

## ÉVALUER L'EFFICACITÉ D'UN PROGRAMME: UNE QUESTION DE RÉFÉRENTS?

Marie Gervais, PhD  
Faculté de médecine  
Université Laval, Ste-Foy, Québec

Jacques Plante, PhD  
Faculté des Sciences de l'Éducation  
Université Laval, Ste-Foy, Québec

Chantal Jeanrie, PhD  
Faculté des Sciences de l'Éducation  
Université Laval, Ste-Foy, Québec

**Résumé:** Une étude conduite auprès de 73 acteurs provenant de deux organisations du domaine de la santé de la région de Québec a permis: 1) de vérifier si différents groupes d'acteurs associés à un même programme (décideurs, intervenants, usagers) partagent une même conception de l'efficacité d'un programme; et 2) de vérifier dans quelle mesure une conception de l'efficacité d'un programme se trouve influencée par des différents facteurs temporels, organisationnels et contextuels. Une approche non traditionnelle a été utilisée. Les acteurs de l'étude ont jugé l'efficacité de 20 programmes fictifs sur une échelle en quatre points. Les analyses de variance furent les principales analyses utilisées. Les éléments de convergence et de divergence se dégageant des évaluations produites par les différents groupes d'acteurs sont examinés et les facteurs explicatifs associés sont explorés.

**Abstract:** The main objectives of this study carried out with 73 people from two organizations in the health field in the Quebec City region are: 1) to verify if several groups of actors associated with the same program (decision makers, participants, users) have identical concepts of the effectiveness of a program; and 2) to verify to what extent the concept of program effectiveness is influenced by temporal, organizational and contextual variables. A non-traditional approach was used. Participants in the study assessed the effectiveness of 20 fictional programs on a four-point scale. Analyses of variance were the main forms of analysis. The elements of convergence and divergence in the evaluations produced by the different groups of participants are presented. Explicative factors are explored.

L'abondance de la documentation associée à l'efficacité met clairement en évidence la popularité et l'importance accordées à ce phénomène dans l'étude des programmes et des organisations canadiennes et américaines. Toutefois, malgré sa présence dans la majorité des modèles d'évaluation recensés, l'efficacité demeure une notion ambiguë se prêtant à multiples définitions selon la perspective théorique adoptée, les acteurs en présence et les opportunités et les contraintes caractérisant le contexte d'évaluation. En fait, il existe presque autant de façons de définir, mesurer et évaluer l'efficacité d'un programme qu'il y a d'études sur le sujet. Effectivement, même si traditionnellement l'efficacité est définie selon la capacité d'un programme à atteindre un ensemble prédéterminé d'objectifs, les aspects fonctionnaliste, systémique et humain du programme sont de plus en plus considérés lors d'une opération d'évaluation de l'efficacité d'un programme. Il devient ainsi souhaitable de tenir compte des acteurs en présence et des systèmes en place tout en intégrant à la fois les aspects politiques et techniques de l'évaluation. Mais qu'en est-il exactement? La notion d'objectifs à atteindre est-elle à ce point centrale quand il est question d'efficacité? Que cherche-t-on à évaluer lorsqu'il est question d'évaluation de l'efficacité d'un programme et que mesure-t-on véritablement? Et comment peut-on s'assurer que ce qui est mesuré est effectivement l'efficacité d'un programme? L'efficacité est-elle à ce point une notion malléable et dynamique? Comment en arriver à prendre une mesure valide et fiable de l'efficacité d'un programme? Une importante problématique liée à la définition de l'efficacité d'un programme et subséquemment à sa mesure et à son évaluation transparaît clairement des écrits sur le sujet (Gervais, 1998). La présente étude cherchera par conséquent à explorer les problèmes conceptuels et de mesure associés à une opération d'évaluation de l'efficacité d'un programme de façon à pouvoir y apporter un éclairage nouveau et constructif.

## LES CARACTÉRISTIQUES DE L'EFFICACITÉ

Quatre caractéristiques associées à l'efficacité d'un programme font actuellement consensus dans les écrits. Elles mettent en lumière la difficulté qu'ont les praticiens et les chercheurs dans le domaine à cerner correctement la réalité et la globalité du phénomène d'efficacité. Elles permettent également de mieux comprendre les problèmes de mesure qui en découlent et de mieux saisir les répercussions qu'elles entraînent sur le processus même d'élaboration d'un jugement d'efficacité. Il est pertinent ici de les présenter.

## L'efficacité en tant que construit multidimensionnel

Les écrits s'entendent pour considérer l'efficacité comme un phénomène abstrait, n'ayant aucune réalité objective et ne pouvant être directement observé et mesuré (Cameron, 1981a; Quinn & Rohrbaugh, 1983; Steers, 1975). L'efficacité doit être traitée comme un construit élaboré à partir de concepts multiples, ce qui conduit à la reconnaissance de la nature multidimensionnelle de l'efficacité. Toutefois des difficultés associées notamment au choix des concepts à intégrer dans le construit, à la spécification des liens unissant les concepts et à la précision de l'importance de chaque concept face au construit global contribuent à mettre en lumière la complexité du construit et à en rendre plus instable sa mesure. En effet, comment en arriver à circonscrire l'efficacité en tant que phénomène non tangible? Comment également en arriver à délimiter ses frontières, considérant que chaque approche théorique y accorde un sens différent selon un choix dissemblable de concepts à considérer? D'ailleurs, Cameron (1981a, 1984) ainsi que Quinn et Rohrbaugh (1983) rappellent que l'imprécision des frontières de la notion d'efficacité entraîne une part inhérente d'erreur dans toute mesure de l'efficacité, chaque mesure de l'efficacité devenant une approximation, une mesure dérivée imparfaite de l'efficacité.

Cela amène un problème additionnel, comment opérationnaliser l'efficacité si une confusion conceptuelle existe de fait? Cameron (1980) mentionne par ailleurs que c'est d'après les indicateurs retenus que sont très souvent définies les limites du phénomène. Toutefois, une sélection appropriée d'indicateurs d'efficacité ne semble pas assurer pour autant la validité d'une évaluation de l'efficacité d'un programme. Peu d'information est en effet retrouvée dans la documentation quant à la façon dont s'effectue habituellement le choix des indicateurs d'efficacité. Certains éléments particulièrement problématiques doivent être soulevés, à savoir, d'où viennent les indicateurs utilisés? Comment ont-ils été développés? L'ont-ils été de façon inductive ou déductive? Représentent-ils les valeurs, les intérêts, les biais d'un groupe quelconque? Ont-ils été choisis par opportunisme, ou encore, selon une tradition d'évaluation en vigueur dans le milieu? Que dire finalement de cette tendance constatée dans les écrits à rechercher des indicateurs passe-partout et applicables à toute situation d'évaluation? Ou encore de cette autre tendance qui consiste pour certaines études à se fier sur un indicateur unique pour mesurer l'efficacité d'un programme? Cela contrevient à la conception multidimensionnelle de l'efficacité maintenant admise.

## L'efficacité en tant que phénomène subjectif

Que l'efficacité soit reconnue dans la documentation consultée comme appuyée, en partie du moins, sur les valeurs véhiculées par les acteurs en présence et sur leurs préférences quant aux objectifs à atteindre par le programme constitue, pour certains auteurs, un autre élément pouvant influencer l'évaluation de l'efficacité d'un programme. En effet, Cameron (1981a, 1981b, 1984, 1986) ainsi que Cameron et Whetten (1983a) affirment que l'efficacité est une notion fondamentalement subjective. Joubert (1991) ainsi que Moynagh (1993) mentionnent de plus que la formation des valeurs évolue dans le temps et dans l'espace. Ainsi ce qui est appelé "efficacité" se voit influencé par les différents enjeux existant dans le programme, enjeux qui sont particulièrement en lien avec les ressources, la légitimité, le pouvoir, le contrôle et la satisfaction. D'ailleurs pour Guba et Lincoln (1981, 1986), les acteurs peuvent avoir des intérêts concurrentiels et même divergents face à l'organisation et être ainsi motivés à diffuser des versions différentes de la réalité du programme. La mouvance sociopolitique de l'efficacité apparaît maintenant avec plus d'acuité.

