

## La simulation : une approche applicable au domaine de l'évaluation de programme

Marie Gervais

Volume 21, numéro 1, 1998

URI : <https://id.erudit.org/iderudit/1091358ar>

DOI : <https://doi.org/10.7202/1091358ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

ADMEE-Canada - Université Laval

ISSN

0823-3993 (imprimé)

2368-2000 (numérique)

[Découvrir la revue](#)

Citer cet article

Gervais, M. (1998). La simulation : une approche applicable au domaine de l'évaluation de programme. *Mesure et évaluation en éducation*, 21(1), 61–83. <https://doi.org/10.7202/1091358ar>

Résumé de l'article

L'analyse de la documentation a permis d'établir une importante variabilité dans la façon de définir et de mesurer l'efficacité d'un programme. Comment alors vérifier si différents acteurs d'un même programme conçoivent de façon similaire l'efficacité de ce programme ? Grâce à la collaboration de 35 acteurs provenant d'une organisation du domaine de la santé de la région de Québec, 20 programmes fictifs ont été élaborés. Le cadre conceptuel utilisé est présenté, les programmes fictifs élaborés sont décrits, les différentes étapes de construction de ces programmes fictifs sont détaillées et une synthèse des résultats obtenus lors de leur préexpérimentation est présentée. Les résultats démontrent que ces programmes fictifs constituent un outil valide et pratique permettant d'observer les comportements des acteurs en situation d'évaluation, et de là, ils permettent de saisir la conception de l'efficacité privilégiée par ces mêmes acteurs. Les applications possibles de l'approche par simulation dans d'autres contextes d'évaluation sont également présentées.

## La simulation : une approche applicable au domaine de l'évaluation de programme

Marie Gervais

Université Laval

MOTS-CLÉS : Évaluation de programme, simulation, référents, efficacité, indicateurs

*L'analyse de la documentation a permis d'établir une importante variabilité dans la façon de définir et de mesurer l'efficacité d'un programme. Comment alors vérifier si différents acteurs d'un même programme conçoivent de façon similaire l'efficacité de ce programme? Grâce à la collaboration de 35 acteurs provenant d'une organisation du domaine de la santé de la région de Québec, 20 programmes fictifs ont été élaborés. Le cadre conceptuel utilisé est présenté, les programmes fictifs élaborés sont décrits, les différentes étapes de construction de ces programmes fictifs sont détaillées et une synthèse des résultats obtenus lors de leur préexpérimentation est présentée. Les résultats démontrent que ces programmes fictifs constituent un outil valide et pratique permettant d'observer les comportements des acteurs en situation d'évaluation, et de là, ils permettent de saisir la conception de l'efficacité privilégiée par ces mêmes acteurs. Les applications possibles de l'approche par simulation dans d'autres contextes d'évaluation sont également présentées.*

KEY WORDS : Program evaluation, simulation, referents, effectiveness, indicators

*The analysis of the documentation on effectiveness allows to establish an important variability in the way the effectiveness of a program is defined and measured. How can we verify if different actors of a same program conceptualize the effectiveness of a program in the same way? Following the collaboration of 35 actors from an organization in the health field in the Québec City area, 20 simulated programs were developed. An overview of the conceptual framework that served as the basis for the 20 simulations, a description of the simulations including a summary of the process used to develop it and a report of the data regarding the preexperimentation of the program simulations are presented. The 20 simulations were found to be valid and a practical instrument that allows*

---

Note de l'auteur : Cet article constituant une partie de thèse de doctorat, l'auteur aimerait remercier monsieur Jacques Plante et madame Chantal Jeanrie qui ont respectivement agi à titre de directeur et de codirectrice de thèse.

*the observation of the behavior of the actors in a situation of evaluation of the effectiveness of a program, and from there, capturing the conception of the effectiveness favoured by these actors. The possible applications of the simulation approach in other contexts of evaluation will be presented.*

## Introduction

Cette étude constitue une étape préliminaire essentielle à la réalisation d'une recherche de plus grande envergure (Gervais, 1996). Cette recherche désire contribuer à clarifier la notion d'efficacité de programme, notion qui lors de la consultation des écrits sur le sujet est rapidement apparue comme étant empreinte d'une importante ambiguïté conceptuelle. Il importe d'entrée de jeu de s'assurer d'une différenciation claire entre l'évaluation de l'efficacité d'un programme, qui est l'objet de cette étude, et les autres types d'évaluation de programme proposés par Plante (1994).

Pour cet auteur, l'évaluation de la pertinence d'un programme cherche à vérifier si les objectifs d'un programme sont en lien direct avec les besoins à combler. L'évaluation de la cohérence d'un programme examine dans quelle mesure les principaux éléments du programme (objectifs, moyens, activités, ressources et acteurs) s'unissent les uns aux autres pour constituer un seul et même outil au service des objectifs visés. L'évaluation de l'à-propos d'un programme s'intéresse à l'étude des contraintes qui peuvent influencer le programme. L'évaluation de l'efficacité d'un programme examine la possibilité d'augmenter la quantité ou la qualité des résultats en agencant ou en gérant différemment les ressources, les moyens, les activités, et les acteurs d'un programmes. Elle s'intéresse également à la possibilité de diminuer la quantité de ressources, de moyens, d'activités et d'acteurs d'un programme sans changer la qualité des résultats obtenus. L'évaluation du rendement d'un programme s'intéresse aux différents éléments de l'environnement interne du programme qui peuvent venir affecter la productivité des acteurs. L'évaluation de l'impact d'un programme permet de repérer les effets non prévus du programme (effets bénéfiques ou indésirables) et permet d'en apprécier la magnitude. L'évaluation du maintien, quant à elle, revient à poser l'évaluation de l'efficacité mais en la situant dans le temps. Il s'agit donc de l'efficacité mais étalée dans le temps, à moyen et à long termes.

Considérant plus spécifiquement l'efficacité d'un programme, aucune définition de l'efficacité n'est reconnue aujourd'hui universellement. Les nombreuses définitions recensées semblent plutôt liées à un investissement

préférentiel de certaines composantes du programme selon la perspective théorique privilégiée et à la non-considération de certaines autres composantes jugées de moindre importance. Si on considère, par exemple, les principales approches d'évaluation de l'efficacité recensées dans les écrits, la première, l'approche centrée sur les objectifs, définit l'efficacité d'un programme comme étant fonction du degré d'atteinte d'un ensemble prédéterminé d'objectifs (Cameron, 1980; Etzioni, 1960; Gibson, Ivancevich & Donnelly, 1991; Goodman & Pennings, 1977; Keeley, 1984; Molnar & Rogers, 1976; Price, 1968; Zammuto, 1982). La deuxième approche, l'approche systémique, définit l'efficacité selon la capacité du programme à assurer sa survie, c'est-à-dire selon sa capacité à exploiter et à s'adapter à son environnement tout en préservant les éléments de base de son système, soit ses ressources, ses moyens, ses activités et ses acteurs (Gibson et al., 1991; Hannan & Freeman, 1977; Yutchman & Seashore, 1967). Selon la troisième approche, l'approche fonctionnelle, l'efficacité est synonyme de santé interne du programme (Edwards, Faerman & McGrath, 1986) et d'optimisation de ses aspects formels (structure et processus) et non formels (ressources humaines). La dernière approche, l'approche sociopolitique, définit l'efficacité d'un programme selon sa capacité à satisfaire certains groupes d'acteurs privilégiés par le pouvoir qu'ils détiennent ou par la position stratégique qu'ils occupent dans l'environnement du programme (Connolly, Conlon & Deutsch, 1980; Keeley, 1978; Pennings & Goodman, 1977; Pfeffer & Salancik, 1978; Zammuto, 1984).

