

Reflets

Revue d'intervention sociale et communautaire



Étude de cas d'une évaluation participative : compromis ou atteinte à la validité scientifique ?

Véronique Chadillon-Farinacci, PhD, Frédéric Saint-Pierre et Marie-Andrée Pelland

Volume 28, numéro 2, automne 2022

Au-delà de la reddition de compte : l'évaluation dans les organismes communautaires et d'économie sociale

URI : <https://id.erudit.org/iderudit/1106956ar>

DOI : <https://doi.org/10.7202/1106956ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

Reflets, Revue d'intervention sociale et communautaire

ISSN

1203-4576 (imprimé)

1712-8498 (numérique)

[Découvrir la revue](#)

Citer cet article

Chadillon-Farinacci, V., Saint-Pierre, F. & Pelland, M.-A. (2022). Étude de cas d'une évaluation participative : compromis ou atteinte à la validité scientifique ? *Reflets*, 28(2), 44-68. <https://doi.org/10.7202/1106956ar>

Résumé de l'article

L'évaluation dépasse la reddition de compte dans le milieu associatif. Plusieurs auteurs rapportent qu'impliquer les parties prenantes dans le processus d'évaluation est une pratique novatrice. Cette étude s'interroge sur la présence de biais générés par cette pratique. Elle pose la question suivante : dans quelle mesure cette implication peut-elle rivaliser avec l'amélioration du programme, voire la production de savoirs empiriques ? Pour y répondre, le cas de l'évaluation d'un programme de prévention de la conduite avec facultés affaiblies chez les jeunes de 10^e année dans quatre provinces du Canada atlantique est présenté. À l'aide des données produites par cette étude de cas et de données secondaires, on compare les caractéristiques de trois groupes de jeunes illustrant différentes modalités d'implication des parties prenantes. Les résultats suscitent une discussion relative à la qualité et à la scientificité des données découlant de l'évaluation.

Étude de cas d'une évaluation participative : compromis ou atteinte à la validité scientifique ?

Véronique Chadillon-Farinacci, PhD

*Professeure, Département de sociologie et de criminologie
Université de Moncton*

Frédéric Saint-Pierre

*Étudiant au baccalauréat en criminologie
Université de Moncton*

Marie-Andrée Pelland

*Professeure, Département de sociologie et de criminologie
Université de Moncton*

Résumé

L'évaluation dépasse la reddition de compte dans le milieu associatif. Plusieurs auteurs rapportent qu'impliquer les parties prenantes dans le processus d'évaluation est une pratique novatrice. Cette étude s'interroge sur la présence de biais générés par cette pratique. Elle pose la question suivante: dans quelle mesure cette implication peut-elle rivaliser avec l'amélioration du programme, voire la production de savoirs empiriques? Pour y répondre, le cas de l'évaluation d'un programme de prévention de la conduite avec facultés affaiblies chez les jeunes de 10^e année dans quatre provinces du Canada atlantique est présenté. À l'aide des données produites par cette étude de cas et de données secondaires, on compare les caractéristiques de trois groupes de jeunes illustrant différentes modalités d'implication des parties prenantes. Les résultats suscitent une discussion relative à la qualité et à la scientificité des données découlant de l'évaluation.

Mots-clés : Biais de sélection, évaluation de programme, conduite avec facultés affaiblies par le cannabis, prévention chez les jeunes, risque perçu pour la santé, consommation autorévélee, implication des parties prenantes, validité interne.

Introduction

Dans le milieu associatif, l'évaluation dépasse la reddition de compte; elle s'inscrit dans un processus d'appréciation de l'atteinte des objectifs du programme évalué et de sa valeur ajoutée (Midy et coll., 1998; Senanayake et coll., 2021; Serafini et coll., 2015). L'évaluation d'un programme mesure aussi les moyens déployés pour atteindre ses objectifs (Lavoie et Panet-Raymond, 2020; White, 2017). Pour préserver la fiabilité des résultats et s'assurer de l'efficacité des programmes, de nombreux organismes subventionnaires exigent que les organismes communautaires s'associent à des experts pour conduire les évaluations (Torres-Cuello et coll., 2018; Tougas et coll., 2021; White, 2017).

L'idée d'impliquer les parties prenantes dans le processus d'évaluation est une pratique décrite par certains auteurs comme incontournable (Cousins et coll., 1996; Fukunaga, 2015; Torres et Preskill, 1999; Torres-Cuello et coll., 2018; Weaver et Cousins, 2004). L'évaluation collaborative permet entre autres de faciliter l'accès aux personnes d'intérêt, d'assurer la qualité de l'information recueillie et de soutenir la résolution d'écueils. De plus, cette collaboration peut favoriser la réceptivité des recommandations formulées tout au long du processus d'évaluation par les partenaires de terrain (Fetterman et coll., 2017; Rodríguez-Campos et Rincones-Gómez, 2012).

Dans ce contexte, les « *parties prenantes* » englobent toutes les personnes qui bénéficient du programme évalué ou qui y contribuent. Elles incluent les bailleurs de fonds, les clients, les employés offrant le programme, les administrateurs, les représentants communautaires, etc. Autrement dit, toute personne ou organisation pouvant représenter, par exemple, une perspective nouvelle sur la mise en œuvre du programme peut être considérée comme une partie prenante (Mertens et Wilson, 2018; Yarbrough et coll., 2010).

La présente étude pose un regard critique sur l'implication des parties prenantes dans le processus d'évaluation d'un programme, plus spécifiquement lors de la collecte des données. Cette étape amorce concrètement la mise en œuvre de l'évaluation. De plus, elle pose les bases des étapes subséquentes : une collecte post-test, l'analyse des résultats, la diffusion de ceux-ci, etc.

Des données biaisées produiront ainsi des résultats biaisés. Or, sensibiliser les évaluateurs à ces questions n'est pas suffisant pour documenter cet enjeu. La présente étude propose une façon de mesurer les biais introduits à travers l'implication des parties prenantes et leurs conséquences sur la qualité des données colligées pour évaluer l'efficacité du programme évalué. Dans quelle mesure ces biais potentiels s'opposent-ils à l'amélioration du programme, voire à la production de savoirs empiriques ?

Pourquoi une telle question quand nombre d'auteurs considèrent cette pratique comme une partie essentielle à la méthodologie de l'évaluation de programmes (Fetterman et coll., 2017; Rodríguez-Campos et Rincones-Gómez, 2012)? En 2020-2021, l'équipe de recherche a participé à l'évaluation d'un programme de prévention de la conduite avec facultés affaiblies chez des jeunes de 10^e année des quatre provinces du Canada atlantique. Au cours de l'analyse des résultats, on a constaté que les participant.e.s au programme, contrairement aux adolescents de deux groupes de référence, affirmaient en plus grandes proportions avoir consommé du cannabis et en proportions inférieures avoir consommé de l'alcool au cours de leur vie. Ce résultat était en soi inattendu : l'alcool tend à être consommé plus tôt par les adolescents que les autres substances (Gallegos et coll., 2021; Spilka et coll., 2021). Cette observation préliminaire a mené à la présente étude. En se questionnant empiriquement sur la présence de biais, il est possible de discuter de leur portée aux étapes subséquentes de l'évaluation.

Pour aborder ce sujet, une recension des études abordant l'implication des parties prenantes dans la qualité des devis évaluatifs a été effectuée. Par la suite sera présenté le cas de l'évaluation du programme « Défi jeunesse de l'Atlantique » conduite à la fin de l'année scolaire 2020-2021. À l'aide des données produites par cette étude de cas ainsi que des données secondaires, les caractéristiques de trois groupes de jeunes illustrant différentes modalités d'implication des parties prenantes à l'étape de la collecte des données sont comparées : les jeunes du premier groupe ont été recrutés par l'organisme dont le programme était évalué, le deuxième, par des directions d'écoles secondaires sollicitées par l'équipe de recherche et le troisième groupe était composé d'un sous-échantillon de l'*Enquête canadienne sur le tabac, l'alcool et les drogues chez les élèves* (2016-2017) (ECTADÉ).

