actions qui respectaient les particularités de leur population et des services offerts sur leur territoire. Le pari des membres de l'équipe ProsPère consistait en ceci que les « façons de faire » émergeraient progressivement de l'action sur le terrain, ou plus particulièrement d'une interaction constante entre les actions et la réflexion critique à l'égard de ces actions.

À partir de 1995-1996 et pour les quatre années subséquentes, des plans d'action annuels ont été élaborés par le comité de pilotage de chacune des communautés, en conformité avec les principaux objectifs poursuivis par le modèle ProsPère.

Tableau 7.1

### Activités réalisées selon la stratégie d'action

| Site A   | Site B  |
|--|---|
| <b>Activités de sensibilisation de la collectivité et d'influence des décideurs et des intervenants</b>  |   |
| Kiosque d'information dans un centre commercial  | Promotion de l'engagement paternel au magasin Partage                                 |
| Exposition itinérante « Dessine-moi un papa »  | Projection de films sur la question du père dans le cadre de la semaine de la famille |
| Participation à des émissions de télévision  |   |
| Les Cent Pères<br><i>Activité de financement auprès d'hommes d'affaires, de commerçants et leaders de la communauté</i>  |   |
| Publication d'articles et rédaction de chroniques dans les journaux locaux et dans les bulletins d'information de la communauté et des organismes communautaires |   |

Tableau 7.1 (suite)

| Site A  | Site B   |
|---|--|
| Présentations dans des colloques  |  |
| Diffusion d'outils promotionnels (T-shirt, affiches, dépliants, cahier d'accompagnement de l'exposition de dessins) lors d'événements publics |  |
| Accroche-pères (8 parutions)  |  |
| <b>Activités de soutien direct aux pères</b>  |  |
| Accompagnement individuel   | Père-à-Père<br>Accompagnement individuel   |
| Groupe de pères<br><i>Groupe orienté vers l'action, l'entraide et le développement d'activités pères-enfants</i>                              | À deux vitesses<br><i>Atelier père-enfant et groupe de discussion</i>                                  |
|   | Cours papa-cuisine<br><i>Initiation aux techniques de base en cuisine pour pères de jeunes enfants</i> |
| <b>Le soutien aux familles en faveur de l'engagement paternel</b>   |  |
| Rencontres mixtes sur le couple<br><i>Soupers-causeries sur les attentes respectives quant à l'exercice du rôle parental</i>                  | Conte animé ou pièce de théâtre<br><i>Papa trésor</i> à la bibliothèque locale                         |
| Fête des neiges (avec activités pères-enfants)  |  |
| Journée de la famille dans le cadre de la semaine de la famille (avec activités pères-enfants)  | Grande fête familiale dans le cadre de la semaine de la famille (avec activités pères-enfants)         |
| Pêche sur la glace pères-enfants  |  |
| <b>Les actions sur les ressources du milieu</b>   |  |
| Activités de sensibilisation des décideurs (rencontres ponctuelles)   | Rencontres prénatales entre papas et futurs papas  |
|   | Rencontre retrouvailles (rencontre postnatale)   |
|   | Ateliers sur la paternité dans le cadre des groupes YAPP (Y'a personne de parfait)                     |
|   | Séminaires-échanges<br>Formations aux intervenants   |
| Distribution d'affiches et de dépliants dans les milieux d'intervention   |  |

Au fil des ans, entre 12 et 20 activités par année ont été offertes dans les deux communautés (voir l'annexe 2). Les activités étaient variées et complémentaires, permettant ainsi d'agir sur différents fronts (multicibles et multistratégies), tout en exploitant des lieux divers (par exemple, organismes communautaires, parcs, centres commerciaux, etc.).

## 2. REGARDS ÉVALUATIFS POSÉS SUR LE PROJET

Les modalités d'évaluation, quant à elles, ont été déployées tout au long de la réalisation du projet. Trois volets ont ainsi été documentés : l'analyse de milieu, l'implantation et les impacts. Nous présenterons brièvement les principaux éléments constituant ces différents volets évaluatifs ainsi que leur contribution respective. Pour une description plus détaillée des paramètres méthodologiques et des résultats obtenus par les évaluations d'implantation et d'impacts de Prospère, le lecteur est invité à consulter Dubeau, Turcotte, Ouellet et Coutu (à paraître).

### 2.1. Analyse du milieu

Le premier volet évaluatif réalisé par l'équipe consistait en une analyse de milieu des deux communautés participant au projet (Forget, 1995, 1996). Cette évaluation visait à faire état du degré d'ouverture des services offerts aux jeunes familles ainsi qu'à documenter les perceptions des intervenants et de la population à l'égard de l'engagement paternel. Cette étude descriptive a été réalisée à l'aide d'entrevues individuelles, d'une analyse documentaire, d'observations et de groupes de discussion menés auprès d'intervenants, de pères et de mères de la communauté. Les conclusions issues de cette analyse ont permis de mieux cibler les activités à intégrer dans les plans d'action annuels de chacun des comités de pilotage des communautés participantes. La présentation des résultats obtenus a également alimenté des discussions qui ont par la suite contribué au développement de la connaissance mutuelle et de la solidarité de groupe. L'extrait qui suit traduit bien l'impression générale qui se dégageait des membres de l'équipe après avoir participé à l'analyse de milieu *« C'est une étape qui était facilitante parce que cela a permis d'avoir des discussions sur l'image qu'on se faisait, la conception qu'on avait de l'engagement paternel, alors cela a permis aux gens de se jauger puis de se connaître [...] »*

## 2.2. Évaluation de l'implantation

Le second volet de l'évaluation s'attardait aux éléments liés à l'implantation<sup>4</sup>. Ce volet couvre la période de l'automne 1995 à l'automne 2000. Cette analyse fait état de l'ensemble des facteurs associés au déploiement et à la mise en œuvre d'une intervention. L'étude a été réalisée selon la méthode des cas multiples (Yin, 1994, 1998) en utilisant notamment : *a*) des entrevues semi-structurées auprès des principaux acteurs (34 entrevues individuelles et 7 entrevues de groupe); *b*) une analyse documentaire (tous les documents produits depuis 1993 par les sites et les promoteurs); et *c*) l'observation participante des comités et l'observation directe des activités. Les résultats obtenus ont permis, d'une part, de consolider certains aspects de l'intervention et, d'autre part, de revoir ou de réorienter certaines pratiques. Comme nous le verrons plus loin, les résultats de cette évaluation d'implantation se sont avérés particulièrement utiles à plus long terme, et ce, afin de mettre en contexte les résultats mesurés lors de l'évaluation des impacts (voir l'annexe 3 sur le processus de production des effets). Enfin, les données d'implantation ont permis de dégager les éléments à retenir dans l'éventualité d'une dissémination de l'intervention à d'autres communautés. Deux rapports d'implantation ont été publiés dans la foulée de ces premières évaluations (Ouellet, Turcotte et Desjardins, 2001; Turcotte, Desjardins et Ouellet, 2001).

## 2.3. Évaluation des impacts

Le troisième volet évaluatif porte sur l'analyse des impacts communautaires du projet<sup>5</sup>. Deux schèmes de recherche ont été utilisés à cette fin : 1) un schème quasi-expérimental (pré- et posttests avec groupes de comparaison non équivalents); et 2) un schème préexpérimental (posttests avec groupes de comparaison non équivalents). Les collectes de données ont eu lieu durant deux périodes principales, soit en 1995-1996 pour les mesures utilisées en prétest et de 1999 et 2001 pour les mesures de posttest. Plus de 800 parents (mères et pères) et de 100 intervenants ont été rencontrés en entrevue pour cette évaluation (incluant les participants provenant des territoires de Mont-Laurier et de Verdun, qui ont fait office de groupes de comparaison). Des données quantitatives et qualitatives ont été obtenues à la suite de l'administration de questionnaires standardisés et d'instruments élaborés par l'équipe. Les mesures d'impacts ont été regroupées en fonction des clientèles ou des personnes vivant dans les milieux visés (population

---

4. L'équipe responsable de l'évaluation d'implantation est composée de F. Ouellet, G. Turcotte et N. Desjardins. Cette étude a été subventionnée par le Conseil québécois de la recherche sociale (CQRS).

5. D. Dubeau, G. Turcotte et S. Coutu ont assumé les responsabilités liées à l'évaluation d'impacts communautaires de ProsPère, une démarche subventionnée par le CQRS.

en général, les services, les intervenants et les décideurs, les familles, et les pères), de même qu'en regard du type de résultats attendus : intermédiaires (création ou réorganisation de services, nombre et qualité de la collaboration entre les intervenants ou les services, changements dans les attitudes et les pratiques des intervenants) ou ultimes (compétence parentale du père, engagement paternel, attitudes éducatives en regard de la socialisation des enfants, qualité des relations familiales).

## 2.4. Bilans de l'évaluation des impacts : résultats et considérations d'ordre méthodologique

Trois considérations nous permettent d'évaluer positivement les résultats obtenus. Ces résultats témoignent des impacts du projet sur différentes dimensions du rôle des pères. Premièrement, les différences statistiquement significatives observées entre les groupes (expérimentaux et de comparaison) et les temps de mesure (pré- et posttest) respectent toutes, sauf une, la direction des effets attendus (améliorations pour les groupes où il y a eu des interventions). Deuxièmement, l'utilisation de sites témoins s'est avéré un atout important nous permettant de nuancer les résultats qui, dans certains cas, indiquaient un effet principal lié au temps (augmentation/diminution lors du posttest). Dans ces cas, il est plus difficile d'attribuer l'effet à l'intervention puisque l'augmentation/diminution est observée autant dans les sites « interventions » que dans les sites « témoins ». Troisièmement, le patron des résultats obtenus démontre des effets touchant à la fois la sphère cognitive (attitudes éducatives, perception des rôles parentaux) et la sphère comportementale (engagement paternel).

Au-delà de ces résultats, la réalisation de cette étude d'impacts nous a permis de porter un regard critique sur les défis méthodologiques associés à ce type d'évaluation. À titre d'exemple, un des éléments témoignant de la complexité de ce type d'évaluation est sûrement ici la durée du processus. L'écart de cinq années entre les deux collectes de données (prétest et posttest), bien que requis pour assurer une mise en œuvre adéquate de l'ensemble des activités proposées<sup>6</sup>, pose des défis considérables à l'évaluateur, notamment en regard des choix que celui-ci doit faire concernant la procédure d'échantillonnage et les instruments de mesure. Dans ce contexte, les mesures pré- et posttests n'ont pas pu être prises auprès des mêmes parents (échantillons indépendants de familles). L'utilisation d'un même échantillon, suivi de façon longitudinale, aurait possiblement démontré des changements

---

6. Selon Goodman, Wheeler et Lee (1995), de cinq à dix années sont requises à toute initiative d'action concertée pour produire les changements escomptés sur le phénomène visé.

sur la mesure d'engagement paternel (par exemple, le niveau d'engagement paternel lorsque l'enfant a 2 ans est fort possiblement différent de celui évalué cinq années plus tard lorsque l'enfant a 7 ans). Cependant, ces changements pourraient tout autant être attribuables au développement de l'enfant qu'à un quelconque effet de l'intervention.

De façon rétrospective, on ne peut que constater la richesse des données générées par l'adoption de ces trois angles évaluatifs posés sur un même objet d'étude. À notre connaissance, rares sont les projets ou les programmes d'intervention qui disposent de cette diversité de renseignements recueillis aux différentes étapes de leur développement. Les données colligées au fil des années et les conclusions issues de leur analyse ont permis, d'une part, de consolider les bases d'une intervention communautaire visant à favoriser l'engagement paternel en alimentant, en guidant, en questionnant et en validant les actions concrètes mises en place dans chacun des sites. D'autre part, sur le plan de la recherche, et malgré les objectifs distincts poursuivis par les différents volets évaluatifs, les méthodologies adoptées et les principaux résultats obtenus par chacun contribuent à une meilleure adéquation des devis d'évaluation coconstruits avec les principaux acteurs du projet, tout en rendant possible une interprétation nuancée des résultats obtenus dans les différents volets.

### **3. VERS UNE MÉTAÉVALUATION**

Tout comme l'évaluation nous permet de porter un jugement sur la valeur intrinsèque d'un programme, il nous apparaît pertinent également de porter à un niveau supérieur ce regard critique mais, cette fois-ci, sur le processus d'évaluation que nous avons adopté à ProsPère, ce que nous qualifierons de «métaévaluation». Trois constats principaux seront discutés.

#### **3.1. L'évaluation, une composante intégrante d'un programme**

L'évaluation n'est plus aujourd'hui considérée comme une étape qui arrive seulement en fin de parcours du développement d'un programme (Tourigny et Dagenais, 2005). Pendant longtemps, cette opération a été vue comme un test ultime auquel certains programmes se soumettaient (parfois) afin de démontrer leur efficacité. Les avancées importantes observées dans le domaine de l'évaluation depuis les trente dernières années ont permis, entre autres, d'identifier une diversité des formes d'évaluation et d'en démontrer leur pertinence. Il en est ainsi de l'évaluation d'implantation : l'on reconnaît désormais que la mise en œuvre d'un programme peut nécessiter des

adaptations de manière à mieux prendre en compte les réalités des milieux, entraînant ainsi des écarts entre la description écrite du programme et ce qui est effectivement réalisé. Bien que l'interprétation de ces écarts varie selon les auteurs<sup>7</sup>, tous s'entendent sur l'incidence potentielle de ces données d'implantation sur l'évaluation de l'efficacité d'un programme.

De nombreux auteurs situent aujourd'hui l'évaluation dans une perspective résolument dynamique: l'évaluation est vue comme un processus plutôt qu'un produit qui prend des formes différentes selon les étapes du développement d'un programme (Chen, 1996; Owen, 2004). Le modèle intégrateur proposé dans cet ouvrage (Avant – Pendant – Après) s'inscrit tout à fait en conformité avec cette perspective. Le caractère dynamique tient aussi du fait que, par ses collectes d'information, l'évaluation alimente le programme et le transforme. Le développement d'un programme est donc fonction de son évaluation, ce qui nous amène à considérer celle-ci comme une composante intégrante du programme et non comme un processus distinct qui suivrait une trajectoire parallèle au programme lui-même.

Comme nous l'avons vu, le projet ProsPère a bénéficié de l'apport de trois volets évaluatifs qui l'ont alimenté tout au long de sa réalisation. *L'analyse de milieu* a ainsi permis de répertorier l'ensemble des ressources œuvrant auprès des jeunes familles qui étaient disponibles dans la communauté en identifiant plus particulièrement celles intéressées par la thématique de l'engagement paternel. Ces données ont contribué à mieux circonscrire les collaborations potentielles entre les organismes et les institutions de la communauté pour la réalisation d'activités. Le mandat respectif des différents établissements a aussi influencé le type d'activités mis en place (par exemple, le partenariat avec l'organisme communautaire Bouffe Action s'est concrétisé par une activité du plan d'action nommée « Cours papa cuisine »). Dans l'ensemble, l'analyse de milieu a permis de guider l'élaboration des premiers plans d'action annuels par l'identification d'activités qui tenaient compte des caractéristiques et des particularités des populations cibles.

*L'évaluation d'implantation* a, quant à elle, contribué à 1) valider et bonifier les activités offertes; 2) vérifier leur conformité quant aux objectifs principaux du modèle d'intervention proposé aux communautés par

---

7. En évaluation d'implantation, deux approches sont généralement adoptées. La première, qualifiée de logique dite « de diffusion », met l'accent sur l'écart entre le modèle planifié et celui implanté en recherchant l'explication de ces écarts (Contandriopoulos, Champagne, Denis et Denis, 1992). Selon cette approche, le programme à implanter est un produit achevé, stable dans le temps et dans l'espace. Toute modification du produit original est ainsi vue comme un handicap qui altère l'efficacité. La seconde approche en évaluation d'implantation est qualifiée de logique dite « de traduction », qui met en évidence le processus d'adaptation mutuelle du programme à implanter et du contexte d'implantation. Selon ce modèle, la transformation du programme est inhérente à tout processus d'implantation (Callon et Latour, 1986).

ProsPère; et 3) analyser l'adéquation et l'évolution des plans d'action au fil des années. La documentation colligée lors de l'implantation (les activités réalisées, les objectifs poursuivis, les taux de participation, l'appréciation des principaux acteurs concernés) a été d'une aide précieuse lors de l'élaboration du devis d'évaluation des impacts. Cette dernière contribution apporte un éclairage intéressant quant aux dimensions formatives et sommatives de l'évaluation. L'erreur fréquente commise par certaines personnes est d'associer ces dimensions à des formes particulières d'évaluation (évaluation d'implantation = formatif, évaluation d'efficacité = sommatif). Or, selon Stufflebeam (2004) ainsi qu'à la lumière du modèle intégrateur présenté dans cet ouvrage, les dimensions formatives et sommatives peuvent être intégrées au sein d'une même démarche d'évaluation. ProsPère nous semble ici un bel exemple de cette intégration. Lors de l'évaluation d'implantation, les données recueillies chaque année permettaient d'identifier les forces et les difficultés rencontrées quant à des composantes particulières de la mise en œuvre du projet (par exemple, recrutement, clientèle rejointe, objectifs poursuivis, qualité du partenariat, etc.) afin de mieux guider les modifications à apporter. Il s'agit de la dimension formative de l'évaluation où l'objectif poursuivi est de colliger des données utiles pour améliorer le programme/projet afin qu'il puisse mieux répondre aux besoins de la clientèle desservie. La dimension sommative, quant à elle, vise à apprécier la performance ou le rendement d'un programme ou d'une (ou plusieurs) composantes de ce programme. Ainsi, le suivi sur plusieurs années de l'implantation de ProsPère, ainsi que les données colligées quant aux effets perçus à la suite de la participation au projet, ont contribué à formuler et vérifier des hypothèses quant au processus de production des effets (dimension sommative). Ces données ont été utiles pour l'élaboration du modèle d'évaluation des impacts du projet.

En regard de l'évaluation d'impacts, les dimensions sommatives (impacts réels proximaux et ultimes du projet) et formatives (adéquation des activités de transfert des connaissances, dissémination possible à d'autres sites) ont de nouveau été considérées. À notre avis, le regard critique porté sur les différents paramètres méthodologiques de l'évaluation peut également se rapporter à la dimension formative d'une évaluation d'impacts. Démontrer les effets d'un programme (sommatif) consiste essentiellement à se centrer sur le « quoi ». Toutefois, l'évaluation peut aussi documenter certaines caractéristiques relatives à la méthodologie ou à l'intervention qui permettraient de mieux comprendre le « pourquoi » de ces effets (formatif). À titre d'exemple, les analyses plus fines réalisées afin de vérifier les effets potentiels liés à la désirabilité sociale, aux caractéristiques d'appariement des sites, aux modèles plus particuliers adoptés dans chacun des sites (actions intersectorielles versus développement communautaire) constituent des questionnements pertinents.

En résumé, concevoir l'évaluation comme une partie intégrante d'un programme, c'est adopter une perspective élargie qui considère à la fois les dimensions formatives et sommatives, et ce, aux différentes étapes de développement d'un programme.

### 3.2. L'évaluation : une question de contexte

Étudier ou évaluer un phénomène en adoptant une perspective développementale – qu'il s'agisse d'un individu ou d'un programme – sous-entend que l'on admette que ce développement ne se produit pas en vase clos. Il est le fruit d'interactions complexes exercées par les caractéristiques des environnements (proximaux et distaux) dans lesquels se déroule l'évaluation. Le contexte prend donc ici toute son sens et devient un aspect incontournable pour de nombreux auteurs dans le domaine de l'évaluation. Le modèle classique d'évaluation « *Context – Input – Process – Product (CIPP)* » élaboré par Stufflebeam (2003) le démontre bien. Selon cet auteur, le contexte, surtout évalué dans les phases initiales de développement d'un programme, réfère à l'identification des principaux besoins auxquels devraient répondre le programme ainsi qu'à une appréciation des atouts et des occasions offertes par sa mise en œuvre. Ces descriptions auraient pour principales fonctions de guider ultérieurement les acteurs concernés par le programme quant au jugement porté sur les buts poursuivis, les priorités établies ainsi qu'en regard des résultats observés (efficacité-efficience).

Dans le cadre du projet ProsPère, un élément important du contexte qui a teinté l'approche évaluative choisie concerne la perspective épistémologique sous-jacente au modèle d'intervention préconisé. Le choix de la recherche-action<sup>8</sup> a ainsi fortement influencé l'arrimage entre la recherche et la pratique, deux milieux soumis à des impératifs souvent bien distincts et susceptibles de susciter des tensions (par exemple, délai pour la diffusion des résultats, date butoir à respecter, financements inégaux pour les activités d'intervention et d'évaluation, etc.). Une des manières d'atténuer des conflits potentiels entre des partenaires issus de milieux différents consiste à allouer des plages de temps destinées à échanger sur leurs réalités respectives tout en favorisant un transfert réciproque des expertises de chacun. Or, de nos jours, le temps est une denrée bien rare qui peut constituer davantage un obstacle qu'un atout pour les personnes concernées. À cet obstacle s'ajoute celui de la mobilité des personnels (départs, congés, mutations, etc.) qui entrave le développement des relations entre les partenaires et modifie l'éta-

8. Selon Ouellet (à paraître), la recherche-action se caractérise par « l'intention de faire surgir un savoir à partir du développement d'une intervention dans un contexte social donné ; le réinvestissement en cours de projet, des connaissances acquises dans l'action et un esprit de collaboration étroite et continue entre chercheurs et acteurs de terrain ».

blissement des consensus communs établis entre eux quant à leur vision de la thématique et des actions à poser. Tout comme nous l'avons énoncé précédemment, près d'une année d'échanges a été allouée avant l'élaboration des premiers plans d'action. Les résultats de l'évaluation d'implantation démontrent que cette période relativement longue a toutefois eu l'avantage de se concrétiser par des plans d'action annuels qui s'inscrivaient en conformité avec les objectifs du modèle d'intervention proposé.

L'expérience de partenariat vécue entre les milieux de la recherche et de la pratique lors de l'évaluation d'implantation de ProSPère a facilité l'étape ultérieure liée à l'évaluation des impacts. Un nouveau chercheur a été intégré à l'équipe pour cette évaluation, ce qui a permis de dissocier davantage les deux formes d'évaluation tout en favorisant une certaine prise de recul quant au projet. Une approche collaborative a été adoptée pour la formulation du devis d'évaluation, favorisant ainsi la participation des différents agents du projet. La rédaction du devis a nécessité plusieurs rencontres sur une durée de six mois entre les comités de pilotage et de la recherche (l'annexe 4 décrit la démarche de réalisation du devis). À nouveau, la période relativement longue allouée aux échanges a contribué à mieux cerner les objectifs poursuivis par le projet, à établir une priorité quant à l'atteinte de ces objectifs, tout en permettant un transfert d'expertise quant aux paramètres méthodologiques requis pour l'évaluation. Dans ce contexte, le rôle de l'évaluateur s'apparente à celui d'un éducateur de l'évaluation (Greene, 2005). Pour certains, l'investissement en temps consenti au transfert d'expertise serait contrebalancé par les impacts positifs associés à une sensibilisation accrue des acteurs du milieu de la pratique à poser un regard critique sur leurs actions et sur l'évaluation (King, 1998). Ultimement, dans certains contextes, cette démarche pourrait amener à ce que Fetterman (2001) désigne comme une prise en charge progressive de l'évaluation par le milieu lui-même (*empowerment evaluation*).

### **3.3. La convergence des subjectivités ne devient-elle pas objectivité ?**

À la lecture des écrits dans le domaine de l'évaluation, on note la dichotomie, si ce n'est l'opposition clairement affirmée, entre le subjectif et l'objectif (cela est également vrai pour tout ce qui touche la recherche en général). On pourrait qualifier grossièrement le «subjectif» par la prise en compte du point de vue des différents acteurs selon une méthode de collecte des données moins structurée, permettant ainsi à chacun d'exprimer librement ses perceptions du phénomène étudié à la lumière de ses valeurs, opinions et expériences passées. En contrepartie, ce qui est qualifié «d'objectif» représente une perception de sens communs qui tentent de réduire

au maximum les éléments affectifs ou personnels qui influenceraient le jugement des personnes interrogées. L'utilisation de méthodologies quantitatives et d'instrumentation validée est ainsi privilégiée.

Le principal problème n'est pas tant l'existence de cette dichotomie mais le jugement de valeur que l'on porte à l'égard de chacun quant à sa contribution à l'avancement des connaissances (savoir, savoir pratique, savoirs expérientiels). Or deux approches distinctes caractérisent la recherche dont l'objectif ultime est l'avancement des connaissances (Deslauriers, 1991). La première approche, qualifiée de quantitative (déductive), vise à infirmer ou à confirmer des hypothèses formulées par le chercheur. Dans cette approche, le chercheur s'appuie sur les connaissances existantes pour suggérer et tester des liens théoriquement plausibles entre les variables qui l'intéressent. La seconde approche est quant à elle de nature qualitative (inductive), et le chercheur est souvent partie prenante du processus; celui-ci joue un rôle d'agent facilitateur visant à faire émerger les connaissances à partir du discours des principales personnes concernées par la thématique à l'étude. Chacune de ces approches, en fonction des objectifs distincts qu'elles poursuivent (inductif versus déductif), oriente inévitablement vers des choix différents quant aux composantes méthodologiques utilisés (dont la nature subjective ou objective des données) et au rôle joué par le chercheur.

L'évaluateur n'est pourtant pas exempt de valeurs susceptibles de guider le choix du paradigme de recherche adopté. Ce choix est également influencé par le type d'évaluation réalisé selon l'étape de développement du programme (évaluation d'implantation versus des effets). Plus fréquemment, on adoptera une approche qualitative en évaluation d'implantation afin de recueillir le point de vue le plus large possible des différents acteurs en fonction de diverses composantes ou processus du programme.

Dans le cadre du projet ProsPère, nous avons vu que diverses méthodologies (quantitatives et qualitatives) ont été utilisées pour la réalisation des évaluations. Ce processus de triangulation a ainsi permis de vérifier la convergence des données provenant de différents répondants et de diverses méthodologies issues à la fois d'approches quantitatives et qualitatives. Dans le cadre de l'évaluation d'implantation, la prise en compte du point de vue des différents acteurs concernés par le programme dans une approche plus approfondie (subjective) s'est avérée particulièrement riche en permettant de faire ressortir des nuances subtiles pertinentes principalement quant aux effets perçus et au processus de production de ces effets. Ces connaissances ont subséquentement permis d'élaborer un devis quantitatif pour l'évaluation d'impacts du projet. Toutefois, la mise en place d'un schème pré- posttests avec groupes de comparaison non équivalents nécessitait dès le début du projet de colliger les données éventuellement pertinentes pour l'évaluation d'impacts. De façon concrète, ces données ont été utilisées lors

de l'implantation, qui s'est traduite par la distribution d'accroche-porte à l'ensemble des domiciles de la communauté pour sensibiliser et informer la population sur différentes facettes de la paternité et sur les activités d'intervention offertes. La reconnaissance des contributions respectives des différentes approches (qualitative et quantitative) et de leur synergie s'est avérée un atout précieux, attribuable en grande partie à une équipe de chercheurs relativement stable tout au long du projet et dont les expertises étaient diversifiées.

En résumé, nous croyons qu'il est peu pertinent en regard de l'avancement des connaissances d'opposer l'objectif et le subjectif. L'un et l'autre ne représenteraient-ils pas la face cachée et éclairée d'un même objet pourtant bien réel? Nous préférons associer ces regards au sein d'un même continuum qui intègre le point de vue des différents acteurs recueilli au moyen de méthodologies diversifiées. L'accumulation des données et des résultats au fil des années permettrait progressivement de dégager des zones de convergence qui définiraient ainsi l'objectif (sens commun issus des multiples regards subjectifs portés au phénomène étudié). Cette objectivité rejoindrait ainsi, et possiblement fort mieux, les préoccupations des différents acteurs concernés par l'évaluation favorisant une meilleure utilisation des résultats et, dans certains cas, l'appropriation même du processus évaluatif. Les contraintes plus grandes vécues dans le domaine de l'évaluation de programmes comparativement à la recherche incitent les évaluateurs à la créativité quant à des méthodologies qui respectent les critères de rigueur, de faisabilité et d'utilité. La combinaison de regards multiples qui permettent l'identification de zones de convergence apparaît ainsi une avenue des plus prometteuses et pertinentes.

## 4. RETOUR SUR LE PROJET PROSPÈRE

Nous avons fait le choix de centrer les contenus abordés sur certains aspects plus particuliers que le projet ProsPère nous a permis d'explorer, prenant en compte la durée du projet (plus de 10 ans) et les multiples regards évaluatifs posés. Il s'agit d'une expérience relativement unique ayant alimenté de nombreuses discussions sur le plan de l'évaluation. Nous aimerions ainsi conclure en retraçant brièvement les principaux éléments de ces discussions que nous dégageons quant aux facteurs ayant facilité ou fait obstacle aux évaluations réalisées dans le cadre de ce projet. Nous croyons que ces pistes de réflexion seront d'un intérêt certain pour tout individu intéressé à se lancer dans cette belle aventure qu'est l'évaluation.

## 4.1. Facteurs facilitants

### 4.1.1. *La thématique du programme d'intervention*

Au début des années 1990, il y avait peu de choses faites en regard de l'intervention auprès des pères et on notait une certaine volonté politique à investir dans ce domaine. L'engagement paternel est ainsi apparu rapidement comme un thème mobilisateur pour réunir différents partenaires provenant de divers secteurs (éducation, services de santé et services sociaux, etc.). C'était une thématique également qui interpellait facilement les personnes intéressées puisque nous avons tous eu un père, nous sommes mère ou père ou en voie de le devenir...

### 4.1.2. *Un attrait pour un modèle souple et pour la recherche action*

La perspective épistémologique adoptée permettait de reconnaître les expertises propres aux milieux de la recherche et de la pratique. Le respect de ces zones d'expertise a ainsi optimisé leur apport au projet en favorisant le développement progressif d'un capital relationnel entre ces deux milieux. La souplesse du modèle d'intervention proposé se traduisait par l'élaboration des plans d'action annuels qui prenaient en compte les caractéristiques de la population et les ressources disponibles pour chacune des communautés.

### 4.1.3. *Des personnes clés au départ*

La mobilisation des personnes ou des partenaires autour d'une thématique d'action réside dans ses étapes initiales par la présence d'acteurs clés, de porteurs de flambeau. Dans le cadre du projet ProsPère, nous avons eu la chance de bénéficier du dynamisme, de la crédibilité et de la visibilité d'organismes communautaires provenant du CLSC de chacune des communautés.

### 4.1.4. *Diversité de l'équipe des chercheurs*

Au fil des années, force a été de reconnaître la contribution apportée par une équipe de recherche diversifiée provenant des milieux universitaires (Université du Québec à Montréal, Université du Québec en Outaouais, Université de Montréal) et institutionnels (Direction de la santé publique Montréal-Centre, Centre jeunesse de Montréal). Les mandats distincts de ces établissements ont favorisé des apports complémentaires sur les plans du financement, des stratégies de diffusion ainsi qu'en regard des approches d'évaluation. La présence d'étudiants aux cycles supérieurs permettait également d'approfondir certains aspects liés plus particulièrement à l'avancement des connaissances.

#### **4.1.5. *Stabilité des chercheurs et des membres des comités de pilotage au fil des années***

Comme nous l'avons vu plus haut, une période relativement longue a été allouée aux échanges entre les principaux acteurs concernés afin de s'assurer d'une vision commune de la thématique et de la solution à adopter. Ces consensus sont rapidement remis en question lors du départ ou de l'arrivée de nouvelles personnes au sein de l'équipe. De façon surprenante, peu de changements ont été observés au sein de l'équipe ProsPère, et ce, autant pour les milieux de la recherche que de la pratique. Nous croyons qu'il s'avère nécessaire que les acteurs se commettent en début de projet afin d'assurer leur engagement sur une période minimale de temps.

#### **4.1.6. *Apports rapides et concrets pour les milieux de leur partenariat établi avec les chercheurs***

De nombreux préjugés existent de part et d'autre en regard de la recherche et de la pratique. L'arrimage de ces milieux se fait graduellement et nécessite, à notre avis, que durant les premières phases de réalisation du projet, il y ait un apport rapide du partenariat établi. Cet apport peut prendre différentes formes. À ProsPère, la présence d'un chercheur au sein du comité de pilotage s'est concrétisée dans un premier temps par la rédaction des procès-verbaux des rencontres. Ce rôle était jugé utile par les milieux de la pratique et pertinent pour les chercheurs qui gardaient ainsi une trace sur des renseignements éventuellement requis pour l'évaluation d'implantation. Dans les étapes ultérieures, les données colligées par les chercheurs étaient diffusées par différents médiums contribuant ainsi au rayonnement et à la diffusion du projet d'intervention (par exemple grâce à l'accroche-porte).

#### **4.1.7. *Obtention de financement pour les volets évaluatifs***

En tout dernier lieu, nous ne pouvons omettre de souligner l'apport important reçu pour le financement des volets évaluatifs (127 460\$ pour l'évaluation d'implantation et 176 270\$ pour l'évaluation d'impacts).

## **4.2. Les défis rencontrés**

### **4.2.1. *Peu de sources de financement pour les activités d'intervention***

Actuellement au Québec, les sources de financement pour l'intervention sont peu nombreuses. Dans le cadre du projet ProsPère, les activités prévues au plan d'action annuel étaient réalisées avec un maigre financement de 20 000\$ par année. Parmi les conséquences observées par cette situation, on retrouve l'investissement de certaines énergies dans la recherche de financement

plutôt que dans l'intervention, et la présence d'irritants suscités par l'écart noté entre le financement des volets évaluatifs (recherche) et de l'intervention. Toutefois, l'équipe relativement nombreuse et diversifiée associée au projet a permis de puiser à même les différents réseaux pour trouver le financement minimal requis.

#### **4.2.2. Difficultés de concilier les impératifs parfois distincts entre les milieux de la recherche et de la pratique (échéanciers, centrations, priorités)**

L'un des défis constants tout au long de la réalisation du projet a été celui de concilier les agendas souvent différents existant entre les milieux de la recherche et de la pratique. La mise en place de mécanismes d'échanges réguliers, et ce, dès le début du projet (rencontre des comités de pilotage, etc.) a permis une familiarisation progressive aux réalités de chacun. Il demeure toutefois certaines sources d'insatisfaction dont celles, entre autres, associées au délai nécessaire pour l'analyse et la diffusion des résultats d'impacts (rigueur et contrôle des différentes variables analysées – diffusion des résultats le plus rapidement possible en appui à l'obtention de financement).

#### **4.2.3. Enjeux éthiques liés à la participation de plusieurs sites quant à la diffusion des résultats (résultats individuels vs de groupe)**

Tout comme il en est de la recherche, l'évaluation soulève de belles questions sur le plan de l'éthique. Dans le cadre du projet ProsPère, deux communautés ont fait l'objet d'interventions. Nous avons privilégié d'analyser de façon combinée les données issues de ces deux sites afin de valider le modèle d'intervention commun qui a été proposé. Toutefois, certaines demandes ont été formulées quant à une analyse comparative de chacun des sites. La réponse à ces demandes requiert une analyse éthique qui pondère les avantages et les désavantages associés à cette stratégie.

#### **4.2.4. La fin du projet**

Il est difficile d'identifier le moment où prendra fin le partenariat établi entre certains acteurs, surtout si ce projet est mené sur plusieurs années. L'identification de ces acteurs lors des étapes initiales de consolidation du projet permet de baliser l'engagement respectif de chacun et ainsi de préparer la période transitoire pendant laquelle s'amorcera (s'il y a lieu) de nouveaux partenariats. La fin d'un projet doit être préparée à l'avance, bien que l'on constate qu'il existe peu de documentation à ce sujet. À ProsPère, le projet continue d'être implanté dans une communauté dans une forme relativement similaire à celle proposée initialement, sans toutefois la

présence des chercheurs. Un nouveau partenariat a été établi dans la seconde communauté dans le cadre d'un autre projet qui vise à rejoindre davantage les pères vulnérables.

En terminant, nous espérons que le plaisir que nous avons eu à cheminer tout au long du projet ProsPère ait contaminé votre lecture de ce chapitre. Durant ce parcours, de nombreuses questions ont accompagné nos démarches évaluatives. Certaines ont trouvé des réponses et d'autres non. L'évaluation est un domaine relativement nouveau qui incite à la créativité pour le développement de méthodes rigoureuses qui respectent les réalités des milieux de l'intervention pour lesquelles il est bien souvent difficiles d'exercer un contrôle. Nous espérons que notre expérience saura servir d'inspiration aux chercheurs, intervenants et toute autre personne intéressée à relever les multiples défis de l'évaluation des programmes d'intervention communautaire.

Annexe I

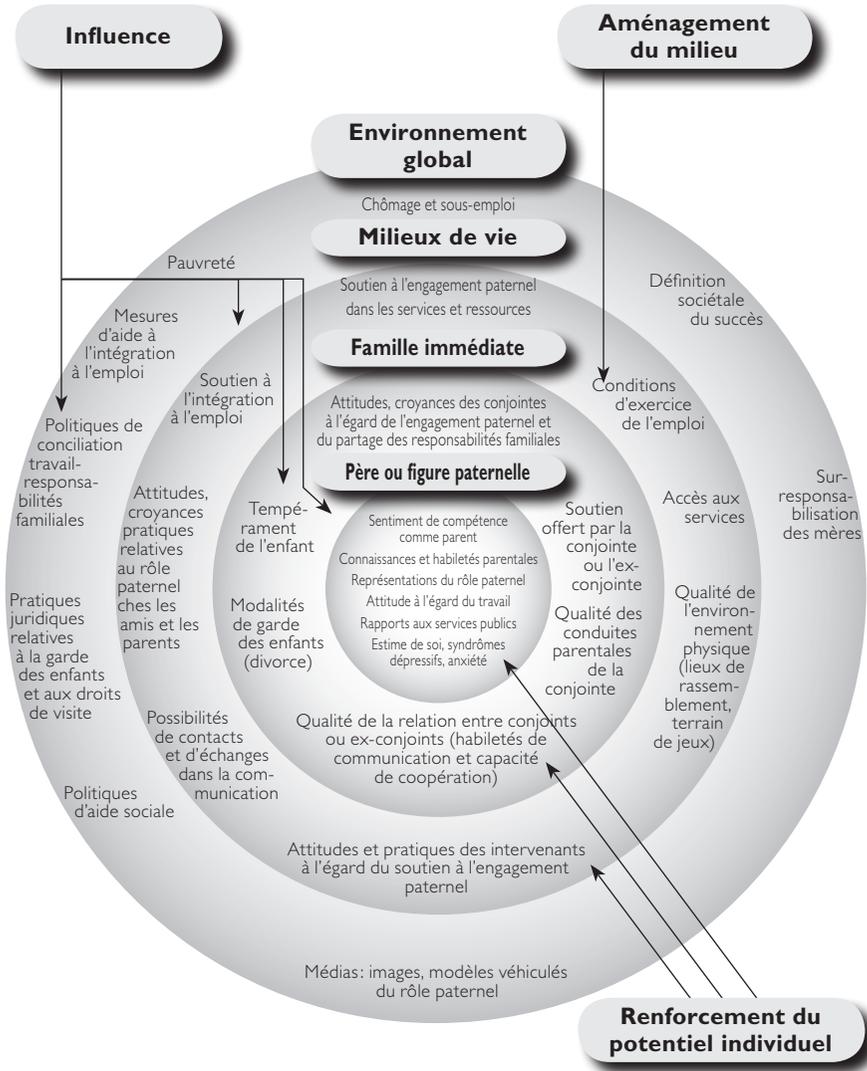
## **Stratégies et cibles d'intervention pour un regroupement local visant à promouvoir l'engagement paternel**

| <b>Stratégies d'intervention</b>     | <b>Description</b>  |
|--------------------------------------|---|
| Influence                            | Ensemble d'activités de sensibilisation allant de la diffusion d'informations à des efforts plus structurés de persuasion visant à renforcer les attitudes et les connaissances d'une personne, d'un groupe ou d'une communauté à l'égard du rôle paternel et du soutien à l'engagement paternel.   |
| Renforcement du potentiel individuel | Activités visant à accroître les habiletés du père ou des personnes significatives (conjointes, amis, parents, intervenants) afin que les pères se sentent plus compétents comme hommes et comme pères, et prennent un rôle actif auprès de leurs enfants.  |
| Aménagement du milieu                | Activités visant à modifier les milieux de vie (services publics, organismes communautaires, garderies, écoles, lieux de loisirs, milieux de travail, etc.) des pères et de leur famille. <ul style="list-style-type: none"> <li>• de façon à offrir des ressources et des activités mieux adaptées aux besoins des pères, plus favorables à l'engagement paternel et à l'amélioration de leur qualité de vie (p. ex., transformation de l'environnement, amélioration des conditions de vie, création de ressources, changement organisationnel);</li> <li>• de façon à permettre aux pères et à la communauté de se regrouper pour définir leurs propres objectifs et choisir leurs propres moyens d'action.</li> </ul> |

Source: Ouellet et Turcotte, 1994.

Figure 1.1

### Déterminants de l'engagement paternel et stratégies d'action



Source: Cette figure a été élaborée pour le groupe ProsPère par Francine Ouellet et Geneviève Turcotte.

