



Higher Education
Quality Council
of Ontario

An agency of the Government of Ontario

Manuel d'évaluation des programmes du COQES

Programmes liés à l'accès et à la persévérance scolaire

Helen Tewolde et Jess McKeown



Publié par

Le Conseil ontarien de la qualité de l'enseignement supérieur

1, rue Yonge, bureau 2402
Toronto (Ontario) Canada M5E 1E5

Téléphone : 416-212-3893
Télécopieur : 416-212-3899
Web : www.heqco.ca
Courriel : info@heqco.ca

Format à utiliser à des fins de citation

Tewolde, H. et McKeown, J. (2018). *Manuel d'évaluation des programmes du COQES — Programmes liés à l'accès et à la persévérance scolaire*. Toronto. Conseil ontarien de la qualité de l'enseignement supérieur.



Les opinions exprimées dans le présent document de recherche sont celles des auteurs et ne représentent pas nécessairement les points de vue ou les politiques officiels du Conseil ontarien de la qualité de l'enseignement supérieur ou d'autres organismes ou organisations qui pourraient avoir fourni un soutien, financier ou autre, pour ce projet. © Imprimeur de la Reine pour l'Ontario, 2018.

Remerciements

Nous tenons à remercier nos examinateurs externes, Heather Smith Fowler (directrice de recherche, Société de recherche sociale appliquée) et Kate Feightner (Feightner & Associates), ainsi que les contributions d'autres membres du personnel du COQES, Sophie Lanthier et Maya Gunnarsson.

Introduction

Le Conseil ontarien de la qualité de l'enseignement supérieur (COQES) est un organisme du gouvernement de l'Ontario qui fait progresser la recherche fondée sur des données probantes pour l'amélioration continue du système d'éducation postsecondaire (EPS) en Ontario. Afin d'améliorer l'accès, la qualité et la responsabilisation des collèges et universités de l'Ontario, le COQES appuie divers projets d'évaluation de programmes dans le secteur de l'EPS afin de s'assurer que les pratiques prometteuses sont reconnues et promues par l'élaboration de programmes et de politiques.

L'évaluation des programmes est une priorité du COQES pour de nombreuses raisons. Le gouvernement de l'Ontario dépense des millions de dollars chaque année pour des programmes et des initiatives visant à favoriser l'accès et la persévérance scolaire des étudiants dans les établissements postsecondaires. Nous devons nous poser les questions suivantes : Qu'est-ce qui fonctionne? Quelle est notre cible? Que mesurons-nous? Sans une évaluation uniforme des programmes financés par l'État, il est impossible de répondre à ces questions.

La Société canadienne d'évaluation définit l'évaluation de programmes comme l'évaluation systématique de la conception, de la mise en œuvre ou des résultats d'une initiative à des fins d'apprentissage ou de prise de décisions¹. Les évaluations de programme peuvent aider une organisation ou un établissement à déterminer a) si, comment et pourquoi un programme fonctionne; b) si un programme atteint les résultats escomptés; c) comment un programme pourrait être amélioré ou d) si un programme peut ou devrait être adapté à d'autres contextes. Avec une définition aussi large, il est facile de comprendre pourquoi les évaluations de programme différeront radicalement d'un contexte à l'autre.

Le présent manuel vise à fournir un aperçu concis des considérations pratiques et un processus dirigé pour ceux qui connaissent moins bien la planification et la mise en œuvre d'une évaluation de programme. Le présent manuel se veut principalement à l'usage des promoteurs du Consortium sur l'accès et la persévérance scolaire (CAPS) du COQES, une communauté de pratique pour ceux qui évaluent les programmes qui améliorent la participation et la réussite des étudiants aux EPS. Le présent manuel peut aussi être utilisé de façon plus générale par ceux qui désirent évaluer des programmes connexes dans l'ensemble du secteur de l'enseignement supérieur. Le manuel est divisé en différentes sections :

A. Déterminer le but et la portée d'une évaluation

1. Se préparer à une évaluation en élaborant des questions de recherche
2. Définir des variables clés
3. Déterminer la portée du projet et qui devrait participer à l'évaluation
4. Choisir une approche d'évaluation
5. Choisir une conception d'étude

B. Effectuer une évaluation

6. Collecter des données
7. Échantillonner la population
8. Gérer, stocker et analyser les données
9. Présenter des résultats

1 Pour plus de détails sur cette définition de l'évaluation de programme, consultez le [site Web](#) de la Société canadienne d'évaluation.

C. Autres facteurs à prendre en considération pour l'évaluation

10. Respecter les normes d'éthique
11. Établir un budget

Les sections 1 à 9 ont été ordonnancées comme une séquence possible pour la planification d'une évaluation, mais ces étapes se dérouleront rarement de façon linéaire. Les évaluateurs constateront que l'évaluation de programme est un processus extrêmement itératif. Les mesures choisies pendant le processus de planification peuvent être prises simultanément ou dans un ordre différent de celui présenté ici. Selon l'expérience d'une personne en matière d'évaluation et son rôle éventuel au sein d'une équipe d'évaluation plus vaste, certaines sections seront plus pertinentes que d'autres. Que les évaluateurs soient nouveaux sur le terrain ou qu'ils cumulent des années d'expérience en évaluation, il est possible qu'ils utilisent une séquence en sens inverse, qu'ils reviennent sur des étapes précédentes pour clarifier des travaux ultérieurs et qu'ils modifient leur plan d'exécution initial. Il est essentiel de prévoir du temps supplémentaire pour la nature itérative de l'évaluation de programme, surtout du point de vue de la gestion de projet.

Avant de plonger dans le contenu du manuel, nous désirons formuler quelques mises en garde importantes. Premièrement, quelle que soit l'approche que l'on privilégie en matière d'évaluation de programme, les questions de recherche constituent l'élément le plus central et essentiel de l'évaluation. Les questions de recherche, parfois appelées questions d'évaluation, serviront à orienter l'ensemble du processus d'évaluation, de la planification à la mise en œuvre en passant par la production de rapports. Il convient donc de préciser un point important au sujet de la langue. La sémantique entourant l'évaluation d'un programme peut varier d'un évaluateur à l'autre, en particulier entre différentes disciplines et différents secteurs (p. ex., santé, psychologie, éducation, etc.). Il faut reconnaître le fait qu'il existe souvent plusieurs façons d'exprimer la même idée. Aussi, dans certains cas, nous avons regroupé des concepts semblables pour faciliter la compréhension des néophytes. Par souci de simplicité, nous nous sommes efforcés d'harmoniser le langage utilisé dans ce manuel avec le langage le plus largement utilisé dans le domaine de l'éducation. Par ailleurs, il convient de préciser que nous utiliserons le terme « programme » pour désigner tout programme, toute intervention ou toute stratégie faisant l'objet d'une évaluation. Enfin, compte tenu de la complexité de l'évaluation des programmes en tant que domaine d'études et de pratique, il importe de souligner que le présent manuel ne se veut pas exhaustif et ne le deviendra jamais. Même si ce manuel se veut à l'usage de néophytes en simplifiant autant que possible le concept d'évaluation, le caractère intrinsèquement désordonné de l'évaluation est quand même représenté. Il existe toujours plusieurs façons de percevoir les choses et l'évaluation ne sera jamais aussi simple qu'elle pourrait le paraître sur papier. Nous n'avons pour seule ambition que ce manuel devienne une ressource utile pour les néophytes en matière d'évaluation, et nous aiguillerons le lecteur vers des ressources additionnelles dans le cas de certains sujets lorsque cela s'avèrera à propos.

A Déterminer le but et la portée d'une évaluation

Partie 1 : Se préparer à une évaluation en élaborant des questions de recherche

Les questions de recherche constituent la pierre angulaire de chaque évaluation de programme. Les questions de recherche sont essentielles pour encadrer une évaluation parce qu'elles sont axées sur l'objectif. Elles éclaireront donc toutes les décisions futures qui seront prises tout au long du projet. Les personnes participant à l'évaluation compareront les nombreux avantages et inconvénients rattachés à toutes les activités de recherche connexes (p. ex., la conception de l'étude, la collecte des données, les techniques d'échantillonnage, l'analyse des données, etc.) et s'en remettront en fin de compte aux questions de recherche pour justifier leurs choix. Le nombre, l'orientation et la spécificité des questions d'évaluation dépendront de l'expérience de l'équipe d'évaluation (Partie 3). Une poignée de questions de recherche de haut niveau peut suffire pour les évaluateurs chevronnés, tandis que les nouveaux évaluateurs auront peut-être avantage à élaborer des questions plus précises. Les questions suivantes représentent des exemples généraux de questions de recherche :

Questions de recherche
Questions fondamentales qui expliquent pourquoi un programme fait l'objet d'une évaluation et orientent les activités de recherche.

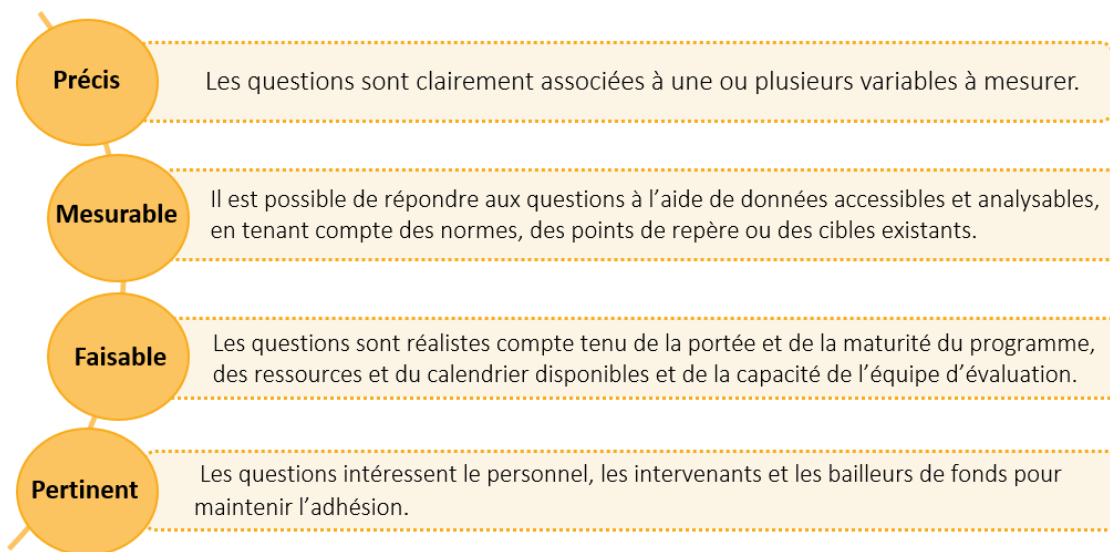
- ✓ Le problème que le programme tente de régler est-il bien établi et le programme est-il conforme au mandat ou à la mission de l'organisation?
- ✓ Dans quelle mesure le programme s'harmonise-t-il avec la recherche et les connaissances pratiques au sujet du problème et de la façon de le régler?
- ✓ Le programme a-t-il été mis en œuvre conformément au plan et au groupe cible?
- ✓ Comment et pourquoi le programme fonctionne-t-il ainsi?
- ✓ Le programme atteint-il les résultats ou les buts prévus (à court, moyen ou long terme)?
- ✓ Le programme entraîne-t-il des conséquences imprévues ou indésirables?
- ✓ Comment le programme pourrait-il être amélioré à l'avenir?
- ✓ Dans quelle mesure le programme est-il rentable comparativement à d'autres programmes ayant des objectifs semblables?
- ✓ Le programme peut-il ou devrait-il être transféré à d'autres contextes ou élargi?
- ✓ Comment des programmes semblables dans le secteur pourraient-ils tirer des leçons des résultats de l'évaluation?

Il est très rare qu'une évaluation puisse porter sur tous les domaines d'intérêt potentiels, de sorte que les évaluateurs devraient réfléchir attentivement aux aspects les plus importants d'un programme à évaluer, et peut-être consulter les intervenants pertinents s'il y a lieu (Partie 3)². Pour les évaluations qui comportent des

² Pour de plus amples renseignements sur le choix des questions de recherche, consultez le site Web de la boîte à outils communautaire [ici](#).

questions de recherche multiples, nous recommandons que les questions soient distinctes, mais complémentaires, et que les évaluateurs établissent les liens et la progression chronologique dans chaque cas. Idéalement, des questions de recherche bien articulées répondront aux critères de la Figure 1³. Ces critères seront revus tout au long du présent document.

Figure 1 : Critères pour des questions de recherche bien rédigées.



Le moment et la façon d'élaborer les questions de recherche dépendront de la nature du projet. Par exemple, les questions peuvent avoir été élaborées avant le début de l'évaluation, élaborées de façon organique au moyen de conversations avec le personnel du programme ou élaborées au moyen de discussions importantes avec un groupe de directeurs principaux. Pour appuyer l'élaboration des questions de recherche, il peut être utile d'élaborer une théorie du changement et un modèle logique⁴ pour le programme.

Théorie du changement

Un résumé succinct de la façon dont les activités d'un programme entraînent des changements à court, moyen et long terme.

Modèle logique

Un « instantané » visuel à jour des composantes du programme pour simplifier le programme et éclairer la planification de l'évaluation.

Ces outils peuvent fournir des éclaircissements au personnel, aux intervenants, aux évaluateurs et aux bailleurs de fonds sur un certain nombre de facteurs, y compris les hypothèses sur lesquelles repose le programme, le contexte dans lequel il est mis en œuvre et les aspects fondamentaux de la fonction du programme. Les modèles logiques sont particulièrement utiles pour établir, distinguer et orienter les relations entre les composantes, les

3 Pour de plus amples renseignements sur les critères pour des questions de recherche bien rédigées, voir la ressource de Wingate & Schroeter (2007) [ici](#).

4 Pour obtenir de plus amples renseignements sur les stratégies et les approches visant à élaborer une théorie du changement et des modèles logiques, consultez le site Web de BetterEvaluation [ici](#) ou celui du gouvernement du Canada [ici](#).

activités, les extrants et les résultats du programme pour toutes les parties participant à une évaluation (Figure 2). Il est essentiel de distinguer ces composantes du programme et de cerner d'autres facteurs externes qui ont une incidence sur la mise en œuvre du programme pour réfléchir à ce qui sera mesuré au cours d'une évaluation. Il convient de noter qu'il existe de nombreuses variantes de modèles logiques qui peuvent être utilisées pour visualiser un programme sous des angles uniques et novateurs. Il est plus important de s'entendre sur la compréhension conceptuelle du programme que sur la conception particulière utilisée. Les figures 3 et 4 présentent des exemples de modèles logiques traditionnels pour les programmes liés Consortium sur l'accès et la persévérance scolaire (CAPS).