La comparaison des réponses issues du même questionnaire permet de qualifier ces groupes. Deux thématiques inhérentes au programme sont présentées : les perceptions de risques pour la santé associées à l'alcool et au cannabis et les comportements associés à la conduite avec facultés affaiblies ainsi qu'à la consommation de cannabis. Les résultats réaffirment l'importance de discuter du niveau de scientificité ou de fiabilité des données découlant des évaluations.

Regards sur la diversité de l'implication des parties prenantes

Les écrits scientifiques sur l'évaluation participative attestent la diversité des formes d'implications des parties prenantes; celle-ci module les rôles des évaluateurs et des

utilisateurs (Patton et LaBossière, 2012). Fetterman et ses collègues (2017) identifient trois principales formes : une collaboration fondée sur une rétroaction des parties prenantes, un partenariat où les parties prenantes et les chercheurs sont égaux et une coopération basée sur l'autonomisation des parties prenantes. Suivant cette troisième forme, les parties prenantes dirigent l'évaluation, et les chercheurs prodiguent des conseils.

Ces principales formes de collaboration se déclinent selon plusieurs paramètres modulant les rôles de tous dans le processus, y compris la sélection des parties à impliquer, le partage du pouvoir entre celles-ci et leur niveau d'implication (Cousins et Whitmore, 1998; Fukunaga, 2015; Taut, 2008; Torres et Preskill, 1999). Ensemble, ces paramètres définissent différentes modalités d'implication à situer sur un continuum. Prenons la prise de décision comme exemple. D'un côté, les décisions techniques sont imposées par les évaluateurs. De l'autre, elles sont prises exclusivement par les parties prenantes (Cousins et Whitmore, 1998; Fetterman et coll., 2014).

La revue systématique de Brandon et Fukunaga (2014) propose un éclairage analytique sur l'implication des parties prenantes dans le processus d'évaluation. Les auteurs ont identifié 41 études sur le sujet à l'aide des critères de sélection suivants : aborder l'implication des parties prenantes dans un contexte évaluatif, produire des résultats empiriques découlant de cette implication, décrire exhaustivement ses modalités, avoir été révisées par des pairs et avoir été publiées entre 1985 et 2010. Deux observations peuvent être extraites de cette revue. Premièrement, les parties prenantes les plus couramment impliquées dans les évaluations sont responsables de la mise en œuvre du programme. Celles-ci sont mentionnées dans près de la moitié des écrits recensés par les auteurs (n=19; 46 %). Deuxièmement, l'implication des parties prenantes est surtout abordée par la bande; une minorité des études porte directement sur l'implication des parties prenantes dans le cadre d'évaluations (n=11; 27 %). Cette deuxième observation indique qu'il s'agit d'un sujet peu étudié.

Bien qu'il soit, à notre connaissance, peu documenté empiriquement, l'enjeu de l'implication des parties prenantes et de son influence sur la qualité de l'évaluation est évoqué par de nombreux chercheurs (Braverman, 2013; Fukunaga, 2015; Scriven, 2003; Taut, 2008). Malgré le soutien dont bénéficie l'approche participative, plusieurs déplorent le manque d'évaluations de la qualité des devis (Brandon et Fukunaga, 2014; Cousins, 2003; Torres et Preskill, 1999) : « Dans quelle mesure est-il éthique d'éroder en partie la qualité technique d'une évaluation afin de favoriser le développement des compétences et l'appropriation psychologique du projet par les parties prenantes? (traduction libre) (Morris, 2003, p. 315). »

Il n'y a pas de réponse simple. On propose d'explorer cette question à travers deux remarques. La première aborde les biais générés par l'implication des parties prenantes, et la deuxième, la responsabilité de l'évaluateur vis-à-vis ces biais. Ces deux remarques conduisent l'évaluateur à une impasse où les intérêts scientifiques de la démarche prennent une place secondaire. D'un côté, l'implication des parties prenantes s'inscrit dans une optique de compromis : l'affaiblissement de la qualité du devis n'invalide pas les savoirs produits. De l'autre, elle risque d'affaiblir la qualité du devis de façon à constituer une atteinte à sa validité.

Deux remarques sur l'évaluation participative

Plusieurs travaux sur l'évaluation participative mettent l'accent sur la validité accrue des résultats produits (Brandon, 1998; Crishna, 2006). Un exemple de stratégie mise de l'avant par ces travaux est le fait de tirer parti de l'expertise des parties prenantes de façon à effectuer une collecte d'informations complète. De cette manière, l'évaluateur diminuerait, en principe, les risques d'atteintes à la validité des résultats (Brandon, 1998).

Dans un premier temps, sans nier ses avantages avérés, l'implication des parties prenantes peut aussi nuire à la qualité de l'évaluation en introduisant des biais pernicieux. La revue systématique citée précédemment présente deux résultats appuyant cette première remarque. D'abord, les auteurs observent qu'une minorité d'études aborde la qualité technique des devis d'évaluation. Lorsque cette question est posée, les résultats ne sont pas en faveur d'une amélioration de celle-ci. Des 3 études sur 41 ayant traité cette question, deux d'entre elles présentent des résultats négatifs ou mitigés, c'est-à-dire qu'une amélioration de la qualité n'est pas clairement observée (Brandon et Fukunaga, 2014).

Ensuite, peu d'études recensées par les auteurs abordent les risques d'introduction de biais par l'évaluation participative. Lorsque ces biais sont évoqués, le portrait est peu reluisant. Des 7 études sur 41 qui en font mention, 6 d'entre elles présentent des résultats négatifs ou mitigés (Brandon et Fukunaga, 2014). Ces études rapportent, entre autres, des préoccupations relatives au contrôle des administrateurs, en particulier lorsqu'ils détiennent un pouvoir décisionnel important sur le déroulement de l'évaluation.

Ces deux indices portent à penser que l'évaluation participative n'est pas la panacée : le risque d'affecter négativement la validité des résultats est réel. Si ce risque ne fait pas l'objet d'une attention empirique, il sera plus difficile pour les évaluateurs d'intervenir.

Dans un deuxième temps, la recherche évaluative s'appuie sur l'a priori que la responsabilité de l'assurance-qualité du devis appartient exclusivement aux évaluateurs (King et Ehlert, 2008 ; Morris, 2003). Bien que les connaissances et les compétences pour réaliser l'évaluation soient préalables (Brisolara, 1998 ; Brisson, 2007 ; Crishna, 2006 ; King et Ehlert, 2008 ; Morris, 2003), l'attribution de cette responsabilité implique que l'évaluateur a le contrôle sur tout le processus d'évaluation.

Dans les faits, l'évaluateur valse entre la faisabilité d'un projet et sa rigueur méthodologique (Braverman, 2013 ; Eck, 2006). Ce compromis s'inscrit dans une négociation avec les parties prenantes lorsque l'évaluation est participative. Il arrive que les parties prenantes préconisent une méthodologie moins rigoureuse (Braverman, 2013). La méthode scientifique peut alors être érodée pour satisfaire le contexte. Par conséquent, plus les parties prenantes sont impliquées dans l'évaluation, plus la responsabilité éthique des évaluateurs augmente vis-à-vis la fiabilité et la diffusion des résultats (Morris, 2003). Ainsi, bien que l'implication des parties prenantes soit conceptuellement prisée, elle reste peu étudiée empiriquement. Ce compromis est central à la présente étude.

La méthode scientifique en second plan : compromis ou risque pour l'intégrité scientifique ?

L'évaluation d'un programme est utile à plusieurs égards. Son utilité est instrumentale lorsqu'elle est utilisée comme outil d'amélioration (Lemay et coll., 2013). Advenant que l'intervention des parties prenantes ébranle la qualité de la démarche, les évaluations participatives de programmes peuvent entraîner des problèmes de validité scientifique. Ces problèmes compromettent l'utilité instrumentale de ces dernières. En ce sens, Farrington (2006) propose une échelle de qualité méthodologique des évaluations. Celle-ci s'appuie sur cinq types de validité, dont la validité interne tirée des travaux de Campbell et Stanley (1963). Par « validité interne », les auteurs indiquent que l'évaluateur a contrôlé les effets d'autres variables pour affirmer que les résultats observés sont bien associés à la participation au programme.