## Annexe 2

**Analyse des plans d'action annuels de CoopÈRE**

| Soutien direct aux pères  | Mobili-<br>sation<br>1995-1996 | Plan d'action     |                   |                |                   |
|---|--------------------------------|-------------------|-------------------|----------------|-------------------|
|   |                                | 1<br>1996-1997    | 2<br>1997-1998    | 3<br>1998-1999 | 4<br>1999-2000    |
| Atelier père-enfant : À deux vitesses   |                                |                   |                   |                |                   |
| Cours Papa-cuisine  |                                | n.d. <sup>1</sup> | n.d. <sup>1</sup> |                |                   |
| Père-à-père : accompagnement individuel   |                                |                   |                   |                |                   |
| Entre-Pères. Rencontres thématiques sur la réalité des pères                      |                                |                   |                   |                |                   |
| Nombre d'activités  |                                | 1                 | 1                 | 3              | 4                 |
| <b>Soutien aux familles en faveur de l'engagement paternel</b>                    |                                |                   |                   |                |                   |
| Focus groups avec des pères et des mères  |                                |                   |                   |                |                   |
| Rencontre prénatale entre papas et futurs papas                                   |                                |                   |                   |                |                   |
| Rencontre-retrouvailles   |                                |                   |                   |                |                   |
| Activité de sensibilisation sur l'engagement paternel à Bouffe-Action             |                                |                   |                   |                |                   |
| Grande fête familiale au parc Le Pélican dans la semaine de la famille            |                                |                   |                   |                |                   |
| Atelier spécial « pères » pour les groupes Y'APP et relations parents-enfants     |                                |                   |                   |                | n.d. <sup>1</sup> |
| Conte animé ou pièce de théâtre <i>Papa trésor</i> à la bibliothèque de Rosemont  |                                |                   |                   |                |                   |
| Je sors avec ma garderie... voir <i>Papa trésor</i> à l'Entregens                 |                                |                   |                   |                |                   |
| Parc d'hiver de Rosemont. Bienvenue aux pères                                     |                                |                   |                   |                |                   |
| Sortie au Parc Safari   |                                |                   |                   |                |                   |
| Nombre d'activités  |                                | 3                 | 6                 | 6              | 8                 |
| <b>Sensibilisation de la communauté à l'engagement paternel</b>                   |                                |                   |                   |                |                   |
| Enquête auprès des familles   |                                |                   |                   |                |                   |
| Lancement du calendrier des activités de CoopÈRE                                  |                                |                   |                   |                |                   |
| Élaboration/diffusion d'outils de communication (dépliants, affiches, calendrier) |                                |                   |                   |                |                   |
| Patins papas et petits : kiosque d'information                                    |                                |                   |                   |                |                   |
| Mon père et moi : activités de dessin dans les écoles                             |                                |                   |                   |                |                   |
| Promotion de l'engagement paternel au Magasin Partage                             |                                |                   |                   |                |                   |
| Projection de film, conférence dans le cadre de la semaine de la famille          |                                |                   |                   |                |                   |
| Articles dans le journal de l'Entre-Gens  |                                |                   |                   |                |                   |
| Articles dans le journal de Rosemont  |                                |                   |                   |                |                   |
| Distribution des accroche-porte   |                                |                   |                   |                |                   |
| Nombre d'activités  |                                | 5                 | 5                 | 5              | 5                 |
| <b>Actions sur les ressources</b>   |                                |                   |                   |                |                   |
| Analyse des services  |                                |                   |                   |                |                   |
| Réunions mensuelles des membres de CoopÈRE  |                                |                   |                   |                |                   |
| Séminaires-échange/ateliers de formation  |                                |                   |                   |                |                   |
| Articles dans la revue <i>Le Croqué vif</i>                                       |                                |                   |                   |                |                   |
| Nombre d'activités  |                                | 2                 | 2                 | 3              | 3                 |
| <b>Total des activités</b>  |                                | 11                | 14                | 17             | 20                |

I. Non défini

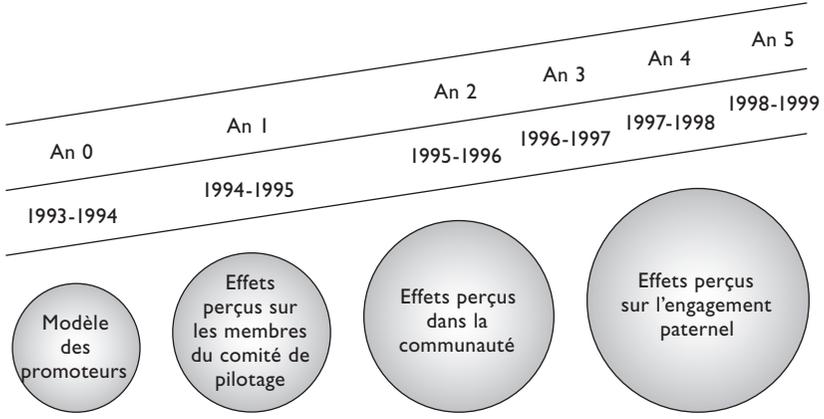
Annexe 3

### Processus de production des effets

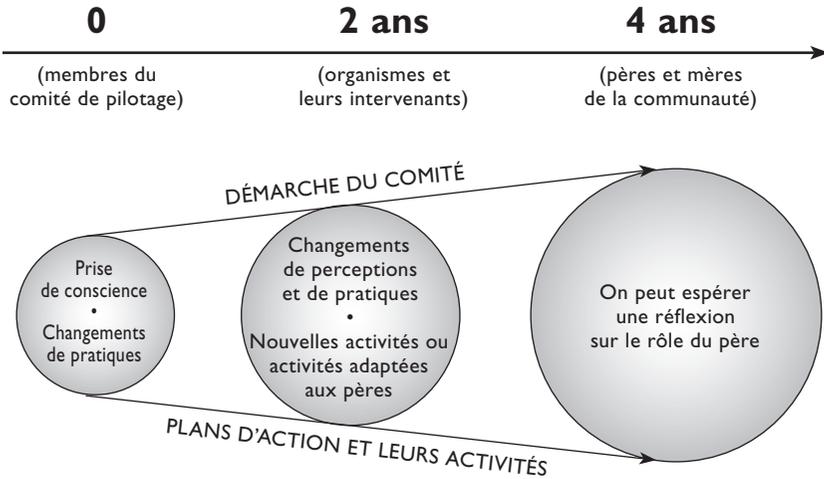
Figures I.2 et I.3

#### Hypothèses quant au processus de production des effets

Chaîne des effets susceptibles de mener à l'engagement paternel à Pointe-Calumet, 1995 à 2000



#### Processus de production des effets



## Annexe 4

**Démarches pour l'élaboration du devis d'évaluation d'impacts de ProsPère et modèle d'évaluation adopté**

Un processus d'évaluation participative (Bernoux, 2004)

**Étape 1: *Identification des dimensions évaluées******(démarche parallèle pour chacune des communautés)***

1. Analyse des plans d'actions pour chacune des années de réalisation du projet et consultation des rapports d'évaluation d'implantation.
2. Liste exhaustive des objectifs implicites et explicites poursuivis par les activités.
3. Classification des objectifs poursuivis par les activités en fonction des quatre objectifs principaux de ProsPère, soit : 1) aménager un environnement sympathique aux pères, 2) influencer la population en général, 3) influencer le réseau familial et 4) renforcer le potentiel individuel du père.
4. Une première rencontre avec chacun des comités de pilotage afin de : a) clarifier le rôle de l'évaluateur ; b) identifier les attentes quant à l'évaluation d'impacts ; et c) valider le système de classification des objectifs des activités.
5. Rédaction d'une seconde version du système de classification des objectifs à la suite des discussions.
6. Deuxième rencontre avec le Comité de pilotage pour vérifier si les modifications apportées étaient conformes aux discussions. Cette rencontre visait également à identifier, par consensus, un ordre de priorité quant aux objectifs évalués.

**Étape 2: *Convergence entre les deux communautés des systèmes de classification des objectifs***

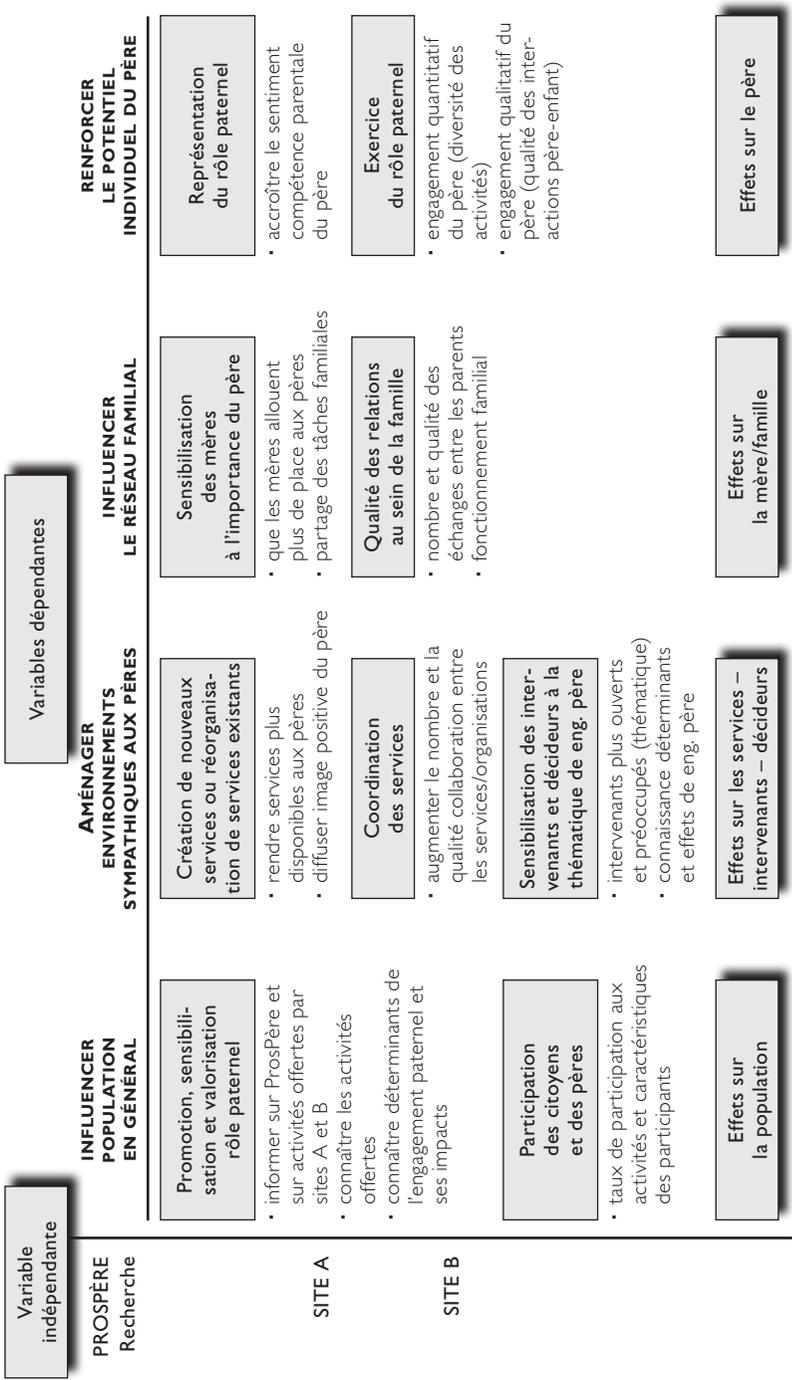
1. Analyse de la convergence (60 % des objectifs identifiés étaient similaires pour les deux communautés).
2. Rencontre des deux comités de pilotage pour leur communiquer les résultats et faire les ajouts (s'il y a lieu).
3. Rencontre du comité de la recherche pour identifier leurs attentes quant à l'évaluation d'impacts et présenter les résultats à la suite des consultations des comités de pilotage. Discussion sur les lignes directrices du modèle d'évaluation.

### Étape 3: *Mise à jour du modèle d'évaluation*

1. Élaboration du modèle d'évaluation à la suite des discussions avec le comité de la recherche.
2. Rencontre des comités de pilotage pour vérifier l'adéquation du modèle d'évaluation et discuter des paramètres méthodologiques proposés dans le devis d'évaluation (procédure, échéancier, instruments de mesure).
3. Rédaction de la seconde version du modèle d'évaluation.

La figure I.4 présente le modèle d'évaluation élaboré à la suite de ces démarches. On peut regrouper les impacts (variables dépendantes) sous quatre catégories d'effets attendus en fonction de la clientèle visée. Ils respectent les quatre principaux objectifs poursuivis par ProsPère qui ont guidé les plans d'actions locaux. Les principaux impacts attendus inscrits dans les encadrés sont opérationnalisés par les objectifs que les comités de pilotage ont considéré comme prioritaires en regard de l'évaluation.

Figure 1.4  
**Modèle d'évaluation**



## BIBLIOGRAPHIE

- ALKIN, M.C. et C.A. CHRISTIE (2004). « An evaluation theory tree », dans M.C. Alkin (dir.), *Evaluation Roots: Tracing Theorists' Views and Influences*, Thousand Oaks, Sage, p. 12-65.
- BERNOUX, J.F. (2004). *L'évaluation participative au service du développement social*, Paris, Dunod.
- CALLON, M. et B. LATOUR (1986). « Les paradoxes de la modernité. Comment concevoir les innovations ? », *Perspective et santé*, vol. 36, p. 13-29.
- CHEN, H.T. (1996). « A comprehensive typology for program evaluation », *Evaluation Practice*, vol. 17, p. 121-130.
- CONTANDRIOPOULOS, A.P. et al. (1992). « L'évaluation dans le domaine de la santé : concepts et méthodes », Version révisée du texte paru dans T. Lebrun, J.C. Saily et M. Amouretti (dir.), *Actes du colloque « L'évaluation en matière de santé : des concepts à la pratique »*, Lille, CREGE, p. 14-32.
- DESLAURIERS, J.-P. (1991). *Recherche qualitative : guide pratique*, Montréal, Théma.
- DUBEAU, D. et al. (à paraître). « L'évaluation ou l'épreuve de la complexité », dans D. Dubeau, A. Devault et G. Forget (dir.) *ProsPère la paternité du XXI<sup>e</sup> siècle*, Québec, Presses de l'Université Laval.
- FETTERMAN, D.M. (2001). *Foundations of Empowerment Evaluation*, Thousand Oaks, Sage.
- FORGET, G. (1995). *Un modèle communautaire de soutien à l'engagement paternel : analyse du milieu à Pointe-Calumet*, Montréal, Direction de la santé publique de Montréal-Centre.
- FORGET, G. (1996). *Défis et opportunités pour la promotion de l'engagement paternel dans le quartier Rosemont*, Montréal, Direction de la santé publique de Montréal-Centre.
- GREENE, J.C., H. KEIDER et E. MAYER (2005). « Combining qualitative and quantitative methods in social inquiry », dans B. Somekh et C. Lewin (dir.), *Research Methods in the Social Sciences*, Londres, Sage, p. 274-281.
- GOODMAN, R.M. et al. (1996). « An ecological assessment of community-based interventions for prevention and health promotion: Approaches to measuring community coalitions », *American Journal of Community Psychology*, vol. 24, p. 33-61.
- GOODMAN, R.M., F.C. WHELLER et P.R. LEE (1995). « Evaluation of the heart to heart project: Lessons from a community-based chronic disease prevention project », *American Journal of Health Promotion*, vol. 9, p. 443-455.
- KING, J.A. (1998). « Making sense of participatory evaluation practice », dans E. Withmore (dir.), *Understanding and Practicing Participatory Evaluation: New Directions for Evaluation*, San Francisco, Jossey-Bass.
- OUELLET, F. (à paraître). « Ma pratique de la recherche action. De la promotion de la santé à la promotion de l'engagement paternel », dans D. Dubeau, A. Devault et G. Forget (dir.) *ProsPère la paternité du XXI<sup>e</sup> siècle*, Québec, Presses de l'Université Laval.

- OUELLET, F., G. TURCOTTE et N. DESJARDINS (2001). *À Rosemont, ça CooPère: analyse d'implantation d'un projet d'action intersectorielle sur la paternité*, Montréal, Régie régionale de la santé et des services sociaux de Montréal-Centre et Direction de la santé publique de Montréal-Centre.
- OUELLET, F., M. PAIEMENT et P.H. TREMBLAY (1995). *L'action intersectorielle, un jeu d'équipe*, Montréal, Direction de la santé publique du Montréal-métropolitain et CECOM de l'Hôpital Rivière-des-Prairies.
- OWEN, J.M. (2004). «Evaluation forms: Toward an inclusive framework for evaluation practice», dans M.C. Alkin (dir.), *Evaluation Roots: Tracing Theorists' Views and Influences*, Thousand Oaks, Sage, p. 356-369.
- ROSSI, P.H., M.W. LIPSEY et H.E. FREEMAN (2004). *Evaluation: A Systematic Approach*, 7<sup>e</sup> éd., Thousand Oaks, Sage.
- STUFFLEBEAM, D.L. (2003). «The CIPP model for evaluation», dans T. Kellaghan et D.L. Stufflebeam (dir.), *The International Handbook of Educational Evaluation*, Boston, Kluwer.
- STUFFLEBEAM, D.L. (2004). «The 21st-century CIPP model. Origins, development, and use», dans M.C. Alkin (dir.), *Evaluation Roots: Tracing Theorists' Views and Influences*, Thousand Oaks, Sage, p. 245-266.
- TOURIGNY, M. et C. DAGENAIS (2005). «Introduction à la recherche évaluative: la recherche au service des intervenants et des gestionnaires», dans S. Bouchard et C. Cyr (dir.), *Recherche psychosociale pour harmoniser recherche et pratique*, Québec, Presses de l'Université du Québec.
- TURCOTTE, G. *et al.* (2001). «Pourquoi certains pères sont-ils plus engagés que d'autres auprès de leurs enfants? Une revue des déterminants de l'engagement paternel», *Revue canadienne de psychoéducation*, vol. 30, n° 10, p. 65-93.
- TURCOTTE, G., N. DESJARDINS et F. OUELLET (2001). *À Pointe-Calumet, on fait place aux pères: analyse d'implantation d'Initiative Place-O-Pères, un projet d'action communautaire sur l'engagement paternel*, Montréal, Régie régionale de la santé et des services sociaux de Montréal-Centre et Direction de la santé publique de Montréal-Centre.
- YIN, R.K. (1994). *Cases study research: Design and methods*, 2<sup>e</sup> éd., Thousand Oaks, Sage Publications.
- YIN, R.K. (1998). «The abridged version of case study research: Design and method», dans L. Bickman et D.J. Rog (dir.), *Handbook of Applied Social Research Methods*, Thousand Oaks, Sage Publications.



# PARTIE 4

---

**DU BILAN  
POSTIMPLANTATION  
À LA MESURE DES  
RÉSULTATS ET EFFETS À  
COURT, MOYEN ET LONG  
TERMES D'UN PROGRAMME  
D'INTERVENTION,  
LA PHASE APRÈS**

8. La perspective plus classique de la mesure des résultats
9. Les apports de la recherche qualitative en évaluation de programmes
10. La dimension formative du bilan d'implantation
11. Les ressources disponibles en évaluation de programme
12. Les programmes se cachent pour mourir

# CHAPITRE 8

---

## **La perspective plus classique de la mesure des résultats** Éléments de méthode, le passage de la recherche scientifique à son adaptation aux réalités de la recherche terrain et de l'évaluation des effets d'un programme

Danny Dessureault  
Département de psychoéducation  
Université du Québec à Trois-Rivières

Valérie Caron  
Département de psychoéducation  
Université du Québec à Trois-Rivières

Le présent chapitre s'inscrit dans la troisième et dernière étape du modèle intégrateur en cube de l'évaluation de programmes. Essentiellement, cette étape vise à déterminer quels sont les effets du programme à court, moyen et long termes. Jusqu'à quel point les personnes ciblées ont-elles vu leur condition de base être modifiée positivement par le programme ? Est-ce que les effets produits par le programme persistent dans le temps ? Voilà autant de questions auxquelles doit répondre l'évaluation de programme afin de statuer sur l'efficacité et l'impact du programme mis en place. En référence à la phase Après du cube, le type d'évaluation présenté dans ce chapitre se situe essentiellement au plan sommatif objectif et subjectif. En fait, il s'agit de tout le domaine de la mesure des résultats concrets du programme et de l'atteinte des buts visés par celui-ci au moment de la planification et de l'évaluation des besoins.

L'évaluation de programmes dans sa phase sommative, c'est-à-dire la mesure des effets, porte sur l'importance des retombées du programme mis en place par une instance. De même, elle permet d'établir le niveau d'atteinte des objectifs poursuivis par le programme auprès d'une population cible. Ce genre d'évaluation de programme est, à l'heure actuelle, assez courant mais aussi particulier à certains domaines plus qu'à d'autres. À titre d'exemple, l'évaluation des pratiques de pointe mises en place par les services de protection à la jeunesse s'appuie sur ce type d'évaluation. En effet, bien que celle-ci soit souhaitable pour tous les programmes implantés, il n'en demeure pas moins qu'elle est très exigeante sur le plan de la méthode et des ressources nécessaires, et c'est pour cette raison qu'elle n'est ainsi pas accessible à tous les programmes.

Dans ses grandes lignes, l'évaluation des effets consiste à établir un lien de causalité entre le programme offert et les changements observés chez la clientèle ou une population. Il faudra donc, d'une part, mettre en place une méthodologie rigoureuse pour parvenir à bien démontrer ce lien de causalité et, d'autre part, recourir la plupart du temps à des mesures objectives du changement. Cette stratégie a pour but de cerner les répercussions d'un programme, qu'elles soient positives ou négatives. Cet aspect est fort important car il peut arriver qu'un programme qui atteint ses objectifs induise en même temps des conséquences négatives sur la population visée. À titre d'exemple, un programme de vaccination est mis en place pour endiguer une maladie particulière, mais provoque, à moyen et long termes, des problèmes importants de santé malgré le fait que l'objectif du programme d'irradier la maladie ait été atteint. Dans ce cas, une évaluation des effets pourrait démontrer que le programme de vaccination atteint sa cible, mais ne fournirait aucune information quant à savoir si les bénéfices obtenus sont supérieurs aux inconvénients vécus. Néanmoins, l'évaluation des effets est pertinente lorsqu'on veut mesurer les effets ou les impacts d'un programme

mis en œuvre depuis quelque temps ou encore lorsqu'on veut s'assurer que les changements manifestés par les personnes découlent directement des activités du programme et non pas d'autres interventions concomitantes ou de tendances naturelles (par exemple, évolution des conditions socioéconomiques, développement normal de la personne, maturation physiologique ou psychologique, rémission spontanée ou naturelle, etc.).

En raison des exigences méthodologiques et des coûts associés à ce type d'évaluation des effets, certains préalables doivent être respectés avant de procéder. Ainsi, le programme évalué doit, d'une part, cibler des objectifs précis et mesurables et, d'autre part, s'appuyer sur un modèle, une logique ou encore sur un processus d'intervention clair, si on désire établir un lien de causalité entre le programme offert et les effets escomptés. Par conséquent, tous les programmes ne se prêtent pas à une évaluation des effets : tout dépend de l'état du programme (phase d'élaboration, d'implantation, d'expérimentation ou d'ajustement). De plus, et tel que le mentionne Potvin (chapitre 5 du présent ouvrage), il importe qu'un programme ait atteint une certaine maturité<sup>1</sup> pour qu'il puisse véritablement se prêter à l'exercice de l'évaluation de ses effets escomptés. Finalement, une bonne connaissance du programme et du milieu dans lequel il est offert devrait permettre de juger non seulement de la pertinence d'une évaluation des effets, mais aussi de sa faisabilité et des éléments sur lesquels l'évaluation peut porter.

## 1. LES APPROCHES THÉORIQUES EN ÉVALUATION DE PROGRAMMES

Au cours des dernières décennies, de nombreuses théories en évaluation de programmes ont été proposées. L'objectif de la présente section n'est pas d'offrir un inventaire exhaustif des approches existantes, ce que tente en partie le premier chapitre du présent ouvrage, mais de fournir un aperçu des grands courants actuels susceptibles de guider une démarche évaluative.

En 1986, Cook et Shadish proposaient une typologie en trois grands groupes d'approches en évaluation de programmes :

1. l'approche classique ou centrée sur les effets ;
2. l'approche centrée sur la théorie de programmes ;
3. l'approche centrée sur les utilisateurs.

---

1. La question de l'évaluation du niveau de maturité d'un programme est également touchée dans la contribution de Joly, Touchette et Pausé, chapitre 6 du présent ouvrage.

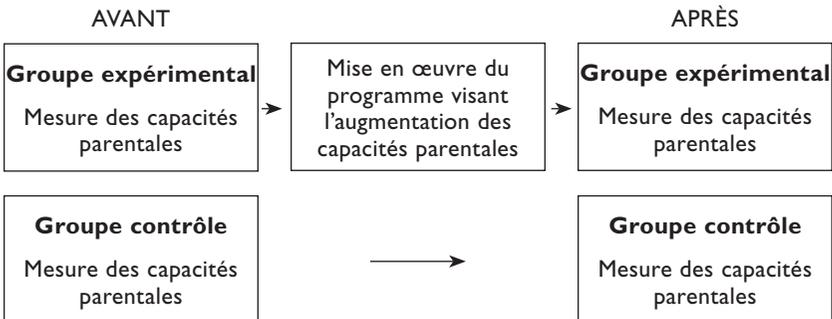
Les deux premières approches s'inscrivent dans la dimension sommative de la troisième face du modèle en cube, alors que l'approche centrée sur les utilisateurs réfère davantage à la dimension formative. Les prochains paragraphes décrivent sommairement chacune de ces trois approches théoriques en évaluation de programmes.

## 1.1. L'approche classique ou centrée sur les effets

L'approche centrée sur les effets vise essentiellement à tester des hypothèses de relations causales entre une intervention ou un programme et des résultats donnés. Pour bien établir la causalité entre le programme et les résultats, le devis expérimental est la méthode la plus adéquate (voir la figure 8.1). Ce devis repose sur la composition de deux groupes identiques de participants dont un est exposé au programme et l'autre non. Une mesure avant et après l'application du programme est faite sur les deux groupes. Si une différence statistique est décelée dans la mesure Après entre les deux groupes et que les biais méthodologiques ont tous été bien contrôlés, on pourra alors conclure que les effets observés sont, selon un ordre de probabilité supérieur au hasard, dus au programme offert. Toutefois, dans la réalité, l'imposition de telles conditions et d'un tel niveau de contrôle demeure généralement impossible à réaliser, tout autant pour des raisons éthiques évidentes que pour des raisons plus pragmatiques liées au manque de moyens et de ressources. Dans de tels cas, on peut choisir un groupe de comparaison qui correspond le plus possible au groupe exposé au programme. On parlera alors d'un devis quasi expérimental, mais il faudra analyser avec prudence les effets observés.

Figure 8.1

### Le modèle expérimental



L'objectif de l'approche centrée sur les effets est de rendre un jugement valide et fiable sur les rapports de causalité entre le programme offert et les effets mesurés. Dans cette approche, la rigueur scientifique est donc recherchée et elle laisse de côté tout le domaine de la compréhension des mécanismes sous-jacents à l'efficacité du programme, c'est-à-dire les questions qui portent sur le pourquoi et le comment se rapportant à la mise en œuvre du programme. Parmi les lacunes de cette approche, il y a, premièrement, son incapacité à déterminer dans quelle mesure une absence d'effet est liée ou non à des difficultés d'implantation et de mise en œuvre du programme (voir le chapitre 6 du présent ouvrage). Deuxièmement, une telle évaluation est onéreuse non seulement en termes de ressources financières et humaines mais aussi en termes de temps destiné à la réalisation. Finalement, l'évolution des considérations éthiques soulève depuis quelques années un questionnement à l'égard de l'utilisation de cette approche.

Bien que ce type d'évaluation soit souhaitable pour tous les programmes implantés, il n'en demeure pas moins qu'il est très exigeant sur le plan de la méthode et des ressources nécessaires à sa réalisation. De plus, l'utilisation de l'approche expérimentale, appliquée de façon stricte à l'évaluation des programmes sociaux, pose des problèmes d'ordre méthodologique et éthique importants liés à la sélection et à la répartition des participants. Par exemple, il est peu fréquent que les participants à un programme puissent être sélectionnés et assignés de manière aléatoire à l'un des deux groupes, expérimental ou de contrôle, le nombre de participants n'étant généralement pas très élevé. De même, on imagine facilement que le choix de soustraire une partie des participants à l'intervention requise par leur état, afin de composer un groupe contrôle, pose des problèmes éthiques ou s'avère souvent impossible dans certains contextes d'intervention. À ces difficultés, s'ajoute le fait que les programmes offerts en milieu naturel offrent peu de possibilités de contrôle sur les facteurs externes qui peuvent biaiser les résultats de l'intervention. En effet, des événements qui se produisent dans la communauté lors du déroulement du programme peuvent agir sur les participants et interférer sur les changements observés lors de l'évaluation du programme. Enfin, l'examen de la relation causale entre l'intervention et les résultats observés n'apporte en soi aucune explication sur les processus associés aux changements observés, ce qui limite considérablement les conclusions de l'évaluation (Paquette et Chagnon, 2000). C'est ce que l'on appelle dans le jargon de l'évaluation de programme le phénomène de la « boîte noire ». En fait, il est possible d'établir que les changements observés sont significatifs sans pouvoir expliquer les processus en cause, ce qui s'avère peu utile pour des recommandations de modifications au programme ou de généralisation à d'autres problématiques comparables (Shadish *et al.*, 1991).

## 1.2. L'approche centrée sur la théorie de programmes

L'approche centrée sur la théorie des programmes vise, bien sûr, à cerner les liens de causalité entre les effets observés et le programme, mais aussi à tenir compte de facteurs extérieurs qui pourraient être à l'origine de ces mêmes effets. Contrairement à l'approche classique, l'approche centrée sur la théorie de programmes s'intéresse à plusieurs questions à la fois, telles que l'évaluation des objectifs, la qualité de l'implantation et les effets souhaités ou pas par les programmes. Dans cette perspective, trois sources d'information peuvent être utilisées pour cerner les motifs et les théories sous-jacentes à la mise en œuvre des programmes : 1) le point de vue des différents acteurs ; 2) les théories existantes ; 3) les recherches exploratoires menées au stade de l'identification des besoins, de la planification ou de l'évaluation de l'implantation (Lipsey et Pollard, 1989). Le recours à ces diverses sources d'information va permettre d'élaborer un modèle théorique pour rendre compte du fonctionnement du programme et de ses effets afin d'anticiper et de préciser les diverses sources d'échec ou de réussite du programme. De ce fait, l'approche centrée sur la théorie du programme se distingue de l'approche plus classique par l'importance qu'elle accorde à l'évaluation de l'implantation du programme et par la participation plus active des divers acteurs intéressés à la démarche évaluative.

Pour Donaldson (cité dans Alkin et Christie, 2005), qui a réalisé plusieurs travaux d'évaluation de programmes en s'appuyant sur l'approche proposée par Chen et Rossi en 1983, l'approche centrée sur la théorie des programmes comporte trois phases distinctes. La première phase consiste à l'élaboration de la théorie du programme. La seconde est la formulation des questions permettant d'évaluer le programme, et la dernière phase n'est autre que la conduite comme telle de l'évaluation du programme. Dans cette perspective, Donaldson considère que l'évaluateur joue, tout au long du processus d'évaluation, un rôle de facilitateur de la discussion. Les discussions ont lieu dans un premier temps avec les responsables de l'application du programme afin d'établir avec eux le modèle théorique du programme, et cela, en tenant compte de toutes les particularités de sa mise en œuvre. Pour cet auteur, la première phase doit être comprise comme étant un processus interactif au cours duquel les responsables et l'évaluateur s'entendent sur les hypothèses ou les trajectoires que devraient suivre les participants pendant l'application du programme. En fait, leur condition de départ devrait se transformer tout au long de la mise en œuvre du programme pour atteindre une nouvelle condition à la fin de celui-ci et rencontrer de ce fait les objectifs du programme. Ainsi, soutenus par la littérature et par les discussions, les responsables et l'évaluateur développent conjointement la théorie du programme. La seconde phase, portant sur la formulation des questions, est réalisée en fonction des éléments présents dans le modèle théorique

du programme, c'est-à-dire la clientèle, les ressources impliquées, les activités d'intervention ainsi que les résultats prévus par le programme. Enfin, la dernière phase comme telle est l'évaluation du programme qui se base sur les réponses aux questions posées et obtenues lors de la cueillette de l'information. Pour Donaldson, le modèle théorique du programme doit guider les choix méthodologiques et non le contraire. Conséquemment, l'évaluateur joue un rôle d'expert auprès des responsables du programme en leur proposant divers moyens pour obtenir de l'information relative à leur programme tout en les informant des avantages et des inconvénients de chacun de ces moyens afin qu'ils puissent faire des choix méthodologiques de façon éclairée. En ce sens, Donaldson (cité par Alkin et Christie, 2005) affirme que l'un des meilleurs exemples de ce type d'évaluation exécutée selon l'approche centrée sur la théorie de programme est celui proposé par le Center for Disease Control (CDC). Selon lui, le modèle d'évaluation du CDC est un bon exemple, non seulement parce qu'il est bien développé, structuré et instructif pour les évaluateurs, mais aussi parce qu'il a été adopté comme guide de référence de l'évaluation des programmes fédéraux de santé partout aux États-Unis.

En terme d'avantages, cette approche permet: 1) d'isoler les sources d'échec et de réussites des programmes évalués; 2) de mettre à jour des effets non prévus; et 3) de mieux comprendre le programme évalué dans sa totalité. Parmi les inconvénients, notons qu'elle exige beaucoup de temps et de ressources; elle demande de grandes habiletés à élaborer des modèles théoriques concernant les liens entre les différentes composantes du programme, et nécessite l'aide d'un expert-conseil en évaluation de programme pour mener à bien certaines tâches.

### 1.3. L'approche centrée sur les utilisateurs

Cette dernière approche regroupe des modèles théoriques dits pluralistes tels que *Responsive evaluation* (Stake, 1983), *Stakeholder-based evaluation* (Weiss, 1983), *Utilization-focused evaluation* (Patton, 1978) ou encore *Practical evaluation* (Patton, 1982). Tous ces modèles privilégient une approche centrée sur l'analyse du programme dans son milieu naturel. Plus particulièrement, l'accent est mis sur l'utilisation des résultats de l'évaluation par les différents acteurs et sur les besoins des utilisateurs. En regard de ce processus, le rôle de l'évaluateur consiste à soutenir les groupes intéressés à se représenter le programme mis en œuvre puisque, comme son nom l'indique, cette approche se caractérise par une participation active des principaux acteurs.

Le recours à l'approche centrée sur les utilisateurs permet, d'une part, de faire ressortir des éléments importants durant le processus d'évaluation et, d'autre part, de permettre aux utilisateurs de s'approprier les résultats de l'évaluation, et conséquemment de favoriser la mise en œuvre des améliorations suggérées. Par contre, la qualité de la démarche et de l'analyse des résultats est fonction du choix des personnes en cause et des questions qui les préoccupent. Et par ricochet, les résultats peuvent parfois être difficilement généralisables. À titre d'exemple de démarches de cet ordre, le lecteur pourra référer à la contribution de Hamel et Cousineau, troisième chapitre du présent ouvrage.

## 1.4. Le modèle logique

Dans le cadre d'un processus d'évaluation, il convient d'obtenir une image qui puisse représenter l'ensemble des éléments entrant dans la mise en œuvre du programme et menant au changement souhaité. L'intérêt d'élaborer un modèle logique réside dans le fait de créer une image relativement simple qui montre comment le programme fonctionne et pourquoi il produit les impacts attendus. Le modèle logique doit découler d'un rationnel issu d'une hypothèse d'explication du problème ou d'une théorie sur les causes du problème ainsi que d'un rationnel sur la façon de le résoudre (Paquette et Chagnon, 2000). Les modèles logiques peuvent prendre des formes diverses suivant les programmes pour lesquels ils sont élaborés. Habituellement, ils se présentent sous la forme de diagrammes qui illustrent de quelles manières un ensemble d'activités est censé donner lieu aux résultats souhaités (Cooksy, Gill et Kelly, 2001). La figure 8.2 présente un exemple de modèle logique pour l'évaluation d'un programme visant l'augmentation des capacités parentales à la suite d'un signalement aux services de protection à l'enfance.

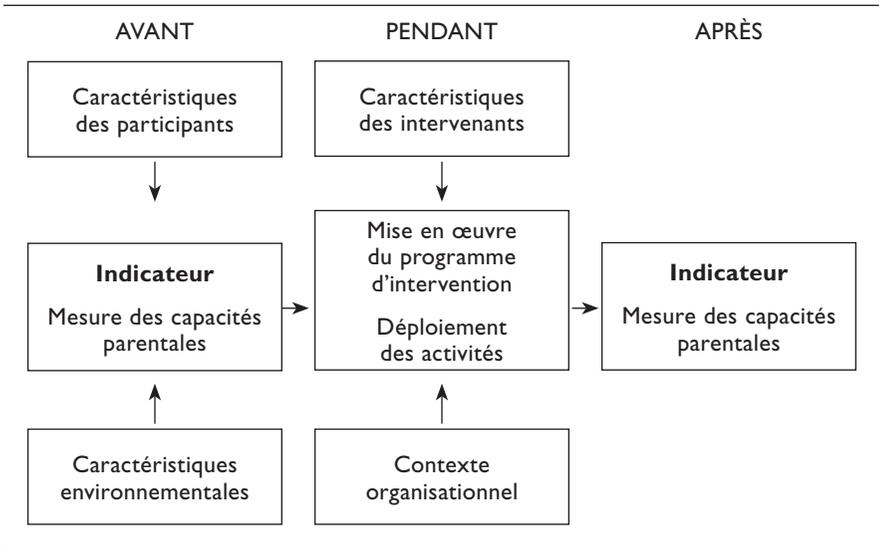
Lors du processus d'élaboration d'un modèle logique pour l'évaluation de programme, certaines questions importantes concernant la planification et l'évaluation du programme doivent être abordées. Ainsi, il faut : 1) définir le problème à résoudre, 2) identifier les besoins et les ressources qui justifient la détermination à résoudre le problème, 3) décrire les résultats attendus, 4) énumérer les facteurs qui pourraient influencer l'introduction des changements, 5) décrire la stratégie de changement proposée par le programme en s'appuyant sur les connaissances tirées des recherches sur les meilleures pratiques, et 6) énoncer des hypothèses détaillant les raisons pour lesquelles les stratégies retenues fonctionneront de la manière prévue lors de l'application dans le milieu.

Pour chaque composante du modèle logique, il est possible d'énumérer un ensemble d'éléments ou caractéristiques qui pourraient jouer un rôle quant à la production des résultats ou de leur absence. Dans la composante

«Caractéristiques des participants», l'évaluateur pourrait tenir compte d'éléments tels que l'âge, le sexe, le type de famille, le revenu, etc. Concernant les caractéristiques organisationnelles, on pourrait inclure les ressources matérielles, financières et humaines nécessaires à la mise en œuvre du programme, la présence d'une restructuration organisationnelle, le non-remplacement des intervenants assignés au programme, etc. Finalement, une attention particulière doit être apportée à la composante «Indicateur» afin de bien choisir les mesures en fonction des résultats souhaités en vue de porter un jugement sur la valeur du lien de causalité entre le programme et les impacts engendrés. Cet aspect est davantage développé dans la prochaine section.