Figure 2 : Description des caractéristiques de chaque élément du modèle logique traditionnel

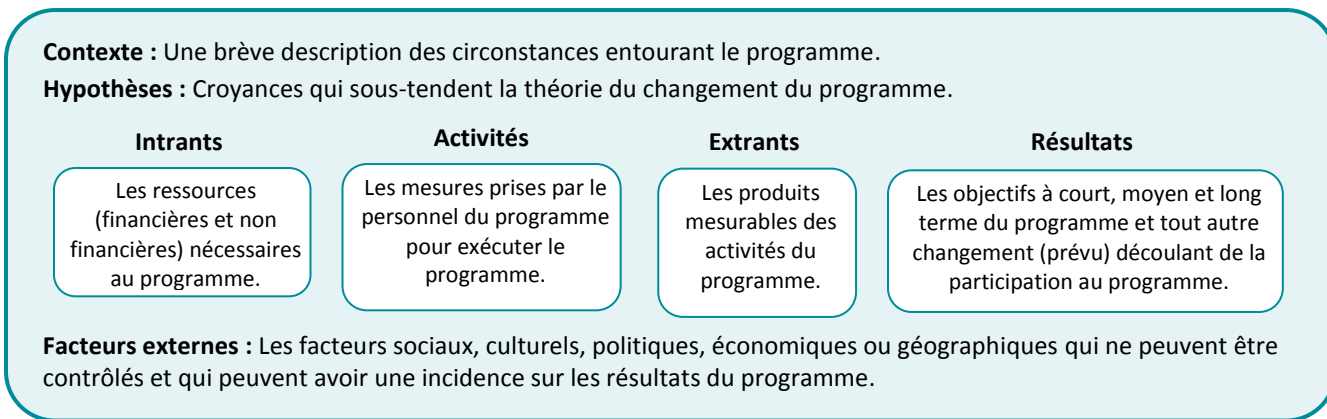


Figure 3 : Un modèle logique pour un programme qui subventionne entièrement les frais de demande d'aide financière aux EPS des étudiants à faible revenu.

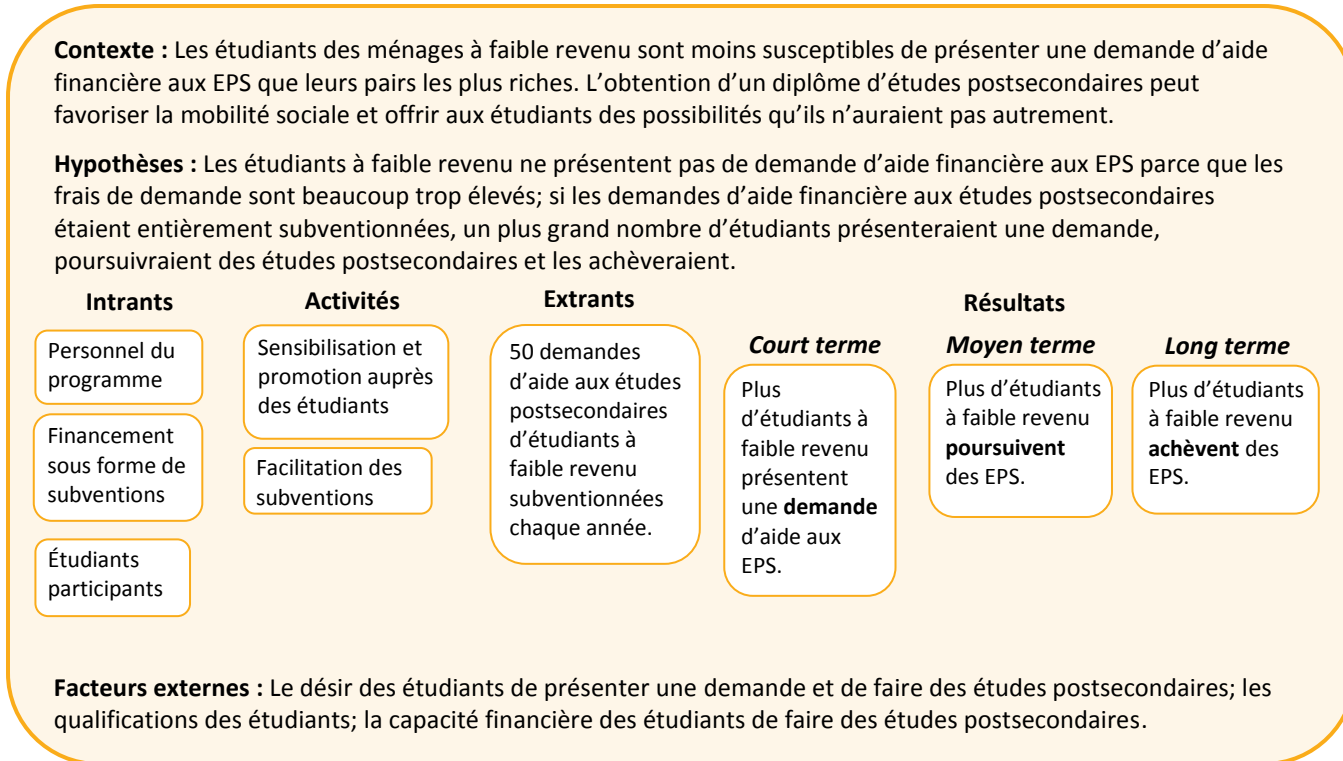
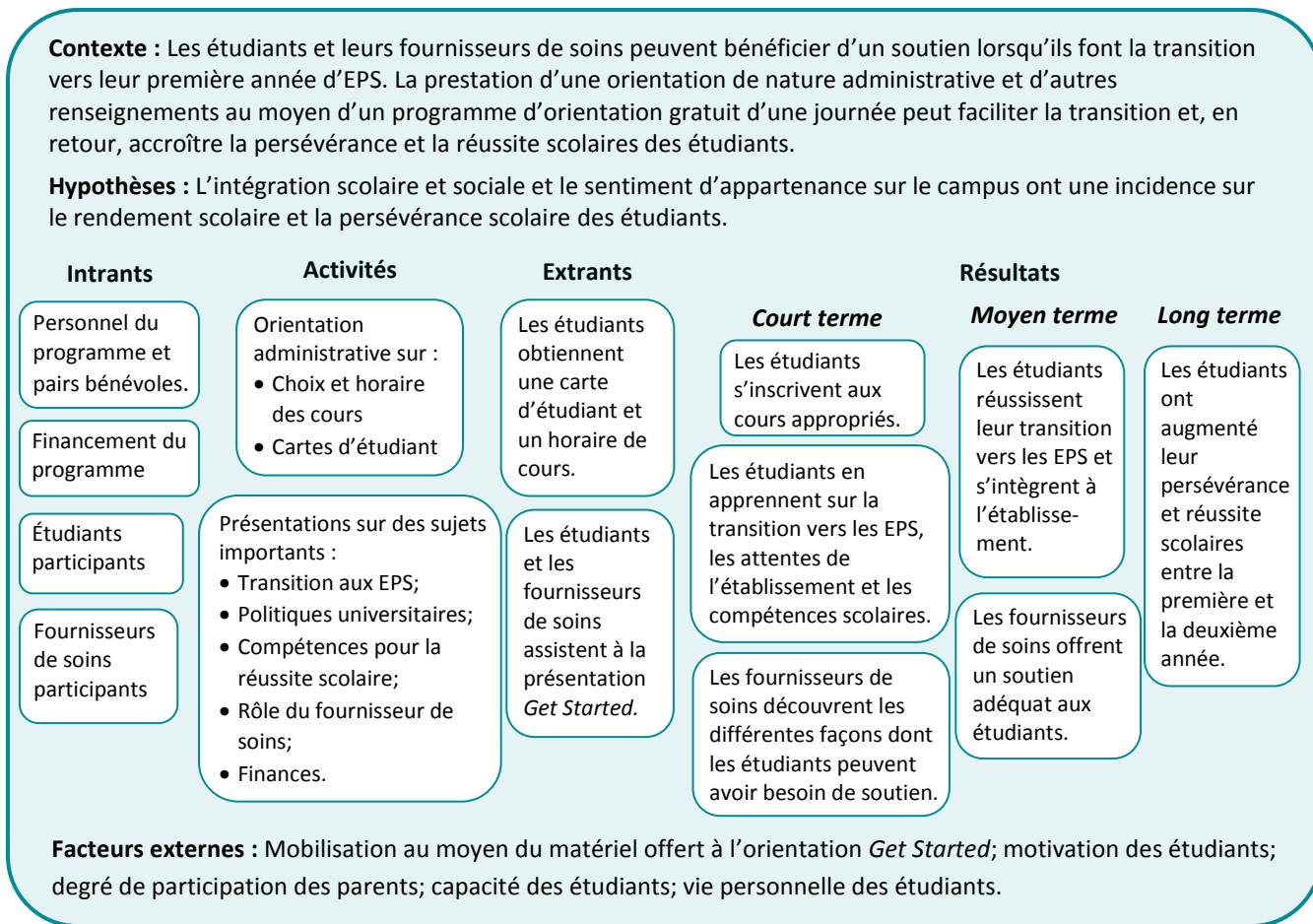


Figure 4 : Un modèle logique pour le projet du CAPS de l'UTSC visant à évaluer le programme d'orientation « Get Started »



Une fois qu'un programme a été schématisé à l'aide d'un modèle logique, le schéma peut servir de guide pour déterminer les composantes du programme qui présentent le plus d'intérêt à des fins d'évaluation. Le moment est peut-être bien choisi pour examiner d'autres évaluations qui pourraient avoir été menées sur des programmes semblables, soit de façon formelle en examinant la documentation, soit de façon informelle au moyen de conversations avec d'autres évaluateurs et le personnel du programme.

Il s'agit d'un moment choisi pour réfléchir à la façon dont les résultats de l'évaluation seront utilisés et au principal public visé par les résultats de l'évaluation (voir la Partie 3). Si l'équipe d'évaluation ne peut s'entendre sur les parties d'un programme à évaluer et sur la façon de structurer les questions de recherche en conséquence, plusieurs stratégies peuvent être utilisées pour susciter les discussions. Par exemple, la réalisation d'une analyse FFPM est une approche fréquente pour se limiter aux composantes d'intérêt d'une évaluation⁵.

Analyse FFPM
Réflexion sur les forces, les faiblesses, les possibilités et les menaces de son programme, qui vise à appuyer l'élaboration de questions de recherche.

⁵ Les personnes qui désirent un exemple d'analyse FFPM peuvent consulter le [site Web](#) de BetterEvaluation.

Dans les prochaines sections, nous discuterons de la façon dont les questions de recherche doivent être utilisées comme cadre de référence pour de nombreuses autres décisions relatives à la conception de l'évaluation.

Partie 2 : Définir des variables clés

Avant de procéder à l'évaluation, il faut réfléchir sérieusement à la façon de répondre à chaque question de recherche. Pour ce faire, des variables clés doivent être sélectionnées et définies.

Variables

Indicateurs tangibles du concept ou du résultat mesuré qui peuvent être classés comme directs ou indirects.

Comme il est mentionné à la Figure 1, les questions de recherche et les variables connexes peuvent être contrôlées au regard des quatre critères suivants : précis, mesurable, faisable et pertinent. Le processus d'identification des variables consiste à vérifier si les questions de recherche sont précises et mesurables (Figure 1). Pour certaines questions de recherche, il peut y avoir une variable clé facile à mesurer, tandis que dans d'autres cas, les questions plus complexes peuvent exiger une combinaison de variables multiples ou indirectes. Par exemple, les concepts subjectifs comme la satisfaction des étudiants à l'égard de leur expérience à la première année seraient plus difficiles à mesurer que les concepts objectifs, comme le taux de rétention d'un établissement entre la première et la deuxième année. L'utilisation de variables multiples ou indirectes est assez courante dans l'évaluation de programme, car il est souvent difficile de cerner des variables uniques. Cela dit, les évaluateurs doivent éviter d'utiliser plus de quatre variables de substitution par question de recherche, car cela pourrait laisser entendre que le concept ou le résultat évalué devrait être mieux défini.

Précis

Mesurable

Faisable

Pertinent

Étude de cas : Déterminer des variables pour répondre à la question de recherche

Disons qu'une évaluation du programme de demande de subventions à partir de la figure 3 vise à évaluer le résultat à court terme du programme selon l'augmentation du nombre d'étudiants à faible revenu qui présentent une demande d'aide financière aux EPS. Les données sur les demandes d'aide comme variable sont simples et faciles à extraire, surtout lorsque ces données sont centralisées. Cependant, il est plus difficile s'il s'agit d'étudiants à faible revenu, car nombre d'entre eux sont susceptibles de ne pas connaître le revenu de leur ménage. Pour mesurer indirectement cette information, les évaluateurs pourraient créer une variable substitutive qui leur permettrait de recueillir les codes postaux des étudiants et de les apparier aux données du recensement afin de déterminer un revenu moyen du ménage par indicatif régional (voir le tableau 1).

Lorsqu'on utilise des variables multiples ou indirectes, il peut être particulièrement utile de déterminer quelles variables serviront à répondre à chaque question de recherche. Ce processus pourrait simultanément susciter une réflexion sur la population d'intérêt ainsi que sur les méthodes de collecte et les sources des données. Quoi qu'il en soit, ces sujets seront abordés plus en détail dans les parties 6 et 7. Le Tableau 1 utilise une question de recherche potentielle de l'évaluation décrite à la Figure 3 à des fins de démonstration.

Tableau 1 : Détermination des variables utiles pour les questions de recherche⁶

Question de recherche	Variable(s)	Méthode de collecte et source des données
Le programme de demande de subventions a-t-il contribué à une augmentation du nombre d'étudiants à faible revenu qui présentent une demande d'aide aux EPS?	Nombre de demandes d'aide aux EPS.	Analyse des bases de données sur les demandes des collèges et des universités de l'Ontario (SACO et OUAC).
	Quintile de revenu selon : a) le code postal; b) les données du recensement.	Analyse documentaire : a) des données administratives sur les EPS; b) des données du recensement de Statistique Canada.