Bien que ces considérations méthodologiques ne soient pas exclusives aux évaluations participatives, les enjeux de collecte des données peuvent être exacerbés par la présence de parties prenantes pouvant influencer les décisions. L'implication des parties prenantes étant une caractéristique distinctive des évaluations participatives, l'évaluateur a moins de contrôle sur le processus que pour d'autres types d'évaluations.

Dans la présente étude, l'intervention des parties prenantes dans le processus d'évaluation soulève des questions de validité interne. Ce type de validité concerne la relation entre les variables dépendantes et la variable indépendante (McDavid et coll., 2018). La variable indépendante est le programme et les variables dépendantes sont les effets attendus (Bouchard et Cyr, 2005; Campbell et Stanley, 1963; Mertens et Wilson, 2018; Rossi et Freeman, 1993). Autrement dit, introduire des facteurs susceptibles d'influencer la relation entre le programme et ses effets attendus diminue la capacité des évaluateurs à étudier cette relation.

L'écart entre le programme prévu et le programme mis en œuvre peut aussi constituer un risque à la validité interne de l'évaluation. En effet, ces « aléas de l'implantation » sont parfois hors du contrôle des évaluateurs et des parties prenantes : « les intervenants n'obtiennent pas toutes les ressources nécessaires; ils se heurtent à des résistances imprévues; ils ne réussissent pas à attirer la clientèle anticipée ou, encore, ils ne peuvent pas maintenir la participation au niveau souhaité » (Cusson et coll., 1994, p. 68). La pandémie de COVID-19 est un exemple d'aléa imprévu ayant perturbé la mise en œuvre du programme et de son évaluation.

Les menaces à la validité interne sont mieux documentées que ses conséquences (Bouchard et Cyr, 2005; McDavid et coll., 2018; Mertens et Wilson, 2018; Stufflebeam et Coryn, 2014; Taut, 2008; Weiss, 1998) travail social, éducation, etc.. La présente étude documente ces questions en comparant la composition de différents échantillons dont l'apport des parties prenantes dans le recrutement des participant.e.s varie. L'hypothèse est que l'implication de l'organisme dans le recrutement des participant.e.s exacerbe le risque de biais de la validité du devis.

La présente étude

L'étude documente les biais introduits par l'implication variable de parties prenantes dans le recrutement d'élèves ayant répondu au même sondage. Elle s'appuie en partie sur les données recueillies à travers l'évaluation du « Défi jeunesse de l'Atlantique ». Ce programme vise à prévenir la conduite avec des facultés affaiblies par le cannabis chez des élèves du secondaire. L'évaluation s'est déroulée à la fin de l'année scolaire 2020-2021 en pleine pandémie de COVID-19.

Le « Défi jeunesse de l'Atlantique » vise à favoriser le développement d'aptitudes relatives à la prise de décision chez les jeunes. La décision encouragée par le programme

consiste à ne pas conduire ou à ne pas être passager d'un véhicule conduit par une personne dont les facultés sont affaiblies. Le programme consiste à impliquer de jeunes leaders de 10^e année dans la création d'une campagne de sensibilisation et à tenir une série d'activités en présentiel. Suivant l'idée de la prévention menée par les pairs (*peer-led*), le programme prévoit que les leaders exercent une influence positive sur leurs camarades.

Le programme est mis en œuvre par un organisme de lutte contre les inégalités affectant la qualité de vie des personnes. Par ailleurs, le « Défi jeunesse de l'Atlantique » ne cible pas une clientèle vulnérable. Ce programme visait à l'origine les jeunes de 10^e année des écoles secondaires des quatre provinces de l'Atlantique.

L'évaluation du programme a été conçue conjointement avec l'organisme. Le devis d'origine a incorporé le point de vue de l'organisme à travers plusieurs rondes de consultation et d'échanges. L'évaluation tentait de répondre à la question suivante : dans quelle mesure le « Défi jeunesse de l'Atlantique » est-il responsable d'une diminution des comportements de conduites avec facultés affaiblies par le cannabis chez les jeunes ?

Un devis quasi expérimental était prévu, puisqu'une répartition aléatoire des participant.e.s n'était pas possible. Or, le devis prévu n'a pas pu être réalisé intégralement. Seules les mesures du prétest ont été analysées, car la participation des jeunes au posttest était trop faible pour capter le changement découlant de l'intervention. Cet imprévu s'est ajouté à celui de la crise sanitaire.

En effet, le déroulement de l'évaluation s'est caractérisé par une forte implication de l'organisme offrant le programme en raison de la pandémie de COVID-19. D'abord, l'organisme a dû revoir sa programmation; aucune activité en présentiel n'a pu être réalisée, car l'accès aux écoles secondaires a été suspendu. Les jeunes ont donc été invité.e.s à participer au programme par les intervenants de l'organisme sur une base volontaire par un recrutement de bouche à oreille et par la méthode de la boule de neige. L'organisme n'est intervenu ni dans l'analyse des données ni dans la rédaction du rapport final.

Pour respecter les consignes de santé publique, l'équipe de recherche n'a pas participé à la distribution ni à la passation des questionnaires en ligne auprès des jeunes. L'organisme s'est chargé de la collecte des données auprès du groupe d'intervention. Sur les 320 jeunes identifié.e.s par les intervenants de l'organisme, près du quart d'entre eux ont répondu au questionnaire conçu pour évaluer le programme (83/320; 25,93 %). Les participant.e.s au questionnaire ont été sollicité.e.s de la même façon que ceux qui ont participé au programme, c'est-à-dire par les intervenants contraints par le contexte sanitaire.

La présente étude documente l'introduction de deux biais à la validité interne. D'une part, le choix des intervenants de se limiter aux « jeunes leaders » introduit un biais de sélection hors du contrôle des évaluateurs. D'autre part, l'implication des intervenants dans la distribution et dans la passation des questionnaires peut encourager un effet de désirabilité sociale au moment de répondre à certaines questions (Panizza et coll., 2021).

Les jeunes de deux groupes de référence n'ont pas participé au programme. Le premier groupe de référence comprend des élèves de 10^e année ayant été recrutés sur une base volontaire par l'équipe de recherche. Près de 200 écoles secondaires des provinces de l'Atlantique ont été contactées suivant un échantillonnage par strates : le nombre d'écoles a été déterminé en tenant compte du poids de la population totale des provinces, et les écoles sollicitées par province ont été déterminées aléatoirement. Le tableau 1 synthétise le processus de sélection des écoles et la participation de celles-ci. Il indique que sur les 183 écoles secondaires contactées, 19 d'entre elles se sont montrées réceptives à l'étude et 4 d'entre elles figurent dans les données¹.

Tableau 1 — Tableau récapitulatif du processus de sélection du premier groupe de référence

Province	Nombre total d'écoles secondaires	Population (2019)	Nombre d'écoles contactées	Nombre de réponses positives	Nombre d'écoles participantes
Nouvelle-Écosse	93	971 395	75	2	0
Nouveau-Brunswick	71	776 827	50	3	1
Terre-Neuve-et-Labrador	132	521 542	43	7	2
Île-du-Prince-Édouard	15	156 947	15	7	1
Total	311	2 426 711	183	19	4

Puisque le recrutement des jeunes de ces deux groupes a suivi deux protocoles distincts, les différences observées entre ceux-ci ne peuvent être occasionnées que par le programme. C'est pourquoi un deuxième groupe de référence a été envisagé.

Données et méthodes

Données

L'étude compare les réponses de trois groupes de jeunes ayant répondu aux mêmes questions. Le premier groupe représente le groupe expérimental de l'évaluation de programme (n=83). Ces jeunes ont été recrutés par les intervenants du « Défi jeunesse de l'Atlantique »; l'appellation « groupe d'intervention » y fait référence. Le taux de participation chiffré par l'organisme ayant offert le programme indique que 25,93 % des participant.e.s au programme ont rempli le questionnaire d'évaluation (83/320). Ces élèves sont réparti.e.s dans neuf écoles secondaires du Nouveau-Brunswick et de Terre-Neuve-et-Labrador. Ils sont âgé.e.s de 14 à 17 ans (moyenne=16,16 ans). Près des deux tiers d'entre eux sont des filles (52/83; 63,41 %).