Ce qui est mesuré, ce ne serait donc pas l'efficacité réelle d'un programme mais plutôt les représentations de l'efficacité que se font ceux qui évaluent le programme. Il s'agit plus spécifiquement de représentations partiales de l'efficacité d'un programme, représentations qui s'avèrent fondées sur l'expérience personnelle, sur les valeurs et sur les intérêts éprouvés face au programme, sur le rôle et la position tenus par rapport au programme de même que sur les enjeux conjoncturels en présence. Cela rejoint la position de Lewin et Minton (1986) pour qui l'efficacité doit être déterminée en terme relatifs. Seashore (1983) avance également que toute mesure de l'efficacité d'un programme doit être vue comme une représentation partielle et transitoire du construit qu'elle représente. On est loin maintenant d'une définition unique et stable de l'efficacité. Ce qui est d'ailleurs appuyé par Seashore (1983), pour qui les estimés d'efficacité s'avèrent toujours pluriels, potentiellement différents, mais égaux en termes de validité pour chaque acteur d'un programme. Moynagh (1993) ajoute de plus que la façon dont l'efficacité est perçue semble dépendre largement de qui s'y intéresse, chaque point de vue étant porteur de sa propre validité. Le caractère polysémique de l'efficacité est ici clairement mis en évidence. Ce qui fait dire à Cameron et Whetten (1983b) que l'efficacité ne doit pas être comprise et évaluée qu'à partir d'un seul point de vue.

## L'efficacité en tant que phénomène flexible, dynamique et transitoire

Les écrits reconnaissent à l'efficacité à la fois ses aspects flexibles, dynamiques et transitoires, ce qui constitue des éléments additionnels à prendre en considération lors d'une mesure de l'efficacité d'un programme. En effet, l'efficacité est présentement perçue comme un phénomène qui doit s'ajuster à une réalité toujours en mouvement. Les notions d'évolution et de changement sont implicitement actives dans toute évaluation de l'efficacité, ce qui permet de reconnaître la dimension temporelle du phénomène. Les travaux de Cameron (1980), Gibson, Ivancevich et Donnely (1991), Harrison (1987), Lewin et Minton (1986), Miner (1988), Schroder (1989) et de Steers (1984) présentent de fait l'efficacité comme étant modulée par les effets combinés d'influences endogènes et exogènes au programme. Par influences endogènes, il est question d'influences provenant du programme lui-même, c'est-à-dire des changements pouvant se produire en rapport à la mission et aux objectifs du programme, en rapport à sa culture et aux valeurs qui y sont prônées, en rapport à la structure, aux ressources, aux processus, aux activités et aux comportements de ses membres, ou encore en rapport à ses pratiques de gestion. Par influences exogènes, il faut mentionner les influences provenant de l'environnement externe au programme, donc des conditions économiques, sociales et politiques qui caractérisent cet environnement. L'évaluation de l'efficacité d'un programme s'avère par conséquent contextuelle. Son opérationnalisation devrait tenir compte des caractéristiques spécifiques à ses environnements interne et externe tout comme elle devrait tenir compte de la réalité complexe, interactive et organisée dans le temps qui s'en dégage. Moynagh (1993) résume d'ailleurs cette position en affirmant que l'efficacité est une notion qui se définit selon les situations et qui tient compte des contraintes et des opportunités propres à ces situations. Chaque définition opérationnelle de l'efficacité en vient, par conséquent, à posséder une utilité et une stabilité toutes relatives en lien avec la situation d'évaluation.

L'énoncé de ces différents problèmes conceptuels et de mesure liés à la notion d'efficacité dénote l'absence de consensus sur ce qui constituerait une mesure significative, suffisante, consensuelle et unique de l'efficacité d'un programme. La problématique entourant le choix du référent à utiliser lors d'une évaluation de l'efficacité d'un programme apparaît avec plus d'acuité et constitue un facteur explicatif important face à la grande variabilité des résultats de mesure d'efficacité obtenus jusqu'à présent.

Ainsi, malgré cet effort de convergence des écrits quant à la précision d'un cadre d'évaluation acceptable à l'évaluation de l'efficacité d'un programme, le débat sur la véritable nature de l'efficacité demeure. L'importance de l'acteur dans le processus d'évaluation est également mise en évidence et, avec cela, plusieurs questions demeurent. En effet, la présence de logiques différentes et concurrentes chez différents groupes d'acteurs associés au programme et oeuvrant dans le même espace spatio-temporel laisse-t-elle présager un choix différent face au référent à considérer lors de l'évaluation de l'efficacité d'un programme? Ce choix de référent, ou conception de l'efficacité, se trouve-t-il influencé par des facteurs temporels et contextuels? D'autres facteurs tels le rôle ou la position tenus par les acteurs face au programme peuvent-ils influencer réellement la composante de l'efficacité alors privilégiée? Le partage d'une même culture et la présence de valeurs et d'intérêts semblables chez les acteurs appartenant à un même groupe entraînent-ils une vision commune de l'efficacité chez ce groupe?

## LES OBJECTIFS DE L'ÉTUDE

Cette étude cherchera à cerner et à confronter les référents utilisés par différents types d'acteurs en situation d'évaluation de l'efficacité d'un programme. Elle cherchera à explorer le lien potentiel entre ces deux éléments, à savoir, le QUI (type d'acteurs) évalué à partir DE QUOI (référents) et leur influence potentielle sur les résultats d'évaluation. La notion de référent est considérée dans l'étude comme reflétant implicitement la conception de l'efficacité privilégiée par un acteur. Cette étude cherchera également à vérifier si différents facteurs contextuels viennent influencer l'utilisation de ces référents et modifier les évaluations d'efficacité produites.

Plus spécifiquement, la constance dans le temps d'évaluations produites par des acteurs de programmes, leur stabilité ainsi que leur uniformité seront examinées. Il s'agira alors de vérifier si, à des moments différents et confrontés à des résultats de mesure similaires, ces acteurs évaluent de la même façon l'efficacité d'un programme. L'influence du facteur temps sur les évaluations produites sera ainsi examinée. La concordance théorique et pratique des référents utilisés par ces acteurs pour conduire à terme une évaluation de l'efficacité d'un programme sera également explorée. Autrement dit, il sera vérifié s'il existe une différence entre le référent que les acteurs disent utiliser de façon préférentielle et le référent qu'ils utilisent réellement une fois placés en situation d'évaluation de l'ef-

ficacité d'un programme. Les référents privilégiés par certains acteurs d'un programme seront par la suite explorés. Cela permettra de vérifier si ces acteurs recourent effectivement à une conception multidimensionnelle de l'efficacité lors de leur jugement d'un programme. Il s'agira également d'explorer s'il existe des différences intragroupes et intergroupes par rapport aux référents utilisés, de préciser la nature de ces différences, s'il y a lieu, et de vérifier s'il existe un lien entre le rôle joué par un groupe d'acteurs face au programme et le référent que ce groupe utilise préférentiellement lors d'une évaluation.

