

LE RÔLE DES INDICATEURS DANS L'ÉVALUATION DES MESURES D'ENCADREMENT

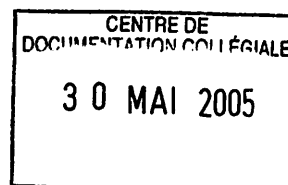
**Recherche réalisée pour le Carrefour de la réussite
au collégial**

par

Nathalie Prévost

Mars 2005

Centre de documentation collégiale
1111, rue Lapierre
Lasalle (Québec)
H8N 2J4



LE RÔLE DES INDICATEURS DANS L'ÉVALUATION DES MESURES D'ENCADREMENT

**Recherche réalisée pour le Carrefour de la réussite
au collégial**

par

Nathalie Prévost

Mars 2005

Rédaction
Nathalie Prévost

Soutien à la recherche
Marie-Johanne Lacroix
Jean-Paul Michaud

Travaux de secrétariat
Julie Desormeaux

Révision linguistique
Rolande LeBlanc Vadeboncœur

23755
785536

Pour ne pas alourdir le texte, le masculin est utilisé à titre épïcène.

TABLE DES MATIÈRES

INTRODUCTION	7
1. LES DONNÉES.....	11
1.1 Données quantitatives.....	11
1.2 Données qualitatives	12
2. LES INDICATEURS EN ÉDUCATION	13
2.1 Indicateurs quantitatifs.....	13
2.2 Indicateurs qualitatifs.....	14
2.3 Indicateurs pour le programme de tutorat.....	15
3. CHOISIR UN INDICATEUR.....	17
4. PLAN D'ÉVALUATION D'UNE MESURE D'ENCADREMENT.....	19
5. TABLEAU DE BORD DES INDICATEURS.....	21
6. TECHNIQUES ET INSTRUMENTS DE COLLECTE DE DONNÉES.....	25
6.1 Techniques et instruments pour recueillir des données quantitatives	25
6.2 Techniques pour recueillir des données qualitatives	25
6.3 Quantitatif ou qualitatif?.....	27
6.4 Avantages et inconvénients de quelques techniques	28
6.5 Instruments de collecte de données répertoriés par Monaghan et Chaloux.....	30
6.6 Techniques et instruments employés dans l'évaluation du programme de tutorat ...	33
6.6.1 <i>Groupe de recherche sur le tutorat (GRT)</i>	34
6.6.2 <i>AmeriCorps Tutoring Outcomes Study</i>	34
6.6.3 <i>Les chercheurs Larose, Soucy et Duchesne</i>	35
6.6.4 <i>Ehly et Larsen</i>	35
6.6.5 <i>Désy</i>	36
6.6.6 <i>Cyrenne et Lacombe</i>	36
7. TYPOLOGIE RETENUE	37
7.1 Les indicateurs et l'évaluation de contexte	41
7.2 Les indicateurs et l'évaluation d'intrants	42
7.3 Les indicateurs et l'évaluation de processus	44
7.4 Les indicateurs et l'évaluation de produit	45
8. STANDARD.....	51
9. EFFICACITÉ.....	53
CONCLUSION	57
ANNEXE 1	
LISTE DE VÉRIFICATION DES QUALITÉS DES INDICATEURS	61
ANNEXE 2	
TABLEAU DE BORD DES INDICATEURS	65
ANNEXE 3	
D'AUTRES INDICATEURS	69
LISTE DES TABLEAUX ET DES FIGURES	73
SOURCES DOCUMENTAIRES	75

INTRODUCTION

Le réseau des collèges publics québécois possède une longue tradition en matière de reddition de comptes et d'analyse institutionnelle¹. On peut même aller jusqu'à affirmer qu'il a été en quelque sorte un chef de file dans ce domaine, notamment sous la poussée initiale des travaux des chercheurs du Centre d'animation, de développement et de recherche en éducation (CADRE), au milieu des années 70. Par après, dans la suite du rapport de la Commission Nadeau, les pratiques d'évaluation institutionnelle des cégeps se sont développées dans un cadre plus formel, soit celui de la Commission d'évaluation du Conseil des collèges, nouvellement créé depuis dix ans.

Par ailleurs, l'ordre collégial, durant sa courte existence d'un peu plus de 35 ans, a été très souvent l'objet d'interpellations. On lui a demandé, à de nombreuses reprises, de faire la preuve de son utilité et de son efficacité. En ce sens, l'évaluation n'est donc pas une préoccupation nouvelle pour le réseau collégial public.

Malgré ce fait, chaque interpellation dont les collèges font l'objet n'est pas sans complications. Tout d'abord, elle est souvent perçue comme nouvelle, très peu d'établissements ayant pu consacrer les ressources nécessaires au développement de systèmes de suivi et à une pratique continue de l'évaluation.

En conséquence, à l'égard de chaque nouvelle demande, les collèges, à des degrés divers, se sentent démunis tant sur le plan des démarches d'évaluation que sur le plan des procédés et des instruments pour les conduire.

En outre, la vision de l'évaluation axée sur une approche perçue comme presque exclusivement quantitative, vision empruntée au domaine de l'administration, est aux yeux de plusieurs trop réductrice lorsqu'il s'agit de rendre compte de plusieurs mérites de la formation, ce qui n'est pas sans créer un certain malaise et même une certaine résistance; également, des techniques servant à analyser, dans cette perspective d'« accountability », des données quantitatives peuvent nécessiter, de la part des utilisateurs, des connaissances poussées en statistiques.

Comme illustration de situations dans lesquelles les collèges peuvent se trouver à l'égard d'une interpellation, la réalisation, à l'automne 2003, des bilans des plans de réussite, un travail d'analyse qui incluait l'appréciation de quelques mesures d'aide que renferment les plans, a mis en évidence le fait que l'évaluation demeure un processus difficile malgré tous les efforts qui peuvent être déployés pour le simplifier, à cause particulièrement du nombre et de la nature des facteurs concernés. C'est un processus qui comporte de nombreux choix et qui nécessite de

¹ Dans ce très bref rappel historique qui constitue cette première partie de l'introduction, nous utilisons indifféremment analyse institutionnelle et évaluation institutionnelle. À ce moment-ci, il n'est pas apparu nécessaire non plus de faire des distinctions entre l'évaluation globale d'un collège et l'évaluation d'un de ses programmes, d'un de ses services ou même d'une de ses activités, et encore moins de distinguer l'évaluation d'une mesure d'aide de l'évaluation de son efficacité. Ces distinctions seront faites plus loin dans le texte lorsque ce sera nécessaire. (Ce rappel a été rédigé par Jean-Paul Michaud, membre du Carrefour de la réussite au collégial.)

la rigueur dans la collecte et dans l'analyse des données. Ce processus est rendu encore plus complexe par le manque d'instruments et de cadre de référence.

Pour contribuer à surmonter les difficultés inhérentes au processus d'évaluation, et dans la poursuite de son mandat de soutien aux collèges relativement au dossier de la réussite, le Carrefour de la réussite au collégial a, tout d'abord, soutenu le déploiement d'un système de suivi des cohortes.

Ensuite, il a financé une recherche sur les indicateurs utilisés en éducation dans les provinces canadiennes et dans quelques pays occidentaux.

Enfin, il présente aujourd'hui ce document, *Le rôle des indicateurs dans l'évaluation des mesures d'encadrement*², dont l'objectif est de présenter un cadre de référence qui permettrait d'éclairer les décisions que des directeurs des études doivent prendre par rapport au processus d'évaluation.

Plus précisément, ce document abordera, par un rappel de connaissances de base sur les indicateurs, plusieurs volets du domaine de l'évaluation dans l'optique du modèle de Stufflebeam, un cadre de référence permettant la coexistence de deux courants en évaluation :

1. Satisfaire aux besoins d'exécuter l'évaluation en réponse à une interpellation;
2. Tirer avantage du processus d'évaluation et porter au mieux la mesure d'encadrement.

Des situations pour exemplifier plusieurs des principes traités dans ce document seront puisées dans les univers virtuel et authentique du tutorat à travers deux de ses modalités : le tutorat maître-élève et le tutorat par les pairs.

Pourquoi les indicateurs?

Le domaine de l'évaluation mérite d'être bien traité, car c'est sur la base de données et d'analyses que reposeront les décisions relatives à une mesure d'encadrement, des décisions qui influenceront grandement sur le parcours scolaire des étudiants. Un certain niveau de précaution méthodologique est donc requis en évaluation, et est atteignable sans que nous soyons de grands spécialistes en la matière. Mais avant de nous engager dans des techniques de collecte et d'analyse de données, nous devons en premier saisir le domaine des indicateurs dans son intégralité.

² Étant donné que ce document porte sur l'évaluation, l'emploi du mot « mesure » comme substantif du verbe « mesurer » est inévitable. Nous nous rapportons aux auteurs Raymond Boudon et Paul Lazarsfeld pour assigner à ce vocable, qui se retrouve par-ci, par-là dans ce document, le sens relatif à « l'acte d'attribuer des propriétés à un objet ». (Raymond Boudon et Paul Lazarsfeld, *Le vocabulaire des sciences sociales*, Paris, 1965, p. 27.) Dans ce sens, il se distingue du mot « mesure » qui s'emploie comme abrégé de l'expression « mesure d'encadrement ». Nous resterons donc vigilants et utiliserons « mesure d'encadrement », et non « mesure » tout court, quand nous parlerons d'une intervention.

Sans abreuver les lecteurs d'interminables détails, nous voulons en plus souligner que « mesurer » diffère du terme « évaluer » : le premier ne permet pas de poser un jugement, car il constitue une action neutre, tandis que le second est relié à une appréciation de la valeur de l'objet d'évaluation. (Kerlinger et Nadeau, rapporté par France Fontaine, *Guide pour la création et la révision d'un programme d'études*, 1992, Montréal, p. 45.)

Or, les indicateurs servent en tant qu'outils d'orientation et d'action, pouvant ainsi nous guider dans notre choix du type de données, d'instruments et d'analyses appropriés à l'évaluation que nous désirons mener. L'évaluation ne se réduisant pas à une analyse de taux de réussite ou de notes attribuées aux étudiants, différents types d'indicateurs sont requis à diverses fins.

Les intervenants ont déjà recours à des indicateurs autres que ceux ayant trait au rendement scolaire : d'une part, il y a des indicateurs de type quantitatif, comme le nombre de participants masculins, le taux d'absentéisme à un cours; d'autre part, il y a des indicateurs de type qualitatif, tels que les caractéristiques des étudiants, la perception par les enseignants des retombées d'une intervention sur les étudiants.

Pourquoi le modèle de Stufflebeam?

Pour nous aider à reconnaître les rôles que viennent jouer les indicateurs en évaluation, une typologie qui marque nettement les différences entre les catégories qui la composent est indispensable. Le modèle C.I.P.P.³ de Stufflebeam, un modèle typologique qui répond à ces critères, servira d'assise à la typologie d'indicateurs que nous présenterons dans ce document, et privilégie quatre catégories d'évaluation, à savoir l'évaluation de contexte, l'évaluation d'intrants, l'évaluation de processus et l'évaluation de produit.

Des indicateurs concourant à nous guider dans le type de données et d'instruments de collecte de données pour chacune de ces catégories d'évaluation forment alors cette typologie d'indicateurs, popularisée par divers auteurs en évaluation. Les catégories de ce système articulé d'indicateurs sont donc : indicateurs de contexte, indicateurs d'intrants, indicateurs de processus et indicateurs de produit. Ces derniers, des indicateurs très usuels en éducation, sont ceux qui renvoient couramment aux taux de réussite, aux scores à des tests, aux résultats à des épreuves que font passer des autorités.

Pourquoi le tutorat maître-élève et le tutorat par les pairs?

Afin que les lecteurs puissent rapidement établir un parallèle entre les composantes théoriques de l'évaluation et la pratique, nous voulons illustrer, dans un contexte familier, plusieurs des notions. Outre le fait, d'après l'étude menée par Désy en 1996, que plus de 33 établissements collégiaux au Québec offraient depuis 1989 le programme de tutorat par les pairs⁴, les collèges doivent y consacrer une part de l'allocation versée par le ministère de l'Éducation et destinée à les soutenir dans la mise en œuvre de leur plan de réussite. Ces réalités nous ont fait reconnaître le tutorat par les pairs comme une intervention largement répandue dans le réseau collégial québécois.

Quant au modèle maître-élève, étant donné qu'il est également l'objet d'évaluations dans le réseau collégial, comme le rapporte le sondage de Larose, Soucy et Duchesne⁵, nous nous

³ Les initiales formant le sigle C.I.P.P. signifient « contexte », « intrants », « processus » et « produit ».

⁴ Jocelyne Désy, *Le tutorat par les pairs tel que perçu par les élèves*, Sainte-Foy, 1996.

⁵ Simon Larose, Nathalie Soucy et Stéphane Duchesne, « Examen des programmes de tutorat maître-élève dans les collèges du réseau québécois », *Pédagogie collégiale*, vol. 13, n° 4, mai 2000, p. 11-37 — récupéré de www.cdc.qc.ca/ped_coll/soucy_13_4.html.

servirons de l'information couvrant ce modèle pour multiplier les possibilités en matière de techniques d'évaluation.

Bref, ce document repose sur les éléments suivants :

1. Des notions relatives aux indicateurs et des moyens de les exploiter;
2. Des indications concernant le tableau de bord des indicateurs, des techniques de collecte de données et d'autres actions faisant partie du processus d'évaluation d'une mesure d'encadrement;
3. Une exploration du modèle C.I.P.P. et des parties qui le composent et qui l'unifient en un outil pratique;
4. Diverses significations attribuées au concept d'efficacité.

Nous voulons traiter, de façon très succincte, des principaux éléments constituant l'opération d'évaluation, et former (ou reformer) le dialogue entre les membres responsables de l'évaluation et un ensemble théorique d'éléments appliqué à ce domaine.

À propos de la terminologie

Mais, avant de poursuivre, nous voulons tout d'abord énoncer une remarque importante qui concerne ce document. Par souci de conformité au vocabulaire employé dans le réseau collégial, nous proposons de réserver les vocables « mesure d'encadrement » et « intervention » pour une activité, un projet ou un programme⁶, c'est-à-dire pour une unité où les intervenants font des actions qui influent d'une façon pratique sur le parcours scolaire des étudiants au collégial. Pour se référer au tutorat, nous recommandons le vocable « programme de tutorat⁷ » qui renvoie simultanément aux deux modèles : le modèle maître-élève et le modèle par les pairs.

⁶ Jacques Plante a établi des différences entre ces termes : « Au regard de l'institution, le programme constitue une unité plus petite, ancrée sur les finalités et les objectifs poursuivis par l'institution responsable [...] c'est un assemblage réfléchi d'idées, de stratégies et de personnes mis en place pour transformer des choses ou l'état d'une chose. »

« Le terme projet dénote un sens très voisin de celui de programme. [...] On reconnaît au programme une durée de vie plus longue que celle d'un projet au sein d'une organisation. Un programme peut même recouvrir plusieurs projets qui eux sont généralement à caractère transitoire. »

« L'activité se réfère à tout ce qui bouge, ce qui est en mouvement, ce qui occupe. On désignera souvent sous ce chapeau les stratégies, méthodes et procédés mis en place pour atteindre des objectifs [...] Par rapport à un programme et à un projet, elle a une durée de vie plus courte, elle est plus facilement transformable et adaptable. » (Jacques Plante, *Évaluation institutionnelle*, Département de mesure et évaluation, Université Laval, juillet 1994, p. 6 et 7, « document inédit ».)

Cette « hiérarchie » a constitué un point d'ancrage dans la rédaction de notre document.

⁷ Il serait difficile de déterminer catégoriquement si le tutorat est un « programme » ou un « projet » ou même une « activité ». Quoi qu'il en soit, nous sommes d'avis que les notions présentées dans ce document peuvent s'appliquer dans le cadre de l'évaluation de toute unité.

⁹ France Fontaine, *Guide pour la création et la révision d'un programme d'études*, 1992, Montréal, p. 45.

1. LES DONNÉES

Les définitions de l'évaluation foisonnent dans la littérature éducationnelle. L'une d'entre elles qui a attiré notre attention est, bien entendu, celle que Stufflebeam lui-même formule, et que France Fontaine a interprétée comme suit : « l'évaluation en éducation est le processus par lequel on délimite, obtient et fournit des informations utiles permettant de juger des décisions possibles⁹ ». Nous remarquons, par cette définition, que les décisions sont au cœur du processus d'évaluation, et nous reviendrons plus tard sur cet aspect quand il sera question de la typologie de Stufflebeam, car nous voulons, avant d'aborder cette dernière, établir des distinctions entre plusieurs notions courantes en évaluation.

Les méthodes d'évaluation d'une mesure d'encadrement exigent un investissement de ressources (humaines, financières, matérielles) qui est déterminé en fonction de la portée de cette mesure d'encadrement : plus la portée est grande, plus les ressources investies doivent être importantes. Il suffit de prendre pour exemples une activité comme une journée d'accueil collectif et un programme comme le tutorat pour illustrer la différence de niveau d'implication qu'exigerait l'évaluation de chacune de ces deux mesures d'encadrement, que ce soit sur le plan des instruments, sur celui des analyses ou sur celui des techniques de collecte de données.

Mais, quelle que soit la portée, certains éléments doivent être considérés dans une évaluation; or, quand les indicateurs sont clairement formulés et que les moyens de collecte de données sont méthodiques et rigoureux, l'évaluation (que Gaudreau nomme « évaluation systématique¹⁰ ») contribue à fournir des données qui permettent aux intervenants de tirer des conclusions solides de leur processus d'évaluation.

Ces données peuvent être de deux types : quantitatives ou qualitatives. Nous entamons cette section avec des descriptions succinctes de ces deux types et nous n'aborderons que l'essentiel qui caractérise chacun, puisque ces éléments seront repris plus loin dans le document.