Le deuxième groupe de jeunes n'a pas participé au programme; il comprend des volontaires de 10^e année sollicité.e.s par leurs directions d'écoles à travers les démarches de l'équipe de recherche (n=48). Ce groupe est appelé le « premier » groupe de référence. Ces élèves se répartissent dans quatre écoles secondaires du Canada atlantique. Les jeunes du premier groupe de référence sont tous âgés de 15 ou de 16 ans. Comme pour le groupe d'intervention, ce groupe est composé à 64,10 % de filles (25/48). Le groupe d'intervention et le premier groupe de référence ont répondu au questionnaire en ligne à l'été 2021, soit près de deux ans après la légalisation du cannabis au Canada.

Le troisième groupe provient de l'*Enquête canadienne sur le tabac, l'alcool et les drogues chez les élèves* (2016-2017) (ECTADÉ). Ces données secondaires comprennent les jeunes de la Nouvelle-Écosse, de l'Île-du-Prince-Édouard et de Terre-Neuve-et-Labrador ayant répondu au questionnaire durant l'année scolaire 2016-2017². Les participant.e.s se distribuent dans des classes de 7^e à 12^e année scolaire variant de 11,80 % à 20,40 % selon l'année. La moitié des participant.e.s de l'ECTADÉ sont des participantes (7746/15400; 50,30 %). Ce « deuxième » groupe de référence constitue une base comparative supplémentaire; la méthode d'échantillonnage préconisé pour cette enquête est la stratification sur deux dimensions (Burkhalter et coll., 2017). Deux raisons expliquent pourquoi l'ECTADÉ est d'intérêt : la sélection des participant.e.s de l'ECTADÉ est indépendante, d'une part du programme « Défi jeunesse de l'Atlantique », d'autre part, de la crise sanitaire.

Le tableau 2 ci-dessous présente les trois groupes de jeunes faisant l'objet de la présente étude. Il comprend le nom des groupes de jeunes, les méthodes d'échantillonnage, les

sources de données, la taille des échantillons et les périodes durant lesquelles ils ont passé le questionnaire.

Tableau 2 — Nom du groupe selon la méthode d'échantillonnage, la source de données et le temps de collecte de données

Nom du groupe	Méthode d'échantillonnage	Source de données (n)	Temps de collecte (crise sanitaire : oui/non)
Groupe d'intervention	Volontaires recruté.e.s par les intervenants de l'organisme offrant le programme	Jeunes qui s'apprentent à suivre le programme « Défi jeunesse de l'Atlantique » (n=82)	Été 2021 (oui)
1^{er} groupe de référence	Volontaires recruté.e.s par les directions d'écoles à travers l'équipe de recherche	Élèves de 10 ^e année ne suivant pas le programme (n=48)	Été 2021 (oui)
2^e groupe de référence	Stratification à deux dimensions	Élèves de la NÉ, de l'IPÉ et de TNL ayant participé à l'ECTADÉ (n=15400)	Année scolaire 2016-2017 (non)

Note: NÉ : Nouvelle-Écosse; IPÉ : Île-du-Prince-Édouard; TNL : Terre-Neuve-Labrador

Dans les résultats, puisque les jeunes avaient le choix de répondre ou non à chacune des questions, seuls les pourcentages valides sont rapportés. C'est pourquoi le nombre d'observations varie selon la question.

Outil et variables

L'outil de collecte est inspiré du questionnaire de l'ECTADÉ. Des 43 questions du sondage, plus de la moitié d'entre elles étaient tirées de cette enquête (24/43; 55,81 %). Une consigne de départ similaire figurait en préface du questionnaire (voir Annexe A). Cette dernière présente aux élèves l'objectif de l'étude, l'équipe de recherche et un mot sur la confidentialité des réponses. Les jeunes ont répondu au questionnaire en ligne.

Trois thématiques communes au questionnaire de l'ECTADÉ et au nôtre sont pertinentes à la présente étude : les caractéristiques individuelles des répondant.e.s, les perceptions de risque par rapport à la santé ainsi que les comportements associés à la conduite avec facultés affaiblies et à la consommation de cannabis.

Le tableau 3 synthétise les variables incluses dans ces trois thématiques, c'est-à-dire celles utilisées pour comparer les groupes, leurs échelles et leurs modalités. Certaines

modalités ont été dichotomisées pour remédier au petit nombre d’observations dans certains groupes. Par exemple, les perceptions de risque comprenaient les modalités « aucun risque », « risque minime », « risque modéré » et « risque élevé ». Dans les analyses, elles comprennent les modalités « aucun risque / risque minime » et « risque modéré / risque élevé ».

Tableau 3 — Principales thématiques du questionnaire, leurs variables, les échelles et les modalités des variables

Thématique	Variables (modalités)
Caractéristiques individuelles	Sexe (1=filles, 0=garçon); âge
Perceptions de risque par rapport à la santé associée à la consommation d’alcool et de cannabis	Niveau de risque pour la santé pour les personnes qui s’adonnent aux 6 activités suivantes ? : 1) boire de l’alcool de temps en temps 2) boire de l’alcool régulièrement 3) fumer de la marijuana ou du cannabis de temps en temps 4) fumer de la marijuana ou du cannabis régulièrement 5) consommer de la marijuana ou du cannabis d’une autre façon qu’en le fumant de temps en temps 6) consommer de la marijuana ou du cannabis d’une autre façon qu’en le fumant régulièrement (0=aucun risque ou risque minime; 1=risque modéré ou élevé).
Comportements de conduite avec facultés affaiblies et de consommation autorévélee de cannabis	<ul style="list-style-type: none"> – A déjà conduit sous l’influence de l’alcool et du cannabis, comme conducteur.trice et comme passager.ère (0=Non, jamais; 1=Oui, mais pas dans les 30 derniers jours et Oui, dans les 30 derniers jours); – expérimentation de l’alcool et du cannabis (0=Non, jamais; 1=Oui); – fréquence de consommation de cannabis dans les 12 derniers mois (1=Pas dans les 12 derniers mois à 8=Tous les jours) et dans les 30 derniers jours (1=Pas dans les 30 derniers jours à 6=Tous les jours); – âge à la première consommation d’alcool et de cannabis.

Portrait des échantillons

Les prochaines lignes présentent les principales différences observées entre les groupes par rapport au sexe des participant.e.s, à leur propension à l’expérimentation de l’alcool et du cannabis, et à la précocité de ces comportements.

D’emblée, on observe une surreprésentation d’adolescentes dans les deux groupes faisant partie de l’évaluation de programme. Alors que l’ECTADÉ est équilibré au

niveau du sexe de ses participant.e.s (7746/15400; 50,30 % de sexe féminin), le groupe d'intervention et le premier groupe de référence présentent des proportions plus élevées de participantes (51/81; 62,96 % de filles dans le groupe d'intervention et 25/39; 64,10 % de filles dans le premier groupe de référence).

Ensuite, les trois groupes se distinguent entre eux par rapport à l'expérimentation de l'alcool et du cannabis. Le premier groupe de référence rapporte avoir essayé l'alcool en plus grandes proportions par rapport aux deux autres groupes: 87,50 % des élèves du premier groupe de référence l'affirment (42/48). Cette proportion diminue à 59,76 % pour le groupe d'intervention (49/82) et à 49,47 % pour le sous-échantillon de l'ECTADÉ (7523/15208).

En ce qui a trait au cannabis, le premier groupe de référence a peu expérimenté cette substance par rapport aux deux autres groupes: 18,75 % des jeunes du premier groupe de référence ont répondu par l'affirmative à la question « As-tu déjà consommé ou essayé de la marijuana ou du cannabis (un joint, du pot, du hasch, du haschisch ou de l'huile de cannabis)? » (9/48). Cette proportion augmente à 24,71 % du côté des élèves de l'ECTADÉ (3753/15187). Toutes proportions gardées, les jeunes du groupe d'intervention se trouvent être plus nombreux à rapporter avoir essayé le cannabis: près du tiers de ces jeunes l'affirme (32,10 %; 26/81). Ces caractéristiques sont inversées pour les jeunes du premier groupe d'intervention: ils sont plus nombreux à avoir expérimenté avec l'alcool et moins nombreux avec le cannabis.