## LE CADRE CONCEPTUEL

Le cadre conceptuel utilisé dans l'étude reconnaît la nature multidimensionnelle de l'efficacité et invite à une lecture plurielle de l'efficacité par le truchement d'une grille en cinq dimensions (Gervais, 1998). Ce cadre conceptuel a été construit selon une perspective intégrative à partir des composantes couvertes par les principales approches traditionnelles d'évaluation de l'efficacité d'un programme (approche centrée sur les objectifs, approche systémique, approche fonctionnelle, approche socio-politique) et sur les efforts faits par plusieurs auteurs pour intégrer les connaissances contemporaines sur l'efficacité. Il constitue donc un reflet et une synthèse des conceptions véhiculées historiquement sur l'efficacité (Gervais, 1998). La documentation organisationnelle a servi de toile de fond pour le développement de ce cadre conceptuel. Toutefois, dans le cadre de cette étude, le cadre conceptuel a été appliqué à l'échelle d'un programme. Celui-ci est en effet considéré comme un sous-système de l'organisation en ce sens que, tout comme l'organisation, il est construit à partir de règles, de ressources et de pratiques qui interagissent et se transforment afin de permettre la réalisation de la mission qui constitue sa raison d'être. Le programme mobilise, de plus, les acteurs de l'organisation. Par conséquent, il est plausible de considérer un programme comme enchâssé dans l'organisation dans laquelle il prend place, tout comme l'organisation se trouve elle-même enchâssée dans le système de santé canadien et que ce dernier est, à son tour, partie intégrante du système sociétal nord-américain.

Les dimensions du cadre conceptuel s'articulent autour des cinq principaux domaines d'activités d'un programme ou domaines de référence (Figure 1). Il s'agit de: 1) la dimension structurelle concernée par les ressources et la structure du programme; 2) la dimension

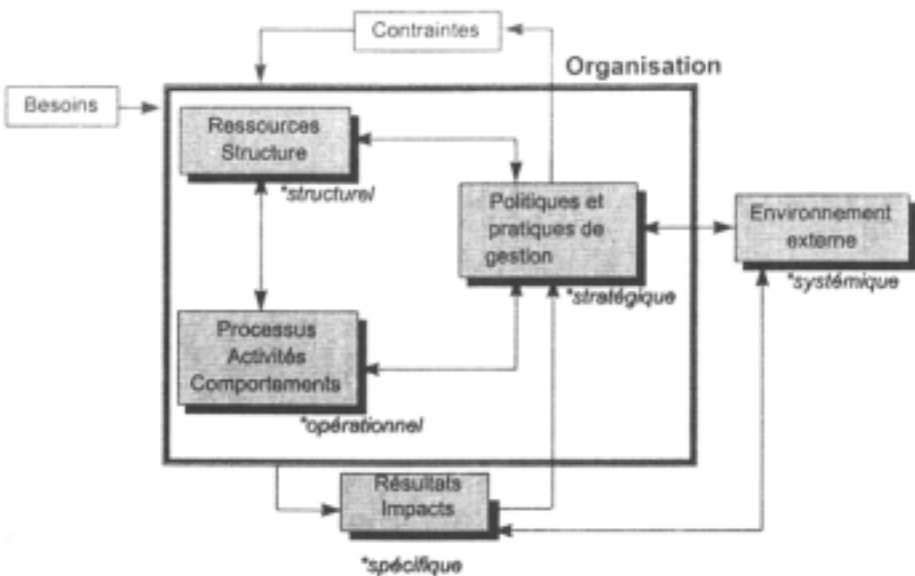
opérationnelle liée aux processus et aux activités du programme ainsi qu'aux comportements de ses membres; 3) la dimension systémique caractérisée par l'environnement externe du programme; 4) la dimension stratégique concernée par les politiques et les pratiques de gestion; et 5) la dimension spécifique préoccupée par les résultats et les impacts produits par le programme. Ces cinq dimensions de l'efficacité représentent différentes conceptions de l'efficacité d'un programme qui peuvent elles-mêmes être considérées comme des référents partiels à partir desquels il est possible d'évaluer l'efficacité d'un programme. Les domaines de référence du programme agiront donc dans l'étude comme une voie d'accès aux représentations de l'efficacité que se font différents acteurs d'un programme. Ils seront utilisés comme capteurs de ces représentations.

## LA MÉTHODOLOGIE

### Le choix de l'approche méthodologique

Une approche par simulation constitue la stratégie méthodologique sélectionnée (Gervais, 1996; Gervais, sous presse). Cette approche permet d'observer et d'analyser les comportements d'évaluation de

**Figure 1. Modèle conceptuel**





plusieurs groupes d'acteurs dans des situations d'évaluation concrètes mais simulées. Elle renseigne sur la façon dont ces acteurs intègrent les résultats de mesure présentés pour en venir à formuler un jugement par rapport à un objet d'évaluation, soit l'efficacité d'un programme.

### L'échantillon

Quatre-vingt acteurs ont été recrutés sur une base volontaire. Il s'agit de décideurs, d'intervenants et d'usagers des programmes administrés par deux organisations du domaine de la santé de la région de Québec. Les décideurs étaient gestionnaires d'un programme ou responsables de la coordination, de l'évaluation ou des prises de décision relatives à un programme de l'organisation. Les intervenants appartenaient à une profession productrice de soins ou de services (médecin, infirmier, ergothérapeute, psychologue, orthophoniste, travailleur social, éducateur physique, etc.). Les usagers étaient âgés de plus de 18 ans et étaient inscrits dans un des programmes de l'organisation ou l'avaient quitté depuis moins de deux ans. Les usagers de l'organisation A étaient atteints d'un trouble mental sévère et persistant (schizophrénie, psychose maniaco-dépressive), n'étaient pas déficients intellectuels et avaient un état mental stabilisé. Les usagers de l'organisation B présentaient une déficience motrice, neurologique ou visuelle. Les données provenant de 73 acteurs ont été considérées pour les analyses statistiques. Ce groupe représente 91,25% de l'échantillon initial. Le tableau suivant présente la composition finale des groupes d'acteurs selon le type et selon l'organisation.

**Tableau 1**  
**Répartition des acteurs selon le type et selon l'organisation**

Types	Organisations		
	A	B	Total
Décideurs	11	12	23
Intervenants	11	15	26
Usagers	11	13	24
<b>Total</b>	<b>33</b>	<b>40</b>	<b>73</b>

Les données socio-démographiques associées aux acteurs de l'étude sont présentées au Tableau 2. Outre les données habituelles, le nombre d'années de travail à l'intérieur de l'organisation se trouve mentionné pour les groupes de décideurs et d'intervenants tandis que le nombre d'années de fréquentation des programmes ainsi que la position relative face au marché du travail sont mentionnés pour les groupes d'utilisateurs.

La construction des instruments de mesure

Deux instruments de mesure furent construits aux fins de l'étude, soit un instrument contenant les programmes fictifs et un instru-

**Tableau 2**  
**Données socio-démographiques associées aux acteurs de l'étude**

Décideurs								
Org	Sexe		Âge		Formation		Travail <sup>1</sup>	
	F	M	Moyenne	é.-t.	Type	%	Moyenne	é.-t.
A	6	5	47.46	5.26	baccalauréat	46	25.00	5.31
B	6	6	46.92	4.66	baccalauréat	67	13.26	6.66
Intervenants								
Org	Sexe		Âge		Formation		Travail <sup>1</sup>	
	F	M	Moyenne	é.-t.	Type	%	Moyenne	é.-t.
A	8	3	41.82	8.57	baccalauréat	64	17.18	7.10
B	8	7	40.60	7.82	baccalauréat	93	11.20	6.99
Utilisateurs								
Org	Sexe		Âge		Présence <sup>2</sup>		Occupation <sup>3</sup>	
	F	M	Moyenne	é.-t.	Moyenne	é.-t.		%
A	5	6	39.55	10.82	5.02	6.19	actif	37
B	5	8	41.00	12.08	5.81	9.33	actif	62

1 Nombre d'années de travail dans l'organisation

2 Nombre d'années de fréquentation des programmes de l'organisation

3 Position face au marché du travail (étudiant, travailleur et bénévole)

ment permettant la hiérarchisation théorique des domaines. Ces deux instruments ont été élaborés à partir du même matériel de base, soit les indicateurs d'efficacité d'un programme. Le terme indicateur réfère dans l'étude à la définition proposée par Péron et Strohmenger (1985) pour qui un indicateur constitue une mesure dont la valeur est utilisée comme point de repère dans l'appréciation de l'état ou de l'évolution d'un phénomène non quantifiable. Les deux instruments ont été pré-expérimentés auprès de quatorze experts provenant de l'entreprise privée, du système de santé québécois, des milieux universitaire et gouvernemental ainsi que de six acteurs provenant des organisations cibles (décideurs, intervenants, usagers).

