

Guide de l'utilisateur des microdonnées

Enquête longitudinale nationale sur les enfants et les jeunes

Cycle 8

Septembre 2008 à juillet 2009



Statistique
Canada

Statistics
Canada

Canada

Table des matières

1.0	Introduction	9
2.0	Contexte	11
3.0	Objectifs	13
4.0	Principaux changements apportés au cycle 8 depuis le cycle 7	15
4.1	Changements apportés au contenu	15
4.2	Changements méthodologiques	15
5.0	Méthodologie de l'enquête — Échantillon	17
5.1	Aperçu du plan d'échantillonnage — Cycles 1 à 8	17
5.2	Enquête sur la population active	18
5.2.1	Population cible	19
5.2.2	Stratification	19
5.2.3	Délimitation et sélection des grappes	20
5.2.4	Sélection des logements	20
5.2.5	Changements introduits dans le cadre du remaniement de 2004	21
5.2.6	Renouvellement de l'échantillon	21
5.2.7	Membres des ménages admissibles aux fins de l'Enquête sur la population active	22
5.3	Précisions concernant l'échantillon de l'Enquête longitudinale nationale sur les enfants et les jeunes	22
5.3.1	Cohorte initiale — Cycles 1 à 8	23
5.3.2	Participation des cohortes du Développement de la petite enfance au cycle 8	24
5.4	Populations de référence longitudinale et transversale pour l'établissement des poids de l'Enquête longitudinale nationale sur les enfants et les jeunes	29
5.4.1	Populations longitudinales des cohortes	30
5.4.2	Cohortes et leurs populations transversales	31
6.0	Collecte des données	35
6.1	Composante « Ménage »	35
6.2	Composante « Enfant »	36
6.3	Composante « Adulte »	36
6.4	Composante « Jeune »	36
6.5	Tailles de l'échantillon au cycle 8	36
6.6	Mesures directes	38
6.7	Questionnaires autoadministrés (14 à 17 ans)	38
6.8	Personnel de collecte (formation, supervision et contrôle)	39
6.8.1	Interview effectuée dans une langue autre qu'une des langues officielles	39
7.0	Traitement des données	41
7.1	Vérification informatique	41
7.2	Saisie des données	42
7.3	Nettoyage	42
7.4	Vérification de l'âge et du genre	43
7.5	Vérification des liens	43
7.6	Vérification préliminaire	43
7.7	Vérification de cheminement	43
7.8	Codage des questions ouvertes	43
7.9	Vérification de la cohérence	44
7.10	Indicateurs d'imputation	45
7.11	Établissement de variables dérivées et vérifications longitudinales	45
7.12	Structure de codage des variables	46
7.13	Nomenclature	47

7.14	Exemples de noms de variables	49
7.15	Siglaion des sections du questionnaire	49
7.16	Fichiers de traitement final et fichiers principaux	52
8.0	Contenu de l'enquête	53
8.1	Composantes de l'enquête	53
8.2	Variables démographiques	54
8.3	Questionnaire de l'adulte	54
8.4	Questionnaire de l'enfant	56
8.5	Questionnaire du jeune (16 à 25 ans)	60
8.6	Questionnaires autoadministrés (14 à 17 ans)	64
9.0	Validation des échelles d'enquête	69
9.1	Validation des données d'échelle	69
9.1.1	Définition de l'échelle	69
9.1.2	Échelles et calculs	69
9.1.3	Évaluation des données de l'échelle	69
9.2	Analyse factorielle	70
9.2.1	Analyse factorielle des échelles	70
9.2.2	Transformation des données au moyen de l'échelonnage optimal	71
9.3	Calcul des scores et imputation des réponses	71
9.3.1	Calcul des scores pour chaque facteur	71
9.3.2	Exemple de calcul des scores factoriels	71
9.3.3	Poids factoriel négatif	72
9.3.4	Codes de non-réponse	72
9.3.5	Items bruts	72
9.4	Mesures de fiabilité des échelles	72
9.4.1	Coefficient alpha de Cronbach	73
9.4.2	Interprétations du coefficient alpha de Cronbach	73
9.4.3	Qu'est-ce qu'un niveau satisfaisant de fiabilité?	73
9.5	Échelles relatives dans le questionnaire du parent	73
9.5.1	Échelle de dépression	74
9.5.2	Échelle de fonctionnement de la famille	74
9.5.3	Échelle de sécurité du quartier	75
9.5.4	Échelle du soutien social	76
9.5.5	Échelles de comportement	77
9.5.6	Échelle du développement moteur et social	80
9.5.7	Échelles du rôle parental	81
9.5.8	Échelle relative dans les étapes du développement	83
9.6	Échelles relatives dans le questionnaire des jeunes	83
9.6.1	Échelle de dépression	83
9.6.2	Échelle de structure du quartier	84
9.6.3	Échelle de l'estime globale de soi	84
9.6.4	Échelle relative au quotient émotionnel	85
9.6.5	Échelle du soutien social	89
9.6.6	Échelle des amis	89
9.6.7	Échelles « Mes parents et moi »	90
9.6.8	Échelle de résolution des conflits	90
9.6.9	Échelles de comportement	91
9.7	Résumé des échelles du cycle 8	93
9.7.1	Échelles relatives dans le questionnaire du parent	93
9.7.2	Échelles relatives dans le questionnaire des enfants et des jeunes (autoadministré)	94
9.7.3	Échelles relatives au questionnaire des jeunes (autoadministré)	95