Pourtant, bien que le nombre d'observations soit faible, les âges moyens à la première consommation d'alcool et de cannabis sont similaires entre ces deux groupes de jeunes. En moyenne, les jeunes du groupe d'intervention ont consommé pour la première fois de l'alcool à 14,37 ans et du cannabis à 14,70 ans. Du côté des jeunes du premier groupe de référence, l'âge moyen rapporté pour leur première consommation d'alcool est de 14,62 ans et de 14,75 ans pour le cannabis.

Ces observations ont initialement suscité notre intérêt à réaliser la présente étude. En effet, l'alcool tend à être consommé plus tôt par les adolescent.e.s en plus d'être davantage consommé que les autres substances (Gallegos et coll., 2021; Spilka et coll., 2021). Ce constat remet en question la validité des réponses du groupe d'intervention.

Afin de capter les différences entre les groupes, des analyses bivariées ont été réalisées. Des tests d'indépendance du chi-carré et des tests de moyennes ont été privilégiés selon le type de variables analysées et le respect des postulats d'utilisation (Field et coll., 2012; Tabachnick et Fidell, 2014). La structuration des données et leur analyse ont été réalisées à l'aide des bibliothèques *gmodels* (version 2.18.1) et *Tidyverse* et *Readxl* (version 1.3.1) de R

(R Core Team, 2016), de l'environnement de développement *RStudio* (RStudio Team, 2020) et du logiciel *JASP* (JASP Team, 2021).

Perceptions de risques pour la santé

Les prochains paragraphes présentent les différences de perceptions entre les groupes vis-à-vis six activités de consommation. Ces activités varient selon deux paramètres : la substance consommée — l'alcool, le cannabis fumé et le cannabis consommé autrement — et la fréquence de consommation — de temps en temps ou régulièrement. À cet égard, le graphique 1 illustre les réponses en pourcentages de chacun des trois groupes pour la modalité « Aucun risque ou risque minime » à ces six questions.

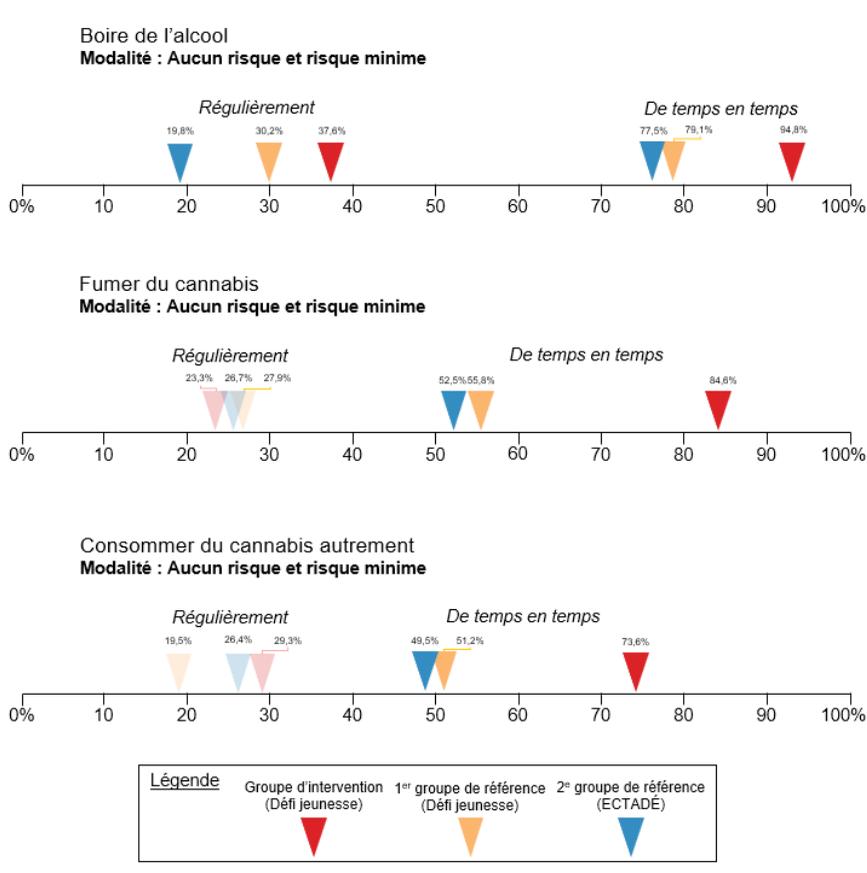
Le graphique illustre l'écart entre les trois groupes selon les six activités listées dans le questionnaire. Chaque groupe d'élèves est représenté par un triangle : le groupe d'intervention est en rouge, le triangle du premier groupe de référence est jaune et le deuxième groupe de référence est représenté par un triangle bleu. Les triangles clairs indiquent que le résultat du test d'indépendance du chi-carré n'est pas statistiquement significatif. Ainsi, on ne détecte pas de différence dans la distribution des fréquences entre les groupes.

Les premiers résultats concernent les perceptions de risque pour la santé au regard de la consommation d'alcool. À ce sujet, la distribution des fréquences, que la consommation soit occasionnelle ou régulière, n'est pas homogène à travers les groupes (consommation occasionnelle : $\chi^2(2, N = 14\,239) = 13,46, p = .0011$; consommation régulière : $\chi^2(2, N = 14\,226) = 18,16, p = .0001$). Par exemple, les jeunes du groupe d'intervention considèrent en plus grandes proportions que la consommation d'alcool régulière ne représente aucun risque ou un risque minime pour la santé par rapport aux deux groupes de référence. Un peu plus du tiers des jeunes du groupe d'intervention perçoit ce risque comme étant inexistant ou minime (29/77; 37,67 %) tandis qu'un peu moins du quart du premier groupe de référence (13/43; 23,23 %) et un peu moins du cinquième du deuxième groupe de référence partagent cet avis (1554/8610; 18,04 %). Ces résultats témoignent d'une minimisation du risque pour la santé associée à la consommation d'alcool de la part des jeunes du groupe d'intervention. Cette observation est similaire, que la consommation d'alcool soit occasionnelle ou fréquente.

S'agissant de la consommation de cannabis, les perceptions de risque différencient aussi les groupes par rapport à la fréquence, mais pas vis-à-vis la méthode de consommation. Que le cannabis soit fumé ou consommé autrement, les élèves des trois groupes rapportent presque unanimement qu'une consommation régulière représente un risque modéré

ou élevé pour la santé. Le test d'indépendance du chi-carré n'indique pas de différence statistiquement significative pour ces deux méthodes de consommation (cannabis fumé : $\chi^2(2, N = 13\,820) = 0,48, p = .79$); cannabis consommé autrement : $\chi^2(2, N = 13\,212) = 1,33, p = .51$). Ces proportions varient de 70,67 % à 80,49 % selon le groupe et la méthode visés par le point du questionnaire. Ainsi, les différences des échantillons ne semblent pas différencier les réponses des élèves pour ces deux activités.

Graphique 1 — Proportion des élèves affirmant que boire de l'alcool, fumer du cannabis et consommer du cannabis autrement ne représente « aucun risque » ou un « risque minime » pour la santé selon la fréquence de l'activité, régulièrement ou de temps en temps, selon l'appartenance aux groupes



En revanche, les différences sont notables lorsqu'il est question du risque associé à la consommation occasionnelle de cannabis, peu importe sa méthode. En effet, le test d'indépendance du chi-carré détecte une différence statistiquement significative indiquant une distribution hétérogène des jeunes à travers les groupes (cannabis fumé : $\chi^2(2, N = 13\ 146) = 17,70, p = .0001$); cannabis consommé autrement : $\chi^2(2, N = 13\ 212) = 32,12, p < .0001$). Par exemple, les jeunes du groupe d'intervention considèrent en plus grandes proportions que fumer du cannabis de temps en temps ne représente aucun risque ou un risque minime pour la santé par rapport aux deux groupes de référence.

Ces élèves (66/78) du groupe d'intervention sont près de 1,5 fois plus nombreux, toutes proportions gardées, à considérer la consommation de cannabis occasionnelle comme ne posant aucun risque ou un niveau minimal par rapport aux deux groupes de référence (24/43; 7238/13769). Ces résultats témoignent, comme pour la consommation d'alcool, d'une minimisation du risque pour la santé des jeunes du groupe d'intervention. Deux observations complètent ce constat.