### La sélection des indicateurs d'efficacité

Le processus de sélection des indicateurs s'est effectué selon la technique consensuelle TRIAGE (Technique de Recherche d'Information par Animation d'un Groupe Expert: Plante & Côté, 1993). Cinquante-trois nouveaux acteurs ont été recrutés sur une base volontaire pour participer à TRIAGE. Ces acteurs étaient des décideurs, des intervenants et des usagers des programmes appartenant aux organisations cibles et répondant aux mêmes critères de sélection que ceux mentionnés précédemment. Six groupes homogènes d'acteurs ont été formés, soit un groupe de décideurs, un groupe d'intervenants et un groupe d'usagers pour chaque organisation.

L'application de TRIAGE s'est déroulée en deux phases distinctes. Une première phase de production individuelle a permis de recueillir 820 propositions d'indicateurs de la part des participants. Ceux-ci devaient en effet proposer individuellement des indicateurs pour chaque domaine d'activités d'un programme. Ces propositions ont par la suite été réparties de façon à créer, pour chaque groupe, une banque d'indicateurs d'efficacité propre à chaque domaine d'activité du programme. Ces banques d'indicateurs ont par la suite été utilisées lors de l'étape de production collective. Dans cette deuxième étape, les participants ont été rencontrés par groupe et ont étudié chaque banque selon la même procédure. Travaillant par approximations successives, chaque groupe a évolué vers un consensus et a sélectionné les indicateurs d'efficacité jugés les plus appropriés aux domaines de référence. Ces indicateurs ont par la suite reçu une deuxième validation par la consultation de 12 experts appartenant aux organisations cibles. Dix indicateurs d'efficacité ont ainsi été

retenus pour chaque organisation, soit deux indicateurs pour chaque domaine d'activités d'un programme (Tableau 3).

L'instrument contenant les programmes fictifs

Deux séries de 20 programmes fictifs ont été élaborées, soit une série pour l'organisation A et une série pour l'organisation B. Ces séries

**Tableau 3**  
**Indicateurs d'efficacité par domaine et par organisation**

<b>Organisation A</b>	
<b>Domaines</b>	<b>Indicateurs</b>
<b>Structurel</b>	Répartition de personnel en fonction des besoins Adéquation de la structure en fonction des besoins
<b>Opérationnel</b>	Capacité du personnel à répondre aux besoins de l'utilisateur Cohésion de l'équipe multidisciplinaire
<b>Systémique</b>	Complémentarité avec les autres programmes et services Partenariat avec les ressources de la région
<b>Stratégique</b>	Contenu du programme adapté aux besoins de l'utilisateur Capacité à résoudre des problèmes à l'interne
<b>Spécifique</b>	Satisfaction de l'utilisateur Degré d'atteinte des objectifs du plan d'intervention pour l'utilisateur
<b>Organisation B</b>	
<b>Domaines</b>	<b>Indicateurs</b>
<b>Structurel</b>	Répartition des ressources en fonction des besoins Compétence des ressources humaines
<b>Opérationnel</b>	Implication de l'utilisateur dans son plan d'intervention Pertinence des activités pour l'utilisateur
<b>Systémique</b>	Adaptation du programme à l'évolution des besoins Complémentarité du travail avec les ressources du milieu
<b>Stratégique</b>	Respect de la mission de l'établissement Actualisation d'un processus d'évaluation de programmes
<b>Spécifique</b>	Degré d'atteinte des objectifs Satisfaction de la clientèle

ont été construites selon les mêmes règles. Seuls ont différé les indicateurs représentant les domaines de référence du programme. L'information contenue dans un programme fictif est présentée sous plusieurs formes (Tableau 4). Un programme fictif possède comme principales caractéristiques: 1) un ensemble de 10 indicateurs d'efficacité présentés de façon aléatoire; 2) une échelle de mesure en trois points utilisée pour présenter les cotes obtenues par le programme en rapport à chaque indicateur d'efficacité; 3) une échelle de cotation en quatre points nécessaire à l'attribution du degré d'efficacité à attribuer au programme fictif; et 4) des instructions quant à la façon de répondre.

**Tableau 4**  
**Exemple de programme fictif**

		Cote obtenue par le programme		
		Faible	Moyenne	Élevée
I	Complémentarité avec les autres programmes et services		X	
N	Contenu du programme adapté aux besoins de l'utilisateur		X	
D	Cohésion de l'équipe multidisciplinaire		X	
I	Degré d'atteinte des objectifs du plan d'intervention pour l'utilisateur	X		
C	Adéquation de la structure en fonction des besoins			X
A	Capacité du personnel à répondre aux besoins des usagers		X	
T	Capacité à résoudre des problèmes à l'interne		X	
E	Satisfaction de l'utilisateur	X		
U	Répartition de personnel en fonction des besoins			X
R	Partenariat avec les ressources de la région		X	

En vous basant sur l'information présentée ci-dessus, veuillez juger si ce programme est "très inefficace", "inefficace", "efficace" ou "très efficace". Placez un "X" à l'endroit approprié sur cette échelle.

Programme très inefficace	Programme inefficace	Programme efficace	Programme très efficace
_____ :	_____ :	_____ :	_____ :

La force attribuée à chacun des domaines a fait l'objet d'une manipulation intentionnelle. Ainsi à l'intérieur de chaque programme fictif, deux domaines ont été mis intentionnellement en évidence: un domaine cote élevé et un autre cote faible. Les trois autres domaines ont été neutralisés par l'attribution de cotes moyennes. Ces cotes "élevée, moyenne et faible" ont été assignées selon une méthode aléatoire stratifiée.

Pour chaque organisation, il y a eu répartition des 20 vingt programmes fictifs en deux séries comparables devant être cotées par les acteurs en des temps différents. Également, deux programmes fictifs ont été répétés à l'intérieur de chaque série de façon à vérifier si les acteurs cotent de façon identique un ensemble de stimuli comparables. Par exemple, le programme 16 a été présenté deux fois à l'intérieur de la première série et une fois à l'intérieur de la deuxième série.

Outre le fait que les indicateurs d'efficacité sélectionnés l'ont été à partir des propositions émises par les acteurs mêmes des organisations cibles, la nature des consignes données aux participants a également contribué au réalisme des programmes fictifs. Ces consignes stipulaient que les programmes présentés poursuivent les mêmes objectifs et ont les mêmes activités, les mêmes ressources et la même clientèle que ceux retrouvés dans les programmes auxquels les participants sont associés, ce qui permettait de tenir compte implicitement des enjeux prévalant dans les programmes d'appartenance des participants.

Le choix de d'offrir aux participants de l'étude un ensemble de 20 programmes fictifs comprenant un nombre limité d'indicateurs ( $N = 10$ ) s'est trouvé justifié par le danger de créer une surcharge cognitive chez le participant s'il devait traiter simultanément un nombre élevé d'informations. Cela risquerait d'entraîner des changements non désirés dans les stratégies d'évaluation du participant, lesquels à leur tour rendraient les résultats obtenus non interprétables. Plusieurs vérifications avaient été effectuées en ce sens par l'auteur lors de la construction et la pré-expérimentation (Gervais, 1996; Gervais, sous presse). Il faut également noter que les instruments devront être accessibles à des groupes différents d'acteurs, dont à des usagers (personnes ayant des incapacités physiques, neurologiques, sensorielles ou psychiatriques), lesquels possèdent en général une connaissance très limitée de l'évaluation.

## L'instrument associé à la hiérarchisation théorique des domaines

Cet instrument devait permettre aux participants d'ordonner les différents regroupements d'indicateurs selon le degré d'importance à leur accorder lors d'une opération d'évaluation de l'efficacité d'un programme. Il contient donc comme principales composantes: 1) dix indicateurs d'efficacité présentés deux par deux selon leur domaine d'origine sans que ceux-ci ne soient identifiés; 2) une échelle de cotation en cinq points; et 3) des directives quant à la façon de répondre.