Une fois que l'équipe d'évaluation est convaincue que les questions de recherche et les variables connexes sont *précises et mesurables*, le prochain facteur à considérer porte sur la *faisabilité* de recueillir des données sur chaque variable (Figure 1). Autrement dit, les chercheurs devraient tenir compte de la portée et de la maturité du programme, des ressources et du calendrier disponibles, ainsi que de la capacité de l'équipe d'évaluation. Il est très important de tenir compte de la portée et de la maturité du programme : en règle générale, les programmes qui en sont aux premières étapes de la mise en œuvre se prêteront mieux aux évaluations des processus, tandis que les programmes matures pourraient être évalués par rapport au processus ou aux résultats (Partie 4).

Étude de cas : Examen de la faisabilité portant sur les questions de recherche et les variables connexes

Supposons que le COQES reçoit le mandat d'évaluer les résultats du programme de demande de subventions décrit à la figure 3 à l'échelle de la province. Si le COQES devait comparer la faisabilité de l'examen des résultats des programmes à court, moyen et long terme (c.-à-d. les demandes d'aide aux EPS des étudiants, le taux de fréquentation des programmes d'EPS et les taux d'achèvement des études postsecondaires, respectivement), il y aurait plusieurs facteurs à prendre en considération. Sur le plan logistique, la collecte de données sur les demandes d'aide financière aux EPS à partir d'une base de données centralisée (OUAC/SACO) pourrait sembler beaucoup moins exigeante en ressources que la collecte de données sur la fréquentation/l'obtention du diplôme de chaque établissement. Cependant, la faisabilité dépend aussi des personnes autorisées à accéder à ces données. Ainsi, les relations avec les intervenants et les ententes de partage de données connexes ne peuvent être négligées. En pareil cas, le COQES s'attendrait à ce que les chercheurs et les administrateurs des établissements aient un accès plus facile aux données sur la fréquentation des programmes et l'obtention du diplôme, tandis que les employés de l'OUAC ou du SACO pourraient avoir un accès plus facile aux données sur les demandes. De toute évidence, le choix des questions de recherche réalisables et des variables connexes dépendrait non seulement des ressources disponibles, mais aussi des relations avec les personnes qui ont accès aux données requises.

⁶ Pour de plus amples renseignements sur l'identification des variables utiles pour les questions de recherche, consultez le site Web du gouvernement du Canada [ici](#).

Le dernier facteur à considérer en ce qui concerne les questions de recherche et les variables connexes consiste à en déterminer la pertinence; c.-à-d. l'intérêt et l'importance pour le public cible. Une fois que les évaluateurs ont une idée des données à leur disposition, ils peuvent examiner les questions auxquelles ils peuvent répondre à l'aide de celles-ci. Les données *quantitatives* (c.-à-d. numériques) sont souvent utiles pour comprendre le fonctionnement d'un programme, tandis que les données *qualitatives* (c.-à-d. narratives/anecdotiques) peuvent être considérées comme plus utiles pour expliquer l'effet ou la raison d'être du mode de fonctionnement d'un programme. La sélection des variables peut prendre la forme d'un processus itératif dans le cadre duquel les évaluateurs peuvent consulter leur public pour connaître ses préférences en matière de données, car ses perceptions à l'égard de celles-ci peuvent influencer considérablement sa perception à l'égard de la crédibilité des constatations. Bien que ces deux types de données — quantitatives et qualitatives — puissent être avantageux dans des contextes différents, la meilleure façon de répondre adéquatement à toutes les questions de recherche consiste habituellement à tirer parti de ces deux types d'information. La prochaine section traitera plus en détail de la mobilisation des intervenants, au cours de laquelle la pertinence des questions de recherche et des variables connexes pourra être examinée plus en profondeur.

Partie 3 : Déterminer la portée du projet et qui devrait participer à une évaluation

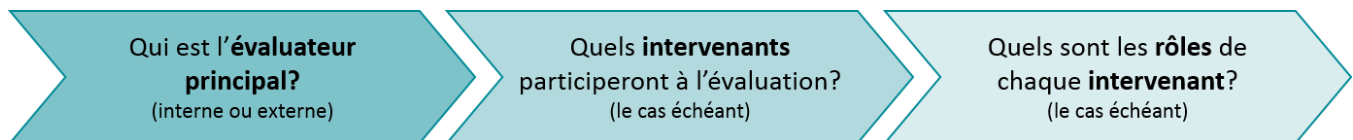
Une fois les questions de recherche et les variables connexes établies, il est opportun de déterminer qui participera au projet d'évaluation. Les intervenants peuvent être définis différemment selon les contextes, mais aux fins du présent manuel, ils incluront tous ceux qui sont touchés par un programme ou qui s'y intéressent, qu'ils œuvrent à l'intérieur ou à l'extérieur du programme. Il est important d'évaluer l'intérêt des intervenants à l'égard d'une évaluation et de définir leurs rôles dans le cadre du projet afin d'examiner certaines questions de recherche pour déterminer si elles sont pertinentes. Cela sera particulièrement important pour les partenariats institutionnels comportant des ententes de partage de données, dans le cadre desquels les données et les ressources sont combinées pour répondre à diverses questions d'intérêt pour différents groupes. Quoi qu'il en soit, les évaluateurs peuvent constater que les questions de recherche initiales seront considérablement peaufinées une fois qu'elles auront été examinées par les intervenants concernés, ce qui démontre la nature itérative de la planification de l'évaluation. Avant d'aborder les facteurs à considérer relatifs aux intervenants, il est recommandé de déterminer la portée du projet d'évaluation. Pour ce faire, il peut être utile de répondre aux questions suivantes :

- ✓ Qui utilisera les résultats de l'évaluation? Dans quel but les résultats seront-ils utilisés?
- ✓ Quels sont l'échéancier et le budget approximatifs? Quelles autres ressources pertinentes sont disponibles?
- ✓ Quelle est la capacité de l'équipe d'évaluation ou de quoi a-t-elle besoin pour réaliser de l'évaluation?
- ✓ Qui doit participer à l'évaluation et de quelle façon, pour que les résultats soient utiles?

La détermination de la portée et la définition de la capacité requise pour une évaluation circonscrivent ce qui est réellement réalisable dans un délai et un budget précis (voir la Partie 11 pour de plus amples renseignements sur le budget). Enfin, il est important d'éviter de se laisser distraire par un idéalisme ou une surestimation de ce qui peut être accompli, des pièges fréquents dans les projets d'évaluation.

Au moment de déterminer qui doit participer à une évaluation, le principal facteur à considérer est l'objectif de l'évaluation. Dans le cas des programmes qui visent à contrôler les résultats des programmes après l'évaluation sur une base continue, il est essentiel d'obtenir rapidement l'adhésion du personnel des programmes et de s'adjoindre ses capacités. Si le but est d'obtenir des résultats utiles et utilisables pour un intervenant en particulier, alors le fait de les inclure dès le départ peut aider à atteindre cet objectif. Si le but de l'évaluation est plus théorique ou s'il n'exige pas l'adhésion des intervenants, il peut être plus utile de gagner du temps et d'épargner des efforts sur la mobilisation des intervenants pour se consacrer à d'autres éléments de l'évaluation.

On peut déterminer qui devrait participer à une évaluation en répondant aux questions suivantes :



Pour désigner un évaluateur principal, la première décision consiste à déterminer si cette personne œuvrera à l'intérieur ou à l'extérieur du programme évalué. Les évaluateurs internes sont choisis au sein d'une organisation ou d'un établissement pour une ou plusieurs des raisons suivantes : 1) la capacité d'évaluation interne existe; 2) le programme est simple; 3) le budget de l'évaluation est limité ou 4) l'évaluation vise principalement l'amélioration du programme plutôt que la responsabilisation (voir la Partie 4).

Évaluateur principal
Responsable d'un projet d'évaluation

Autrement, on peut retenir les services d'évaluateurs à l'extérieur de l'organisation ou de l'établissement afin de les consulter à titre d'experts en évaluation impartiaux lorsque 1) une perspective externe est essentielle; 2) le programme est vaste ou complexe; 3) il y a un budget souple pour l'évaluation ou 4) l'évaluation vise principalement la responsabilisation plutôt que l'amélioration (voir la Partie 4). Après avoir déterminé les questions de recherche ainsi que la portée et la capacité de l'évaluation, les évaluateurs principaux peuvent élaborer un plan de travail afin de déterminer l'ordre de priorités des différentes composantes de leur mission, puis trouver ou renforcer la capacité de répondre à d'autres objectifs figurant sur la liste des objectifs de l'évaluation du programme.

Une fois que l'évaluateur principal a été choisi, il faut ensuite déterminer si d'autres intervenants participeront à l'évaluation et quelle sera la mesure de leur participation. La participation des intervenants peut varier considérablement d'une région à l'autre, en grande partie en fonction des questions de recherche, de la portée du projet et de la façon dont les résultats seront utilisés. Les petits programmes ayant des ressources limitées peuvent avoir une liste relativement courte d'intervenants à prendre en considération, tandis que les programmes plus vastes ou plus complexes peuvent avoir une longue liste d'intervenants qui valent la peine d'être mis à contribution. Si l'évaluation est trop difficile à coordonner par une seule personne, une équipe d'évaluation composée d'intervenants pertinents pourrait être créée pour appuyer l'évaluateur principal. Le type d'intervenants dont il est question dépendra également du type d'organisme ou d'établissement dans lequel l'évaluation est menée (p. ex., groupes communautaires, collèges ou universités). La Figure 5 propose certains intervenants qui pourraient être inclus dans les évaluations des programmes liés à l'accès ou à la persévérance scolaire.

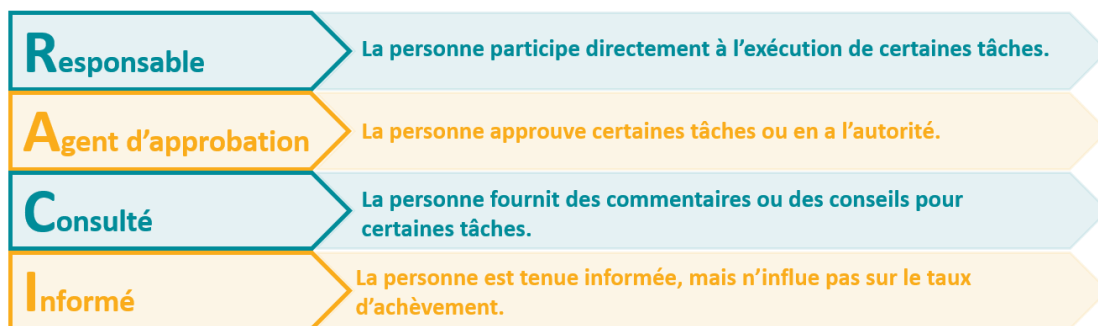
Intervenants
Personnes ou groupes qui ont un intérêt direct à l'égard du processus ou des résultats d'un projet d'évaluation.

Figure 5 : Intervenants à inclure éventuellement dans une évaluation de programmes importants ou complexes



Quelle que soit la taille d'un programme ou de l'équipe d'évaluation correspondante, il sera utile de déterminer les rôles de chaque intervenant en catégorisant sa contribution attendue à des tâches importantes⁷. La Figure 6 représente une approche commune de la caractérisation des rôles des intervenants.

Figure 6 : Caractérisation de la participation des intervenants à l'aide de l'acronyme RACI.



Le processus de définition des rôles des évaluateurs et des intervenants relève plus de l'art que de la science. À l'aide des questions de recherche et de la portée de l'évaluation, les évaluateurs devraient pouvoir justifier l'inclusion ou l'exclusion des intervenants dans le projet d'évaluation. Lorsque les intervenants sont mobilisés,

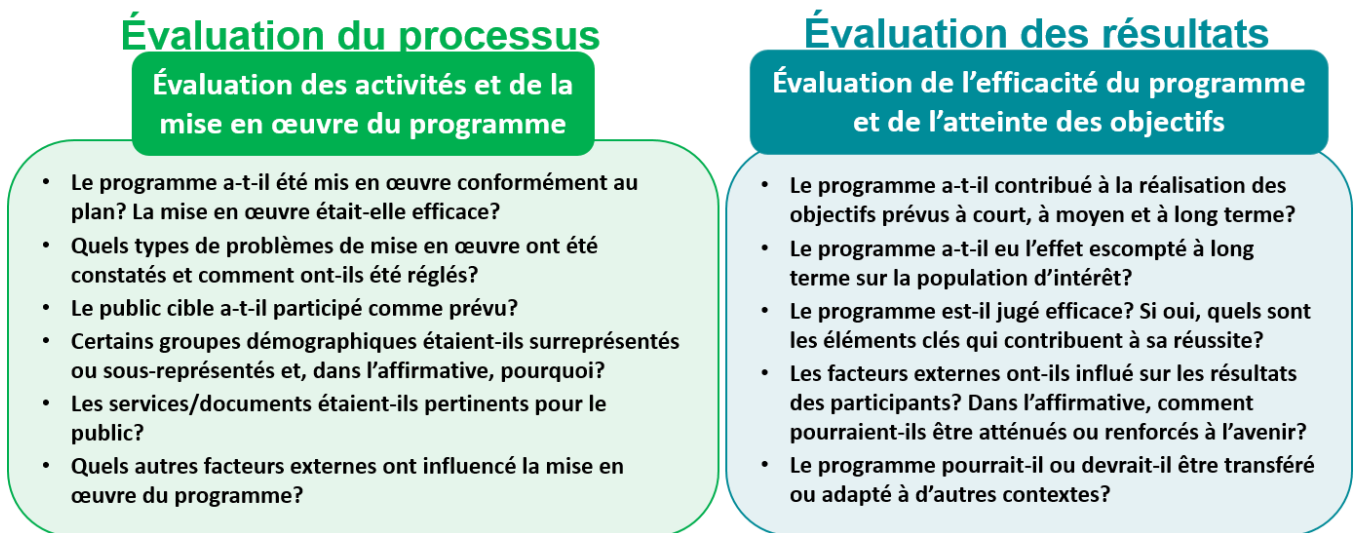
⁷ Pour tout intervenant qui ne sera pas responsable, agent d'approbation, consulté ou directement informé d'une évaluation (p. ex., contribuables, futurs participants), ces points de vue peuvent quand même être pris en compte lorsque les évaluateurs déterminent si le plan de mise en œuvre et de production de rapports d'évaluation est conforme à l'éthique pour toutes les parties (voir la Partie 10).

cela peut accroître la probabilité de mettre à contribution leur expertise, faciliter leur acceptation des résultats et inspirer leur action pour l'amélioration des programmes. Toutefois, il n'est pas toujours possible ou avantageux de maximiser la participation des intervenants, selon la nature du projet. Parfois, il peut être nuisible d'avoir trop de cuistots dans la cuisine, et les évaluateurs doivent éviter de faire participer inutilement ou trop tôt les intervenants à un projet. En fin de compte, que les intervenants soient sciemment inclus ou exclus, les évaluateurs doivent s'efforcer de communiquer de façon claire, cohérente et respectueuse.