D'abord, cette minimisation du risque est constante, que le cannabis soit fumé ou consommé autrement. Ensuite, l'éloignement du groupe d'intervention par rapport aux deux autres groupes est plus prononcé pour la consommation occasionnelle de cannabis que pour la consommation d'alcool. Ainsi, le paramètre de la substance consommée, en l'occurrence le cannabis, distingue davantage les perceptions rapportées du groupe d'intervention par rapport à celles des deux autres groupes.

En résumé, les résultats indiquent que pour quatre activités sur six, le groupe d'intervention est tenu à l'écart des deux groupes de référence. Ainsi, les jeunes recrutés par les intervenants du « Défi jeunesse de l'Atlantique » se distinguent des autres pour les questions relatives aux perceptions de risque pour la santé. Ces résultats témoignent d'une polarisation entre le groupe d'intervention et les deux groupes de références. Fait intéressant, le premier et le deuxième groupe de référence ont des réponses similaires, malgré le recours à des stratégies d'échantillonnage et un contexte sanitaire dissemblables.

Comportements de conduite avec facultés affaiblies et de consommation de cannabis

Cette section compare les groupes vis-à-vis la conduite autorapportée avec facultés affaiblies, comme conducteur et comme passager. Les variations de consommation autorapportée de cannabis complètent ces résultats; la conduite avec facultés affaiblies par le cannabis a comme condition préalable la consommation de cette substance.

D'abord, les élèves ayant affirmé avoir pris le volant d'un véhicule avec les facultés affaiblies sont rares pour tous les groupes, qu'il s'agisse de l'alcool ou du cannabis. Les proportions des personnes qui se sont déjà adonnées à cette pratique varient de 6,25 % à 12,77 % pour l'alcool et de 2,56 % à 6,89 % pour le cannabis selon les groupes. En ce qui a trait au fait d'avoir été passager d'un véhicule conduit par une personne ayant ses facultés affaiblies, que le conducteur ait consommé de l'alcool ou du cannabis, la relation n'est pas statistiquement significative.

Les habitudes autorévéloées de consommation de cannabis des jeunes vont dans le même sens, car la prévalence de la consommation de cannabis au sein des échantillons est faible. En effet, la plupart des jeunes interrogés affirment ne jamais consommer de cannabis. Le test indépendant du chi-carré ne capte pas de différence entre les groupes : 32,10 % des jeunes du groupe d'intervention rapportent avoir consommé ou essayé du cannabis (26/81) contre 18,75 % des élèves du premier groupe de référence (9/48) et 24,70 % du deuxième (3753/15187). Dans le même ordre d'idée, le test de différence de moyennes Kruskal-Wallis ne capte pas de différence significative des fréquences de consommation de cannabis des 12 derniers mois et des 30 derniers jours (($H(2)=3,9223$; $p=.14$) et ($H(2)=4,9536$; $p=.08$)). Ainsi, une faible proportion de jeunes, tous groupes confondus, rapportent avoir déjà consommé du cannabis au cours de leur vie. Ce constat se traduit par une difficulté à capter des différences intergroupes vis-à-vis l'expérimentation et la fréquence de consommation de cette substance.

En bref, à première vue, les comportements visés par le « Défi jeunesse de l'Atlantique » tels que rapportés par les jeunes ne semblent pas caractérisés ni par la manière dont les participant.e.s ont été recrutés ni par la pandémie de COVID-19. Cependant, qu'il s'agisse de conduite avec facultés affaiblies ou de consommation de cannabis, ce sont dans les deux cas des comportements peu fréquents. Ainsi, le niveau de puissance statistique n'est pas suffisant pour capter des différences. En effet, la rareté d'un comportement peut augmenter les risques d'erreur de type II, c'est-à-dire de faux négatif (Bourque, 2016).

Discussion

Le recours à l'évaluation participative va de pair avec son potentiel vanté par la communauté scientifique (Cousins et coll., 1996; Torres et Preskill, 1999; Weaver et Cousins, 2004). Or, il est difficile d'être sans équivoque sur ses avantages devant sa diversité de formes. Celle-ci pose des questions de partage du pouvoir (Cousins et Whitmore, 1998; Fukunaga, 2015; Taut, 2008; Torres et Preskill, 1999).

L'approche participative demeure sujette à caution en raison du manque d'évaluations de la qualité des devis qui en découlent (Brandon et Fukunaga, 2014 ; Cousins, 2003 ; Torres et Preskill, 1999). La présente étude contribue à cette lacune par la présentation d'une étude de cas. Elle explore l'étape de la collecte des données à travers les caractéristiques de trois échantillons qui se distinguent par la méthode et le contexte de composition de ceux-ci.

Dans la même veine que les travaux de Brandon et Fukunaga (2014), l'étude se penche sur les parties prenantes les plus couramment impliquées dans les évaluations de programmes. Lors du « Défi jeunesse de l'Atlantique », les responsables du programme ont réalisé la collecte des données du groupe d'intervention et cette pratique n'est pas rare (Thayer et Fine, 2001).

Dans l'étude de cas présentée, l'implication des parties prenantes semble avoir introduit un biais de sélection plus fort dans le groupe d'intervention que dans l'échantillon comportant le plus faible nombre de jeunes volontaires, c'est-à-dire le premier groupe de référence. Ce constat n'est observable qu'en le mettant en rapport avec un troisième groupe issu d'une collecte de données antérieures de plus grande envergure, soit le sous-échantillon de l'ECTADÉ.

L'étude de cas met en lumière qu'une partie de l'assurance-qualité peut être hors du contrôle des évaluateurs. Par ailleurs, les trois groupes se distinguent les uns des autres différemment. Le faible taux de participation des écoles du premier groupe de référence le distingue des deux autres. Le fait de ne pas avoir été colligées dans un contexte de pandémie de COVID-19 et de cannabis légalisé distingue les données issues du deuxième groupe de référence par rapport aux deux autres. Malgré ces différences, le groupe d'intervention sort presque systématiquement du lot en ce qui a trait aux perceptions rapportées des risques pour la santé par les jeunes. Alors que l'évaluateur doit concilier la faisabilité d'un projet et sa rigueur méthodologique (Braverman, 2013 ; Eck, 2006), l'importance de ce constat transcende le contexte sanitaire ayant bouleversé le programme.

Les adaptations à la crise sanitaire ont réduit le « Défi jeunesse de l'Atlantique » à la seule activité de campagne virtuelle avec les jeunes leaders. Leur vulnérabilité pourrait-elle donc expliquer qu'ils se distinguent des autres ? Deux éléments indiquent que les jeunes du groupe d'intervention ne semblent pas plus vulnérables que ceux du premier groupe de référence : d'abord, les intervenants n'avaient pas de directives particulières de cibler des jeunes vulnérables dans la sollicitation des participant.e.s au programme. Ensuite, l'analyse préliminaire des résultats du sondage indique certaines incohérences avec cette hypothèse : l'expérimentation autorévélee avec le cannabis est plus prévalente du côté des

jeunes du groupe d'intervention, mais cette expérimentation n'est pas plus précoce. En parallèle, des études canadiennes récentes n'indiquent pas une augmentation de l'intérêt des jeunes envers le cannabis depuis la légalisation en 2018 ou avec les mesures de confinement de la crise sanitaire (Leatherdale et coll., 2021; Zuckermann et coll., 2021).

Les jeunes du premier groupe de référence auraient-ils pu sous-rapporter leur expérimentation avec le cannabis? La question se pose aussi auprès de l'échantillon de l'ECTADÉ. Pour ces deux groupes, le recrutement s'est fait à travers les écoles, ce biais peut être distribué de façon similaire entre ces deux groupes.

Ainsi, les résultats de l'étude portent à croire que le biais de sélection introduit par le recrutement des intervenants est saillant dans les perceptions de risques rapportées. La méthode d'échantillonnage ressort de cette observation. Les différences sont moindres pour le premier groupe de référence, le groupe le moins nombreux, dont la distribution des réponses se situe plus proche du deuxième groupe de référence. Nous concluons que, pour augmenter la qualité d'une évaluation issue d'une pratique participative, la mise en œuvre de celle-ci nécessite un certain contrôle de la part des évaluateurs (Scriven, 2003).