## La procédure de collecte des données

Quatre-vingt quatorze pourcent des acteurs contactés par téléphone ont accepté de participer à l'étude. La collecte des données s'est déroulée en trois phases distinctes et de façon simultanée dans les deux organisations. Les acteurs ont reçu par le courrier les documents spécifiques à chacune des étapes. Les acteurs ont d'abord reçu une lettre descriptive de l'étude (objectifs et modalités de participation), la formule de consentement concernant leur participation à l'étude et une première série de programmes fictifs à évaluer. Les acteurs ont reçu par la suite la deuxième série de programmes fictifs à évaluer. Le document associé à la hiérarchisation théorique des domaines et le formulaire visant à recueillir les données socio-démographiques ont été joints au troisième envoi.

## Le choix des analyses statistiques

Un modèle statistique univarié a été principalement retenu (analyse de variance univariée selon un devis expérimental à unités divisées, *split-plot*). Ce modèle statistique a d'abord été utilisé pour vérifier la constance dans le temps des évaluations, leur l'homogénéité chez un même groupe d'acteurs et pour distinguer une utilisation différentielle des domaines de référence par les acteurs de l'étude. Des comparaisons multiples ont été effectuées suite aux analyses de variance. Le coefficient de corrélation empirique de Spearman a été utilisé pour vérifier la consistance et la stabilité des évaluations de l'efficacité d'un programme. Les résultats obtenus ont par la suite été confirmés par le test de signification des corrélations basé sur la transformation de Fisher (Sokal & Rohlf, 1981) et par la statistique de Kappa (Fleiss, 1981). Le test généralisé du rapport de vraisemblance maximale (*generalized likelihood ratio test*

*statistic*) a finalement permis de vérifier la concordance entre les référents privilégiés de façon théorique par les acteurs et les référents utilisés de façon préférentielle en situation d'évaluation (Milliken & Johnson, 1989; Seber, 1984). La présentation détaillée des résultats statistiques se retrouvent dans Gervais (1996).

## LES RÉSULTATS

### La constance dans le temps d'une évaluation de l'efficacité d'un programme

Il s'agissait ici de vérifier si les acteurs, confrontés à des résultats de mesure similaires, ont évalué de la même façon l'efficacité de deux séries comparables de programmes fictifs présentés en des temps différents de l'étude. Des résultats obtenus, il est permis de conclure que les variations observées dans les cotations se sont avérées insuffisantes et que, par conséquent, les évaluations produites par les acteurs aux deux temps de l'étude sont jugées constantes.

D'intéressantes pistes de recherche futures peuvent ici être mentionnées. Il serait effectivement à propos d'explorer, à l'aide d'une stratégie méthodologique qualitative, le processus d'adaptation de l'acteur à la tâche évaluative demandée ainsi que les stratégies cognitives qu'il a mises en œuvre pour produire ses jugements d'efficacité. A-t-il, par exemple, utilisé une approche davantage intuitive lorsque confronté une première fois aux programmes à évaluer? Cette approche s'est-elle modifiée suite à une meilleure connaissance de la tâche à accomplir et une clarification du script cognitif nécessaire à la résolution de la tâche évaluative, ce qui l'aurait conduit vers une approche davantage rationnelle et analytique?

### La stabilité et la consistance d'une évaluation de l'efficacité d'un programme

Il s'agissait d'examiner si les acteurs cotent de façon identique l'efficacité d'un même programme fictif présenté dans des conditions d'évaluation différentes. Appelés à évaluer deux fois le même programme à l'intérieur d'une même série, les acteurs ont fait preuve de consistance dans leurs évaluations. Et lorsque appelés à évaluer un même programme en deux temps différents (1<sup>ère</sup> série, 2<sup>e</sup> série), les acteurs ont démontré de la stabilité dans leurs évaluations. En ces deux occasions, les programmes cibles ont reçu des cotes d'effi-



capacité comparables. Les évaluations ainsi produites se sont avérées fidèles ( $.32 < r < .85$ ). Ces résultats appuient ceux obtenus en rapport avec la constance d'une évaluation de l'efficacité d'un programme.

La concordance entre les référents adoptés par les acteurs en situation d'évaluation et ceux privilégiés de façon théorique

Il importait ici de comparer les données provenant de la cotation des programmes fictifs et les données obtenues suite à la hiérarchisation théorique des domaines de référence. Des résultats obtenus, il est apparu une concordance, chez tous les groupes d'acteurs de l'étude, entre les domaines de référence utilisés de façon préférentielle en situation d'évaluation et les domaines théoriquement identifiés comme plus importants. En d'autres termes, les référents associés à l'évaluation de l'efficacité d'un programme semblent les mêmes implicitement et explicitement pour chacun des groupes. Les acteurs ont évalué en conformité avec ce qu'ils disent qu'ils auraient fait en situation d'évaluation. C'est donc dire que les représentations théoriques et pratiques du construit d'efficacité concordent chez les acteurs d'un programme. Une correspondance existe entre ce que les acteurs disent qu'ils feraient une fois placés en situation d'évaluation et ce qu'ils font effectivement dans cette circonstance. La stabilité des référents utilisés par les acteurs est à nouveau démontrée.

L'ensemble des résultats ci-dessus mentionnés permettent de percevoir les acteurs d'un programme comme des êtres capables d'identifier les éléments qui guident leurs jugements. Ces acteurs apparaissent de plus comme des êtres cohérents et conséquents dont les discours sont constants et intégrés dans les pratiques évaluatives. Ces éléments vont à l'encontre de certains tenants des théories de la décision qui affirment que les acteurs ne savent pas toujours sur quoi ils basent leurs jugements, leurs décisions et qui stipulent qu'il existe une différence entre ce que les acteurs disent qu'ils préfèrent (*theory in action*) et ce que leurs comportements disent qu'ils préfèrent (*theory in use*) (Argyris & Schon, 1978; Nisbet & Wilson, 1977; Slovic & Lichtenstein, 1971).

Ces mêmes résultats n'ont toutefois pas permis d'apprécier la fluidité attendue des valeurs et des intérêts des acteurs dans le temps ainsi que la mouvance consécutive de leur choix de référents pour mener à terme une évaluation. L'intervalle temporel ici retenu est

apparu insuffisant pour apprécier l'interférence possible des éléments humains et sociaux dans la rationalité attendue d'une évaluation de l'efficacité d'un programme. Dans une perspective de court terme, cette étude a cependant démontré que ces éléments n'influaient pas sur la fidélité d'une évaluation. Il serait fort intéressant de reprendre cette étude à moyen terme, dans un ou deux ans par exemple, afin de vérifier si les mêmes résultats seraient alors obtenus.

### L'homogénéité des évaluations de l'efficacité d'un programme

Cette autre vérification avait pour but de déterminer si les acteurs appartenant à un même groupe évaluent de façon similaire l'efficacité d'un programme. Il faut principalement retenir des résultats obtenus qu'il existe à la fois une homogénéité dans les évaluations produites par les acteurs lorsque regroupés par type (décideurs, intervenants ou usagers), et ce peu importe leur organisation d'origine, et une homogénéité dans les évaluations produites par les acteurs regroupés cette fois par organisation, et ce sans tenir compte du type d'acteurs auquel chacun s'associe. Il est donc possible de dégager un patron homogène d'évaluation de l'efficacité d'un programme chez les acteurs de l'étude.