Figure 8.2

### Un modèle logique



## 2. LES MÉTHODES POUR CONDUIRE UNE ÉVALUATION DE PROGRAMME SOMMATIVE

L'objectif principal d'une évaluation sommative est de déterminer l'ampleur des effets engendrés par la mise en œuvre d'un programme. En termes conceptuels, les effets ou les impacts d'un programme réfèrent à un ou des changements attendus (comportements, attitudes, conditions sociales, etc.) chez les participants et qui sont en lien avec les objectifs identifiés lors de la

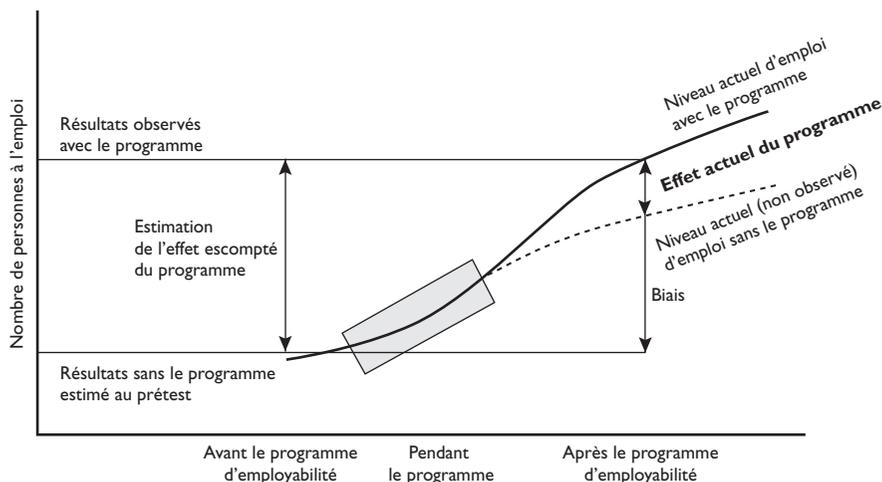
mise en œuvre du programme. Par conséquent, on ne devrait pas observer de changements si le programme n'est pas offert. Méthodologiquement, le défi d'une évaluation sommative est de parvenir à établir avec certitude que c'est uniquement le programme qui engendre les effets observés. En sciences humaines et sociales, les relations causales sont exprimées en termes de probabilités. Ainsi, l'argument « A induit B » veut tout simplement dire que si on introduit A, alors B a beaucoup plus de chance de se produire que si A n'est pas introduit. Toutefois, l'argument ne signifie pas que B est toujours le résultat de A, ou encore que B apparaît seulement si A est présent. À titre d'exemple, considérons un programme de reclassement de la main-d'œuvre (A) pour réduire le taux de chômage (B) dans une population donnée. Si le programme est efficace, il devrait faire augmenter les probabilités chez les participants de se trouver un nouvel emploi, mais cela n'induit pas pour autant que tous les participants au programme vont décrocher un emploi. En effet, la possibilité de se trouver un emploi est aussi reliée à d'autres facteurs qui ne sont pas en lien avec la pertinence et l'efficacité du programme, par exemple, les conditions économiques de la région, la diversité industrielle ou encore le taux d'investissement. Dans une telle perspective, un certain nombre de participants aurait retrouvé de toute façon un emploi sans une participation au programme de reclassement. Mais le fait de participer au programme augmente-t-il de manière significative le nombre de personnes qui se sont trouvé un nouvel emploi ? C'est à cette question que doit répondre l'évaluation sommative. La figure 8.3 montre graphiquement le niveau de l'écart qui peut exister entre l'effet escompté du programme et l'effet réel engendré par ce dernier. L'écart, dans ce cas-ci, n'est autre chose que l'effet du temps et du contexte socioéconomique sur l'employabilité des gens.

## 2.1. Les devis de recherche

Pour y parvenir, l'évaluateur devra élaborer un devis de recherche dans lequel on retrouve une série de procédures méthodologiques qui permettront d'estimer les effets engendrés par un programme. Le premier élément à déterminer est le type de devis sur lequel toute la démarche d'évaluation reposera car l'évaluation des effets se fait essentiellement sous un mode comparatif. En ce sens, déterminer les impacts d'un programme repose sur la comparaison des résultats obtenus par les individus du groupe expérimental participant à un programme à ceux observés dans un groupe contrôle, où l'assignation à chacun des groupes est aléatoire. Si elle ne peut être réalisée, on parlera alors d'un devis de recherche quasi expérimental.

Figure 8.3

### Biais dans l'estimation des effets d'un programme basé sur le changement avant et après sa mise en œuvre



Source: Tiré et adapté de Rossi, Lipsey et Freeman, 2004.

#### 2.1.1. Qu'est-ce qu'un résultat?

Les résultats reflètent un changement entre la condition de départ (avant l'intervention) et la condition à la fin du programme (après l'intervention). Toutefois, ici, une distinction importante doit être faite. L'évaluation des effets d'un programme ne doit jamais porter sur les services, les activités ou les moyens mis en place pour engendrer le changement escompté. En fait, il ne faut pas confondre les services offerts aux participants par le programme (dans la littérature anglophone on parle de *outputs*) des changements observés chez ces derniers (les *outcomes*). C'est ce dernier élément qui doit être uniquement considéré comme un résultat et sur lequel doit porter l'évaluation des impacts d'un programme. Par contre, si le résultat s'applique aux éléments mis en place pour offrir le programme, il ne s'agit plus d'une évaluation de programme mais d'une évaluation d'implantation de programme. Ces deux types d'évaluation, bien que complémentaires dans une évaluation globale, portent sur des aspects et des objectifs distincts. Donc les résultats, au sens où nous l'entendons, peuvent se manifester par des modifications d'habitudes de vie, d'attitudes, de comportements, de perceptions, etc. Généralement, ces changements ne résultent pas d'un lien causal mais plutôt d'une série de plusieurs petits changements qui se sont produits graduellement durant la mise en œuvre du programme. Ce sont ces petits changements qui nous serviront éventuellement d'indicateurs pour l'évaluation.

### 2.1.2. *Qu'est-ce qu'un indicateur?*

Un indicateur est une mesure dont les valeurs sont utilisées comme points de repère dans l'appréciation de l'état ou de l'évolution d'un phénomène (Saucier, 1995). En fait, l'indicateur permet de classer un objet dans une catégorie par rapport à une caractéristique donnée. Il peut prendre la forme de comportements mesurables, de caractéristiques sociodémographiques, comme l'âge, le sexe, le revenu, etc., d'éléments tirés d'un test ou d'un questionnaire. Par exemple, l'augmentation des capacités parentales et la continuité des services sont des phénomènes qui ne sont pas directement mesurables et quantifiables. On peut cependant établir une série de concepts permettant de décrire ces phénomènes, choisir certains de ces concepts et, par la suite, concevoir ou recourir à des mesures permettant d'évaluer ces aspects. En règle générale, un seul indicateur peut rarement rendre compte de toutes les facettes du phénomène étudié et, pour mieux le cerner, l'utilisation de plusieurs indicateurs est recommandée.

À titre d'exemple, un programme destiné aux parents d'enfants d'âge scolaire issus de milieux défavorisés pourrait avoir comme objectif d'augmenter leurs compétences parentales. Un concept comme celui des compétences parentales est à la fois général et abstrait, et peut être perçu de différentes façons en plus d'être abordé par plusieurs moyens différents. Ainsi, le premier travail de l'évaluateur est d'examiner si les objectifs du programme sont clairs et opérationnels et, s'ils ne le sont pas, de les opérationnaliser afin de le rendre plus concrets, observables et mesurables. En ce sens, l'objectif d'augmenter les compétences parentales pourrait être redéfini par plusieurs compétences observables et mesurables que le parent devrait acquérir au terme du programme. Ainsi, après six mois de participation à un programme visant l'augmentation des compétences parentales, ce dernier devrait avoir atteint les objectifs suivants :

1. acquérir des connaissances sur le développement psychophysiologique de l'enfant ;
2. préparer trois repas équilibrés par jour ;
3. prodiguer les soins nécessaires aux enfants ;
4. participer aux rencontres scolaires ;
5. superviser les leçons et les devoirs des enfants et y participer activement ;
6. etc.

Ainsi, ces objectifs observables et mesurables pourraient être transformés par l'évaluateur en indicateurs tels que le nombre d'heures de formation suivies par le parent, le nombre de repas offerts à l'enfant par jour, le nombre de soins d'hygiène, le nombre de participations aux rencontres scolaires ou le temps consacré aux travaux scolaires. Il est essentiel que ces

mesures soient accessibles à l'évaluateur avant et après la mise en œuvre du programme d'intervention afin qu'il puisse les comparer entre elles dans le but de mesurer les impacts réels du programme.

Pour cibler les bons indicateurs, l'évaluateur doit, dans un premier temps, effectuer une recension de la littérature afin de prendre connaissance des indicateurs qui ont été utilisés pour des programmes similaires. Ainsi, la connaissance acquise à l'égard de l'évaluation de ces programmes est susceptible d'éclairer l'évaluateur dans le choix des indicateurs et des instruments de mesure. Finalement, les indicateurs retenus doivent être ceux qui sont les plus pertinents et les plus représentatifs des objectifs poursuivis par le programme à évaluer.

Dans leur ensemble, les programmes d'intervention psychosociale réfèrent la plupart du temps à des concepts plutôt abstraits tels que les troubles de comportements, les capacités parentales, les relations familiales, l'estime de soi, etc. Néanmoins, de nombreux outils sont disponibles pour les mesurer. Cependant, ils doivent être choisis en fonction, d'une part, des objectifs de l'évaluation, et d'autre part, de leurs valeurs psychométriques. Si les objectifs du programme sont bien définis opérationnellement, mais que les instruments nous fournissent une mesure autre que celle attendue, on ne pourra rien conclure à l'égard du dit programme. Parmi les valeurs psychométriques, la *validité* et la *fidélité* d'une mesure nous apparaissent fondamentale, deux éléments sur lesquels nous ne revenons pas ici tant ils sont déjà bien abordés dans des ouvrages à portée plus méthodologique. Tout au long du processus d'évaluation du programme, l'évaluateur doit se référer au rationnel du programme. Il doit garder en tête « pour qui » et « pourquoi » le programme a été mis en place, de même que ce qu'il est supposé engendrer comme transformation chez les participants.

Il arrive parfois qu'un programme n'atteigne pas ses objectifs mais en atteint d'autres qui n'étaient pas prévus. Dans ce cas-ci, ce ne sont pas les résultats indirects qui doivent être mis en évidence par l'évaluation, mais bien les résultats attendus. À titre d'exemple, citons la première évaluation du programme Head Start (Cicarreli, Cooper et Granger, 1969), un programme préscolaire destiné aux familles défavorisées. Une des mesures utilisées pour évaluer l'impact du programme a été l'amélioration des habiletés intellectuelles, alors que Lazar (1981) affirme que l'un des objectifs fondamentaux du programme Head Start était l'implication des parents dans l'éducation des enfants. Bien que cette évaluation ait pu relever des effets significatifs dans l'amélioration des habiletés intellectuelles des enfants issus de milieux défavorisés, il n'en demeure pas moins que ce n'était pas l'objectif du programme. Dans ce cas-ci, l'amélioration des habiletés intellectuelles aurait pu faire partie de l'évaluation comme l'un des indicateurs mais aurait dû être bonifiée par des mesures représentatives de l'implication

parentale dans l'éducation des enfants puisque c'était là le principal objectif. Dans cet exemple, les indicateurs pour l'évaluation du programme auraient pu être plus nombreux et, surtout, porter sur les objectifs du programme eux-mêmes.

## 2.2. Collecte de données et analyses statistiques

Dans le cadre d'une évaluation sommative de type quantitatif, la réalisation d'une évaluation de programme comprend toujours les activités de collecte de données et d'analyses statistiques. Les données doivent être recueillies et compilées en vue d'aboutir à des conclusions sur les questions posées ou encore sur l'efficacité du programme. La façon de procéder lors de la cueillette de données est fonction du type de devis de recherche retenu pour réaliser l'évaluation afin d'estimer le niveau de causalité entre la mise en œuvre du programme et les impacts observés. Généralement, il s'agit d'une étape assez longue demandant rigueur et constance ; il faut veiller à ce que la compilation des données soit cohérente puisque toute la stratégie d'analyses statistiques repose sur celles-ci.

Le choix des méthodes de collecte de données doit idéalement avoir été fait au moment de la préparation de l'évaluation. À cet égard, différents facteurs doivent être considérés : la nature de l'information désirée, les sources de renseignements disponibles, l'approche ou le modèle d'évaluation adopté, les indicateurs servant à représenter les critères retenus, le temps et les ressources financières disponibles pour réaliser l'évaluation, etc. Comme la plupart des évaluations s'intéressent à plus d'une question à la fois et comme les diverses méthodes de collecte de données comportent des forces et des faiblesses, il est généralement recommandé de s'appuyer sur plus d'une méthode de collecte de données afin d'être en mesure de poser un jugement éclairé sur le programme. Dans ce sens, la triangulation des sources d'information offre une avenue intéressante pour bien établir la causalité entre le programme et ses effets, mais des coûts supplémentaires en termes de ressources financières et humaines sont à prévoir.

## 3. INTERPRÉTATION DES RÉSULTATS

L'interprétation des résultats sert, d'une part, à dégager un sens des résultats statistiques obtenus et, d'autre part, à nuancer à l'aide des divers indicateurs le lien de causalité entre le programme et ses impacts. Plus on s'éloigne d'un devis de recherche expérimental, plus il conviendra d'être prudent quant à la portée réelle du programme par rapport aux impacts observés et mesurés.

### 3.1. Dans le contexte d'un devis expérimental

Dans la mesure où l'évaluation du programme repose sur un devis expérimental, c'est-à-dire deux groupes similaires dont l'un est assujéti au programme d'intervention et l'autre non, et comme nous l'avons vu plus haut, le lien de causalité observé entre le programme et ses effets est valide d'un point de vue statistique. Mais avant de confirmer le tout, des vérifications doivent être faites par l'évaluateur en tenant compte du modèle logique développé pour la réalisation de l'évaluation. L'idée consiste ici non pas à reprendre les éléments de l'évaluation de l'implantation (c'est-à-dire la mesure des écarts entre ce que l'on planifiait de faire et ce que l'on fait réellement), mais bien à tenir compte des écarts observés comme autant d'éléments de compréhension supplémentaires susceptibles d'éclairer la teneur des résultats observés et mesurés par l'évaluation sommative. Si la vérification indique que tout a été réalisé comme il se doit, alors les effets observés sont dus au programme, et ses mêmes effets seront, théoriquement, observables de nouveau si le programme est appliqué une nouvelle fois tel quel. Par contre, si la vérification laisse entrevoir que certains éléments auraient pu induire un biais pouvant influencer les effets, l'évaluateur doit en tenir compte lors de l'interprétation des résultats et le mentionner dans son rapport d'évaluation, à savoir si le biais a favorisé ou défavorisé les mesures des effets. Dans ce contexte, l'évaluateur doit tenir compte de ces biais pour nuancer ses conclusions avant d'émettre ou de formuler ses recommandations quant aux diverses pistes d'amélioration du programme. À cet égard, il faut bien saisir, malgré toutes les précautions méthodologiques prises durant le processus d'évaluation, que l'évaluateur ne travaille jamais dans un contexte de connaissance totale du programme, mais toujours dans un contexte de connaissance partielle. Cette situation révèle en soi toute la complexité de l'évaluation de programme. Par conséquent, le « bon » évaluateur doit demeurer humble dans son rôle de porter un jugement sur la valeur intrinsèque du programme qu'il évalue<sup>2</sup>.

2. Au regard de ce dernier point, Guba et Lincoln (1989) reprochaient aux approches de la troisième génération d'être très orientées sur un jugement de valeur du programme. Pour eux, l'évaluation ne devrait pas être menée en vase clos avec les responsables du programme mais plutôt réalisée de façon interactive avec les différents acteurs (professionnels, clients, famille, etc.) concernés par le programme. Ainsi, un changement chez un individu pourrait être lié en partie au programme, sachant que ce même individu a pu être à la fois influencé par son environnement et par le fait qu'il participait à un programme. Pour ces auteurs, le rôle de l'évaluateur ne se résume pas à se prononcer, à donner une opinion ou à porter un jugement sur l'efficacité du programme. Pour eux, il s'agit d'entreprendre un processus interactif en impliquant les acteurs concernés par le programme pour en dresser un portrait global et, par la suite, tenter de l'améliorer afin de mieux répondre aux besoins des usagers de manière réaliste tout en tenant compte du contexte dans lequel se déroule l'intervention. Guba et Lincoln (1989) considèrent que l'évaluateur doit être sensible et à l'écoute des différentes perspectives qui pourraient lui apporter un éclairage sur le programme et ne

### 3.2. Dans le contexte d'un devis quasi expérimental

Dans la situation où le devis est non expérimental, l'interprétation des effets engendrés par le programme est plus délicate. N'ayant pas de groupe contrôle, on ne peut affirmer de prime abord qu'il existe un lien de causalité entre la mise en œuvre du programme et les effets observés. Toutefois, il est possible d'établir si des changements ont eu lieu entre le début et la fin de la mise en œuvre du programme, sans pouvoir nécessairement les attribuer exclusivement au programme. En termes méthodologiques, il s'agit d'une situation «test-retest ou encore prétest-posttest», où une mesure de l'effet escompté du programme est prise auprès des participants avant l'introduction du programme et une seconde fois après l'avoir terminé. En comparant les deux mesures, l'évaluateur peut déterminer de manière statistique s'il existe une différence significative entre ces deux mesures, et par la suite conclure que cette différence est possiblement due au programme. À titre d'exemple, un programme d'intervention est destiné aux parents d'enfants d'âge scolaire issus de milieux défavorisés, dont les objectifs pour les parents sont les suivants :

1. avoir suivi une formation de 15 heures concernant le développement de l'enfant ;
2. offrir à son enfant trois repas équilibrés tous les jours ;
3. voir aux soins corporels de son enfant tous les jours ;
4. offrir un encadrement scolaire à son enfant en lui consacrant au moins 30 minutes par jour pour les travaux scolaires.

Ainsi, en comparant les résultats avant et après l'introduction du programme, on pourrait observer que celui-ci n'induit pas énormément de changement. Mais en poussant davantage son investigation, l'évaluateur pourrait observer qu'un certain nombre de participants font montre d'un changement significatif puisqu'ils ont atteint tous les objectifs visés. Une autre partie des participants peut avoir atteint partiellement les résultats, alors que les autres ne peuvent avoir connu aucun changement à la suite du programme. En regard de tels résultats, l'évaluateur ne doit pas conclure immédiatement que le programme est peu efficace puisqu'il n'a permis qu'à un nombre restreint de participants d'atteindre tous les objectifs. Les motifs pour lesquels les autres participants n'ont pas atteint tous les objectifs du programme peuvent être très nombreux et différents d'un individu à un autre, des éléments qui, en eux-mêmes, pourront faire l'objet d'une réflexion évaluative.

---

pas se limiter uniquement à celles des responsables du programme. Pour eux, l'écart entre la condition de l'individu avant et après le programme ne résulte pas d'un lien linéaire mais plutôt d'un processus interactif entre les différents éléments qui ont pu l'influencer parallèlement avec le programme

En fait, l'évaluateur doit se questionner sur les différents facteurs qui ont pu être présents lors de la mise en œuvre du programme et, par le fait même, influencer ces résultats. Ces facteurs peuvent être liés au contexte organisationnel du milieu dans lequel est implanté le programme, à des facteurs économiques, culturels ou humains. Pour pouvoir les déterminer, il devra regrouper les caractéristiques communes propres aux participants de chacun de ces trois groupes afin de les comparer, en vue de faire ressortir ce qui les distingue malgré le fait que tous les participants répondaient aux critères d'inclusion pour participer au programme. Éventuellement, ce regroupement des caractéristiques pourrait servir à établir une typologie des participants en vue de leur offrir un programme plus adapté à leurs besoins ou à leur réalité.

Pour conclure, on doit retenir que l'interprétation des résultats dans un contexte d'évaluation de programme doit être réalisée avec autant de rigueur méthodologique et scientifique que possible afin d'éviter que des conclusions à l'égard des effets réels de celui-ci soient faussement émises.

## BIBLIOGRAPHIE

- AGENCE DE DÉVELOPPEMENT DE RÉSEAUX LOCAUX DE SERVICES DE SANTÉ ET DE SERVICES SOCIAUX (2004). *Utiliser les modèles logiques pour coordonner la planification, l'action et l'évaluation*, Estrie, Guide d'Élaboration de modèles logiques de programmes.
- ALKIN, M.C. et C.A. CHRISTIE (2005). «Theorists' models in action», *New Directions for Evaluation*, vol. 106, p. 114-115.
- CHEN, H.T. et P.W. ROSSI (1983). «Evaluating with sense: The theory-driven approach», *Evaluation Review*, vol. 7, p. 283-302.
- CICARELLI, V.G., W.H. COOPER et R.L. GRANGER (1969). *The Impact of Head Start: An Evaluation of the Effects of Head Start on Children's Cognitive and Affective Development*, Westinghouse Learning Corporation, OEO Contract No. B 89 4536.
- COOK, T. et W. SHADISH (1986). «Program evaluation: The worldly science», *Annual Review of Psychology*, vol. 37, p. 192-232.
- COOKSY, L.J., P. GILL et P.A. KELLY (2001). «The program logic model as an integrative framework for a multimethod evaluation», *Evaluation and Program Planning*, vol. 24, n° 2, p. 215-235.
- GOUVERNEMENT DU QUÉBEC (1998). *Cadre de pratique pour l'évaluation des programmes. Applications en promotion de la santé et en toxicomanie*, Québec, Ministère de la Santé et des Services sociaux.
- GUBA, E.G. et Y.S. LINCOLN (1989). *Fourth Generation Evaluation*, Newbury Park, Sage.

- LAZAR, I. (1982). *Early Intervention is Effective*, Londres, Educational Leadership, p. 303-305.
- LIPSEY, M.W. et J.A. POLLARD (1989). «Driving toward theory in program evaluation: More models to choose from», *Evaluation and Program Planning*, vol. 12, p. 317-328.
- MONETTE, M., M. CHARRETTE et I. JOBIN (1995). *L'élaboration de programme*, Cap-Rouge, Presses Inter Universitaires.
- PAQUETTE, F. et F. CHAGNON (2000). *Cadre de référence pour le développement et l'évaluation de programmes au Centre jeunesse de Montréal*, Montréal, Institut universitaire Centre jeunesse de Montréal.
- PATTON, M.Q. (1978). *Utilization-Focused Evaluation*, Beverly Hills, Sage.
- PATTON, M.Q. (1982). *Practical Evaluation*, Beverly Hills, Sage.
- PLANTE, J. (1994). *Évaluation de programme*, Québec, Presses de l'Université Laval.
- POSAVAC, E.J. et R.G. CAREY (2007). *Program Evaluation: Method and Case Studies*, 7<sup>e</sup> éd., New Jersey, Pearson Education.
- ROSSI, P.H, H.E. FREEMAN et M.W. LIPSEY (1998). *Evaluation, A Systematic Approach*, 6<sup>e</sup> éd., Thousand Oaks, Sage Publications.
- ROSSI, P.H, H.E. FREEMAN et M.W. LIPSEY (2004). *Evaluation, A Systematic Approach*, 7<sup>e</sup> éd., Thousand Oaks, Sage Publications.
- ROYCE, D. et al. (2001). *Program Evaluation: An Introduction*, 3<sup>e</sup> éd., Wadworth, Brooks Cole Social Work.
- SAUCIER, A. (1995). *Les indicateurs et la gestion par résultats*, Québec, Ministère de la Santé et des Services Sociaux, coll. «Méthodologie et Instruments».
- SHADISH, W.R., T.D. COOK et L.C. LEVITON (1991). *Foundations of Program Evaluation, Theories of Practice*, Thousand Oaks, Sage Publications.
- STAKE, R.E. (1983). «Program evaluation, particularly responsive evaluation», dans G.F. Madaus, M.F. Scriven et D.L. Stufflebeam (dir.), *Evaluation Models: Viewpoints on Educational and Human Services Evaluation*, Boston, Kluwer-Nijhoff.
- WEISS, C.H. (1983). «The stakeholder approach to evaluation: Origins and promise», *New Directions for Program Evaluation*, vol. 17, p. 3-14.
- WEISS, C.H. (1997). «How can theory-based evaluation make greater headway?», *Evaluation Review*, vol. 21, n<sup>o</sup> 4, p. 501-524.
- WEISS, C.H. (1998). *Evaluation*, 2<sup>e</sup> éd., Upper Saddle River, Prentice-Hall.
- WHOLEY, J.S. (1979). *Evaluation: Promise and performance*, Washington, Urban Institute.
- WHOLEY, J.S. (1994). «Assessing the feasibility and likely usefulness of evaluation», dans J.S. Wholey, H.P. Hatry et K.E. Newcomer (dir.), *Handbook of Practical Program Evaluation*, San Francisco, Jossey-Bass, p. 15-39.

# CHAPITRE 9

---

## **Les apports de la recherche qualitative en évaluation de programmes**

Daniel Turcotte  
École de service social  
Université Laval

Isabelle F.-Dufour  
École de service social  
Université Laval

Marie-Christine Saint-Jacques  
École de service social  
Université Laval

Certains auteurs s'entendent pour dire qu'un des grands changements à s'être produit dans le domaine des sciences sociales au cours des vingt-cinq dernières années est l'engouement grandissant de la communauté scientifique pour la recherche qualitative (Morgan, 2007; Guba et Lincoln, 2005). En offrant la possibilité de répondre à des questions demeurerées sans réponses jusque-là et en produisant de nouvelles formes de connaissances, les méthodes qualitatives ont acquis une nouvelle légitimité dans plusieurs secteurs d'activités. Dans le champ de l'évaluation de programmes, bien que cette position ne fasse pas l'unanimité (Guba et Lincoln, 2005), on conçoit généralement la coexistence des méthodes qualitatives et quantitatives (Morgan, 2007). En effet, il est de plus en plus admis que ce qui doit déterminer l'utilisation de l'une ou de l'autre des méthodes, ou leur combinaison, est davantage le type de question que l'on se pose que l'orientation paradigmatique du chercheur (Morgan, 2007; Patton, 2002; Turcotte, 2000). En ce qui a trait particulièrement à la recherche qualitative, il existe plusieurs tendances auxquelles le chercheur peut se référer. Néanmoins, il est possible d'identifier des traits communs qui peuvent servir de points de repère dans l'application de ce type de recherche pour l'évaluation de programmes. Ce chapitre tentera de mettre en évidence ces traits communs en permettant au lecteur : 1) d'avoir une vue d'ensemble de ce qui est commun à la recherche qualitative, 2) de saisir les situations dans lesquelles les méthodes qualitatives sont particulièrement pertinentes pour l'évaluation de programmes, 3) de se familiariser avec les paramètres méthodologiques sous-jacents à l'utilisation de ces méthodes, et 4) de prévoir des stratégies pour assurer la pertinence scientifique de cette démarche de recherche. Finalement, une illustration de recherche évaluative basées sur des méthodes qualitatives sera présentée. Il est important de préciser que ces repères sont présentés afin de faciliter la démarche de la recherche qualitative. Ils ne doivent pas devenir des obstacles à la créativité, à l'originalité et à la flexibilité qui caractérisent cette méthode.

## 1. LA RECHERCHE QUALITATIVE

La recherche qualitative est ainsi désignée parce qu'elle se base sur des données qui ne sont pas numériques (Riessman, 1994). Généralement, elle découle d'une conception de la réalité (ontologie) et d'une façon de l'appréhender (épistémologie) qui la distingue de la recherche quantitative (Spencer, Ritchie, Lewis et Dillon, 2003). De fait, les recherches qualitatives sont souvent l'expression méthodologique d'une lecture constructiviste de la réalité, alors que la recherche quantitative a plus de compatibilité avec le paradigme positiviste. La recherche qualitative est également singulière par sa souplesse sur le plan méthodologique, par sa sensibilité au contexte, par

son approche inductive plutôt que déductive, par l'utilisation de méthodes de collecte et d'analyse de données plus flexibles, par sa quête d'informations riches et complexes, par l'importance accordée à la présence prolongée du chercheur dans le milieu « naturel » des acteurs et par la reconnaissance explicite de l'influence de la perspective du chercheur sur la démarche et les résultats de la recherche. Le principal objectif de la recherche qualitative est de saisir en profondeur l'expérience humaine à partir de la signification subjective des acteurs selon le contexte dans lequel ils évoluent et selon leurs trajectoires personnelles. La recherche qualitative ne poursuit pas l'objectif de dégager des lois universelles, mais tente plutôt de faire émerger différentes constructions de l'objet à l'étude (Deslauriers, 1991).

Bien qu'il existe plusieurs types de recherches qualitatives, il y a certains dénominateurs communs par lesquels ils se rejoignent. Ainsi, aux termes « scientifique, positivisme, réalisme, objectivisme et universalisme », qui sont généralement associés à la recherche quantitative, sont opposés les termes « naturaliste, interprétative, constructiviste, subjective et contextuelle » pour désigner la recherche qualitative (Spencer *et al.*, 2003). Les champs d'application de la recherche qualitative sont évidemment multiples, mais on lui reconnaît de plus en plus une contribution particulière dans le domaine de l'évaluation, qu'elle soit utilisée seule ou en conjonction avec des méthodes quantitatives (Spencer, Ritchie, Lewis et Dillon, 2003).

## 2. LE DOMAINE DE L'ÉVALUATION

Selon Patton (2002), Shaw (1999) et Spencer et ses collaborateurs (2003), le domaine de l'évaluation couvre un large spectre d'activités qui incorpore l'évaluation de programmes, la recherche sur les politiques sociales ainsi que l'analyse des pratiques. Cependant, la notion d'évaluation de programmes est souvent utilisée dans un sens générique pour décrire la production de connaissances pratiques nécessaires pour permettre aux praticiens et aux gestionnaires de prendre des décisions pour améliorer les services offerts à la population. L'évaluation de programmes permet donc généralement de produire trois types de connaissances : elle permet de découvrir « ce qui a fonctionné », « ce qui semble fonctionner » ou « ce qui a le potentiel de fonctionner » dans un service ou un programme (Weinbach, 2005). Selon Patton (1990), la recherche évaluative est essentiellement un effort visant à maximiser l'efficacité humaine par le biais d'une analyse systématique des données ; son objectif est de permettre une prise de décision utile pour les usagers.

## 2.1. Les courants en évaluation de programmes

Selon Greene (1994), le premier courant en évaluation de programmes était principalement orienté sur la mesure de l'efficacité et l'efficacité des programmes et reposait sur des devis expérimentaux misant sur des méthodes quantitatives. Ce courant, qui est qualifié actuellement de post-positiviste, compte toujours de nombreux adeptes, même s'il fait l'objet de critiques, notamment quant à sa capacité de fournir un éclairage sur les facteurs qui contribuent au succès ou à l'échec d'un programme (Pawson et Tilley, 1997). Un deuxième courant, surtout influencé par les travaux de Patton (1982, 1987, 1990, 2002), est apparu pour répondre aux besoins des gestionnaires d'obtenir une information rapide et utile à la prise de décision, ce que permettent rarement les devis expérimentaux. Ce courant pragmatique mise sur une combinaison de méthodes qualitatives et quantitatives, en fonction de l'information recherchée par ceux qui demandent l'évaluation. Selon Patton (1990), la recherche qualitative est utile pour faire des descriptions, pour expliquer ou pour donner une signification aux résultats, mais elle ne permet pas les inférences causales. Les travaux de Lincoln et Guba (1985) illustrent le troisième courant où l'évaluation est abordée comme un processus de construction consensuelle de la réalité par le biais de la dialectique et de l'herméneutique. Cette approche évaluative permet aux différents acteurs concernés par un programme de partager leurs constructions subjectives, leurs expériences et leurs attentes quant au programme. Il est donc naturel de privilégier les méthodes qualitatives pour ce type d'évaluation. Le dernier courant s'est développé sous l'influence des courants féministes, néomarxistes et autres théories critiques. Il met de l'avant l'importance de cerner les facteurs historiques et structureaux qui sont à la base des phénomènes sociaux et qui influencent la distribution des ressources et le partage du pouvoir (Greene, 1994). D'où l'importance de permettre aux participants d'un programme de l'améliorer par un processus de réflexion et d'autoévaluation; le rôle du chercheur consiste alors à agir comme collaborateur ou comme facilitateur de la démarche d'évaluation dont les participants eux-mêmes sont les premiers responsables (Fetterman, 1996). Ces courants sont résumés dans le tableau 9.1.

Pitman et Maxwell (1992) situent les travaux de Miles et Huberman (1984) à mi-chemin entre l'orientation pragmatique de Patton (1982, 1987, 1990) et l'orientation interprétative de Guba et Lincoln (1989). Se qualifiant eux-mêmes de «néopositivistes réalistes», Huberman et Miles (1985) avancent que les phénomènes sociaux existent objectivement et qu'ils présentent certaines régularités. L'objectif de l'évaluation est donc de mettre en évidence une «microcausalité» ou encore une compréhension des relations et des processus qui se produisent dans un contexte précis. Pour cerner ces relations, ils suggèrent l'utilisation d'un cadre conceptuel initial qui délimite

les procédures de collecte et d'analyse des données. Selon leur vision, les méthodes qualitatives permettent non seulement de saisir la signification des événements à partir du point de vue des participants, mais elles permettent la modélisation d'une causalité contextuelle. Ce survol rapide illustre qu'au sein même des auteurs qui privilégient l'utilisation des méthodes qualitatives, il existe des conceptions diversifiées de l'évaluation.

Tableau 9.1

### Les principaux courants en recherche qualitative

| Paradigme                        | Principale conception idéologique/valeurs                                       | Questions habituelles   | Méthodes de recherche  |
|----------------------------------|---|---|--|
| Post-positiviste                 | Réalisme critique / efficacité, reddition de comptes, connaissances prédictives | <ul style="list-style-type: none"> <li>Est-ce que les effets désirables observés sont attribuables au programme ?</li> <li>Est-ce que ce programme est le plus efficient/efficace ?</li> </ul>  | <i>Quantitatives</i> : expérimentale et quasi expérimentale, modélisation causale, vérification d'hypothèses, analyse des coûts bénéfiques                                     |
| Pragmatique                      | Gestion/ fonctionnalité, contrôle de qualité, utilité                           | <ul style="list-style-type: none"> <li>Qu'est-ce qui fonctionne bien dans le programme et qu'est-ce qui doit être ajusté ?</li> <li>À quel point le programme est efficace par rapport aux objectifs de l'organisation ? par rapport aux besoins des bénéficiaires ?</li> </ul> | <i>Éclectiques, mixtes</i> : questionnaire structuré ou semi-structuré, entrevues, observations  |
| Interprétatif                    | Pluralisme/ compréhension, subjectivité, diversité, solidarité                  | <ul style="list-style-type: none"> <li>Quelle est l'expérience vécue par les utilisateurs du programme ?</li> </ul>   | <i>Qualitatives</i> : études de cas, entrevues, observations, analyse documentaire   |
| Critique de la science normative | Émancipation/ empowerment, changement social, participation sociale             | <ul style="list-style-type: none"> <li>De quelle façon les prémisses, objectifs ou activités du programme contribuent-ils à maintenir un accès inégal aux ressources et au pouvoir dans la société ?</li> </ul>   | <i>Participatives</i> : les utilisateurs du programme participent au choix de la stratégie d'évaluation et à la sélection de la méthode, analyse historique ; critique sociale |

Source: Tiré de Greene, 1994.

## 2.2. La pertinence des méthodes qualitatives en évaluation de programmes

Selon le modèle du «cube», il existe trois moments où il est particulièrement utile de mener des recherches évaluatives: 1) avant l'implantation du programme; 2) pendant sa réalisation, notamment pour s'assurer que le programme correspond aux besoins des utilisateurs ou aux objectifs de départ; 3) et après pour en mesurer les effets sous l'angle de l'efficacité ou de l'efficience. Les auteurs s'entendent en général pour reconnaître que la recherche quantitative est plus pertinente pour mesurer des effets, notamment en raison de l'utilisation de statistiques inférencielles qui permettent de mesurer les différences avant-après (Royse, 2008; Royse *et al.*, 2006).

Les méthodes qualitatives, quant à elles, présentent des avantages pour appréhender plusieurs aspects qui sont mesurés avant et pendant le programme. Ainsi, avant le programme, les questions relatives à la nature des besoins, aux ressources à mobiliser, à l'ouverture du milieu quant au programme, ou à la conjonction entre ce qui est souhaité et ce qui est possible, peuvent être abordées avec plus de profondeur avec des méthodes qualitatives.

Pendant le programme, les données relatives à la mobilisation des acteurs, à l'adaptation du programme aux changements et aux imprévus, à la description des processus par lesquels les actions sont réalisées, à la perception des acteurs quant au fonctionnement du programme et à l'adéquation des ressources consacrées au programme présentent généralement une plus grande sensibilité au milieu si elles s'appuient sur une démarche qualitative.

Les méthodes qualitatives peuvent aussi être appropriées après l'implantation d'un programme si l'objectif est de cerner le point de vue des participants sur l'amélioration de leur condition et pour déceler les effets imprévus ou non recherchés. Le tableau 9.2 résume les principales dimensions de l'évaluation de programmes qui peuvent être particulièrement bien documentées à l'aide d'une perspective qualitative.

D'autres situations justifient aussi l'utilisation d'une approche qualitative. En effet, par son approche flexible et inductive des objets, elle permet d'explorer des dimensions sans que ces dernières ne soient nécessairement prévues à l'avance. De fait, l'évaluation qualitative est particulièrement appropriée lorsque l'on dispose de peu d'informations sur le programme, lorsqu'il est difficile de prévoir ce qui doit être mesuré, lorsqu'il s'agit d'un sujet délicat ou d'un sujet sur lequel les connaissances sont peu développées ou, encore, lorsque la recherche porte sur des enjeux, sur des liens informels ou non structurés ou encore sur des programmes évolutifs, transitifs ou en changement. On peut aussi y avoir recours pour obtenir des informations

étouffées sur le contexte ou sur les dynamiques internes du programme, saisir l'expérience selon le point de vue des utilisateurs, compléter des données quantitatives ou encore lorsque les variables que l'on veut mesurer n'ont pas encore été opérationnalisées. C'est le cas également lorsque le contexte de la recherche fait en sorte qu'il est difficile, voire impossible, de conduire une recherche quantitative car les phénomènes ou les processus sont complexes, lorsque l'on a besoin de mesures non intrusives, comme l'observation plutôt que le questionnaire, ou lorsque l'on dispose d'une trop grande diversité de sources de données. Enfin, l'évaluation qualitative est appropriée lorsque l'on ne peut conduire une recherche expérimentale pour des raisons pratiques ou éthiques, comme c'est le cas lorsque le fait de composer un groupe de contrôle empêche des utilisateurs de bénéficier du programme et que cela est susceptible de leur créer un préjudice (Patton, 1987; Rossi, Lipsey et Freeman, 2004; Royse *et al.*, 2006; Spencer *et al.*, 2003; Weinbach, 2005; Deslauriers et Kérésit, 1997; Padgett, 1998; Royse, 2008).

Tableau 9.2

### Apports de la perspective qualitative à l'évaluation d'un programme

| Avant  | Pendant  | Après  |
|--|--|--|
| <ul style="list-style-type: none"> <li>• Nature des besoins</li> <li>• Ressources à mobiliser</li> <li>• Ouverture du milieu</li> <li>• Faisabilité</li> </ul> | <ul style="list-style-type: none"> <li>• Mobilisation des acteurs</li> <li>• Adaptations du programme</li> <li>• Processus sous-tendant les actions</li> <li>• Perceptions des acteurs</li> <li>• Adéquation des ressources</li> </ul> | <ul style="list-style-type: none"> <li>• Changements perçus par les participants</li> <li>• Effets non attendus</li> </ul> |

## 3. LA MÉTHODOLOGIE DE L'ÉVALUATION DE PROGRAMMES BASÉE SUR DES MÉTHODES QUALITATIVES

Dans une démarche évaluative basée sur un devis expérimental ou quasi expérimental, il est essentiel que le protocole de recherche soit appliqué uniformément dans toutes les situations. Il est donc incontournable de rédiger un devis précis qui prévoit toutes les particularités de la démarche de recherche : population, échantillon, instruments, procédures de collecte des données et stratégies d'analyse. En recherche qualitative, la situation est différente ; bien qu'on retrouve un devis précisant les principaux aspects de la méthodologie envisagée, ce dernier ne se présente pas comme une démarche

linéaire qui doit être suivie à la lettre, mais plutôt comme un processus permettant des ajustements. Ainsi, comme les données peuvent être analysées, ou du moins explorées, au fur et à mesure qu'elles sont collectées, le chercheur peut rectifier sa démarche s'il constate que celle-ci ne lui permet pas de répondre à ses questions de recherche. Cette flexibilité ne veut pas dire pour autant que la recherche évaluative est une démarche improvisée, bien au contraire (Royse *et al.*, 2006). Même si la démarche doit être souple, il est conseillé, particulièrement aux nouveaux chercheurs, de la prévoir avec le plus de précision possible, tout en acceptant que cette planification puisse être ajustée en cours de réalisation.

### 3.1. Le contenu du devis de recherche

Patton (1987) identifie deux types d'éléments à aborder dans la planification d'une évaluation basée sur des méthodes qualitatives : les éléments conceptuels et les éléments techniques. Les premiers portent sur la façon dont l'évaluation est envisagée. Ils font référence à des aspects comme les personnes qui demandent l'évaluation, le but de l'évaluation, les questions ou les enjeux de l'évaluation, les éléments politiques à prendre en considération et les ressources disponibles. Les éléments techniques, quant à eux, font référence au plan de collecte et d'analyse des données. Tard, Beaudoin, Turcotte et Ouellet (1997) formulent quatre questions qui peuvent servir d'assise à l'élaboration d'un devis d'évaluation : Que cherche-t-on à savoir ? De quelle information a-t-on besoin ? Qui peut fournir l'information ? Comment obtenir l'information ? Deslauriers et Kérésit (1997) distinguent certains thèmes à clarifier dans un devis de recherche qualitative : la construction de l'objet de recherche, les propositions de recherche, l'échantillon, la collecte des données, l'analyse des données et le rapport. Ces différents modèles sont utilisés dans cette section pour structurer l'information relative à la démarche de recherche évaluative basée sur des méthodes qualitatives.

#### 3.1.1. Les éléments conceptuels

Lorsque s'agit d'évaluation de programmes, le chercheur définit rarement son objet « dans le silence de son bureau » ; il s'agit avant tout d'un « objet négocié, dépendant à la fois de circonstances particulières et de facteurs structurels » (Deslauriers et Kérésit, 1997, p. 91-92). Est-ce que le but de l'évaluation est d'identifier les besoins de la population cible d'un programme ? de décrire comment les différentes activités du programme sont mises en œuvre ? de cerner la trajectoire, de l'entrée dans le programme à la sortie, des personnes qui reçoivent des services ? de clarifier les facteurs qui influencent la réalisation ou les résultats du programme ? C'est sur la base de cet objet principal

que le chercheur peut formuler des questions précises qui lui permettront d'établir ensuite quelle est la nature des informations dont il a besoin pour répondre à ces questions.