1.1 Données quantitatives

En ce qui concerne les données quantitatives, elles peuvent correspondre à la taille des individus, aux scores obtenus à un test de connaissances ou à tout autre résultat d'une mesure quelconque¹¹.

L'exemple suivant, un élément que l'on retrouverait dans un questionnaire de sondage, présente des données quantitatives, qui sont les chiffres se trouvant sur l'échelle de mesure allant de 1 à 4¹².

¹⁰ Louise Gaudreau, *Les indicateurs et les critères*, t. 1 d'*Évaluer pour évoluer*, Montréal, 2001, p. 16 et 17.

L'auteure énumère d'autres critères que nous devons prendre en compte quand nous évaluons. Ils seront abordés plus loin dans ce document.

¹¹ David C. Howell, *Méthodes statistiques en sciences humaines*, Paris, 1999, p. 4.

¹² David C. Howell, *Méthodes statistiques en sciences humaines*, Paris, 1999, p. 4.

« L'atelier a été offert à un moment approprié de l'année scolaire. »

1. Complètement en désaccord
2. Plutôt en désaccord
3. Plutôt d'accord
4. Complètement d'accord

1.2 Données qualitatives

Les données qualitatives sont des données brutes du monde empirique, c'est-à-dire qu'elles sont recueillies sans avoir été classées dans des catégories prédéterminées, ce qui n'est pas le cas quand nous cherchons à obtenir des données quantitatives, celles-ci se retrouvant habituellement dans des catégories telles que les réponses possibles à un test de connaissances ou à un questionnaire à questions fermées. Les données qualitatives pourvoient les intervenants de renseignements détaillés qui vont au fond des choses¹³.

La question ouverte suivante, qui pourrait être posée à des membres après qu'ils ont participé à un atelier, permettrait de récolter de telles données :

Question :

« Qu'avez-vous apprécié le plus de cet atelier? »

Réponses possibles :

« Les explications étaient claires et précises. »

« J'ai bien aimé le fait que l'on pouvait souvent poser des questions. »

Les données qualitatives se présentent sous forme d'événements, d'idées, d'histoires, etc. Elles sont obtenues à partir d'entrevues, de l'analyse de documents et par la technique de l'observation. Une fois obtenues, elles doivent être réparties en catégories¹⁴. Des fréquences pour les données dans chaque catégorie sont habituellement calculées.

À titre d'exemple, après avoir rassemblé et rangé par thèmes des données qualitatives, on pourrait faire le constat suivant : « Cinq personnes ont cité l'apprentissage, trois, la conscience et neuf, l'émotivité. » Dans cet exemple, les données qualitatives ont été regroupées en trois catégories, soit « l'apprentissage », « la conscience » et « l'émotivité ».

¹³ Inspirée de l'explication offerte par Michael Quinn Patton, *Qualitative Evaluation Methods*, Beverly Hills, 1980, p. 22.

¹⁴ Santé Canada, *Guide d'évaluation de projet : Une démarche participative*, La Direction de la santé de la population, août 1996, p. 22 — récupéré de www.hc-sc.gc.ca/hppb/violencefamiliale/html/fvprojevaluation_f.html.

2. LES INDICATEURS EN ÉDUCATION

(Pour plus de clarté, nous distinguerons, dans ce document, un indicateur de son contenu : l'indicateur est l'énoncé même, et son contenu, dressé après la collecte de données, indique un résultat, une caractéristique qui a été observée ou mesurée¹⁵.)

Dans l'introduction, il est mentionné que les indicateurs servent à nous orienter dans la sélection du type : des données à recueillir; des techniques nécessaires pour la collecte des données (quantitatives ou qualitatives); ainsi que de l'analyse des données. Un indicateur, de par sa nature, est un informateur, et son contenu a comme fonction de fournir des indications sur l'état d'un phénomène.

Plusieurs définitions d'un indicateur, concept emprunté à l'économie¹⁶, abondent dans les revues éducatives. Mais, même si la majorité des documents que nous avons consultés traitaient exclusivement les indicateurs de type quantitatif, ce document abordera aussi l'aspect qualitatif que peut prendre un indicateur.

En effet, Gaudreau affirme qu'en plus de faire usage d'indicateurs quantitatifs, on a reconnu récemment dans plusieurs domaines, incluant l'éducation, qu'un indicateur peut également être de type qualitatif¹⁷. Nadeau et son équipe de recherche, dans un projet visant à reconnaître les indicateurs de qualité pour les collèges et les universités au Canada, ont également repéré des indicateurs qualitatifs¹⁸.

2.1 Indicateurs quantitatifs

Un indicateur quantitatif se rapporte à une quantité de faits, à une mesure déterminée par tout type d'échelle. Nous reconnaissons un indicateur quantitatif par son contenu qui présente des pourcentages, des proportions, des moyennes, etc.¹⁹ Les instruments pour recueillir ces données sont des questionnaires à questions fermées, des tests, des registres statistiques, etc.

Voici un exemple d'un indicateur quantitatif :

- Le taux de réussite au cours de Calcul différentiel (201-103) pour la session d'automne.

¹⁵ C'est à partir des deux tomes de Louise Gaudreau que nous avons décidé d'appliquer pareillement cette distinction.

¹⁶ Paul-Émile Gingras, *Le concept d'« indicateur » appliqué en éducation*, coll. « L'analyse institutionnelle », Montréal, CADRE, 1977, p. 8.

¹⁷ Louise Gaudreau, *Les indicateurs et les critères*, t. 1 d'Évaluer pour évoluer, Montréal, 2001, p. 20.

¹⁸ Gilles Nadeau, *Critères et indicateurs de qualité et d'excellence dans les collèges et les universités du Canada : bilan des trois phases du projet*, Winnipeg, janvier 1995, p. 28.

¹⁹ Notre explication résume celle présentée par Louise Gaudreau, *Les indicateurs et les critères*, t. 1 d'Évaluer pour évoluer, Montréal, 2001, p. 20.

Après avoir recueilli l'information pertinente, le contenu de cet indicateur pourrait être énoncé comme suit :

- Le taux de réussite à ce cours est de 67 %.

Comme autres exemples d'indicateurs quantitatifs, mentionnons :

- Le taux d'absentéisme (= pourcentage d'absences);
- L'effectif étudiant;
- Le rendement moyen à un test de connaissances;
- Le taux de satisfaction;
- Le nombre de décrocheurs;
- Le nombre de tuteurs;
- Les coûts associés au fonctionnement du programme de tutorat.

Dans ces exemples, nous voyons que chacun de ces indicateurs révèle clairement que le type de données que nous devons récolter est quantitatif.

2.2 Indicateurs qualitatifs

Le contenu d'un indicateur qualitatif présente des opinions, des événements, des idées, soit des données représentées sous forme d'énoncés et que nous pouvons obtenir à partir de techniques comme l'observation, l'entretien, l'analyse de documents. Son contenu peut, toutefois, faire appel à des chiffres²⁰.

Comme indicateur qualitatif, nous pourrions retrouver : « les dires positifs des tuteurs qui ont participé à l'atelier de formation. »

Après qu'on a obtenu des données qualitatives et qu'elles ont été regroupées par thèmes, le contenu de l'indicateur pourrait s'écrire de la façon suivante :

- Un aspect positif de l'atelier et souligné par 75 % des participants est le temps alloué pour poser des questions.

(Ce ne serait pas le seul aspect positif.)

Les exemples suivants illustrent d'autres indicateurs qualitatifs :

- Les attitudes des étudiants envers une matière donnée;
- La perception des effets du programme de tutorat par les enseignants qui enseignent à des étudiants qui y ont participé;
- Les bénéfices affectifs des participants.

²⁰ Louise Gaudreau, *Les indicateurs et les critères*, t. 1 d'*Évaluer pour évoluer*, Montréal, 2001, p. 20 et 21.

2.3 Indicateurs pour le programme de tutorat

Larose *et al.* ont relevé, à travers les résultats de leur sondage, les indicateurs quantitatifs et les indicateurs qualitatifs (qu'eux désignent sous le nom de « variables ») les plus employés par les collèges dans le cadre de l'évaluation du programme de tutorat maître-élève²¹ :

- Échec ou réussite relativement à des cours;
- Satisfaction et utilité du tutorat;
- Moyenne générale de l'étudiant pour la session;
- Habiletés d'études;
- Habiletés sociales;
- Habiletés vocationnelles;
- Taux d'abandon des études.

Onze des douze collèges qui avaient participé au sondage avaient répondu qu'ils font usage de plus qu'un indicateur (variable) dans leur processus d'évaluation²².

Ces mêmes chercheurs croient que, dans l'évaluation d'un programme de tutorat (maître-élève), des variables (indicateurs) autres que celles rattachées aux résultats scolaires des étudiants devraient être retenues, comme des variables socioaffectives qui incluent le désir d'accomplissement par les tuteurs, le degré d'autonomie atteint par les « tutorés » (ceux qui reçoivent de l'aide), etc.²³

(Les lecteurs disposent en annexe d'une liste d'indicateurs destinés à l'évaluation de diverses mesures d'encadrement, dont le tutorat.)

²¹ Simon Larose, Nathalie Soucy et Stéphane Duchesne, « Examen des programmes de tutorat maître-élève dans les collèges du réseau québécois », *Pédagogie collégiale*, vol. 13, n° 4, mai 2000, p. 11-37 — récupéré de www.cdc.qc.ca/ped_coll/soucy_13_4.html.

²² Simon Larose, Nathalie Soucy et Stéphane Duchesne, « Examen des programmes de tutorat maître-élève dans les collèges du réseau québécois », *Pédagogie collégiale*, vol. 13, n° 4, mai 2000, p. 11-37 — récupéré de www.cdc.qc.ca/ped_coll/soucy_13_4.html.

²³ Simon Larose, Nathalie Soucy et Stéphane Duchesne, « Examen des programmes de tutorat maître-élève dans les collèges du réseau québécois », *Pédagogie collégiale*, vol. 13, n° 4, mai 2000, p. 11-37 — récupéré de www.cdc.qc.ca/ped_coll/soucy_13_4.html.

3. CHOISIR UN INDICATEUR

Nous pouvons, en éducation, avoir en main un ensemble d'indicateurs quantitatifs ou qualitatifs parmi lesquels nous voulons choisir celui ou ceux qui conviendraient pour les besoins de l'évaluation d'une mesure d'encadrement.

Gaudreau présente un tableau, *Liste de vérification des qualités des indicateurs*, qui met en évidence des critères (qu'elle nomme « qualités ») pouvant nous servir de guide dans notre sélection d'indicateurs prioritaires²⁴.

Nous pouvons déduire de ce tableau que les ressources tant financières qu'humaines dont dispose le collège, de même que le temps alloué pour la collecte et le traitement des données, doivent être des facteurs à considérer dans le choix d'un indicateur.

Nous reproduisons un extrait de ce tableau, extrait qui ne retient que les critères répondant aux objectifs de notre présent document. Une notice expliquant les moyens de se servir de ce tableau a été légèrement adaptée et accompagne l'extrait. (L'extrait ainsi que le mode d'emploi sont en annexe pour ceux qui aimeraient les reprographier.)

TABLEAU 3.1

Liste de vérification des qualités des indicateurs

INDICATEURS :				
1. Faisabilité pour colliger et traiter les données : <ul style="list-style-type: none"> • Disponibilité des sources • Facilité pour colliger et traiter les données • Faible coût pour recueillir et traiter les données • Court laps de temps pour les recueillir et les traiter • Peu de personnes nécessaires pour les colliger et les traiter 				
2. Facile de compréhension				
3. Facilement comparable aux autres				
4. Spécificité				
5. Sensibilité				
6. Validité				

Source : Louise Gaudreau, 2001, *Les indicateurs et les critères*, t. 1 d'Évaluer pour évoluer, p. 59.

²⁴ Louise Gaudreau, *Les indicateurs et les critères*, t. 1 d'Évaluer pour évoluer, Montréal, 2001, p. 32 et 33.

MODE D'EMPLOI

1. En haut, inscrire le nom des indicateurs pressentis (ou leur numéro ou leur abréviation).
2. Comparer chacun des indicateurs à chacune des qualités : inscrire
 - un « √ » sous l'indicateur vis-à-vis de la qualité qu'il possède;
 - un « + ou - » pour la qualité qu'il possède peu;
 - un « non » s'il ne possède pas cette qualité.
3. Conserver les indicateurs qui obtiennent le plus de qualités (le plus de « √ », « + ou - »), mais qui répondent aussi à un minimum de faisabilité (sinon, on ne pourra pas s'en servir).

Source : Louise Gaudreau, 2001, *Les indicateurs et les critères*, t. 1 d'Évaluer pour évoluer, p. 59.

On trouvera ci-dessous la description²⁵ de quelques-unes des qualités qui apparaissent dans le tableau, permettant ainsi d'appliquer ces qualités de façon univoque.

Qualité n° 2

Facile de compréhension : l'indicateur doit être compris même par des non-spécialistes en évaluation.

Qualité n° 4

Spécificité : l'indicateur ne mesure qu'un seul aspect à la fois; le contenu de l'indicateur se modifie seulement quand cet aspect se transforme.

Qualité n° 5

Sensibilité : l'indicateur détient cette propriété de changer très vite et beaucoup quand l'aspect qu'il contient varie.

Qualité n° 6

Validité : l'indicateur doit « représenter vraiment ce qu'il est supposé représenter²⁶ ».

²⁵ D'après Louise Gaudreau, *Les indicateurs et les critères*, t. 1 d'Évaluer pour évoluer, Montréal, 2001, p. 32.

²⁶ Louise Gaudreau, *Les indicateurs et les critères*, t. 1 d'Évaluer pour évoluer, Montréal, 2001, p. 32.

4. PLAN D'ÉVALUATION D'UNE MESURE D'ENCADREMENT

L'évaluation étant un travail comprenant de multiples tâches, tout exercice d'évaluation doit être planifié d'avance afin que nous puissions bien cerner les moyens propres à atteindre, dans des délais prévus, les visées de l'évaluation d'une mesure d'encadrement, que ce soit pour des visées de reddition de comptes, de régulation de la mesure d'encadrement, ou autres. Si cette planification n'est pas établie, l'évaluation peut facilement aller à vau-l'eau.

Le plan d'évaluation est un document qui décrit les principaux éléments du processus d'évaluation; il est un guide indispensable pour évaluer scrupuleusement toute mesure d'encadrement.

Comme éléments de ce plan, nous pouvons citer²⁷ :

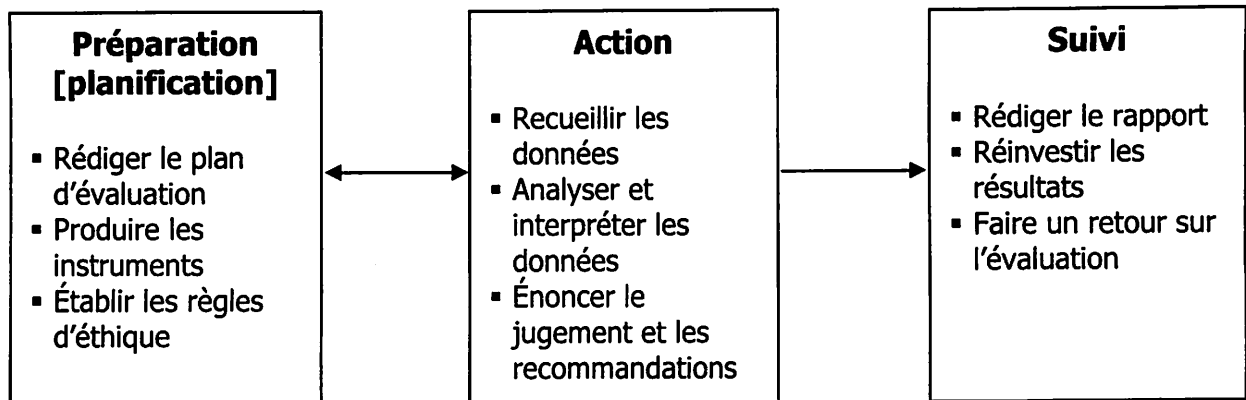
1. l'objet d'évaluation (il faut savoir ce qui est évalué : la mesure d'encadrement; et ce que l'on en évalue : ses objectifs, ses effets)
2. les indicateurs
3. la ou les questions évaluatives (pour juger la valeur d'une intervention)
4. les objectifs de l'évaluation
5. les critères²⁸, leur norme respective et leur pondération
6. la méthodologie (les sources de données, le type de données, les techniques et les instruments pour la collecte de données, les modalités d'analyse à appliquer, etc.)
7. les moyens envisagés pour l'utilisation des résultats de l'évaluation

Pour mieux situer l'étape de la rédaction de ce plan dans le processus d'évaluation d'une mesure d'encadrement, nous fournissons le schéma suivant, qui expose les trois phases principales de ce processus²⁹ :

²⁷ Louise Gaudreau, *Les étapes d'une évaluation de programme ou de projet*, t. 2 d'*Évaluer pour évoluer*, Montréal, 2001, p. 22-28.

²⁸ Un critère se définit comme un référent utilisé dans l'appréciation de renseignements extraits des indicateurs retenus (Diane Charlebois, atelier n° 805 : *Considérations pratiques sur les indicateurs de programmes*, conférence donnée à Saint-Hyacinthe lors du 24^e colloque annuel de l'AQPC, le 10 juin 2004).

²⁹ Louise Gaudreau, *Les étapes d'une évaluation de programme ou de projet*, t. 2 d'*Évaluer pour évoluer*, Montréal, 2001, p. 18.

FIGURE 4.1**La démarche générale d'une évaluation**

Source : Louise Gaudreau, 2001, *Les étapes d'une évaluation de programme ou de projet*, t. 2 d'*Évaluer pour évoluer*, Montréal, p. 18.