Plusieurs définitions de l'efficacité coexistent donc dans les écrits. De plus, aucune des approches évaluatives recensées ne peut non plus être considérée supérieure et aucun ensemble d'indicateurs ne fait l'unanimité. Il semble d'ailleurs y avoir autant de façon d'opérationnaliser l'efficacité qu'il y a d'études sur le sujet. Cette multiplicité théorique et, de là, l'absence d'utilisation d'un même référent lors de l'évaluation de l'efficacité d'un programme apparaissent clairement comme un élément majeur contribuant à alimenter la confusion entourant l'étude du phénomène d'efficacité.

## **Les objectifs de l'étude**

Cette étude de type exploratoire a pour objectif principal l'élaboration et la validation d'une approche par simulation nécessaire à l'exploration des comportements d'évaluation manifestés par différents types d'acteurs lorsqu'ils sont confrontés à une situation d'évaluation de l'efficacité d'un

programme. Ainsi, cette méthode devra permettre d'examiner si différents types d'acteurs d'un programme (décideurs, intervenants et usagers) possèdent une même vision de l'efficacité d'un programme et recourent conséquemment au même référent pour porter un jugement d'efficacité sur ce programme. Elle devra également permettre de vérifier dans quelle mesure le choix de ces référents se trouve influencé par des variables contextuelles et temporelles.

L'approche par simulation adoptée permettra donc d'observer et d'analyser les comportements d'évaluation de plusieurs groupes d'acteurs dans des situations d'évaluation concrètes mais simulées. Elle renseignera 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. De là, elle informera sur le choix des référents privilégiés pour mener à terme leur opération d'évaluation.

Une présentation détaillée des différentes étapes de construction de l'approche par simulation sera par conséquent effectuée. Le contexte d'application de cette démarche, retenu à titre illustratif, correspond à un milieu de réadaptation comportant des programmes d'interventions s'adressant à une clientèle présentant des troubles mentaux. Le choix de ce milieu se justifie par le fait qu'il s'agit d'un des sites d'expérimentation déjà ciblés pour la recherche principale. Toutefois, il aurait été tout aussi possible de choisir pour cible de cette validation méthodologique un milieu de formation en éducation, celui-ci comportant tout autant des programmes d'interventions se prêtant à l'évaluation.

### **Le modèle conceptuel**

Le modèle conceptuel élaboré pour l'étude se présente comme une synthèse des enseignements dégagés des approches traditionnelles d'évaluation de l'efficacité d'un programme et des efforts faits par plusieurs auteurs pour intégrer les connaissances contemporaines sur l'efficacité. Ce modèle présente une conception multidimensionnelle de l'efficacité (Gervais, 1996). En effet, cinq grandes dimensions à l'efficacité sont reconnues à l'intérieur de ce modèle, dimensions qui s'articulent autour des cinq principaux domaines d'activités d'un programme (figure 1). Il s'agit des dimensions 1) structurelle; 2) opérationnelle; 3) systémique; 4) stratégique; 5) spécifique. Il convient maintenant d'examiner le caractère distinctif des cinq domaines d'activités d'un programme et de présenter, pour chacun, les composantes privilégiées du programme et les principales tâches ou activités associées.

Le domaine structurel couvre l'ensemble des activités du programme liées à ses ressources physiques, matérielles, financières, informationnelles et humaines et couvre l'ensemble des activités associées à sa structure. Il s'agit particulièrement de la division, de l'aménagement et de la hiérarchisation des tâches et des procédures de travail, du partage des responsabilités, des procédures de coordination et de contrôle, des relations entre les individus, les groupes et les unités ainsi que des réseaux de communication et d'information.

Le domaine opérationnel couvre les activités du programme associées à ses processus, à ses activités et aux comportements de ses membres. Ce domaine s'intéresse à l'environnement de travail et aux dynamiques internes du programme. Se trouvent associés à ce domaine les activités en lien avec la pratique professionnelle, la programmation des activités, la coordination, la prestation de services, la communication, le climat de travail, la motivation des membres du programme ainsi que leur performance, leur attachement et leur engagement face à celui-ci. Les activités rattachées au développement du personnel sont également considérées.

Le domaine systémique comprend les activités du programme en lien avec son environnement extérieur. Il est question des activités assurant à la fois la légitimité du programme et son ouverture vers l'extérieur. Ce domaine couvre les activités du programme en rapport à sa responsabilité sociale, à son rôle et à son positionnement dans le réseau. Ce sont donc les aspects interactif et adaptatif du programme avec son environnement qui sont mis en évidence dans le domaine systémique.

Le domaine stratégique contient les activités du programme associées à ses politiques et à ses pratiques de management. La notion de régulation est particulièrement présente à l'intérieur des activités touchées par ce domaine. Ainsi, les activités associées à la gestion du programme et de ses contraintes, à la médiation entre les différentes instances organisationnelles, à la planification stratégique, à la gestion de l'horizon temporel, au processus de prise de décision, à la supervision, au leadership, à l'évaluation, à la création d'un environnement de travail propice à la performance, au respect des normes et des règlements et à l'éthique sont principalement considérées par ce domaine.

Et finalement, le domaine spécifique s'intéresse aux activités du programme orientées vers l'obtention de résultats. Il est principalement question d'optimisation des objectifs du programme, de satisfaction des attentes et de production de services. Ce domaine englobe également les impacts non

anticipés du programme lors de la formulation des objectifs, ces impacts pouvant être bénéfiques ou indésirables.