Partie 4 : Choisir une approche d'évaluation

Après avoir établi les questions de recherche et choisi l'évaluateur ou l'équipe d'évaluation, la prochaine étape pourrait consister à définir une approche d'évaluation globale. L'approche adoptée pour l'évaluation d'un programme porte sur l'objet de l'évaluation et devrait être révélée par les questions de recherche. Le plus souvent, l'approche d'évaluation est classée selon qu'il s'agit d'une évaluation du processus ou d'une évaluation des résultats (Figure 7)⁸. Dans le cas des évaluations comportant des questions de recherche multiples, les évaluateurs peuvent adopter une approche hybride qui incorpore des composantes de l'évaluation du processus et de l'évaluation des résultats. Il est très important de distinguer les questions qui se prêtent à chaque type d'approche, car cela dictera d'autres choix à faire éventuellement.

Figure 7 : Comparaison entre les évaluations portant sur les processus et les évaluations portant sur les résultats et des questions de recherche connexes



Source : Adapté de Chen (2014) et Innovation Network (s.d.).

⁸ Certains évaluateurs pourraient préférer le terme « évaluation de la mise en œuvre » à « évaluation du processus ».

Les évaluateurs devront peut-être aussi préciser l'objet de l'évaluation du programme (c.-à-d. comment les résultats seront utilisés). Encore une fois, le but de l'évaluation devrait être clairement extrapolé à partir de questions de recherche bien rédigées, qui peuvent être classées en deux métacatégories, constructive ou concluante.



Évaluations constructives

Évaluations entreprises pour améliorer un programme (parfois appelées formatives ou axées sur l'apprentissage)



Évaluations concluantes

Évaluations entreprises pour prouver l'efficacité d'un programme (parfois appelées « sommatives » ou « axées sur la responsabilisation »)

Bien qu'il puisse sembler que les évaluations de processus seraient associées à des questions d'évaluation constructive, et les évaluations de résultats à des questions d'évaluation concluante, ce n'est pas toujours le cas. Les résultats d'un processus ou d'une évaluation des résultats peuvent être utilisés à des fins constructives ou concluantes. Et les évaluations comportant plusieurs questions de recherche pourraient s'inscrire dans une autre catégorie hybride, c'est-à-dire qui réunit des questions d'évaluation constructive et des questions d'évaluation concluante. Ce qui est important ici, c'est que les évaluateurs sachent comment distinguer les objectifs de chaque question de recherche, car cela sera crucial pour choisir une conception d'étude (Partie 5).

Étude de cas : Exemples d'évaluation constructive axée sur les résultats et d'évaluation concluante axée sur le processus

Comparons quelques approches différentes pour évaluer le programme de demande de subventions à la figure 3. Une évaluation constructive des résultats pourrait sonder les bénéficiaires de subventions un an après la demande pour déterminer s'ils poursuivent des EPS (un résultat à moyen terme) et s'ils ont des commentaires constructifs au sujet de la promotion du programme pour les prochaines cohortes d'étudiants. Autrement, au lieu d'attendre un an, on pourrait procéder à une évaluation concluante du processus pour évaluer l'efficacité de la stratégie de promotion pendant la mise en œuvre initiale du programme de subvention. Si cette évaluation révélait que le plan de promotion était complètement inefficace et que la plupart des étudiants de la population cible n'avaient pas fait de demande de subvention, les intervenants pourraient abolir le programme. Cela permettrait de réaffecter les ressources afin de mieux servir la population cible en temps opportun.

En plus d'être constructives ou concluantes, les évaluations pourraient aussi viser un objectif principal qui ne correspond pas parfaitement à ces deux métacatégories. Ces évaluations pourraient avoir pour but d'établir la confiance et la légitimité auprès des intervenants, d'enrichir une base de données probantes plus vaste, ou d'évaluer la possibilité d'élargir ou de transférer un programme à d'autres contextes, etc.

Étude de cas : Projet du CAPS de la Hamilton Community Foundation au moyen d'une évaluation hybride à phases multiples

Pour évaluer un programme pilote au niveau de l'école intermédiaire pour les étudiants confrontés à des obstacles multiples à l'accès aux EPS, la Hamilton Community Foundation utilise une approche d'évaluation hybride à phases multiples. Étant donné qu'il s'agit d'un nouveau programme et qu'il repose sur une nouvelle théorie du changement, l'évaluation s'amorce par une approche constructive axée sur le processus visant à cerner les composantes de base du programme ainsi que les buts et les objectifs du projet pilote. Après avoir répondu à ces questions, la Hamilton Community Foundation a l'intention de se doter d'une approche d'évaluation plus concluante axée sur les résultats afin de cerner les résultats du programme (prévus et imprévus) pour les participants, ainsi que l'évolutivité du programme pour d'autres contextes. L'utilisation d'approches et questions d'évaluation multiples exige une conception d'étude à méthodes mixtes.

En fin de compte, peu importe la complexité de l'évaluation, il est important d'harmoniser chaque question de recherche avec 1) une enquête sur le ou les processus ou les résultats, et 2) un objectif distinct, constructif, concluant ou un autre facteur principal. Sans nous attarder sur la sémantique des types d'évaluation hybrides, ou sur les nombreuses approches d'évaluation uniques qui existent, nous espérons que le fait de définir quelle partie du programme est évaluée et comment les résultats seront utilisés aidera les évaluateurs à établir leur approche d'évaluation globale. En établissant ces paramètres, il devrait être beaucoup plus simple de choisir une conception d'étude appropriée.

Partie 5 : Choisir une conception d'étude

Après avoir mis la dernière main aux questions de recherche et établi une approche d'évaluation, l'étape suivante consiste à choisir une conception d'étude appropriée. Certaines questions de recherche nécessiteront (ou empêcheront) l'utilisation de conceptions d'étude particulières. La conception d'étude est étroitement liée à l'objectif de l'évaluation, de sorte que ce choix dépendra de l'utilisation des résultats à des fins constructives (axées sur l'amélioration) ou concluantes (axées sur la responsabilisation). De plus, les évaluations des résultats exigeront généralement l'utilisation d'expériences ou de quasi-expériences, tandis que les évaluations des processus peuvent utiliser n'importe quelle conception, mais le plus souvent, elles s'harmonisent avec des conceptions non expérimentales⁹. Les trois conceptions sont décrites ci-après.

⁹ Pour de plus amples renseignements sur le choix d'une conception d'étude, consultez le site Web du Pell Institute [ici](#).

Les conceptions expérimentales évaluent les effets objectifs d'un programme en attribuant des participants au hasard à un groupe témoin ou à un groupe d'intervention afin de générer des groupes équivalents à des fins de comparaison. Chaque groupe est habituellement évalué au moins à deux reprises : avant et après le programme ou l'intervention.

Groupe d'intervention

Personnes qui participent au programme (l'affectation aléatoire dépend de la conception d'étude).

Groupe témoin

Personnes affectées au hasard pour ne pas participer au programme.

Également connues sous le nom d'essais contrôlés randomisés (ECR), les expériences ont toujours été considérées comme l'étalon-or en matière d'évaluation de programme parce qu'on pense qu'elles fournissent la preuve la plus crédible d'une relation causale entre un programme et un résultat.

Étude de cas : Projet d'ARC du Mohawk College utilisant un essai contrôlé randomisé

Cette étude a utilisé un modèle expérimental pour mesurer les effets des conseils proactifs, un service étudiant dans le cadre duquel le personnel aide les étudiants à établir des objectifs personnels et scolaires tout au long de leur première année. Dans le cadre de cette étude, environ 4 500 étudiants ont été affectés à un groupe témoin ou à l'un des deux groupes d'intervention (pour recevoir des conseils de groupe ou des conseils individuels). Cette conception a permis aux chercheurs d'effectuer des analyses statistiques pour comparer les effets des conseils de groupe et des conseils individuels aux données de base générées par le groupe témoin (c.-à-d. aucun conseil).

Les conceptions quasi expérimentales évaluent les effets d'un programme en analysant les participants des groupes de comparaison et d'intervention. Les groupes d'intervention et de comparaison sont habituellement évalués à deux reprises : avant et après le programme. Les évaluateurs qui utilisent des conceptions quasi expérimentales doivent se rappeler qu'une analyse statistique avancée est habituellement nécessaire pour cette méthode afin de contrôler les différences entre les groupes, de manière à isoler les résultats du programme¹⁰.

Groupe de comparaison

Une version adaptée d'un groupe témoin qui évite la randomisation en sélectionnant un groupe de non-participants aussi semblable que possible au groupe d'intervention afin de réduire au minimum les biais.

¹⁰ Même si le groupe témoin n'a pas reçu l'intervention, il est tout de même évalué en même temps que le groupe d'intervention dans des quasi-expériences.

Étude de cas : Projet du CAPS de l'Université de Toronto à Scarborough au moyen d'une quasi-expérience

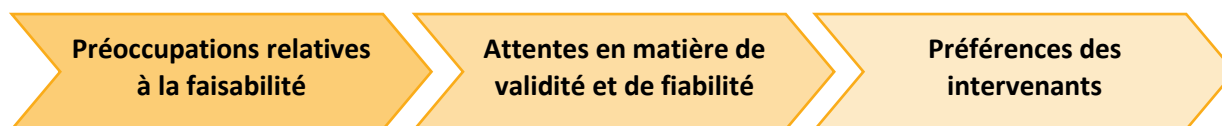
Cette étude utilisera une conception quasi expérimentale pour mesurer l'impact du programme d'orientation de la première année « *Get Started* » (décrit à la figure 4). À l'aide d'une méthode d'appariement des coefficients de propension, une technique statistique qui crée un groupe de comparaison en appariant les caractéristiques démographiques au groupe d'intervention, les chercheurs évalueront si la participation à un programme a une incidence sur la persévérance et la réussite scolaires des participants, en mettant particulièrement l'accent sur les effets pour les étudiants de première génération et les étudiants à faible revenu.

Les conceptions non expérimentales évaluent le fonctionnement d'un programme en analysant principalement le groupe d'intervention. Certains modèles peuvent suivre les participants au fil du temps, tandis que d'autres peuvent n'évaluer les participants qu'une seule fois, après le programme ou l'intervention. Plutôt que d'évaluer les effets d'un programme, on utilise souvent une évaluation non expérimentale comme outil descriptif pour comprendre les éléments d'un programme. Dans certains contextes, cette conception d'étude semble la plus réaliste, surtout lorsque la création d'un groupe témoin ou de comparaison est impossible, contraire à l'éthique ou inabordable. Malgré tout, les efforts déployés pour trouver un moyen de comparer les résultats des participants au programme (p. ex., l'analyse comparative par rapport aux normes ou aux cibles, ou l'utilisation des données du recensement) peuvent servir à l'établissement d'un point de référence utile.

Étude de cas : Projet du CAPS « Success Beyond Limits » à l'aide d'une évaluation non expérimentale

Pour évaluer son programme de mentorat dirigé par des jeunes, Success Beyond Limits (SBL) mise sur la Recherche-Action Participation pour les jeunes. Cette approche permettra aux anciens de SBL de mener des entrevues avec leurs anciens pairs afin de mieux comprendre le rôle du programme SBL sur les trajectoires et les résultats des EPS. Cette approche peut être bénéfique lorsqu'il s'agit de recueillir des données de façon rétroactive, car elle peut permettre de recueillir des données auprès des participants même après qu'ils ont participé au programme.

En plus d'avoir une compréhension générale de ce que chacune de ces conceptions d'étude comporte, les évaluateurs devront également garder à l'esprit les facteurs à considérer suivants :



Pour ce qui est de la faisabilité, la mise en œuvre des ECR et des quasi-expériences peut exiger beaucoup de ressources. Outre le temps et le budget requis, ces conceptions exigent habituellement des évaluateurs expérimentés possédant une vaste expertise en analyse de données. De plus, les considérations éthiques peuvent constituer un obstacle important, car la randomisation des participants à certains programmes peut ne pas être justifiable (voir la Partie 10). Une exception à cette règle concerne les programmes pour lesquels il existe une base de données limitée, car le recours aux ECR semble plus facilement justifié lorsqu'il s'agit de connaître les effets d'un programme pour la première fois.

Sur le plan de la validité et de la fiabilité, les ECR ont toujours eu tendance à produire les meilleurs résultats, suivis de près par les quasi-expériences. Toutefois, la validité inhérente des ECR a été remise en question dans le cadre des évaluations de programme. Bien qu'il soit bien établi que les expériences traditionnelles ont la plus grande validité interne en ce qui concerne toute conception d'étude, sur le plan de la comparabilité, ces conceptions d'étude ont une validité externe inférieure. Cela s'explique par le fait que les facteurs externes normalement pris en compte ou atténués par le personnel du programme pour assurer la réussite d'un programme dans son contexte réel sont éliminés dans la conception de l'étude expérimentale. Par conséquent, les résultats des ECR ne sont pas toujours généralisables au contexte même dans lequel le programme est exécuté normalement.

Validité interne

Mesure dans laquelle une étude permet de cerner la validité de son objet.

Validité externe

Généralisabilité des résultats d'une étude à d'autres populations, contextes, variables et mesures.

Le dernier facteur à considérer lors du choix d'une conception d'étude pourrait consister à concilier les préférences des intervenants, au besoin. Quels que soient les autres facteurs à considérer, les évaluateurs peuvent être limités par les opinions préconçues des intervenants quant à la rigueur de certaines conceptions d'étude.