Ce tableau fait ressortir que la première phase du processus d'évaluation est entièrement réservée à la planification de l'évaluation; une bonne planification mène infailliblement à un bon déroulement des étapes ultérieures³⁰. Pour les six premières étapes, Gaudreau explique qu'elles ne nécessitent pas d'être franchies dans cet ordre d'une façon rigide; par exemple, pendant la collecte des données, on pourrait reprendre une étape précédente et modifier un instrument³¹.

Selon cette auteure, la rédaction du plan d'évaluation et la construction des instruments sont les étapes qui requièrent le plus de temps et le plus de connaissances techniques, mais qui, lorsqu'elles sont bien réalisées, permettent plus facilement de franchir les étapes suivantes³².

³⁰ Louise Gaudreau, *Les étapes d'une évaluation de programme ou de projet*, t. 2 d'*Évaluer pour évoluer*, Montréal, 2001, p. 22.

³¹ Louise Gaudreau, *Les étapes d'une évaluation de programme ou de projet*, t. 2 d'*Évaluer pour évoluer*, Montréal, 2001, p. 19.

³² Louise Gaudreau, *Les étapes d'une évaluation de programme ou de projet*, t. 2 d'*Évaluer pour évoluer*, Montréal, 2001, p. 19.

5. TABLEAU DE BORD DES INDICATEURS

Un autre outil, un « tableau de bord des indicateurs », peut s'avérer fort utile en évaluation. Il résulte de la première phase du processus d'évaluation, soit la préparation (voir la section précédente), et résume le plan d'évaluation composé par les responsables de l'évaluation³³. Il jouit de plusieurs utilités, dont les suivantes :

1. Considéré comme un résumé du plan d'évaluation, il sert à prévoir les opérations du processus d'évaluation et facilite ainsi le passage de la phase de préparation à la phase d'action (voir figure 4.1)³⁴;
2. Il peut également tenir lieu d'un abrégé du processus d'évaluation qui a été appliqué, en présentant des éléments essentiels tels que les indicateurs qui ont été retenus pour l'évaluation, les moyens utilisés pour récolter les données, ainsi que des conclusions découlant des analyses et des interprétations³⁵. Une fois rempli, le tableau est un document particulièrement précieux, pouvant servir de référence pour des évaluations ultérieures.

Nous présenterons prochainement une légère adaptation de ce tableau; pour que les lecteurs puissent mieux saisir les divers thèmes, nous y avons inséré des exemples, chacun représentant une situation fictive après que des données eurent été recueillies et analysées.

(Ce tableau dénué d'exemples est en annexe pour ceux qui voudraient le reproduire.)

³³ Louise Gaudreau, *Les étapes d'une évaluation de programme ou de projet*, t. 2 d'*Évaluer pour évoluer*, Montréal, 2001, p. 22.

³⁴ Louise Gaudreau, *Les étapes d'une évaluation de programme ou de projet*, t. 2 d'*Évaluer pour évoluer*, Montréal, 2001, p. 58.

³⁵ Louise Gaudreau, *Les indicateurs et les critères*, t. 1 d'*Évaluer pour évoluer*, Montréal, 2001, p. 34 et 35.

TABLEAU 5.1

Tableau de bord des indicateurs

Indicateurs	Types de données	Sources des données	Techniques et instruments pour collecte des données	Analyses	Valeurs calculées ou nature des observations obtenues	Jugements et décisions
Taux de satisfaction des tutorés ³⁶	Numérique et citations écrites	Étudiants ayant participé au programme de tutorat (au moins 9 heures ³⁷) en tant que tutorés.	Un sondage avec un questionnaire composé de 20 questions fermées (échelle allant de 1 à 6) et de 3 questions ouvertes.	Effectuer des calculs statistiques (simples) sur les réponses données aux questions fermées et catégoriser par thèmes celles des questions ouvertes.	D'après les résultats, 75 % des étudiants (tutorés) considèrent que l'inscription au programme de tutorat est trop longue et complexe. En outre, 23 % des tutorés trouvaient que leur tuteur n'était pas très patient et qu'il n'avait pas une écoute très attentive. Et 58 % se sont plaints du fait que la salle de travail était mal éclairée.	Modifier le formulaire d'inscription au programme de tutorat avant le début du prochain trimestre. Mettre plus d'accent sur des habiletés d'écoute attentive au cours de la formation des tuteurs. Trouver une autre salle de travail pouvant accueillir les tuteurs et tutorés.
Dires des participants sur l'utilité du programme de tutorat	Citations orales	Étudiants ayant participé en tant que tutorés ou tuteurs au programme de tutorat.	Cinq entretiens individuels avec des tuteurs, à l'aide d'un questionnaire renfermant quatre questions ouvertes. Six entretiens individuels avec des tutorés, à partir de l'utilisation d'un questionnaire à cinq questions ouvertes.	À partir des transcriptions des entretiens menés auprès des tutorés et des tuteurs, découper les propos en thèmes et faire des analyses.	Des entretiens avec les tuteurs, quatre ont affirmé gagner de la confiance en soi; trois ont dit éprouver désormais de la sympathie pour les enseignants; cinq ont admis mieux comprendre la matière. Des entretiens avec les tutorés, quatre disent aimer mieux recevoir des explications d'un étudiant plutôt que d'un enseignant, de peur que celui-ci ne les juge ou ne perde patience.	Afin de voir si d'autres participants au programme éprouvent des sentiments semblables, élaborer deux questionnaires (un pour les tuteurs et un pour les tutorés) à questions fermées, et enquêter auprès de tous les participants. Ce questionnaire renfermerait des questions portant sur les thèmes et propos dégagés des analyses des entretiens.

³⁶ L'expression « tutoré » est employée dans ce document pour désigner, rappelons-le, un étudiant qui reçoit de l'aide d'un tuteur.

³⁷ Il est à noter que nous avons choisi librement le nombre d'heures, mais qu'il aurait pu être repéré dans des études attestant le nombre minimal d'heures que les tutorés doivent consacrer à des activités tutorales avant qu'ils puissent en retirer des avantages.

Indicateurs	Types de données	Sources des données	Techniques et instruments pour collecte des données	Analyses	Valeurs calculées ou nature des observations obtenues	Jugements et décisions
Raisons principales pour la non-participation au programme de tutorat ³⁸	Citations écrites	Étudiants faisant partie d'une population cible (ceux dont la moyenne dans une matière donnée est faible) et ne participant pas en tant que tutorés au programme.	Sondage avec un questionnaire qui comporte : 1. une liste de raisons parmi lesquelles les répondants font un choix; 2. une section où les répondants pourraient insérer d'autres raisons; 3. une section dans laquelle les répondants peuvent donner des suggestions quant à des moyens pour améliorer la participation de cette population au programme de tutorat.	Regrouper par thèmes les données recueillies et faire les analyses appropriées.	Des répondants, 72 % ont donné comme raison principale pour leur non-participation au tutorat : « réussir aux études est ma responsabilité »; par ailleurs, 55 % ont suggéré de mieux publiciser les retombées du programme de tutorat.	Afin que les étudiants de cette population reconnaissent qu'ils ne sont pas les seuls à éprouver de la difficulté dans une matière donnée, diffuser plus d'information au sujet du programme de tutorat, notamment préciser les bénéfices que les étudiants peuvent en retirer et décrire le profil des participants y ayant pris part.
Rendement sur le plan de lecture	Numérique	Étudiants ayant des problèmes relatifs à la lecture en français	Deux tests mesurant le rendement des étudiants : un avant que la session d'activités tutorales débute et un après la session (ce dernier moment peut survenir avant la fin d'une session scolaire)	Appliquer des calculs statistiques appropriés (tel ANOVA) et comparer le rendement aux deux tests	Les scores obtenus des « posttests » ont excédé ceux des « prétests » par une moyenne de 4,5 points	Poursuivre les activités de formation des tuteurs, en insistant sur des techniques qu'ils pourraient employer pour consolider les habiletés en lecture chez les tutorés

Source : Louise Gaudreau, *Les indicateurs et les critères*, t. 1 d'*Évaluer pour évoluer*, Montréal, 2001, p. 35 et 61.

³⁸ Cet exemple est inspiré du rapport de recherche de Jocelyne Désy, *Le tutorat par les pairs tel que perçu par les élèves*, Sainte-Foy, 1996.

6. TECHNIQUES ET INSTRUMENTS DE COLLECTE DE DONNÉES

La section « Plan d'évaluation d'une mesure d'encadrement » a indiqué que la planification (préparation) d'une évaluation d'une mesure d'encadrement comprend la rédaction d'un plan d'évaluation, lequel renferme les indicateurs retenus pour l'évaluation, ainsi que la production des instruments de collecte de données.

Et, comme nous l'avons déjà fait remarquer, un indicateur nous permet de connaître *a priori* le type de données à récolter, ce qui nous oriente grandement dans le choix du type d'instruments à concevoir ou à se procurer. Si l'on veut, par exemple, mesurer le taux de satisfaction chez les participants, on n'aura guère recours à une grille d'observation. En outre, la fiabilité des données recueillies d'une évaluation dépend, entre autres choses, de la qualité des instruments utilisés. Il importe donc de bien choisir (ou construire) l'instrument qui servira à recueillir nos données.

Une multitude de techniques et d'instruments étant disponibles pour rassembler des données, nous ne présenterons ici qu'un bref aperçu de ceux qui semblent être les plus usuels en éducation.

6.1 Techniques et instruments pour recueillir des données quantitatives

Les taux de réussite à un cours, les notes aux examens de fin de session et les moyennes des étudiants à une session donnée sont toutes des données quantitatives qui peuvent être facilement obtenues pendant l'année (à partir de registres statistiques ou d'enseignants)³⁹.

L'obtention d'autres données quantitatives peut requérir la construction d'instruments spécifiques⁴⁰, tels que : un test sur les connaissances dans une discipline donnée; un questionnaire à questions fermées utilisé dans le cadre d'un sondage sur les attitudes des étudiants ayant participé à un programme de tutorat (voir section 6.5 pour des instruments spécifiques de collecte de données quantitatives). Les questionnaires d'un sondage peuvent être envoyés par la poste, diffusés sur l'intranet d'un établissement, remplis par soi-même ou utilisés à l'occasion d'un sondage téléphonique.

6.2 Techniques pour recueillir des données qualitatives

Nous avons mentionné dans l'introduction que les approches strictement quantitatives, approches souvent réservées à des fins administratives (comme celles qui incorporent l'utilisation stricte d'indicateurs ayant trait au rendement scolaire), sont parfois considérées comme insuffisantes lorsqu'on veut reconnaître d'autres avantages et utilités propres d'une mesure d'encadrement. Pour compenser cette difficulté, nous pouvons ajouter à notre processus d'évaluation une approche qualitative. Patton nous fait part de trois techniques

³⁹ Jean-Guy Blais, Michel Laurier et Guy Pelletier, « Regards sur la problématique des indicateurs en éducation », *Mesure et évaluation en éducation*, vol. 22, n^{os} 2 et 3, 1999, p. 47-69.

⁴⁰ Jean-Guy Blais, Michel Laurier et Guy Pelletier, « Regards sur la problématique des indicateurs en éducation », *Mesure et évaluation en éducation*, vol. 22, n^{os} 2 et 3, 1999, p. 47-69.

utilisées dans l'approche qualitative ainsi que du type de données qualitatives que l'on pourrait en obtenir⁴¹ :

1. Entrevues à questions ouvertes, où les données sont des citations orales obtenues directement des personnes et portant sur leurs expériences, leurs opinions, leurs sentiments et leurs connaissances.
2. Observations directes, à partir desquelles les données recueillies sont des descriptions détaillées des éléments suivants : des activités d'une intervention; des comportements des participants; des actions de la part du personnel; et d'interactions qui peuvent faire partie des expériences vécues au sein d'une intervention.
3. Documents écrits, incluant les questions ouvertes dans un questionnaire, les agendas personnels, les rapports officiels, la correspondance. Les données que l'on peut retirer de l'analyse de ces éléments peuvent être des extraits, des citations ou des passages entiers.

(Les lecteurs ont sûrement constaté que ce sont les mêmes techniques qui ont formé des éléments constituant les définitions antérieures de données qualitatives et d'indicateurs qualitatifs.)

Il existe divers types d'entrevues : des entrevues de type dyade, où le participant et la personne qui mène l'entrevue sont les seuls présents; des entrevues de groupe; des groupes de discussion. Les entrevues peuvent se dérouler par téléphone ou dans un contexte de face-à-face.

La validité et la fiabilité des données qualitatives dépendent en grande partie des compétences et de la formation de l'évaluateur, ainsi que de la rigueur de la méthodologie employée dans l'évaluation⁴².

Mais les données quantitatives et les données qualitatives ne sont pas recueillies exclusivement à partir des techniques que nous avons décrites : des données qualitatives peuvent être recueillies dans des recherches dites expérimentales⁴³, et des données quantitatives peuvent être obtenues au moyen d'approches naturalistes, comme les entrevues. Néanmoins, des recherches expérimentales sont essentiellement menées pour des analyses statistiques de données quantitatives, et le but de mener des entrevues consiste principalement à récolter des données qualitatives⁴⁴.

⁴¹ Michael Quinn Patton, *How to use Qualitative Methods in Evaluation*, 6th Printing, Newbury Park, 1990, p. 7 – traduction libre de Nathalie Prévost.

⁴² Michael Quinn Patton, *How to use Qualitative Methods in Evaluation*, 6th Printing, Newbury Park, 1990, p. 8 – traduction libre de Nathalie Prévost.

⁴³ Dans un schème expérimental, les sujets sont répartis en deux (ou plusieurs) groupes selon des techniques aléatoires. L'un des groupes est soumis à un traitement quelconque, alors que l'autre sert de groupe de contrôle. Ce schème favorise la formation de groupes équivalents avant une expérimentation.

⁴⁴ Michael Quinn Patton, *How to use Qualitative Methods in Evaluation*, 6th Printing, Newbury Park, 1990, p. 14 – traduction libre de Nathalie Prévost.

6.3 Quantitatif ou qualitatif?

Dans une approche qualitative, on cherche à savoir comment des gens perçoivent une situation et quel sens ils y accordent; on y parvient en analysant des documents (lettres, procès-verbaux, etc.), en donnant la chance à un petit nombre de personnes de manifester ouvertement et amplement leur pensée à l'égard d'une situation donnée, etc. Tandis qu'avec une approche quantitative, on recherche plutôt à quantifier des phénomènes et à analyser un grand nombre de données numériques⁴⁵.

Mais quelle approche, quantitative ou qualitative, prime en évaluation? Certes, cette question a fait l'objet de plusieurs débats en évaluation. À ce sujet, Lapan déclare que, même si les partisans de l'approche qualitative prétendent que des descriptions d'un phénomène qui vont au fond des choses sont supérieures à des données quantitatives, tels des scores à un test, et que ceux de l'autre approche avancent que ce sont ces scores et autres résultats numériques qui fournissent plus d'évidences objectives des effets d'un programme (comme le définit Plante; voir note en bas de la page 10), tous ces partisans reconnaissent maintenant que chaque approche occupe une place importante dans l'évaluation, et que les meilleures évaluations sont celles qui incorporent les deux⁴⁶.

Pour abonder dans le même sens que cet auteur, nous appuyons le fait qu'une diversification de techniques peut fournir aux intervenants un ensemble plus complet des avantages et des inconvénients rattachés à une mesure d'encadrement. Bien entendu, cette multiplicité peut entraîner avec elle des dépenses substantielles en matière d'instruments et d'analyse.

Toutefois, un même questionnaire qui recueillerait des données quantitatives et des données qualitatives pourrait se qualifier comme un instrument qui intègre les deux approches, ne serait pas trop coûteux et permettrait effectivement aux intervenants de dresser un meilleur portrait des mérites et des écueils d'une mesure d'encadrement.

Mais ce ne sont pas toutes les situations qui se prêtent facilement à ces deux approches. Ainsi, d'après Patton, pour certains phénomènes, telle la créativité, il peut s'avérer difficile d'utiliser un instrument (comme un questionnaire) ayant pour but de recueillir des données quantitatives sur le phénomène, alors que des données qualitatives seraient plus faciles à recueillir. Il soutient que, lorsqu'un instrument de mesure n'a pas été conçu de façon rigoureuse et est fabriqué avec une échelle de mesure dont la validité et la fiabilité sont douteuses, il vaudrait mieux recueillir de l'information descriptive moyennant une approche qualitative⁴⁷.

⁴⁵ Sylvain Giroux et Ginette Tremblay, *Méthodologie en sciences humaines — La recherche en action*, 2^e éd., coll. « ERPI Sciences humaines », Montréal, 2002, p. 23.

⁴⁶ Stephen D. Lapan, *Foundations of Research: Methods of Inquiry in Education and the Social Sciences*, Northern Arizona University [www4.nau.edu/cee/ci_doc/current/resources/13_lapan.pdf] — traduction libre de Nathalie Prévost.

Au fait, Stufflebeam partage aussi cet avis.

⁴⁷ Michael Quinn Patton, *How to use Qualitative Methods in Evaluation*, 6th Printing, Newbury Park, 1990, p. 37 — traduction libre de Nathalie Prévost.

6.4 Avantages et inconvénients de quelques techniques

Toute technique de collecte de données, qu'elle soit qualitative ou quantitative, comporte des avantages ainsi que des inconvénients; la connaissance de ces éléments peut nous aider :

1. à mieux sélectionner la ou les techniques à retenir dans une évaluation;
2. à prévoir les difficultés qui seront engendrées par l'utilisation d'une technique spécifique.

Le tableau suivant⁴⁸ fait état de ces éléments liés à des techniques qui s'associent à l'une des deux approches (mais pas nécessairement de façon exclusive) : à la recension des écrits (approche qualitative), à l'entrevue (approche qualitative) et à la passation d'un questionnaire dans le cadre d'un sondage (approche quantitative).