Ces évaluations similaires semblent référer à une convergence des représentations individuelles de l'efficacité en une vision commune de l'efficacité d'un programme chez les acteurs concernés. Effectivement, ces acteurs semblent partager un schème similaire d'interprétation des événements se produisant en rapport à un programme, ce qui entraîne l'élaboration de jugements d'efficacité comparables. Cette même vision de l'efficacité peut être explorée sous un angle développemental. En effet, considérant que tous les groupes ont été en contact avec les programmes auxquels ils sont affiliés pendant un minimum de cinq années, ceci permet de croire que le fait d'avoir partagé le même espace spatio-temporel et d'avoir été imprégné de la culture spécifique au programme conduit à l'expression d'une rationalité commune chez les acteurs. Ce partage d'une même culture semblerait donc conditionner la vision de l'efficacité véhiculée par les acteurs d'un programme. Il serait intéressant d'examiner dans le futur comment se développe la conception de l'efficacité d'un programme chez un acteur et de vérifier le rôle que peut effectivement jouer le temps de contact avec le programme dans la vision de l'efficacité de l'acteur. Les acteurs nouvellement intégrés à un programme poseraient-ils les mêmes jugements d'efficacité?

## L'utilisation différentielle des domaines de référence selon certains acteurs d'un programme

Il était ici question de déterminer si les acteurs de l'étude font preuve d'une utilisation différentielle des domaines de références en situation d'évaluation. En d'autres termes, il s'agissait d'examiner si les acteurs, regroupés par type et par organisation, utilisent les mêmes référents pour produire leurs évaluations de l'efficacité d'un programme. Il est apparu que le domaine spécifique a été coté plus élevé et qu'il a été habituellement suivi du domaine opérationnel. Les domaines structurel et stratégique ont présenté une importance plutôt moyenne pour les acteurs et ont démontré une plus grande instabilité. Le domaine systémique s'est vu accorder le moins d'importance. Le domaine spécifique s'est donc positionné comme le domaine de référence privilégié par l'ensemble des groupes d'acteurs étudiés, à l'exception du groupe d'utilisateurs qui accorda une importance comparable au domaine opérationnel. Il faut également noter que les liens entre les autres domaines sont nombreux et que ces domaines ont souvent été traités de façon indifférenciée par les acteurs. C'est le cas notamment du domaine opérationnel qui s'est trouvé confondu, chez l'ensemble des groupes, à la fois avec le domaine stratégique et avec le domaine structurel. Les domaines structurel et stratégique ont été traités de façon semblable par cinq des six groupes. Les liens dégagés entre le domaine systémique et les autres domaines sont apparus moins clairs.

Les résultats permettent maintenant d'affirmer que, lorsque placés en situation d'évaluation d'un programme, les acteurs se positionnent de façon identique et sélectionnent la même dimension de l'efficacité comme référent pour mener à terme cette évaluation. Cela va à l'encontre d'une utilisation différentielle selon le type d'acteurs puisque transparaît plutôt un consensus quant au domaine à considérer de façon prioritaire dans un contexte d'évaluation de l'efficacité d'un programme. Les acteurs s'entendent donc sur la façon dont s'exprime un programme efficace. Le domaine spécifique apparaît ainsi clairement comme le référent qu'ils privilégient et ce sont par conséquent les résultats obtenus par le programme en rapport avec ce domaine d'activités qui sont prioritairement considérés lors de l'élaboration d'un jugement d'efficacité (ou d'inefficacité) d'un programme. Cela témoigne d'une même vision de l'efficacité chez l'ensemble des acteurs de l'étude, vision qui s'articule sur ce que devrait avoir accompli le programme en rapport avec ses objectifs.

Il est toutefois intéressant de constater que la diversité des rôles joués dans le programme par les décideurs, les intervenants et les usagers ainsi que la multiplicité de leurs intérêts et des objectifs qu'ils poursuivent face à un même programme, ne semblent pas affecter le jugement final qu'ils posent sur l'efficacité d'un programme. En effet, l'efficacité dite spécifique constitue ce qui est prioritairement recherché à la fois par les décideurs, par les intervenants et par les usagers. Ces résultats surprennent puisqu'ils vont à l'encontre de ce qui est habituellement véhiculé dans les écrits. En effet, la sélection des groupes d'acteurs à intégrer dans un processus d'évaluation de l'efficacité d'un programme est souvent présentée comme critique: la perspective retenue par ces groupes étant perçue comme ayant une influence majeure sur la façon dont l'efficacité est définie et opérationnalisée.

Ce choix collectif de retenir le domaine spécifique comme référent le plus important pourrait s'expliquer par le fait que ce domaine représente le mieux l'aspect instrumental du programme et que la notion d'objectifs de résultats y est centrale. Les objectifs étant plus facilement mesurables que les processus et les comportements, le domaine spécifique se prêterait plus facilement à un consensus chez les acteurs en rapport à ce qui est perçu comme valable ou désirable pour un programme. Il s'avérerait conséquemment la sphère du programme la moins sujette à la politisation. Ce choix de privilégier le même référent pour réaliser une évaluation de l'efficacité d'un programme pourrait de plus témoigner de l'attachement des groupes d'acteurs aux objectifs du programme et de leur implication dans la réalisation d'un projet commun (la mission du programme) en vue de l'obtention de bénéfices mutuels. Il serait intéressant d'examiner de façon plus exhaustive l'ajustement collectif des intérêts des différents groupes pour en venir à la fabrication d'un sens commun accordé à l'efficacité et, de là, en venir au choix privilégié du même référent lors d'une évaluation de l'efficacité d'un programme.

Il importe également d'examiner ce que deviennent les autres dimensions de l'efficacité d'un programme soit, les dimensions structurelle, opérationnelle, systémique et stratégique. Il convient de questionner la place qu'elles détiennent effectivement lors d'une évaluation de l'efficacité d'un programme. Ces dimensions apparaissent aujourd'hui comme ayant une importance secondaire dans l'élaboration du jugement final concernant l'efficacité d'un programme. Tous les groupes d'acteurs de l'étude, à l'exception du groupe d'usagers, ont en effet accordé une considération moindre à ces dimen-

sions lors de leurs évaluations des programmes fictifs. Le groupe d'utilisateurs a accordé, quant à lui, une importance comparable aux domaines spécifique et opérationnel. Ceci suggère que ce groupe semble investir le programme non seulement pour ses objectifs explicites (domaine spécifique), mais également pour d'autres raisons associées cette fois au vécu expérimenté par l'utilisateur à l'intérieur du programme (domaine opérationnel).

Il importe également de rappeler l'existence des liens multiples entre les domaines structurel, opérationnel, systémique et stratégique et la difficulté rencontrée lors des analyses à distinguer un traitement différent de ces domaines par les acteurs lors de la cotation des programmes fictifs. Il faut mentionner que ces domaines proviennent d'un découpage artificiel du construit d'efficacité qui a été proposé suite à l'analyse de la documentation en lien avec les approches traditionnelles d'évaluation. Pris globalement, ces quatre domaines semblent représenter l'aspect fonctionnel du programme. Ils apparaissent comme des éléments qui contribuent davantage à l'atteinte de l'efficacité dite spécifique par un programme plutôt que d'être une finalité en soi. Ces domaines constituent des composantes dynamiquement interreliées d'un processus non linéaire conduisant à l'efficacité d'un programme. Ils se positionnent plus précisément comme des déterminants (ou des obstacles) à l'efficacité d'un programme. Les domaines structurel, opérationnel, systémique et stratégique reflètent en fait les opportunités et les contraintes technologiques, organisationnelles, sociales et environnementales qui viennent conditionner l'atteinte des objectifs d'un programme dans un contexte donné. En cela, ils permettent d'expliquer la relative efficacité ou inefficacité d'un programme et fournissent l'information capable d'éclairer sur les changements à apporter dans un programme pour en accroître l'efficacité.