L'élément le plus important est que les évaluateurs utilisent la question de recherche et l'approche d'évaluation connexe pour orienter leur choix d'une conception d'étude appropriée. En cernant la mesure dans laquelle les avantages et les inconvénients de la conception d'étude choisie s'harmonisent avec les objectifs de l'évaluation, ils peuvent acquérir un certain degré de confiance quant à leur capacité de justifier leur choix pour les étapes suivantes.

B Effectuer une évaluation

Partie 6 : Collecte des données

Une fois que la conception de l'étude a été choisie, il faut planifier la méthode de collecte des données. En se reportant au Tableau 1 (Partie 2), les évaluateurs peuvent catégoriser les variables requises pour chaque question de recherche selon leur nature quantitative ou qualitative. Si un seul type de variable est nécessaire, les évaluateurs utiliseront uniquement des méthodes quantitatives ou uniquement qualitatives pour répondre à la question. En revanche, si les évaluateurs doivent s'en remettre à deux types de variables pour répondre aux questions de recherche, ils préconiseront une approche reposant sur des méthodes mixtes.

Les méthodes mixtes gagnent en popularité dans le domaine de l'évaluation des programmes parce qu'elles permettent de surmonter les inconvénients associés à l'utilisation de sources de données uniques. Les méthodes mixtes qui reposent sur des sources de données multiples (c.-à-d. variables) et les techniques de collecte (p. ex., sondages, entrevues) offrent une meilleure validité et fiabilité. Non seulement les sources de données multiples permettent la triangulation des données, mais elles peuvent aussi éclairer les résultats de différentes façons et, par conséquent, accroître la profondeur de l'analyse et la fiabilité des rapports.

Triangulation des données

La validation des données par la vérification croisée de deux sources ou techniques ou plus.

Bien que nous puissions classer des techniques de collecte de données particulières en fonction du type de données qu'elles produisent, cet exercice est compliqué par le fait que de nombreuses techniques peuvent être utilisées pour produire des formes de données multiples (c.-à-d. des sondages peuvent utiliser des questions ouvertes ou fermées et produire des données qualitatives ou quantitatives, ou les deux). Pour cette raison, les techniques de collecte des données sont souvent classées selon la source des données, de la façon suivante¹¹ :

- **Examen de la documentation existante** — examen des bases de données, des statistiques officielles, des archives, des dossiers financiers, de la recension des écrits, des dossiers de projet, des résultats d'évaluation antérieurs, etc.
- **Observation** — observation directe ou indirecte, visites sur place, examen de photos ou de vidéos, etc.
- **Collecte de renseignements auprès de personnes** — entrevues (dirigées ou non dirigées)¹², journaux/registres, sondages (questions ouvertes ou fermées), etc.
- **Collecte de renseignements auprès de groupes** — groupes de discussion, schématisation des concepts, débats, etc.

Dans chaque catégorie susmentionnée, plusieurs facteurs seront pris en considération pour choisir la technique de collecte des données la plus appropriée. L'analyse des documents internes d'un établissement ou d'une organisation peut permettre aux évaluateurs d'en comprendre les ressources, les valeurs, les processus, les priorités et les préoccupations afin de cerner les forces et les inefficacités. L'analyse des dossiers publics peut se révéler essentielle pour établir des comparaisons à grande échelle entre les groupes et les collectivités. L'observation des programmes en action peut fournir aux évaluateurs une perspective holistique du contexte

11 Pour de plus amples renseignements sur les techniques de collecte de données, consulter le site Web de BetterEvaluation [ici](#).

12 Dans le cas des entrevues et des groupes de discussion, les évaluateurs devraient déterminer comment saisir des renseignements détaillés avec fiabilité. S'il n'est pas possible de prendre des notes détaillées, il est recommandé d'utiliser un appareil d'enregistrement, surtout s'il est pertinent de citer directement les participants.

dans lequel le programme est exécuté et peut être utile pour cerner des problèmes que le personnel ou les participants ne connaissent pas, veulent éviter ou ne veulent pas aborder dans une entrevue. La collecte de renseignements auprès de personnes ou de groupes comporte plusieurs avantages et inconvénients. Ceux-ci sont résumés au Tableau 2.

Tableau 2 : Comparaison des techniques de collecte de renseignements auprès des personnes

Description	Avantages	Inconvénients
<p>Sondages à questions fermées</p> <p>Recueil de l'information auprès de grands groupes sur des sujets qui ne sont ni complexes ni délicats à l'aide de questions clairement définies et de réponses prédéterminées.</p>	<ul style="list-style-type: none"> Grande fiabilité Collecte des données efficace Analyse des données simple Faible coût 	<ul style="list-style-type: none"> Difficile à élaborer L'autodéclaration peut causer des biais Impersonnel et manque de contexte Validité limitée Ne convient pas aux sujets complexes ou délicats
<p>Sondages à questions ouvertes¹³</p> <p>Recueil de l'information auprès de groupes de taille moyenne pour obtenir des réponses ouvertes à des questions clairement définies sur des sujets qui peuvent être assez complexes ou délicats.</p>	<ul style="list-style-type: none"> Validité modérée Collecte des données efficace Coût modéré 	<ul style="list-style-type: none"> Difficile à élaborer L'autodéclaration peut causer des biais Impersonnel Analyse des données difficile Fiabilité limitée
<p>Entrevues</p> <p>Recueil de riches renseignements contextuels ou comportementaux sur les sentiments ou les opinions d'un petit échantillon de personnes bien informées.</p>	<ul style="list-style-type: none"> Produit des données instructives Permet un contact interpersonnel Grande validité Convient aux répondants complexes, sensibles et de haut niveau 	<ul style="list-style-type: none"> Exige des intervieweurs qualifiés Collecte de données laborieuse Analyse des données difficile Fiabilité limitée Coût élevé
<p>Groupes de discussion</p> <p>Recueil des renseignements relativement non délicats auprès de groupes de 8 à 12 personnes lorsque les interactions de groupe peuvent être exploitées pour stimuler des réponses riches en renseignements.</p>	<ul style="list-style-type: none"> Produit des données instructives Permet un contact interpersonnel Grande validité Collecte de données moyennement efficace Coût modéré 	<ul style="list-style-type: none"> Exige des intervieweurs qualifiés Analyse des données difficile Fiabilité limitée
<p>Observation</p> <p>Recueil des données sur une vaste gamme de comportements afin de cerner des problèmes que le personnel du programme ou les participants ignorent ou dont ils ne peuvent ou ne veulent pas discuter.</p>	<ul style="list-style-type: none"> Produit des données directes et globales sur les comportements Fournit un bon sens du contexte à l'évaluateur Se déroule dans des contextes naturels et non dirigés 	<ul style="list-style-type: none"> Exige des observateurs qualifiés Risque de biais dans les données (p. ex., participants méfiants, perception sélective) Analyse des données difficile Fiabilité limitée Coût élevé

Source : Adapté de National Science Foundation, 2002; Baker, s.d.

¹³ Les questions fermées et ouvertes peuvent être utilisées efficacement dans un même outil d'enquête.

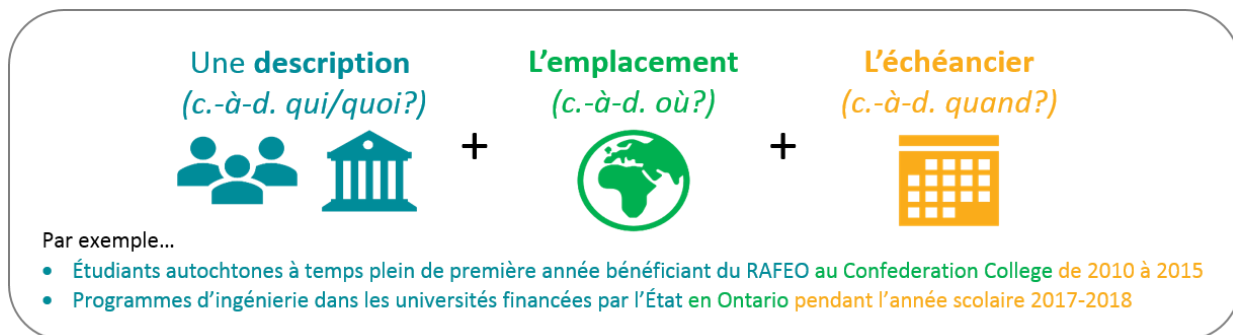
Après avoir convenu de la façon de recueillir toutes les données d'une évaluation, les évaluateurs peuvent créer un plan de travail relatif aux données afin de maintenir l'organisation du processus de collecte et d'analyse des données. Les structures de répartition du travail sont un outil couramment utilisé en gestion de projet et peuvent servir à déterminer les produits livrables et à identifier les responsables de chaque tâche à accomplir pendant l'évaluation¹⁴. Il sera également essentiel d'établir un calendrier pour s'assurer que l'équipe dispose de suffisamment de temps pour procéder à l'analyse des données et à la rédaction de rapports. Un diagramme de Gantt est un outil couramment utilisé pour visualiser les échéanciers de projets complexes dans lesquels plusieurs personnes sont mobilisées dans le cadre de diverses activités assorties d'échéances précises¹⁵. Ces types d'outils peuvent appuyer d'autres aspects de la planification de l'évaluation en ce qui concerne le budget et l'affectation des ressources (voir la [Partie 11](#)).

Partie 7 : Échantillonnage de la population

Une fois que la conception de l'étude et les méthodes de collecte des données ont été établies, l'étape suivante consiste à préciser l'objet et le bien-fondé de l'évaluation. Les évaluateurs devraient utiliser les questions de recherche et la portée de l'évaluation comme base pour établir leur population d'intérêt particulière. Pour définir une population d'intérêt, les évaluateurs doivent inclure les trois composantes décrites à la Figure 8.

Population d'intérêt
Le groupe sur lequel recueillir des renseignements; composé de personnes ou de toute autre unité de données (p. ex., établissements, organisations, dossiers, etc.)

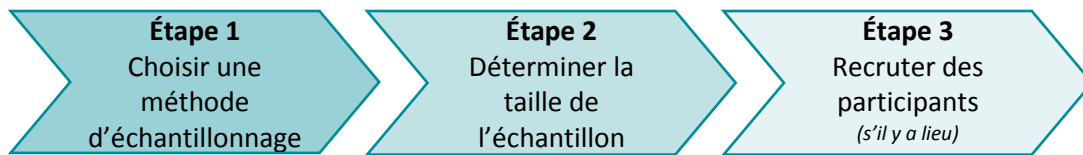
Figure 8 : Définition d'une population d'intérêt



Une fois la population d'intérêt définie, les évaluateurs détermineront quels cas (c. - à - d. membres ou unités) pourraient être recrutés pour participer à leur étude. Si la population d'intérêt est de petite taille ou si la question de recherche exige que tous les participants au programme soient inclus, les évaluateurs peuvent inclure tous les cas dans leur étude. Par exemple, si un évaluateur veut examiner l'incidence d'une séance de formation particulière sur les conseillères et conseillers de résidence dans des résidences réservées aux femmes, il pourrait inclure dans son étude tous les conseillers qui ont suivi ce programme de formation. Pour ces types d'évaluations qui comprennent l'ensemble de la population d'intérêt, les évaluateurs peuvent passer à l'étape 3 ci-dessous. Autrement, si les évaluateurs ont des populations d'intérêt de grande taille, ou s'ils veulent étudier un sous-ensemble particulier de leur population d'intérêt, ils peuvent suivre les étapes suivantes :

14 Pour de plus amples renseignements sur les structures de répartition du travail, consultez le [site Web](#) de Project Management Docs.

15 Pour de plus amples renseignements sur les diagrammes de Gantt, consultez le [site Web](#) de Project Manager.



Étape 1 : Choisir une méthode d'échantillonnage

Pour sélectionner un sous-ensemble de la population d'intérêt, les évaluateurs peuvent choisir entre un échantillonnage probabiliste et un échantillonnage non probabiliste¹⁶. Le choix de la méthode d'échantillonnage dépendra des objectifs de l'évaluation, du plan de recherche et de la faisabilité en ce qui concerne la capacité, le temps, le budget et les contraintes éthiques.

Échantillonnage probabiliste

Sélectionne un sous-ensemble en utilisant des techniques aléatoires ou quasi aléatoires pour dégager des inférences généralisées au sujet de la population d'intérêt plus générale.

Échantillonnage probabiliste. Pour utiliser l'échantillonnage probabiliste, les évaluateurs doivent disposer d'un moyen de repérer chaque cas dans la population qui est exact, exhaustif (c.-à-d. incluant l'ensemble de la population d'intérêt), à jour, sans redondance et sans dynamique quant à la façon dont les noms des membres ou des unités apparaissent dans une liste. Les échantillons peuvent ensuite être sélectionnés par :

- **Échantillonnage aléatoire** — Chaque membre ou unité de l'échantillon est choisi au hasard parmi la population et a la même probabilité d'être choisi pour constituer l'échantillon.
 - Par exemple, un programme informatique sélectionne 30 étudiants parmi une liste de tous les diplômés en psychologie de l'Université Ryerson.
- **Échantillonnage par stratification** — La population est d'abord divisée en groupes distincts pour maximiser la représentation, puis les cas sont sélectionnés au hasard dans chaque groupe.
 - Par exemple, les noms de tous les diplômés en psychologie inscrits dans les universités de l'Ontario sont d'abord triés par université, puis un programme informatique sélectionne 30 étudiants de chaque établissement.

Bien que l'échantillonnage probabiliste soit idéal pour dégager des inférences au sujet de la population d'intérêt, il n'est pas toujours nécessaire (p. ex., si la priorité consiste à dégager des renseignements sur chacun des cas) ou faisable (p. ex., il peut être coûteux ou difficile d'obtenir une liste complète de tous les membres d'une population).

¹⁶ Pour de plus amples renseignements sur les méthodes d'échantillonnage, consultez le site Web de BetterEvaluation [ici](#) ou celui du Pell Institute [ici](#).

Échantillonnage non probabiliste

Sélectionne un sous-ensemble en utilisant des techniques non aléatoires, ce qui peut permettre ou non de dégager des inférences au sujet de la population d'intérêt plus générale.