C'est aux lecteurs d'examiner les éléments présentés et d'y appliquer leurs propres réflexions quant à la meilleure technique des trois, en retenant toutefois deux principes :

- Que les meilleurs procédés d'évaluation sont ceux qui incorporent plus qu'une technique;
- Que le choix d'une technique repose non seulement sur le contenu de ce tableau, mais également sur des facteurs étant parallèles à ceux qui influent sur le choix des indicateurs (voir la section 3 sous le critère de la faisabilité).

⁴⁸ Tiré intégralement du document *Évaluation institutionnelle* par Jacques Plante, juillet 1994, p. 82 et 83, « document inédit ».

TABLEAU 6.1**Avantages et inconvénients des techniques**

Technique	Avantages	Inconvénients
La recension des écrits ⁴⁹	<ul style="list-style-type: none"> ▪ Permet une analyse en profondeur ▪ Couvre des aspects divers ▪ Se prête à des analyses répétées ▪ Se fait facilement dans un même lieu 	<ul style="list-style-type: none"> ▪ On ne peut rien y ajouter ▪ Demande beaucoup de temps ▪ Difficile à analyser ▪ Sensible aux intérêts de l'analyste ▪ Beaucoup de documents se révèlent peu pertinents ▪ Données collectées difficiles à traiter statistiquement ou qualitativement ▪ Difficile d'estimer à l'avance les coûts
L'entrevue à questions ouvertes	<ul style="list-style-type: none"> ▪ Contact personnalisé ▪ Permet d'étudier l'essentiel d'un phénomène ▪ Se prête à une grande flexibilité ▪ Permet de s'adapter à la personne et à la situation ▪ Permet de réinterroger pour obtenir des précisions ▪ Questions relativement faciles à formuler ▪ Questions faciles à poser ▪ La liberté d'expression est favorisée 	<ul style="list-style-type: none"> ▪ Nombre de questions limité ▪ Aspects mesurés peu nombreux ▪ Personne à personne ou petit groupe seulement ▪ Favorise la verbalisation qui risque d'être vide de sens ou étrangère à la question ▪ Difficile de tout noter ▪ Données difficiles à analyser et à traiter ▪ Laisse place à la subjectivité de la part de la personne qui analyse et interprète les réponses ▪ Peut se révéler coûteuse ▪ L'enquête peut être influencé par la personne qui interroge ▪ Exigeant de la part de l'enquête ▪ Validité et fiabilité peuvent être remises en question
Le questionnaire fermé utilisé dans le cadre d'un sondage	<ul style="list-style-type: none"> ▪ Peut être utilisé pour mesurer à peu près n'importe quoi ▪ Facile d'accroître sa validité et sa fiabilité ▪ Facile à répondre ▪ Réponses faciles à traiter statistiquement et à analyser ▪ Peut s'adresser à un grand nombre de personnes dans un délai donné ▪ Économique ▪ Les questions peuvent être nombreuses ▪ Questions courtes, réponses courtes ▪ Se prête à une communication rapide des résultats ▪ Peut couvrir plusieurs sujets à la fois 	<ul style="list-style-type: none"> ▪ Questions difficiles à formuler ▪ Réduit la liberté d'expression des enquêtés ▪ Favorise les réponses mécaniques ▪ Difficile de s'assurer que toutes les questions essentielles sont posées ▪ Laisse souvent place à des questions d'intérêt particulier à certaines personnes plutôt qu'à l'enquête

⁴⁹ La recension des écrits se réfère aux archives, statistiques publiées, banques de données, journaux, documents personnels, annuaires et bottins (Plante, p. 69).

6.5 Instruments de collecte de données répertoriés par Monaghan et Chaloux

Des instruments pour recueillir des données (plutôt quantitatives) ne sont pas toujours faciles à concevoir, et peuvent exiger de la part des intervenants beaucoup d'efforts et de temps; nous n'avons qu'à penser à ceux qui mesurent chez un étudiant des dimensions qui touchent la motivation ou la confiance en soi. Pour surmonter cette difficulté, des établissements ont recours à des instruments qui sont préfabriqués, et les utilisent tels quels ou les modifient.

Or, il existe une grande variété d'instruments possibles pour la mesure d'indicateurs en éducation. Le tableau suivant⁵⁰ expose divers outils de mesure, leur description ainsi que les indicateurs (que les chercheurs ont nommés « variables ») servant en tant que facteurs mesurables par l'outil correspondant. Certains de ces outils constituent des instruments diagnostiques, comme l'IAP et le Processus Résultats Plus. Les auteurs de ce rapport de recherche affirment que tous les instruments inscrits dans ce tableau sont des instruments valides⁵¹.

Sans insister sur les détails, nous voulons toutefois mentionner que, si nous utilisons des instruments déjà construits, nous devons en connaître au moins les modalités d'administration de manière à reproduire les mêmes circonstances que celles dans lesquelles l'instrument s'est avéré utile.

⁵⁰ Extrait intégral du rapport de recherche *Évaluation d'un programme d'aide à la réussite*, une recherche subventionnée par le ministère de l'Éducation dans le cadre du Programme d'aide à la recherche sur l'enseignement et l'apprentissage (PAREA) et réalisée par Denis Monaghan et Natalie Chaloux du Cégep de Sainte-Foy. Ce rapport, ainsi que plusieurs autres, est disponible sur le site du CDC (Centre de documentation collégiale) à l'adresse suivante : http://www.clarendeau.qc.ca/ceg_andre/Accueil/affichage.asp?C=0_07_04_&N=2.

⁵¹ Denis Monaghan et Natalie Chaloux, *Évaluation d'un programme d'aide à la réussite*, Cégep de Sainte-Foy, 2004, p. 11.

TABLEAU 6.2
Liste d'instruments

Outils	Description	Variable(s) mesurée(s)
Career Decision Scale (CDS) (Osipow, 1976)	Outil (19 items) qui mesure la réaction au choix de carrière	<ul style="list-style-type: none"> ◆ Confusion ◆ Hésitation ◆ Diversité ◆ Méconnaissance de soi ◆ Contraintes ◆ Niveau d'indécision ◆ Niveau de certitude
College Student Inventory (Stratil, 2002)	Questionnaire conçu spécifiquement pour les nouveaux arrivants au collégial (première session) et composé de 194 items répartis sous 19 échelles	<ul style="list-style-type: none"> ◆ Motivation scolaire ◆ Motivation sociale ◆ « Coping Skills » ◆ Réceptivité aux services d'aide ◆ Impression initiale
Échelle de motivation en éducation (EME) (Vallerant <i>et al.</i> , 1993)	Questionnaire (20 items) qui permet de percevoir les raisons (motivations) que les étudiants se donnent afin de poursuivre leurs études	<ul style="list-style-type: none"> ◆ Types de motivation scolaire (quatre sous-échelles) ◆ Amotivation scolaire
Étudiants Plus (Cliche, 1999)	Questionnaire (64 items) qui permet de cerner les perceptions et les comportements de l'étudiant à l'égard des stratégies d'enseignement et d'apprentissage	<ul style="list-style-type: none"> ◆ Conceptions de l'apprentissage (bonnes ou mauvaises) de l'étudiant
Inventaire de réseau social (IRS) (Trickett et Peri, traduit et adapté par Larose et Roy, 1989)	L'IRS n'est pas un test psychométrique. Il s'agit plutôt d'un outil descriptif qui rend compte des différentes dimensions du réseau social d'un individu	<ul style="list-style-type: none"> ◆ Nombre de personnes dans le réseau social ◆ Statut et niveau scolaire des personnes avec qui l'élève interagit ◆ Fonctions d'aide (ex. : scolaire, sociale) ◆ Sens de l'aide (j'aide l'autre ou je suis aidé par l'autre) et ramifications présentes dans le réseau social au cégep (indice de densité)

Outils	Description	Variable(s) mesurée(s)
Inventaire d'acquis précollégiaux (IAP) (Larose et Roy, 2001)	Outil qui mesure trois dimensions de la réalité des étudiants de niveau collégial : les acquis liés à l'apprentissage (9 échelles), les acquis liés à l'orientation (3 échelles) et les acquis sociaux (5 échelles)	<ul style="list-style-type: none"> ◆ Réactions d'anxiété ◆ Anticipation de l'échec ◆ Préparation aux examens ◆ Recours à l'aide du professeur ◆ Qualité d'attention ◆ Entraide ◆ Priorité aux études ◆ Croyance dans la facilité ◆ Croyance dans les méthodes ◆ Aspirations scolaires ◆ Clarté des choix d'orientation ◆ Engagement à l'égard des choix ◆ Anxiété sociale ◆ Perception du soutien social ◆ Leadership ◆ Implication sociale ◆ Expérience du secondaire
Mesure d'évaluation du soutien interpersonnel (MESI) (Cohen et Hoberman, 1983)	Mesure d'évaluation du soutien interpersonnel développée spécifiquement pour la population des élèves du collège. Elle évalue quatre fonctions du soutien social.	<ul style="list-style-type: none"> ◆ Soutien social
Processus Résultats Plus (Thivierge et Carboneau, 1998)	Questionnaire de 85 items traitant des composantes cognitives, personnelles et affectives susceptibles d'influencer l'apprentissage suite à une évaluation	<ul style="list-style-type: none"> ◆ Attribution causale ◆ Traitement de l'information en classe ◆ Motivation à la matière ◆ Motivation aux études ◆ Traitement en surface de l'information lors de l'étude ◆ Traitement en profondeur de l'information lors de l'étude ◆ Stress lors de l'examen ◆ Traitement de l'information lors de l'examen ◆ Contexte personnel ◆ Temps d'étude ◆ Traitement de l'information en contextes pratiques
Questionnaire d'identification des perceptions individu environnement (QIPIE) (St-Louis et Vigneault, 1984)	Sous-échelles regroupées dans trois dimensions : la connaissance de soi, le monde du travail et l'implication dans les études	<ul style="list-style-type: none"> ◆ Intérêts, capacités, choix et estime de soi ◆ Sens du travail, perception du marché du travail et connaissance du travail ◆ Activité, utilité et appréciation

Outils	Description	Variable(s) mesurée(s)
Questionnaire sur les attributions causales en situations scolaires (QACSS) (Barbeau, 1991)	Questionnaire de douze situations hypothétiques (six positives – six négatives) traitant, en nombre égal, de situations de performance et de situations impliquant des relations interpersonnelles	♦ Système d'attributions causales (interne/externe, stable/modifiable, globale/spécifique) selon le champ étudié (performance scolaire)
Questionnaire sur l'estime de soi de Rosenberg (Rosenberg, 1965)	Questionnaire de 10 items mesurant l'estime de soi globale	♦ Estime de soi
Student Adaptation College Questionnaire (SACQ) (Baker et Siryk, 1989) Instrument validé en français (Larose et Roy)	Outil qui mesure spécifiquement l'ajustement des élèves au contexte des études collégiales (67 items) sous quatre dimensions	♦ Ajustement scolaire, ajustement social, ajustement personnel-émotif et attachement à l'institution
Test mesurant les sources et les indicateurs de la motivation scolaire (TSIMS) (Barbeau, 1995)	Questionnaire (65 items) mesurant deux sources et deux indicateurs de la motivation scolaire	♦ Perception de compétence, perception de l'importance de la tâche, engagement cognitif et participation

Source : Denis Monaghan et Natalie Chaloux, *Évaluation d'un programme d'aide à la réussite*, Cégep de Sainte-Foy, 2004, Annexe 1.

Nous joignons à ce tableau des sites Web où les lecteurs intéressés peuvent trouver des renseignements sur l'achat et sur l'utilité de quelques instruments.

1. Career Decision Scale par Osipow — www.parinc.com/product.cfm?ProductID=34
2. College Student Inventory par Stratil — www.collegecounseling.org/members/login/publications/newsletter_old/csiexperience.html
<http://buros.unl.edu/buros/jsp/reviews.jsp?item=16032591>
3. Student Adaptation College Questionnaire par Baker et Siryk — <http://wbarratt.indstate.edu/dragon/saroi/sa-sacq.htm>

6.6 Techniques et instruments employés dans l'évaluation du programme de tutorat

Rappelons que, pour des considérations telles qu'elles sont énoncées dans l'introduction, nous avons choisi d'illustrer dans un contexte familier des notions relativement à l'évaluation du programme de tutorat, à savoir le tutorat par les pairs et le tutorat maître-élève, étant donné que c'est une mesure d'encadrement couramment mise en place dans le réseau collégial.

On trouvera dans cette section un bref inventaire des moyens utilisés par des établissements ou par des chercheurs pour évaluer l'un ou l'autre des deux modèles de tutorat. Les moyens que nous présentons ne concernent pas tous des programmes de tutorat offerts au collégial; nous avons aussi inséré ceux des programmes de tutorat offerts à d'autres ordres d'enseignement afin de démontrer la diversité des techniques empruntées dans l'évaluation de ce programme.

Or, les rapports de recherche ainsi que les documents que nous avons consultés révèlent que les évaluateurs adoptent une approche quantitative ou une approche qualitative⁵² ou une approche mixte, et que ceux-ci ont recours à des moyens appropriés selon l'approche, des moyens qui contribuent à repérer l'information générale sur les participants, des mérites et des écueils du programme de tutorat, des dimensions du programme à améliorer, à abandonner ou à conserver, des effets prévus ou imprévus, etc. Autrement dit, non seulement une évaluation de produit mais aussi une évaluation d'autres composantes du modèle de Stufflebeam (voir section 7) est conduite par les établissements ou organismes présentés dans cette section.

Nous voulons aussi faire remarquer, en passant, que des éléments peuvent faire obstacle à l'évaluation du programme de tutorat. En effet, certains chercheurs critiquent les établissements éducatifs pour le manque de dispositifs mis en place pour récupérer des données pertinentes, comme dans le cas de rencontres entre le tuteur et le tuteuré qui ne sont pas documentées. Il va sans dire qu'il est difficile d'adopter une méthodologie méticuleuse d'évaluation lorsque ces conditions se retrouvent dans les milieux collégiaux.

6.6.1 Groupe de recherche sur le tutorat (GRT)

Dans leur étude du tutorat par les pairs offert dans différents secteurs (scolaire, universitaire, professionnel, parascolaire telle l'aide aux devoirs), les chercheurs du Groupe de recherche sur le tutorat (GRT) se servent, de façon individuelle ou simultanée, des moyens suivants : des enquêtes par questionnaires, des entretiens avec les acteurs du tutorat, des observations ethnographiques (groupes humains) ainsi que des schèmes expérimentaux⁵³.

(Les tuteurs sont des non-professionnels du domaine de l'éducation ou de la formation, et sont, en général, des personnes plus âgées que les tuteurés.)

6.6.2 AmeriCorps Tutoring Outcomes Study

Les évaluateurs ont employé l'approche intégrant l'utilisation de prétests et de posttests en vue de constater si le rendement d'étudiants qui ont pris part à des activités tutorales en lecture s'était amélioré ou non⁵⁴.

⁵² Voir section 6.3 qui expose des différences entre l'approche qualitative et l'approche quantitative.

⁵³ Groupe de recherche sur le tutorat (GRT) — récupéré de www.u-bordeaux2.fr/grt/page1.html.

⁵⁴ Abt. Associates Inc., *AmeriCorps Tutoring Outcomes Study, Executive Summary*, 2001, p. 2 — récupéré de www.americorps.org.

6.6.3 Les chercheurs Larose, Soucy et Duchesne

D'après les résultats du sondage qu'ont mené les chercheurs Larose *et al.*, certains collègues opteront pour une évaluation du programme de tutorat maître-élève en employant une méthode de type qualitatif, d'autres, pour une évaluation en employant à la fois des méthodes quantitatives et des méthodes qualitatives⁵⁵.

Parmi ces méthodes, certains collègues se sont limités à ne recueillir des données qu'à la fin du programme au moyen, par exemple, d'un questionnaire destiné à connaître la satisfaction de l'étudiant, la perception de l'utilité du programme ou autres objets d'intérêt; d'autres se sont assurés, en plus d'obtenir des données à la fin d'une session, de recueillir des données au début de la session, tels les antécédents scolaires et le test d'inventaire des acquis précollégiaux (IAP)⁵⁶.

De plus, deux collègues ont opté pour former un groupe contrôle : un collègue par une méthode de distribution aléatoire des étudiants, et l'autre par une méthode de distribution non aléatoire⁵⁷. Les résultats de ce sondage ont indiqué aux auteurs que très peu de collègues prennent les moyens pour former des groupes contrôles avant que les activités tutorales ne débutent, quoique plusieurs d'entre eux appliquent une méthode d'évaluation⁵⁸.

Les résultats ont également fait ressortir que l'évaluation des retombées du programme de tutorat sur les tuteurs n'est pas une activité habituelle, et que, lorsqu'elle est réalisée, elle vise à connaître leur taux de satisfaction ainsi que leurs sentiments d'accomplissement et de compétence⁵⁹.

6.6.4 Ehly et Larsen

Certaines activités de tutorat sont dotées d'objectifs très particuliers et sont définies avant le déroulement du programme. Par exemple, dans un des multiples modèles de tutorat aux États-Unis rapporté par Ehly et Larsen, une enseignante du primaire avait assigné des objectifs généraux et précis à un élève qui recevrait de l'aide d'un autre élève.

⁵⁵ Page 15 du texte original de Simon Larose, Nathalie Soucy et Stéphane Duchesne, « Examen des programmes de tutorat maître-élève dans les collèges du réseau québécois », *Pédagogie collégiale*, vol. 13, n° 4, mai 2000, p. 11-37 — récupéré de www.cdc.qc.ca/ped_coll/soucy_13_4.html.

⁵⁶ Page 15 du texte original de Simon Larose, Nathalie Soucy et Stéphane Duchesne, « Examen des programmes de tutorat maître-élève dans les collèges du réseau québécois », *Pédagogie collégiale*, vol. 13, n° 4, mai 2000, p. 11-37 — récupéré de www.cdc.qc.ca/ped_coll/soucy_13_4.html.