Il est également permis de croire que ces quatre domaines soient le lieu où s'expriment les différents enjeux de pouvoir, d'influence et de contrôle davantage associés au rôle tenu par les acteurs à l'intérieur du programme. Ces domaines seraient un espace de dialogue ainsi qu'un lieu d'expression et de négociation des objectifs poursuivis spécifiquement par chaque groupe et ce, parallèlement à leur implication dans l'atteinte des objectifs du programme. La perspective sociopolitique dans laquelle s'inscrivent ces domaines laisse transparaître un risque accru de conflits et de divisions entre les acteurs concernés par l'efficacité d'un programme. Il est donc possible de présager que l'entente relative chez les différents grou-

pes d'acteurs quant à la façon d'évaluer l'efficacité d'un programme pourrait s'effriter lorsque viendra le temps de prendre une décision et d'entreprendre des actions face à ce même programme. Il est effectivement plausible que les moyens suggérés par les différents groupes pour atteindre les objectifs d'un programme varient d'un groupe à l'autre en fonction des valeurs et des enjeux qu'ils sous-tendent et qu'un jeu de négociation s'opère selon le contexte. La présente étude n'ayant pas sondé ce terrain, il deviendrait tout à fait pertinent de vérifier, lors de recherches ultérieures, dans quelle mesure les décisions et les actions suggérées par les différents groupes d'acteurs varient selon la position des acteurs dans un programme.

Une nuance doit toutefois être apportée quant au traitement accordé par les acteurs aux domaines structurel, opérationnel, systémique et stratégique. La présente étude ne permet pas de conclure à la non considération de ces domaines lors d'une évaluation de l'efficacité d'un programme, à leur non pertinence ou à leur élimination souhaitable du cadre d'analyse proposé. Les analyses effectuées ont plutôt permis d'ordonner ces domaines selon l'importance à leur accorder lors d'une évaluation. Par conséquent, il est plausible que ces domaines constituent des référents secondaires en lien avec l'efficacité d'un programme et que leur utilisation soit pondérée en fonction du contexte. De nombreux commentaires émis par les participants vont d'ailleurs en ce sens. Plusieurs ont en effet souligné le fait que chaque domaine apporte sa propre contribution à l'efficacité d'un programme et qu'il doit être considéré comme tel. Également, l'analyse des feuilles de commentaires complétées par les participants à la pré-expérimentation (N = 20) laisse entrevoir qu'en moyenne 4.7 indicateurs provenant de 2.5 domaines ont été considérés par les individus pour produire leur évaluation de l'efficacité des programmes fictifs, ce qui irait dans le sens d'une utilisation simultanée de plusieurs référents s'appuyant sur une conception multidimensionnelle de l'efficacité.

Ces éléments rejoignent également la mise en garde exprimée par Cameron (1978) lequel mentionne qu'il serait fallacieux de réduire l'efficacité uniquement à une question d'atteinte d'objectifs, position qui est également reprise par Moynagh (1993). D'ailleurs pour emprunter à Guba et Lincoln (1981, 1986), l'efficacité en tant qu'objet d'étude composé de multiples relations formant une structure d'ensemble ne peut être comprise que lorsqu'elle est considérée de façon holistique, ce qui implique la prise en considération des éléments

humains, politiques, sociaux, culturels et contextuels présents à l'intérieur des quatre domaines pré-cités.

## LES PRINCIPALES CONTRIBUTIONS DE L'ÉTUDE

Cette étude a d'abord permis de jeter un regard empirique sur un objet d'étude, l'efficacité d'un programme. Elle a révélé une pratique consensuelle chez les différents acteurs de l'étude quant au référent prioritairement utilisé pour compléter l'évaluation de l'efficacité d'un programme. L'efficacité a été de fait évaluée de façon préférentielle d'après les intentions spécifiques du programme plutôt que d'après les valeurs et les préférences particulières des acteurs en présence. La compatibilité ainsi observée entre les perspectives des différents groupes d'acteurs, en ce qui a trait à l'efficacité désirée d'un programme, vient témoigner de l'adhésion de ces groupes aux objectifs visés par le programme. Cette étude laisse également présager que des dissensions pourraient, cette fois, survenir dans les décisions et les actions découlant de l'évaluation de l'efficacité d'un programme selon les types d'acteurs consultés. En cela, elle appuie la nécessité d'une évaluation à caractère partenarial pour améliorer la pertinence du système décisionnel, consécutif à l'évaluation elle-même, ainsi que du système d'action qui en découlera. Le caractère pluraliste d'une telle démarche permettrait le maillage entre les personnes, les groupes et le programme. Ceci permettrait d'assurer la légitimité politique et la viabilité d'une évaluation de l'efficacité d'un programme en plus de contribuer à une meilleure utilisation des résultats d'évaluation.

Cette étude se démarque également des études antérieures par la consultation de trois groupes d'acteurs provenant des programmes participants, soit les décideurs, les intervenants et les usagers. La prise en considération des perspectives spécifiques à ces gens de terrain constitue à notre connaissance la première étude nord-américaine en ce genre. De plus, le fait de solliciter une pluralité d'acteurs en tant qu'experts de la situation a permis de franchir un des obstacles inhérents au métier de chercheur, soit celui de ne pas pouvoir tout comprendre ou tout saisir d'un programme, de sa mise en oeuvre, de son fonctionnement, de son environnement et des préoccupations qu'il soulève. La conception pluraliste de l'évaluation qui a été retenue a, par conséquent, contribué à enrichir la crédibilité et la pertinence de la démarche proposée et constituée, de fait, un exemple de collaboration et d'alliance possible entre la recherche et les lieux de pratique.

De plus, la participation à cette étude des acteurs clés d'un programme a permis une valorisation de leurs expériences respectives comme acteur d'abord, puis comme critique au regard de l'efficacité de programmes similaires aux leurs. Certains éléments viennent appuyer cette répercussion chez l'acteur de sa participation. Il faut d'abord mentionner que plus de la moitié des participants ont retourné la feuille facultative de commentaires qui était jointe au dernier envoi postal. De ce nombre, un peu moins de 50% ont fait explicitement mention de la réflexion qu'a engendrée leur participation à l'étude. Il faut également considérer la démarche vécue par les acteurs impliqués lors des étapes préliminaires de l'étude, soit lors de la construction des instruments de mesure. La démarche de sélection et de validation des indicateurs d'efficacité a nécessité chez ces acteurs une observation du quotidien de leur programme, laquelle permit par la suite l'identification des éléments-clés permettant de dire si ce programme est efficace ou non. Lors de l'étude, l'ensemble des acteurs ont par conséquent eu l'occasion de réfléchir sur leur rôle et sur leurs pratiques à l'intérieur d'un programme et de se positionner face au fonctionnement du programme ainsi qu'aux moyens à mettre en oeuvre pour atteindre les objectifs poursuivis. En cela, cette étude désirait contribuer à sensibiliser les acteurs à la mobilisation individuelle nécessaire pour que surviennent des changements au niveau organisationnel.

La stratégie de recherche utilisée s'avère une autre contribution appréciable de l'étude. Son originalité s'appuie sur l'utilisation de TRIAGE, une technique consensuelle de recherche d'information (Plante et Côté, 1993) et son application simultanée à trois types d'acteurs issus du domaine de la santé ainsi que sur l'emploi de programmes fictifs comme moyen d'accès aux représentations de l'efficacité que se font différents acteurs d'un programme. Les résultats positifs obtenus (acceptabilité par les participants, coût, flexibilité, capacité à répondre aux questions de recherche) rendent la transférabilité d'une telle stratégie à d'autres problématiques d'évaluation fort prometteuse.

## LES LIMITES DE L'ÉTUDE

Au plan conceptuel, le cadre d'analyse de l'efficacité d'un programme utilisé dans l'étude décompose l'efficacité en cinq dimensions distinctes lesquelles semblent couvrir l'espace attribué à ce phénomène par les approches traditionnelles d'évaluation de l'efficacité. Les écrits ne permettent toutefois pas d'affirmer que ces dimensions sont



exhaustives et qu'elles reflètent la réalité du phénomène. Ce cadre analytique doit donc être perçu pour ce qu'il est réellement, soit un effort d'intégration des différentes options théoriques présentées dans la documentation en vue de fournir une réponse satisfaisante davantage aux problèmes de mesure habituellement rencontrés lors de l'évaluation de l'efficacité. Une autre limite conceptuelle est associée au fait que les principales dimensions de l'efficacité ou domaine de référence ont été traités dans l'étude comme des catégories discrètes. La réalité est toutefois autre: ces domaines ne sont pas mutuellement exclusifs et des interactions existent entre eux.