### 3.1.2. *Les éléments techniques*

Dans une démarche d'évaluation, le chercheur doit nécessairement faire un choix parmi l'ensemble des données disponibles. Ce choix fait référence à l'échantillonnage, soit la sélection d'un sous-ensemble d'éléments ou d'unités parmi l'éventail de ce qui est disponible. Dans le cadre de l'évaluation d'un programme, Guba et Lincoln (1989) distinguent trois grandes catégories d'acteurs qui peuvent fournir des points de vue différents sur un programme et dont il peut être avantageux de recueillir les opinions: 1) les agents, soit les personnes qui sont engagées dans la conception et l'implantation du programme, 2) les utilisateurs, qui tirent profit directement ou indirectement du programme, et 3) les victimes, c'est-à-dire les personnes qui sont affectées négativement par le programme soit parce qu'elles en sont exclues ou parce qu'elles en subissent des dommages directs ou indirects. Si ces auteurs estiment important de rencontrer des représentants de toutes ces catégories d'acteurs lorsqu'il s'agit d'évaluer un programme, le choix des répondants relève du chercheur et dépend des objectifs de la recherche. Patton (2002) utilise le terme « intentionnel » (*purposeful*) pour caractériser la démarche d'échantillonnage en recherche qualitative. L'idée sous-jacente à l'usage de ce qualificatif est que comme la profondeur est privilégiée au détriment de l'étendue, il ne s'agit pas de rencontrer le plus de personnes possible, mais plutôt d'aller le plus loin possible dans l'information recherchée, en misant sur les personnes qui sont le plus à même de la fournir.

Quelle que soit la classification utilisée, l'idée de base qui guide l'échantillonnage en recherche qualitative c'est que, contrairement aux études quantitatives, le hasard constitue l'option la moins intéressante lorsqu'il s'agit d'obtenir des informations en profondeur. Comme il s'en remet à un nombre limité de sources d'information, le chercheur a intérêt à les sélectionner en s'appuyant sur des critères précis et explicites, dont il peut justifier la pertinence en fonction de ses objectifs de recherche.

Contrairement aux études quantitatives qui s'appuient sur la signification statistique pour fixer la taille de l'échantillon, en recherche qualitative il n'y a pas de limite préétablie. C'est la saturation qui détermine quand l'information collectée est suffisante, à moins que tous les informateurs aient été rencontrés ou que toutes les sources de données aient été explorées (Royse, 2008). Le point de saturation est atteint quand l'analyse des données présente des répétitions et des redondances ou quand les nouvelles

données confirment les résultats déjà obtenus, plutôt que de les élargir<sup>1</sup>. Pour établir si le point de saturation est atteint, il est évidemment nécessaire que la collecte et l'analyse des données soient réalisées simultanément ; c'est une condition essentielle pour que le chercheur puisse juger de l'apport de nouvelles données dans sa connaissance du phénomène à l'étude.

### 3.2. La collecte de données

En recherche évaluative, on distingue généralement trois types de données sur lesquels peut s'appuyer l'évaluation d'un programme : 1) les documents et autres sources de données existantes, 2) les notes de terrain résultant des observations du chercheur et 3) les transcriptions et autres documents générés par des entrevues avec des informateurs. C'est la nature des questions de recherche et la disponibilité des données qui déterminent la pertinence de recourir à chacun de ces types de données. Cependant, comme le mentionnent Royse *et al.* (2006), quand c'est possible, il est souhaitable de s'appuyer sur plus d'une source de données, en prenant soin de préciser quelles données sont nécessaires en fonction des différentes questions de recherche. Cette diversification des sources d'information correspond à la triangulation, une stratégie par laquelle le chercheur diversifie son approche du phénomène à l'étude pour enrichir ou valider son interprétation<sup>2</sup>.

Lorsque la recherche vise à documenter les processus d'un programme ou le fonctionnement d'un organisme, il est généralement utile de commencer la collecte des données en consultant les documents qui existent. Avoir une connaissance minimale du programme ou de l'organisme avant de rencontrer des informateurs constitue une forme de respect pour ces personnes et ce qu'elles font, tout en permettant d'avoir un aperçu des informations à colliger. Les documents utiles pour la recherche peuvent prendre différentes formes : correspondance, procès-verbaux, demandes de financement, statuts et règlements, énoncé de mission, charte, documents promotionnels, vidéos,

- 
1. S'appuyant sur les écrits de Glaser et Strauss (1967), Pirès (1997) distingue deux types de saturation : la saturation théorique, qui s'applique à un concept, et qui est atteinte lorsque les données n'ajoutent aucune propriété nouvelle au concept, et la saturation empirique, qui désigne « le phénomène par lequel le chercheur juge que les derniers documents, entrevues ou observations n'apportent plus d'informations suffisamment nouvelles ou différentes pour justifier une augmentation du matériel empirique » (p. 157).
  2. Denzin (1989) distingue quatre types de triangulation : la triangulation des données, qui peut être relative aux temps de collecte des données, aux contextes dans lesquels l'information est recueillie ou aux sources de données ; la triangulation des chercheurs, qui consiste à associer plusieurs personnes à la réalisation de la recherche ; la triangulation théorique, qui implique le recours à plusieurs cadres théoriques ; la triangulation méthodologique qui consiste à combiner des méthodes qualitatives et des mesures quantitatives.

bref, tout matériel se rapportant au programme et à son fonctionnement. L'avantage des documents écrits, c'est que leur contenu n'est pas influencé par la présence du chercheur ; par contre, lorsqu'il s'agit de documents officiels, leur contenu peut avoir été épuré et ainsi ne refléter que la perspective officielle du programme. C'est pourquoi la documentation écrite doit être complétée par d'autres types de données.

L'observation est une modalité très utilisée en recherche qualitative, car elle permet de voir comment les personnes se comportent dans la réalisation de leurs activités habituelles et de documenter comment les choses se passent au quotidien. Le chercheur peut ainsi s'imprégner de la culture des personnes qu'il observe (Roysse, 2008). Cet avantage est cependant contrebalancé par le fait que c'est une procédure qui exige beaucoup de temps ; en effet, il faut généralement une présence prolongée dans un milieu pour avoir une vue d'ensemble des activités qui s'y déroulent et surtout pour être suffisamment imprégné de sa culture pour interpréter correctement les phénomènes qui sont observés<sup>3</sup>.

Le troisième type de données sur lequel le chercheur peut s'appuyer provient des entrevues réalisées avec des informateurs. L'objectif principal de ces entrevues, qu'elles soient individuelles ou de groupe, est de permettre aux répondants de s'exprimer librement dans leurs propres mots. Cependant, pour faire en sorte que l'information obtenue soit judicieuse, la démarche doit être encadrée par l'utilisation d'un guide d'entrevue qui précise les principaux thèmes que le chercheur veut couvrir. Généralement, ce guide est composé d'un nombre limité de questions très larges qui laissent place à l'expression des idées de la personne interrogée. Dans la conduite des entrevues, le chercheur doit arriver à maintenir un équilibre entre son désir de couvrir les thèmes prévus dans le guide d'entrevue et la nécessité d'être ouvert à de nouvelles informations (Padgett, 1998). Roysse *et al.* (2006) distinguent quatre types d'information qui peuvent être abordées dans un guide d'entrevue : les attitudes (ou opinions), les sentiments/émotions, les connaissances et les comportements<sup>4</sup>.

- 
3. Il existe une très grande variation dans le degré d'engagement de l'observateur dans le milieu. À un extrême du continuum d'engagement, le chercheur adopte une position en retrait, essayant d'avoir le moins possible d'influence sur ce qui se déroule. À l'autre extrême, celle de l'observation participante, il prend part aux activités et il tisse des liens de collaboration avec les gens qui l'entourent.
  4. Selon deMarras (2004), la personne qui fait les entrevues doit éviter les questions longues et compliquées, les questions auxquelles il est possible de répondre par « oui » ou « non », les questions vagues qui invitent à des réponses très générales ainsi que les questions basées sur les théories ou les postulats du chercheur et qui suggèrent une orientation à la personne interrogée. Le chercheur doit par ailleurs être à l'affût des indices subtils qui laissent transparaître que la personne pourrait partager davantage d'informations si elle était invitée à le faire.

Bien que l'entrevue individuelle soit la modalité de collecte de données la plus répandue en recherche qualitative, de plus en plus de chercheurs recourent à l'entrevue de groupe ou au groupe de discussion (*focus group*). Le groupe de discussion est une modalité qui rassemble une dizaine de personnes, qui sont invitées à exprimer leurs points de vue. Simard (1989) le définit comme une méthode de recherche sociale qui consiste à recruter de 6 à 12 personnes répondant à des critères particuliers, à susciter une discussion à partir d'un guide de discussion définissant les thèmes à aborder et à faire une analyse-synthèse du contenu pour relever les principaux messages de même que les points de convergence et de divergence ainsi que les points de vue exceptionnels.

Cette modalité présente l'avantage de pouvoir obtenir la perception subjective de plusieurs personnes simultanément dans un contexte de synergie où elles s'alimentent mutuellement. Cependant, elle est généralement exigeante sur le plan de la préparation, car il n'est pas facile de réunir une dizaine de personnes disponibles simultanément pendant 2 ou 3 heures pour participer à une activité de recherche. De plus, cette modalité exige du chercheur qu'il possède non seulement des habiletés d'intervieweur, mais également des compétences en animation de groupe pour maintenir le cap sur son objectif qui est d'obtenir de l'information. De l'avis de Lee (1999), la compétence de l'animateur à gérer les interactions au sein du groupe est le principal déterminant du succès d'une entrevue de groupe. D'ailleurs, il est recommandé que les groupes de discussion soient animés par deux personnes : l'une menant l'entrevue et l'autre observant les interactions et les comportements non verbaux (Krueger, 1993). Cet observateur a aussi pour fonction de résumer l'essentiel des propos tenus par le groupe sur chacun des points soulevés et de proposer ce résumé aux participants afin que ces derniers l'approuvent ou y apportent les nuances nécessaires. Il s'agit ainsi d'une manière de valider la qualité de l'information colligée et son interprétation<sup>5</sup>.

### 3.3. L'analyse des données

Dans la réalisation d'une recherche évaluative de type qualitatif, l'analyse des données représente souvent l'étape la plus ardue du processus, entre autres parce qu'elle s'appuie essentiellement sur le jugement du chercheur, parce

---

5. Unrau, Gabor et Grinnell (2006) identifient quatre tâches principales de l'animateur d'un groupe de discussion : éviter qu'une personne ou un petit groupe domine la discussion, faciliter l'expression des personnes plus silencieuses, obtenir des réponses de l'ensemble des participants pour s'assurer du plus large éventail possible de points de vue et maintenir un équilibre entre le rôle de modérateur de la discussion et d'intervieweur.

qu'il n'existe pas de procédure unique applicable à toutes les situations pour la réaliser et parce qu'elle passe par la réduction d'une quantité importante de matériel dont la nature est souvent diversifiée. L'analyse des données qualitatives est une tâche qui exige à la fois créativité, esprit de synthèse et introspection (Royse *et al.*, 2006). Parmi les différentes modélisations possibles, celle de Taylor-Powell et Renner (2006) nous apparaît à la fois simple et complète. Ces auteurs suggèrent cinq étapes qui couvrent l'essentiel des opérations nécessaires pour faire une analyse approfondie des données qualitatives : 1) la connaissance d'ensemble des données, 2) l'organisation de l'analyse, qui peut se faire en fonction du contenu (questions, sujets, période ou événements) ou en fonction des répondants (cas, individus, groupe), 3) la catégorisation de l'information en fonction de catégories prédéterminées et émergentes, 4) l'identification de phénomènes récurrents (*patterns*) et de relations à l'intérieur et entre les catégories d'information, et 5) l'interprétation, par la mise en commun de l'ensemble des données. Dans ce processus d'analyse des données qualitatives, il est important de dépasser la stricte énumération des catégories d'information pour en dégager une image plus globale ; c'est cette perspective plus large qui ouvre à l'interprétation. Pour y arriver, deux opérations doivent faire l'objet d'une attention particulière : la catégorisation de l'information, qui exige également de la hiérarchiser, et la mise en relation de l'information. Dans le premier cas, il s'agit d'en arriver à proposer une structure qui regroupe les grandes catégories d'information se rapportant à un même phénomène selon une organisation hiérarchique. Par exemple, un chercheur qui s'intéresserait aux motifs qui incitent les adolescents à participer aux activités d'une maison de jeunes pourrait en arriver à distinguer deux grandes catégories de motifs selon qu'ils se réfèrent à des facteurs intrinsèques (par exemple, le goût de participer à un des projets mis de l'avant à la maison de jeunes) ou à des facteurs extrinsèques (se réfugier à la maison de jeune pour fuir des problèmes familiaux). Dans le second cas, le chercheur tente de déceler la présence de relations entre des informations se rapportant à des éléments différents. Si l'on reprend le même exemple, dans ce cas, le chercheur pourrait examiner s'il existe une relation entre les motifs de participation et l'engagement réel dans les activités de la maison de jeunes.

Miles et Huberman (1994) formulent de très nombreux conseils pour guider l'analyse des données qualitatives. Nous en reprenons ici quelques-uns en fonction des différentes étapes du processus d'analyse des données : 1) appuyer la recherche sur un cadre conceptuel basé sur les connaissances disponibles, en le présentant de façon graphique plutôt que sous forme de texte, 2) réaliser le codage à partir d'un ensemble de thèmes qui s'inscrivent dans une structure, c'est-à-dire qui sont reliés aux autres ou s'en distinguent sur des aspects importants pour la recherche, 3) ne pas hésiter à rédiger des mémos sur les idées, images, impressions, émotions, hypothèses qui viennent

en tête pendant la recherche, 4) privilégier l'usage de graphiques (*réseau de concepts*) et de matrices pour synthétiser les données recueillies, 5) recourir à des matrices chronologiques lorsque la recherche explore un phénomène qui se situe sur une certaine période de temps, 6) utiliser des diagrammes pour modéliser les relations entre les variables. Certains logiciels d'analyse qualitative contiennent des fonctionnalités qui facilitent la réalisation de matrices et de diagrammes; c'est le cas notamment du logiciel N'Vivo.

### 3.4. Le rapport

La rédaction du rapport de recherche est une étape capitale en recherche qualitative car, contrairement à la recherche quantitative où les chiffres et les tableaux de résultats constituent l'essentiel du contenu, en recherche qualitative, un rapport mal rédigé peut invalider tout le travail qui a été fait antérieurement (Royse *et al.*, 2006). Si tout rapport de recherche doit fournir au lecteur les informations permettant de savoir comment la recherche a été faite, en recherche qualitative on s'attend en plus à ce que ce rapport contienne suffisamment d'informations sur le milieu ou le programme à l'étude pour pouvoir juger de la transférabilité des résultats. Ainsi, l'énoncé du problème devrait fournir une description détaillée du programme ou de l'organisme qui fait l'objet d'une évaluation, notamment sous l'angle de l'évolution, du contexte de réalisation et des acteurs concernés (Miles et Huberman, 1994). Un autre enjeu important de la rédaction du rapport est de maintenir l'équilibre entre la présentation de descriptions détaillées et d'extraits de documents, et la place relative accordée aux analyses. Comme le souligne Patton (1987), les descriptions et les citations sont des éléments essentiels en recherche qualitative pour permettre au lecteur de se familiariser avec la situation à l'étude et s'imprégner des points de vue des personnes représentées dans le rapport; cependant, le lecteur n'a pas à prendre connaissance de tout ce qui a été dit ou a été fait. Un rapport intéressant contient suffisamment d'informations descriptives pour permettre au lecteur de suivre les analyses et suffisamment d'analyse pour lui permettre de comprendre les descriptions. Le recours à des matrices, diagrammes et graphiques peut généralement se révéler efficace pour résumer les résultats et pour compléter les descriptions présentées dans le texte. En outre, comme les données obtenues en recherche qualitative ne peuvent être mises à l'épreuve par les tests de signification statistique, il est important que le rapport établisse une distinction entre les résultats qui ont davantage de crédibilité et ceux qui sont plus ambigus. Il revient au chercheur de qualifier la valeur de ses conclusions en spécifiant si elles sont fortement appuyées par les données ou si, au contraire, elles sont plus faibles. Taylor-Powell et Renner (2006) identifient trois précautions à prendre en recherche qualita-

tive: éviter les généralisations, sélectionner soigneusement les citations ou extraits de document qui sont rapportés dans le rapport et préciser les limites de la recherche de même que les explications alternatives aux résultats.

Un élément majeur à prendre en considération lors de la rédaction du rapport et la diffusion des résultats d'une recherche est l'orientation du message (Royse, 2008). La recherche évaluative, qu'elle soit de nature qualitative ou quantitative, ne peut être isolée du contexte dans laquelle elle est produite; en effet, c'est une démarche ayant une portée « politique » en raison des applications qui peuvent en être tirées. Il est donc important que les résultats d'évaluation trouvent leur auditoire, que ce soit auprès des praticiens, du public général, des gestionnaires, des groupes de défense des droits ou, encore, dans certains cas, des médias. Dans le cas des évaluations basées sur des données qualitatives, il est encore plus important de s'assurer que les propos des répondants ne soient pas déformés ou décrits hors contexte, et de faire en sorte que les résultats servent concrètement à l'amélioration des programmes. Il faut aussi diffuser les résultats de la recherche évaluative par le biais d'articles scientifiques à la fois pour contribuer à l'avancement des connaissances, pour faciliter la réplique des programmes dans des contextes différents ou avec une population différente, ainsi que pour faire évoluer la théorisation et la pratique de la recherche évaluative (Royse *et al.*, 2008).

#### 4. LA VALEUR SCIENTIFIQUE DE LA RECHERCHE QUALITATIVE

La question de la rigueur scientifique est parfois considérée comme le talon d'Achille de la recherche qualitative (Padgett, 1998), dans la mesure où il est parfois difficile d'établir le degré d'authenticité des données et la crédibilité des interprétations qui en découlent (Lincoln et Guba, 1985). Cette rigueur, qui correspond à la scientificité de la recherche qualitative, est un sujet qui fait l'objet de débats au sein de la communauté scientifique (Guba et Lincoln, 2005). Certains chercheurs (Schwandt, 1996; Scheurich, 1997 et Smith, 1993, tous dans Guba et Lincoln, 2005) croient qu'il faut rejeter toute tentative pour établir des critères de scientificité de la recherche qualitative, puisque ces derniers sont une menace à l'essence même de cette recherche qui se doit d'être souple et fluide. Cette position ne fait toutefois pas l'unanimité; la plupart des tenants de la recherche qualitative reconnaissent l'importance d'établir des critères pour établir l'authenticité du processus et des résultats, sans pour autant sombrer dans une application simple et mécanique d'une « méthode » de validation (Guba et Lincoln, 2005). Ainsi, lorsqu'il est question des critères de scientificité de la recherche qualitative, il est important de retenir que ces derniers doivent servir de points de repères et ne pas

prendre la forme d'une procédure rigide qui risquerait de compromettre la nature même de la recherche qualitative. Ces points de repère peuvent prendre la forme de principes directeurs qui contribuent à assurer la qualité, la rigueur, l'authenticité et la transparence de la démarche évaluative.

Un ensemble de questions et de techniques sont disponibles pour s'assurer que chaque étape de la recherche évaluative basée sur des méthodes qualitatives est conforme aux critères de validité. Nous proposons une série de techniques permettant de répondre à chacun des critères de la valeur scientifique. Ces techniques, sont tirées du travail de divers auteurs, notamment Shek, Tang et Han (2005), Spencer *et al.* (2003) et Padgett (1998). Elles sont reprises dans le tableau 9.3 pour permettre de saisir les liens qu'elles entretiennent avec chaque critère de scientificité de la recherche qualitative.

## 5. ILLUSTRATION D'UNE DÉMARCHÉ DE RECHERCHE ÉVALUATIVE COMBINANT DONNÉES QUANTITATIVES ET DONNÉES QUALITATIVES

Le Programme d'action communautaire pour les enfants (PACE) est un programme fédéral d'aide financière qui a été mis sur pied au début des années 1990 pour aider les organismes communautaires et les établissements publics à répondre aux besoins des enfants les plus exposés à des facteurs de risque susceptibles d'entraver leur développement. Pour documenter la mise en place et les effets de ce programme, une vaste démarche d'évaluation a été mise en place pour examiner le processus d'allocation du financement aux projets du PACE, les populations rejointes par ces projets, la nature des actions entreprises, le profil des biens et services offerts, les expériences de partenariat vécues dans la réalisation des projets, l'impact du programme sur l'accroissement des services à l'enfance, les facteurs d'influence sur la réalisation des projets et l'efficacité des projets auprès de la clientèle (Turcotte, Simard, Beaudoin et Turgeon, 2000). Certains de ces aspects ont été évalués à l'aide de méthodes quantitatives: la population rejointe, la nature des actions, le profil des biens et services offerts, l'incidence sur l'accroissement des services et l'efficacité des projets. D'autres ont été abordés à l'aide de méthodes qualitatives: le processus d'allocation du financement, les expériences de partenariat et les facteurs d'influence sur la réalisation des projets.

Pour cerner les facteurs d'influence sur la réalisation des projets, une étude approfondie de 19 projets a été conduite (Turcotte, Samson, Lessard et Beaudoin, 1997; 2000). S'appuyant sur la recommandation de Miles et Huberman (1994), dans un premier temps, une revue des écrits sur les

Tableau 9.3

**Questions et techniques permettant d'examiner la scientificité de la recherche qualitative**

| Critères primaires | Questions   | Techniques   |
|--------------------|---|--|
| Authenticité       | <p>Est-ce que la représentation de la perspective subjective démontre une préoccupation pour les différences subtiles dans les discours de chaque participant ?</p> | <ul style="list-style-type: none"> <li>• Maximiser l'inclusion (ajustement du langage, pleine participation de tous les répondants, etc.).</li> <li>• Expliciter pourquoi certaines personnes n'ont pas participé ou n'ont pas été sélectionnées dans l'échantillon.</li> <li>• Détailler la méthode de recrutement des participants.</li> <li>• Illustrer les propos des répondants.</li> </ul>   |
| Crédibilité        | <p>Est-ce que les résultats de la recherche reflètent l'expérience des participants et le contexte de façon vraisemblable ?</p>                                     | <ul style="list-style-type: none"> <li>• Préciser la démarche qui a conduit aux résultats/conclusions.</li> <li>• Vérifier que les résultats/conclusions ont du « sens », sont cohérents.</li> <li>• Vérifier si les conclusions ont une résonance avec les expériences/connaissances d'autres personnes.</li> <li>• Triangler, corroborer ou approfondir les conclusions avec d'autres sources (autres études/données différentes/contextes différents).</li> </ul> |
| Réflexivité        | <p>Est-ce que le processus de recherche témoigne d'une analyse critique ?</p>   | <ul style="list-style-type: none"> <li>• Justifier le type de devis et les choix effectués (méthode de collecte des données, durée, choix de l'échantillon).</li> <li>• Justifier le choix du programme et du contexte de l'évaluation.</li> <li>• Explorer des pistes explicatives rivales.</li> <li>• Démontrer une capacité à se remettre en question.</li> </ul>   |

Tableau 9.3 (suite)

| Critères Primaires   | Questions  | Techniques  |
|----------------------|--|---|
| Intégrité            | <p>Est-ce que la recherche démontre un effort soutenu et répété pour assurer la validité de la démarche, une analyse rigoureuse des données ainsi qu'une présentation prudente des résultats ?</p> | <p>Lors de la collecte de données :</p> <ul style="list-style-type: none"> <li>• répondre aux questions suivantes : qui/comment/dans quel contexte/avec quel outil ?</li> <li>• s'assurer d'avoir un engagement prolongé dans le milieu</li> <li>• décrire avec précision le contexte des observations</li> <li>• tenir un journal de bord avec des notes et des mémos sur la démarche</li> <li>• rendre disponible la transcription des verbatim, sur demande</li> </ul> <p>Lors de l'analyse :</p> <ul style="list-style-type: none"> <li>• présenter le cadre conceptuel</li> <li>• faire la description des catégories analytiques</li> <li>• faire des vérifications des analyses auprès de pairs/participants/experts</li> <li>• analyser les cas « négatifs »</li> </ul> <p>Lors de la présentation des résultats :</p> <ul style="list-style-type: none"> <li>• identifier les possibles sources de biais ou d'erreur et comment elles ont été corrigées, le cas échéant</li> <li>• réfléchir à l'influence du chercheur sur les résultats</li> </ul> |
| Critères secondaires | Questions  | Techniques  |
| Transparence         | <p>Est-ce que les décisions méthodologiques, les biais possibles et les interprétations alternatives ont été abordées de façon explicite ?</p>   | <ul style="list-style-type: none"> <li>• Exposer comment les décisions ont été prises (par qui, sur quelle base?)</li> <li>• Mettre en évidence les idéologies/valeurs et perspectives de l'équipe de recherche et leur impact sur la démarche de recherche.</li> <li>• Offrir une description détaillée de la nature et de la procédure de sélection des critères de succès</li> <li>• Expliciter les points de vue divergents et leur origine</li> <li>• Présenter les effets imprévus, leur incidence sur les participants et les raisons pour lesquelles ils se sont présentés</li> </ul>   |

|                   |  |
|-------------------|--|
| <p>Vivacité</p>   | <ul style="list-style-type: none"> <li>• Offrir des descriptions détaillées, étoffées et fidèles de la situation.</li> <li>• Faire valider les descriptions par des personnes extérieures à la recherche pour évaluer leur conformité avec le phénomène/contexte à l'étude.</li> <li>• Fournir des informations suffisamment détaillées pour permettre au lecteur de faire des comparaisons avec d'autres contextes/phénomènes.</li> </ul>   |
| <p>Créativité</p> | <ul style="list-style-type: none"> <li>• Présenter les résultats de façon à offrir un angle d'appréhension imaginaire et original.</li> <li>• Explorer différentes options pour présenter les résultats.</li> <li>• Prendre en note ses impressions, émotions, réactions et les utiliser pour l'analyse des données.</li> </ul>  |
| <p>Profondeur</p> | <ul style="list-style-type: none"> <li>• Préciser en quoi la saturation des données est atteinte.</li> <li>• Se référer à la littérature scientifique pour situer les résultats au regard des connaissances questions/enjeux actuels.</li> <li>• Identifier comment les résultats concourent à l'avancement des connaissances.</li> <li>• Préciser comment les résultats peuvent servir pour l'amélioration de la pratique/de la théorie.</li> <li>• Présenter clairement les limites de la recherche et les nouvelles pistes de recherche qui s'en dégagent.</li> </ul>   |
| <p>Congruence</p> | <ul style="list-style-type: none"> <li>• Exposer clairement le but et les objectifs de la recherche.</li> <li>• S'assurer que les résultats sont clairement en lien avec le but, les objectifs et le programme.</li> <li>• Expliquer, le cas échéant, pourquoi certains objectifs n'ont pu être atteints.</li> <li>• S'assurer du lien entre le cadre conceptuel, l'analyse et les résultats.</li> <li>• Établir des liens avec les recherches antérieures.</li> <li>• Expliciter pour quelle population ou à quel contexte les résultats de l'étude peuvent être transférables.</li> <li>• Reconnaître les caractéristiques de la recherche qui restreignent la transférabilité des résultats.</li> </ul> |

Tableau 9.3 (suite)

| Critères secondaires | Questions   | Techniques   |
|----------------------|---|--|
| Sensibilité          | Est-ce que la recherche a été réalisée en étant sensible aux contextes humains, sociaux et culturels? | <ul style="list-style-type: none"> <li>• Présenter la façon dont les participants ont été recrutés et ont donné leur consentement.</li> <li>• Décrire les mesures prises pour assurer la confidentialité des données et l'anonymat des participants.</li> <li>• Offrir du soutien aux participants lorsque nécessaire (information/services/conseils).</li> <li>• Anticiper les effets négatifs potentiels de la recherche et préciser comment ils ont été abordés.</li> <li>• Tenir compte des préoccupations des groupes opprimés ou marginalisés.</li> <li>• Voir à ce que chaque participant puisse tirer des avantages de sa participation à la recherche.</li> <li>• S'assurer que les résultats de l'étude servent à l'amélioration des services offerts aux populations étudiées.</li> </ul> |

Source : Ce tableau a été élaboré à partir de Whittemore *et al.* (2001), Shek, Tang et Han (2005), Spencer *et al.* (2003) et Padgett (1998).

facteurs de réussite des programmes de prévention et d'intervention a été réalisée. Cette revue des écrits a conduit à élaborer un cadre conceptuel à double entrée distinguant trois types de facteurs (intervenants, projet, organisme) susceptibles d'avoir une influence aux trois étapes principales de la réalisation d'un projet : la conception, la mise en œuvre, l'évaluation des effets. Les 19 projets ont été sélectionnés avec un souci de diversification, en tenant compte de cinq éléments : 1) les axes d'intervention (services et objectifs) ; 2) la région géographique ; 3) le type de milieu (urbain, régional, semi-rural) ; 4) le type d'organisme responsable de la gestion du projet (organisme communautaire et établissement du réseau) et 5) le budget alloué.

La collecte des données a été effectuée selon trois modalités : l'analyse documentaire, l'entrevue et l'observation. Ainsi, avant de se rendre sur le terrain, les intervieweurs ont passé en revue les documents consignés dans le dossier des projets (demande de subvention, rapports périodiques, sommaire du projet). Cet exercice visait à développer une première connaissance du projet afin d'être bien préparé aux entrevues. Des entrevues semi-dirigées, dont le contenu a été guidé par le modèle conceptuel, ont par la suite été réalisées auprès de différents types d'informateurs. Au total, 93 personnes ont été interviewées, 36 participants, 20 partenaires, 15 responsables de projets, 14 intervenants et 8 personnes qui agissaient à la fois comme responsables et comme intervenantes. Les données ont été analysées en se référant aux techniques proposées par Miles et Huberman (1994) pour les analyses de cas (*within-case display*) et les analyses intersites (*cross-case display*). Dans un premier temps, l'évolution de chaque projet a été reconstitué depuis sa conception jusqu'à la collecte des données, en faisant ressortir les facteurs qui ont influencé cette évolution. À l'étape subséquente, les segments d'information issus de chacun des projets ont été superposés afin de voir si certains *patterns* pouvaient en être dégagés.

Les données indiquent que l'étape de la conception est particulièrement exigeante pour les intervenants plus habitués à l'action directe qu'à la planification à long terme. Cette tâche leur est d'autant plus difficile qu'ils doivent composer avec des moyens limités à l'intérieur d'un échancier serré. La possibilité de s'associer à des partenaires du milieu est une façon efficace de contourner ces difficultés. Leur contribution facilite la structuration de la demande de subvention et elle est susceptible de conduire à une meilleure définition des besoins et à un choix de stratégies d'intervention plus appropriées. Au moment de la mise en œuvre, les principaux obstacles sont liés aux limites financières et à l'imprécision des objectifs du projet particulièrement lorsque le personnel chargé de l'implantation n'a pas été associé à la conception. Les données font également ressortir l'importance de la précision, de l'adéquation, du réalisme et de la souplesse des modalités d'action, de même que l'importance de la relation entre les participants et

les intervenants. Cette relation apparaît tributaire de trois caractéristiques des intervenants : leur connaissance de la problématique, leurs habiletés d'intervention et leur engagement dans le projet. Le cadre organisationnel est également très important : lorsque l'intervenant entretient des relations positives et de soutien avec les autres membres de l'équipe, le projet est facilité. Par contre, s'il travaille de façon isolée, lorsqu'il y a des tensions internes ou que le personnel change régulièrement, le travail est alourdi et compliqué. En ce qui concerne l'évaluation, les données indiquent que, pour les intervenants et les responsables des projets, l'évaluation sommative est difficile à actualiser, d'une part, parce que leurs objectifs ne sont pas toujours formulés en termes mesurables et, d'autre part, parce qu'ils ne disposent pas des connaissances et des outils nécessaires. Les intervenants s'appuient essentiellement sur leurs propres observations et sur les commentaires des participants pour porter un jugement sur leur action.

Les données suggèrent quelques avenues pour améliorer l'intervention auprès des enfants en difficulté : 1) bien cerner les besoins de la population visée, 2) élaborer une intervention ciblée et orientée vers des objectifs opérationnels et cohérents, 3) travailler en concertation avec les autres organismes de la communauté et 4) disposer d'une stratégie d'évaluation formative permettant de porter un jugement continu sur le déroulement du projet ainsi qu'une procédure d'évaluation sommative pour vérifier si le projet, tel qu'il est réalisé, conduit à l'atteinte des objectifs visés.

## BIBLIOGRAPHIE

- COMEAU, Y. (2001). *L'économie sociale et le plan d'action du Sommet sur l'économie et l'emploi*, Québec, École nationale d'administration publique, Centre de recherche sur les services communautaires, Université Laval.
- DEMARRAIS, K. (2004). « Qualitative interview studies: Learning through experience », dans K. DeMarrais et S.D. Laplan (dir.), *Foundations for Research: Methods of Inquiry in Education and the Social Sciences*, Mahwah, Lawrence Erlbaum Associates Publishers, p. 51-68.
- DENZIN, N.K. (1989). *The Research Act: A Theoretical Introduction to Sociological Methods*, 3<sup>e</sup> éd., Englewood Cliffs, Prentice Hall.
- DESLAURIERS, J.P. (1991). *Recherche qualitative. Guide pratique*, Montréal, McGraw-Hill.
- DESLAURIERS, J.P. et M. KÉRÉSIT (1997). « Le devis de recherche qualitative », dans P. Poupart (dir.), *La recherche qualitative, Enjeux épistémologiques et méthodologiques*, Montréal, Gaëtan Morin éditeur, p. 85-112.

- FETTERMAN, D.M. (1996). «Introduction», dans D.M. Fetterman, S.J. Kaftarian et A. Wandersman (dir.), *Empowerment Evaluation, Knowledge and Tools for Self-Assessment & Accountability*, Thousand Oaks, Sage Publications, p. 3-48.
- GALL, M.D., W.R. BORG et J.P. GALL (1996). *Educational Research: An Introduction*, White Plains, Longman Publishers
- GLASER, B.G. et A.L. STRAUSS (1967). *The Discovery of Grounded Theory; Strategies for Qualitative Research*, Chicago, Aldine Publications.
- GREENE, J.C. (1994). «Qualitative program evaluation: Practice and promise», dans N.K. Denzin et Y.S. Lincoln (dir.), *Handbook of Qualitative Research*, Thousand Oaks, Sage Publications, p. 530-544.
- GUBA, E.G. et Y.S. LINCOLN (1989). *Fourth Generation Evaluation*, Newbury Park, Sage Publications.
- GUBA, E.G. et Y.S. LINCOLN (2005). «Paradigmatic controverses, contradictions, and emerging confluences», dans N.K. Denzin et Y.S. Lincoln (dir.). *Handbook of Qualitative Research*, 3<sup>e</sup> éd., Thousand Oaks, Sage Publications, p. 191-215.
- HUBERMAN, A.M. et M. MILES (1985). «Assessing local causality in qualitative research», dans D.N. Berg et K.K. Smith (dir.), *Exploring Clinical Methods for Social Research*, Thousand Oaks, Sage Publications, p. 351-381.
- KRUEGER, R.K. (1993). «Quality control in Focus Group research», dans D.L. Morgam (dir.), *Successful Focus Groups. Advancing State of the Art*, Newbury Park, Sage Publications, p. 65-88.
- L'ÉCUYER, R. (1990). *Méthodologie de l'analyse développementale de contenu: méthode GPS et concept de soi*, Québec, Presses de l'Université du Québec.
- LEE, T.W. (1999). *Using qualitative methods in organizational research*, Thousand Oaks, Sage Publications.
- LINCOLN, Y.S. et E.G. GUBA (1985). *Naturalistic Inquiry*, Beverly Hills, Sage.
- MAYER, R. et J.P. DESLAURIERS (2000). «Quelques éléments d'analyse qualitative», dans R. Mayer et al. (dir.), *Méthodes de recherche en intervention sociale*, Montréal, Gaëtan Morin éditeur, p. 159-190.
- MILES, M.B. et A.M. HUBERMAN (1984). *Qualitative Data Analysis*, Thousand Oaks, Sage Publications.
- MILES, M.B. et A.M. HUBERMAN (1994). *Qualitative Data Analysis. An Extended Sourcebook*, 2<sup>e</sup> éd., Thousand Oaks, Sage Publications.
- MORGAN, D.L. (2007). «Paradigms lost and pragmatism regained: Methodological implications of combining qualitative and quantitative methods», *Journal of Mixed Methods Research*, vol. 1, n° 1, p. 48-76.
- OUELLET, F. et M.C. SAINT-JACQUES (2000). «Les techniques d'échantillonnage», dans R. Mayer et al. (dir.), *Méthodes de recherche en intervention sociale*, Montréal, Gaëtan Morin éditeur, p. 71-90.
- PADGETT, D. (1998). *Qualitative Methods in Social Work Research: Challenges and Rewards*, Thousand Oaks, Sage Publications.
- PATTON, M.Q. (1982). *Practical Evaluation*, Bervely Hills, Sage Publications.

- PATTON, M.Q. (1987). *How to Use Qualitative Methods in Evaluation*, Newbury Park, Sage Publications.
- PATTON, M.Q. (1990). *Qualitative Evaluation and Research Methods*, Newbury Park, Sage Publications.
- PATTON, M.Q. (1997). *Utilization-focuses Evaluation: The New Century Text*, Thousand Oaks, Sage Publications.
- PATTON, M.Q. (2002). *Qualitative Evaluation and Evaluation Methods*, 3<sup>e</sup> éd., Thousand Oaks, Sage Publications.
- PAWSON, R et N. TILLEY (1997). *Realistic Evaluation*, Thousand Oaks, Sage Publications.
- PIRÈS, A.P. (1997). «Échantillonnage et recherche qualitative: essai théorique et méthodologique», dans P. Poupard et al. (dir.), *La recherche qualitative. Enjeux épistémologiques et méthodologiques*, Montréal, Gaëtan Morin éditeur, p. 113-172.
- PITMAN, M.A. et J.A. MAXWELL (1992). «Qualitative approaches to evaluation: Models and methods», dans M.D. LeCompte, L. Milroy et J. Preissle (dir.), *The Handbook of Qualitative Research in Education*, San Diego, Academic Press Inc., p. 729-770.
- RISSMAN, C.K. (1994). *Qualitative Studies in Social Work Research*, Thousand Oaks, Sage.
- ROSSI, P.H., M.W. LIPSEY et H.E. FREEMAN (2004) *Evaluation: A systematic approach*, 7<sup>e</sup> éd., Thousand Oaks, Sage Publications.
- ROSSMAN, G. et S. RALLIS (2003). *Learning in the Field: An Introduction to Qualitative Research*, 2<sup>e</sup> éd., Thousand Oaks, Sage Publications.
- ROYSE, D. (2008). *Research Methods in Social Work*, Toronto, Brooks/Cole-Thomson Learning.
- ROYSE, D. et al. (2006). *Program Evaluation: An Introduction*, 4<sup>e</sup> éd., Toronto, Thomson Brooks/Cole.
- RUBIN, A. et E.R. BABBIE (2008). *Research Methods for Social Work*, Belmont, Thompson/Brooks/Cole.
- SHAW, I. (1999). *Qualitative Evaluation*, Thousand Oaks, Sage Publications.
- SHEK, D., V.M. TANG et X.Y. HAN (2005). «Evaluation of evaluation studies using qualitative research methods in the social work literature (1990-2003): Evidence that constitutes a wake-up call», *Research on Social Work Practice*, vol. 15, n° 3, p. 180-194.
- SIMARD, G. (1989). *Animer, planifier et évaluer l'action: la méthode du «Focus group»*, Laval, Mondia.
- SPENCER, L. et al. (2003). *Quality in Qualitative Evaluation: A Framework for Assessing the Evidence*, Londres, National Centre for social Research, Cabinet office.
- TARD, C. et al. (1997). *Outils d'évaluations pour l'amélioration de la qualité des programmes*, Québec, Centre de recherche sur les services communautaires, Université Laval.

- TAYLOR-POWELL, E. et M. RENNER (2006). « Analyzing qualitative data », dans Y.A. Unrau, P.A. Gabor et R.M. Grinnell (dir.), *Evaluation in Social Work. The Art and Science of Practice*, 4<sup>e</sup> éd., New York, Oxford University Press, p. 341-360.
- TURCOTTE, D. (2000). « Le processus de la recherche sociale », dans R. Mayer *et al.* (dir.), *Méthodes de recherche en intervention sociale*, Montréal, Gaëtan Morin éditeur, p. 39-68.
- TURCOTTE, D. *et al.* (1997). *La dynamique de réalisation des projets du PACE*, Québec, Centre de recherche sur les services communautaires, Université Laval.
- TURCOTTE, D. *et al.* (2000). « Les facteurs d'influence sur la réalisation des projets du PACE », dans D. Turcotte *et al.*, *L'éducation familiale en milieu défavorisé : portrait québécois du Programme d'action communautaire pour les enfants*, Québec, Éditions S. Harvey, p. 125-144.
- UNRAU, Y.A., P.A. GABOR et R.M. GRINELL (2006). *Evaluation in Social Work: The Art and Science of Practice*, 4<sup>e</sup> éd., New York, Oxford University Press.
- WEINBACH, R.W. (2005). *Evaluating Social Work Services and Programs*, Boston, Pearson/Allyn and Bacon.