Dans l'étude de cas présentée, la collaboration aurait pu être améliorée à deux niveaux: durant l'adaptation aux contraintes imposées par la santé publique et pendant l'offre du programme. D'une part, le programme a été adapté sans la collaboration des chercheuses. D'autre part, durant l'offre du programme, les échanges auraient pu être plus fréquents entre la coordonnatrice du programme et l'équipe de chercheuses. En effet, les échanges avec les intervenants de premières lignes n'ont eu lieu qu'à la fin du programme. Un lien plus étroit avec les formateurs aurait pu réduire les biais dans le recrutement des participant.e.s au programme et au questionnaire.

La présente étude visite le concept de compromis dans un contexte d'évaluation participative. L'intervention de l'organisme a pu compromettre la qualité scientifique du devis. En effet, les données issues du groupe d'intervention semblent être les moins fiables des trois groupes. Il y a lieu de se demander si ces données ne compromettent pas l'utilité instrumentale de l'évaluation. Dans ces conditions, la question de l'intervention des parties prenantes ne peut être évoquée sans parler de la qualité du processus d'évaluation.

Conclusion

La présente étude de cas porte sur les biais de sélection des participant.e.s dans l'évaluation d'un programme de prévention de la conduite avec facultés affaiblies par le cannabis chez

les jeunes élèves des provinces de l'Atlantique. Pour mettre en contraste les caractéristiques des jeunes du groupe d'intervention et du premier groupe de référence, tous deux interrogés à l'été 2021, un sous-échantillon de l'ECTADÉ (2016-2017) a été joint à l'analyse formant ainsi un deuxième groupe de référence. Ces trois groupes d'élèves ont été comparés sur la base de deux sections du questionnaire : les perceptions de risque sur la santé ainsi que les comportements associés à la conduite avec facultés affaiblies et à la consommation de cannabis.

La présente étude comporte deux principaux résultats. Le premier résultat concerne les perceptions de risques selon l'appartenance aux groupes d'élèves. Il indique que, pour une majorité des activités proposées, le groupe d'intervention sort du lot par rapport aux groupes de référence. Les réponses des jeunes recruté.e.s par les intervenants du « Défi jeunesse de l'Atlantique » minimisent davantage les risques de santé associés à la consommation occasionnelle d'alcool et de cannabis. On retient également que les deux groupes de référence sont similaires pour cette thématique, malgré des stratégies d'échantillonnage et un contexte sanitaire différents.

Le deuxième résultat révèle que les réponses concernant la conduite avec facultés affaiblies et la consommation de cannabis ne semblent pas avoir été affectées par les stratégies d'échantillonnage ou le contexte sanitaire différenciant les trois groupes d'élèves. On postule que la rareté de ces comportements soulève des enjeux de puissance statistique. Toutefois, la rareté de la présence de comportements de conduite avec facultés affaiblies porte à réfléchir sur la cible initiale du programme.

La présente étude comprend plusieurs limites. D'abord, on ne peut pas inférer de liens de causalité entre les groupes; la taille des échantillons et les analyses réalisées ne permettent pas de le faire. Ni le groupe d'intervention ni le premier groupe de référence n'est représentatif des jeunes de l'Atlantique. Ensuite, le contexte de l'évaluation n'a permis ni la composition d'échantillons de grande taille ni l'affectation aléatoire des jeunes dans les groupes. Enfin, le cas présenté ne documente que la collecte des données et non pas le processus d'évaluation dans son ensemble. C'est pourquoi les résultats s'inscrivent dans la présentation d'une étude de cas. Elle revêt un potentiel heuristique, malgré la petite taille des échantillons (Eck, 2006) — surtout si les évaluateurs sont sensibles aux questions de biais.

Annexe A

Consigne de départ du questionnaire en ligne

Dans le cadre de ce projet d'évaluation, nous demandons à des jeunes comme toi de compléter l'étude sur les habitudes et les attitudes par rapport à la consommation de substances afin de permettre d'évaluer comment les activités auxquelles tu participeras durant le Défi de l'Atlantique pour prévenir la conduite sous influence ont des effets sur tes connaissances, tes décisions et tes comportements liés à la conduite avec les facultés affaiblies par le cannabis. Pour arriver à ce résultat, nous voulons aussi connaître tes habitudes de consommations. Ce projet est mené par [noms, fonctions et institutions d'attache des personnes autrices du manuscrit]. Nous te remercions de nous aider dans nos efforts. Même si tu ne consommes pas de tabac, d'alcool ou de drogues, nous te demandons de répondre à chaque question, afin que tous les élèves prennent autant de temps pour répondre au questionnaire. Ce questionnaire N'EST PAS un test. Toutes tes réponses demeureront confidentielles. Personne ne connaîtra tes réponses, pas même tes parents ni tes professeurs. Nous te demandons donc de répondre honnêtement aux questions. S'il y a une question que tu ne comprends pas ou à laquelle tu ne désires pas répondre, n'y réponds pas et passe à la question suivante. Choisis l'option qui se rapproche le plus de ce que tu penses ou de ce qui s'applique à ta situation selon toi. **TES RÉPONSES SONT CONFIDENTIELLES.**

Références bibliographiques

- Bouchard, S. et Cyr, C. (2005). *Recherche psychosociale : Pour harmoniser recherche et pratique* (2^e éd.). Presses de l'Université du Québec. <http://ebookcentral.proquest.com/lib/umontreal-ebooks/detail.action?docID=599920>
- Bourque, J. (2016). *Manuel d'introduction à la statistique appliquée aux sciences sociales*. Presses de l'Université Laval.
- Brandon, P. R. (1998). Stakeholder participation for the purpose of helping ensure evaluation validity : Bridging the gap between collaborative and non-collaborative evaluations. *American Journal of Evaluation*, 19(3), 325-337. <https://doi.org/10.1177/109821409801900305>
- Brandon, P. R. et Fukunaga, L. L. (2014). The state of the empirical research literature on stakeholder involvement in program evaluation. *American Journal of Evaluation*, 35(1), 26-44. <https://doi.org/10.1177/1098214013503699>
- Braverman, M. T. (2013). Negotiating measurement : Methodological and interpersonal considerations in the choice and interpretation of instruments. *American Journal of Evaluation*, 34(1), 99-114. <https://doi.org/10.1177/1098214012460565>
- Brisolara, S. (1998). The history of participatory evaluation and current debates in the field. *New Directions for Evaluation*, 1998(80), 25-41. <https://doi.org/10.1002/ev.1115>
- Brisson, D. (2007). Collaborative evaluation in a community change initiative : Dilemmas of control over technical decision making. *The Canadian Journal of Program Evaluation*, 22(2), 21-39.
- Burkhalter, R., Thompson-Haile, A., Rynard, V. et Manske, S. (2017) *Guide de l'utilisateur des microdonnées de l'Enquête canadienne sur le tabac, l'alcool et les drogues chez les élèves de 2016-2017*. Waterloo, Ontario : Centre pour l'avancement de la santé des populations Propel, Université de Waterloo, 1-59.
- Campbell, D. T. et Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally College Publishing.
- Cousins, J. B. (2003). Utilization effects of participatory evaluation. Dans T. Kellaghan et D. L. Stufflebeam (dir.), *International Handbook of Educational Evaluation* (p. 245-265). Springer Netherlands. https://doi.org/10.1007/978-94-010-0309-4_16
- Cousins, J. B., Donohue, J. J. et Bloom, G. A. (1996). Collaborative evaluation in North America : Evaluators' self-reported opinions, practices and consequences. *Evaluation Practice*, 17(3), 207-226. [https://doi.org/10.1016/S0886-1633\(96\)90002-1](https://doi.org/10.1016/S0886-1633(96)90002-1)
- Cousins, J. B. et Whitmore, E. (1998). Framing participatory evaluation. *New Directions for Evaluation*, 1998(80), 5-23. <https://doi.org/10.1002/ev.1114>