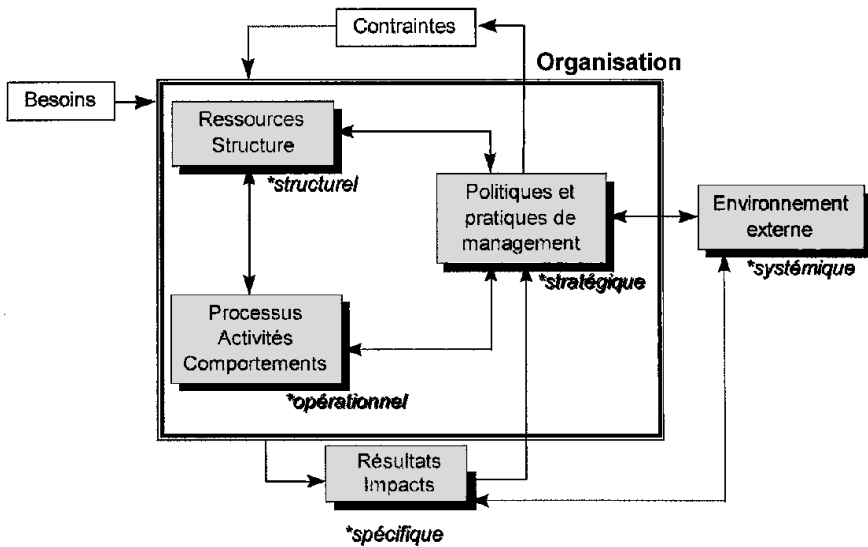


Figure 1. *Cadre d'analyse de l'efficacité d'un programme*

Ce modèle permet donc de classifier l'efficacité et de considérer ces cinq dimensions comme des référents partiels disponibles pour une opération d'évaluation de l'efficacité d'un programme. Les domaines d'activités 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.

### L'approche méthodologique

L'approche par simulation préconisée s'est inspirée de la procédure mise au point par Slovic et ses collaborateurs (Slovic & Lichtenstein, 1971 ; Slovic, Fischhoff & Lichtenstein, 1977). Cette procédure, connue sous l'appellation de *policy capturing*, a été employée lors de travaux ayant comme principal intérêt les modèles décisionnels utilisés dans les organisations (Hitt & Barr, 1989 ; Hitt, Ireland, Keats & Vianna, 1983 ; Hitt & Middlemist, 1979 ; Hitt & Tyler, 1991). D'importantes modifications ont dû être apportées à cette procédure pour permettre l'arrimage avec les besoins spécifiques de cette étude conduite dans le domaine de l'évaluation de programme.

L'objectif visé consiste à créer des simulations ou programmes fictifs qui constitueront des scénarios différents du comportement d'un programme sur une certaine période de temps. Elles seront obtenues par suite de la manipulation des paramètres structuraux du modèle conceptuel construit pour représenter l'efficacité d'un programme. Cette manipulation permettra l'attribution de scores de performance différents pour le programme en rapport à chacun de ses domaines d'activités. Un descriptif de ces programmes fictifs est présenté un peu plus loin.

Pour procéder à la construction de ces programmes fictifs, il a été décidé de recourir à des indicateurs d'efficacité. Cela se justifie par le désir d'offrir aux acteurs des stimuli concrets, quantifiables, en lien direct avec l'objet à évaluer et en rapport avec les pratiques évaluatives habituelles. Le terme «indicateur» utilisé dans l'étude se réfère à la définition proposée par Péron et Strohmenger (1985). Pour ces auteurs, un indicateur est 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. Un indicateur donne une indication, c'est-à-dire une information incomplète, mais significative et utile, sur le phénomène étudié (Saucier, 1995). Ces définitions rejoignent d'ailleurs celle de De Landsheere (1979), un indicateur s'avérant pour cet auteur un phénomène qui témoigne de la présence d'un autre. Considérant que les indicateurs détiennent un rôle-clé dans la construction des instruments de mesure et considérant qu'il a été convenu de travailler à partir d'indicateurs d'efficacité proposés par les acteurs des organisations plutôt que d'utiliser des indicateurs puisés à même les écrits, l'établissement de ces indicateurs a fait l'objet d'une démarche particulière qui sera maintenant explicitée.

### ***Le choix de la méthode de sélection des indicateurs d'efficacité***

La méthode de sélection retenue devait permettre le choix des deux meilleurs indicateurs d'efficacité pour chaque domaine d'activités d'un programme. Le processus de collecte de données s'est effectué selon la technique consensuelle TRIAGE (*T*echnique de *r*echerche d'*i*nformation par *a*nimation d'un groupe *e*xpert: Plante & Côté, 1993) en raison de la possibilité qu'offre cette technique de traiter plus d'un thème à la fois. La procédure utilisée a donc permis de traiter simultanément les cinq dimensions du cadre conceptuel de l'étude. De plus, alors que traditionnellement TRIAGE travaille avec un seul groupe expert, le fonctionnement habituel de cette technique a été ici adapté afin de recourir à l'expertise de trois types d'acteurs, soit des décideurs,



des intervenants et des usagers de programmes de réadaptation. La considération de plus d'un type d'acteurs visait à éviter que les données résultantes ne soient biaisées par les valeurs et par les intérêts propres à un type d'acteurs en particulier. Cela a d'ailleurs permis que les indicateurs choisis soient davantage représentatifs de l'organisation ciblée. L'application de la technique TRIAGE s'est déroulée en deux phases distinctes, une phase de production individuelle et une phase de production collective. Les indicateurs résultant de TRIAGE ont par la suite subi une deuxième validation par la consultation d'un groupe d'experts.

### ***Les participants à TRIAGE***

Vingt-neuf acteurs ont été recrutés sur une base volontaire pour participer à TRIAGE. Ces acteurs étaient des décideurs, des intervenants et des usagers des programmes administrés par une organisation du domaine de la santé et des services sociaux 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. Ils étaient inscrits dans un des programmes de l'organisation ou l'avaient quitté depuis moins de deux ans. Les usagers étaient atteints d'un trouble mental sévère et persistant (schizophrénie, psychose maniaco-dépressive), n'étaient pas déficients intellectuels et avaient un état mental stabilisé.

Tableau 1

#### ***Données sociodémographiques associées aux acteurs de TRIAGE***

| Acteurs      | Sexe |   | Âge     |       | Formation    |     | Travail <sup>1</sup> |       |
|--------------|------|---|---------|-------|--------------|-----|----------------------|-------|
|              | F    | M | Moyenne | é.-t. | Type         | %   | Moyenne              | é.-t. |
| Décideurs    | 5    | 0 | 44,80   | 3,96  | Baccalauréat | 67  | 22,00                | 2,92  |
| Intervenants | 5    | 4 | 42,11   | 6,33  | Baccalauréat | 56  | 17,22                | 5,93  |
| Experts      | 4    | 2 | 37,67   | 5,39  | Maîtrise     | 100 | 10,25                | 5,19  |

| Acteurs | Sexe |   | Âge     |       | Présence <sup>2</sup> |       | Occupation <sup>3</sup> |  |
|---------|------|---|---------|-------|-----------------------|-------|-------------------------|--|
|         | F    | M | Moyenne | é.-t. | Moyenne               | é.-t. | %<br>Actif              |  |
| Usagers | 6    | 9 | 43,87   | 7,81  | 19,21                 | 9,74  | 67                      |  |

Trois groupes homogènes d'acteurs ont été formés, soit un groupe de décideurs, un groupe d'intervenants et un groupe d'utilisateurs. Les données sociodémographiques de ces acteurs sont présentées au tableau 1. 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.