# CHAPITRE 10

---

## **La dimension formative du bilan d'implantation** Des suites d'un programme à l'état des acquis subjectivement perçus par les intervenants au programme

Carl Lacharité  
Département de psychologie  
Université du Québec à Trois-Rivières

Dans le domaine des services psychologiques, sociaux et éducatifs, un programme repose étroitement sur la qualité de l'appropriation, par les agents responsables de son application<sup>1</sup>, des valeurs, des concepts, des principes et des stratégies d'action qui le charpentent. Dans ce sens, l'évaluation d'un programme, même si parfois elle se concentre, voire se limite, à des éléments objectifs, n'est jamais complète si elle ne tient pas compte de la dimension subjective chez les agents<sup>2</sup> de celui-ci. Or c'est probablement cette dimension subjective qui recèle les informations les plus immédiatement disponibles ainsi que les plus nuancées auxquelles les agents de programme ont directement accès à propos des processus mis en branle par le programme, justement parce qu'ils en sont les dépositaires privilégiés.

Par contre, la subjectivité a souvent mauvaise réputation dans les démarches usuelles de gestion et d'évaluation d'un programme. Depuis près de deux siècles, la perspective biomédicale constitue le modèle dominant utilisé pour « penser » les programmes sociaux et éducatifs, un peu comme si la matérialité des corps et les enjeux qui la concernent pouvaient servir de cadre de référence par excellence pour définir et comprendre la substance de ce qui est psychologique, éducatif, social ou culturel, de même que les enjeux qui traversent ces plans de la réalité humaine (Foucault, 1972; Miller et Crabtree, 2008). Notamment, cette perspective nous a appris (ou convaincu) que raison et subjectivité ne font pas bon ménage: un point de vue subjectif est d'emblée suspecté de partialité, d'émotivité et de manque de rigueur. Il faut toutefois comprendre que ce que l'on appelle la « raison » n'est pas une démarche désincarnée et décontextualisée. Elle constitue un phénomène essentiellement contesté auquel s'affrontent diverses formes de rationalité. Dans le domaine des activités humaines, la forme actuelle de rationalité hégémonique repose sur 1) une démarche d'objectivation ou de construction abstraite des personnes et des processus interpersonnels et socioculturels, et 2) une rupture radicale entre les personnes ainsi objectivées et les personnes qui réalisent cette objectivation. Au-delà de cette hégémonie, d'autres formes de rationalité existent et se sont développées, qui visent non pas à mettre de côté ou à annihiler la position subjective des personnes à propos d'elles-mêmes et des conditions qui façonnent leur vie quotidienne

- 
1. L'expression « agents de programme » réfère, tout au long du texte, aux intervenants, gestionnaires, bénévoles et autres acteurs *directement* impliqués dans la réalisation des actions découlant d'un programme. Il ne s'agit donc pas uniquement des personnes ayant un contact avec les usagers du programme, mais également des cadres intermédiaires et des superviseurs ou conseillers cliniques ayant un mandat d'encadrement de la pratique des personnes en contact avec les usagers à l'intérieur du programme.
  2. La dimension subjective touche également le point de vue des usagers d'un programme. Le présent chapitre s'attardera uniquement à la contribution que les agents d'un programme peuvent avoir dans l'évaluation de ce dernier. Il faut toutefois noter que des processus très semblables permettent de tenir compte du point de vue subjectif des usagers dans l'évaluation d'un programme.

mais plutôt à comprendre cette position et à partager cette compréhension avec les personnes concernées. Plusieurs courants en sciences humaines et sociales reposent essentiellement sur une telle prise en compte formelle de la subjectivité des individus: l'anthropologie (Geertz, 1973, 1983; Turner et Bruner, 1986) l'ethnographie (Conquergood, 1991; Geertz, 1973, 1983; Mauss, 1902/2002; Smith, 1999, 2005), l'ethnométhodologie (De Fornel, Ogien et Quéré, 2001; Garfinkel, 2007) et le constructionnisme social en psychologie (Billig, 1987, 1991; Gergen, 1982, 1985; Shotter, 1993). Les concepts et les méthodes développés dans ces champs d'études contribuent à structurer des formes de rationalité qui s'appuient sur des pratiques de collaboration sociale permettant aux personnes de prendre du recul par rapport à leur expérience immédiate.

L'économie d'un travail d'analyse du point de vue subjectif des agents d'un programme constitue probablement le principal facteur de rupture entre les résultats qui découlent habituellement de son évaluation et sa transformation effective sur le terrain. La démarche d'appropriation, par les agents d'un programme, des constats d'une évaluation de ce dernier ne peut reposer uniquement sur la lecture (même attentive) d'un rapport d'évaluation ou l'exposition à des présentations publiques (aussi nombreuses et détaillées soient-elles) par les responsables de l'évaluation. Cette appropriation doit être conçue, dès le départ, comme une activité s'appuyant sur la prise en compte formelle du point de vue subjectif des agents du programme (si possible, l'ensemble de ceux-ci; sinon, un sous-groupe de personnes clés).

L'objectif de ce chapitre est de présenter les principaux éléments d'une démarche ethnographique appliquée à la prise en compte du point de vue des agents dans l'évaluation d'un programme. Pour ce faire, je partirai d'une situation réelle d'évaluation d'un programme s'adressant à des familles vivant des difficultés majeures à répondre aux besoins de leurs enfants. Le programme en question, Programme d'aide personnelle, familiale et communautaire (PAPFC: Lacharité, Fafard et Bourassa, 2005), est actuellement utilisé comme cadre de référence dans plusieurs régions au Québec de même qu'en France et au Brésil. La prochaine section présente de manière schématique les éléments constitutifs de la démarche de bilan des processus et des acquis d'un programme par les agents. La section suivante décrit l'application de cette démarche dans l'évaluation de l'implantation du PAPFC sur un territoire donné. Le chapitre conclut par une description des principaux enjeux dans l'application d'une telle démarche.

## 1. DÉMARCHE ETHNOGRAPHIQUE DU BILAN DES PROCESSUS ET DES ACQUIS D'UN PROGRAMME PAR LES AGENTS

Dans le cadre d'une évaluation de programme, ce type de démarche vise trois objectifs: 1) de mettre en place des conditions qui favorisent, chez les agents de programme, une prise de distance par rapport aux décisions et aux actions qu'ils ont réalisées dans le cadre de ce programme au cours d'une période donnée, 2) de produire une synthèse collective, sous forme de bilan, des points forts et des aspects à corriger et 3) de produire un plan d'action pour la poursuite de l'implantation ou pour la suite des opérations. Il s'agit d'une démarche dont les retombées sont prioritairement dirigées vers les agents de programmes eux-mêmes. Ces retombées ont toutefois la possibilité d'être également dirigées vers d'autres interlocuteurs (en particulier, les personnes ayant un mandat de reddition de comptes face au programme ou des personnes dans d'autres territoires souhaitant implanter le programme).

Les deux aspects qui organisent l'exploration et l'analyse du point de vue des agents du programme sont le *plan organisationnel* de celui-ci et le *plan d'utilisation des services* par les usagers (Rossi, Lipsey et Freeman, 2004). Le plan organisationnel d'un programme a trait aux éléments essentiels permettant de le mettre en œuvre (les ressources humaines, financières et matérielles, la définition des rôles et tâches, les structures de gestion et d'imputabilité, la structure de partenariat, etc.). Le plan d'utilisation des services réfère à l'ensemble des interfaces entre le programme et les usagers (critères d'inclusion et d'exclusion des usagers, procédures de recrutement ou de sélection, procédures de tenue de dossier incluant une description des actions particulières posées par les agents de programmes auprès des usagers, procédures permettant aux usagers de rétroagir à propos de ces actions, etc.).

Sur le plan opérationnel, une démarche ethnographique met l'accent sur la capacité des individus impliqués dans une situation à attribuer des significations à ce qu'ils vivent, à développer des savoirs et des concepts leur permettant de comprendre leur expérience et de dialoguer au sujet de celle-ci, c'est-à-dire de partager avec d'autres et de réfléchir conjointement. Une des principales conditions à mettre en place pour réaliser une telle démarche est la constitution d'un comité local composé soit de l'ensemble des agents du programme (si cela est faisable) ou de personnes clés. Ce comité local peut être créé de toutes pièces pour les besoins de la procédure d'évaluation du programme. Il s'agit alors d'un groupe de discussion si le nombre de personnes est restreint (par exemple, vingt ou moins) ou d'un forum si le nombre de personnes devient plus important. Ce comité n'a alors

que cette fonction. Il peut aussi s'agir d'un comité existant et ayant d'autres fonctions (par exemple, comité de coordination, équipe clinique, table de concertation, table d'orientation, etc.) auquel peut s'ajouter d'autres acteurs considérés comme étant pertinents. Le but visé est de créer un espace collectif à l'intérieur duquel les agents du programme ont la possibilité d'exprimer et de partager leur propre expérience<sup>3</sup> du programme.

La démarche de réflexion collective nécessite habituellement au moins deux rencontres d'environ trois heures, ou une rencontre s'étalant sur une journée complète. Dans certains cas, elle peut nécessiter plus de temps. L'animation de cette ou de ces rencontres peut être assumée par une ou deux personnes du comité, mais il est cependant intéressant qu'elle soit assumée par une ou deux personnes externes au comité (consultants, chercheurs, etc.). Le ou les animateurs externes doivent toutefois posséder une connaissance fonctionnelle du programme (cadre conceptuel, principes, modèle logique, stratégies, pratiques, etc.) et une compréhension de la problématique vécue par les usagers visés par celui-ci. Les animateurs, qu'ils soient internes ou externes, doivent avoir un minimum d'expérience à mener des entretiens ethnographiques ou expérientiels (voir à ce sujet Campbell et McGregor, 2002; Kaufman, 2007).

Les participants à ces rencontres doivent avoir fait un travail de réflexion préalable à propos des principaux éléments du plan organisationnel et du plan d'utilisation des services du programme. Ce travail préparatoire peut être individuel ou en sous-groupe. Il peut se faire de manière non structurée (par exemple, points retranscrits sur une feuille résumant les principales idées du ou des participants à propos de leur expérience du programme) ou structurée (par exemple, questionnaire). Dans l'illustration présentée à la section suivante, une liste de questions a été utilisée par les participants pour soutenir leur réflexion.

Les rencontres sont organisées en trois temps: 1) expression et partage des expériences, 2) bilan synthèse et 3) plan d'action. La première étape vise à permettre aux agents de programme qui participent aux rencontres de mettre en relief et d'enrichir leur expérience du programme. L'expérience d'aucun agent ou catégorie d'agents ne doit être privilégiée aux dépens des autres. Les hiérarchies institutionnelles ou intersectorielles peuvent ici jouer de mauvais tours en provoquant un aplatissage des points de vue qui

---

3. Il faut noter ici que le terme « expérience » dépasse largement la référence aux sentiments et aux opinions des personnes. Il englobent quatre catégories d'éléments subjectifs: 1) les valeurs, savoirs, concepts et représentations, 2) les préoccupations, inquiétudes, malaises et problèmes perçus, 3) les intentions, aspirations et projets et 4) les initiatives, actions et affects. Il n'est pas possible, dans le cadre de ce chapitre, de décrire de manière exhaustive cette notion. Pour plus de détails, le lecteur est référé au texte de Lacharité (sous presse).

ne correspondent pas à ceux qui jouissent d'un privilège tacite particulier. Par exemple, les préoccupations ayant trait aux coûts d'un programme et les exigences d'équilibre budgétaire qui sont l'apanage de certains gestionnaires de programme peuvent, aussi légitimes soient-elles, être placées au-dessus des formes d'expression de l'expérience des autres agents et agir ainsi comme un frein à leur expression ou comme commentaire à propos de celles-ci lorsqu'elles sont exprimée. Autre exemple, le point de vue des agents professionnels peut facilement être considéré comme étant plus pertinent et central que le point des agents non professionnels (bénévoles, intervenants communautaires, etc.). C'est face à de tels phénomènes que la dimension subjective a acquis sa mauvaise réputation. Dans un tel cas, la subjectivité est maladroitement conçue comme un produit raffiné qu'il faut prendre tel quel. Au contraire, les points de vue exprimés spontanément pas les participants doivent plutôt être conçus comme un matériau brut sur lequel ils ont à travailler. C'est le rôle des animateurs de guider ce mouvement individuel et collectif de distanciation réflexive par rapport aux expressions initiales et spontanées des points de vue. Bourdieu (2002) parle ici d'un processus « d'objectivation participante ».

Un autre piège fréquent lors de la première étape est que les participants comprennent la démarche comme étant un moment où ils doivent exprimer leurs « opinions » à propos du programme (est-il efficace ou inefficace, etc.). La démarche ethnographique présentée ici se situe très loin de l'enquête d'opinions. À n'en pas douter, le contexte des rencontres constitue un cadre où des opinions sont spontanément exprimées. Par contre, ce sont les expériences vécues qui fondent ces opinions qui constituent la raison d'être de la démarche. Dans le cas des opinions, il s'agit pour les animateurs de faciliter leur incarnation et leur contextualisation à l'intérieur d'expériences particulières vécues par les participants qui les expriment.

De plus, les agents de programme ayant un contact direct avec les usagers ont habituellement une représentation riche et nuancée de l'expérience que ces derniers peuvent avoir des activités ou services auxquels ils ont participé (Lacharité, sous presse). Les animateurs doivent porter une attention particulière à rendre audible la voix ou visible la présence des usagers à travers l'expérience qu'en ont ces agents de programmes lors des rencontres<sup>4</sup>.

4. Idéalement, des usagers devraient avoir la possibilité d'exprimer formellement leur point de vue à propos du programme. Cela constitue un autre type de démarche que celle décrite dans ce chapitre. Toutefois, les agents étant en contact direct avec les usagers ont une responsabilité d'être des porte-parole de ces derniers dans les lieux et les moments où des décisions sont prises qui pourraient avoir une incidence sur leur vie (Waldegrave, Tamasese, Tuhaka et Campbell, 2003). La mise en évidence de cette responsabilité de porte-parole fait partie du travail préparatoire inhérent à la démarche

Les étapes du bilan synthèse et de la confection d'un plan d'action ne sont pas différentes de ce que les agents de programme connaissent et pratiquent régulièrement dans le cadre de leurs fonctions. Il s'agit ici de relever les idées majeures qui ressortent de la réflexion collective et de faire des liens avec les éléments qui constituent le plan organisationnel et le plan d'utilisation des services du programme. Ce bilan peut servir d'ébauche de rapport formel, incluant des recommandations, pour le bénéfice de personnes concernées à l'extérieur du programme (les décideurs institutionnels, les bailleurs de fonds, etc.). Le plan d'action décrit, quant à lui, les objectifs, les cibles, les moyens et les acteurs qui seront impliqués dans la poursuite de l'implantation du programme ou dans sa transformation au cours d'une période donnée (par exemple, la prochaine année).

## 2. LE POINT DE VUE DES ACTEURS DANS L'ÉVALUATION DE L'IMPLANTATION D'UN PROGRAMME DESTINÉ AUX FAMILLES VIVANT EN SITUATION DE VULNÉRABILITÉ

Cette section illustre brièvement l'application de la démarche décrite plus haut dans le cadre de l'implantation d'une nouvelle génération du Programme d'aide personnelle, familiale et communautaire (PAPFC) sur le territoire d'un centre de santé et de services sociaux (CSSS) au Québec. Les détails de cette opération se trouvent dans deux rapports (Lacharité, Bourassa et Le Marois, 2004; Fafard et Lacharité, 2007).

Le programme est décrit dans un manuel (Lacharité *et al.*, 2005). Le modèle logique est présenté à la figure 10.1. Il s'agit d'un programme reposant sur un partenariat territorial entre un centre jeunesse dont le mandat est la protection des enfants dans le cadre de la Loi sur la protection de la jeunesse, un centre de santé et de services sociaux dont le mandat est d'offrir des services de première ligne et un ou des organismes communautaires autonomes dont la mission est d'offrir du soutien aux familles. Il s'appuie sur un cadre de référence écosystémique permettant de comprendre le phénomène des négligences graves envers les enfants (Lacharité, Éthier et Nolin, 2006) et s'inscrit dans une logique de services intégrée (Lacharité, Pinard, Giroux et Cossette, 2007). Sa structure de gestion repose sur un *comité de coordination local* où sont représentés des gestionnaires et des intervenants ayant un contact direct avec les familles (en particulier, ceux œuvrant dans les volets d'animation de groupes de parents, d'activités collectives auprès des familles et de soutien paraprofessionnel) de chaque établissement ou

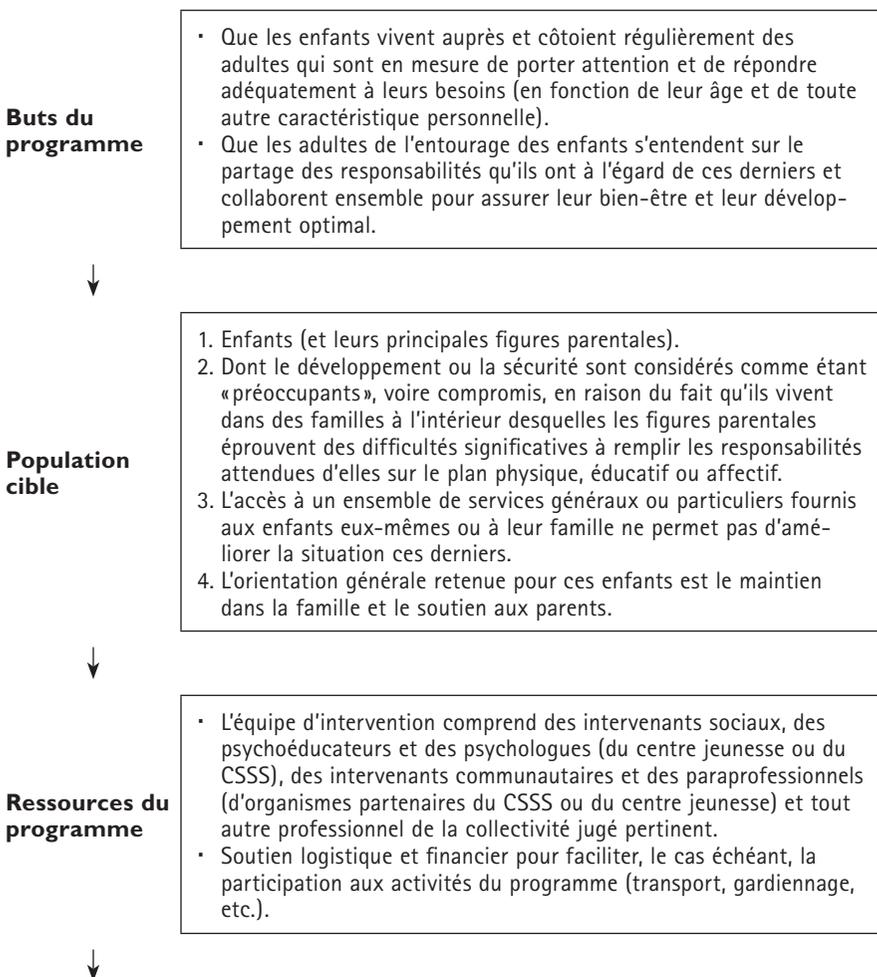
---

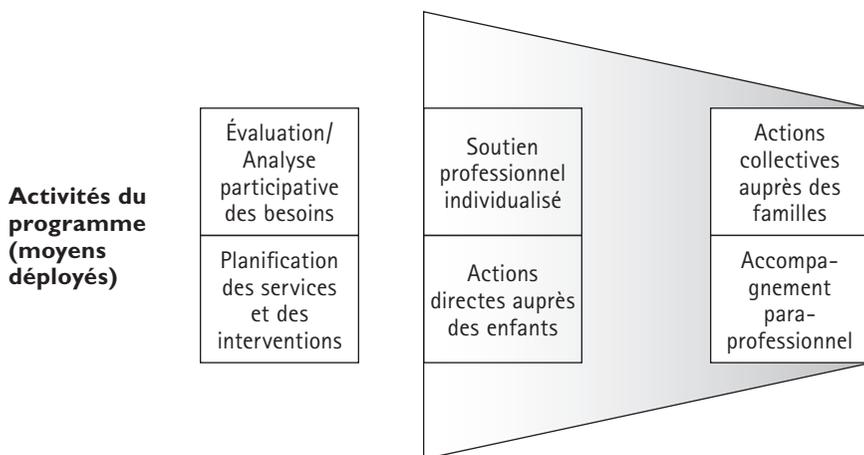
globale décrite ici. Le point de vue des intervenants sur l'expérience vécue par les usagers des services qu'ils reçoivent constitue toujours un aspect important dans la prise en compte de la dimension subjective lors de l'évaluation d'un programme.

organisme partenaire. Ce comité a le mandat de soutenir l'implantation du programme. Un intervenant de chaque établissement ou organisme partenaire (habituellement les mêmes personnes qui siègent au comité de coordination) constituent le *comité de programme* (ou équipe clinique) auquel se joint les intervenants œuvrant dans les autres volets du programme (analyse participative des besoins des enfants et des familles, suivi individualisé des familles, actions directes envers les enfants). Ce comité a pour mandat de faire le suivi de chaque famille participant au programme.

Figure 10.1

### Modèle logique du PAPFC





**Objectifs liés au processus d'intervention**

- Réaliser, en collaboration avec les parents et les autres adultes significatifs de l'entourage, une évaluation et une analyse des besoins des enfants et des parents en termes des trois fonctions du PAPFC.
- Élaborer des plans d'intervention et, le cas échéant, un plan de services qui contiennent des objectifs abordant chacune des trois fonctions du PAPFC.
- Fournir aux parents un soutien professionnel individualisé basé sur l'application de principes d'aide efficace avec des personnes vulnérables.
- Diriger les enfants et les parents vers des services généraux/particuliers, spécialisés ou surspécialisés correspondant aux besoins identifiés et offrir un accompagnement individualisé dans la provision de ces services.
- Mettre en place des rencontres de groupe de parents abordant des thématiques associées aux trois fonctions du PAPFC et fondées sur des principes d'animation efficace avec des personnes vulnérables.
- Mettre en place des activités collectives d'entraide et de socialisation pour les parents et les enfants.
- Mettre en place un service d'accompagnement paraprofessionnel des familles (par exemple, parents soutiens).
- Mettre en place une structure de gestion permettant de coordonner les services et les interventions et de les ajuster en fonction des réponses des enfants et des parents.



Figure 10.1 (suite)

## Modèle logique du PAPFC

### Objectifs liés aux résultats proximaux du programme (court terme)

- Les conditions d'existence des familles s'améliorent pour permettre aux enfants et aux parents de disposer de ressources nécessaires au maintien de leur santé et de leur bien-être et d'avoir du pouvoir sur leur vie.
- Les figures parentales disposent du soutien social nécessaire à l'exercice de leurs responsabilités auprès de leurs enfants. Ce soutien correspond à la nature des besoins manifestés par les parents (donc suffisamment intense, cohérent et continu).
- La disponibilité psychologique des figures parentales de même que les conduites responsables et sensibles à l'égard des besoins de leurs enfants s'améliorent.
- Le fonctionnement psychosocial et cognitif des enfants s'améliore à l'intérieur de leurs divers milieux de vie.



### Objectifs liés aux résultats distaux du programme (long terme)

- Les figures parentales démontrent, de manière stable, une compétence à répondre aux besoins de leurs enfants et à mobiliser les ressources (sur le plan personnel, social, économique), le soutien et l'aide nécessaires en fonction des besoins de ces derniers et de leurs propres besoins personnels.
- Le développement global et l'adaptation psychosociale des enfants s'améliorent.
- Le cycle de la négligence et de l'abus est brisé au moment où les enfants deviennent eux-mêmes des adultes et des parents.

Le comité de coordination local a été désigné comme étant le principal lieu pour la démarche d'évaluation décrite dans ce chapitre. La position stratégique des membres de ce comité à l'intérieur de chacun des établissements ou organismes partenaires de même que l'habitude de travailler ensemble ont été les principales raisons qui ont motivé ce choix. Sur le plan formel, c'est aussi ce comité qui avait la responsabilité de produire un bilan de l'implantation et un plan d'action.

Les membres du comité ont reçu préalablement un questionnaire leur permettant de faire le point sur l'implantation du programme sous l'angle de leur établissement ou organisme respectif (voir l'annexe pour une description sommaire des questions). Ils avaient la possibilité d'utiliser ce questionnaire de la façon qu'ils considéraient la mieux adaptée. Ils pouvaient consulter d'autres membres de leur institution s'il souhaitait obtenir des descriptions plus riches ou détaillées de certains aspects du programme.

Trois rencontres ont été réalisées pour mettre en commun les informations recueillies. L'animation a été effectuée par une personne agissant comme observateur participant aux activités de ce comité et ayant un statut de chercheur. Les sections du questionnaire (plutôt que les questions particulières) ont été utilisées comme thématiques dans le canevas d'animation des rencontres. Ces thématiques (partenariat interétablissement, ressources, structures de gestion, etc.) ont servi de déclencheurs à l'expression et au partage de points de vue et d'expériences vécus au cours des douze mois précédents. Il faut noter que, sur ce territoire, une première génération de PAPFC était implantée depuis une dizaine d'années. Les partenaires institutionnels souhaitent implanter la seconde génération du programme. L'objectif central dans l'implantation de la nouvelle génération était donc de soutenir la migration des pratiques.

Le bilan a notamment permis de mettre en relief les points suivants :

- Le partenariat entre établissements et organismes est fort et s'appuie sur une compréhension commune du cadre de référence et des principes du programme. Les effets réels sur les usagers (parents) de certaines pratiques institutionnelles (utilisation de formulaires) sont cependant questionnés par le partenaire communautaire.
- Considérant la nature multidimensionnelle du programme, trois volets avaient été priorisés dans l'implantation de la nouvelle génération du programme : les rencontres de groupe de parents, l'accompagnement paraprofessionnel des familles ainsi que les activités collectives. Pour ces volets, le constat est que les pratiques s'y référant ont pris, de façon significative, un virage dans le sens des principes du programme de nouvelle génération. Dans l'ensemble, ces activités ont une base suffisamment solide pour permettre de croire que, si les agents de programme qui les actualisent continuent d'être adéquatement soutenus, la qualité et la cohérence des activités continueront d'évoluer de façon positive.
- Le taux de participation des familles à ces activités est élevé. De même, l'expérience qu'en ont les parents s'avère particulièrement positive. Les parents semblent apprécier et considérer comme utiles ces activités. De plus, les agents de programmes jugent que, pour la plupart des usagers, ceux-ci évoluent dans la direction des objectifs identifiés au plan d'intervention.
- Les agents de programme qui ont été les plus directement engagés dans les structures de gestion du programme (comité de coordination et équipe clinique) ont un niveau d'appropriation des principes et des stratégies d'action du programme plus élevé et plus homogène que les agents de programme ayant peu participé à ces structures (en particulier, les intervenants responsables de l'analyse des besoins et du

suivi individualisé auprès des familles). Une grande hétérogénéité est observée dans l'appropriation des principes et des stratégies d'action parmi cette catégorie d'agents.

- Le changement de pratique auprès des familles ayant des comportements négligents est une préoccupation à la fois des gestionnaires des établissements, des membres du comité de coordination, des agents de programme impliqués directement dans les équipes cliniques ainsi que de la part de l'équipe de recherche. À ce sujet, le guide de programme comporte plusieurs éléments permettant d'orienter ce changement. Il s'agit ici de placer ces éléments dans le contexte particulier de chaque établissement.

Le plan d'action qui a découlé de ce type de réflexion et de ce bilan a mis l'accent sur l'élargissement de l'appropriation des concepts et des principes du programme par les intervenants responsables du suivi individualisé auprès des familles. De multiples défis ont pu être identifiés dans la mise en œuvre de cet objectif au cours de l'année suivante (augmentation de l'intégration du programme à l'intérieur des processus cliniques utilisés dans chaque établissement, identification des besoins de formation des professionnels, mise en place de structures de supervision clinique alignées sur les principes du programme, mise en place de procédures d'imputabilité face à l'adoption de pratiques ciblées par le programme, etc.).

## CONCLUSION

L'application d'une démarche de type ethnographique en évaluation de programme soulève divers enjeux. Le premier de ces enjeux a trait à l'appropriation d'une démarche d'évaluation de programme par les acteurs eux-mêmes de ce programme. Dans cette perspective, évaluer un programme ne doit plus être considéré uniquement comme une activité sporadique, voire exceptionnelle, ni comme une activité qui doit être prise en charge par une instance externe. Il s'agit plutôt d'une activité continue s'apparentant à une pratique réflexive telle que décrite par Argyris et Schön (1974) et Schön (1983). Les principales personnes concernées par le programme (ses agents, ses usagers) constituent des sources d'information inestimables sur l'incidence de celui-ci dans leur vie professionnelle et personnelle. Ces personnes ont la capacité de se centrer sur leur expérience et leur point de vue pour analyser cette incidence.

Un autre enjeu réside dans la légitimité accordée à cette forme de démarche. Il existe beaucoup de préjugés face à la validité d'une approche reposant sur l'exploitation du matériel subjectif. Il est donc essentiel de bien comprendre les indications et les limites non seulement de ce type

de démarche, mais également d'une démarche centrée sur la quantification et l'analyse quantitative d'informations à propos d'un programme. Une démarche ethnographique telle que celle que nous avons décrite ici n'est pas incompatible avec une approche plus objective. Au contraire, les constats objectifs, au sujet de la théorie d'un programme, de son implantation, de ces effets et de son efficacité (rapport coût-bénéfice), peuvent aisément être intégrés à la réflexion expérimentielle des agents de ce programme. C'est le sens que donnent ces agents à ces constats objectifs qui constitue le véhicule privilégié pour faire vivre et faire évoluer un programme.

## ANNEXE

### **Bilan des processus d'implantation du PAPFC (Fafard et Lacharité, 2007)**

---

#### **Plan organisationnel du programme**

##### **Ententes de services et partenariat**

1. Une entente de partenariat entre les établissements et les organismes impliqués a-t-elle été approuvée par les autorités institutionnelles appropriées (CA, cadres supérieurs) ?
2. Est-ce que chacun des gestionnaires des établissements et des organismes a été bien informé de la nature du programme et du rôle attendu, **avant d'accepter de s'impliquer** ?
3. Est-ce que les rôles de chacun des établissements et des organismes sont **actuellement** bien définis à l'intérieur du programme ?
4. Existe-t-il des ententes particulières en lien avec la répartition des coûts entre les établissements et les organismes (matériel, locaux, garderie, frais de déplacements, collation, etc.) ?
5. Est-ce que le climat de travail, en lien avec l'implantation du programme, est satisfaisant entre les établissements et les organismes partenaires, notamment au sein du comité de coordination, de l'équipe clinique, de l'équipe de paraprofessionnels et des rencontres de groupe ?

##### **Les ressources humaines et financières**

6. Des cadres et des intervenants de chacun des établissements et des organismes impliqués, et le cas échéant des paraprofessionnels, ont-ils été dûment identifiés et assignés à l'implantation du programme ?

7. L'expérience et l'expertise des agents de programme<sup>5</sup> ciblés correspondent-elles aux buts et objectifs visés par le programme ?
8. Est-ce que, pour chacun des agents de programme dégagés pour l'implantation, le nombre d'heures, leur description de tâches et l'implication demandée ont été clairement définis et précisés à ces derniers ?
9. La description de tâches des agents de programme correspond-elle aux buts et objectifs visés par le programme ?
10. Est-ce que chaque gestionnaire a les outils lui permettant de faire une évaluation (au moins annuellement) du fonctionnement de chaque agent de programme sous son autorité, en fonction des quatre processus d'implantation (processus d'objectivation, d'appropriation, de réflexion qualifiante et rétroaction ?
11. Est-ce que l'évaluation du fonctionnement de chaque agent de programme permet de démontrer que la majorité de ceux-ci ont une maîtrise fonctionnelle des quatre processus d'implantation du programme ?
12. L'estimation des coûts (en temps) pour le personnel salarié et les bénévoles a-t-il été fait pour chacun des établissements et des organismes ?
13. Les budgets nécessaires en lien avec chacune des activités du programme, sont-ils dégagés et reconduits annuellement par chacune des organisations ?

### Structure de gestion de programme

14. Une structure de gestion (comité de coordination et équipe clinique) qui correspond aux buts et objectifs visés par le programme a-t-elle été mise en place ?
15. Les mandats, les rôles et les responsabilités des différents comités sont-ils clairement définis et connus de tous les agents de programme ?
16. Est-ce que le comité de coordination et l'équipe clinique assument leurs mandats et responsabilités de façon cohérente avec l'esprit du programme ?
17. Est-ce que le nombre de rencontres du comité de coordination et de l'équipe clinique est suffisant afin de répondre aux besoins des agents de programme ?

---

5. Un agent de programme réfère à toute personne qui est liée de près au programme, c'est-à-dire les intervenants, les gestionnaires, les paraprofessionnels, les animateurs de groupe ainsi que les membres des différents comités.

18. Est-ce que des activités de formation ont été planifiées et réalisées pour chacune des activités du programme<sup>6</sup> ?
19. Est-ce que le contenu des formations est cohérent avec l'esprit du programme ?
20. Est-ce que le nombre de formation est suffisant afin de répondre aux besoins des agents de programme ?
21. Est-ce que des activités de supervision ont été planifiées et réalisées pour chacune des activités du programme ?
22. Est-ce que le contenu des supervisions est cohérent avec l'esprit du programme ?
23. Est-ce que le nombre de supervisions est suffisant afin de répondre aux besoins des agents de programme ?
24. Est-ce qu'il existe une procédure à l'intérieur même du programme permettant d'offrir du soutien aux agents de programme ?
25. Si une procédure existe, les agents de programme se sentent-ils soutenus en regard des tâches et des responsabilités qui leur incombent ?

### Plan d'action

26. Est-ce qu'un plan d'action<sup>7</sup> annuel, qui correspond aux buts et objectifs visés par le programme, a été produit en partenariat par les établissements et les organismes impliqués ?
27. Est-ce que chacun des établissements et des organismes impliqués a un plan d'action particulier en lien avec le programme ?
28. Est-ce qu'une évaluation annuelle **du fonctionnement** du programme est prévue pour les établissements et les organismes par l'entremise du comité de coordination ?
29. Est-ce qu'une évaluation annuelle **des effets** du programme est prévue pour les établissements et les organismes par l'entremise du comité de coordination ?

---

6. Les activités du programme sont : l'évaluation et l'analyse des besoins des enfants et des familles, le suivi professionnel individualisé, l'animation de groupe de parents, l'accompagnement paraprofessionnel des parents et les actions directes auprès des enfants.

7. Un plan d'action est défini comme étant un ensemble d'objectifs prioritaires, accompagné d'une description des moyens utilisés pour les atteindre, des responsables ciblés et de l'échéancier à respecter.

### Outils de gestion du programme<sup>8</sup>

30. Existe-t-il des procédures pour recueillir et archiver les informations recueillies à travers les outils de gestion, et ce, pour chacun des établissements et des organismes ?
31. Existe-il des comptes rendus rédigés après les réunions des comités de coordination et de l'équipe clinique ?
32. Les comptes rendus permettent-ils d'obtenir et d'échanger toutes les informations permettant au comité de coordination d'avoir le pouls du fonctionnement du programme pour chacune des activités ?
33. Les comptes rendus permettent-ils de partager les informations qui sont essentielles aux agents de programme ?
34. Est-ce qu'un protocole de référence/recrutement des familles a été élaboré à partir des critères retenus ?

### Outils servant à recueillir les informations sur la situation des familles participantes.

35. Est-ce qu'au moins 90 % des familles acceptées au programme ont fait l'objet d'un plan de services ?
36. Est-ce que les plans d'intervention ont été révisés aux trois mois ?
37. Est-ce que les plans de services ont été révisés annuellement ?
38. Est-ce que dans au moins 90 % des cas référés, une évaluation et une analyse des besoins développementaux des enfants ont été réalisées par les intervenants responsables et à laquelle **au moins un des parents** a participé ?
39. Les résultats de l'évaluation des besoins de familles sont-ils partagés de façon régulière entre les agents de programme concernés ?
40. Est-ce qu'au moins 90 % des plans d'intervention et des plans de services comportent des informations sur les objectifs à atteindre en termes des besoins développementaux des enfants, des capacités des adultes de l'entourage à répondre à ces besoins et des conditions familiales et sociales ?

---

8. Ces outils de gestion servent à gérer les informations sur le programme.

## Plan d'utilisation des services par les usagers

### Procédures permettant de rejoindre les parents potentiels dans chaque établissement

41. Est-ce que des activités d'information et de sensibilisation sont régulièrement réalisées auprès des intervenants et des gestionnaires dans chaque établissement référant ?
42. Est-ce que des documents ont été élaborés pour décrire et expliquer le programme aux intervenants et aux gestionnaires concernés de chacun des établissements et des organismes partenaires ?
43. Est-ce que les cadres et les intervenants directement concernés par la référence et le recrutement des familles ont une connaissance fonctionnelle des buts, des objectifs et des activités relatives au programme ?
44. Est-ce que des documents ont été élaborés pour décrire et expliquer le programme aux paraprofessionnels ?
45. Est-ce que des documents, ou d'autres stratégies, ont été élaborés pour décrire et expliquer le programme aux parents concernés (leur expliquer la nature du programme et ce qu'il implique comme engagement) ?
46. Des moyens particuliers ont-ils été mis en place afin de recruter et maintenir les pères au sein du programme ?

### Description des critères d'inclusion et d'exclusion des familles

47. Les critères d'inclusion et d'exclusion des familles sont-ils connus et formalisés par les agents de programme ?
48. Y a-t-il un registre des références incluant des informations telles : les critères d'inclusion et d'exclusion pour chaque famille référée, le nombre de familles ayant accepté et ayant refusé de participer ainsi que les raisons du refus ?
49. Est-ce qu'un profil des parents ayant abandonné le programme a été réalisé ?
50. Est-ce qu'au moins 25 % des familles répondant aux critères de sélection ont été référées au programme par chacun des établissements référents ?
51. Est-ce qu'au moins 90 % des familles acceptées au programme répondent aux critères de sélection ?
52. Est-ce que le taux d'acceptation à participer au programme par les familles est d'au moins 80 % pour chacun des établissements référents ?

53. Est-ce que le taux de rétention des familles à l'intérieur du programme est d'au moins 80 % après un an pour chacun des établissements référents ?

#### **Outils et procédures formelles utilisés avec les enfants ou les parents**

54. Y a-t-il une recension décrivant les outils, les questionnaires, les formulaires ainsi que les procédures formelles utilisées avec les enfants ou les parents ?
55. Existe-il des documents ou des questionnaires remplis par les parents dans le cours de leur participation au programme ?
56. Y a-t-il des procédures pour recueillir et archiver les documents et les questionnaires complétés par les parents ?
57. Est-ce que les familles participantes ont formellement des occasions de rétroagir à propos des divers services qu'ils reçoivent au sein du programme ?
58. Y a-t-il une recension des outils visant le processus de rétroaction aux parents ?
59. Les outils de rétroaction sont-ils utilisés au sein de chacune des activités du programme ?
60. Les résultats des activités de rétroaction sont-ils compilés et utilisés par les agents de programme ?

#### **Création d'un dossier clientèle commun**

61. Est-ce qu'il y a un dossier clientèle commun pour l'ensemble des établissements impliqués dans le programme ?
62. Est-ce que l'ensemble des informations concernant la famille participante est classé à l'intérieur d'un dossier central (description des services fournis, décisions institutionnelles à propos des familles, contenu des plans d'interventions et des plans de services) ?

### **Questions de synthèse**

63. Est-ce que le suivi professionnel de la famille, notamment en ce qui concerne l'accompagnement de la famille dans sa collectivité (gestion de cas) et l'aide psychosociale apportée aux parents par l'intervenant responsable répond, en termes d'intensité et de contenu, aux principes proposés par le programme ?

64. Est-ce que les actions directes auprès des enfants répondent, en termes d'intensité et de contenu, aux principes proposés par le programme ?
65. Est-ce que l'animation de groupes de parents répond, en termes d'intensité et de contenu, aux principes proposés par le programme ?
66. Est-ce que la réalisation d'activités collectives répond, en termes d'intensité et de contenu, aux principes proposés par le programme ?
67. Est-ce que l'accompagnement paraprofessionnel répond, en termes d'intensité et de contenu, aux principes proposés par le programme ?