- Crishna, B. (2006). Participatory evaluation (II) – translating concepts of reliability and validity in fieldwork. *Child: Care, Health and Development*, 33(3), 224-229. <https://doi.org/10.1111/j.1365-2214.2006.00658.x>
- Cusson, M., Tremblay, P., L.-Biron, L., Ouimet, M. et Grandmaison, R. (1994). *La prévention du crime. Guide de planification d'évaluation* (p. 116). École de criminologie de l'université de Montréal.
- Eck, J. E. (2006). When is a bologna sandwich better than sex? A defense of small-n case study evaluations. *Journal of Experimental Criminology*, 2(3), 345-362. <https://doi.org/10.1007/s11292-006-9014-9>
- Ekblom, P. et Pease, K. (1995). Evaluating crime prevention. *Crime and Justice*, 19, 585-662. <https://doi.org/10.1086/449238>
- Farrington, D. P. (2006). Methodological quality and the evaluation of anti-crime programs. *Journal of Experimental Criminology*, 2(3), 329-337. <https://doi.org/10.1007/s11292-006-9012-y>
- Fetterman, D. M., Rodríguez-Campos, L., Zukoski, A. P. et Contributors. (2017). *Collaborative, participatory, and empowerment evaluation : Stakeholder involvement approaches* (1st edition). The Guilford Press.
- Fetterman, D., Rodríguez-Campos, L., Wandersman, A. et O'Sullivan, R. G. (2014). Collaborative, participatory, and empowerment evaluation : Building a strong conceptual foundation for stakeholder involvement approaches to evaluation (A response to Cousins, Whitmore, and Shulha, 2013). *American Journal of Evaluation*, 35(1), 144-148. <https://doi.org/10.1177/1098214013509875>
- Field, A., Miles, J. et Field, Z. (2012). *Discovering statistics using R*. Sage publications.
- Fukunaga, L. (2015). *The theory of planned behavior and stakeholder involvement in program evaluation : Evaluator decisions about diversity, depth, and control* [Thesis, [Honolulu] : [University of Hawaii at Manoa], [May 2015]]. <http://scholarspace.manoa.hawaii.edu/handle/10125/50975>
- Gallegos, M. I., Zaring-Hinkle, B., Wang, N. et Bray, J. H. (2021). Detachment, peer pressure, and age of first substance use as gateways to later substance use. *Drug and Alcohol Dependence*, 218, 108352. <https://doi.org/10.1016/j.drugalcdep.2020.108352>
- JASP Team. (2021). *JASP (Version 0.16)[Computer software]*. <https://jasp-stats.org/>
- King, J. A. et Ehlert, J. C. (2008). What we learned from three evaluations that involved stakeholders. *Studies in Educational Evaluation*, 34(4), 194-200. <https://doi.org/10.1016/j.stueduc.2008.10.003>
- Lavoie, J. et Panet-Raymond, J. (2020). *La pratique de l'action communautaire* (4^e éd.). Presses de l'Université du Québec.
- Lemay, J.-F., Leclerc, L., Bélanger, S. et Messikh, D. (2013). Les particularités de l'évaluation pour les groupes communautaires : Réflexions et enjeux. *Bulletins de la Société québécoise d'évaluation de programme*, 25(1), 8-10.

- McDavid, J. C., Huse, I. et Hawthorn, L. R. L. (2018). *Program evaluation and performance measurement : An introduction to practice*. SAGE Publications.
- Mertens, D. M. et Wilson, A. T. (2018). *Program evaluation theory and practice, Second Edition : A comprehensive guide*. Guilford Publications. <http://ebookcentral.proquest.com/lib/umontreal-ebooks/detail.action?docID=5601866>
- Midy, F., Vanier, C. et Grant, M. (1998). *Guide d'évaluation participative et de négociation*.
- Morris, M. (2003). Ethical considerations in evaluation. Dans T. Kellaghan et D. L. Stufflebeam (dir.), *International Handbook of Educational Evaluation* (p. 303-327). Springer Netherlands. https://doi.org/10.1007/978-94-010-0309-4_19
- Panizza, F., Vostroknutov, A. et Coricelli, G. (2021). How conformity can lead to polarised social behaviour. *PLoS Computational Biology*, 17(10). <https://doi.org/10.1371/journal.pcbi.1009530>
- Patton, M. Q. et LaBossière, F. (2012). L'évaluation axée sur l'utilisation. Dans V. Ridde et C. Dagenais (dir.), *Approches et pratiques en évaluation de programmes* (Les Presses de l'Université de Montréal, p. 145-160).
- R Core Team. (2016). R: A language and environment for statistical computing. Dans *R Foundation for Statistical Computing*.
- Rodríguez-Campos, L. et Rincones-Gómez, R. (2012). *Collaborative Evaluations : Step-by-Step, Second Edition* (2 edition). Stanford Business Books.
- Rossi, P. H. et Freeman, H. E. (1993). *Evaluation : A systematic approach*. Newbury Park, CA: Sage.
- Scriven, M. (2003). Evaluation theory and metatheory. Dans T. Kellaghan et D. L. Stufflebeam (dir.), *International Handbook of Educational Evaluation* (p. 15-30). Springer Netherlands. https://doi.org/10.1007/978-94-010-0309-4_3
- Senanayake, S., Abell, B., Novick, M., Exley, H., Dolejs, W., Hutchinson, K., McPhail, S. et Kularatna, S. (2021). Impact and outcome evaluation of HealthPathways : A scoping review of published methodologies. *Journal of Primary Health Care*, 13(3), 260-273. <https://doi.org/10.1071/HC21067>
- Serafini, E. J., Lake, J. B. et Long, M. H. (2015). Needs analysis for specialized learner populations : Essential methodological improvements. *English for Specific Purposes*, 40, 11-26. <https://doi.org/10.1016/j.esp.2015.05.002>
- Spilka, S., Philippon, A., Le Nézet, O., Janssen, E., Eroukmanoff, V. et Godeau, E. (2021). Usages d'alcool, de tabac et de cannabis chez les élèves de 3e en 2021. *Tendances*, 1(148), 1-4.
- Stufflebeam, D. L. et Coryn, C. L. S. (2014). *Evaluation theory, models, and applications* (2nd edition). Jossey-Bass.
- Tabachnick, B. G. et Fidell, L. S. (2014). *Using multivariate statistics* (Pearson new international edition, sixth edition). Pearson.

- Taut, S. (2008). What have we learned about stakeholder involvement in program evaluation? *Studies in Educational Evaluation*, 7.
- Thayer, C. et Fine, A. (2001). Evaluation and outcome measurement in the non-profit sector : Stakeholder participation. *Evaluation and Program Planning*, 24, 103-108. [https://doi.org/10.1016/S0149-7189\(00\)00051-3](https://doi.org/10.1016/S0149-7189(00)00051-3)
- Torres, R. T. et Preskill, H. (1999). Ethical dimensions of stakeholder participation and evaluation use. *New Directions for Evaluation*, 1999(82), 57-66. <https://doi.org/10.1002/ev.1137>
- Torres-Cuello, M. A., Pinzón-Salcedo, L. et Midgley, G. (2018). Developing a systemic program evaluation methodology : A critical systems perspective. *Systems Research and Behavioral Science*, 35(5), 538-547. <https://doi.org/10.1002/sres.2561>
- Tougas, A.-M., Bérubé, A. et Kishchuk, N. (2021). *Évaluation de programme : De la réflexion à l'action*. Presses de l'Université du Québec.
- Weaver, L. et Cousins, J. B. (2004). Unpacking the participatory process. *Journal of MultiDisciplinary Evaluation*, 1(1), 19-40.
- Weiss, C. H. (1998). *Evaluation : Methods for studying programs and policies* (2nd ed.). Prentice Hall.
- White, H. (2017). Effective targeting of social programmes : An overview of issues. *Journal of Development Effectiveness*, 9(2), 145-161. <https://doi.org/10.1080/19439342.2017.1317282>
- Yarbrough, D. B., Shulha, L. M., Hopson, R. K. et Caruthers, F. A. (2010). *The program evaluation standards : A guide for evaluators and evaluation users* (Third edition). SAGE Publications.

Notes

- ¹ La méthode d'échantillonnage privilégiée ne tient pas compte du nombre d'élèves par école. Elle a toutefois comme avantage de pondérer les invitations à participer à l'étude en donnant plus d'importance aux provinces les plus peuplées.
- ² Pour cette cohorte, le Nouveau-Brunswick n'a pas participé à l'enquête.