⁵⁷ Page 15 du texte original de Simon Larose, Nathalie Soucy et Stéphane Duchesne, « Examen des programmes de tutorat maître-élève dans les collèges du réseau québécois », *Pédagogie collégiale*, vol. 13, n° 4, mai 2000, p. 11-37 — récupéré de www.cdc.qc.ca/ped_coll/soucy_13_4.html.

⁵⁸ Page 16 du texte original de Simon Larose, Nathalie Soucy et Stéphane Duchesne, « Examen des programmes de tutorat maître-élève dans les collèges du réseau québécois », *Pédagogie collégiale*, vol. 13, n° 4, mai 2000, p. 11-37 — récupéré de www.cdc.qc.ca/ped_coll/soucy_13_4.html.

⁵⁹ Page 16 du texte original de Simon Larose, Nathalie Soucy et Stéphane Duchesne, « Examen des programmes de tutorat maître-élève dans les collèges du réseau québécois », *Pédagogie collégiale*, vol. 13, n° 4, mai 2000, p. 11-37 — récupéré de www.cdc.qc.ca/ped_coll/soucy_13_4.html.

Un des objectifs précis était le suivant : « l'élève devrait être capable de prononcer des mots qui débutent avec une consonne particulière. » L'enseignante a évalué le tutoré après six séances de tutorat, et les résultats démontraient qu'il avait maîtrisé la compétence⁶⁰.

6.6.5 Désy

Dans une recherche qui privilégiait une approche qualitative et qui n'était pas menée dans le but de vérifier une hypothèse de recherche, Désy, en recourant à des entrevues individuelles semi-structurées, a cherché à récupérer les raisons pour lesquelles :

- des étudiants se sont inscrits au Tandem (un centre d'aide à l'apprentissage) en tant que tuteurs ou tutorés;
- des étudiants faisant partie d'une population en possession d'un dossier scolaire faible ne se sont pas inscrits au Tandem.

C'est par une méthode non probabiliste suivant quelques critères que les sujets ont été sélectionnés par la chercheuse. Une méthode probabiliste n'était pas exigée, car Désy ne cherchait pas à généraliser ses résultats, mais plutôt à se faire une idée claire des causes et des raisons de la participation ou non d'étudiants au Tandem. Elle a mené une seule entrevue d'une durée d'environ 50 minutes avec chaque étudiant participant à l'étude, pour laquelle un montant de 5 \$ a été versé à chacun⁶¹.

6.6.6 Cyrenne et Lacombe

Les chercheurs Cyrenne et Lacombe ont pu isoler les effets imputables au seul programme de tutorat (maître-élève) dans le cadre d'une recherche où sont intervenus plusieurs groupes expérimentaux et un groupe contrôle⁶².

⁶⁰ Stewart W. Ehly et Stephen C. Larsen, *Peer Tutoring for Individualized Instruction*, Boston, 1980, p. 5 — traduction libre de Nathalie Prévost.

⁶¹ Jocelyne Désy, *Le tutorat par les pairs tel que perçu par les élèves*, Sainte-Foy, 1996.

⁶² Rapporté par Diane Charlebois, *Analyse des activités d'encadrement — Rapport final*, 1998, p. 17.

À titre d'information supplémentaire, Charlebois présente dans ce même rapport une analyse de l'efficacité d'une mesure d'encadrement mise en place au Collège Édouard-Montpetit; c'est une évaluation qu'elle a réalisée au moyen d'une approche intégrant des analyses quantitatives et grâce à l'utilisation d'indicateurs (de produit) liés au rendement scolaire (Charlebois, p. 106).

7. TYPOLOGIE RETENUE

Quand vient le temps d'évaluer une mesure d'encadrement, l'obligation de rendre des comptes peut exiger que les intervenants se préoccupent uniquement de la moyenne à un cours en particulier, du taux de réussite à un cours ou à une session donnée afin de voir s'il y a eu un accroissement (significatif) ou non du contenu de ces indicateurs.

Et, comme nous l'avons indiqué dans l'introduction, certains jugent que ces méthodes sont trop limitées pour signaler d'autres conséquences ou effets de l'encadrement donné aux étudiants.

Toutefois, si nous voulons assurer un bon fonctionnement d'une mesure d'encadrement, nous ne pouvons nous limiter aux réalisations de celle-ci. En effet, l'évaluation peut, en plus, aller au-delà d'une recherche d'effets : elle peut servir à orienter les décisions que nous devons prendre pour améliorer la planification, la gestion et la mise en œuvre d'une mesure d'encadrement.

Or, c'est en réponse à diverses finalités que peut viser une évaluation que nous avons opté pour une typologie d'indicateurs qui traitent, en plus des effets prévus ou imprévus (les réalisations), d'autres aspects liés à une mesure d'encadrement⁶³.

Cette typologie retient quatre catégories d'indicateurs, à savoir⁶⁴ :

1. les indicateurs de contexte;
2. les indicateurs d'intrants;
3. les indicateurs de processus;
4. les indicateurs de produit (le rendement des étudiants à un cours donné constitue un exemple classique d'un indicateur de produit).

Elle est basée sur un modèle d'évaluation ayant été rendu populaire par Stufflebeam et retenu, par la suite, par d'autres auteurs⁶⁵.

Ce modèle de référence de Stufflebeam, que celui-ci nomme « modèle C.I.P.P. », sert de fondement à cette typologie d'indicateurs, et englobe et met en relation quatre catégories d'évaluation, telles⁶⁶ :

- l'évaluation de contexte;

⁶³ Par ailleurs, ceux qui aimeraient faire un tour d'horizon plus large des typologies en éducation sont invités à lire le rapport de recherche, *Les Indicateurs de l'enseignement supérieur*, de Dallaire, Demers et Lescop, 2003. Cette recherche a été réalisée pour le Carrefour de la réussite au collégial.

⁶⁴ *Programme d'indicateurs de la qualité en éducation (PIQÉ) — Le cadre* — récupéré de http://www.eqao.com/pdf_f/01/01P028f.pdf.

Par contre, les auteurs ont substitué le terme « résultats » au terme « produit ».

⁶⁵ Jean-Guy Blais, Michel Laurier et Guy Pelletier, « Regards sur la problématique des indicateurs en éducation », *Mesure et évaluation en éducation*, vol. 22, n^{os} 2 et 3, 1999, p. 47-69.

⁶⁶ Daniel L. Stufflebeam et Anthony J. Shinkfield, *Systematic Evaluation: a self-instructional guide to theory and practice*, coll. « Evaluation in Education and Human Services », Boston, 1985, p. 165-169.

- l'évaluation d'intrants;
- l'évaluation de processus;
- l'évaluation de produit.

Ce n'est pas seulement en éducation que des études s'appuient sur le modèle de Stufflebeam; celui-ci déclare que son modèle jouit d'une renommée internationale et est employé dans divers domaines, autant pour des recherches à long terme que pour des recherches à court terme, de petite ou de grande envergure⁶⁷.

Et, puisque, à la lecture des études qui ont intégré le modèle C.I.P.P., des approches qualitatives ainsi que des approches quantitatives ont constitué les moyens pris par des chercheurs dans l'évaluation d'un programme éducatif, nous pouvons confirmer la maniabilité du modèle de Stufflebeam et sa capacité à servir en tant que cadre de référence pour l'évaluation de toute unité, que ce soit une activité, un projet éducatif, un programme comme le tutorat ou un programme d'études, etc. (C'est pour cette raison que les notions présentées dans cette section peuvent s'appliquer à une mesure d'encadrement, même si les textes de Stufflebeam concernent un programme.)

Bien entendu, d'autres modèles sont également utilisés en évaluation : le modèle de Tyler, axé sur les objectifs d'une mesure d'encadrement et favorisant l'usage de prétests et de posttests; le modèle de Guba, préconisant une approche naturaliste par l'utilisation de méthodes qualitatives; et d'autres encore.

Mais, comme les collèges visent diverses finalités quand ils évaluent leurs mesures d'encadrement, nous voulions nous procurer un modèle qui répondrait à une multitude de besoins relativement à l'évaluation des mesures mises en place dans le réseau collégial.

Fins sommatives et fins formatives

Le modèle C.I.P.P. peut servir à évaluer des mesures d'encadrement pour des fins sommatives ou pour des fins formatives. Qu'elle soit dirigée vers le contexte, les intrants, le processus ou le produit, l'évaluation dans l'optique de ce modèle est^{68, 69} :

1. formative quand le but consiste à guider les efforts et à prendre des décisions qui exerceront une action sur l'unité faisant l'objet de l'évaluation (comme le programme de tutorat); elle est alors orientée vers l'avenir (prospective);
2. sommative quand elle cherche, dans une perspective d'« accountability », à rendre des comptes, c'est-à-dire qu'elle jette un regard sur le passé et récapitule la valeur,

⁶⁷ Daniel L. Stufflebeam, Western Michigan University, *The CIPP Model for Evaluation*, 2003, p. 2, conférence donnée par Stufflebeam lors du *2003 Annual Conference of the Oregon Program Evaluators Network (OPEN)* et pendant laquelle il présente une dernière actualisation de son modèle. (La conférence peut être récupérée de <http://www.wmich.edu/evalctr/pubs/CIPP-ModelOregon10-03.pdf>.)

⁶⁸ Daniel L. Stufflebeam et Anthony J. Shinkfield, *Systematic Evaluation: a self-instructional guide to theory and practice*, coll. « Evaluation in Education and Human Services », Boston, 1985, p. 163-169.

⁶⁹ Daniel L. Stufflebeam, Western Michigan University, *The CIPP Model for Evaluation*, 2003, p. 5-7, conférence présentée lors du *2003 Annual Conference of the Oregon Program Evaluators Network (OPEN)*.

l'importance et les mérites d'une unité (rétrospective). (Le bilan des plans de réussite que les collègues ont soumis à la CEEC au mois de décembre 2003 évoque la finalité sommative à laquelle peut tendre l'évaluation d'une mesure d'encadrement.)

Plus les évaluateurs conduisent des évaluations formatives efficaces et les documentent, plus ils sont en possession d'éléments essentiels pour non seulement enrichir la prise de décision, mais rendre facilement compte, en réponse à une interpellation, de la valeur d'une mesure d'encadrement⁷⁰.

L'évaluation (formative ou sommative) s'appuyant sur ce modèle nous permet d'apporter des réponses aux questions suivantes⁷¹ :

- Contexte
 - Quelles sont les priorités (ou les objectifs) de la mesure d'encadrement (formative)?
 - Est-ce qu'elles rejoignent les raisons pour lesquelles les étudiants y participent (sommative)?
- Intrants
 - Quels sont les stratégies, les ressources, les besoins en personnel qui sont nécessaires pour la mise en place d'une mesure d'encadrement (formative)?
 - Comment se comparent-ils avec les stratégies, les plans et le budget d'autres collègues (compétiteurs) qui offrent les mêmes activités interventionnistes (sommative)?
- Processus
 - D'après des rétroactions (feedback) périodiques sur le déroulement d'une mesure d'encadrement, quels sont les meilleurs moyens pour implanter cette mesure (formative)?
 - Comment les ressources ont-elles été déployées? Est-ce qu'elles ont été utilisées comme il était prévu (sommative)?
- Produit
 - En se référant aux résultats obtenus, est-ce que la mesure d'encadrement devra subir des modifications, ou devra-t-elle être supprimée (formative)?
 - Quels sont les résultats — positifs et négatifs ainsi que prévus et imprévus? Comment la valeur, les avantages et l'utilité des réalisations d'une mesure d'encadrement ont-ils été jugés par les participants (sommative)?

⁷⁰ Daniel L. Stufflebeam, Western Michigan University, *The CIPP Model for Evaluation*, 2003, p. 5, conférence présentée lors du *2003 Annual Conference of the Oregon Program Evaluators Network (OPEN)*.

⁷¹ Daniel L. Stufflebeam et Anthony J. Shinkfield, *Systematic Evaluation: a self-instructional guide to theory and practice*, coll. « Evaluation in Education and Human Services », Boston, 1985, p. 159-165.

Nous pouvons traiter une seule de ces quatre composantes formant le modèle C.I.P.P. ou évaluer, en une seule tentative, des combinaisons de ces composantes⁷². Un même sondage qui vise à repérer des renseignements sur le contexte, les intrants et le processus constituerait un moyen d'évaluer simultanément des éléments à chacune de ces trois composantes.

Mais que l'on évalue les intrants, le produit, le contexte ou le processus, ou des combinaisons, le processus d'évaluation comprendra toujours les étapes principales que nous avons présentées à la section 4 (Plan d'évaluation d'une mesure d'encadrement) :

- Rédiger le plan d'évaluation;
- Produire les instruments;
- Établir les règles d'éthique;
- Recueillir les données;
- Analyser et interpréter les données;
- Énoncer le jugement et les recommandations;
- Rédiger le rapport;
- Réinvestir les résultats;
- Faire un retour sur l'évaluation.

Bref, il faut, avant de s'engager dans des activités propres à la récolte de données, faire des choix au regard de l'information exigée pour la prise de décision et au regard des moyens pour la collecte de l'information, des choix qui affecteront grandement le déroulement du processus d'évaluation. C'est ici que les indicateurs entrent en action et jouent un rôle d'orientation ayant pour but de faciliter ces choix.

Afin que les lecteurs puissent apprécier l'utilité du modèle C.I.P.P., nous présentons dans les pages qui suivent une description de chacun des quatre types d'indicateurs⁷³ (et, parallèlement, la catégorie de l'évaluation qui y correspond), des exemples d'indicateurs pour chacun ainsi qu'un contenu fictif qui pourrait se retrouver dans le cadre de l'évaluation d'un programme de tutorat.

Nous ne détaillerons pas toutefois les deux orientations d'une évaluation (formative ou sommative) pour toutes les catégories (contexte, intrants, processus, produit); les lecteurs, avec les particularités que nous exposerons, devraient être en mesure de discerner une catégorie de l'autre.

De plus, nous ne nous attendons pas à ce que chaque indicateur énuméré dans les pages suivantes suscite le même intérêt dans tous les établissements collégiaux; le choix se fera selon

⁷² Daniel L. Stufflebeam et Anthony J. Shinkfield, *Systematic Evaluation: a self-instructional guide to theory and practice*, coll. « Evaluation in Education and Human Services », Boston, 1985, p. 162.

⁷³ Au fait, Stufflebeam a proposé une cinquième et dernière version de son modèle, dans laquelle la catégorie se rapportant à l'évaluation du produit est scindée en quatre sous-thèmes. Pour en connaître davantage, les lecteurs intéressés peuvent visiter les sites suivants : www.wmich.edu/evalctr/checklists/cippchecklist.htm et www.wmich.edu/evalctr/pubs/CIPP-ModelOregon10-03.pdf.

les besoins que les collègues éprouveront par rapport à l'évaluation de leurs mesures d'encadrement.

7.1 Les indicateurs et l'évaluation de contexte

L'évaluation de contexte sert de fondement pour les autres types d'évaluation. Les avantages que cette évaluation procure sont certes notables^{74, 75} :

1. Elle facilite la prise de décision en vue de la détermination des objectifs d'une intervention; c'est par l'identification des besoins à satisfaire et par le diagnostic des problèmes sous-jacents à ces besoins que nous obtenons l'information nécessaire pour formuler des objectifs.
2. Elle constitue une base d'information à partir de laquelle nous pouvons déterminer si les objectifs d'une intervention déjà établis ou ses priorités sont en concordance avec les besoins de ceux qui participent à l'intervention. En effet, ce que vise une mesure d'encadrement et ce que recherchent les étudiants participants devraient converger vers un but commun; autrement dit, une intervention doit répondre aux attentes des étudiants qui y participent.

Des sondages, des entretiens, des tests diagnostiques et l'analyse de contenu de documents composent un ensemble de techniques plausibles pour la collecte de données.

Comme indicateurs de contexte, nous pouvons alors retrouver⁷⁶ :

- les caractéristiques socioéconomiques des populations étudiées;
- le taux d'étudiants s'exprimant dans des langues autres que le français et l'anglais;
- le taux d'étudiants qui occupent un emploi à temps partiel;
- le taux d'étudiants en besoin d'aide particulière dès l'entrée dans leur programme d'études;
- le taux d'étudiants qui ont un enfant;
- le nombre moyen d'heures que les étudiants consacrent à un emploi durant une session;
- la consommation par les étudiants d'alcool et de drogues illégales;
- le taux d'étudiants provenant de familles à faible revenu;
- le nombre d'heures consacrées à la télévision ainsi que le type d'émissions visionnées;
- les activités de divertissement des étudiants;

⁷⁴ Daniel L. Stufflebeam, W. J. Foley, W. J. Gephart, E. G. Guba, R. L. Hammond, H. O. Merriman et M. M. Provus, *L'évaluation en éducation et la prise de décision*, Victoriaville, Les Éditions N.H.P., 1980, (1974 pour l'édition originale américaine) — traduction de Jules Dumas.

⁷⁵ Daniel L. Stufflebeam et Anthony J. Shinkfield, *Systematic Evaluation: a self-instructional guide to theory and practice*, coll. « Evaluation in Education and Human Services », Boston, 1985, p. 169-173.

⁷⁶ Les indicateurs ont été puisés dans *Programme d'indicateurs de la qualité en éducation (PIQÉ) — Le cadre*, un programme mettant à profit un système d'indicateurs étant également basé sur le modèle de Stufflebeam, *L'évaluation en éducation et la prise de décision* de Stufflebeam, ainsi que dans *Systematic Evaluation: a self-instructional guide to theory and practice*, une autre interprétation par Stufflebeam de son propre modèle.

- le temps imparti aux études;
- le nombre de cours suivis pendant une session;
- les raisons pour lesquelles les étudiants participent à une intervention.