Échantillonnage non probabiliste. Si le but d'une évaluation est de dégager des renseignements sur des cas à des fins autres que la généralisation, alors l'échantillonnage non probabiliste est approprié. Les échantillons non probabilistes peuvent être sélectionnés par :

- **Échantillonnage raisonné** — Choisir un sous-ensemble fondé sur une ou plusieurs caractéristiques prédéterminées (p. ex., sélection des cas qui sont typiques, cruciaux (c.-à-d. importants d'une certaine façon), aberrants ou alignés sur un certain critère ou modèle).
 - Par exemple, recruter les étudiants de troisième année inscrits à la majeure en psychologie à l'Université Ryerson qui ont l'intention de poursuivre des études supérieures.
- **Échantillonnage de commodité** — Choisir des échantillons qui sont facilement disponibles (p. ex., choisir des participants volontaires ou utiliser un échantillonnage en boule de neige).
 - Par exemple, demander à un adjoint à l'enseignement de recruter des étudiants de l'Université Ryerson dans son bassin d'étudiants inscrits à des cours de psychologie de niveau 300.

L'échantillonnage raisonné permet aux évaluateurs de sélectionner des cas qui ont une signifiante intrinsèque par rapport à l'étude afin de dégager des inférences précises ou analytiques au sujet de la population d'intérêt plus générale. L'échantillonnage de commodité ne permet pas de dégager des inférences crédibles au sujet de l'ensemble de la population. Dans ces deux types d'échantillonnage non probabiliste, les biais de l'échantillon empêcheront les généralisations à l'ensemble de la population d'intérêt. Des mises en garde s'imposent donc lors de la présentation des constatations.

Étape 2 : Déterminer la taille de l'échantillon

Au moment de déterminer la taille de l'échantillon, il faudra faire plusieurs compromis entre l'étendue et la profondeur, en tenant compte du type de données recueillies (quantitatives et qualitatives), de l'intention de l'évaluateur de généraliser les constatations et de la faisabilité. Comme la collecte de données qualitatives peut exiger plus de ressources, les évaluateurs devront peut-être se contenter d'échantillons de plus petite taille et miser davantage sur la profondeur plutôt que sur l'étendue pour mener à bien leur mission. En revanche, dans le cas des évaluateurs qui désirent tirer des conclusions applicables à l'ensemble de leur population, la profondeur sera souvent sacrifiée au profit de la portée. En pareils cas, les évaluateurs doivent saisir un échantillon de taille suffisamment grande pour donner un aperçu précis de l'ensemble de la population afin de pouvoir établir des généralisations. Dans certains cas, un évaluateur peut utiliser une analyse de puissance pour déterminer la taille de l'échantillon qui convient afin de cerner avec précision l'effet qu'il désire mesurer¹⁷. Lorsqu'il n'est pas important pour l'évaluateur de généraliser ses résultats, il peut fonder le choix de taille d'échantillon en a) sélectionnant la taille de l'échantillon à l'avance b) échantillonnant au point de redondance ou de saturation, ou c) utilisant un échantillonnage évolutif pour ajouter des cas, au besoin, au fil du temps. Les évaluateurs peuvent même choisir d'étudier un seul cas si cela convient à leur conception d'étude.

17 Pour de plus amples renseignements sur le choix de la taille des échantillons, voir Taylor-Powell (1998) [ici](#).

Étape 3 : Recruter des participants

Pour recueillir des données prospectives auprès des individus, les techniques de recrutement sont un facteur essentiel, surtout lorsque la participation des individus à une étude est entièrement facultative. Cette étape peut souvent exiger beaucoup plus de temps que prévu, mais elle en vaut assurément la peine. Il pourrait être de mise de consulter de nouveau les intervenants à cette étape, car ils pourraient éventuellement agir comme agents de liaison avec les participants ou prêter conseils sur les techniques de recrutement. Le recrutement des participants peut être laborieux selon la taille de la population d'intérêt, les relations entre les membres et l'approche d'échantillonnage. La publicité est souvent utilisée pour faire connaître les occasions de participation, que ce soit en personne, sur papier ou en ligne. Les incitatifs sont également une stratégie utile pour accroître la participation lorsque c'est possible, mais il faut prendre en considération des facteurs d'ordre éthique connexes (voir la Partie 10). Les évaluateurs doivent également prévoir les taux de non-réponse et y réagir, surtout lorsqu'une taille minimale stricte de l'échantillon est requise. Pour atténuer ce problème, les évaluateurs pourraient augmenter le nombre de participants qu'ils recrutent en conséquence. Enfin, les évaluateurs doivent également reconnaître que leurs stratégies de recrutement pourraient fausser les résultats au moment de présenter leurs conclusions.

Une fois que les évaluateurs ont défini leur population d'intérêt, déterminé leur méthode d'échantillonnage, choisi la taille de leur échantillon et recruté leurs participants, ils peuvent enfin amorcer leur conception d'étude et commencer à recueillir les données pour éclairer les résultats de leur évaluation.

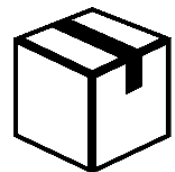
Partie 8 : Gestion, stockage et analyse des données

Il y a plusieurs facteurs importants à prendre en considération lorsqu'il s'agit de stocker, de gérer et d'analyser des données d'évaluation de programme. On recommande aux évaluateurs d'inclure des détails à ce sujet dans leur plan de travail sur les données avant d'entreprendre leur évaluation afin que les tâches puissent être accomplies en temps opportun. Les processus de saisie, d'organisation, de stockage et de nettoyage des données avant l'analyse doivent être clairement définis pour le membre de l'équipe chargé de chaque activité. Selon la capacité de l'équipe d'évaluation, il peut être plus efficace de retenir les services d'un expert-conseil externe lorsque l'expertise est insuffisante à l'interne.

Gestion des données : Les processus de saisie, d'organisation et de nettoyage des données avant l'analyse doivent être clairement définis pour chaque membre d'une équipe d'évaluation. Cela est particulièrement important lorsque plusieurs membres de l'équipe manipuleront les données, car il faut suivre un protocole strict pour s'assurer que toutes les données sont toujours traitées de la même façon. Par exemple, si le même sondage est diffusé en personne et en ligne, l'équipe d'évaluation devra disposer d'un plan pour le codage manuel et la saisie des données à partir des copies physiques afin qu'elles correspondent aux extraits de la version en ligne. Comme il a été mentionné précédemment, les structures de répartition du travail peuvent être très utiles pour l'attribution des tâches et des produits livrables aux membres de l'équipe pendant cette phase de la mission.



Stockage des données : En ce qui concerne le stockage des données, le principal facteur à considérer a trait à la protection des renseignements personnels. Le niveau de sécurité requis pour le stockage des données dépendra du caractère délicat et confidentiel des renseignements, lequel devrait être décrit dans les protocoles en matière d'éthique (voir la Partie 10). Il est essentiel d'établir une stratégie sûre et fiable pour stocker l'information avant qu'elle ait été recueillie afin que les données ne soient jamais vulnérables à la corruption, à la destruction ou à l'abus.



Analyse des données : Contrairement à de nombreuses autres composantes de l'évaluation, les compétences requises pour l'analyse des données sont très précises et parfois difficiles à trouver. Lorsque cette expertise n'est pas disponible dans l'équipe d'évaluation ou dans un réseau plus vaste, il peut être plus rentable de retenir les services d'un analyste de données externe que de former une personne à l'interne pour un seul projet. L'analyste des données peut avoir une expertise particulière en analyse quantitative ou qualitative, car ces méthodes varieront considérablement¹⁸. Il ne faut pas sous-estimer la valeur de ces compétences, car l'analyse des données influera directement sur la crédibilité des résultats. Dans un cas comme dans l'autre, le nettoyage des données est une première étape importante où les erreurs et les incohérences éventuelles dans les données sont éliminées ou corrigées pour assurer l'exactitude de l'ensemble de données.



Analyse des données quantitatives : Il s'agit du processus de transformation des données brutes en données significatives au moyen du calcul des statistiques¹⁹. Les statistiques descriptives (p. ex., la moyenne, la médiane, le maximum et le minimum) permettent aux analystes de résumer et d'examiner les données brutes, tandis que les statistiques inférentielles permettent aux analystes de dégager des inférences au sujet de la population d'intérêt. Les analystes peuvent prendre les mesures suivantes pour analyser leurs données quantitatives :



Étape 1 : Saisie de données dans un logiciel d'analyse

Le logiciel d'analyse sert à recueillir, à stocker, à organiser, à visualiser et à interpréter les données. Le logiciel à utiliser doit être sélectionné avant la collecte des données en tenant compte du type de statistiques à calculer. Par exemple, des programmes comme Microsoft Excel suffiront pour calculer des statistiques descriptives, tandis que des programmes comme SPSS ou STATA pourraient mieux convenir pour calculer des statistiques inférentielles.

Étape 2 : Choisir des statistiques descriptives appropriées

Les deux statistiques descriptives les plus importantes pour résumer les ensembles de données sont la tendance centrale (p. ex., la moyenne, la médiane ou le mode) et la variabilité (p. ex., la fourchette ou l'écart-type). Le choix des statistiques les plus appropriées pour résumer les données dépend de l'échelle de mesure de chaque variable (c.-à-d. la façon dont les variables/chiffres sont définis et catégorisés). Il existe quatre échelles de mesure, qui sont résumées dans le Tableau 3 ci-dessous. Les coefficients de corrélation peuvent également être utilisés pour décrire la nature (p. ex., positive, négative ou neutre) et l'intensité (p. ex., forte ou faible) de la relation entre les deux variables d'intérêt.

¹⁸ Pour de plus amples renseignements sur les méthodes d'analyse des données, voir Cozby & Rawn (2016).

¹⁹ Pour de plus amples renseignements sur l'analyse quantitative, consultez le site Web du Pell Institute [ici](#).

Tableau 3 : Échelles de mesure

Échelle	Caractéristiques	Exemple	Statistiques appropriées
Nominale	Catégories ou numéros utilisés comme identificateurs ou noms.	<ul style="list-style-type: none"> • Classification du sexe ou de la race • Groupe d'intervention/témoin 	<ul style="list-style-type: none"> • Mode • Pourcentages
Ordinale	Données catégorielles classées en ordre, mais sans valeur particulière pour les différences entre les classements.	<ul style="list-style-type: none"> • Classe sociale • Ordre des naissances 	<ul style="list-style-type: none"> • Mode • Médiane • Pourcentages
Intervalle	Les propriétés numériques sont littérales et la différence entre les nombres de l'échelle est égale en taille.	<ul style="list-style-type: none"> • Année • Échelles de Likert 	<ul style="list-style-type: none"> • Moyenne • Écart-type
Ratio	Les propriétés numériques sont littérales avec un zéro réel et la différence entre les nombres de l'échelle est égale.	<ul style="list-style-type: none"> • Temps nécessaire pour effectuer un examen • Âge • Fréquence des comportements 	<ul style="list-style-type: none"> • Moyenne • Écart-type

Étape 3 : Choisir des statistiques inférentielles appropriées

Contrairement aux statistiques descriptives, il existe de nombreuses statistiques inférentielles qui peuvent être calculées pour connaître la population d'intérêt. Bien qu'une liste exhaustive de toutes les analyses possibles dépasse la portée du présent manuel, voici quelques techniques courantes :

- **Régression** — extension de la corrélation utilisée pour examiner et prévoir la relation entre deux variables d'intérêt ou plus.
- **Test t** — utilisé pour déterminer s'il existe une différence statistiquement significative entre les moyennes de deux groupes (p. ex., un groupe d'intervention et un groupe témoin).
- **Analyse de la variance (ANOVA)** — semblable au Test t, elle sert à déterminer s'il existe une différence statistiquement significative entre les moyennes de trois groupes ou plus.
- **Analyse de covariance** — semblable à l'ANOVA, mais permet aussi à l'évaluateur de contrôler les variables confusionnelles lorsqu'il cerne des différences statistiquement significatives entre les moyennes de groupes.

Étude de cas : Utilisation d'une analyse quantitative pour le projet du CAPS du Mohawk College

Cette étude a évalué l'incidence des conseils proactifs au moyen d'une analyse de régression pour déterminer si et dans quelle mesure les interventions permettaient de prédire la persévérance scolaire. Les résultats de la régression ont démontré que les séances proactives de conseils en groupe étaient plus efficaces que les séances proactives de conseils individuelles et que les séances de conseils en groupe ont produit une amélioration statistiquement significative des taux de persévérance scolaire chez les hommes seulement.

L'analyse qualitative des données est le processus qui consiste à dégager une signification à partir de données narratives ou anecdotiques, souvent recueillies au moyen d'entrevues, de groupes de discussion, de sondages à réponse ouverte, d'observations ou d'analyses de documents²⁰. Ce processus est moins normalisé que l'analyse quantitative des données, mais il peut fournir des renseignements très instructifs sur la nature d'un programme et les expériences des participants. Sur le plan pratique, les données peuvent être codées manuellement ou à l'aide de programmes spécialement conçus pour l'analyse qualitative (p. ex., NVivo), ou par une combinaison des deux.



Si le même membre d'une équipe d'évaluation est chargé à la fois de la collecte et de l'analyse des données, ces processus peuvent ne pas se dérouler en vase clos. À l'exception des entrevues dirigées (c.-à-d. des consultations fondées sur des questions prédéterminées), la collecte et l'analyse de données qualitatives peuvent être un processus itératif qui permet aux analystes de comprendre l'information au moment où elle est recueillie. Bien que la nature itérative de ce processus puisse introduire un biais, elle peut être bénéfique, car elle permet aux analystes de cerner les thèmes ou les questions émergents à approfondir dans les séances de consultation de suivi. Une fois que toutes les données ont été recueillies, les analystes devraient utiliser les questions de recherche comme cadre pour procéder à la réduction des données (c.-à-d. simplifier et condenser les données originales sous forme de codes pour se concentrer sur ce qui est le plus significatif pour le projet). Les mêmes données peuvent être analysées selon des points de vue multiples en fonction des questions de recherche traitées²¹.