Les mêmes acteurs devaient prendre part aux deux étapes de TRIAGE. Deux acteurs n'ont toutefois pu être présents au deuxième temps de TRIAGE pour cause de conflits d'horaire. Leurs propositions d'indicateurs d'efficacité produites lors de la première étape ont été considérées lors des rencontres de groupes subséquentes. Ces participations incomplètes représentent moins de 7% de l'échantillon de départ de TRIAGE.

### ***Le déroulement de TRIAGE***

Lors de la phase de production individuelle de TRIAGE, les participants ont reçu par le courrier les documents de travail nécessaires à leur participation. Une mise en situation leur était offerte. Les participants devaient s'imaginer préparer l'évaluation de l'efficacité d'un programme semblable, pour ce qui est des objectifs, des ressources et de la clientèle, à celui qu'ils côtoient actuellement ou ont déjà côtoyé dans leur organisation respective. Il leur était demandé de considérer les cinq domaines d'activités d'un programme et de rechercher, pour chacun de ces domaines, les indicateurs d'efficacité jugés les plus pertinents. Une définition des domaines leur était alors fournie. Un maximum de cinq indicateurs par domaine était recherché. Une période de deux semaines était accordée aux acteurs pour réaliser les tâches spécifiques à cette étape.

Quatre cent quatre-vingt-huit propositions d'indicateurs ont été reçues. Le groupe de décideurs a produit le plus de propositions d'indicateurs (N = 178) suivi du groupe d'intervenants (N = 171) et du groupe d'utilisateurs (N = 139). Ces propositions ont par la suite été colligées de façon à créer, pour chaque groupe, une banque d'indicateurs d'efficacité propre à chaque domaine d'activités du programme. Ces banques d'indicateurs ont servi de point de départ pour l'étape de production collective, deuxième temps de TRIAGE.

L'étape de production collective de TRIAGE est une étape structurée qui mise sur l'interaction des participants. Des séances de travail d'une durée

approximative de trois heures ont été mises à l'horaire pour chaque groupe dans les locaux de l'organisation. Pour chaque groupe de TRIAGE, cinq banques d'indicateurs d'efficacité étaient disponibles, soit une banque pour chaque domaine d'activités d'un programme. Elles ont été étudiées une à une selon la même procédure. L'objectif était de dégager par consensus, pour chaque domaine d'activités d'un programme, les cinq indicateurs d'efficacité jugés les plus importants.

Un tri a été effectué parmi les indicateurs proposés. Les acteurs ont pu reformuler les indicateurs, les regrouper, les subdiviser, les rejeter ou les reporter lors de l'étude d'un autre domaine d'activités du programme. Ils ont également pu proposer de nouveaux indicateurs. Travaillant par approximations successives, il y eut évolution vers un consensus de groupe et sélection des indicateurs d'efficacité jugés les plus appropriés. Chaque groupe d'acteurs a sélectionné un total de 25 indicateurs d'efficacité, soit cinq indicateurs pour chaque domaine d'efficacité d'un programme. Ces indicateurs étaient dès lors disponibles pour une opération de validation.

### ***La validation des indicateurs d'efficacité***

Cette étape devait permettre de déterminer, à partir des listes produites à l'étape TRIAGE, les indicateurs qui pouvaient le mieux témoigner de l'efficacité d'un programme dans ses cinq principaux domaines d'activités. Six experts ont été recrutés de façon intentionnelle. Ils avaient été préalablement reconnus comme experts par leurs pairs. Ils devaient répondre à certains critères modifiés pour cette étude d'après ceux proposés par Rossi et Freeman (1989). Ces experts devaient avoir une connaissance suffisante de l'organisation et des différents programmes offerts. Ils devaient posséder une formation ou une expérience en évaluation et être actifs dans ce domaine. Ils devaient finalement être familiers avec les évaluations conduites à l'intérieur de leur organisation respective. Les caractéristiques sociodémographiques de ces experts sont également présentées au tableau 1.

L'instrument de mesure construit pour l'étape de validation était composé de cinq listes d'indicateurs d'efficacité, soit une liste par domaine d'activités d'un programme. Les indicateurs d'efficacité proposés à l'étape TRIAGE par les décideurs, les intervenants et les usagers ont été regroupés par domaine, d'où un maximum possible de 15 indicateurs par liste. Ces listes d'indicateurs ont été présentées dans un ordre différent pour chacun des experts consultés. L'ordre d'apparition des indicateurs à l'intérieur de ces listes a également varié aléatoirement d'un expert à l'autre.

Adoptant un rôle de juge, les experts devaient consulter ces listes d'indicateurs puis préciser, sur une échelle en 6 points, l'importance qu'ils accorderaient à chaque indicateur lors de l'évaluation de l'efficacité d'un programme (1 = très peu important, à 6 = très important).

Pour chaque indicateur présenté, la moyenne des cotes accordées par les experts a été calculée. Les indicateurs d'efficacité obtenant les deux meilleures moyennes ont été considérés comme les indicateurs les plus représentatifs des domaines auxquels ils sont associés et comme les meilleurs témoins de l'efficacité d'un programme en rapport à ce domaine. Un total de dix indicateurs a ainsi été obtenu, soit deux indicateurs par domaine d'activités d'un programme (tableau 2). Ces indicateurs étant sélectionnés par l'intermédiaire d'une démarche de double validation (TRIAGE, experts), leur validité pour l'organisation se trouvait mieux assurée. Il devenait alors possible de procéder à la construction des programmes fictifs.

Tableau 2  
*Indicateurs d'efficacité par domaine*

| Domaines     | Indicateurs   |
|--------------|---|
| Structurel   | Répartition du personnel en fonction des besoins<br>Adéquation de la structure en fonction des besoins    |
| Opérationnel | Capacité du personnel à répondre aux besoins de l'utilisateur<br>Cohésion de l'équipe multidisciplinaire  |
| Systémique   | Complémentarité avec les autres programmes et services<br>Partenariat avec les ressources de la région    |
| Stratégique  | Contenu du programme adapté aux besoins de l'utilisateur<br>Capacité à résoudre des problèmes à l'interne |
| Spécifique   | Satisfaction de l'utilisateur<br>Degré d'atteinte des objectifs du plan d'intervention pour l'utilisateur |

### *La construction des programmes fictifs*

Vingt programmes fictifs ont été élaborés. L'information contenue dans un programme fictif est présentée sous plusieurs formes (tableau 3). Un programme fictif possède comme principales caractéristiques: 1) un ensemble de dix indicateurs d'efficacité; 2) une échelle de mesure en trois points; 3) une échelle de cotation en quatre points; 4) des instructions quant à la façon de répondre.