7.2 Les indicateurs et l'évaluation d'intrants^{77, 78}

Disons que, par l'évaluation de contexte, un département reconnaît que des étudiants se plaignent d'éprouver de la difficulté à lire le matériel pédagogique. Des intervenants décident alors de mettre sur pied des activités tutorales en vue d'améliorer la performance des étudiants en lecture. Les moyens qu'ils prendraient pour réaliser de telles activités seraient définis à partir d'une évaluation d'intrants.

Or, les intrants représentent essentiellement les composantes qui entrent dans le processus et qui influent sur le déroulement d'une mesure d'encadrement. Par ces indicateurs, une équipe peut déterminer les stratégies à prendre pour atteindre au maximum et à un coût raisonnable les objectifs ou les priorités d'une mesure d'encadrement. Elle peut également prévoir les coûts qu'engendrera la mise en œuvre d'une mesure d'encadrement ainsi que les obstacles qui pourraient survenir au moment du déroulement de celle-ci.

La raison d'être des intrants vient de l'évaluation de contexte, dans laquelle des objectifs réalisables ont été élaborés à la suite d'un diagnostic de besoins et de problèmes, ou bien il est apparu que la mesure d'encadrement répond légèrement aux attentes des étudiants participants et que des changements devront y être apportés.

L'évaluation d'intrants précise les ressources matérielles et financières exigées pour l'implantation d'une mesure d'encadrement, ainsi que les besoins en personnel.

Pour mettre en évidence les indicateurs d'intrants, les collèges élaborent des plans d'action qui retiennent les stratégies à appliquer et qui esquissent les exigences de personnel (enseignant ou non), de temps et de financement. La préparation d'une présentation de demande de subvention pour des projets de recherche appliquée constitue aussi un exemple d'application de l'évaluation d'intrants.

⁷⁷ Daniel L. Stufflebeam et Anthony J. Shinkfield, *Systematic Evaluation: a self-instructional guide to theory and practice*, coll. « Evaluation in Education and Human Services », Boston, 1985, p. 173 et 174.

⁷⁸ Daniel L. Stufflebeam, W. J. Foley, W. J. Gephart, E. G. Guba, R. L. Hammond, H. O. Merriman et M. M. Provus, *L'évaluation en éducation et la prise de décision*, Victoriaville, Les Éditions N.H.P., 1980, p. 282-290 (1974 pour l'édition originale américaine) — traduction de Jules Dumas.

⁸¹ Les indicateurs ont été puisés dans *Programme d'indicateurs de la qualité en éducation (PIQÉ) — Le cadre*, un programme mettant à profit un système d'indicateurs étant également basé sur le modèle de Stufflebeam, *L'évaluation en éducation et la prise de décision* de Stufflebeam, ainsi que dans *Systematic Evaluation: a self-instructional guide to theory and practice*, une autre interprétation par Stufflebeam de son propre modèle.

Enfin, les indicateurs d'intrants relatifs à une mesure d'encadrement sont ceux qui peuvent correspondre aux éléments suivants⁸¹ :

- les caractéristiques et les compétences des enseignants, du personnel;
- les problèmes éventuels au moment de la mise en œuvre d'une intervention;
- les contraintes à envisager
 - (*contenu possible* : les tuteurs ne peuvent excéder 12 séances avec un même tuteur);
- la formation des tuteurs
 - (*contenu possible* : les tuteurs recevront chacun 4 heures de formation);
- les ressources disponibles, telles que le matériel, l'équipement, le système informatique pour la rentrée du nombre de sessions par tuteur, le personnel
 - (*contenu possible* : deux professionnels seront chargés de la coordination du programme de tutorat);
- les installations
 - (*contenu possible* : quatre salles où les dyades pourront se rencontrer seront disponibles);
- le financement prévu
 - (*contenu possible* : un montant de 5 000 \$ est disponible pour payer les tuteurs);
- le profil des classes (ex. : le nombre moyen d'étudiants par enseignant);
- le pourcentage de la population totale d'étudiants inscrits au programme de tutorat;
- le pourcentage d'étudiants dans leur première année d'études au collégial et inscrits au programme de tutorat;
- le pourcentage d'étudiants inscrits à des sessions d'accueil et de transition;
- le pourcentage d'étudiants de sexe masculin
 - (*contenu possible* : seulement 37 % des étudiants inscrits en tant que tuteurs sont de sexe masculin).

7.3 Les indicateurs et l'évaluation de processus^{82, 83}

L'évaluation de processus sert à déceler ou à prédire des écueils dans l'implantation de l'intervention. Elle cherche également à connaître si les dispositifs qui ont été proposés pour assurer le bon fonctionnement d'une mesure d'encadrement ont été effectivement mis en place par les intervenants.

Une des méthodes pour évaluer le processus d'une mesure d'encadrement est de décrire ce qui s'est réellement passé. Avec cette information, les responsables sont plus en mesure d'interpréter les résultats d'une mesure d'encadrement, prévus ou imprévus, ou d'apporter des explications quant à l'atteinte ou non des objectifs.

Des techniques ou des instruments pouvant nous servir dans le recueil de données sur le processus d'une mesure d'encadrement peuvent être :

- l'analyse des échanges entre diverses personnes;
- des feuilles de route;
- des entretiens;
- des journaux de bord tenus à jour (par les tuteurs, par exemple);
- des comptes rendus.

Or, la rétroaction obtenue des participants (incluant le personnel de la mesure d'encadrement) et portant sur le déroulement d'une intervention (et non sur ses retombées) constitue une riche source de données, car c'est avec ces données que nous pouvons reconnaître les aspects de la mesure d'encadrement qui doivent être modifiés ou maintenus tels quels.

De plus, les indicateurs de processus renseignent ceux qui désireraient implanter dans leur collège une même mesure d'encadrement et qui demandent à connaître les moyens à prendre.

L'évaluation de processus peut se réaliser durant l'étape de la mise en œuvre d'une intervention, et non seulement à la fin d'une session. Toutefois, les plans pour évaluer le processus seraient dressés avant la mise en place des activités interventionnistes.

⁸² Daniel L. Stufflebeam et Anthony J. Shinkfield, *Systematic Evaluation: a self-instructional guide to theory and practice*, coll. « Evaluation in Education and Human Services », Boston, 1985, p. 174-176.

⁸³ Daniel L. Stufflebeam, W. J. Foley, W. J. Gephart, E. G. Guba, R. L. Hammond, H. O. Merriman et M. M. Provus, *L'évaluation en éducation et la prise de décision*, Victoriaville, Les Éditions N.H.P., 1980, p. 290-293 (1974 pour l'édition originale américaine) — traduction de Jules Dumas.

⁸⁶ Les indicateurs ont été puisés dans *Programme d'indicateurs de la qualité en éducation (PIQÉ) — Le cadre*, un programme mettant à profit un système d'indicateurs étant également basé sur le modèle de Stufflebeam, *L'évaluation en éducation et la prise de décision* de Stufflebeam, ainsi que dans *Systematic Evaluation: a self-instructional guide to theory and practice*, une autre interprétation par Stufflebeam de son propre modèle.

Même si les indicateurs de produit manifestent une certaine preuve d'un effet positif de la mesure d'encadrement, les causes des gains devraient être repérées afin que les intervenants puissent recourir, dans la même ligne, à des moyens dont l'utilité aurait été confirmée.

Les indicateurs de processus ont donc rapport⁸⁶ :

- aux conditions dans lesquelles se sont réalisées les activités
 - (*contenu possible* : les salles dans lesquelles les dyades travaillaient n'étaient pas bien chauffées ni suffisamment éclairées);
- aux relations interpersonnelles entre le personnel et les étudiants;
- aux coûts engagés pour avoir mis en œuvre une mesure d'encadrement;
- au nombre d'heures passées avec un tuteur durant le trimestre
 - (*contenu possible* : les tuteurs ont passé en moyenne 8 heures avec leur tuteur);
- aux stratégies pédagogiques utilisées par les tuteurs avec leur tuteur
 - (*contenu possible* : les tuteurs possédaient une grille leur permettant d'évaluer leur tuteur après chaque session; cette grille était ensuite remise au responsable du programme de tutorat);
- aux méthodes d'évaluation employées par les enseignants;
- aux activités d'apprentissage réalisées;
- à un ensemble de jugements de la part des étudiants à l'égard des activités tutorales;
- au niveau de satisfaction des étudiants ayant participé au programme de tutorat;
- aux perceptions qu'ont les étudiants de la qualité des efforts qui ont été investis par le personnel.

7.4 Les indicateurs et l'évaluation de produit

Le but, en menant une évaluation de produit, est de mesurer, d'interpréter et de juger les réalisations d'une mesure d'encadrement⁸⁷. Par une évaluation de produit, nous pouvons vérifier jusqu'à quel point les besoins des étudiants visés par la mesure d'encadrement ont été comblés. Une évaluation de produit devrait explorer les résultats d'une mesure d'encadrement, incluant

⁸⁷ Il y a certes une confusion qui règne en éducation, car nous ne pouvons isoler des descriptions (ou explications) qui établissent nettement une différenciation entre les termes « effet », « impact », « retombée », « résultat », « conséquence », « répercussion », « extrants », « produit », soit des termes qui ont trait à un indicateur de produit. (Stufflebeam emploie « outcome » et « effect » de façon interchangeable.) Nous préconiserons alors le vocable « réalisation » qui inclut tous ces termes, c'est-à-dire « toute chose » qui s'est réalisée à la suite d'une intervention, qu'elle soit prévue ou imprévue, positive ou négative. Quant aux autres termes, nous n'offrirons pas de définitions, mais leur sens sera évoqué d'après les contextes dans lesquels ils seront insérés.

les effets prévus et imprévus, positifs et négatifs⁸⁸. Une variété de techniques, quantitatives comme qualitatives, peuvent être appliquées dans l'évaluation de produit.

L'évaluation de produit répond très bien aux besoins ayant trait à la responsabilité de rendre des comptes; quand il y a évidence de réalisations positives, elle encourage des instances particulières à continuer à subventionner une intervention⁸⁹.

L'évaluation des effets prévus

Les effets prévus sont très souvent liés au rendement scolaire, mais offrent aussi la possibilité d'être détectés à partir d'indicateurs d'ordre social⁹⁰. Des observateurs, des employeurs et surtout ceux qui ont bénéficié d'une mesure d'encadrement peuvent servir en tant que sources de données quand nous évaluons les effets prévus⁹¹. Aussi, nous retrouvons, dans l'une des premières versions du modèle de Stufflebeam, une technique que plusieurs chercheurs empruntent dans leur recherche, soit la technique qui consiste en une comparaison entre le rendement à un test et l'un des éléments suivants⁹² :

- le rendement à un prétest;
- des normes;
- des standards prédéterminés;
- le rendement d'un groupe de comparaison.

Étant donné que les intervenants n'ont pas toujours à leur disposition le temps nécessaire pour concevoir, tester et valider des tests, que des tests valides ne sont pas toujours facilement disponibles et que certains groupes d'étudiants ne se mesurent pas facilement par des tests standardisés, Stufflebeam reconnaît aujourd'hui que très souvent on doit se contenter de « tests maison »⁹³. (Les chercheurs Monaghan et Chaloux ont effectivement élaboré un

⁸⁸ Daniel L. Stufflebeam et Anthony J. Shinkfield, *Systematic Evaluation: a self-instructional guide to theory and practice*, coll. « Evaluation in Education and Human Services », Boston, 1985, p. 176-178.

⁸⁹ Daniel L. Stufflebeam et Anthony J. Shinkfield, *Systematic Evaluation: a self-instructional guide to theory and practice*, coll. « Evaluation in Education and Human Services », Boston, 1985, p. 176-178.

⁹⁰ En effet, les chercheurs Monaghan et Chaloux ont évalué des effets prévus en prenant des indicateurs d'ordre social. Voir Denis Monaghan et Natalie Chaloux, *Évaluation d'un programme d'aide à la réussite*, Cégep de Sainte-Foy, 2004, p. 3.

⁹¹ Daniel L. Stufflebeam et Anthony J. Shinkfield, *Systematic Evaluation: a self-instructional guide to theory and practice*, coll. « Evaluation in Education and Human Services », Boston, 1985, p. 176-178.

⁹² Daniel L. Stufflebeam et Anthony J. Shinkfield, *Systematic Evaluation: a self-instructional guide to theory and practice*, coll. « Evaluation in Education and Human Services », Boston, 1985, p. 176-178.

⁹³ Daniel L. Stufflebeam, Western Michigan University, *The CIPP Model for Evaluation*, 2003, conférence présentée lors du *2003 Annual Conference of the Oregon Program Evaluators Network (OPEN)*.

questionnaire qui permettait de relever, à l'intérieur de leur recherche, des données sur la perception qu'ont les sujets de leur connaissance de soi⁹⁴.)

D'après Stufflebeam, le schème expérimental était le moyen traditionnel dans l'évaluation des effets prévus; avec cette démarche, des objectifs clairs et opérationnels⁹⁵ sont formulés, et une analyse du rendement des groupes soumis à ce schème permet de fournir des indices quant à l'efficacité ou l'inefficacité d'un traitement quelconque (telle une intervention)⁹⁶.

Stufflebeam affirme cependant que des difficultés peuvent venir entraver une telle méthode; par exemple, il est difficile de respecter tous les principes exigés par l'approche intégrant le plan strictement expérimental, comme la norme de hasard (une distribution aléatoire d'enfants provenant de milieux défavorisés peut être illégale aux États-Unis). Également, l'utilisation de ce schème ne permet pas toujours aux évaluateurs de déceler des différences significatives entre les rendements^{97, 98}.

Puisqu'il leur était difficile de respecter la condition de hasardisation et de constituer des groupes de comparaison de façon purement aléatoire, les chercheurs Monaghan et Chaloux ont, quant à eux, mené leur recherche dans le cadre d'un plan quasi expérimental (qui diffère d'un plan expérimental)⁹⁹.

Ils ont alors formé un groupe Zénith et un groupe témoin par un autre procédé (autre que par hasardisation), ont quand même formulé des objectifs et des hypothèses de recherche et ont vérifié, au moyen de tests d'hypothèses et en s'appuyant sur le modèle de Vitaro, l'efficacité du cheminement Zénith, soit une mesure d'encadrement ayant comme but d'aider les étudiants qui

⁹⁴ Denis Monaghan et Natalie Chaloux, *Évaluation d'un programme d'aide à la réussite*, Cégep de Sainte-Foy, 2004, p. 28.

Les chercheurs voulaient vérifier si le cheminement Zénith avait atteint l'objectif suivant : « à la fin de la session d'encadrement, les élèves du groupe Zénith obtiendront des résultats supérieurs à ceux du groupe témoin pour le questionnaire de la connaissance de soi. » (Monaghan et Chaloux, p. 24)

⁹⁵ Quand un objectif contient des termes abstraits, comme « qualité », « réussite », « étudiant à risque », l'objectif en question est difficilement mesurable (ou observable). La méthode qui consiste essentiellement à convertir en termes concrets et mesurables (ou observables) les concepts abstraits que renferme un objectif se nomme « l'analyse des concepts » (voir Giroux et Tremblay, 2002, p. 58-60). Cette méthode rend opérationnels les objectifs d'une mesure d'encadrement.

⁹⁶ Daniel L. Stufflebeam, W. J. Foley, W. J. Gephart, E. G. Guba, R. L. Hammond, H. O. Merriman et M. M. Provus, *L'évaluation en éducation et la prise de décision*, Victoriaville, Les Éditions N.H.P., 1980, p. 295 et 296 (1974 pour l'édition originale américaine) — traduction de Jules Dumas.

⁹⁷ Daniel L. Stufflebeam, W. J. Foley, W. J. Gephart, E. G. Guba, R. L. Hammond, H. O. Merriman et M. M. Provus, *L'évaluation en éducation et la prise de décision*, Victoriaville, Les Éditions N.H.P., 1980, p. 295 et 296 (1974 pour l'édition originale américaine) — traduction de Jules Dumas.

⁹⁸ Daniel L. Stufflebeam, Western Michigan University, *The CIPP Model for Evaluation*, 2003, p. 30, conférence présentée lors du *2003 Annual Conference of the Oregon Program Evaluators Network* (OPEN).

⁹⁹ Denis Monaghan et Natalie Chaloux, *Évaluation d'un programme d'aide à la réussite*, Cégep de Sainte-Foy, 2004, p. 33.

ont échoué à plusieurs cours lors de leur première ou deuxième session au Cégep de Sainte-Foy, malgré que ces étudiants aient obtenu de bonnes notes au secondaire¹⁰⁰.

Ces chercheurs ont cependant fait part de certains inconvénients quant à la méthode qu'ils ont adoptée pour évaluer l'efficacité de cette intervention : le modèle de Vitaro requiert un nombre important de sujets, étant donné qu'il incorpore des analyses statistiques avancées; les utilisateurs d'une telle méthode doivent être dotés de connaissances poussées en tests statistiques¹⁰¹.

Difficultés à réaliser une évaluation des effets prévus

Un nombre très élevé de participants à une mesure d'encadrement constitue une occasion rare dans le réseau collégial, à moins de vouloir effectuer une étude de type longitudinale ou une étude qui englobe plusieurs collèges. De plus, une formation aléatoire de groupes peut aller à l'encontre de certaines valeurs, ainsi que d'un projet éducatif où le principe d'égalité ou d'équité est de mise; la formation non aléatoire des groupes de comparaison peut constituer une solution de rechange, mais l'équivalence des groupes doit être démontrée, un exercice ne manquant pas de justesse et pouvant aussi exiger l'application de tests statistiques avancés. Enfin, tous les collèges n'ont pas à leur disposition des personnes formées en analyses statistiques avancées. Bref, ces affirmations peuvent corroborer le fait qu'il peut être difficile pour les collèges de réaliser une évaluation des effets prévus.