Une fois que les analystes ont réduit leurs données jusqu'à un niveau gérable, les données codées peuvent être analysées de diverses façons pour en déterminer la signification. Cette signification peut évoluer directement à partir des questions de recherche ou émerger naturellement au fur et à mesure de l'étude. Les données sont le plus souvent codées à la fois en fonction du contenu (p. ex., en dégageant les mots et les expressions clés) et des thèmes (p. ex., sentiments positifs et négatifs). Il est également important de noter la fréquence relative à laquelle les différents sujets sont abordés et l'intensité avec laquelle ils sont exprimés. Comme il a été mentionné précédemment, la triangulation des données peut être utilisée pour accroître la crédibilité des constatations présentées, car l'utilisation de méthodes/sources multiples peut renforcer les résultats d'évaluation. La triangulation des données s'applique même si les deux méthodes de collecte des données sont qualitatives, comme lorsque les données d'observation peuvent être corroborées par les données d'entrevue. Toutefois, si les évaluateurs constatent des écarts entre les diverses méthodes de collecte des données (p. ex., les données des observations ne concordent pas avec les données des entrevues), ils doivent utiliser leur expertise pour soupeser les éléments probants ou approfondir leurs recherches afin de comprendre les différences. Les intervenants peuvent participer de nouveau à cette étape pour fournir les détails ou nuances nécessaires à l'analyse et pour discuter des résultats inattendus ou surprenants obtenus à partir des données avant la production du rapport.

20 Pour de plus amples renseignements sur l'analyse qualitative, consultez le site Web du Pell Institute [ici](#).

21 Si un analyste de données externe qui ne connaît pas bien le programme est chargé de réduire les données, il pourrait recevoir l'aide d'une personne qui participe au programme à cette étape.

Étude de cas : Utilisation de l'analyse qualitative pour le projet du CAPS de l'Université Queen's

L'Université Queen's a cherché à comparer les versions en personne et en ligne de son programme de transition de première année, le programme « Q-success », et les répercussions différentielles chez les étudiants historiquement sous-représentés. L'analyse des réponses qualitatives du questionnaire sur le quotient de développement positif (QDP) a révélé une diminution du développement positif chez les étudiants qui ont suivi le programme, et chez les étudiants qui n'y ont pas participé, ce qui donne à penser que le QDP n'était pas un outil efficace pour documenter l'incidence de ce programme. Les réponses écrites des étudiants ont révélé que les conflits d'horaire et d'autres engagements les dissuadaient de participer au programme complet.

Partie 9 : Rapport sur les résultats

Une fois toutes les données recueillies et analysées, il est temps de réfléchir à la façon dont les résultats de l'évaluation seront communiqués à son auditoire. Le principe primordial sous-jacent à la présentation des rapports est celui de la pertinence. Les évaluateurs doivent réfléchir à la façon dont les constatations doivent être présentées pour être pertinentes pour l'éventail des intervenants desservis par l'évaluation. Cela exigera vraisemblablement un exercice d'équilibre, car les évaluations sont souvent conçues de façon assez holistique et comportent un certain nombre d'éléments interreliés et intégrés dont il faut rendre compte. La plupart des intervenants auront un intérêt direct dans un aspect particulier d'une évaluation et ils voudront que ces aspects soient mis en évidence de façon appropriée lorsque les résultats seront communiqués. Dans cette optique, il est recommandé aux évaluateurs de réfléchir à la façon dont les résultats seront communiqués aux intervenants pendant les étapes de planification de l'évaluation. Selon le type d'entente de partenariat avec divers intervenants, il peut être nécessaire de veiller à ce que toutes les parties concernées soient informées de la façon dont les résultats seront diffusés et de la personne qui en est responsable. La façon dont les résultats sont communiqués peut dépendre du protocole en matière d'éthique d'une partie concernée (voir la Partie 10).

Il est également important de discuter à l'avance de la façon dont les constatations pourraient avoir une incidence sur les actions des différents intervenants. Plus particulièrement, il faut s'entendre sur la façon dont les résultats nuls ou négatifs seront déclarés. Une préoccupation importante souvent exprimée par les intervenants concerne le risque que les constatations de l'évaluation puissent être déformées, prises hors contexte ou mal comprises. Ainsi, il est primordial que les évaluateurs examinent comment, quand, avec qui et à quelles fins les résultats sont communiqués. Certains intervenants peuvent vouloir un rapport nuancé qui raconte une histoire, tandis que d'autres peuvent préférer un rapport concis qui met principalement l'accent sur la mesure numérique des résultats. La gestion des perceptions est également importante pour que les programmes puissent maintenir leur financement, surtout lorsque des fonds publics sont en jeu, car les contribuables peuvent avoir des opinions bien arrêtées sur ces investissements.



Étude de cas : Prise en considération des points de vue des intervenants pour la production de rapports sur les résultats de l'évaluation

Si nous évaluons le programme de demande de subventions de la Figure 3 afin de comprendre les résultats à court, à moyen et à long terme du programme, il pourrait y avoir plusieurs intervenants clés ayant chacun leurs intérêts particuliers. Les étudiants à faible revenu et les entités qui défendent leurs intérêts (p. ex., parents, groupes et associations d'étudiants) se préoccupent probablement de réduire les obstacles financiers aux EPS. Bien que les données probantes tirées de l'évaluation puissent indiquer que les subventions pour frais de scolarité ne soutiennent pas en fait une augmentation de la fréquentation chez les étudiants à faible revenu, le maintien de l'option de présenter une demande sans payer de frais pourrait devenir un point litigieux. Le personnel du programme voudra peut-être comprendre l'incidence des demandes de subvention isolément des autres interventions qu'il peut offrir aux étudiants (p. ex., soutien individuel, sensibilisation et mobilisation). Pour leur part, les bailleurs de fonds/donateurs pourraient être plus préoccupés par le rendement du capital investi, en regardant les résultats à moyen et à long terme en ce qui a trait au nombre d'étudiants à faible revenu qui ont accès à la subvention et poursuivent des EPS et obtiennent leur diplôme.

Compte tenu des différents intérêts et investissements en cause, il est important que les évaluateurs déterminent si des approches différentes seront adoptées pour rendre des comptes à divers publics. Les résultats peuvent être présentés sous diverses formes (p. ex., rapports écrits, présentations, vidéos, affiches, notes de service, etc.) et devraient être adaptés à chaque public cible. Il est ici question d'une orientation à privilégier et de faire preuve de jugement pour déterminer l'information importante à partager, avec qui, à quel moment et dans quel but. De même, la profondeur et la fréquence des rapports devront être déterminées au début du projet. Les contrats conclus avec les bailleurs de fonds préciseront probablement les exigences en matière de rapports, mais il sera également important pour les évaluateurs de soupeser les besoins de tous les intervenants lorsqu'ils rendront compte des résultats. Il est recommandé que les évaluateurs accordent la priorité au maintien d'une bonne relation avec les bailleurs de fonds et les donateurs, tout en les tenant informés des besoins de tous les autres intervenants, afin d'assurer l'intégrité des rapports et de l'évaluation dans son ensemble.



Autres facteurs à prendre en considération pour une évaluation

Partie 10 : Respect des normes d'éthique

L'un des éléments essentiels des évaluations est qu'elles respectent les normes d'éthique nécessaires. Cela est particulièrement important pour la recherche et les évaluations impliquant des sujets humains. Les évaluateurs doivent veiller à ce que toutes les activités menées tout au long de l'évaluation, de la sélection et du recrutement des participants, jusqu'à la collecte et au stockage des données en passant par l'interprétation et la communication des résultats, soient réalisées en tenant compte des normes d'éthique. Selon l'organisation ou l'institution, il peut y avoir des processus ou des lignes directrices internes que les évaluateurs doivent respecter. Dans le contexte des EPS, tous les établissements postsecondaires ont un processus d'approbation par le Comité d'éthique de la recherche (CÉR) qui pourrait devoir être suivi²².

Comité d'éthique de la recherche

Un groupe de chercheurs ou d'administrateurs expérimentés qui examinent en profondeur l'objectif et l'approche méthodologique de tout projet de recherche/évaluation impliquant des sujets humains.

Les CÉR veillent à ce que la recherche respecte des normes scientifiques et d'éthique élevées qui respectent et protègent les participants conformément à l'[Énoncé de politique des trois Conseils : Éthique de la recherche avec des êtres humains](#)²³. L'énoncé des trois Conseils exige que « la recherche avec des êtres humains soit menée de manière à tenir compte de la valeur intrinsèque de tous les êtres humains ainsi que du respect et de la considération qui leur sont dus ». Tout préjudice potentiel prévu pour les participants à l'étude (qu'il soit d'ordre physique ou psychologique) peut être justifié en fonction des avantages qu'eux-mêmes, d'autres et la société en général en tireront éventuellement²⁴. Les chercheurs qui ont l'intention de faire de la recherche avec des participants humains sont priés de suivre le [didacticiel EPTC 2](#) (et peuvent être tenus de présenter leur certificat d'achèvement avec une demande).

Les composantes typiques d'une demande d'un CÉR comprennent souvent les suivantes :

- | | |
|---|---|
| ✓ Contexte/description de l'étude | ✓ Sécurité des données |
| ✓ Question de recherche | ✓ Confidentialité et protection de l'anonymat |
| ✓ Conception de l'étude/méthodologie proposée | ✓ Diffusion |
| ✓ Description des participants | ✓ Conflit d'intérêts |
| ✓ Plan de recrutement | ✓ Risques |
| ✓ Processus de consentement | ✓ Avantages possibles |

²² Les conseils scolaires ont également des protocoles d'éthique de la recherche à consulter lorsque des recherches sont menées avec des élèves de la maternelle à la 12^e année.

²³ Pour ceux qui s'intéressent à une approche plus appliquée de l'orientation de l'énoncé des trois Conseils, consulter la [Formation en éthique de la recherche du gouvernement du Canada](#).

²⁴ Pour de plus amples renseignements sur les préjudices et les avantages potentiels d'une étude, consulter le [site Web](#) du Conseil national de recherches du Canada.

En plus des établissements postsecondaires, les évaluateurs d'autres types d'organismes peuvent encore avoir besoin d'obtenir l'approbation du CÉR ou de suivre un autre processus d'approbation de l'éthique prescrit. Pour tout organisme qui n'a pas de processus d'approbation d'éthique de la recherche (p. ex., les groupes communautaires), les évaluateurs peuvent examiner un processus typique du CÉR afin de décrire les lignes directrices en matière d'éthique auxquelles le projet peut souscrire.

Parmi les considérations d'ordre éthique abordées dans un processus d'approbation d'éthique typique, il y a quelques points particulièrement pertinents qui s'appliquent aux programmes liés à l'accès et à la persévérance scolaire et qui méritent une discussion plus approfondie.

La première est d'obtenir le consentement éclairé des participants. Tous les participants à une évaluation doivent être informés de l'objectif de l'évaluation, de ce qu'implique leur participation, de tous les risques liés à la participation et de la façon dont leurs renseignements seront protégés. Il devrait également être clair pour les participants qu'ils peuvent retirer leur consentement à participer en tout temps pendant l'étude. Si l'autonomie des participants potentiels à donner leur consentement n'est pas totale, les participants peuvent avoir besoin d'une protection ou d'un consentement supplémentaires de la part d'un tiers qui est chargé de prendre des décisions pour le participant. L'exigence d'un consentement éclairé permet aux évaluateurs de communiquer aux participants les renseignements qu'ils doivent savoir au sujet de l'évaluation d'une manière claire et concise, en utilisant un langage approprié, afin que ces renseignements puissent servir de base à toute question ou discussion au sujet de leur consentement²⁵.

La prochaine considération d'ordre éthique concerne le recours à des incitatifs pour accroître la taille de l'échantillon. Selon l'énoncé des trois Conseils, les incitatifs ne sont ni recommandés ni découragés, et il incombe à l'évaluateur de justifier le recours aux incitatifs. Qu'il s'agisse d'incitatifs financiers ou d'une autre forme de récompense, les évaluateurs doivent faire attention à ce que les incitatifs soient appropriés quant à la nature et au montant. Les incitatifs ne devraient pas être importants au point où ils pourraient encourager une négligence irresponsable de tenir compte des risques associés à la participation. Ils devraient également être fournis de façon équitable aux participants, avec une rémunération égale pour une participation égale.

Une autre considération d'ordre éthique pour les évaluateurs est de savoir s'ils travailleront avec des membres de populations vulnérables. La politique des trois Conseils stipule que « la vulnérabilité résulte souvent d'une capacité décisionnelle limitée ou d'un accès limité à des biens sociaux comme des droits, des opportunités de développement, et du pouvoir ». De plus, cette politique souligne que « les personnes ou les groupes qui peuvent se retrouver en situation de vulnérabilité dans le contexte d'un projet de recherche ne devraient être ni intégrés indûment au projet ni automatiquement exclus en raison de leur situation ». De nombreux programmes d'accès et de persévérance scolaire ciblent spécifiquement les populations vulnérables afin d'accroître leur participation et leur réussite aux EPS. Lorsqu'on évalue des programmes qui desservent des membres de ces populations, il est particulièrement important que les chercheurs s'assurent que leur bien-être est respecté et qu'ils sont traités de façon équitable.

Populations vulnérables
Tout groupe incapable de protéger pleinement les intérêts de ses membres dans le contexte d'un projet de recherche (p. ex., mineurs, personnes âgées, prisonniers, minorités ethnoculturelles, personnes handicapées, personnes économiquement ou socialement défavorisées, etc.)

25 Pour obtenir des renseignements sur la création d'un formulaire de consentement pour les participants éventuels, consulter le [site Web](#) du Conseil national de recherches Canada.

Étude de cas : Approche axée sur la participation de Success Beyond Limits

Comme il a été mentionné précédemment, Success Beyond Limits (SBL) effectue une Recherche-Action Participation pour les jeunes afin d'évaluer ses programmes pour les jeunes du secteur Jane-Finch de Toronto. En adoptant une approche axée sur la participation, les membres de la population évaluée seront chargés de diriger l'évaluation. Cette approche permettra aux anciens de SBL de participer à la conception et à la mise en œuvre de l'évaluation afin qu'ils puissent comprendre le rôle du programme SBL sur les trajectoires et les résultats des EPS. Ce type de conception de recherche tient compte du rôle du chercheur ou de l'évaluateur dans l'étude des populations vulnérables.