## BIBLIOGRAPHIE

- ARGYRIS, C. et D.A. SCHÖN, (1974). *Theory in Practice: Increasing Professional Effectiveness*, San Francisco, Jossey-Bass.
- BILLIG, M. (1987). *Arguing and Thinking: A Rhetorical Approach to Social Psychology*, Cambridge, Cambridge University Press.
- BILLIG, M. (1991). *Ideology, Rhetoric, and Opinions*, Londres, Sage.
- BOURDIEU, P. (2002). *Réponses*, Paris, Seuil.
- CAMPBELL, M. et F. MCGREGOR (2002). *Mapping Social Relations: A Primer in Doing Institutional Ethnography*, Toronto, Garamond.
- CONQUERGOOD, D. (1991). «Rethinking ethnography: Towards a critical cultural politics», *Communication Monographs*, vol. 58, p. 179-194.
- DE FORNEL, M., A. OGIEN et L. QUÉRÉ (2001). *L'ethnométhodologie: une sociologie radicale*, Paris, La Découverte.
- FAFARD, G. et C. LACHARITÉ (2007). *Évaluation d'implantation du PAPFC-2: aspects organisationnels*, Document inédit, Groupe de recherche et d'intervention en négligence, Trois-Rivières, Université du Québec à Trois-Rivières.
- FOUCAULT, M. (1972). *La naissance de la clinique: une archéologie du regard médical*, Paris, Presses universitaires de France.
- GARFINKEL, H. (2007). *Recherches en ethnométhodologie*, Paris, Presses universitaires de France.
- GEERTZ, C. (1973). *The Interpretation of Cultures: Selected Essays*, New York, Basic Books.
- GEERTZ, C. (1983). *Local Knowledge: Further Essays in Interpretative Anthropology*, New York, Basic Books.
- GERGEN, K.J. (1982). *Towards Transformation in Social Knowledge*, New York, Springer.
- GERGEN, K.J. (1985). «The social constructionist movement in modern psychology», *American Psychologist*, vol. 40, p. 266-275.
- KAUFMAN, J.C. (2007). *L'entretien compréhensif*, 2<sup>e</sup> éd., Paris, Armand Colin.

- LACHARITÉ, C., L. BOURASSA et E. LE MAROIS (2004). *Évaluation des références au PAPFC : territoire de Victoriaville*, Rapport déposé à l'Agence régionale de santé et de services sociaux de la Mauricie et du Centre-du-Québec.
- LACHARITÉ, C., G. FAFARD et L. BOURASSA (2005). *Programme d'aide personnelle, familiale et communautaire : nouvelle génération (PAPFC-2)*, Trois-Rivières, Groupe de recherche et d'intervention en négligence, Université du Québec à Trois-Rivières.
- LACHARITÉ, C., L.S. ÉTHIER et P. NOLIN (2006). « Vers une théorie écosystémique de la négligence envers les enfants », *Bulletin de psychologie*, vol. 59, n° 4, p. 381-394.
- LACHARITÉ, C. (sous presse). « Approche participative auprès des familles », dans C. Lacharité et J.P. Gagnier (dir.), *À la rencontre des familles : réalités plurielles et regards interdisciplinaires*, Montréal, Chenelière.
- LACHARITÉ, C. et al. (2007). « Faire la courte échelle : développement d'un programme de services intégrés pour contrer la négligence », dans C. Chamberland, N. Trocmé et S. Léveillé, *Protéger les enfants, aider les adultes : deux univers à rapprocher*, Québec, Presses de l'Université du Québec.
- LACHARITÉ, C. et G. FAFARD (sous presse). « Research-practice partnerships in the development of programs for child neglect » dans N. Trocmé, S. Léveillé et C. Chamberland (dir.), *Research-Community Partnerships in Child Welfare*, Montréal, McGill University Press.
- MAUSS, M. (2002). « Métier d'ethnographe, méthode sociologique », <classiques.uqac.ca/classiques/mauss\_marcel/œuvres\_3/œuvres\_3\_09/metier\_ethnographe>.
- MILLER, W.L. et B.F. CRABTREE (2008). « Clinical research », dans N.K. Denzin et Y.S. Lincoln (dir.), *Strategies of Qualitative Inquiry*, Thousand Oaks, Sage.
- ROSSI, P.H., M.W. LIPSEY et H.E. FREEMAN (2004), *Evaluation: A Systematic Approach*, 7<sup>e</sup> éd., Thousand Oaks, Sage.
- SCHÖN, D.A. (1983). *The Reflexive Practitioner: How Professionals Think in Action*, New York, Basic Books.
- SHOTTER, J. (1993). *Conversational Realities: Constructing Life Through Language*, Thousand Oaks, Sage.
- SMITH, D.S. (1999). *Writing the Social: Critique, Theory, and Investigations*, Toronto, University of Toronto Press.
- SMITH, D.S. (2005). *Institutional Ethnography: A Sociology for People*, Toronto, Altamira Press.
- TURNER, V.W. et E.M. BRUNER (1986). *The Anthropology of Experience*, Chicago, University of Illinois Press.
- WALDEGRAVE, C. et al. (2003). *Just Therapy: A Journey*, Adelaïde, Dulwich Centre Publications.

# CHAPITRE 11

---

## **Les ressources disponibles en évaluation de programme**

Marc Alain

Département de psychoéducation  
Université du Québec à Trois-Rivières

Pierre Potvin

Département de psychoéducation  
Université du Québec à Trois-Rivières

La question de l'état des ressources disponibles en matière d'évaluation de programme pourrait occuper à elle seule l'ensemble de pages dévolues à l'ensemble des sujets abordés jusqu'ici. Il va donc de soi que nous ne ferons qu'effleurer le domaine, sans compter que, d'une manière ou d'une autre, les ressources, notamment celles qui sont accessibles par Internet, changent rapidement au point de rendre l'exercice tout aussi rapidement caduc. Nous nous en tiendrons donc aux grandes lignes, des grandes lignes susceptibles de servir de fils conducteur pour tout lecteur désirant en savoir plus, qu'il s'agisse des évaluateurs déjà au fait des diverses modalités et formes de l'évaluation ou des évaluateurs en devenir. Une classification sommaire de ces ressources fait ressortir :

- des organismes de référence en évaluation accessibles par Internet ;
- des revues et périodiques spécialisés ;
- des sites exclusivement consacrés à des questions méthodologiques ;
- des ouvrages de référence.

Dans ce dernier cas, nous ne traiterons que les ouvrages que la notoriété acquise et le temps ont rendus à peu près incontournables ; pour ce qui est de ceux qui abordent des thématiques plus pointues – sur le plan des méthodologies, par exemple –, notons qu'ils émaillent les bibliographies de chacun des chapitres précédents et fournissent ainsi autant de points de repère utilisables par les lecteurs. Ainsi, par exemple, on retrouvera, pour qui s'intéresse plus particulièrement aux outils et techniques d'évaluation de besoins, des références pertinentes dans les bibliographies des chapitres 3 et 4 ; pour qui s'intéresse plutôt aux outils et techniques quantitatives d'évaluation des effets, c'est au chapitre 8 qu'il trouvera son compte, tandis que le lecteur qui privilégie une approche plus qualitative de la mesure des effets ira chercher ces mêmes outils dans les références du chapitre 9, et ainsi de suite.

## 1. LES ORGANISMES NATIONAUX ET DE RÉFÉRENCE EN ÉVALUATION ACCESSIBLES PAR INTERNET

Compte tenu que notre premier lectorat se situe au Québec, nous amorcerons logiquement notre revue des organismes de référence par la Société québécoise d'évaluation de programme. La SQEQ définit ainsi sa mission (<[www.sqep.ca/pages/p\\_mission.htm](http://www.sqep.ca/pages/p_mission.htm)>) :

La **Société québécoise d'évaluation de programme** est incorporée depuis 1989 et a pour objectifs de :

- permettre aux personnes intéressées par l'évaluation de programme d'échanger entre elles ;

- contribuer au développement de l'évaluation de programme, notamment en encourageant la recherche et le développement en évaluation de programme;
- constituer un centre de référence, d'action, d'information et de formation pour les intervenants en évaluation de programme au Québec;
- prendre position sur des questions touchant l'évaluation de programme.

La Société québécoise d'évaluation de programme (SQÉP) est incorporée en vertu de la partie II de la Loi sur les corporations canadiennes. Elle est affiliée à la Société canadienne d'évaluation (SCÉ) et représente la Section du Québec au sein de cette dernière. Ses règlements sont rédigés en conformité avec cette loi et en respect de l'entente qui la lie à la SCÉ.

Le site Internet de l'organisme (<[www.sqep.ca/](http://www.sqep.ca/)>) propose divers bulletins et nouvelles des activités tenues par la Société elle-même ou par l'un ou l'autre de ses nombreux partenaires, qui se situent en Europe (France, Belgique, Suisse) ou aux États-Unis (l'American Evaluation Association, par exemple). Les activités que la SQEP propose sont nombreuses et variées; elles vont de la présentation de congrès et de colloques plus spécialisés à une offre très détaillée de formations et d'ateliers donnés par des spécialistes reconnus de l'évaluation. L'un des attraits du site réside dans sa collection d'archives de présentations (généralement en format Powerpoint) faites au fil des ans dans divers congrès et colloques; ces archives, fort bien référencées et parmi lesquelles il est possible de réaliser des recherches par mot clé, sont accessibles à tous. Ce fait marque d'ailleurs une volonté d'ouverture de l'organisme que l'on ne retrouve pas nécessairement partout ailleurs où seuls les membres en règle peuvent accéder aux documents en ligne.

La **Société canadienne d'évaluation** (<[www.evaluationcanada.ca/site.cgi?s=1](http://www.evaluationcanada.ca/site.cgi?s=1)>) se présente comme l'organisme pancanadien qui chapeaute, en quelque sorte, les différentes organisations provinciales (toutes les provinces canadiennes ont leur propre société d'évaluation, et c'est également le cas pour les Territoires du Nord-Ouest)<sup>1</sup>. On y offre la possibilité de s'inscrire et de participer à une multitude de formations diverses, de l'analyse des résultats aux considérations éthiques en matière d'évaluation, en passant par des formations très pointues sur la communication des résultats et l'utilisation de méthodes statistiques avancées. La question des normes d'évaluation de programmes est abordée de manière très fouillée en quatre sections touchant 1) les normes d'**utilité** (la notion de répondre aux besoins d'information des

---

1. On retrouve une telle structure au sein de certaines des sociétés internationales d'évaluation; c'est notamment le cas de l'African Evaluation Association (<[www.afrea.org/home/index.cfm](http://www.afrea.org/home/index.cfm)>) qui accueille la plupart des sociétés nationales africaines d'évaluation.

utilisateurs de l'évaluation), 2) les normes de **faisabilité** (faire de l'évaluation dans un rapport coûts/bénéfices aussi efficace et efficient que possible), 3) les normes de **rectitude** (dans le respect des lois, de l'éthique et du souci optimal du bien-être des personnes susceptibles d'être affectées par les résultats de l'évaluation), et 4), les normes d'**exactitude** (comprenant ici l'ensemble des soucis propres à l'exécution impartiale des plus hauts standards de recherche appliquée)<sup>2</sup>. On retrouve également sur le site une liste de liens qui nous renvoient aux divers organismes et ministères fédéraux qui ont réalisé des démarches d'évaluation ou qui présentent différents outils et rapports. Mais si cette dernière section demeure relativement pauvre en termes de contenu, c'est indéniablement la liste des sites Internet tenue à jour sur le site de la Société canadienne qui constitue probablement l'un de ses plus grands attraits. On y liste pas moins de cent sites liés directement ou indirectement à l'évaluation de programmes, dont entre autre, un lien qui mènera les intéressés à d'autres pages de liens internet. C'est également la Société canadienne d'évaluation qui publie la *Revue canadienne d'évaluation de programme*, seule publication dans les deux langues dans le domaine au Canada et sur laquelle nous reviendrons un peu plus bas. Notons, pour terminer, que la plupart des documents les plus intéressants ne sont disponibles que pour les membres en règle de la Société.

Toujours sur le plan des organisations nationales, l'**American Evaluation Association** (AEA) présente elle aussi une impressionnante liste de ressources et de liens. On peut même trouver sur le site (<[www.eval.org/](http://www.eval.org/)>) une liste d'évaluateurs reconnus avec lesquels il est possible de faire affaire, aux États-Unis et ailleurs. Le site de l'association américaine demeure sans contredit l'un des plus complets qui existe et, à ce titre, sa consultation est une quasi-obligation pour qui veut tenter de demeurer à jour dans ce domaine qui, à l'instar des autres champs de connaissance en science sociale, évolue de façon pratiquement exponentielle. Au chapitre des ressources, le site de l'Association présente 16 possibilités d'entrées constamment mises à jour, allant d'une liste de personnes ressources à une liste de textes accessibles gratuitement, en passant par une liste des programmes universitaires américains qui offrent des cours et des cursus complets en évaluation (on y retrouve d'ailleurs la mention de l'un des programmes de second cycle entièrement dédié à l'évaluation existant au Canada, soit le certificat de

---

2. Notons qu'à peu près toutes les sociétés nationales d'évaluation proposent une forme ou une autre de cadre standard qui sont, à peu de chose près, des variantes de ce que l'Association canadienne propose. À titre d'exemple, la Société allemande d'évaluation propose des standards en 4 catégories très similaires à celles que propose la Société canadienne, soit les normes: 1) d'utilité, 2) de la faisabilité, 3) de convenance, et 4) d'exactitude (<[www.degeval.de/calimero/tools/proxy.php?id=72](http://www.degeval.de/calimero/tools/proxy.php?id=72)>).

second cycle en évaluation de programme de l'Université d'Ottawa<sup>3</sup>). De ces 16 entrées possibles, on retiendra celle qui a trait aux instruments et outils recommandés par les membres de l'Association, de même que celle qui regroupe une liste des liens d'accès à d'autres sites similaires à celui de l'Association; dans ces deux cas, chacune des entrées est commentée en fonction du public cible, de la portée de ce qui est proposé, des coûts lorsque c'est le cas, etc.

Le **Western Michigan University Evaluation Center** constitue, pratiquement au même titre que l'AEA, un incontournable (<[www.wmich.edu/evalctr/](http://www.wmich.edu/evalctr/)>). Le site est d'ailleurs bien connu des évaluateurs par son impressionnante section des listes de contrôle – les *checklist* – où l'on retrouve la plupart des listes développées par Michael Scriven et Daniel Stufflebeam (on y retrouve, signe potentiel de l'exhaustivité de ces auteurs en matière de liste de contrôle, une liste de contrôle pour établir des listes de contrôle – *checklists development checklist*). Un périodique en ligne, le *Journal of Multi-Disciplinary Evaluation*, y est également hébergé et accessible à tous. On y retrouve aussi une bibliothèque virtuelle (<[ec.wmich.edu/library/index.htm](http://ec.wmich.edu/library/index.htm)>) où l'on peut à loisir effectuer des recherches dans les quelques 4 711 titres que comprend la collection.

Le site du **Penn State University Evaluation Program Center** (<[www.extension.psu.edu/evaluation/](http://www.extension.psu.edu/evaluation/)>), bien que sensiblement moins fourni que le précédent, présente néanmoins une liste fort utile pour les praticiens de l'évaluation de feuilles de conseils («*tipsheets*»). Il s'agit de courts documents qui synthétisent très adroitement différentes étapes importantes d'une démarche d'évaluation, de l'analyse préliminaire des besoins à l'évaluation des résultats, en passant par la communication de ces mêmes résultats aux parties prenantes. On retrouve également, toujours dans cette liste de feuilles de conseil, plusieurs éléments de méthodologie, et des conseils très pertinents quand vient le temps d'élaborer des outils de cueillette de données adaptées aux particularités des sites et programmes évalués.

En français, et c'est ce qui en fait une ressource d'autant plus intéressante, on retrouvera le **Portail francophone de l'évaluation** (<[evaluation.francophonie.org/](http://evaluation.francophonie.org/)>), une porte d'entrée pour qui veut consulter, en ligne, des manuels de référence, des outils méthodologiques, des rapports d'évaluation, voire des appels d'offre internationaux francophones pour des contrats d'évaluation. Le portail se veut une plate-forme de soutien aux pays en développement qui mise sur le réseautage et le partage d'expertises et d'expériences; à

---

3. Plusieurs autres programmes existent cependant, notamment à l'ENAP (un micro-programme de deuxième cycle et deux cheminements de concentration spécialisés en évaluation de programmes publics dans le programme de maîtrise en administration publique).

ce titre, on retrouve une liste de 113 organisations qui, dans le monde, sont directement ou indirectement engagées dans des activités d'évaluation. Pour qui recherche des sites francophones, c'est l'une des portes d'entrée les plus intéressantes. Une impressionnante liste de documents est aussi présentée. Il est cependant à noter que la très grande majorité de ces documents émane de la Banque mondiale et sont, pour la plupart, en anglais. La liste des revues spécialisées que l'on y présente est assez exhaustive et on y retrouvera les revues tant anglophones que francophones.

## 2. LES REVUES ET PÉRIODIQUES SPÉCIALISÉS

Ici, contrairement à ce qui est le cas pour les sites Internet, dont la seule énumération constituerait un exercice tout à fait fastidieux, la liste est plus courte. Si l'on exclut les périodiques qui émanent des sociétés nationales d'évaluation, comme c'est le cas de la *Revue canadienne d'évaluation de programmes* et son pendant américain, le *American Journal of Evaluation*, deux périodiques scientifiques se démarquent du lot par leurs importants apports sur le plan théorique et sur celui du développement de la discipline. Il s'agit, d'une part, de la revue *New Directions for Evaluation* et, d'autre part, de la revue *Evaluation and Program Planning*. Commençons notre survol par les revues des sociétés nationales canadiennes et américaines.

### 2.1. La *Revue canadienne d'évaluation de programme* (<[evaluationcanada.ca](http://evaluationcanada.ca)>)

Les objectifs de la revue sont de favoriser à la fois la réflexion théorique et la pratique de l'évaluation, et ce, en français ou en anglais, comme c'est le cas de plusieurs périodiques qui sont sous la responsabilité de sociétés savantes canadiennes. On y retrouve donc des articles qui abordent l'une ou l'autre de ces deux considérations, voire les deux à la fois dans certains cas. La revue publie également, en plus des classiques revues des ouvrages qui sont publiés dans le domaine, une section dédiée à des notes de recherches appliquées. L'objectif plus particulier de cette partie consiste à fournir aux lecteurs des cas types qui illustrent, de manière concise – on privilégie des articles de 1 500 mots et moins – des pratiques innovantes en évaluation. La revue en est maintenant à sa vingt et unième année et est publiée deux fois l'an, soit au printemps et à l'automne. On peut accéder à l'ensemble des titres et des résumés d'articles qui ont été publiés à partir du site Internet de la Société canadienne d'évaluation, mais il faut être abonné pour accéder aux textes complets (ou alors, bien sûr, passer par une bibliothèque qui l'est).

## 2.2. *L'American Journal of Evaluation* ([www.eval.org/Publications/AJE.asp](http://www.eval.org/Publications/AJE.asp))

Plus jeune que sa consœur canadienne (le premier numéro date de 1996), son titre original, lors des deux premières années de publication, soit *Evaluation Practice*, donnait le ton de son contenu. En effet, l'AJE propose à ses lecteurs d'horizons très divers (de la psychologie à la sociologie et au travail social, en passant par le management) des articles essentiellement centrés sur les techniques, méthodes et modalités d'évaluation. On se propose en fait d'assister et d'aider les évaluateurs à améliorer leurs approches, leurs compétences, à encourager le dialogue et à améliorer leurs bases de connaissances. La revue est publiée quatre fois l'an et est accessible uniquement par voie d'abonnement.

## 2.3. *New Direction for Evaluation* ([www.eval.org/Publications/NDE.asp](http://www.eval.org/Publications/NDE.asp))

NDE, tout comme son nom l'indique, est essentiellement consacrée aux questions de développement de la discipline. Les numéros sont tous à thèmes; ainsi, par exemple, le tout premier numéro de la revue dans sa forme actuelle (celui du printemps 1994) était entièrement consacré au débat entourant l'intégration des approches qualitatives et quantitatives. La plupart des chercheurs qui ont, d'une manière ou d'une autre, contribué au développement de la discipline de l'évaluation, tant sur le plan théorique que sur les plans de la méthode et des pratiques, ont écrit dans NDE. Le tableau suivant présente chacun des thèmes abordés dans les numéros de NDE depuis ses débuts; le lecteur pourra ainsi référer à l'un ou l'autre des numéros de la revue qui traite du thème qui l'intéresse.

Tableau 11.1

### **Les thèmes de tous les numéros de la revue *New Directions for Evaluation* depuis le début de la parution jusqu'au numéro de l'automne 2008**

| Année | Numéro | Thème abordé   |
|-------|--------|--|
| 1994  | 61     | Relation entre la recherche qualitative et la recherche quantitative |
|       | 62     | Le rôle de l'évaluation en recherche et la formation des évaluateurs |
|       | 63     | Validité de l'échantillonnage aléatoire                              |
|       | 64     | Les mauvais usages de l'évaluation                                   |
| 1995  | 65     | Le rôle de l'évaluateur dans la réforme des sciences de l'éducation  |
|       | 66     | Les lignes directrices pour les évaluateurs                          |
|       | 67     | Nouvelles dimensions en évaluation                                   |
|       | 68     | La logique et le raisonnement en évaluation                          |

Tableau 11.1 (suite)

**Les thèmes de tous les numéros de la revue  
New Directions for Evaluation depuis le début  
de la parution jusqu'au numéro de l'automne 2008**

| Année | Numéro  | Thème abordé  |
|-------|---|---|
| 1996  | 69  | Évaluation des initiatives des services intégrés  |
|       | 70  | L'implication de la population dans l'évaluation  |
|       | 71  | Vérification et évaluation  |
| 1997  | 72  | Utilisation de patrons comme outils pour l'évaluation formative                                       |
|       | 73  | Utilisation de graphiques   |
|       | 74  | Méthodes mixtes   |
|       | 75  | Mesure de la performance  |
| 1998  | 76  | Le progrès de l'utilisation de l'évaluation en recherche  |
|       | 77  | Le développement de l'évaluation  |
|       | 78  | L'évaluation réaliste   |
|       | 79  | Évaluation et revenus   |
| 1999  | 80  | L'évaluation participative  |
|       | 81  | Utilisation législative de l'évaluation   |
|       | 82  | L'éthique de l'évaluation   |
|       | 83  | L'importance de l'évaluation pour la communauté   |
| 2000  | 84  | L'évaluation et la technologie  |
|       | 85  | L'évaluation délibérative et démocratique   |
|       | 86  | Comment et pourquoi le langage est un facteur qui compte en évaluation                                |
|       | 87  | L'évaluation et la théorie  |
| 2001  | 88  | L'utilisation de l'évaluation   |
|       | 89  | Modèles d'évaluation  |
|       | 90  | L'évaluation des services publics   |
|       | 91  | Programmes d'aide sociale, évaluation des effets du programme   |
| 2002  | 92  | L'évaluation réceptive ( <i>Responsive Evaluation</i> )   |
|       | 93  | L'application de l'évaluation du processus de construction des capacités dans plusieurs organisations |
|       | 94  | L'évaluation sur plusieurs sites ( <i>Multisite evaluation</i> )                                      |
|       | 95  | Évaluation des relations interorganisationnelles  |
| 2003  | 96  | L'évaluation féministe  |
|       | 97  | La relation entre la pratique et la théorie en évaluation   |
|       | 98  | Le rôle des jeunes dans l'avancement de l'évaluation participative et de la recherche                 |
|       | 99  | L'intégration de l'évaluation dans l'ensemble des activités des organisations                         |
| 100   | Utilisation de l'enquête d'appréciation en évaluation |   |

| Année | Numéro | Thème abordé  |
|-------|--------|---|
| 2004  | 101    | La coconstruction des modalités d'évaluation réceptive dans l'évaluation des résultats scolaires  |
|       | 102    | Intégration de notions de culture dans la théorie, la politique et la méthodologie de l'évaluation de l'évaluation  |
|       | 103    | Contrôle et évaluation du VIH/SIDA  |
|       | 104    | La communauté internationale d'évaluation et le développement du Programme des standards d'évaluation   |
| 2005  | 105    | L'utilisation de la méthode de cas en évaluation de l'enseignement  |
|       | 106    | Théorie de l'évaluation selon l'approche de théoriciens et instructions générales pour une bonne analyse comparative d'approches d'évaluation             |
|       | 107    | Application d'analyse de réseau social ( <i>Social Network Analysis</i> ) et analyse de données à la pratique de l'évaluation                             |
| 2006  | 108    | Évaluation de l'éducation informelle  |
|       | 109    | Évaluation du programme de sciences, technologies, génie et mathématiques (STEM) en éducation publique  |
|       | 110    | Analyse de méthodes d'évaluation au niveau de programmes de prévention auprès de la population adolescente  |
|       | 111    | La consultation indépendante en évaluation  |
|       | 112    | Le rôle de l'évaluation en politique  |
| 2007  | 113    | Politiques et controverses courantes au gouvernement fédéral américain sur la méthodologie en évaluation  |
|       | 114    | Vingt ans de collaboration entre NDE et l'American Evaluation Association   |
|       | 115    | Utilisation d'études faites en ligne et le choix de l'échantillon   |
| 2008  | 116    | Utilisation de processus  |
|       | 117    | Les conséquences de la politique du <i>No Child Left Behind</i> sur l'évaluation de la performance scolaire   |
|       | 118    | La logique et la stratégie de l'évaluation en recherche   |
|       | 119    | Le rôle des organisations à but non lucratif dans les services publics, l'évaluation organisationnelle et l'importance de développer la culture du savoir |

## 2.4. Evaluation and Program Planning

EPP constitue probablement l'une des plus anciennes revues spécialisées dans le domaine; son premier numéro remonte en effet à 1978. Comme son nom l'indique, on y retrouve tout autant des articles relatifs aux thèmes généraux de l'évaluation de programme mais aussi, des contributions plus particulièrement liées à la planification des programmes. Bien que la plupart des numéros de la revue proposent, dans une forme assez classique, des articles théoriques, des réflexions méthodologiques ou théoriques, des études de cas et des revues de livres, à partir de 2003, certains numéros – généralement une fois l'an, mais parfois plus aussi – présentent deux sections distinctes :

une première section est consacrée à des articles d'intérêt général, tandis qu'une seconde est particulièrement consacrée à un thème donné. Le tableau suivant présente ces différents thèmes et le numéro de la revue qui y est consacré.

Tableau 11.2

### **Les numéros à thème de la revue *Evaluation and Program Planning*, de 2003 à 2008**

| <b>Année</b> | <b>Numéro de la revue</b> | <b>Thème abordé</b>  |
|--------------|---------------------------|--|
| 2003         | Vol. 26, n° 4             | L'étude de cas en évaluation de programme  |
|              | Vol. 26, n° 3             | Évaluation longitudinale des programmes de désintoxication et de traitement des dépendances  |
|              | Vol. 26, n° 2             | Évaluation des technologies de l'éducation   |
| 2004         | Vol. 27, n° 3             | Considérations éthiques de l'évaluation dans les entreprises à profits   |
|              | Vol. 27, n° 2             |  |
|              | Vol. 27, n° 1             | Les programmes de traitement à domicile des mères et futures mères éprouvant des problèmes de dépendance<br>L'étude de cas en évaluation de programme, 2 <sup>e</sup> partie |
| 2005         | Vol. 28, n° 3             | Évaluation du Center for Substance Abuse Treatment   |
| 2006         | Vol. 29, n° 4             | Évaluation de l'effet des sites Internet destinés à répondre aux questions relatives à la santé (eHealth)  |
|              | Vol. 29, n° 3             | Évaluation des programmes communautaires de promotion de l'activité physique   |
|              | Vol. 29, n° 2             | La capacité de résistance des programmes   |
|              | Vol. 29, n° 1             | L'apprentissage organisationnel  |
| 2007         | Vol. 30, n° 4             | Éthique et évaluation  |
| 2008         | Vol. 31, n° 3             | Le «systems thinking» appliqué à l'évaluation  |

Notons, pour terminer, que le comité éditorial de la revue est constitué d'auteurs qui se situent essentiellement dans la foulée des théoriciens du «*theory-driven evaluation*» et de l'utilisation des modèles logiques, que l'on pense à Huey T. Chen et à Carol H. Weiss, par exemple, tous deux membres de ce comité. Tout comme pour les autres revues, on y accède par des bibliothèques qui sont abonnées et qui donnent ainsi électroniquement accès au contenu de chacun des numéros.

## 2.5. Quelques autres publications accessibles par Internet ou par bibliothèque abonnée

Sans prétendre à l'exhaustivité et après avoir touché quelques-uns des périodiques importants en matière d'évaluation, nous présentons ici une simple liste d'autres publications disponibles dans lesquelles il est également possible de réaliser des recherches.

- Bulletin SQEP, Société québécoise d'évaluation de programme, <[www.sqep.ca/pages/p\\_bulletin.htm](http://www.sqep.ca/pages/p_bulletin.htm)>. Français
- Mesure et évaluation en éducation, Association pour le développement de la mesure et de l'évaluation en éducation, <[www.umoncton.ca/raicheg/sitemee/](http://www.umoncton.ca/raicheg/sitemee/)>. Français/Anglais
- Journal of Multidisciplinary Evaluation, Evaluation Center, Western Michigan University, <[survey.ate.wmich.edu/jmde/index.php/jmde\\_1](http://survey.ate.wmich.edu/jmde/index.php/jmde_1)>. Anglais
- The Evaluation Exchange, Harvard Family Research Project, <[www.gse.harvard.edu/hfrp/eval.html](http://www.gse.harvard.edu/hfrp/eval.html)>. Anglais
- Aid evaluation and effectiveness, euforic (Forum européen de coopération internationale), <[www.euforic.org](http://www.euforic.org)>. Éditorial en ligne sur l'évaluation de l'aide Plusieurs langues
- Electronic Journal of Information Systems Evaluation (EJISE), <[www.ejise.com/index.htm](http://www.ejise.com/index.htm)>. *The Electronic Journal of Information Systems Evaluation is intended to provide critical perspectives on topics relevant to Information Systems Evaluation, with an emphasis on the organisational and management implications.* Anglais
- *Télescope*; Revue d'analyse comparée en administration publique, Observatoire de l'administration publique du Canada, <[www.observatoire.enap.ca/fr/index.aspx?sortcode=1.11.13](http://www.observatoire.enap.ca/fr/index.aspx?sortcode=1.11.13)>. Français

### 3. QUELQUES SITES INTERNET CONSACRÉS À DES QUESTIONS DE MÉTHODOLOGIES APPLIQUÉES À L'ÉVALUATION

Notons tout d'abord ici que l'opération de lister de tels sites demeure relativement périlleuse; en effet, d'une part, ces sites vivent et meurent parfois très rapidement tandis que, d'autre part, il n'est jamais vraiment certain que l'on ait touché les sites les plus pertinents, étant entendu que la question de la pertinence dépend de critères liés à telle ou telle démarche évaluative, voire aux préférences d'un évaluateur pour une approche plutôt qu'une autre. Ces quelques repères pourront constituer une base de départ qui pourra par la suite être bonifiée par la consultation des sites des sociétés nationales et autres que nous avons évoqué plus haut. À ce titre, par exemple, la consultation des bibliothèques virtuelles, telle que celle du World Wide Evaluation Information Gateway <[www.policy-evaluation.org/](http://www.policy-evaluation.org/)>, constituera une première démarche d'exploration. Voici donc une courte liste d'intérêt général.

- Le site de l'Université du Wisconsin: Program Development and Evaluation, <[www.uwex.edu/ces/pdande/evaluation/evaldocs.html](http://www.uwex.edu/ces/pdande/evaluation/evaldocs.html)>.
- La partie du site du Center for Disease Control dédiée à l'évaluation de programme, <[www.cdc.gov/eval/framework.htm](http://www.cdc.gov/eval/framework.htm)>.
- Le *Evaluation Handbook* de la Kellogg Foundation, <[www.wkkf.org/Pubs/Tools/Evaluation/Pub770.pdf](http://www.wkkf.org/Pubs/Tools/Evaluation/Pub770.pdf)>.
- Le site de conseils statistiques en ligne Hyperstat, <[davidmlane.com/hyperstat/](http://davidmlane.com/hyperstat/)>.
- Le site de recherche action essentiellement consacré aux approches qualitatives, Action Research Resources, <[www.scu.edu.au/schools/gcm/ar/arhome.html](http://www.scu.edu.au/schools/gcm/ar/arhome.html)>.
- Le site d'aide et de conseils à l'utilisation des méthodes statistiques de la compagnie de logiciels Statsoft, <[www.statsoft.com/textbook/stathome.html](http://www.statsoft.com/textbook/stathome.html)>.
- Le Instructional Assessment Resources de l'Université du Texas à Austin, <[www.utexas.edu/academic/diia/assessment/iar/programs/](http://www.utexas.edu/academic/diia/assessment/iar/programs/)>.
- Le site consacré aux approches qualitatives de George C. Boeree, <[webpace.ship.edu/cgboer/qualmeth.html](http://webpace.ship.edu/cgboer/qualmeth.html)>.
- Le site consacré à la réalisation de questionnaires et à leur analyse subséquente de Alison Galloway, <[webpace.ship.edu/cgboer/qualmeth.html](http://webpace.ship.edu/cgboer/qualmeth.html)>.
- Le *handbook* de l'évaluation de la National Science Foundation, <[www.nsf.gov/pubs/2002/nsf02057/start.htm](http://www.nsf.gov/pubs/2002/nsf02057/start.htm)>.
- Un site dédié aux questions d'échantillonnage et de sondage: The Survey System, <[www.surveysystem.com/sscalc.htm](http://www.surveysystem.com/sscalc.htm)>.

- Un site portail qui donne accès à un très grand nombre d'autres sites et ressources en évaluation: le World Wide Evaluation Information Gateway, <[www.policy-evaluation.org/](http://www.policy-evaluation.org/)>.
- Le site de l'évaluation pour le changement social, un site essentiellement destiné aux organisations à but non lucratif: le Innovation Network, <[www.innonet.org/](http://www.innonet.org/)>.
- L'Observatoire international en réussite scolaire du Centre de recherche et d'intervention sur la réussite scolaire (CRIRES); l'Observatoire propose un système de veille documentaire sur des pratiques et des programmes éducatifs qui ont fait l'objet d'une évaluation. Ces évaluations, les méthodes employées et les principaux résultats sont également accessibles, <[www.crires-oirs.ulaval.ca/](http://www.crires-oirs.ulaval.ca/)>.

## 4. QUELQUES OUVRAGES DE RÉFÉRENCE IMPORTANTS

Un peu comme dans le cas de la liste des sites Internet, la question du recensement des ouvrages importants en évaluation pourrait prêter flanc à la critique de la partialité: en vertu de quoi, en fait, mentionnerait-on tel ouvrage plutôt qu'un autre? Une consultation des syllabus des cours qui se donnent actuellement en évaluation de programme dans les universités québécoises permet cependant un recoupement fort utile qui ne garantira pas nécessairement une couverture exhaustive, mais nous donnera à tout le moins une base de départ. Nous présentons donc ici ces quelques ouvrages qui peuvent être considérés comme pratiquement incontournables pour qui exprimerait, par exemple, le désir de se constituer une bibliothèque en évaluation. Nous proposons ici une classification sommaire en trois catégories: 1) les ouvrages plus théoriques, 2) ceux qui sont essentiellement centrés sur des modalités de pratiques de l'évaluation et, finalement, 3) les «*textbooks*» qui se veulent des compendium à la fois d'aspects théoriques et de considérations pratiques.

### 4.1. Les ouvrages théoriques importants

C'est en 1989 que Egon Guba et Yvonna Lincoln publient *Fourth Generation Evaluation*, un ouvrage qui allait marquer une rupture fondamentale dans l'approche et les fondements mêmes de l'évaluation. C'est à partir de ce moment que la question du rôle de l'évaluateur et de sa relation au programme évalué passe au premier rang. Pour Guba et Lincoln, la recherche de l'impartialité est dorénavant un leurre; l'évaluateur doit donc jouer un rôle non seulement actif dans le programme mais, qui plus est, en être l'un des éléments moteurs.

Se présentant de manière peut-être un peu plus pragmatique, Huey-Tsyh Chen publie, en 1990, *Theory-Driven Evaluation*. Sans nécessairement créer une rupture de l'ordre de l'onde de choc provoquée par Guba et Lincoln, Chen inscrit tout de même un principe qui suscitera un repositionnement important de l'évaluation: en effet, la question de l'approche méthodologique passe dorénavant au second plan. Elle est, en quelque sorte, tributaire de ce dont le programme évalué est réellement constitué. C'est dans cet ouvrage que Chen proposait l'approche maintenant très utilisée du modèle logique comme une manière pour les évaluateurs d'en arriver à une compréhension mutuelle avec les parties prenantes des différentes dimensions, indicateurs et enjeux du programme évalué. Le même auteur publiera, en 2005, *Practical Program Evaluation*, toujours chez Sage, un ouvrage peut-être un peu plus proche des préoccupations pratiques exprimées par les praticiens de l'évaluation quant aux éléments fondamentaux du «comment faire» de l'évaluation.

Fetterman, Kaftarian et Wandersman présentaient, en 1996, dans l'ouvrage collectif qu'ils dirigeaient, la notion de l'évaluation comme moteur du développement des programmes. *Empowerment Evaluation* propose de faire en sorte que la démarche évaluative ne soit plus qu'affaire de spécialistes, mais qu'elle soit, au fur et à mesure qu'elle se déroule, transférée aux parties prenantes elles-mêmes. De ce fait, on passe de la notion d'évaluation extérieure à celle de l'autoévaluation.

## 4.2. Quelques ouvrages qui portent essentiellement sur la pratique de l'évaluation

Notons, tout d'abord, et en français, l'excellent manuel produit par l'Institut universitaire des Centres jeunesse de Montréal: *Cadre de référence pour le développement et l'évaluation des programmes aux Centres jeunesse de Montréal*. L'ouvrage, édité en 2000, est réalisé par Francine Paquette et François Chagnon. Même si l'ouvrage vise essentiellement les promoteurs de programmes d'intervention en centre jeunesse, ses valeurs heuristique et pratique en font un ouvrage de référence fort utile pour ce qui est de trouver des démarches types d'évaluation.

Quelques autres titres similaires à celui-ci ont également été produits par divers ministères et agences gouvernementales; indiquons, à titre d'exemples, les références suivantes:

- La stratégie nationale de prévention de la criminalité offre, à partir de son site Internet, toute une panoplie de ressources pratiques en évaluation, <[www.publicsafety.gc.ca/prg/cp/ncpc\\_tr-fra.aspx#a03](http://www.publicsafety.gc.ca/prg/cp/ncpc_tr-fra.aspx#a03)>.

- Sur Internet également, le Centre international de prévention de la criminalité propose quelques outils pratiques <[www.crime-prevention-intl.org/kb\\_tools\\_search.php](http://www.crime-prevention-intl.org/kb_tools_search.php)>, tout comme sa contrepartie américaine, la National Crime Prevention Council <[www.ncpc.org/](http://www.ncpc.org/)> et le ministère québécois de la Sécurité publique <[www.msp.gouv.qc.ca./prevention/prevention.asp?txtSection=publicat&txtCategorie=guide\\_eppapc&txtNomAutreFichier=chap5\\_elab\\_plan.htm](http://www.msp.gouv.qc.ca./prevention/prevention.asp?txtSection=publicat&txtCategorie=guide_eppapc&txtNomAutreFichier=chap5_elab_plan.htm)>.

Deux ouvrages américains relativement récents proposent des approches novatrices sur des thèmes clés de la démarche évaluative. Il s'agit, en premier lieu, du livre de Rita O'Sullivan (2004): *Practicing Evaluation, A Collaborative Approach*, publié chez Sage. On trouve ensuite cet ouvrage collectif très adapté à la pratique de l'évaluation de programme dans le contexte de la pratique en service social, et dirigé par Ian Shaw et Joyce Lishman (1999): *Evaluation and Social Work Practice*, aussi publié chez Sage.

Signalons un ouvrage important qui porte particulièrement sur les aspects éthiques, parfois très complexes, liés à l'évaluation, celui de Dianna L. Newman et de Roberts D. Brown, publié en 1996 chez Sage: *Applied Ethics for Program Evaluation*. Et, finalement, un dernier ouvrage important et en français de surcroît écrit par Jean-François Bernoux, *L'évaluation participative au service du développement social*, paru en 2004 aux Éditions Dunod.

### 4.3. Les textbooks

Nous terminerons notre tour d'horizon par une courte liste de ces ouvrages qui ont essentiellement comme cible les étudiants inscrits à des cours ou programmes en évaluation dans les universités américaines offrant ces formations. Ces ouvrages couvrent à la fois les fondements théoriques et considérations plus pratiques des démarches d'élaboration et d'évaluation des programmes d'intervention. Ils constituent donc d'excellentes introductions à ces deux domaines, qu'il est ensuite possible d'approfondir, qui par des articles de revue spécialisée, qui par des ouvrages plus pointus. Notons que, dans la plupart des cas, ces ouvrages sont périodiquement revus et réédités, comme c'est le cas des deux premiers ouvrages cités ici, et dont il s'agit de la septième édition.

- Rossi, P.H. *et al.* (2004). *Evaluation: A Systematic Approach*, 7<sup>e</sup> éd., Thousand Oaks, Sage.
- Posovac, E.J. et G.C. Raymond (2007). *Program Evaluation, Methods and Case Study*, 7<sup>e</sup> éd., Upper Saddle River, Prentice Hall.
- Royse, D. *et al.* (2001). *Program Evaluation, An Introduction*, 3<sup>e</sup> éd., Belmont, Wadsworth & Thompson Learning.