Des effets imprévus

Toutefois, l'évaluation de produit ne se réduit pas à une analyse des « effets prévus ». Comme l'affirme Stufflebeam, le jugement de l'utilité d'une mesure d'encadrement peut se baser sur d'autres éléments, tels des « effets imprévus », positifs et négatifs¹⁰².

C'est à partir d'une vaste sélection d'informateurs que nous pouvons récupérer leurs opinions quant aux effets imprévus de l'intervention. Certaines techniques sont alors proposées par Stufflebeam¹⁰³ :

¹⁰⁰ Denis Monaghan et Natalie Chaloux, *Évaluation d'un programme d'aide à la réussite*, Cégep de Sainte-Foy, 2004, p. 24, 27, 69 et 82.

¹⁰¹ Denis Monaghan et Natalie Chaloux, *Évaluation d'un programme d'aide à la réussite*, Cégep de Sainte-Foy, 2004, p. 82 et 85.

Les chercheurs Monaghan et Chaloux se sont aussi servis d'autres indicateurs pour mesurer d'autres effets prévus, comme la grandeur et la densité du réseau social des élèves ciblés. C'est l'instrument *Inventaire de réseau social (IRS)* (voir section 6.5) qui a permis à ces chercheurs d'examiner comment le cheminement Zénith a pu influencer sur l'intégration sociale des élèves.

¹⁰² Daniel L. Stufflebeam et Anthony J. Shinkfield, *Systematic Evaluation: a self-instructional guide to theory and practice*, coll. « Evaluation in Education and Human Services », Boston, 1985, p. 176-178.

Nous voulons ajouter que des effets imprévus pourvoient les intervenants d'une information très valable, pourvu que ces derniers soient imprégnés d'objectivité dans leur processus de collecte et d'analyse de données, quantitatives ou qualitatives.

¹⁰³ Daniel L. Stufflebeam et Anthony J. Shinkfield, *Systematic Evaluation: a self-instructional guide to theory and practice*, coll. « Evaluation in Education and Human Services », Boston, 1985, p. 176-178.

- des entretiens individuels de groupe;
- des entretiens individuels avec les participants afin de connaître leurs perceptions des effets imprévus;
- des études de cas sur les expériences vécues par un petit échantillon de participants;
- des sondages par téléphone ou par la poste afin de savoir comment les participants ont jugé la mesure d'encadrement et ce qu'ils en ont retiré.

Ces effets imprévus peuvent consister en des conséquences, positives ou négatives, perçues par des personnes autres que celles qui ont participé activement à l'intervention. Gaudreau utilise, comme indicateur pour les mesurer, les « indicateurs d'impact »; « les dires des enseignants sur la motivation de leurs étudiants qui ont participé au tutorat » illustrent bien un indicateur d'impact¹⁰⁴.

Comme exemples additionnels, nous voulons mentionner que Charlebois, à partir de rapports rédigés par le département de philosophie, a rapporté que, même si des effets directs du CAP¹⁰⁵ sur les taux de réussite n'étaient pas évidents de façon immédiate, des effets indirects ont été reconnus par des étudiants et par des enseignants :

- des enseignants ont affirmé que, depuis l'implantation de ce centre, une plus grande concertation entre des enseignants sur les objectifs des cours et sur les pratiques d'évaluation s'est établie dans le département de philosophie;
- quant aux étudiants qui étaient en contact avec le centre, ils affirment qu'en travaillant avec divers enseignants, autres que celui ou celle qui dispense l'enseignement de leur cours de philosophie, ils ont agrandi, pour traiter des sujets qui relèvent du domaine de la philosophie, leur vision de ceux-ci et élargi leur capacité de réfléchir aux questions d'ordre philosophique¹⁰⁶.

Pour résumer

En somme, les indicateurs de produit fournissent l'information nécessaire pour que nous puissions reconnaître quels ont été les effets, prévus ou imprévus, positifs ou négatifs, et jusqu'à quel point les besoins des étudiants ont été comblés (l'évaluation de processus juge de la valeur du déroulement d'une intervention).

Une interprétation et une compréhension judicieuses des réalisations doivent se baser sur l'information que nous détenons sur le contexte, les intrants ou le processus. En effet, l'évaluation de produit conjuguée à l'évaluation d'une ou d'autres catégories nous permet d'enrichir nos interprétations des résultats obtenus.

Cette alliance se traduit ainsi par de meilleures décisions et diminue le risque de compromettre une mesure d'encadrement : nous ne voudrions pas faire de gestes inutiles, comme supprimer

¹⁰⁴ Louise Gaudreau, *Les indicateurs et les critères*, t. 1 d'Évaluer pour évoluer, Montréal, 2001, p. 24, 31 et 35.

¹⁰⁵ Centre d'aide en philosophie, où l'assistance à des étudiants en difficulté est assumée par des enseignants du département et non par des étudiants (Charlebois, p. 43).

¹⁰⁶ Diane Charlebois, *Analyse des activités d'encadrement — Rapport final*, 1998, p. 45.

une mesure d'encadrement, sans avoir jeté un sérieux regard sur les autres catégories et reconnaître que, par exemple, une diminution du taux de réussite peut possiblement être attribuée au fait que des étudiants beaucoup plus faibles ont fréquenté un centre d'aide pendant la session servant de comparaison, et que l'intervention mérite toujours l'estime qui lui a déjà été accordée par les responsables.

NOTA BENE :

Nous avons constaté que, dans les documents consultés, diverses appellations sont employées pour désigner les indicateurs appartenant à la catégorie de produit telle qu'elle est définie par Stufflebeam : certains auteurs utilisent « indicateurs d'impacts », d'autres, « indicateurs d'effets » et d'autres encore, « indicateurs de résultats ».

Par ailleurs, parmi les auteurs préconisant le modèle C.I.P.P. de Stufflebeam, certains désignent un indicateur donné sous une catégorie, alors que d'autres classifient ce même indicateur sous une autre catégorie. À titre d'exemple, nous avons rencontré l'indicateur relié au financement comme appartenant à la catégorie d'intrants chez un auteur, et à la catégorie de contexte chez un autre auteur. Il n'est donc pas aisé de présenter une classification catégorique des indicateurs utilisés en éducation.

8. STANDARD

Dans la documentation recensée, très peu d'auteurs et de chercheurs abordent explicitement le concept de « standard ». Pour cette partie-ci, nous avons dû nous appuyer sur les travaux de Gadbois. D'après cet auteur, le standard signifie ordinairement une cible que nous visons parmi toutes les données que le contenu d'un indicateur puisse compter¹⁰⁷.

Pour déterminer le standard à atteindre, il faut connaître la situation initiale du phénomène à l'étude. Illustrons tout cela avec un exemple. À la fin de l'année scolaire 2001-2002, dans un collège donné, des intervenants constatent que seulement 10 étudiants ont participé en tant que tuteurs (situation initiale). Ils trouvent que ce nombre n'est pas assez élevé et qu'un nombre plus élevé d'étudiants devraient profiter de ce programme. Ils disent alors que, pour l'année suivante (2002-2003), le nombre de tuteurs (l'indicateur) devra, d'après les intrants qu'ils pourront investir dans le tutorat, être au moins de 25. Ce nombre représenterait alors le standard, cette cible visée par ces responsables.

Mais ce ne sont pas tous les indicateurs pour lesquels il est aussi aisé de déterminer une cible (ou standard) à atteindre. Nous n'avons qu'à rappeler que c'est par un examen de tendance et par une comparaison avec le réseau (avec des données fournies par Chesco) que des collèges publics au Québec se sont fixés, en 2000 (ou 2001), des cibles à atteindre en 2004 relativement à des indicateurs de produit tels le taux de réussite, le taux de diplomation et le taux de persévérance.

¹⁰⁷ Louis Gadbois, *Introduction aux indicateurs de l'éducation*, coll. « L'analyse institutionnelle », Centre d'animation, de développement et de recherche en éducation (CADRE), Montréal, 1988, p. 15.

9. EFFICACITÉ

Les lecteurs ont sûrement constaté que, dans ce document, nous nous sommes gardés de distinguer « l'évaluation » de « l'évaluation de l'efficacité ». Certes, nous avons évoqué le vocable « efficacité » quand il a été question de l'évaluation de produit, mais c'était parce que les chercheurs ou les auteurs en avaient fait une mention explicite dans leurs écrits (voir section 7.4).

Il ne faudrait pas que les lecteurs en déduisent que ce terme est uniquement réservé aux études qui portent sur l'évaluation de produit¹⁰⁸; il faut le savoir, ce terme est source d'une divergence d'interprétation, et des personnes avanceront qu'elles ont rendu témoignage de l'efficacité à partir d'indicateurs n'étant pas liés au produit d'une mesure d'encadrement.

Pour amorcer cette dernière constatation, nous voulons mentionner ce que Désy a noté après avoir scruté des études sur le tutorat par les pairs : en plus du rendement scolaire (comme la moyenne générale ou le résultat final à un cours, le taux de réussite à une session donnée et le taux de persévérance) et du sentiment d'appartenance, des statistiques de fréquentation ainsi que la satisfaction des tutorés et des enseignants sont aussi des indicateurs qui ont servi à des chercheurs ayant affirmé avoir examiné, au moyen de ces éléments, l'efficacité de ce programme¹⁰⁹.

Les auteurs Saucier et Brunelle, dans un document qu'ils ont rédigé pour le milieu de la santé, mentionnent que la notion d'efficacité s'allie très souvent à une mesure de l'atteinte des objectifs (de santé) et proposent une définition du concept d'efficacité qui est axée sur la comparaison des résultats obtenus à des objectifs, normes ou standards. Ils soulignent que cette définition s'utilise aussi dans les cas où les objectifs visent à augmenter un phénomène autre qu'un résultat (comme une ressource ou des intrants que nous voulons investir dans une mesure d'encadrement, telle la participation d'étudiants en tant que tuteurs), et qu'il y aura efficacité si l'objectif visé est atteint¹¹⁰.

En outre, des indicateurs autres que ceux du rendement scolaire, à savoir le nombre de participants et l'opinion des étudiants et des intervenants, ont du être retenus par les collègues dans la réalisation de leur bilan qu'ils ont soumis à la CEEC au mois de décembre 2003; les collègues devaient, d'après les consignes de la CEEC, porter un jugement sur les mesures d'encadrement les plus efficaces en se basant sur ces indicateurs (et sur d'autres, s'ils le désiraient).

De multiples sens

Bref, il appert que de multiples significations sont attribuées à ce concept, et, à moins que les utilisateurs de ce terme n'aient bien circonscrit son sens et précisé de façon explicite ce à quoi il

¹⁰⁸ Certains chercheurs mentionnent que l'efficacité d'une unité ne se mesure qu'avec des indicateurs relatifs au rendement scolaire.

¹⁰⁹ Rapporté par Diane Charlebois, *Analyse des activités d'encadrement — Rapport final*, 1998, p. 14 et 105.

¹¹⁰ Alain Saucier et Yvon Brunelle, *Les indicateurs et la gestion par résultats*, gouvernement du Québec, ministère de la Santé et des Services sociaux, Direction générale de la planification et de l'évaluation, août 1995, p. 11.

renvoie, nous ne pouvons assumer qu'il s'allie à une augmentation d'un rendement scolaire. Si nous voulons accorder au concept d'efficacité une signification particulière dans une évaluation, nous aurions tout intérêt à éviter de noyer le poisson et à clarifier *a priori* les indicateurs en tête de l'évaluation.

Des pistes de réflexion

Puisqu'une évaluation qui tend à déterminer l'efficacité d'une mesure d'encadrement n'est pas toujours facilement réalisable quand l'efficacité est jugée au moyen d'une analyse du rendement scolaire, ou que le contexte dans lequel se déroule une mesure d'encadrement ne se prête pas favorablement à une telle évaluation, il faut être vigilant si nous voulons adopter une définition d'efficacité axée sur le rendement scolaire.

Au terme de cette section, il convient de signaler les raisons pour lesquelles l'évaluation de l'efficacité d'une mesure d'encadrement (et non d'un plan de réussite) à partir du rendement scolaire doit s'effectuer avec des mécanismes circonspects, et que les moyens qui sont parfois utilisés pour en apprécier l'efficacité doivent être repensés :

1. Une étude de type longitudinal (considérer plusieurs sessions pendant lesquelles des étudiants prennent part à une intervention dans un même cours) peut être envisagée pour évaluer l'efficacité d'une mesure d'encadrement. Toutefois, quand le nombre d'étudiants participants est faible, l'analyse comparative demeure très difficile à faire.
2. Le schème expérimental, où les sujets ont été répartis de façon aléatoire dans un groupe de contrôle et dans un groupe expérimental, assure une équivalence de groupes.

Mais, comme des mesures d'encadrement n'ont pas été élaborées dans la perspective d'être évaluées par une telle méthode¹¹¹, et que la formation de ces groupes peut soulever des problèmes d'autre nature, des chercheurs opteront pour des plans quasi expérimentaux, avec lesquels le groupe témoin et le groupe de contrôle auront été constitués à partir de moyens qui assurent le plus possible l'équivalence des groupes.

Le problème maintenant revient à démontrer cette équivalence, un exercice dont les tentatives sont nombreuses, mais qui ne demeure pas facilement exécutable.

3. Des intervenants d'une mesure d'encadrement dans laquelle les étudiants participants sont peu nombreux ont certainement constaté la difficulté de donner des significations claires aux résultats qu'ils ont obtenus, surtout à partir des données renvoyant à des indicateurs de produit comme le taux de réussite. Mais le fait qu'une intervention soit difficilement analysable ne signifie pas qu'elle est inefficace.
4. Si nous voulions comparer un taux de réussite d'un cours où des étudiants ont participé au tutorat avec, disons, le taux de réussite de ce même cours offert auparavant durant une ou deux sessions où l'intervention était inexistante, cette comparaison nécessiterait, au moins, que tous les étudiants de ce cours offert à des sessions différentes aient été évalués selon une même méthodologie d'évaluation, sinon l'effet (si nous pouvons le

¹¹¹ Diane Charlebois, *Analyse des activités d'encadrement — Rapport final*, 1998, p. 108.

discerner) pourrait être attribué au fait qu'une pratique différente a été employée pour évaluer les étudiants, et non à l'intervention¹¹².

5. Il n'est pas aisé, dans une évaluation d'efficacité, de relever tous les facteurs qui peuvent influencer sur le rendement scolaire, et donc problématique d'attribuer à une seule mesure d'encadrement un phénomène mesuré et lié au rendement scolaire. Ce n'est pas parce que nous constatons, sans une analyse statistique poussée, une augmentation d'un taux quelconque que nous pouvons l'imputer indubitablement à l'intervention; de même, si les effets recherchés ne se sont pas manifestés au « premier coup d'œil », nous ne pouvons conclure, pour autant, que l'intervention a été infructueuse¹¹³.
6. Évaluer l'efficacité d'une mesure d'encadrement sur le rendement scolaire quand les étudiants n'ont participé qu'à deux ou trois activités tutorales, chacune d'une durée de 50 minutes, ne se fait pas sans confusion. De même, analyser des taux de réussite des cours dans lesquels le nombre d'étudiants ayant participé à une intervention est négligeable peut donner lieu à des conclusions erronées.
7. Enfin, des analyses de données quantitatives peuvent nécessiter un très grand nombre de sujets, car c'est par la nature même de plusieurs tests statistiques que de petites différences significatives sont plus faciles à déceler quand il y a un nombre important de sujets¹¹⁴. Mais des analyses statistiques avancées ne sont pas une fin en soi...

¹¹² Argument inspiré de celui présenté par Diane Charlebois, *Analyse des activités d'encadrement — Rapport final*, 1998, p. 107.

¹¹³ Explication inspirée de celle de François-Marie Gérard, « L'évaluation de l'efficacité d'une formation », *Gestion 2000*, 2000, vol. 20, n° 3, p. 13-33.

¹¹⁴ Andrée Lamoureux, Jean-Pierre Archambault, François Berthiaume et Nathalie Fréchette, *Une démarche scientifique en sciences humaines : méthodologie*, Laval, 1992, p. 331.

CONCLUSION

Les indicateurs annoncent les phénomènes qui seront observés (ou mesurés) au cours d'une évaluation d'une mesure d'encadrement; ils possèdent l'atout d'aider les intervenants à cerner le type de données pour les besoins de l'évaluation d'une mesure d'encadrement et à faciliter le choix du type d'instruments de collecte qui serviront à recueillir les données appropriées. Même si des indicateurs quantitatifs pullulent dans les revues éducatives, les indicateurs qualitatifs, eux, y font leur entrée graduellement.

Le tableau de bord des indicateurs, la liste de vérification des qualités des indicateurs et le plan d'évaluation composent un matériel qui s'avère indispensable en évaluation. Idéalement, pour que le processus d'évaluation se déroule en conformité avec certains principes théoriques et qu'il n'aille pas à la dérive, un plan d'évaluation devrait s'élaborer, et ce, bien avant la collecte de données.

Comme l'évaluation ne se réduit pas à analyser des données relatives à un taux de réussite ou à des moyennes, une méthode qui s'inscrit très souvent dans une perspective d'« accountability », et qu'elle peut aussi viser à reconnaître des effets imprévus (positifs ou négatifs), à améliorer toute façon de procéder et à écarter la possibilité de mettre une intervention en situation critique, des indicateurs autres que ceux liés à un rendement scolaire sont requis.

Le modèle de Stufflebeam, un modèle classique et exploité en évaluation de programmes (ou de projets ou d'activités), a servi comme fondement à notre typologie d'indicateurs : une catégorie d'indicateurs répondait à chacun des quatre catégories d'évaluation, et nous retrouvons alors dans notre système de classification les indicateurs de contexte, les indicateurs d'intrants, les indicateurs de processus et les indicateurs de produit.