Tableau 3  
*Information contenue dans un programme fictif*

|   |  | COTE OBTENUES 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 des personnels 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 |
|---------------------------|----------------------|--------------------|-------------------------|
| _____                     | _____                | _____              | _____                   |

Les dix indicateurs représentant les cinq domaines d'activités d'un programme proviennent de l'étape de validation précédente (TRIAGE, experts). Ils ont été présentés aléatoirement à l'intérieur de chaque programme fictif. Une échelle qualitative en trois points a été utilisée pour présenter les cotes obtenues par le programme en rapport à chaque indicateur d'efficacité. Une situation fictive était offerte aux participants. Il leur était mentionné qu'un groupe d'experts s'était intéressé aux 12 derniers mois d'opérations des programmes et qu'ils avaient mesuré leur valeur en rapport aux dix indicateurs d'efficacité. Ces experts ont utilisé à cette fin trois possibilités de cotes, soit une cote élevée, une cote moyenne et une cote faible.

Lors de la construction des programmes fictifs, une cote identique a été attribuée aux indicateurs représentant un même domaine. La force attribuée à chacun des domaines a fait l'objet d'une manipulation intentionnelle et varie, de fait, d'un programme fictif à l'autre. Dans chaque programme fictif, deux domaines ont été mis en évidence : un domaine à cote élevée et un autre à cote faible. Les trois autres domaines ont été neutralisés par l'attribution de cotes

moyennes. Par exemple, lors de l'élaboration du programme fictif illustré au tableau 3, une cote élevée a été attribuée aux deux indicateurs associés au domaine structurel, une cote faible a été accordée aux deux indicateurs associés au domaine spécifique et une cote moyenne a été attribuée aux indicateurs représentant les domaines opérationnel, systémique et stratégique. Ces cotes «élevée, moyenne et faible» ont été assignées selon une méthode aléatoire stratifiée. Dans l'ensemble des 20 programmes fictifs produits, chaque domaine a reçu le même nombre de cotes élevées, de cotes moyennes et de cotes faibles. Considérant, par exemple, le domaine structurel, une cote élevée lui a été attribuée dans quatre programmes fictifs, une cote moyenne dans 12 programmes fictifs et une cote faible dans quatre programmes fictifs. La même répartition a été respectée pour les quatre autres domaines.

Une échelle en quatre points a finalement été utilisée pour juger de l'efficacité des programmes fictifs. Ainsi, un programme fictif pouvait recevoir la cote très inefficace, inefficace, efficace ou très efficace. Tenant compte de l'ensemble de l'information présentée, il était demandé aux acteurs de se prononcer sur le degré d'efficacité à accorder à chacun des programmes et d'apposer un X dans la case correspondant au degré d'efficacité attribué à chacun d'eux.

## **La pré-expérimentation**

Préalablement à la préexpérimentation, plusieurs précautions ont été prises de façon à s'assurer que les programmes élaborés s'apparentent le plus possible à la réalité des programmes mis en place dans les organisations cibles. 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 des programmes auxquels les participants sont associés. Une vérification informelle a également été effectuée auprès d'acteurs œuvrant dans le domaine de la santé afin de s'assurer que l'information présentée (nature, quantité, format) permette effectivement aux acteurs de poser un jugement sur l'efficacité des programmes fictifs.

Une préexpérimentation formelle a par la suite été mise à l'horaire. Cette préexpérimentation avait comme objectifs: 1) de vérifier si l'instrument de

mesure proposé était capable de dégager des tendances, des préférences dans les comportements d'évaluation des acteurs; 2) de dépister et de permettre de résoudre les problèmes associés à l'administration de l'instrument de mesure.

Quatorze experts provenant de l'entreprise privée, du système de santé québécois, des milieux universitaire et gouvernemental ainsi que trois acteurs provenant de l'organisation cible, soit un décideur, un intervenant et un usager, ont été consultés lors de cette préexpérimentation. Ces acteurs devaient répondre aux mêmes critères de sélection que ceux mentionnés précédemment. Les experts étaient majoritairement des femmes (57%) et leur moyenne d'âge était de 48,86 ans (é.t. = 3,06). Parmi eux 71% détenaient une maîtrise ou un doctorat. Ces experts travaillaient en moyenne depuis 15,42 années dans leur organisation d'affiliation (é.t. = 6,40). Le décideur (50 ans) et l'intervenant (37 ans) étaient des femmes détentrices d'un baccalauréat et à l'emploi de l'organisation depuis respectivement 30 et 12 ans. L'usager était un homme de 37 ans qui avait été inscrit pendant deux ans dans les programmes de l'organisation et qui était alors actif sur le marché du travail.

Dans un seul envoi postal, les experts et les acteurs ont reçu l'ensemble des programmes fictifs. Ces derniers ont été présentés dans un ordre aléatoirement différent pour chaque participant. Les 20 participants ont tous accompli la tâche qui leur a été demandée dans l'intervalle de deux semaines qui leur avait été alloué à cet effet.

Un retour a été effectué sur l'expérience vécue par les participants. Une feuille de commentaires jointe aux documents associés à leur participation à l'étude a d'abord été utilisée. Elle invitait les participants à exprimer leurs observations, leurs critiques et leurs suggestions relativement à l'instrument et à la procédure utilisés. Elle leur demandait également de spécifier la stratégie utilisée pour coter les programmes fictifs ainsi que les indicateurs privilégiés pour réaliser la tâche demandée. Les participants ont par la suite été contactés par téléphone. Quatre-vingts pourcent des experts et 100% des acteurs ont répondu à ces entretiens semi-directifs, ce qui a permis de recueillir des commentaires additionnels sur l'approche méthodologique employée.