10.0	Imputation	97
10.1	Imputation du revenu du ménage	97
10.2	Imputation du revenu des jeunes — 16 et 17 ans	101
10.3	Imputation du revenu des jeunes — 18 ans et plus	102
10.4	Imputation de l'échelle de Développement moteur et social (DMS)	104
11.0	Pondération et traitement de la non-réponse	107
11.1	Poids disponibles au cycle 8	107
11.1.1	Poids longitudinaux : poids entonnoir (variable HWTCWd1L) et poids non-entonnoir (variable HWTCW01L)	107
11.1.2	Populations longitudinales	108
11.1.3	Poids transversaux (variable HWTCW01C)	109
11.2	Méthode de pondération	109
11.2.1	Poids déterminé par le plan de l'échantillonnage de l'Enquête longitudinale nationale sur les enfants et les jeunes	109
11.2.2	Premier ajustement : ajustement pour la non-réponse	110
11.2.3	Second ajustement : stratification a posteriori	111
11.2.4	Différences entre la méthode de pondération des cycles 6, 7 et 8 et celle des cycles précédents	111
11.3	Application de la méthode de pondération	111
11.3.1	Pondération longitudinale non-entonnoir	111
11.3.2	Pondération longitudinale entonnoir	112
11.3.3	Pondération transversale	113
12.0	Qualité des données, taux de réponse et couverture	115
12.1	Erreur d'échantillonnage	115
12.2	Erreurs non dues à l'échantillonnage	115
12.3	Non-réponse totale et biais de non-réponse	115
12.3.1	Définitions de la réponse	116
12.3.2	Taux de réponse transversaux	118
12.3.3	Taux d'attrition longitudinaux pour la cohorte du cycle 1	119
12.3.4	Taux de réponse longitudinaux pour les enfants sélectionnés aux cycles 5 et 6	123
12.4	Non-réponse partielle	123
12.4.1	Composante de l'enfant	124
12.4.2	Composante de la personne la mieux renseignée	125
12.4.3	Composante du conjoint	125
12.4.4	Composante du jeune	126
12.4.5	Échelle de vocabulaire en images de Peabody	126
12.4.6	Connaissance des nombres	126
12.4.7	Qui suis-je?	127
12.4.8	Tests de mathématiques	127
12.4.9	Résolution de problèmes	127
12.4.10	Évaluation de la littératie	128
12.4.11	Évaluation de la numératie	128
12.4.12	Questionnaires autoadministrés	128
12.5	Non-réponse de cycle	129
12.6	Incidence des erreurs de réponse pour les caractéristiques rares	129
12.7	Erreurs de réponse liées à des comportements déviants ou à des questions de nature délicate	129
12.8	Erreurs de réponse dues à des approximations	130
12.9	Erreurs de réponse dues à des erreurs de mémoire	130
12.10	Erreurs de réponse dues aux réponses de la personne interposée	130
12.11	Schémas de réponse dont les catégories de réponse sont mal définies	131
12.12	Langue de l'interview	131
12.13	Renseignements contradictoires	131
12.14	Qualité des données de l'indice de masse corporelle	132

12.14.1	Indice de masse corporelle	132
12.14.2	Indice de masse corporelle — Centers for Disease Control	133
12.14.3	Indice de masse corporelle : seuils internationaux	133
12.14.4	Indice de masse corporelle — Qualité des données	133
12.15	Biais d'accoutumance	134
12.16	Personne la mieux renseignée	134
12.17	Couverture des enfants canadiens par l'échantillon de l'ELNEJ	135
12.17.1	Exclusions de l'EPA	135
12.17.2	Couverture des immigrants récents	135
12.17.3	Couverture selon le mois de naissance	136
12.17.4	Couverture selon l'ordre de naissance dans la cohorte originale	136
12.18	Validation des données	136
12.19	Conclusion	137
13.0	Estimation de la variance	139
13.1	Termes liés à l'erreur et à la variance d'échantillonnage	140
13.2	Coefficient de variation	141
13.3	Importance de déclarer la variance d'échantillonnage	143
13.4	Calcul de la variance d'échantillonnage	144
13.4.1	Chiffrier de variances d'échantillonnage approximatives pour des proportions	146
13.4.2	Macros SAS et SPSS pour calculer la variance d'échantillonnage à partir des poids bootstrap : Bootvar	147
13.4.3	Autres outils informatiques	147
13.4.4	Linéarisation de Taylor et autres techniques	148
14.0	Mesures directes	149
14.1	Échelle de vocabulaire en images de Peabody	149
14.1.1	Propriétés psychométriques des scores	150
14.2	Évaluation de la connaissance des nombres	151
14.3	Qui suis-je?	153
14.4	Exercice de calculs mathématiques	156
14.5	Exercice de résolution de problèmes (16 et 17 ans)	157
14.6	Évaluation de la littératie (18 et 19 ans)	159
14.7	Évaluation de la numératie (20 et 21 ans)	160
15.0	Enjeux analytiques	161
15.1	Incidence de l'utilisation d'un plan de sondage complexe sur l'analyse	161
15.2	Unité d'analyse	162
15.3	Analyse longitudinale et analyse transversale	162
15.4	Estimations pondérées simples (totaux, moyennes et proportions)	163
15.4.1	Estimation d'un total de la population	163
15.4.2	Estimation d'une moyenne de population	164
15.4.3	Estimation d'une proportion de la population	164
15.4.4	Estimation de domaines de la population	164
15.5	Poids normalisés	165
15.6	Mesures répétées	168
15.7	Regroupement des données	169
15.8	Non-réponse	171
15.9	Autres sources d'erreurs non dues à l'échantillonnage	172
15.10	Calcul de la variance avec certains logiciels	172
15.11	Coefficients de variation pour des proportions	173
15.12	Écart type et erreur type	173
15.13	Comprendre la différence entre « Non déclaré », « Ne sait pas », « Refus » et « Enchaînement valide »	173