# CHAPITRE 12

---

**Les programmes  
se cachent pour mourir**  
Le bouclage de l'évaluation  
et de la programmation  
dans la modernité avancée

Denis Allard

Un programme d'intervention qui vise l'élimination d'un problème social repose sur l'idée simple d'une boucle récursive où l'obtention des incidences souhaitées devrait produire un effet en retour sur l'existence même du programme, à la limite mener à sa disparition. Dans cette grande boucle, l'évaluation est introduite comme une petite boucle réflexive qui permet de constater les incidences du programme et de prendre les décisions concernant sa poursuite. Telle est l'image fondatrice de l'évaluation lorsqu'elle vient en appui aux grands programmes sociaux américains dans les années 1960-1970, une image de « société d'expérimentation » (Airasian, 1987) où l'évaluateur, sur la base de la science et à l'aide de devis expérimentaux, devient le conseiller de l'État dans le choix et la rétention des programmes, ce qui lui donne en quelque sorte un « droit de vie ou de mort ».

En ce début de XXI<sup>e</sup> siècle, sans avoir perdu de sa pertinence comme soutien réflexif pour les concepteurs et les promoteurs des programmes, l'évaluation est bien loin de ce phantasme initial. La grande boucle de la programmation est plutôt considérée comme un vaste enchevêtrement de processus de régulation sociale à l'intérieur duquel l'évaluation hétéronome (celle des spécialistes) n'est qu'une boucle réflexive parmi d'autres. Celle-ci a une influence très variable sur la transformation des programmes, la plupart du temps bien loin de celle d'une décision de vie ou de mort. Si on postule que l'identité d'un programme se construit minimalement autour d'une vision claire des finalités poursuivies, des fondements théoriques de l'intervention proposée et de l'organisation du travail nécessaire à sa mise en œuvre, il faut admettre que la pratique contemporaine de l'évaluation se déploie assez souvent dans un univers de quasi-programmes ou de programmes moribonds. En admettant que la mort puisse prendre plusieurs visages, un évaluateur expérimenté saura la reconnaître, dès l'amorce d'une évaluation, sous la forme de « petites morts » (composantes essentielles du programme déjà abandonnées), de « morts nés » (programmes n'arrivant pas à prendre leur envol parce que le contexte y est largement défavorable), de « morts-vivants » (activisme de services qui camoufle une perte de vue des finalités), et bien sûr de « monstres » (méga-programmes recomposés par un État Frankenstein à partir de programmes moribonds). Ce côté obscur de la réalité des programmes existe bel et bien ; il est souvent occulté parce qu'il manifeste des situations bloquées d'où les concepteurs et les promoteurs des programmes ne veulent ou ne peuvent se sortir. Le titre du chapitre, qui paraphrase celui d'un best-seller célèbre (McCullough, 1999), veut exprimer cette occultation. Mais le travail de programmation a aussi un côté lumineux de réactions de vie, fait de reproductions, d'adaptations et d'innovations qui réussissent souvent à rendre les programmes plus pertinents et efficaces. C'est au cœur de ces transformations identitaires continuelles de vie et de

mort des programmes que la contribution de l'évaluation hétéronome prend forme et est débattue, au même titre que celle des autres acteurs du processus de programmation. Nous reviendrons en première partie de ce chapitre, à partir d'un cas exemplaire, sur ce **bouclage de l'évaluation** par le cadre plus large de la programmation et de son environnement.

Les évaluateurs ne sont jamais restés indifférents à cette relativisation de leur potentiel d'influence dans la grande boucle de la programmation. Cette faille est demeurée aux fondements mêmes du développement de la discipline, par des tentatives pour la combler ou la contourner. Et un des effets extrêmes de ce travail est de produire une tendance inverse du bouclage précédent, soit de développer la discipline, de telle sorte que de plus en plus de facettes du processus de programmation tombent sous son œil réflexif. La littérature en évaluation ne cesse de présenter des moyens d'élargir et de différencier le travail des évaluateurs dans la conception, l'implantation, le suivi des résultats et la pérennisation des programmes, de même que dans la mobilisation collective des acteurs pour y parvenir, le travail d'antéroaction (régulation de l'action par anticipation, Mairlot, 1982) prenant le pas sur celui, classique, de rétroaction. La croissance des outils, permettant de questionner en tout temps la vraisemblance d'un programme, en est le principal symptôme. Ce **bouclage de la programmation** par l'évaluation, qui est porté par une volonté de renforcer l'identité du programme pour mieux en maîtriser l'évaluation, aboutit paradoxalement à une représentation du programme qui échappe à une identité figée, une représentation multiple où se côtoient l'idée classique du programme comme traitement ou remède, et celle plus récente du programme comme réseau d'acteurs. Dans la seconde partie du chapitre, en nous limitant à un condensé historique, nous referons ce parcours disciplinaire.

Il est faux de croire, toutefois, que ce bouclage effectué par l'évaluation n'est qu'un effort méthodologique pour combler les failles de la rationalité instrumentale (les lumières de la connaissance et de la science au service de l'action), caractéristique essentielle de la modernité. Ce volet offensif est aussi accompagné d'un volet défensif d'ajustement et de prise en compte d'une « modernité avancée » (Giddens, 1993; Martuccelli, 1999), où le processus de subjectivation (Touraine, 2005), d'autonomisation et d'enracinement affectif et local des acteurs en réaction à la globalisation et à l'abstraction des pouvoirs sociaux pose des défis nouveaux aux disciplines d'intervention sociale. Les bouclages de l'évaluation et de la programmation seraient en fait sous l'influence d'un **bouclage par la subjectivation**. Dans une dernière partie, nous allons examiner cette hypothèse de travail et tenter d'en cerner la signification pour le travail en évaluation.

## 1. BOUCLAGE DE L'ÉVALUATION PAR LA PROGRAMMATION : UN CAS EXEMPLAIRE

Si l'évaluation, dans ses liens avec la programmation, est intégrée à une pluriboucle complexe de régulations sociales qui encadrent et limitent souvent son influence, il doit être possible, en prenant un peu de recul, d'en observer les manifestations dans l'évolution des programmes. La description, même sommaire, de la dynamique historique d'un cas concret peut nous aider. Le cas que nous avons choisi est exemplaire, dans le sens où le programme analysé, tout en faisant preuve d'efficacité, finit tout de même par être mis sur la touche. Il s'agit d'un programme de prévention des maux de dos qu'un organisme paritaire de santé et de sécurité du travail offre aux centres hospitaliers. Ce programme cherche à faire cheminer un comité interne patronal-syndical dans l'analyse, la recherche et l'implantation de solutions pour réduire l'incidence des maux de dos. Dans un centre hospitalier, le programme est démarré dans une unité fortement touchée par ce problème et doit par la suite être répliqué dans d'autres unités. Un retour sur les données d'un projet d'évaluation de ce programme (Allard et Ferron, 2000) permet de voir se déployer les multiples boucles de régulation d'un parcours de programmation.

1. La croissance initiale de la demande pour le programme est d'abord marquée par un contexte favorable. D'une part, l'organisme d'État responsable de la santé et la sécurité du travail exerce des pressions auprès des centres hospitaliers pour qu'ils réduisent par la prévention les coûts d'indemnisation reliés aux maux de dos; d'autre part, la plupart de ces centres ne disposent pas de l'expertise nécessaire pour mettre en place une démarche paritaire et systématique de prévention. Le programme proposé par l'organisme paritaire vient à point. Il implique que ses conseillers accompagnent les comités patronaux-syndicaux dans l'implantation de toutes les étapes du programme et assurent un suivi des données d'incidence des maux de dos et de mise en place des solutions proposées. Ce processus **autoévaluatif** aide à systématiser le programme, à en raffermir les principes, les conditions et les étapes de réalisation, en même temps qu'il donne des munitions pour sa promotion, d'autant que les résultats sont positifs dans la majorité des unités participantes. Ce début d'implantation du programme peut être associé à une **boucle réflexive** (Morin, 2001), où le programme est ajusté grâce aux retours d'information tirés des premières expériences.
2. La **systématisation** du programme conduit rapidement, toutefois, à des phénomènes de ramification de l'intervention qui, à terme, menacent l'intégrité et la demande pour le programme. Puisque la réalisation du programme implique des recommandations et la mise en œuvre de solu-

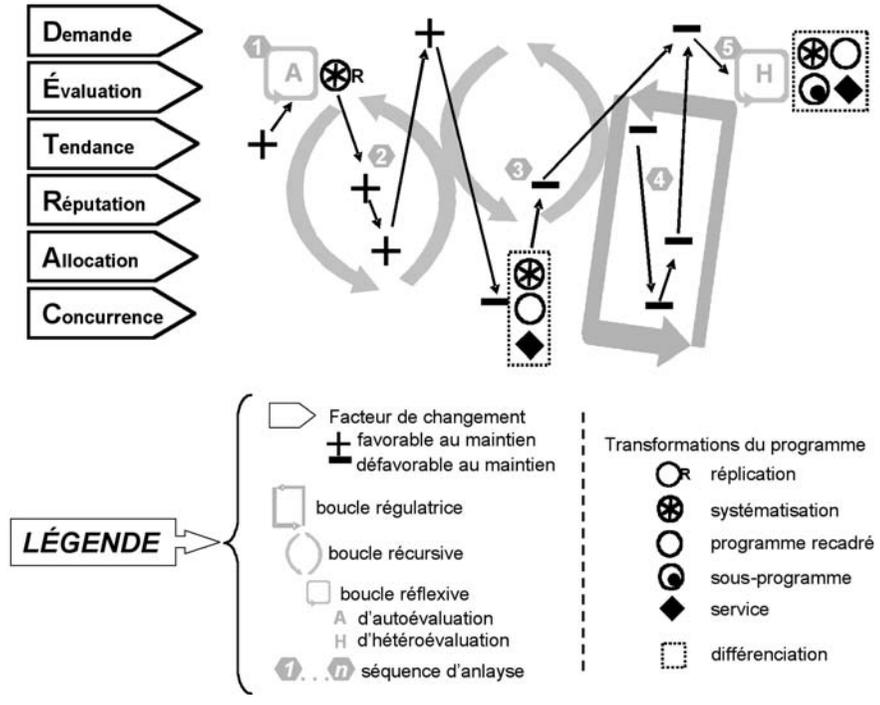
tions pour prévenir les maux de dos, l'organisme paritaire se voit dans l'obligation de participer au développement de certaines solutions qui ne sont pas adéquates ou disponibles sur le marché. C'est ainsi qu'il est appelé à travailler avec un manufacturier à la conception d'un prototype de lit électrique qui correspond aux exigences de prévention identifiées par les comités paritaires dans le cadre de la réalisation du programme. Ce premier succès en appelle d'autres qui servent d'arguments de vente du programme. Son attrait pour les centres hospitaliers devient de plus en plus lié à cette dimension technique de déplacement sécuritaire des patients, à un point tel que l'adoption des solutions développées finit par précéder, si ce n'est évincer, le processus d'analyse des problèmes que suppose le programme. L'idée de répliquer le programme en entier dans d'autres unités, avec sa démarche longue et coûteuse en temps-ressources, devient moins alléchante pour les institutions les moins fortunées. Déjà, certains conseillers proposent une version recadrée et condensée du programme, sous-titrée « soutien technique », centrée sur un diagnostic rapide d'un expert (le conseiller lui-même) et la référence à une banque de solutions déjà éprouvées dans les milieux précédents. La demande augmente, mais elle n'est plus tout à fait la même. On peut y voir une **boucle réursive** (Morin, 1977) où certains produits du programme (recommandations, innovations techniques) commencent à agir sur sa nature même.

3. La version recadrée du programme n'est que l'amorce d'un processus de différenciation du programme qui prolonge l'effet du virage technique déjà amorcé. Ce virage force dans son sillage une redéfinition des compétences professionnelles des conseillers de l'organisme promoteur. De généralistes en prévention centrés sur l'accompagnement de la démarche paritaire du programme, les conseillers en viennent à former un groupe composé de plus en plus de spécialistes en ergonomie. Les embauches et les incitatifs financiers au recyclage professionnel favorisent cette tangente. En contact fréquent avec les acteurs des milieux hospitaliers, les conseillers finissent par faire dériver une partie de la demande sur des besoins particuliers d'interventions ergonomiques qui ne nécessitent pas l'appareillage du programme. La demande pour le programme complet a donc tendance à diminuer alors qu'apparaît une différenciation en quasi-programmes de prévention (**version recadrée** du programme initial orientée vers une sortie rapide sur les solutions, **sous-programmes** d'accompagnement des comités paritaires dans leur planification de la prévention pour l'ensemble du centre hospitalier ou de formation sur le déplacement sécuritaire des patients) et en **services** spécialisés d'ergonomie (études de postes particuliers, soutien ergonomique à la conception architecturale de nouveaux espaces de soins).

4. Les effets récurifs de la bifurcation technique du programme ne rendent pas compte toutefois de l'ensemble du phénomène de baisse de la demande. Il faut pour cela examiner des processus régulateurs (Barel, 1989) de plus grande ampleur et principalement externes au programme. La marginalisation du programme est aussi reliée à l'apparition de stratégies parallèles non préventives de réduction des coûts d'indemnisation associés aux maux de dos, stratégies largement favorisées par l'organisme d'État responsable de la santé et de la sécurité du travail. Cet organisme décide, entre autres, de moduler les cotisations des centres hospitaliers en fonction des statistiques d'accidents de chaque institution plutôt que sur celles de l'ensemble du secteur. Certains centres hospitaliers y voient la possibilité, en agissant rapidement, de faire des économies par la réduction du nombre et de la durée des accidents. D'où le recours à des mécanismes de suivis médicaux plus serrés auprès de travailleurs accidentés et l'accélération du retour au travail en réaffectant ceux-ci dans des postes moins dangereux. De là, aussi, le renforcement de la boucle réursive précédente de demande d'interventions plus rapides et spécialisées en ergonomie des postes de travail, ce qui entre en concurrence avec la démarche plus lourde du programme complet. À cela s'ajoute le fait que la situation financière de plus en plus précaire de nombreux centres hospitaliers rend difficile l'utilisation des ressources humaines pour une participation à un programme de longue durée.
5. La question de la baisse de la demande soulève des questionnements chez les dirigeants et le personnel de l'organisme paritaire. Elle aboutit à une évaluation hétéronome, réalisée par des évaluateurs externes à l'organisation, qui a pour objectifs de valider les résultats déjà obtenus par l'autoévaluation et d'expliquer cette baisse d'intérêt des centres hospitaliers. Les résultats de cette évaluation permettent d'organiser et de donner un sens global à un ensemble de constats séparés qui circulent déjà parmi les acteurs du programme. L'organisme obtient une réponse mieux documentée sur les raisons cette baisse de la demande (voir les points précédents) et constate l'inéluctabilité de la différenciation de son programme dans le cadre des transformations de la problématique de la santé et de la sécurité du travail. Le programme complet n'est pas déqualifié dans ses fondements, ce sont les conditions de sa réplication généralisée qui ne sont plus présentes.

Figure 12.1

### Bouclage de l'évaluation par la programmation : un cas exemplaire



Cet exemple montre l'encerclement des potentiels régulateurs de l'évaluation sur le long terme d'un processus de programmation. La figure 12.1 en donne une formalisation graphique qui resitue cette dynamique par rapport aux principaux facteurs de changement évoqués dans l'analyse précédente du cas. Le premier facteur est la **demande** pour le programme, inséparable d'ailleurs de l'offre qui en est faite. Le second est l'**évaluation**, soit autonome (produite par les promoteurs ou les acteurs du programme), soit hétéronome (produite par des spécialistes externes). Le troisième est la **tendance** prise par la problématisation de l'intervention, c'est-à-dire le lien défini entre l'évolution du problème ciblé et sa représentation sociale, d'une part, et l'orientation de l'intervention, d'autre part. Le quatrième tient en la **réputation** du programme auprès des clients ou des bénéficiaires, réels ou potentiels ; elle est constituée de jugements globaux sur la qualité du programme qui n'ont pas nécessairement de liens systématiques avec les évaluations formelles. Le cinquième est l'**allocation** des ressources nécessaires à la mise en œuvre et la continuité du programme. Le sixième est la **concurrence**, c'est-à-dire la

présence d'autres programmes ou interventions, visant le même problème, susceptibles de jouer sur les facteurs précédents. Tous ces facteurs interagissent pour réguler le programme dans des sens qui soit favorisent son maintien et sa reproduction, soit font dévier sa marche, son fonctionnement et ses mécanismes, littéralement le « détraquent ». Une analyse comparative avec d'autres cas permettrait d'affiner ce modèle DÉTRAC, mais il s'agit d'un autre défi que celui d'illustrer le bouclage de l'évaluation par la programmation, bouclage auquel d'ailleurs la discipline de l'évaluation de programme n'a eu de cesse de réagir au cours de sa brève histoire.

## 2. BOUCLAGE DE LA PROGRAMMATION PAR L'ÉVALUATION : UN CONDENSÉ D'HISTOIRE

La discipline de l'évaluation de programme reste marquée par ses origines et son désir de soutenir la grande boucle de la programmation par une boucle réflexive qui apporte la preuve de l'effet du programme et peut fonder les décisions subséquentes sur la connaissance scientifique. Dans les années 1960-1970, cette preuve a trouvé son champion méthodologique dans la transposition du devis expérimental, largement utilisé en psychologie et en médecine, à l'évaluation des programmes sociaux des grandes réformes sociales (Suchman, 1967; Airasian, 1987; Campbell, 1994). L'idée de base de ce devis est de mesurer l'écart des résultats obtenus entre deux groupes, dans lesquels les individus sont idéalement attribués au hasard, qu'on cherche à rendre différents sur un seul facteur: le fait d'être exposé ou non au programme. L'écart mesuré est parfois appelé l'impact et correspond à l'effet particulier du programme. Cette forme de réflexivité correspond bien aux questionnements des bailleurs de fonds: « Est-ce que le programme fait une différence? » ou, autrement dit: « Est-ce que nous en avons pour notre argent? ». Pour reprendre la distinction de Mark *et al.* (2000), cette logique réflexive est « molaire ». Elle ne cherche pas l'analyse « moléculaire » des multiples variables de la « boîte noire » qu'est le programme et de leur fonctionnalité dans différents contextes d'implantation. Elle veut plutôt l'argument massue qui facilite les choix et les décisions sur l'ensemble du programme.

Cette boucle réflexive se heurte toutefois rapidement à des obstacles qui en limitent l'utilité. Ils sont de trois ordres. Les premiers relèvent de la mise en œuvre du devis expérimental comme tel. L'attribution au hasard des individus, des groupes ou des communautés à un groupe expérimental, qui reçoit le programme, et à un groupe témoin, qui en est exclu, est techniquement et éthiquement difficile (Lam *et al.*, 1994), sinon impossible dans certaines situations (Devine *et al.*, 1994). Une solution est de limiter l'utilisation du devis aux situations qui en permettent une actualisation adéquate

(Orr, 1999). Une autre est de passer outre la condition initiale d'attribution au hasard et d'utiliser des devis quasi-expérimentaux qui maintiennent l'exigence d'une comparaison entre le fait d'être ou de ne pas être exposé au programme, tout en précisant les contextes où leur utilisation est la plus appropriée (Bingham et Felbinger, 1989). Les seconds obstacles relèvent de la qualité de la conception et de l'implantation du programme. Des résultats d'évaluation d'impact obtenus sur un programme mal conçu ou mal implanté ne sont pas des résultats d'évaluation qui résistent à la critique. Pour surmonter ces limites, on propose de faire précéder l'évaluation d'impact d'une étude d'évaluabilité (Smith, 1989) pour s'assurer que les objectifs, les moyens et les ressources du programme sont bien identifiés et agencés. Par ailleurs, il devient crucial aussi de vérifier que le programme bénéficie d'une implantation minimale de ses composantes essentielles avant de se lancer à la conquête coûteuse des données d'impact. Le projet évaluatif prend alors la forme d'un emboîtement où l'évaluation d'impact est précédée par la vérification de l'évaluabilité et celle de l'implantation, regroupées parfois en une seule démarche (Wholey, 1979). Toutefois, cette séquence ne résiste pas aux obstacles d'un troisième ordre relatif à la solidité des résultats d'évaluation d'impact, et à leur utilisation par les acteurs du programme. Ces résultats, qu'ils démontrent clairement ou non l'effet du programme, donnent une vision molaire qui privilégie le questionnement des décideurs politiques ou des bailleurs de fonds, et laissent de côté celle, plus moléculaire, des intervenants de terrain et des bénéficiaires. Or, surtout quand l'effet n'est pas clairement démontré, sur quoi faire porter les correctifs à apporter à un programme si les résultats d'évaluation ne précisent pas les processus qui en sont à l'origine ?

À partir de telles remises en question, les stratégies d'évaluation s'élargissent pour devenir plus sensibles aux besoins des acteurs (« *responsive evaluation* »; Abma et Stakes, 2001) et plus pénétrantes quant aux conditions et processus d'implantation (McLaughlin, 1987; Roberts-Gray et Scheirer, 1988; Yin, 1997). Cet élargissement favorise la pénétration de l'évaluation dans la vie et la réflexivité des acteurs et des organisations. Ce qui était préalable et rattaché à l'évaluation d'impact, soit l'étude d'évaluabilité et l'analyse d'implantation, devient à la fois plus autonome et fondamental. À un point tel que se produit lentement une inversion de la séquence d'emboîtement et sa restructuration dans un ensemble plus large de questionnements évaluatifs enchevêtrés de ce que nous appelons la performance (le rapport entre finalisation et action), la résilience (le rapport entre conformité et adaptation) et la vraisemblance (le rapport entre anticipation et expérience). De plus, la dynamique de ces questionnements est structurée par ce qui est considéré dorénavant essentiel à la démarche évaluative, soit le développement de la théorie des programmes (Bickman, 1987; Weiss, 1997) et la

participation des acteurs (Gregory, 2001 ; Ryan, 2004). La partie gauche de la figure 12.2 illustre cette inversion qui annonce un mouvement disciplinaire de bouclage de la programmation par l'évaluation.

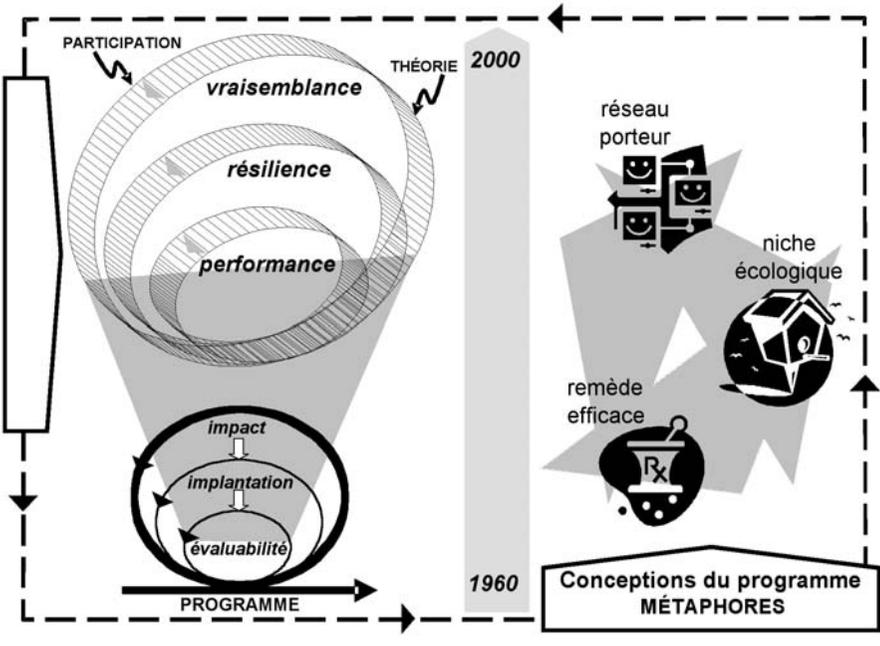
Le signe le plus spectaculaire de ce bouclage est sans doute le développement de l'évaluation de **vraisemblance**, qui est une forme de régulation du programme par anticipation. Elle passe du stade de garant de l'évaluation (Wholey, 1994), avec les études d'évaluabilité, à celui de garant de la programmation, avec une panoplie de techniques associant les acteurs autour d'efforts de conception d'un programme qui fasse collectivement du sens. Cette évolution se cristallise d'abord autour de l'idée de modélisation du programme dans des formes narratives et figuratives (McLaughlin et Jordan, 1999). Ces dernières, par les schémas, deviennent des miroirs (Allard *et al.*, 2008) autour desquels les acteurs peuvent penser le changement comme un cheminement, une séquence d'étapes à franchir, de conditions à réaliser, de résultats à obtenir. Ils peuvent y voir l'importance de montrer les traces du franchissement de ces étapes et appuyer l'évaluation dans sa quête d'indicateurs. Un second saut qualitatif tient à l'approfondissement des fondements et postulats théoriques sur lesquels reposent les passages d'une étape à l'autre, et l'obtention des résultats souhaités (Leeuw, 2003). Autant le modèle de programme fait partie maintenant de l'arsenal de base de l'évaluateur, autant le recours à des synthèses d'évaluations par types d'intervention et par mécanismes de changement (Pawson, 2002) devient un moyen d'argumenter sur les choix du programme anticipé ou du programme remis en question. Les synthèses, reconnaissant les limites des recherches évaluatives, vont plus loin en intégrant parfois les points de vue des acteurs (Oliver *et al.*, 2005) sur le potentiel et les limites contextuelles d'un programme. Les progrès les plus récents de l'évaluation de vraisemblance se concrétisent autour d'un tel dialogue touchant les conditions d'implantation (Rohrbach *et al.*, 2006) et les conditions de pérennisation (Scheirer, 2005). Ce dialogue intègre la prise en compte du réseau de programmes et de politiques sociales dans lequel il va se retrouver, parfois en concurrence. En s'élargissant de la sorte, l'évaluation de vraisemblance devient littéralement une porte ouverte sur celle de la planification organisationnelle (Wandersman *et al.*, 2000 ; Bernoux, 2004 ; Allard *et al.*, 2007), ce qui renforce le processus de bouclage de la programmation.

L'évaluation de vraisemblance stimule et dépend tout à la fois des travaux d'évaluation de résilience et de performance. L'évaluation de **résilience** examine la mise en œuvre des programmes, soit du point de vue de la conformité d'implantation, soit du point de vue de l'adaptation aux contextes, idéalement des deux. Du côté conformité, on voit se développer des démarches évaluatives qui insistent sur l'analyse de la mise en place des composantes essentielles des programmes (Cousins *et al.*, 2004). Du côté

de l'adaptation, la recherche prend de nombreuses directions. L'évaluation réaliste en est un exemple qui cherche à décortiquer les programmes en microthéories de l'intervention (Pawson et Tilley, 1997). Elle constate que les interventions ont des mécanismes sous-jacents qui ne produisent des résultats que pour certains groupes et dans certaines circonstances. Par exemple, un programme de gratuité de détecteurs de fumée pour la prévention des incendies ne produira pas nécessairement les comportements escomptés dans des familles où, faute de ressources, les piles sont récupérées pour des usages alternatifs. D'où l'idée d'identifier, souvent avec les acteurs, des combinatoires contextes-mécanismes-résultats, de les expérimenter sur le terrain, et de faire ainsi évoluer les fondements théoriques de l'intervention à partir de ce travail d'essais-erreurs.

Figure 12.2

### Le bouclage de la programmation par l'évaluation : inversion et superposition



Les évaluateurs réalistes reconnaissent, par ailleurs, que la réalité des programmes est soumise à une causalité générative qui dépasse le cadre opérationnel de l'intervention, d'où la nécessité d'avoir recours à des théories qui débordent ce cadre (Barnes *et al.*, 2003) pour saisir des facteurs qui

concernent l'ensemble et l'au-delà du programme. On a pu ainsi analyser des facteurs comme l'aveuglement volontaire des acteurs (Freidman, 2004), les métamessages contradictoires du programme (Dahler-Larson, 2001) ou les pressions politiques (Davis, 2005). En complémentarité de ces analyses, qui isolent des facteurs structurants, d'autres évaluations se concentrent sur les processus inhérents au développement des programmes. McLaughlin (1987) avait déjà mis la table au milieu des années 1980 en affirmant que le programme est un système d'implantation qui désigne plus un apprentissage qu'une installation, que cet apprentissage, et les résultats qui en découlent, sont médiatisés dans des choix locaux par les acteurs qui sont les plus près du terrain. Les travaux plus récents montrent avec plus de clarté les régulations qui sont au cœur même des transformations des programmes (St-Pierre et Kaltreider, 2004) et la construction-déconstruction des réseaux d'acteurs qui les portent (Bilodeau *et al.*, 2002; Koivisto *et al.*, 2006).

Limitées souvent au cadre des résultats opérationnels du programme («*outputs*»), les évaluations de résilience, puisque reliées aux questionnements de vraisemblance du programme, ne peuvent pas ne pas s'intéresser aux effets du programme sur les groupes et les populations visées («*outcomes*»). Elles sont donc dépendantes ultimement, comme les évaluations de vraisemblance, des travaux d'évaluation de **performance**, qui sont marqués eux aussi par une ouverture des perspectives méthodologiques. La construction de l'argumentation sur l'incidence d'un programme n'est plus limitée qu'aux seuls apports des devis expérimentaux et quasi expérimentaux, eux-mêmes assouplis par le progrès des outils statistiques (Rossi, 1997). Ils reposent de plus en plus sur ce qu'on pourrait appeler la « traçabilité » des effets détectés chez les participants, que ce soit par le biais d'études qualitatives des chaînes de causalité (Mohr, 1999) ou par des observations provoquées, comme dans l'expérimentation à cas unique (Bloom, 1999) ou la recherche-action (Rothman, 2002). Dans leurs retranchements les plus poussés, les effets du programme sont examinés, d'une part, du point de vue du sens qu'ils prennent dans la vie de chaque individu (Kushner, 2000) et, d'autre part, du point de vue des conséquences ignorées, imprévues ou imprévisibles (Morell, 2005) qui sont générées. Dans ce dernier cas, le travail de recherche n'est possible qu'à partir d'hypothèses sur les trajectoires possibles du programme, ce qui renvoie à l'évaluation de vraisemblance. La boucle est ainsi bouclée.

Ce compte rendu historique de l'inversion de l'évaluation, qui s'apparente à un bouclage de la programmation dans le cadre de ses pratiques disciplinaires, demeure évidemment sommaire. Il faudrait le reprendre et l'argumenter avec une chronologie plus détaillée. Mais il est suffisant pour en montrer le mouvement volontariste de consolidation des programmes. Toutefois, il aboutit en même temps, et paradoxalement, à un mouvement

contraire, dans la mesure où les représentations du programme y prennent des formes qui échappent de plus en plus à une identité figée. La manière la plus éclairante que nous avons trouvée pour en parler est l'utilisation de trois métaphores qui correspondent à trois visions qui se succèdent dans le temps (voir la partie droite de la figure 12.2). La première est celle du programme comme **remède efficace**. Elle est à la source même du projet évaluatif et de l'assimilation du programme à un traitement dont ceux qui y ou s'y investissent veulent s'assurer de l'efficacité, de la performance. La seconde est celle du programme comme **niche écologique** qui représente l'adaptation mutuelle d'une espèce vivante (le programme) et de son environnement. L'image du nichoir d'oiseaux veut souligner le fait que la transformation mutuelle du programme et de son environnement n'est possible qu'à l'intérieur de certaines limites. Ce ne sont pas toutes les hirondelles qui utilisent un nichoir, ni tous les nichoirs qui conviennent aux hirondelles pouvant en utiliser un. La troisième métaphore est celle du réseau informatique. Elle fait appel à une vision plus sociale du programme, c'est-à-dire une proposition d'action autour de laquelle un réseau d'acteurs doit se constituer et faire consensus (l'image choisie présente un réseau informatique de moniteurs ayant le sourire). Le programme ne peut tenir que s'il se constitue en **réseau porteur** d'un sens partagé. Ce sens, autant que le réseau qui le porte, est en constante négociation. L'identité du programme est alors problématique et son contenu, plus dynamique.

Ce mouvement paradoxal de bouleversement des représentations du programme nous signale le fait que l'évaluation ne s'est pas limitée qu'à effectuer un bouclage de la programmation. Elle a dû réagir à quelque chose de plus fondamental qui relève de l'évolution de la société moderne. L'institution de la programmation, comme celle de l'évaluation d'ailleurs, sont basées sur des rapports sociaux qui cherchent à rationaliser l'action, à intégrer la connaissance rationnelle dans les fibres de l'action collective, ce qui est une des caractéristiques de la modernité. Mais une autre de ces caractéristiques est de faire des individus les sujets de leur action, les bâtisseurs de leur identité, les décideurs des liens sociaux qu'ils veulent privilégier. Ce processus de subjectivation atteint un niveau avancé dans la modernité actuelle alors que la réflexivité individuelle, la capacité de réfléchir sur soi et son action, dispose de moyens sans précédent (par exemple Internet). Ce potentiel de réflexivité nouveau rend certains rapports sociaux plus éloignés, plus abstraits. Par exemple, les soubresauts de l'économie nous touchent directement, mais la conscience qu'on a des acteurs qui en sont la cause demeure relativement floue, et la plupart de acteurs eux-mêmes, inaccessibles. En réaction à cette abstraction, le processus de subjectivation manifeste un mouvement contraire vers le renforcement de liens sociaux plus concrets, plus rapprochés. Le passage de la conception d'un programme comme une imposition, le remède, à une invitation, le réseau, serait la reconnaissance et l'adaptation

du champ disciplinaire de l'évaluation à ce processus de subjectivation, la conception du programme comme réseau permettant à la fois de garantir des liens de proximité et une portée plus grande, puisqu'un réseau peut s'élargir sur les systèmes abstraits. La partie suivante du chapitre cherche à mettre en chantier cette hypothèse de travail et à en voir les implications pour la programmation et l'évaluation.

### 3. BOUCLAGE DES DEUX PAR LA SUBJECTIVATION : UNE HYPOTHÈSE DE TRAVAIL

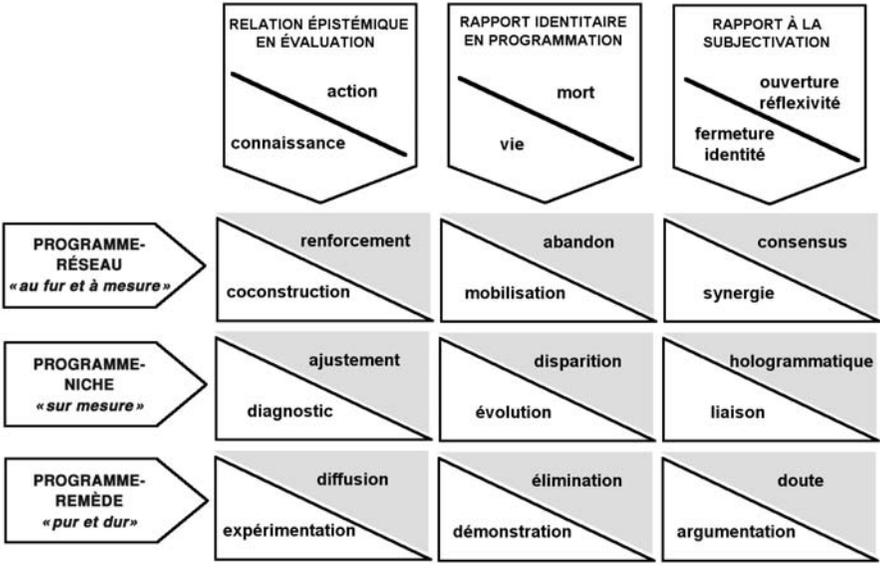
Pour aborder ce chantier, il importe d'abord de bien marquer les différences de signification de ce changement des représentations du programme du point de vue de l'évaluation et de la programmation. Pour l'évaluation, on peut s'en rendre compte dans la modification du lien épistémique qui est tissé entre la connaissance et l'action. Pour la programmation, il est possible d'en avoir des indications dans la manière dont la vie et la mort du programme sont définies, question aussi de revenir sur la thématique de départ. Ces analyses ouvrent la porte, par la suite, à l'examen du rapport de la programmation et de l'évaluation à la subjectivation.

On peut passer rapidement sur le premier point de la relation épistémique entre connaissance et action puisqu'elle est en filigrane de la section précédente. Il suffit de rappeler qu'avec la représentation du programme-remède, la connaissance est de l'ordre de l'**expérimentation** et l'action de celui de la **diffusion** (voir la figure 12.3). La « société d'expérimentation » est celle de la science qui expérimente les programmes et peut indiquer par la suite lesquels ont réussi le test de la performance. Dans la représentation du programme-niche, la connaissance évaluative repose sur le **diagnostic** (Allard, 1996) de conformité aux intentions et d'adaptation aux contraintes environnementales. L'action qui en découle se formule donc en **ajustements**, dans la création d'une niche écologique qui convient à la fois au programme et à son environnement. La représentation du programme-réseau fait de la connaissance une **coconstruction** par les acteurs mêmes du programme. Elle pose l'action comme un **renforcement** de ce réseau pour y maintenir les acteurs déjà en place et y ajouter ceux qui deviennent significatifs avec le temps.

Il est aussi relativement facile de percevoir les conceptions du rapport identitaire vie-mort dans la programmation au travers des mêmes catégories de programme. Avec le programme-remède, le processus de programmation fait de la vie du programme un acte de **démonstration** d'efficacité et d'impact. L'image de la « société d'expérimentation », même si elle s'avère irréalisable,

Figure 12.3

## Le bouclage par la subjectivation dans la modernité avancée



continue de trouver un écho dans le recours à des transitions allant du projet-pilote expérimental, le programme « pur et dur », à des transpositions en milieux naturels de pratique sous la forme de projets de démonstration ou à une implantation généralisée dans un grand nombre de sites (Berk et Rossi, 1990; Straw et Herrell, 2002). À la limite, l'échec de la démonstration ouvre la porte à un processus décisionnel de rupture, d'**élimination** du programme de la sphère des programmes viables. Lorsque la pression sociale pour une intervention rapide n'est pas trop forte, le projet-pilote à petite échelle d'une intervention simple (par exemple, la procuration d'un incitatif financier), qui mobilise temporairement des ressources nouvelles intégrées à un processus de recherche, offre la meilleure possibilité d'agir en ces termes décisionnels de maintien ou d'élimination. Si un des facteurs précédents est modifié, ce qui est la règle plutôt que l'exception, cette possibilité est réduite d'autant (Tilley, 1996). Une intervention complexe (par exemple, le développement communautaire des compétences parentales), utilisant les ressources existantes d'organisations de services, induit une multitude de processus d'appropriation ou de rejet avant même d'en arriver aux actions qui peuvent produire un effet propre sur une population ciblée. On peut penser aux effets de l'introduction d'un tel programme sur le processus de professionnalisation, celui-ci pouvant être, dans certains cas, l'occasion d'une affirmation de

nouvelles compétences ou, dans d'autres, une menace à l'autonomie professionnelle (Lallement, 2007). Que dire, en plus, lorsque la pression sociale et politique pour consolider une intervention sur le terrain se fait omniprésente ? Il peut en résulter l'introduction généralisée d'un programme avant même l'obtention de résultats d'évaluation ou une interprétation biaisée de résultats préliminaires permettant cette extension (Stevens et Dial, 1994). Le potentiel d'impact se heurte à la niche écologique.

Dans la représentation du programme-niche, la programmation est vécue comme un parcours **évolutif** où le programme, comme une espèce vivante, s'adapte aux contextes auxquels il est confronté. Les mutations, différenciations et spécialisations qu'il subit en font un projet « sur mesure ». Il est rare, dans ces situations, de voir l'élimination complète d'un programme. On observe plutôt de multiples formes de transformations qui, dans certains cas, si ce n'était du maintien de l'étiquette (« *label fallacy* », McLaughlin, 1987), pourraient passer pour des **disparitions** tant la nouvelle forme se distancie de la forme initiale. Les identités multiples des programmes manifestent des adaptations qui sont autant, pour reprendre la vieille distinction de Piaget (1967), des « accommodations », des ajustements **à** l'environnement, que des « assimilations », des ajustements **de** l'environnement. Quand l'environnement est immédiat et concret, le potentiel d'assimilation, d'ajustements actifs est plus grand. Quand l'environnement est lointain et abstrait, il y a une probabilité plus grande d'ajustements passifs, sauf si les acteurs du programme élargissent leur réseau pour rejoindre, faire des liens, concrétiser ce qui est loin et abstrait. Le programme se constitue alors en réseau.

Avec la représentation du programme-réseau, les identités multiples deviennent une identité flexible qui se définit « au fur et à mesure » (rappelons que « *fur* » a pour origine le mot latin « forum ») que se structure le réseau qui la porte. Le programme équivaut à un ensemble de relations hologrammatiques (Morin, 1986) où le tout est dans les parties et les parties dans le tout. Il est d'autant plus solide que le tout, le programme où se trouvent **mobilisés** et associés les acteurs, est dans chaque partie, chaque acteur lui-même, et que chaque acteur y reconnaît son association. Le réseau entraîne dans sa dynamique les acteurs qui donnent un sens au programme en même temps que chaque acteur trouve un sens, en synergie avec les autres, à son action dans le programme. Dans cette perspective, la mort d'un programme ne peut être définie que comme un **abandon**, le couple acteurs-réseau n'arrivant plus à faire vivre l'intéressement et la mobilisation autour de l'élaboration d'un consensus d'action.