### ***Les résultats de la pré-expérimentation***

La grande majorité des participants à la préexpérimentation ont utilisé un processus séquentiel de traitement et d'analyse des résultats de mesure attribués aux indicateurs d'efficacité contenus dans les programmes fictifs, processus qui variait en complexité. Il a été effectivement observé que la plupart des participants ont cherché à simplifier le processus décisionnel

demandé en limitant le nombre d'indicateurs considérés et en pondérant certains indicateurs plus que d'autres. Les indicateurs sélectionnés et pondérés ont ainsi eu une influence majeure sur le processus décisionnel subséquent des individus. Plusieurs tendances ont pu être dégagées des façons de faire des participants. Ceux-ci ont utilisé en moyenne 4,7 indicateurs appartenant à 2,5 domaines différents pour produire leur jugement d'efficacité des programmes fictifs. La plupart ont d'abord considérés les cotes accordées à deux ou trois indicateurs préférentiels. Les deux indicateurs les plus fréquemment mentionnés sont la satisfaction de l'utilisateur et le degré d'atteinte des objectifs du plan d'intervention de l'utilisateur. Les cotes associées à ces indicateurs devenaient alors déterminantes pour l'établissement de la cote d'efficacité à apposer à un programme fictif. Par exemple, lorsque les indicateurs privilégiés avaient obtenu une cote élevée, les participants avaient tendance à coter le programme fictif «très efficace». Inversement, lorsque ces indicateurs indiquaient une cote faible, les participants tendaient plutôt vers une cote «très inefficace». Dans le cas d'une cote moyenne attribuée à ces indicateurs, deux ou trois autres indicateurs supplémentaires étaient alors considérés, ce qui permettait de discriminer le degré d'efficacité à attribuer au programme fictif et de distinguer ainsi entre une cote «efficace» et une cote «inefficace». Trois experts ont mentionné avoir considéré l'ensemble des dix indicateurs pour poser leur jugement d'efficacité alors qu'un autre expert a élaboré et utilisé une grille sophistiquée de cotation qui tenait compte à la fois de l'importance qu'il accordait personnellement à chaque indicateur et de la pondération qui lui avait été attribuée à l'intérieur du programme fictif évalué.

De plus, le questionnement qualitatif des participants, effectué par l'intermédiaire de la feuille de commentaires annexée à l'instrument de mesure et de l'entrevue téléphonique, a permis de spécifier que tous les participants ont été en mesure de préciser les éléments à partir desquels ils ont élaboré leur jugement d'efficacité des programmes présentés ainsi que la stratégie qu'ils avaient utilisée. Aucune stratégie dite intuitive ou de hasard n'a été recensée chez les participants. Il faut également préciser qu'au départ, les participants pensaient être en mesure de considérer l'ensemble des résultats associés aux dix indicateurs présentés. Rapidement, vers le deuxième ou troisième programme fictif, la plupart en sont venus à sélectionner un nombre restreint d'indicateurs jugés plus importants. Les dix premiers programmes fictifs semblent donc avoir permis aux participants de raffiner et d'adapter leur stratégie d'évaluation en fonction des exigences et des contraintes propres à la tâche demandée. Les participants ont effectivement mentionné être à quelques



reprises retourné vérifier leurs jugements antérieurs dans la première section de l'instrument de mesure (dix premiers programmes fictifs) et que leur jugement des programmes fictifs suivants était beaucoup plus assuré. Tous les participants ont également mentionné garder l'impression d'une grande cohérence dans l'ensemble des évaluations des programmes fictifs produites.

L'analyse visuelle de la distribution des cotes d'efficacité selon les caractéristiques des programmes fictifs permet de dégager deux autres tendances intéressantes. En effet, les programmes fictifs à l'intérieur desquels une cote élevée était attribuée au domaine spécifique ont dans l'ensemble été jugés « très efficaces ». À l'inverse, les programmes fictifs à l'intérieur desquels une cote faible était attribuée au domaine spécifique ont habituellement reçu une cote « très inefficace ». Dans cette préexpérimentation, le jugement de l'efficacité des programmes fictifs semble donc d'abord suivre la force accordée au domaine spécifique. Il apparaît donc exister une cohérence entre les indicateurs théoriquement retenus par les participants pour produire leurs évaluations, indicateurs reliés surtout au domaine spécifique, et ceux effectivement utilisés en situation d'évaluation, c'est-à-dire lors de la cotation des programmes fictifs. L'examen de cette distribution des cotes d'efficacité n'a toutefois pas permis de départager l'importance accordée aux quatre autres domaines.

Quant aux programmes fictifs eux-mêmes, aucun participant n'a mentionné avoir été confronté à un nombre trop élevé d'indicateurs présentés à l'intérieur de chaque programme fictif ( $N = 10$ ). De plus, l'absence de définitions associées aux indicateurs utilisés n'a pas semblé déranger la performance des participants. Le nombre de programmes fictifs présentés n'a pas non plus contribué à créer une surcharge cognitive qui aurait pu mettre en danger la qualité des évaluations produites ( $N = 20$ ).

L'échelle de mesure utilisée a été jugée facile à interpréter. Il est question ici de la présentation visuelle des cotes – élevée, moyenne ou faible – accordées à chaque indicateur. L'échelle de cotation en quatre points utilisée pour spécifier le degré d'efficacité à associer à chaque programme fictif a également été jugée facile d'utilisation. Cette échelle en quatre points a démontré qu'elle est en mesure d'offrir l'étendue nécessaire pour assurer la progression logique dans le choix de réponses tout en évitant les recouvrements. Elle offre également suffisamment de variance pour permettre de distinguer plusieurs niveaux d'efficacité. Il importe également de rappeler que la simplicité de cette échelle en quatre points a été rendue nécessaire afin de

rendre accessible, dans la recherche principale, l'instrument de mesure à un large éventail d'individus possédant une formation et une expérience en évaluation allant de très limitées ou absentes (usagers) à très importantes (décideurs).

Concernant cette échelle de cotation, il importe de mentionner que la distribution des cotes d'efficacité démontre que les participants ont jugé dans 60% des cas les programmes fictifs comme étant «très efficace» (12%) ou «efficace» (48%) et dans 40% des cas comme étant «très inefficace» (13%) ou «inefficace» (27%). Cela va dans le sens habituel de la théorie de la mesure, à savoir, une prédisposition favorable chez les participants à privilégier lors de leurs cotations la cote «efficace» et une hésitation de leur part à utiliser les extrêmes de l'échelle, soit les cotes «très efficace» et «très inefficace».

### **Domaines d'application de l'approche par simulation**

L'originalité de la stratégie de recherche adoptée s'appuie sur la construction de programmes fictifs comme moyen d'accès aux représentations d'un objet d'évaluation multidimensionnel, ici l'efficacité, que se font différents acteurs d'un programme. Les résultats positifs obtenus (acceptabilité par les participants, coût, flexibilité, capacité à discriminer les référents utilisés) rendent la transférabilité d'une telle stratégie à d'autres problématiques d'évaluation fort prometteuse. L'approche par simulation devient de fait applicable lorsque l'objet d'évaluation ciblé possède plus d'une dimension et qu'il s'avère pertinent de discriminer laquelle ou lesquelles sont utilisées de façon privilégiée pour appuyer la production d'un jugement dans certains contextes d'évaluation.

Considérant la recherche principale dont il était mention dans l'introduction, il importe de mentionner que l'approche par simulation a également été préexpérimentée avec succès dans une autre organisation du domaine de la réadaptation, organisation qui traite cette fois une clientèle présentant des incapacités d'ordre physique, neurologique ou sensorielle. Cela vient démontrer la transférabilité possible de cette approche dans d'autres contextes d'évaluation.