Ce modèle présente l'évaluation comme un instrument avec lequel les intervenants peuvent déterminer les effets prévus d'une mesure d'encadrement, si cet exercice est réalisable ou souhaitable, ou repérer les effets imprévus, les mérites et les écueils d'une mesure d'encadrement, les aspects significatifs en besoin de changement, de continuité ou d'abandon, etc.

Des omissions

Par cette recherche, nous avons tenté de répondre à un besoin que nous avons pressenti relativement au dossier de l'évaluation des mesures d'encadrement, mais en omettant des particularités de l'évaluation, plusieurs jugées essentielles par des spécialistes en évaluation :

1. Nous n'avons pas présenté tous les éléments de la liste des qualités des indicateurs apparaissant à la section 3, notre but avec ce document étant, avant tout, d'effleurer la théorie relative à l'évaluation.

2. Chaque technique de collecte de données comportant des avantages et des inconvénients, nous n'avons pas soulevé tous les débats que suscite chacune des approches (quantitative et qualitative) et n'avons pas précisé que toute méthodologie comporte des biais.
3. Nous avons aussi caractérisé brièvement les indicateurs, et fait part d'une courte liste d'indicateurs utilisés dans le milieu éducatif, mais sans attester de la disponibilité des instruments qui pourraient servir à les mesurer.
4. Nous n'avons pas traité de façon exclusive le code de déontologie, les échelles de mesure, le désistement des sujets, les techniques d'échantillonnage des sujets, la difficulté à assurer, dans la réalité, les critères pratiques et scientifiques en évaluation, comme les validités interne et externe, et la fidélité.
5. Nous n'avons pas expliqué que parfois les bénéfices se réalisent plus tard dans un parcours scolaire, après que les étudiants ont tiré parti d'une intervention, et que, si des intervenants concluaient hâtivement que la mesure d'encadrement n'avait pas produit les résultats escomptés, cette conclusion pourrait donner lieu à des décisions erronées.
6. Nous n'avons pas mis l'accent sur le fait que, s'ils désirent maintenir à la hauteur les mesures d'encadrement mises en place dans leur établissement, les intervenants auraient tout avantage à documenter les procédés qu'ils ont utilisés dans l'évaluation d'une mesure d'encadrement, de les examiner et d'y apporter des changements constructifs. En effet, pour que l'évaluation ne soit pas détournée de sa fin, il importe qu'elle fasse l'objet, à son tour, d'un examen appréciatif.
7. Nous avons seulement rapporté un sommaire de méthodes qualitatives et de méthodes quantitatives qui sont consacrées à l'évaluation du programme de tutorat, sans en expliciter les détails.
8. Nous n'avons pas développé sur l'aspect relié au nombre de sujets requis pour l'emploi des techniques quantitative et qualitative.
9. Nous n'avons pas insisté sur la nécessité d'évaluer, à l'intérieur d'une évaluation de contexte, la pertinence des objectifs d'une mesure d'encadrement, en sachant, d'une part, que ces objectifs (s'ils existent) n'ont pas tous été précédés d'une analyse systématique de besoins et, d'autre part, qu'il serait difficile pour les intervenants d'arriver à un consensus sur les finalités d'une mesure d'encadrement qui est déjà mise en place et qui concourt déjà aussi à combler une multiplicité de besoins chez les étudiants.

De plus, nous avons constaté que, peu importe la visée d'une mesure d'encadrement, celle-ci peut être destinée, par des instances, à faire preuve d'efficacité relativement à des taux de réussite, à des taux de persévérance et à des taux de diplomation.

10. La pratique d'évaluer exigeant une attention d'envergure et un effort d'équipe, nous n'avons pas précisé qu'il importe forcément d'envisager plusieurs aspects dans une évaluation d'une mesure d'encadrement, et que, quand les pratiques d'évaluation ne sont pas fonctionnelles, l'évaluation pourrait perdre son effet et, par ricochet, la mesure d'encadrement pourrait ne plus avoir cours.

Mais l'inclusion de tous ces renseignements aurait largement dépassé le cadre de notre document, et c'est plutôt au cours des prochaines étapes que nous aurons l'occasion de souligner d'autres facteurs qui doivent être considérés par les intervenants dans leur évaluation des mesures d'encadrement.

Bref, ce document ne prétend pas se substituer à des cours de formation avancée en évaluation des mesures d'encadrement; au contraire, il ne faut y voir qu'une somme partielle d'éléments informatifs, servant à alimenter les réflexions sur la problématique de l'évaluation pour que, par la suite, les aspects pratiques à considérer dans une évaluation soient facilement perçus.

Notre vision

Des collègues qui désirent conduire des activités d'évaluation de grande portée (comme celles qui engagent des calculs statistiques avancés, un plan expérimental ou quasi expérimental, ou même des analyses qualitatives poussées) devront avoir recours à des spécialistes en évaluation.

Cependant, tout collègue peut appliquer une méthodologie qui n'exige pas de formation « spécialisée » en évaluation et qui fait preuve d'un niveau certain de rigueur, pourvu que le travail d'évaluation se fasse avec la collaboration de plusieurs collègues.

La satisfaction, les perceptions des participants et d'autres indicateurs peuvent alors, même si ce ne sont pas des indicateurs qui représentent nécessairement un gage de la réussite scolaire, servir à évaluer un degré certain de la valeur des interventions, pourvu que les instruments soient construits et que les procédés d'évaluation se déroulent dans cette finalité.

Nous sommes donc d'avis que, même avec un atelier d'une durée de quelques heures offert dans le but de former des étudiants en recherche documentaire, des renseignements sur le contexte (savoir si l'atelier a répondu aux attentes des étudiants), sur la qualité du matériel fourni pour l'atelier et sur la satisfaction des étudiants par rapport aux activités réalisées durant l'atelier peuvent fournir de précieux indices aux organisateurs de cet atelier, leur permettant ainsi de reconnaître si de poursuivre tel quel est pertinent ou s'il devra faire l'objet de modifications. Quoiqu'ils ne mesurent pas les effets directs de cette activité sur le rendement scolaire des étudiants, les intervenants peuvent, avec les éléments informatifs recueillis, améliorer cet atelier et éviter qu'il devienne caduc.

Et encore, nous jugeons que tout procédé qu'un collègue emprunte pour évaluer des mesures d'encadrement est déjà imprégné d'un certain niveau de rigueur, et que, lorsque les intervenants s'adonnent avec impartialité aux moyens qu'ils prennent pour recueillir les données, tant quantitatives que qualitatives, ils ne font que renforcer les chances de recueillir des données valables.

C'est dans cette optique que le Carrefour de la réussite au collégial désire poursuivre sa recherche; plus précisément, que chaque collègue s'approprie une démarche instrumentée pour évaluer, avec des choix mieux éclairés et une plus grande précision, les mesures d'encadrement mises en place dans leur établissement, une démarche dont la préoccupation est croissante dans le réseau collégial.

ANNEXE 1

Liste de vérification des qualités des indicateurs

ANNEXE 1**LISTE DE VÉRIFICATION DES QUALITÉS DES INDICATEURS**

INDICATEURS :				
1. Faisabilité pour colliger et traiter les données : <ul style="list-style-type: none"> • Disponibilité des sources • Facilité pour colliger et traiter les données • Faible coût pour recueillir et traiter les données • Court laps de temps pour recueillir et traiter les données • Peu de personnes nécessaires pour colliger et traiter les données 				
2. Facile de compréhension				
3. Facilement comparable aux autres				
4. Spécificité				
5. Sensibilité				
6. Validité				

Source : Louise Gaudreau, 2001, *Les indicateurs et les critères*, t. 1 d'Évaluer pour évoluer, p. 59.

MODE D'EMPLOI

1. En haut, inscrire le nom des indicateurs pressentis (ou leur numéro ou leur abréviation)
2. Comparer chacun des indicateurs à chacune des qualités : inscrire
 - un « √ » sous l'indicateur vis-à-vis de la qualité qu'il possède;
 - un « + ou - » pour la qualité qu'il possède peu;
 - un « non » s'il ne possède pas cette qualité.
3. Conserver les indicateurs qui obtiennent le plus de qualités (le plus de « √ », le plus de « + ou - »), mais qui répondent aussi à un minimum de faisabilité (sinon, on ne pourra pas s'en servir).

Source : Louise Gaudreau, 2001, *Les indicateurs et les critères*, t. 1 d'Évaluer pour évoluer, p. 59.

ANNEXE 2

Tableau de bord des indicateurs

ANNEXE 2

TABLEAU DE BORD DES INDICATEURS

Indicateurs	Types de données	Sources des données	Techniques et instruments pour collecte des données	Analyses	Valeurs calculées ou nature des observations obtenues	Jugements et décisions

Source : Louise Gaudreau, 2001, *Les indicateurs et les critères*, t. 1 d'Évaluer pour évoluer, p. 35 et 61.

ANNEXE 3

D'autres indicateurs

ANNEXE 3

D'AUTRES INDICATEURS

D'autres indicateurs ont aussi été relevés à travers les documents que nous avons consultés pour la rédaction de cette recherche : nous les avons plus souvent retrouvés sous l'appellation « variables ». Ils sont présentés ici dans un ordre alphabétique; certains sont quantitatifs, d'autres, qualitatifs.

Même si certains étaient utilisés dans le cadre de l'évaluation de mesures d'encadrement autres que le programme de tutorat, ils peuvent être retenus pour l'évaluation de celui-ci. Bien entendu, il faut faire des choix, et le tableau fourni à la section 3 (Liste de vérification des qualités des indicateurs) peut guider les intervenants dans la sélection d'indicateurs appropriés à leur évaluation.

- Attitude envers l'école
- Attitude envers le tutorat
- Attitude envers une matière donnée
- Autonomie dans l'apprentissage
- Bénéfices affectifs
- Bénéfices émotionnels
- Bénéfices sociaux
- Capacités d'analyse, de synthèse et de pensée critique
- Compréhension des notions
- Concept de soi
- Confiance en soi
- Cote R
- Dépense de fonctionnement : rapport élèves/dépenses
- Estime de soi (chez les tuteurs et les tutorés)
- Intérêt des tutorés envers la matière
- Nature des comportements entre élèves
- Niveau de stress à l'égard d'un examen
- Niveau de motivation
- Nombre d'exercices réalisés par les tutorés avec leur tuteur
- Nombre d'exercices réalisés seul
- Nombre de participants masculins
- Nombre de participants masculins désignés comme étant en difficulté dans une matière donnée

- Nombre de tuteurs
- Nombre de tutorés
- Perception des parents
- Perception par les étudiants de l'environnement éducatif
- Perception du programme de tutorat par un établissement
- Pourcentage des dépenses totales du cégep correspondant aux dépenses reliées au tutorat
- Pourcentage d'unités de cours réussis ou ratés
- Quantité de ressources allouées au tutorat
- Résultat obtenu à un test de connaissances
- Résultats aux prétests et posttests de connaissances
- Résultats des étudiants aux évaluations formatives
- Score obtenu à un test sur la motivation scolaire
- Taux d'absentéisme à un cours (un cours à risque d'échec élevé) dans lequel des étudiants participent au tutorat
- Transfert d'acquis
- Utilité du tutorat

LISTE DES TABLEAUX ET DES FIGURES

Tableau 3.1	Liste de vérification des qualités des indicateurs	p. 17
Figure 4.1	La démarche générale d'une évaluation	p. 20
Tableau 5.1	Tableau de bord des indicateurs	p. 22
Tableau 6.1	Avantages et inconvénients des techniques	p. 29
Tableau 6.2	Liste d'instruments	p. 31

SOURCES DOCUMENTAIRES

- ABT. ASSOCIATES INC., *AmeriCorps Tutoring Outcomes Study*, Executive Summary, 2001, 5 p., www.americorps.org.
- BLAIS, Jean-Guy, Michel LAURIER et Guy PELLETIER, « Regards sur la problématique des indicateurs en éducation », *Mesure et évaluation en éducation*, vol. 22, n^{os} 2 et 3, 1999, p. 47-69.
- BOUDON, Raymond et Paul LAZARSELD, *Le vocabulaire des sciences sociales*, coll. « Méthodes de la sociologie », Paris, Mouton & Co. et Maison des Sciences de l'Homme, 1965, 309 p.
- CHARLEBOIS, Diane, *Considérations pratiques sur les indicateurs de programmes*, conférence donnée à Saint-Hyacinthe lors du 24^e colloque annuel de l'AQPC, Québec, 10 juin 2004.
- CHARLEBOIS, Diane, *Analyse des activités d'encadrement — Rapport final*, 1998, 175 p.
- DÉSY, Jocelyne, *Le tutorat par les pairs tel que perçu par les élèves*, Cégep de Sainte-Foy, 1996, 135 p.
- EHLI, Stewart W. et Stephen C. LARSEN, *Peer Tutoring for Individualized Instruction*, Boston, Allyn and Bacon, Inc., 1980, 247 p.
- FONTAINE, France, *Guide pour la création et la révision d'un programme d'études*, Montréal, Service d'aide à l'enseignement, Université de Montréal, 1992, 81 p.
- GADBOIS, Louis, *Introduction aux indicateurs de l'éducation*, coll. « L'analyse institutionnelle », Montréal, Centre d'animation, de développement et de recherche en éducation (CADRE), 1988, 43 p.
- GAUDREAU, Louise, *Les indicateurs et les critères*, t. 1 d'Évaluer pour évoluer, Outremont, Les Éditions LOGIQUES, 2001, 71 p.
- GAUDREAU, Louise, *Les étapes d'une évaluation de programme ou de projet*, t. 2 d'Évaluer pour évoluer, Outremont, Les Éditions LOGIQUES, 2001, 103 p.
- GÉRARD, François-Marie, « L'évaluation de l'efficacité d'une formation », *Gestion 2000*, 2000, vol. 20, n^o 3, p. 13-33.
- GINGRAS, Paul-Émile, *Le concept d'« indicateur » appliqué en éducation*, coll. « L'analyse institutionnelle », Montréal, Centre d'animation, de développement et de recherche en éducation (CADRE), 1977, 90 p.
- GIROUX, Sylvain et Ginette TREMBLAY, *Méthodologie des sciences humaines : la recherche en action*, 2^e éd., coll. « ERPI Sciences humaines », Saint-Laurent, Éditions du Renouveau Pédagogique inc., 2002, 262 p.

GRUPE DE RECHERCHE SUR LE TUTORAT (GRT), www.u-bordeaux2.fr/grt/page1.html.

HOWELL, David C., *Méthodes statistiques en sciences humaines*, 2^e tirage, trad. de l'anglais par Marylène Rogier, Paris, De Boeck Université; Boston, ITP, 1999, 821 p.

LAMOUREUX, Andrée, Jean-Pierre ARCHAMBAULT, François BERTHIAUME et Nathalie FRÉCHETTE, *Une démarche scientifique en sciences humaines : méthodologie*, Laval, Études Vivantes, 1992, 638 p.

LAPAN, Stephen D., *Foundations of Research: Methods of Inquiry in Education and the Social Sciences*, Northern Arizona University, www4.nau.edu/cee/ci_doc/current/resources/13_lapan.pdf.

LAROSE, Simon, Nathalie SOUCY et Stéphane DUCHESNE, « Examen des programmes de tutorat maître-élève dans les collèges du réseau québécois », *Pédagogie collégiale*, vol. 13, n^o 4, mai 2000, p. 11-37, www.cdc.qc.ca/ped_coll/soucy_13_4.html.

MONAGHAN, Denis et Natalie CHALOUX, *Évaluation d'un programme d'aide à la réussite*, Cégep de Sainte-Foy, 2004, 97 p., http://www.claurendeau.qc.ca/ceg_andre/Accueil/affichage.asp?C=0_07_04_&N=2.

NADEAU, Gilles, *Critères et indicateurs de qualité et d'excellence dans les collèges et les universités au Canada : bilan des trois phases du projet*, Winnipeg, Centre for Higher Education, Research and Development, University of Manitoba, 1995, 268 p.

PATTON, Michael Q., *Qualitative Evaluation Methods*, Beverly Hills, Sage Publications, 1980, 381 p.

PATTON, Michael Q., *How to use Qualitative Methods in Evaluation*, 6th Printing, Newbury Park, Sage Publications, 1990, 176 p.

PLANTE, Jacques, *Évaluation institutionnelle*, document inédit, Département de mesure et évaluation, Université Laval, 1994, 86 p.

Programme d'indicateurs de la qualité en éducation (PIQÉ) — Le cadre, www.eqao.com/pdf_f/01/01P028f.pdf.

SANTÉ CANADA, *Guide d'évaluation de projet : une démarche participative*, Direction de la santé de la population, 1996, 51 p., www.hc-sc.gc.ca/hppb/violencefamiliale/html/fvprojevaluation_f.html.

SAUCIER, Alain et Yvon BRUNELLE, *Les indicateurs et la gestion par résultats*, coll. « Méthodologie et instrumentation », gouvernement du Québec, ministère de la Santé et des Services sociaux, Direction générale de la planification et de l'évaluation, 1995, 33 p.

STUFFLEBEAM, Daniel L., W. J. FOLEY, W. J. GEPHART, E. G. GUBA, R. L. HAMMOND, H. O. MERRIMAN et M. M. PROVUS, *L'évaluation en éducation et la prise de décision*, Victoriaville, Les Éditions N.H.P., 1980, 464 p. (1974 pour l'édition originale américaine) – traduction de Jules Dumas.

STUFFLEBEAM, Daniel L. et Anthony J. SHINKFIELD, 1985, *Systematic Evaluation: a self-instructional guide to theory and practice*, coll. « Evaluation in Education and Human Services », Boston, Kluwer-Nijhoff Publishing, 1985, 347 p.

STUFFLEBEAM, Daniel L., Western Michigan University, *The CIPP Model for Evaluation*, 2003, 68 p., conférence donnée lors du *2003 Annual Conference of the Oregon Program Evaluators Network (OPEN)* — <http://www.wmich.edu/evalctr/pubs/CIPP-ModelOregon10-03.pdf>.