Du panorama de ce qui semble être une transformation de l'évaluation et de la programmation en lien avec les représentations du programme, doit-on conclure que l'aboutissement de l'évaluation est dans un mouvement incessant des consensus et des dissensus portés par des réseaux, et que la

programmation, comme processus identitaire de rationalisation de l'action, se dissout peu à peu dans l'incertitude de ce mouvement ? Il faudrait pour cela admettre que la nouvelle représentation fait maison nette des précédentes et que le programme-réseau se forge dans le vide. Or il n'en est rien. Ce programme part aussi d'une idée, d'une innovation et d'intentions qu'un ou plusieurs promoteurs mettent de l'avant. Pour affirmer leur leadership dans les délibérations des acteurs et en mobiliser de nouveaux, ces promoteurs doivent être porteurs d'une identité pour celui-ci et ne peuvent ignorer, sous peine d'être considérés rétrogrades, les fondements rationnels qui sont fournis par les acteurs scientifiques. Pour aller de l'avant, le programme doit aussi faire un cheminement d'expériences et d'apprentissages (Jenlink, 1994) dans l'écologie des racines historiques et des routines organisationnelles des milieux d'intervention. Tout cela colore les consensus. À la limite, il faut bien admettre que les trois visions successives du programme ne sont pas exclusives l'une de l'autre, mais en forment plutôt une seule en superposition systémique (Barel, 1989). Et même si le processus de subjectivation de la modernité avancée trouve un meilleur écho dans le programme-réseau, il ne peut pas ne pas interagir avec les autres représentations. Il est possible d'en rendre compte en l'examinant dans son caractère paradoxal, souligné par Kaufmann (2001 et 2004), d'une ouverture liée à la réflexivité et d'une fermeture associée à la recherche d'identité.

Dans leur modèle d'une quatrième génération de l'évaluation, qu'ils appellent le cercle herméneutique dialectique, Guba et Lincoln (1989) soumettent un programme à des délibérations poussées entre des acteurs représentant les points de vue les plus variés possibles et leur demandent de négocier une version plus articulée et consensuelle d'un nouveau programme. Ils admettent que le consensus ne sera pas atteignable sur tous les points, que les derniers objets de discordance devront faire l'objet d'un arbitrage. Ils reconnaissent à la recherche scientifique un rôle déterminant dans ce travail. La démarche de Mason et Mitroff (1981), en planification stratégique, se situe dans les mêmes eaux et aboutit aux mêmes conclusions sur la contribution que peut avoir la recherche dans le rapprochement des positions opposées. Coconstruire un programme ou un plan en réseau, comme ils le proposent, n'est donc pas une négation du rôle de la science mais bien son articulation nouvelle entre deux positions : celle de l'ouverture réflexive des sujets qui insufflent un **doute** récurrent dans la connaissance, celle de la fermeture identitaire qui est capable de mettre sur la table la meilleure **argumentation** possible appuyée par des validations empiriques (Giddens, 1993 ; Lipsey, 1997). Toutefois, il faut reconnaître aujourd'hui une difficulté supplémentaire dans la mise en œuvre de conditions favorables à ce dialogue, entre autres par la disponibilité accrue de moyens de connaissance chez les acteurs. Elle peut se mêler aux phénomènes de marchandisation généralisée (Enriquez, 2007) et de vente de soi (Guienne, 2007), en croissance dans la modernité

avancée, qui n'épargnent pas la production-diffusion des connaissances. Le risque est que le réseau ne devienne une vaste plate-forme marketing pour des acteurs et des programmes concurrents.

Le processus de subjectivation influence aussi la programmation-réseau dans son interaction avec la perspective écologique du programme-niche parce qu'elle y trouve le terreau de son identité, les acteurs pouvant y vivre la reconnaissance mutuelle de ce qui fait leur **liaison**. Les traces des compromis et des règles de fonctionnement qui découlent des négociations (Thuderoz, 2000) et du partage intime des difficultés et des succès dans les étapes franchies (Martuccelli, 2002; Dubey, 2001) forment un écosystème qui est une source d'ancrage pour les acteurs du réseau. Comment ne pas interpréter l'utilisation récente de la « recherche appréciative » en évaluation (Coghlan *et al.*, 2003), qui initie sa démarche par le partage entre les acteurs de leurs réussites, comme une actualisation de cette quête d'un enveloppement émotif? Ces liaisons contrebalancent l'éloignement des systèmes abstraits auxquels les acteurs n'ont le plus souvent accès que dans leur réflexivité (Giddens, 1984 et 1991). Il arrive souvent qu'un programme, proposé par l'État ou des promoteurs universitaires, vienne de l'extérieur des milieux de pratique ou de vie. Il est d'abord vécu de l'intérieur, comme imposition ou proposition, par les acteurs des milieux d'implantation. Cette intériorisation peut demeurer très abstraite si les promoteurs ne sont pas connus, ni accessibles, et si le discours sur le programme ne se réfère pas à des expériences familières. La confiance dans l'action proposée ne pourra venir que si ce programme intérieur se nourrit et demeure cohérent avec le développement du programme extérieur, celui des pratiques et des résultats vécus. Cette conscience **hologrammatique**, où le programme est à la fois extérieur et intérieur au sujet, peut se vivre de manière positive ou négative dans un processus de programmation-réseau. Elle est vécue positivement si la programmation permet de repérer des points d'accès aux systèmes abstraits et ainsi d'anticiper le travail de liaison. Le promoteur d'un programme est avantagé dans son travail de mobilisation d'un réseau et de création d'un climat de confiance autour du programme s'il dispose au départ de tels points d'accès. Un promoteur, par exemple, qui a participé à l'élaboration d'une politique sociale dont découle le programme proposé, peut représenter pour les acteurs locaux une liaison concrète avec des acteurs lointains qui demeurent la plupart du temps des inconnus, comme les hauts fonctionnaires ou d'autres technocrates (Bilodeau *et al.*, 2003). Par ailleurs, elle peut être ressentie négativement si elle demeure et se reproduit dans l'abstraction, la perspective du réseau à mobiliser prenant une telle ampleur et une telle complexité qu'elle devient paralysante et mène à un refus d'implication.

L'antidote de ce refus d'implication de certains acteurs est dans le recentrage de la programmation sur l'essence même de la coconstruction en réseau, qui est le consensus et la synergie. Il ne peut s'agir d'un **consensus** global et définitif mais bien d'un enlignement commun qui se nourrit d'étapes négociées qui sont significatives pour la majorité des acteurs. Ce sont de petits « faire sens ensemble » dans la trajectoire d'un plus grand « faire sens ensemble ». Et ce chemin doit être parsemé de résultats **synergétiques** qui permettent de voir clairement ce qu'apporte de plus le travail en association avec d'autres (Lasker *et al.* 2001). En même temps, l'appel à une mobilisation consensuelle et synergétique implique une mise à l'épreuve qui nécessite des ressources dont les acteurs, selon leur position sociale, ne sont pas tous également pourvus. Cet appel peut être vécu comme une injonction, une obligation à la participation, une nouvelle forme de domination, centrée sur la responsabilisation individuelle (Martuccelli, 2004) à laquelle les sujets peuvent réagir encore une fois par un refus, déguisé en participation rituelle ou en repli sur soi. Le rôle de l'évaluation dans l'égalisation des ressources de participation fait partie des défis qui ont traversé les modèles d'évaluation participative ces dernières années (Cousins et Whitmore, 1998 ; Ryan, 2004). Des progrès ont été accomplis, mais la côte reste difficile à monter, en particulier pour ce qui est de la participation des usagers, des bénéficiaires et des clients potentiels puisque le caractère abstrait, sinon invisible, du programme est plus accentué pour eux, le caractère concret étant la plupart du temps délimité à une relation de service. L'organisation-terrain des politiques sociales renforce ce phénomène puisqu'elle repose de plus en plus sur des stratégies gouvernementales nourries de sondages et de gains politiques immédiats. La croissance et la juxtaposition de services répond mieux à ce contexte stratégique que l'approfondissement de démarches programmées. Des approches semblables à celles de Kushner (2000), qui explorent la signification biographique des programmes et des services pour les sujets qui y sont exposés, pourraient aider à mieux saisir ce rapport concret-abstrait chez les participants et renouveler le genre de la programmation.

#### 4. POUR NE PAS BOUCLER TROP VITE

Nous avons montré au début de ce chapitre que les incursions de l'évaluation dans le processus de programmation étaient pertinentes mais fragiles, qu'elles étaient emportées dans le courant de la dynamique de facteurs de changement au parcours souvent imprévisible et incontrôlable. Nous avons vu ensuite que l'évaluation, comme discipline autonome et entière, a travaillé à renforcer la dynamique interne de la programmation pour y développer sa propre force et, à la limite, pratiquement fusionner avec elle.

Cet effort disciplinaire, par ailleurs, l'a conduite à une restructuration de l'idée de programme où il est reconnu que celui-ci ne peut survivre sans la mobilisation d'un réseau social de soutien. Nous avons fait l'hypothèse que cette nouvelle représentation du programme manifeste une adaptation de la programmation-évaluation à l'accélération, dans la modernité avancée, du processus de subjectivation, de création de sujets autonomes par rapport aux cadres sociaux prédéfinis. L'examen de cette hypothèse nous a forcé à redéfinir les articulations entre des représentations de prime abord divergentes du programme ; il a permis de souligner aussi les dangers et les défis nouveaux auxquels est confrontée l'évaluation. Cet examen n'était qu'une ouverture partielle sur un chantier, d'orientation sociologique, qu'il reste à questionner et consolider.

On peut déjà, toutefois, entrevoir des enlignements méthodologiques non pas pour le bouclage mais pour la reproduction réciproque (Barel, 1973) de la programmation et de l'évaluation, qui fait appel à l'élargissement du travail des concepteurs et des évaluateurs. Un premier enlignement, qui n'est pas nouveau, souligne que le travail en réseau est une forme de délocalisation multiple de la réflexivité et qu'il faut, par conséquent, lui créer un espace rhétorique où chacun peut trouver les moyens d'éprouver sa voix (McKie, 2003). Un second revient à situer le programme dans une argumentation à la fois large et profonde qui non seulement relie la science à l'expérience, les synthèses du passé et les concurrents actuels, mais scénarise les échecs autant que les succès pour éviter les ruptures de sens. Un troisième rappelle le potentiel mobilisateur de l'ancrage concret du programme dans les différents paliers de son écosystème, dont deux exemples à l'opposé l'un de l'autre sont l'identification des points d'accès aux systèmes abstraits et la célébration des produits synergétiques. Ces enlignements placent l'évaluateur dans des rôles multiples de médiateur, de pédagogue, de démarcheur, de synthétiseur et de prospectiviste. Devant de telles exigences, pour faire écho à l'humour de Mabry (2002), le rôle le plus sûr pour l'évaluateur postmoderne ne devrait-il pas être celui d'ermite ? Si oui, cela signifierait que l'évaluateur, lui aussi, va se cacher pour mourir. Il faut plutôt souhaiter le contraire.

## BIBLIOGRAPHIE

- ABMA, T.A. et R.E. STAKE (2001). «Stake's responsive evaluation: Core ideas and evolution», *New Directions for Evaluation*, vol. 92, p. 7-22.
- AIRASIAN, P.W. (1987). «Societal experimentation», dans G.F. Madaus, M.S. Scriven et D.L. Stufflebeam (dir.), *Evaluation Models – Viewpoints on Educational and Human Services Evaluation*, Boston, Kluwer-Nijhoff, p. 163-175.
- ALLARD, D. (1996). *De l'évaluation de programme au diagnostic socio-systémique: trajet épistémologique*, Thèse inédite, Montréal, Université de Montréal.

- ALLARD D. et M. FERRON (2000). *Évaluation du programme PAD-PRAT – Vie et reproduction d'un programme*, Montréal, ASSTSAS.
- ALLARD, D., A. BILODEAU et C. LEFEBVRE (2007). «Le travail du planificateur public en situation de partenariat», dans M.-J. Fleury, M. Tremblay et H. Nguyen (dir.), *Le système sociosanitaire au Québec – Gouvernance, régulation et participation*, Montréal, Gaëtan Morin, p. 479-494.
- ALLARD, D., A. BILODEAU et S. GENDRON (2008). «Figurative thinking and models: Tools for participatory evaluation», dans M. Potvin et D.M. McQueen (dir.), *Evaluation Practices in Health Promotion*, New York, Springer.
- BAREL, Y. (1973). *La reproduction sociale – Systèmes vivants, invariance et changement*, Paris, Anthropos.
- BAREL, Y. (1989). *Le paradoxe et le système – Essai sur le fantastique social*, Grenoble, Presses universitaires de Grenoble.
- BARNES, M., E. MATKA et H. SULLIVAN (2003) «Evidence, understanding and complexity – Evaluation in non-linear systems», *Evaluation*, vol. 9, p. 265-284.
- BERK, R.A. et P.H. ROSSI (1990). *Thinking About Program Evaluation*, Newbury Park, Sage.
- BERNOUX, J.F. (2004). *L'évaluation participative au service du développement social*, Paris, Dunod.
- BICKMAN, L. (1987). «The functions of program theory», *New Directions for Program Evaluation*, vol. 33, p. 5-18.
- BILODEAU, A., C. LEFEBVRE et D. ALLARD (2002). *Les priorités nationales de santé publique 1997-2002: une évaluation de l'actualisation de leurs principes directeurs – Le cas des programmes de prévention du VIH-SIDA chez les hommes gais de la région de Montréal-Centre*, Québec, Institut national de santé publique.
- BILODEAU, A., C. LEFEBVRE et D. ALLARD (2003). *Les priorités nationales de santé publique 1997-2002: une évaluation de l'actualisation de leurs principes directeurs – Le cas du déploiement du protocole sociojudiciaire en matière de violence conjugale de la région du Bas St-Laurent*, Québec, Institut national de santé publique.
- BINGHAM, R.D. et C.L. FELBINGER (1989). *Evaluation in Practice – A Methodological Approach*, New York, Longman.
- BLOOM, M. (1999). «Single-system evaluation», dans I. Swan et J. Lishman (dir.), *Evaluation and Social Work Practice*, Londres, Sage, p. 198-217.
- CAMPBELL, D.T. (1994). «Retrospective and prospective on program impact assessment», *Evaluation Practice*, vol. 15, p. 291-298.
- COGHLAN, A.T., H. PRESKILL et C.T. TZAVARAS (2003). «An overview of appreciative inquiry in evaluation», *New Directions for Evaluation*, vol. 100, p. 5-23.
- COUSINS, J.B. et al. (2004). «Using key component profiles for the evaluation of program implementation in intensive mental health case management», *Evaluation and Program Planning*, vol. 27, p. 1-23.

- COUSINS, J.B. et E. WHITMORE (1998). «Framing participatory evaluation», *New Directions for Evaluation*, vol. 80, p. 5-23.
- DAHLER-LARSEN, P. (2001). «From programme theory to constructivism – On tragic, magic and competing programmes», *Evaluation*, vol. 7, p. 331-349.
- DAVIS, P. (2005). «The limits of realist evaluation – Surfacing and exploring assumptions in assessing the best value performance regime», *Evaluation*, vol. 11, p. 275-295.
- DEVINE, J.A., J.D. WRIGHT et L.M. JOYNER (1994). «Issues in implementing a randomized experiment in a field setting», *New Directions for Program Evaluation*, vol. 63, p. 27-40.
- DUBEY, G. (2001). *Le lien social à l'ère du virtuel*, Paris, Presses universitaires de France.
- ENRIQUEZ, E. (2007). «La marchandisation généralisée: du citoyen au consommateur», *Cahiers de recherche sociologique*, vol. 43, p. 23-35.
- FREIDMAN, V.J. (2004). «Designed blindness: An action science perspective on program theory evaluation», *American Journal of Evaluation*, vol. 22, p. 161-181.
- GIDDENS, A. (1984). *The Constitution of Society*, Berkeley, University of California Press.
- GIDDENS, A. (1991). *Modernity and Self-identity – Self and Society in the Late Modern Age*, Stanford, Stanford University Press.
- GIDDENS, A. (1993). «Une théorie critique de la modernité avancée», dans M. Audet et H. Bouchikhi (dir.), *Structuration du social et modernité avancée – Autour des travaux d'Anthony Giddens*, Québec, Presses de l'Université Laval, p. 29-53.
- GREGORY, A. (2001). «Problematizing participation – A critical review of approaches in evaluation theory», *Evaluation*, vol. 6, p. 179-199.
- GUBA, E.G. et Y.S. LINCOLN (1989). *Fourth Generation Evaluation*, Newbury Park, Sage.
- GUIENNE, V. (2007). «Savoir se vendre: qualité sociale et disqualification sociale», *Cahiers de recherche sociologique*, vol. 43, p. 7-20.
- JENLINK, P.M. (1994). «Using evaluation to understand the learning architecture of an organization», *Evaluation and Program Planning*, vol. 17, n° 3, p. 315-325.
- KAUFMANN, J.C. (2001). *Ego – Pour une sociologie de l'individu*, Paris, Nathan.
- KAUFMANN, J.C. (2004). *L'invention de soi – Une théorie de l'identité*, Paris, Armand Collin.
- KOIVISTO, J., K. VATAJA et R. SEPPÄNEN-JÄRVELÄ (2006). *Steps Towards Relational Evaluation of Organizational Development*, Londres, European Evaluation Society Conference.
- KUSHNER, S. (2000). *Personalizing Evaluation*, Londres, Sage.
- LALLEMENT, M. (2007). *Le travail – Une sociologie contemporaine*, Mesnil-sur-l'Estray, Gallimard.

- LAM, J.A., H.S. WILSON et J.F. JEKEL (1994). «I prayed real hard, so I know I'll get in»: Living with randomization», *New Directions for Program Evaluation*, vol. 63, p. 55-66.
- LASKER, R.D., E.S. WEISS et R. MILLER (2001). «Partnership synergy: A practical framework for studying and strengthening the collaborative advantage», *The Milbank Quarterly*, vol. 79, p. 179-205.
- LEEUW, F.L. (2003). «Reconstructing program theories: Methods available and problems to be solved», *American Journal of Evaluation*, vol. 24, p. 5-20.
- LIPSEY, M.W. (1997). «What can you build with thousands of bricks? Musing on the cumulation of knowledge in program evaluation», *New Directions for Evaluation*, vol. 76, p. 7-19.
- MABRY, L. (2002). «Postmodern evaluation – or not?», *American Journal of Evaluation*, vol. 23, p. 141-157.
- MAILOT, F. (1982). *La nouvelle cybernétique – Essai d'épistémologie des systèmes dynamiques*, Bruxelles, Chabassol.
- MARK, M.M., G.T. HENRY et G. JULNES (2000). *Evaluation – An Integrated Framework for Understanding, Guiding and Improving Policies and Programs*, San Francisco, Jossey-Bass
- MARTUCELLI, D. (1999). *Sociologies de la modernité*, Mesnil-sur-l'Estrée, Gallimard.
- MARTUCELLI, D. (2002). *Grammaires de l'individu*, Mesnil-sur-l'Estrée, Gallimard.
- MARTUCELLI, D. (2004) «Figures de la domination», *Revue française de sociologie*, vol. 45, n° 3, p. 469-497.
- MASON, P. et M. BARNES (2007). «Constructing theories of change – Methods and sources», *Evaluation*, vol. 13, p. 151-170.
- MASON, R.O. et I.I. MITROFF (1981). *Challenging Strategic Planning Assumptions – Theory, Cases and Techniques*, New York, Wiley-Interscience.
- MCCULLOUGH, C. (1999). *Les oiseaux se cachent pour mourir*, Paris, J'ai lu.
- MCKIE, L. (2003). «Rhetorical spaces: Participation and pragmatism in the evaluation of community health work», *Evaluation*, vol. 9, p. 307-324.
- MCLAUGHLIN, J.A. et G.B. JORDAN (1999). «Logic models: A tool for telling your program's performance story», *Evaluation and Program Planning*, vol. 22, p. 65-72.
- MCLAUGHLIN, M.W. (1987). «Implementation Realities and Evaluation Design», *Evaluation Studies Review Annual*, vol. 12, p. 73-97.
- MOHR, L.B. (1999). «The qualitative method of impact analysis», *American Journal of Evaluation*, vol. 20, p. 69-84.
- MORELL, J.A. (2005). «Why are there unintended consequences of program action, and what are the implications for doing evaluation?», *American Journal of Evaluation*, vol. 26, p. 444-463.
- MORIN, E. (1977). *La Méthode 1. La nature de la nature*, Paris, Seuil.

- MORIN, E. (1986). *La Méthode 3. La connaissance de la connaissance*, Paris, Seuil.
- MORIN, E. (2001). *La Méthode 5. L'humanité de l'humanité*, Paris, Seuil.
- OLIVER, S., A. HARDEN et R. REES (2005). « An emerging framework for including different types of evidence in systematic reviews for public policy », *Evaluation*, vol. 11, p. 428-446.
- ORR, L. (1999). *Social Experiments – Evaluating Public Programs with Experimental Methods*, Thousand Oaks, Sage.
- PAWSON, R. (2002). « Evidence-based Policy: The Promise of “Realist Synthesis” », *Evaluation*, vol. 8, p. 340-358.
- PAWSON, R. et N. TILLEY (1997). *Realistic Evaluation*, Londres, Sage.
- PIAGET, J. (1967). *Biologie et connaissance*, Paris, Gallimard.
- ROBERTS-GRAY, C. et M.A. SCHEIRER (1988). « Checking the congruence between a program and its organizational Environment », *New Directions for Program Evaluation*, vol. 40, p. 63-82.
- ROHRBACH, L.A. et al. (2006) « Type II translation: Transporting prevention interventions from research to real-world settings », *Evaluation and the Health Professions*, vol. 29, p. 302-333.
- ROSSI, P.H. (1997). « Advances in quantitative evaluation, 1987-1996 », *New Directions for Evaluation*, vol. 76, p. 57-68.
- ROTHMAN, J. (2002). « Action evaluation: Integrating evaluation into the intervention process », <[www.aepro.org/inprint/conference/rothman.html](http://www.aepro.org/inprint/conference/rothman.html)>.
- RYAN, K.E. (2004). « Serving public interests in educational accountability: Alternative approaches to democratic evaluation », *American Journal of Evaluation*, vol. 25, p. 443-460.
- SCHEIRER, M.A. (2005). « Is sustainability possible? A review and commentary of empirical studies of program sustainability », *American Journal of Evaluation*, vol. 26, p. 320-347.
- SMITH, M. F. (1989). *Evaluability Assessment – A Practical Approach*, Boston, Kluwer.
- STEVENS, C. et M. DIAL (1994). « What constitutes misuse? », *New Directions for Program Evaluation*, vol. 64, p. 3-15.
- ST-PIERRE, T.L. et D.L. KALTREIDER (2004). « Tales of refusal, adoption and maintenance: Evidence-based substance abuse prevention via school-extension collaboration », *American Journal of Evaluation*, vol. 25, p. 479-491.
- STRAW, R.B. et J.M. HERRELL (2002). « A framework for understanding and improving multisites evaluation », *New Directions for Evaluation*, vol. 94, p. 5-16.
- SUCHMAN, E.A. (1967). *Evaluative Research – Principles and Practice in Public Service and Social Action Programs*, New York, Russell Sage Foundation.
- THUDEROZ, C. (2000). *Négociations – Essai de sociologie du lien social*, Paris, Presses universitaires de France.

- TILLEY, N. (1996). « Demonstration, exemplification, duplication and replication in evaluation research », *Evaluation*, vol. 2, p. 35-50.
- TOURAINÉ, A. (2005). *Un nouveau paradigme – Pour comprendre le monde d'aujourd'hui*, Paris, Fayard.
- WANDERSMAN, A. *et al.* (2000). *Getting to Outcomes: Methods and Tools for Planning, Self-evaluation and Accountability*, South Carolina, National Center for the Advancement of Prevention.
- WEISS, C.H. (1997). « Theory-based evaluation: Past, present and future », *New Directions for Evaluation*, vol. 76, p. 41-56.
- WHOLEY, J.H. (1979). *Evaluation: Promise and Performance*, Washington, The Urban Institute.
- WHOLEY, J.H. (1994). « Assessing the feasibility and likely usefulness of evaluation », dans J.S. Wholey, H.P. Hatry et K.E. Newcomer (dir.), *Handbook of Practical Program Evaluation*, San Francisco, Jossey-Bass, p. 15-39.
- YIN, R.K. (1997). « Case study evaluation: A decade of progress ? » *New Directions for Evaluation*, vol. 76, p. 69-78.



---

## Notices biographiques

**Marc Alain**, Ph.D. en criminologie, est professeur régulier au Département de psychoéducation de l'Université du Québec à Trois-Rivières (UQTR) depuis 2005 et chercheur rattaché au Centre international de criminologie comparée (CICC) depuis 2002. Il a, entre 2000 et 2005, assumé la direction scientifique du Centre d'intégration et de diffusion de la recherche en activité policière à l'École nationale de police du Québec, institution avec laquelle il poursuit des collaborations en matière de recherche. Depuis 2005, il coopère régulièrement avec divers milieux d'intervention au Québec, essentiellement à titre d'évaluateur où son expertise est mise à profit dans l'élaboration et l'implantation de programmes d'intervention destinés à des clientèles multiples et diversifiées. À titre de responsable de ce secteur au Département de psychoéducation de l'UQTR, il enseigne le cours d'élaboration et évaluation de programmes.

**Denis Allard**, Ph.D. en sociologie, est spécialiste en évaluation de programme. Son doctorat en sociologie (1996) a porté sur l'épistémologie de l'évaluation. La majorité de sa carrière de chercheur (1980-2007) s'est déroulée dans des organismes québécois de santé publique, où il a travaillé à l'évaluation et au monitoring de programmes de prévention en santé au travail, santé des enfants et des jeunes, maladies infectieuses et jeu pathologique. Ses publications les plus récentes (2005-2008) ont porté sur les démarches participatives en évaluation et planification, de même que sur l'utilité méthodologique de la pensée figurative (schémas, dessins, métaphores) dans de tels contextes.

**Valérie Caron**, B. Ed., est candidate à la maîtrise en psychoéducation à l'Université du Québec à Trois-Rivières. Son mémoire porte sur l'évaluation de l'inclusion des enfants présentant une déficience intellectuelle ou un trouble envahissant du développement dans les milieux de garde québécois. Elle est assistante de recherche pour la Chaire de recherche du Canada en intervention précoce de l'Université du Québec à Trois-Rivières.

**Marie-Marthe Cousineau**, Ph.D. en sociologie, est professeure titulaire à l'École de criminologie de l'Université de Montréal et chercheure associée à l'Institut de recherche sur les dépendances, au Centre international de criminologie comparée (CICC) et au Centre de recherche interdisciplinaire sur la violence familiale et la violence faite aux femmes (CRI-VIFF). Ses intérêts de recherche portent, pour une part, sur différentes formes de marginalité et de marginalisation vécues par les jeunes (gangs de rue, prostitution, délinquance, toxicomanie, jeu) et sur les façons d'y réagir, et, pour une autre part, sur les violences touchant les femmes et les réponses sociales à ces violences. Ses recherches emploient des méthodologies qualitatives, quantitatives et évaluatives.

**Sylvain Coutu**, Ph.D. en psychologie, est professeur au Département de psychoéducation et psychologie de l'Université du Québec en Outaouais depuis 1988. Il est chercheur coordonnateur de l'équipe de recherche QEMVIE (Qualité éducative des milieux de vie de l'enfant). Ses travaux de recherche actuels portent sur le développement socioémotionnel des enfants d'âge préscolaire et sur la prévention des problèmes d'adaptation sociale dans les services de garde.

**Danny Dessureault**, Ph.D. en psychologie, est psychoéducateur de formation et est titulaire d'un doctorat en psychologie de l'Université de Montréal. Après des études postdoctorales auprès du Groupe de recherche sur les inadaptations sociales de l'enfance à l'Université de Sherbrooke, il poursuit sa carrière à l'Université du Québec à Trois-Rivières en tant que

professeur régulier. Ses travaux portent sur la maltraitance de même que sur le phénomène de la récurrence et des signalements en protection de la jeunesse.

**Diane Dubeau**, Ph.D. en psychologie, est chercheure associée au GRAVE-ARDEC (Groupe de recherche et d'action sur la victimisation des enfants – Alliance de recherche pour le développement des enfants dans leur communauté) et à QEMVIE (Qualité éducative des milieux de vie de l'enfant). Depuis plus de cinq ans, elle assume la responsabilité scientifique de ProsPère, une équipe de chercheurs universitaires et institutionnels dont les travaux et les actions visent à valoriser l'engagement des pères auprès de leur enfant. Ces projets de recherche se regroupent sous deux niches principales soit la paternité et l'évaluation de programmes d'intervention. Elle possède une formation en psychologie du développement de l'enfant. Elle est titulaire d'une maîtrise en psychologie (spécialisation enfance) de l'Université du Québec à Trois-Rivières et d'un Ph.D. en psychologie de l'Université du Québec à Montréal. Depuis dix-sept ans, elle est professeure au Département de psychoéducation et de psychologie de l'Université du Québec en Outaouais.

**Isabelle F.-Dufour**, M. Serv. Soc., est candidate au doctorat en service social à l'Université Laval. Elle a réalisé un mémoire de maîtrise exploratoire portant sur la réinsertion sociale des personnes contrevenantes soumises à un emprisonnement avec sursis au Québec. Elle poursuit actuellement des études doctorales qui ont pour objectifs d'approfondir et d'enrichir les principaux constats issus de son mémoire. Elle participe également à une recherche qui vise à reconstruire les trajectoires de vie des personnes autochtones incarcérées au Québec. Elle s'intéresse principalement aux questions qui touchent la marginalité, la déviance, l'exclusion sociale et la disqualification sociale.

**Martin Goyette**, Ph.D. en service social, est professeur-chercheur à l'École nationale d'administration publique (ENAP, Université du Québec), où il enseigne l'évaluation de programmes. Il est chercheur au Centre de recherche et d'expertise en évaluation de l'ENAP, chercheur régulier de l'équipe Économie sociale, santé et bien-être (équipe FQRSC) et chercheur régulier de l'Institut universitaire – Centre jeunesse de Montréal. Ses travaux s'intéressent à l'évaluation de programmes sociaux, notamment à l'endroit des jeunes en difficulté, à l'insertion et à l'exclusion sociale des populations démunies et au rôle des réseaux sociaux dans les étapes du cycle de vie.

**Sylvie Hamel**, Ph.D. en psychologie, est professeure au Département de psychoéducation de l'Université du Québec à Trois-Rivières. Elle est également chercheure régulière au Centre international de criminologie comparée (CICC) regroupement UQTR, chercheure régulière au Groupe de recherche et d'intervention sur l'adaptation psychosociale et scolaire (GRIAPS) et chercheure associée à l'Institut de recherche pour le développement social des jeunes (IRDS). Ses champs d'intérêt touchent les adolescents, notamment membres de gang; les processus d'affiliation, de désaffiliation, d'exclusion, de marginalisation; les trajectoires et les déterminants psychosociaux s'attachant à la violence et à la délinquance, de même qu'à l'intervention pouvant s'y adresser; l'intervention familiale et communautaire.

**Jacques Joly**, Ph.D. en psychologie, est professeur au Département de psychoéducation de l'Université de Sherbrooke où il assume l'enseignement de l'évaluation de programme et chercheur au Groupe de recherche sur les inadaptations sociales de l'enfance. Depuis l'obtention de son doctorat en 1981, M. Joly a agi à titre de consultant en recherche et associé de deux firmes de recherche du Québec jusqu'en 2000. Il s'intéresse particulièrement à l'implantation des programmes et ses travaux portent surtout sur la fidélité d'implantation.

**Carl Lacharité**, Ph.D. en psychologie clinique, est engagé depuis plus de vingt ans dans l'exploration et le développement de pratiques cliniques et sociales auprès d'enfants, d'adolescents et d'adultes vivant dans des contextes de négligence, de violence et d'exclusion. Il est psychologue pour enfants et pratique la thérapie familiale depuis 1984. Il est professeur titulaire au Département de psychologie de l'Université du Québec à Trois-Rivières. Il dirige le Centre d'études interdisciplinaires sur le développement de l'enfant et la famille en plus de collaborer aux travaux d'équipes de recherche en France, en Angleterre et au Brésil. Ses travaux ont été publiés dans de multiples revues scientifiques et professionnelles nationales et internationales.

**Line Massé**, Ph.D. en psychologie de l'éducation, est professeure agrégée au Département de psychoéducation de l'Université du Québec à Trois-Rivières. Elle est aussi directrice du Groupe de recherche et d'intervention sur l'adaptation psychosociale et scolaire (GRIAPS) et membre du Centre de recherche et d'intervention sur la réussite scolaire (CRIRES). Ses recherches portent principalement sur l'éducation des élèves doués et sur l'intervention multidimensionnelle auprès des élèves présentant des troubles du comportement, en particulier le trouble du déficit d'attention/hyperactivité et le trouble oppositionnel avec provocation.

**Robert Pauzé**, Ph.D. en psychologie, est psychologue clinicien et spécialisé dans le domaine de la thérapie familiale. Après l'obtention de son doctorat, M. Pauzé a été chargé de cours dans le domaine de la psychologie de l'alimentation au Département de nutrition de l'Université de Montréal et a été psychologue à l'Hôpital Sainte-Justine pendant une dizaine d'années. Il est professeur titulaire et directeur du Département de psychoéducation de l'Université de Sherbrooke. M. Pauzé est également membre du Groupe de recherche sur les inadaptations sociales de l'enfance (GRISE). Ses intérêts en recherche portent principalement sur les caractéristiques des jeunes et des familles desservis par les centres jeunesse du Québec, l'élaboration de programmes et les dysfonctions alimentaires à l'adolescence.

**Pierre Potvin**, Ph.D. en psychopédagogie, est psychoéducateur de profession et titulaire d'un doctorat en psychopédagogie de l'Université Laval. Pierre Potvin est professeur titulaire et associé au Département de psychoéducation de l'Université du Québec à Trois-Rivières (UQTR). Chercheur retraité au Groupe de recherche et d'intervention en adaptation psychosociale et scolaire (GRIAPS) et au Centre de recherche et d'intervention sur la réussite scolaire (CRIRES) de l'UQTR. Il est consultant au Centre de transfert pour la réussite éducative du Québec (CTREQ).

**Marie-Christine Saint-Jacques**, Ph.D. en sciences humaines appliquées, travailleuse sociale, est professeure agrégée à l'École de service social de l'Université Laval. Elle assume la direction scientifique du Centre de recherche sur l'adaptation des jeunes et des familles à risque de même que celle de l'équipe en institut JEFAR. Ses travaux portent sur le développement et l'évaluation des interventions sociales destinées aux jeunes et aux familles en difficulté. Elle mène aussi des recherches sur la thématique des transitions familiales entourant la séparation des parents, notamment, la recomposition familiale.

**Luc Touchette**, Ph.D. en éducation, est professeur au Département de psychoéducation de l'Université de Sherbrooke et chercheur associé au Groupe de recherche sur les inadaptations sociales de l'enfance. Ces travaux de recherche portent sur l'intervention de crise, les processus cliniques d'intervention auprès des jeunes et des familles et l'utilisation d'instruments d'évaluation standardisés à des fins de planification et d'évaluation d'intervention.

**Daniel Turcotte**, Ph.D. en éducation de l'Université du Québec à Chicoutimi, est professeur titulaire à l'École de service social de l'Université Laval. Il est membre du Centre de recherche sur l'adaptation des jeunes et des familles à risque (JEFAR), et directeur des programmes de deuxième et

troisième cycles. Ses travaux portent principalement sur l'évaluation des programmes et des services sociaux s'adressant aux enfants et aux familles en difficulté ainsi que sur la violence familiale.

**Geneviève Turcotte**, M. en sociologie, est chercheure à l'Institut de recherche pour le développement social des jeunes (IRDS) et au Centre jeunesse de Montréal-Institut universitaire. Elle est impliquée depuis près de 15 ans au sein du GRAVE-ARDEC et de son équipe ProsPère, où elle a contribué au développement et à l'évaluation de plusieurs projets de soutien à l'engagement paternel auprès de populations vulnérables de la région de Montréal.

**Martine Vézina**, M. Serv. Soc., a reçu un diplôme de maîtrise en service social de l'Université de Sherbrooke, incluant un cheminement en recherche évaluative. Son mémoire portait sur l'évolution des pratiques des organismes communautaires, dans le contexte de la réorganisation du réseau de la santé et des services sociaux. Elle fut coordonnatrice du projet Jeunesse et gangs de rue pendant trois ans pour l'Institut de recherche pour le développement social des jeunes à titre de professionnelle.



*Timbrés et affranchis*, 2005,  
techniques mixtes sur toile

MIREILLE BOURQUE est une artiste audacieuse qui expérimente, de toile en toile, les médiums les plus diversifiés. Pour elle, le processus de création est semblable à une méditation. Elle est attentive aux émotions qui émergent en regard de ce qui apparaît sur la toile. Il s'agit, la plupart du temps, de personnages car ce thème favorise le rapprochement et lui permet d'instaurer une forme de dialogue. Aussi, depuis quelque temps, des visages apparaissent en premier plan et brisent l'anonymat des personnages : ils ont des choses à dire ! Au-delà de sa démarche artistique, l'implication de Mireille Bourque au sein du programme *Vincent et moi* témoigne de sa volonté de jouer un rôle actif dans la démythification de la maladie mentale.

*Vincent et moi* est un programme d'accompagnement en soutien aux artistes atteints de maladie mentale mis sur pied en mai 2001 au Centre hospitalier Robert-Giffard – Institut universitaire en santé mentale. *Vincent et moi* innove en mettant au premier plan la valeur artistique d'œuvres réalisées par des personnes qui, au-delà de la maladie, s'investissent à l'égal de tout autre artiste dans un processus de création en arts visuels. La principale mission du programme étant de faire connaître et reconnaître leur contribution artistique et culturelle, on favorise la diffusion de leurs œuvres, crée des événements et expose leurs créations.

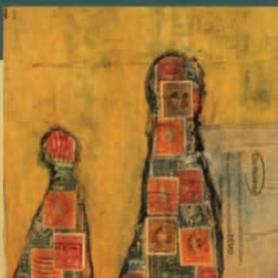
L'exposition annuelle du programme se déroule au centre hospitalier chaque automne, en association avec les Journées de la culture. Ce moment privilégié permet aux artistes de connaître la consécration de leur démarche en arts visuels et favorise la réappropriation de leur dignité grâce, entre autres, à la reconnaissance de leur statut d'artiste. L'exposition se veut un espace de rencontre contribuant à l'abolition des mythes et des préjugés, invitant la communauté à modifier le regard qu'elle porte sur les artistes souffrant de maladie mentale. En complément, tout au long de l'année, le programme présente des expositions thématiques itinérantes dans des lieux culturels et institutionnels, lesquelles permettent à un large public d'être touché par des œuvres rarement accessibles.

*Vincent et moi*, c'est aussi une collection unique ! Dons des artistes au programme, les œuvres de la collection forment un patrimoine artistique exceptionnel à l'exemple des grandes collections européennes. Constituée exclusivement d'œuvres originales conservées et mises en valeur selon les normes muséales, la collection s'enrichit au fil des années.

Galerie virtuelle : [www.rgiffard.qc.ca](http://www.rgiffard.qc.ca)

Courriel : [vincentetmoi@ssss.gouv.qc.ca](mailto:vincentetmoi@ssss.gouv.qc.ca)

Téléphone : 418-663-5321, poste 6440



En une série d'étapes liées les unes aux autres par un modèle intégrateur, les auteurs proposent un regard sur les programmes d'intervention mis en place dans les institutions ayant le mandat d'intervenir sur le plan psychosocial.

De la mesure des besoins à celle des effets à moyen et à long termes, en passant par le monitoring de l'implantation d'un programme, chacun des auteurs présente une façon de réaliser l'une ou l'autre des étapes de ce processus tout en s'appuyant sur des exemples concrets et des mises à l'essai réelles. Les étudiants et évaluateurs en devenir pourront ainsi construire leur propre cheminement en élaboration et évaluation de programmes d'intervention.

MARC ALAIN, Ph.D., est professeur régulier au Département de psychoéducation de l'UQTR depuis 2005, où il enseigne et pratique, en tant que chercheur et évaluateur, l'élaboration et l'évaluation de programmes.

DANNY DESSUREAULT, Ph.D., est professeur régulier au Département de psychoéducation de l'UQTR, où il travaille notamment sur le phénomène des resignements en protection de la jeunesse.

*Ont collaboré à cet ouvrage*

Marc Alain • Denis Allard • Valérie Caron • Marie-Marthe Cousineau • Sylvain Coutu  
Danny Dessureault • Diane Dubeau • Isabelle F.-Dufour • Martin Goyette • Sylvie Hamel  
Jacques Joly • Carl Lacharité • Line Massé • Robert Pauzé • Pierre Potvin • Marie-Christine  
Saint-Jacques • Luc Touchette • Daniel Turcotte • Geneviève Turcotte • Martine Vézina

[www.puq.ca](http://www.puq.ca)



9 782760 523463

ISBN 978-2-7605-2346-3