Une application de cette approche dans le domaine de l'étude de besoins, domaine où il convient de discriminer entre différents besoins ou différentes catégories de besoins à hiérarchiser, serait également envisageable. Il pourrait

également être intéressant d'explorer l'importance relative accordée à certaines valeurs, croyances ou principes dans le choix des besoins à combler de façon prioritaire.

En éducation, les situations où la mise en œuvre de stratégies cognitives par l'étudiant est nécessaire pour résoudre une situation d'apprentissage pourrait constituer un terrain d'application propice de l'approche par simulation. Face à une tâche donnée, quels types d'information privilégie l'étudiant? Comment en vient-il à choisir son information? Comment intègre-t-il les éléments d'information présents pour en venir à mener à terme la tâche demandée? Confronté à de nombreux éléments d'information, comment en vient-il à simplifier son processus décisionnel pour être en mesure de réaliser la tâche demandée? En effet, la discrimination des stratégies cognitives privilégiées par l'étudiant, et de là, la spécification du cheminement qu'il a emprunté pour en arriver à produire son jugement face à une situation donnée peuvent facilement constituer des pistes de recherche intéressantes.

Dans le domaine de l'évaluation, l'utilisation d'une approche par simulation pourrait également permettre de mettre en évidence la façon dont l'individu gère les différentes influences auxquelles il doit faire face avant d'effectuer son jugement. Une telle approche pourrait, par exemple, conduire à explorer les influences que détiennent les différents enjeux contextuels, les jeux politiques en présence, les éléments de pouvoir, d'influence et de contrôle sur le processus perceptuel et évaluatif de l'individu et comment se construit conséquemment sa prise de décision. Cette approche pourrait également permettre de mieux comprendre comment les caractéristiques personnelles de l'évaluateur (style personnel, compétence à évaluer, expériences antérieures, etc.) et ses conflits motivationnels (ambition et objectifs personnels) face à la tâche demandée peuvent venir interférer dans son processus décisionnel.

## **Contributions de l'approche par simulation**

Certains avantages peuvent maintenant être clairement dégagés de l'utilisation d'une approche par simulation. Le premier avantage concerne les contraintes temporelles souvent expérimentées lors d'une opération d'évaluation. Le fait par exemple de travailler à partir de programmes réels obligerait les acteurs à un investissement de temps très important pour se familiariser avec ces programmes et pour être en mesure de porter un jugement approprié sur leur efficacité. Recourir à des programmes fictifs devient

pertinent si on tient compte de ces contraintes temporelles, les programmes fictifs permettant par leur aspect synthétique de focaliser uniquement sur les éléments-clés à prendre en compte pour élaborer un jugement.

Le deuxième avantage de cette approche concerne le contrôle des variables susceptibles de contaminer la situation d'évaluation (apparition d'événements particuliers, changement dans les attitudes et la motivation des acteurs, modification dans le milieu de travail, etc.). En offrant un ensemble de stimuli définis, les programmes fictifs permettent de minimiser l'interférence possible de ces variables contaminantes sur les comportements d'évaluation des acteurs. L'utilisation de programmes fictifs permet par conséquent d'isoler de façon plus satisfaisante l'influence des représentations qu'ont les acteurs face à un objet précis, dans le cas-ci l'efficacité, lors d'une opération d'évaluation.

Cette étude aura finalement permis de dresser une liste d'indicateurs d'efficacité en lien avec les cinq domaines d'activités d'un programme. Élaborés de façon inductive à la suite de la consultation de groupes d'acteurs des organisations cibles, ces indicateurs deviennent maintenant disponibles pour d'éventuelles activités d'évaluation des programmes associés à ces organisations. Ces indicateurs reflètent d'ailleurs le caractère situationnel de l'efficacité, puisqu'ils sont spécifiques aux programmes, à leur environnement tel que perçu par les acteurs et à leur stade développemental. Ces indicateurs expriment également les préoccupations des différents groupes d'acteurs. Ces indicateurs doivent donc être qualifiés de ponctuels et de conjoncturels. Ils doivent également être considérés comme des sources d'information à caractère stratégique permettant de distinguer les cibles responsables des situations jugées favorables ou défavorables au regard de l'efficacité d'un programme. Ils peuvent être utilisés, selon l'expression de Joubert (1991), pour assurer une rétroinformation orientée, laquelle permettra de préciser les modifications à apporter au programme pour en accroître l'efficacité. Ces indicateurs d'efficacité constituent par conséquent des critères d'action.

### **Limites de l'approche par simulation**

La principale limite de cette approche par simulation est liée au fait qu'elle offre aux participants, malgré les nombreuses précautions prises lors de la construction des programmes fictifs, des situations d'évaluation artificielles. Elle demeure donc imparfaite, n'étant pas en mesure de reproduire exactement les conditions d'évaluation qui se retrouvent dans les organi-

sations. Ces éléments mettent à nouveau en évidence la rigueur qui doit caractériser les étapes de construction de ces programmes fictifs de façon qu'ils puissent être valides, crédibles et familiers pour leurs utilisateurs éventuels.

La deuxième limite est cette fois associée à l'instrument de mesure lui-même. Il faut noter l'absence d'appui solide dans les écrits quant aux choix faits en ce qui a trait à la longueur de l'instrument de mesure (20 programmes fictifs) et quant à sa complexité (dix résultats de mesure par programme fictif). En effet, la documentation consultée démontre une grande variabilité quant au nombre de stimuli présentés aux participants. Les études recensées présentent de cinq à 25 programmes fictifs pouvant eux-mêmes comporter de dix à 25 indicateurs chacun. La seule mise en garde relevée traite du danger de surcharge cognitive chez le participant, liée à sa difficulté à traiter une si grande quantité d'information en même temps. Un instrument trop long ou trop complexe risquerait de créer de l'ennui, de la fatigue, du découragement et même un abandon de la part du participant. Un tel instrument pourrait également entraîner des changements non désirés dans les comportements d'évaluation du participant. Des changements de stratégie d'évaluation pourraient être appliqués, comme une cotation effectuée au hasard afin de simplifier et d'accélérer le processus d'évaluation ou encore une réponse selon une stratégie qui n'a rien à voir avec la stratégie évaluative utilisée au départ. Les résultats obtenus deviendraient alors non interprétables. Les vérifications informelles effectuées ainsi que la préexpérimentation ont permis d'assurer la faisabilité de l'approche en choisissant un nombre acceptable de programmes fictifs comportant un nombre limité de stimuli.

La troisième limite de cette approche se trouve associée à l'étape de validation des indicateurs. Les experts ayant individuellement procédé à leur sélection (méthode des juges), il est possible de supposer la possibilité que des résultats différents dans le choix final des indicateurs soient survenus dans la perspective d'une sélection cette fois effectuée dans un contexte d'échange et de choix consensuels.