En fin de compte, lorsqu'ils conçoivent une évaluation qui mobilise des participants humains, les évaluateurs doivent prévoir suffisamment de marge de manœuvre pour planifier une conception soucieuse de l'éthique et demander l'approbation de la conception en question. Il est également recommandé de tenir compte de cette marge de manœuvre d'emblée dans l'échéancier du projet, car le délai de réception de l'approbation du CÉR peut souvent prendre plusieurs semaines ou mois. Heureusement, les évaluateurs patients seront récompensés par la possibilité de recueillir des renseignements instructifs auprès des participants à l'évaluation.

Partie 11 : Établir un budget

Il est essentiel d'établir un budget pendant la phase de planification d'une évaluation, car chaque budget sera minutieusement adapté en fonction de son contexte particulier. Les évaluateurs peuvent d'abord consulter les ressources humaines ou d'autres administrateurs pour comprendre les règles et les règlements régissant l'établissement d'un budget au sein de leur établissement ou organisation. Il conviendra peut-être de se poser les questions suivantes :



- ✓ Quel est le financement disponible pour l'évaluation?
- ✓ Faut-il obtenir des fonds supplémentaires (p. ex., collecte de fonds ou sensibilisation)?
- ✓ Quel soutien en nature sera fourni et par qui (p. ex., personnel, espaces de réunion, stockage de données, matériel, équipement, etc.)?

Après avoir obtenu une idée générale des fonds et des soutiens en nature disponibles, les évaluateurs devraient déterminer tout élément non négociable du budget, sans lequel l'évaluation ne réussirait pas à répondre aux questions de recherche. En fin de compte, il est impératif que les évaluateurs comprennent l'incidence que le budget pourrait avoir sur la conception de l'étude, la collecte des données, les produits livrables et les échéanciers, et inversement.

Voir le Tableau 4 ci-dessous pour un modèle de budget utilisé pour les projets du COQES. Le budget type d'un projet peut comprendre les coûts (réels ou estimés) pour les éléments suivants :

- ✓ Dotation (p. ex., évaluateurs internes et externes, chercheurs, analystes, autres intervenants, etc.);
- ✓ Déplacements requis pour les activités d'évaluation (p. ex., participation à des réunions, collecte de données, etc.);
- ✓ Accès aux données ou stockage de celles-ci;
- ✓ Matériel, fournitures et équipement;
- ✓ Incitatifs pour les participants;
- ✓ Diffusion des connaissances (p. ex., présentations, conférences, ateliers, etc.).

La dotation est habituellement le facteur à considérer le plus coûteux pour une évaluation. Pour déterminer le taux de rémunération par employé, il faut 1) comprendre les qualifications et l'expertise de chaque employé, 2) estimer approximativement le temps que chaque employé consacrera aux activités d'évaluation et 3) avoir une bonne estimation de la durée du projet dans son ensemble. Au départ, les évaluateurs devraient également déterminer qui est responsable de la surveillance du budget et de l'administration des fonds, et s'il y a des coûts supplémentaires associés à cette responsabilité. Idéalement, si les paiements sont directement associés à des produits livrables précis que le personnel doit produire (p. ex., rapports provisoires ou finaux), les évaluateurs devraient également déterminer qui sera responsable de l'approbation de ces produits livrables, combien de temps il faudra, ainsi que les conséquences financières si les produits livrables connexes ne respectent pas les normes convenues. L'estimation des frais de déplacement pour les activités d'évaluation suivra un processus semblable, et les évaluateurs peuvent consulter les ressources humaines et les administrateurs pour déterminer le montant qui peut être remboursé pour les frais de déplacement (p. ex., les frais de transport en commun, les vols, les hôtels et les indemnités de repas).

Au moment d'estimer les coûts relatifs à l'accès aux données ou au stockage des données, certains scénarios seront relativement simples alors que dans d'autres scénarios, les évaluateurs devront peut-être s'en remettre à des ententes de partage de données. Ces ententes précisent les données échangées pour un projet donné, ainsi que les coûts connexes, les autorisations d'utilisation, les exigences en matière de stockage et les privilèges de partage. L'établissement des coûts du matériel, des fournitures, de l'équipement et des incitatifs devrait également être relativement simple. Toutefois, la budgétisation des coûts rattachés à la diffusion des connaissances peut être complexe, surtout si les résultats d'une évaluation ne sont pas prévisibles. Les évaluateurs auraient peut-être intérêt à formuler des hypothèses sur les événements ou les publications auxquels ils désirent participer après avoir dégagé les constatations de l'évaluation.

Tableau 4 : Modèle de budget du COQES

PERSONNEL :				
Nom des membres de l'équipe proposés <i>(ajoutez des lignes au besoin)</i>	Titre et poste	Nombre de jour par semestre	Taux par jour / par semestre	Total : Nombre x taux
Total partiel des coûts de personnel				COQES : <i>En nature</i>
COÛTS DIRECTS :				
Description détaillée des coûts directs (le cas échéant) <i>(ajoutez des lignes au besoin)</i>				Coûts
Total partiel des coûts directs				COQES : <i>En nature</i>
DÉPLACEMENTS ET HÉBERGEMENT :				
Description détaillée des déplacements et de l'hébergement (le cas échéant) <i>(ajoutez des lignes au besoin)</i>				Coûts
Total partiel des coûts de déplacement et d'hébergement				COQES : <i>En nature</i>
COÛTS INDIRECTS :				
Description détaillée des coûts indirects (le cas échéant) <i>(ajoutez des lignes au besoin)</i>			Coûts	10 % des coûts
Total partiel des coûts indirects			COQES : <i>En nature</i>	
FINANCEMENT :				
Total partiel des coûts de personnel				
Total partiel des coûts directs (le cas échéant)				
Total partiel des coûts de déplacement et d'hébergement (le cas échéant)				
10 % du total des coûts indirects (le cas échéant)				
Prix fixe ferme (financement du COQES)				
TVH (13 %) (le cas échéant)				
Total du financement du COQES (prix fixe ferme + TVH)				
Contribution en nature				
Coût total du projet (total du financement du COQES + contribution en nature)				

Conclusion

Ce manuel visait à présenter un aperçu des considérations pratiques pour l'évaluation des différents types de programmes. L'évaluation des programmes est à la fois un art et une science. Comme il en a été question, les approches, les méthodes, les outils et les techniques doivent s'aligner sur les questions de recherche et s'harmoniser avec la vision et les objectifs généraux de l'évaluation. Cet alignement exige beaucoup de réflexion et l'exercice de jugement et nous sommes convaincus que le jeu en vaut la chandelle.

Au début du présent manuel, nous avons souligné l'importance de l'évaluation des programmes pour le COQES, car une évaluation uniforme et de grande qualité des programmes est la seule façon de répondre avec confiance aux questions « Qu'est-ce qui fonctionne? » et « Pourquoi cela fonctionne-t-il? » dans les programmes d'accès et de persévérance scolaire financés par le secteur public. Nous espérons que ce manuel servira de guide utile pour l'élaboration de plans d'évaluation pour les programmes qui visent à améliorer l'accès ou la persévérance scolaire dans les établissements d'enseignement supérieur. Nous sommes impatients de travailler avec les promoteurs du Consortium sur l'accès et la persévérance scolaire (CAPS) du COQES qui utiliseront le contenu du présent manuel afin d'élaborer, d'améliorer ou de peaufiner leurs évaluations de programme. Nous tenons à remercier à l'avance les promoteurs du CAPS de l'intérêt qu'ils portent à notre mission commune d'améliorer l'accès, la persévérance et la réussite scolaires des étudiants du système d'éducation postsecondaire de l'Ontario.

Références

- Baker, A. (s.d.). *Evaluation Services : Making Data Collection Decisions*.
<http://nebula.wsimg.com/79cc3a812b2de18f3dad8393b80b0ea0?AccessKeyId=8A32C507409C3E494DCC&disposition=0&alloworigin=1> [en anglais seulement]
- Better Evaluation (s.d.). *Evaluation Options : Collect and/or Retrieve Data*.
https://www.betterevaluation.org/en/rainbow_framework/define/collect_retrieve_data [en anglais seulement]
- Better Evaluation (s.d.). *Evaluation Options : Develop Programme Theory / Theory of Change*.
https://www.betterevaluation.org/en/rainbow_framework/define/develop_programme_theory [en anglais seulement]
- Better Evaluation (s.d.). *Evaluation Options : Sample*.
https://www.betterevaluation.org/en/rainbow_framework/define/sample [en anglais seulement]
- Better Evaluation (s.d.). *Evaluation Options : SWOT Analysis*.
<https://www.betterevaluation.org/evaluation-options/swotanalysis> [en anglais seulement]
- Société canadienne d'évaluation (s.d.). *Qu'est-ce que l'évaluation?* <https://evaluationcanada.ca/fr/quest-ce-que-évaluation>
- Chen, H. T. (2014). *Practical Program Evaluation: Theory-driven Evaluation and the Integrated Evaluation Perspective*. United States : Sage. [en anglais seulement]
- Cozby, P. C., & Rawn, C. D. (2016). *Methods in Behavioural Research*. Canada : McGraw Hill Education. [en anglais seulement]
- Gouvernement du Canada (2014). *Énoncé de politique des trois Conseils : Éthique de la recherche avec des êtres humains*. http://www.ger.ethique.gc.ca/pdf/fra/eptc2-2014/EPTC_2_FINALE_Web.pdf
- Gouvernement du Canada (s.d.). *Pour soutenir des évaluations efficaces : Guide d'élaboration de stratégies de mesure du rendement*. <https://www.canada.ca/fr/secretariat-conseil-tresor/services/verifications-évaluations/centre-excellence-en-évaluation/guide-elaboration-strategies-mesure-rendement.html>
- Gouvernement du Canada. Conseil national de recherches Canada (2012). *Guide du formulaire de demande d'approbation déontologique*.
https://www.nrc-cnrc.gc.ca/fra/apropos/ethiques_integrite/demande_approbation_deontologique.html
- Gouvernement du Canada. Conseil national de recherches Canada (2012). *Guide du modèle de formulaire de consentement*.
https://www.nrc-cnrc.gc.ca/fra/apropos/ethiques_integrite/modele_formulaire_consentement.html
- Innovation Network (s.d.). *Evaluation Plan Workbook*.
https://www.innonet.org/media/evaluation_plan_workbook.pdf [en anglais seulement]
- National Science Foundation (2002). *User-Friendly Handbook for Project Evaluation: An Overview of Quantitative and Qualitative Data Collection and Methods*. https://www.nsf.gov/pubs/2002/nsf02057/nsf02057_4.pdf [en anglais seulement]
- Project Management Docs. (s.d.). *Work Breakdown Structure*.
<https://www.projectmanagementdocs.com/template/project-planning/work-breakdown-structure/#axzz5Mx5Sj1Ht> [en anglais seulement]
- Project Manager (s.d.). *The Ultimate Guide to: Gantt Charts*. <https://www.projectmanager.com/gantt-chart> [en anglais seulement]

- Taylor-Powell, E. (1988). *Program Development and Evaluation : Sampling*. University of Wisconsin Extension. <https://learningstore.uwex.edu/Assets/pdfs/G3658-03.pdf> [en anglais seulement]
- The Pell Institute (s.d.). *Evaluation Toolkit : Analyzing Qualitative Data*. <http://toolkit.pellinstitute.org/evaluation-guide/analyze/analyze-qualitative-data/> [en anglais seulement]
- The Pell Institute (s.d.). *Evaluation Toolkit : Analyzing Quantitative Data*. <http://toolkit.pellinstitute.org/evaluation-guide/analyze/analyze-quantitative-data/> [en anglais seulement]
- The Pell Institute (s.d.). *Evaluation Toolkit : Choose an Evaluation Design*. <http://toolkit.pellinstitute.org/evaluation-guide/plan-budget/choose-an-evaluation-design/> [en anglais seulement]
- The Pell Institute (s.d.). *Evaluation Toolkit: Understanding Sampling*. <http://toolkit.pellinstitute.org/evaluation-guide/collect-data/understand-sampling/> [en anglais seulement]
- The University of Kansas' Community Toolbox (s.d.). *Operations in Evaluating Community Interventions: Choosing Questions and Planning the Evaluation*. Evaluating Community Programs and Initiatives. <https://ctb.ku.edu/en/table-of-contents/evaluate/evaluate-community-interventions/choose-evaluation-questions/main> [en anglais seulement]
- Wingate, L., & Schroeter, D. (2007). *Evaluation Questions Checklist for Program Evaluation*. <https://wmich.edu/sites/default/files/attachments/u350/2018/eval-questions-wingate%26schroeter.pdf> [en anglais seulement]

Ressources d'évaluation supplémentaires

- Balbach, E.D. (1999). Using Case Studies to Do Program Evaluation. Tufts University. California Department of Health Services. <http://case.edu/affil/healthpromotion/ProgramEvaluation.pdf> [en anglais seulement]
- Better Evaluation (s.d.). *Evaluation Options : Rainbow Framework*. https://www.betterevaluation.org/en/rainbow_framework [en anglais seulement]
- Centre d'excellence de l'Ontario en santé mentale des enfants et des adolescents (s.d.). *Carrefour des ressources*. <http://www.excellencepourenfantsados.ca/carrefour-des-ressources>
- Stufflebeam, D.L. (2004). *Evaluation Design Checklist*. Western Michigan University Evaluation Center. <https://www.wmich.edu/sites/default/files/attachments/u350/2014/evaldesign.pdf> [en anglais seulement]
- The Pell Institute (s.d.). Evaluation Toolkit. <http://toolkit.pellinstitute.org/> [en anglais seulement]
- The University of Kansas' Community Toolbox (s.d.). *Evaluating Community Programs and Initiatives*. <https://ctb.ku.edu/en/evaluating-community-programs-and-initiatives> [en anglais seulement]
- U.S. Department of Health and Human Services Centers for Disease Control and Prevention: Office of the Director, Office of Strategy and Innovation. (2011). *Introduction to Program Evaluation for Public Health Programs: A Self-study Guide*. <https://www.cdc.gov/eval/guide/cdcevalmanual.pdf> [en anglais seulement]
- YouthREX (s.d.). Evaluation Toolkit for Youth Programs. <http://www.youthrex.com/toolkit/> [en anglais seulement]

Groupes d'évaluation

- Evaluation Capacity Network : <http://www.evaluationcapacitynetwork.com/> [en anglais seulement]
- Evaluation Roundtable : <http://www.evaluationroundtable.org/> [en anglais seulement]



Higher Education
Quality Council
of Ontario

An agency of the Government of Ontario