Au plan méthodologique, il faut tout d'abord préciser que l'approche par simulation utilisée offre aux participants des situations d'évaluation artificielles. En cela, et malgré les nombreuses précautions prises lors de la construction de la stratégie méthodologique, elle reste imparfaite, n'étant pas en mesure de reproduire exactement les conditions d'évaluation qui se retrouvent dans les organisations. Le mode de sélection des participants constitue une autre limite de l'étude. Le fait que cette sélection se soit effectuée de façon volontaire laisse sans réponse la question de la représentativité réelle des groupes d'acteurs. Aucune information n'est d'ailleurs disponible relativement aux acteurs qui ont refusé de participer. Il faut toutefois mentionner le taux élevé d'acceptation de participation à l'étude chez les individus contactés (91.25%). De plus, le fait que la collecte des données ne soit pas ponctuelle mais qu'elle s'effectue en plusieurs étapes ne permet qu'un contrôle imparfait du facteur temps. Les conditions dans le programme évoluant constamment, les comportements étudiés demeurent toujours en interaction avec d'autres facteurs difficilement contrôlables. Il faut également mentionner que la passation des instruments s'est effectuée de façon autonome et, qu'en ce sens, aucune indication n'a été donnée aux participants concernant le lieu où procéder à l'évaluation des programmes fictifs (travail, maison), le moment de le faire (pendant la journée, soir), quant à la prise de notes entre les passations et quant à la consultation possible de d'autres personnes. Quant au principal instrument de mesure utilisé dans l'étude, il faut noter l'absence d'appui solide dans les écrits quant aux choix faits en ce qui a trait à sa longueur (20 programmes fictifs) et quant à sa complexité (10 résultats de mesure par programme fictif). L'ensemble de ces limites incite à la prudence et restreint la généralisation des résultats. Il faut toutefois les mettre en lien avec l'aspect exploratoire de l'étude et reconnaître que certaines de ces limites étaient inévitables considérant la spécificité de cette étude.

## RÉFÉRENCES

- Argyris, C., & Schon, D. (1978). *Organizational learning: A theory of action perspective*. Reading, MA: Addison-Wesley.
- Cameron, K.S. (1978). Measuring organizational effectiveness in institutions of higher education. *Administrative Science Quarterly*, 23, 604–632.
- Cameron, K.S. (1980). Critical questions in assessing organizational effectiveness. *Organizational Dynamics*, 9(2), 66–80.
- Cameron, K.S. (1981a). Construct space and subjectivity problems in organizational effectiveness, *Public Productivity Review*, 5(2), 105–121.
- Cameron, K.S. (1981b). The enigma of organizational effectiveness. *New Directions for Program Evaluation*, 11, 1–13.
- Cameron, K.S. (1984). The effectiveness of ineffectiveness. *Research in Organizational Behavior*, 6, 235–285.
- Cameron, K.S. (1986). Effectiveness as paradox: Consensus and conflict in conceptions of organizational effectiveness. *Management Science*, 32(5), 539–553.
- Cameron, K.S., & Whetten, D.A. (1983a). Organizational effectiveness: One model or several. Dans K.S. Cameron & D.A. Whetten (Eds.), *Organizational effectiveness: A comparison of multiple models* (pp. 1–24). New York: Academic Press.
- Cameron, K.S., & Whetten, D.A. (1983b). Some conclusions about organizational effectiveness. Dans K.S. Cameron & D.A. Whetten (Eds.), *Organizational effectiveness: A comparison of multiple models* (pp. 261–277). New York: Academic Press.
- Fleiss, J.F. (1981). *Statistical methods for rates and proportions*. New York: John Wiley.
- Gervais, M. (1996). *Étude exploratoire des domaines de référence utilisés par différents acteurs lors de l'évaluation de l'efficacité de programme*. Thèse de doctorat. Département de mesure et évaluation. Faculté des sciences de l'éducation. Ste-Foy, QC: Université Laval.

- Gervais, M. (1998). Repenser le concept d'évaluation de l'efficacité d'une organisation. *Revue canadienne d'évaluation de programme*, 13(2), 89–112.
- Gervais, M. (sous presse). La simulation: Une approche innovatrice appliquée au domaine de l'évaluation de programme. *Mesure et évaluation en éducation*.
- Gibson, J.L., Ivancevich, J.M., & Donnely, J.H. (1991). *Organizations: Behavior, structure, processes*. Homewood, IL: Irwin.
- Guba, E., & Lincoln, Y. (1981). *Effective evaluation*. Newbury Park, CA: Sage.
- Guba, E., & Lincoln, Y. (1986). The countenance of fourth generation evaluation: Description, judgment and negotiation. *Evaluation Studies Review Annual*, 11. 70–88.
- Harrison, M.I. (1987). *Diagnosing organizations: Methods, models, and processes*. Newbury Park, CA: Sage.
- Joubert, P. (1991). La pratique de l'évaluation dans le contexte de la réforme de la loi sur les services de santé et de services sociaux: Un processus de construction et de reconstruction. *Actes du colloque sur l'avenir de l'évaluation au Québec* (pp. 15–27). Québec, QC: Société québécoise d'évaluation de programme et École nationale d'administration publique.
- Lewin, A.Y., & Minton, J.W. (1986). Determining organizational effectiveness: Another look and an agenda for research. *Management Science*, 32(5), 514–538.
- Milliken, G.A., Johnson, D.E. (1989). *Analysis of messy data: Non replicated experiments*. New York: Van Nostrand/Reinhold.
- Miner, J.B. (1988). Effectiveness in organizations. Dans J.B. Miner (Ed.), *Organizational behavior: Performance and productivity* (pp. 33–71). New York: Random House.
- Moynagh, W.D. (1993). *La présentation et la vérification de l'information sur l'efficacité: De la théorie à la pratique*. Ottawa, ON: Fondation canadienne pour la vérification intégrée.

- Nisbet, R.E., & Wilson, T. (1977). Telling more than we can know: Verbal reports and mental processes. *Psychological Review*, 134, 231–259.
- Péron, Y., & Strohmenger, C. (1985). *Indices démographiques et indicateurs de la santé des populations*, cat 82-543F. Ottawa, ON: Statistiques Canada.
- Plante, J., & Côté, M. (1993). *TRIAGE: Technique d'animation d'un groupe expert*. Texte non-publié.
- Quinn, R.E., & Rohrbaugh, J. (1983). A spatial model of effectiveness criteria: Towards a competing values approach to organizational analysis. *Management Science*, 29(3), 363–377.
- Schroder, H.M. (1989). *Managerial competence: The key to excellence*. Dubuque, IO: Kendall/Hunt.
- Seashore, S.E. (1983). A framework for an integrated model of organizational effectiveness. Dans K.S. Cameron & D.A. Whetten (Eds.), *Organizational effectiveness: A comparison of multiple models* (pp. 55–70). New York: Academic Press.
- Seber, G.A. (1984). *Multivariate observations*. New York, NY: John Wiley.
- Slovic, P., & Lichtenstein, S. (1971). Comparison of Bayesian and regression approaches to the study of information processing in judgment. *Organizational Behavior and Human Performance*, 6, 649–744.
- Sokal, R.R., & Rohlf, F.L. (1981). *Biometry: The principles and practice of statistics in biological research*. New York: Freeman.
- Steers, R.M. (1975). Problems in the measurement of organizational effectiveness. *Administrative Science Quarterly*, 20(4), 546–558.
- Steers, R.M. (1984). *Introduction to organizational behavior*. Greenview, IL: Scott, Foresman and Company.