Finalement, une dernière limite concerne le type de résultats produits. Afin de mieux exploiter la richesse de l'information générée par cette approche, il pourrait être judicieux de compléter par une méthodologie qualitative. Celle-ci pourrait permettre de mieux saisir comment l'individu organise son expérience évaluative, et de là, elle pourrait permettre de suivre la mécanique de choix et de stratégies mise en place par l'individu pour

réaliser la tâche demandée. Elle pourrait également permettre de discerner les motivations ou les cognitions sous-jacentes aux choix effectués. Cela pourrait se concrétiser par l'observation directe de l'individu en situation d'évaluation ou encore par son questionnement oral ou écrit effectué *a posteriori*.

## Conclusion

Cette étude exploratoire a permis d'établir sommairement, par le biais d'une approche par simulation, une description quantitative et qualitative du comportement de choix inhérent à l'évaluation de l'efficacité d'un programme. Elle a également permis de démontrer que la méthode élaborée est suffisamment sensible pour mettre en évidence les référents utilisés préférentiellement par des individus lors d'une activité d'évaluation. En ce sens, cette méthode a pu agir comme capteur des représentations d'un objet d'évaluation que se font ces mêmes individus. Il serait intéressant de poursuivre l'exploration des pistes de recherche se dégageant de cette étude et de valider les résultats obtenus sur un plus grand nombre de participants, ce qui permettrait l'utilisation des statistiques traditionnelles.

## RÉFÉRENCES

- Cameron, K. S. (1980). Critical questions in assessing organizational effectiveness. *Organizational Dynamics*, 9(2), 66-80.
- Connolly, T., Conlon, E. J., & Deutsch, S. J. (1980). Organizational effectiveness: A multiple-constituency approach. *Academy of Management Review*, 5(2), 211-217.
- De Landsheere, G. (1979). *Dictionnaire de l'évaluation et de la recherche en éducation*. Paris, France : Presses universitaires de France.
- Edwards, R. L., Faerman, S. R., & McGrath, M. R. (1986). The competing values approach to organizational effectiveness: A tool for agency administrators. *Administration and Social Work*, 10(4), 1-14.
- Etzioni, A. (1960). Two approaches to organizational analysis: A critique and a suggestion. *Administrative Science Quarterly*, 5(2), 257-278.
- Gervais, M. (1996). *Étude exploratoire des domaines de référence utilisés par différents acteurs lors de l'évaluation de l'efficacité d'un programme*. Thèse de doctorat, Université Laval, Département de mesure et évaluation, Faculté des sciences de l'éducation, Sainte-Foy, QC.
- Gibson, J. L., Ivancevich J. M., & Donnely, J. H. (1991). *Organizations: Behavior, structure, processes*. Homewood, IL: Irwin.
- Goodman, P. S. & Pennings, J. M. (1977). *New perspectives on organizational effectiveness*. San Francisco, CA: Jossey-Bass.

- Hannan, M. T., & Freeman, J. (1977). Obstacles to the comparative studies. In Goodman, P. S., & Pennings, J. M. (éds.), *New perspectives on organizational effectiveness* (pp. 106-131). San Francisco, CA: Jossey-Bass.
- Hitt, M. A., & Barr, S.H. (1989). Managerial selection decision models: Examination of configural cue processing. *Journal of Applied Psychology*, 74(1), 53-61.
- Hitt, M. A., Ireland, R. D., Keats, B. W., & Vianna, A. (1983). Measuring subunit effectiveness. *Decision Sciences*, 14(1), 87-102.
- Hitt, M. A., & Middlemist, R. D. (1979). A methodology to develop the criteria and criteria weightings for assessing subunit effectiveness in organizations. *Academy of Management Journal*, 22(2), 356-374.
- Hitt, M. A., & Tyler, B. B. (1991). Strategic decision models: Integrating different perspectives. *Strategic Management Journal*, 12, 327-351.
- Keeley, M. (1984). Impartiality and participant-interest theories of organizational effectiveness. *Administrative Science Quarterly*, 29(1), 1-25.
- Keeley, M. (1978). A social-justice approach to organizational evaluation. *Administrative Science Quarterly*, 23, 272-292.
- Molnar, J. J., & Rogers, D. L. (1976). Organizational effectiveness: An empirical comparison of the goal and system resource approaches. *Sociological Quarterly*, 17(3), 401-13.
- Morin, E. M. (1989). *Vers une mesure de l'efficacité organisationnelle: Exploration conceptuelle et empirique des représentations*. Thèse de doctorat non publiée, Université de Montréal, Montréal, QC.
- Pennings, P. S., & Goodman, J. M. (1977). Toward workable framework. In Goodman, P. S., & Pennings, J.M. (éds.), *New perspectives on organizational effectiveness* (pp. 146-184). San Francisco, CA: Jossey-Bass.
- Pfeffer, J., & Salancik, G. R. (1978). *The external control of organizations*. New York, NY: Harper & Row.
- Péron, Y., & Strohmenger, C. (1985). *Indices démographiques et indicateurs de la santé des populations* (Cat. 82-543F). Ottawa, ON: Statistiques Canada.
- Plante, J. (1994). *Évaluation de programme*. Département de mesure et évaluation. Université Laval. Sainte-Foy, QC: Auteur.
- Plante, J., & Côté, M. (1993). [TRIAGE: Technique d'animation d'un groupe expert]. Texte non publié.
- Price, J. L. (1968). *Organizational effectiveness: An inventory of propositions*. Homewood, IL: Richard D. Irwin.
- Rossi, P. H., & Freeman, H. E. (1989). *Evaluation: A systematic approach*. Newbury, CA: Sage.
- Saucier, A. (1995). *Les indicateurs et la gestion par résultats*. Direction générale de la planification et de l'évaluation. Ministère de la santé et des services sociaux. Gouvernement du Québec. Québec, QC: Auteur.
- Slovic, P., Fischhoff, B., & Lichtenstein, S. (1977). Behavioral decision theory. *Annual Review of Psychology*, 28, 1-39.
- Slovic, P., & Lichtenstein, S. (1971). Comparaison of Bayesian and regression approaches to the study of information processing in judgement. *Organizational Behavior and Human Performance*, 6, 649-744.

- Yutchman, E., & Seashore, S. E. (1967). A system resource approach to organizational effectiveness. *American Sociological Review*, 32(6), 891-903.
- Zammuto, R. F. (1982). Theory and research on organizational effectiveness. In R. F. Zammuto (éd.), *Assessing organizational effectiveness* (pp. 21-58). Albany, NY: State University of New York Press.
- Zammuto, R. F. (1984). A comparaison of multiple constituency models of organizational effectiveness. *Academy of Management Review*, 9(4), 606-616.