16.0	Lignes directrices concernant la totalisation, l'analyse et la diffusion des données-----	175
16.1	Lignes directrices concernant l'arrondissement-----	175
16.2	Lignes directrices concernant la pondération de l'échantillon en vue de la totalisation ---	176
16.3	Lignes directrices concernant la modélisation statistique -----	176
16.4	Lignes directrices concernant la diffusion -----	176
16.5	Modélisation des données de l'ELNEJ et des poids bootstrap-----	177
Annexe 1	— Lignes directrices pour les chercheurs et les analystes de l'Enquête longitudinale nationale sur les enfants et les jeunes -----	179
Annexe 2	— Analyse de la non-réponse partielle -----	185
Annexe 3	— Concordance entre les noms de variables du traitement et les noms de variables de la diffusion, pour les questionnaires autoadministrés du cycle 8 de l'ELNEJ -----	187

1.0 Introduction

Le cycle 8 de l'Enquête longitudinale nationale sur les enfants et les jeunes (ELNEJ) a été mené entre septembre 2008 et juillet 2009 par Statistique Canada, en collaboration avec Ressources humaines et Développement des compétences Canada.

Le présent manuel a été préparé afin de faciliter la manipulation des fichiers de microdonnées sur les résultats de l'enquête et pour consigner les questions liées à la qualité des données et aux autres questions analytiques concernant l'ELNEJ.

Toutes les questions au sujet de l'ensemble des données ou de son utilisation devraient être adressées à :

Statistique Canada

Services à la clientèle

Division des enquêtes spéciales

Téléphone : 613-951-3321 ou le numéro sans frais : 1 800 461-9050

Télécopieur : 613-951-4527

Courriel : des@statcan.gc.ca

Division des enquêtes spéciales

Statistique Canada

150, promenade Tunney's Pasture

Ottawa (Ontario) K1A 0T6

Ressources humaines et Développement des compétences Canada

Manon Monette

Gestionnaire

Direction de la gestion de la connaissance et des données

Ressources humaines et Développement des compétences Canada

Téléphone : 819-994-9565

Courriel : manon.monette@hrsdc.rhdcc.gc.ca

2.0 Contexte

L'Enquête longitudinale nationale sur les enfants et les jeunes (ELNEJ) est une étude à long terme qui a débuté en 1994. Elle vise principalement à observer le développement et le bien-être des enfants au Canada dans leur cheminement de la naissance à l'âge adulte. L'ELNEJ est menée par Statistique Canada et parrainée par Ressources humaines et Développement des compétences Canada.

L'enquête a été conçue en vue de recueillir des renseignements sur les facteurs qui influent sur le développement social et émotionnel ainsi que sur le comportement des enfants et des jeunes. Elle permet en outre de suivre au fil du temps les effets de ces facteurs sur leur développement.

L'enquête porte sur un large éventail de sujets dont la santé, le développement physique, l'apprentissage et le comportement des enfants et comprend des données sur leur environnement social (famille, amis, école et communauté).

Les résultats de l'ELNEJ seront utilisés par différentes personnes provenant de tous les échelons du gouvernement, des universités et des organismes responsables de l'élaboration des politiques.

Population de l'enquête

Au cycle 8, on a suivi un échantillon représentatif d'enfants de 0 à 7 ans de chacune des provinces du Canada afin de produire des estimations longitudinales et transversales. On a suivi la cohorte longitudinale d'enfants et de jeunes de 14 à 25 ans en vue de produire des estimations longitudinales.

Population cible

Les objectifs de l'ELNEJ sont axés sur la production d'estimations tant longitudinales que transversales. Par conséquent, plusieurs populations cibles sont visées par l'échantillon du cycle 8. Veuillez consulter le chapitre 5.0 du présent guide pour obtenir plus de renseignements au sujet de l'échantillon.

- Transversalement, l'échantillon du cycle 8 représente les enfants qui, le 31 décembre 2008, avaient de 0 à 7 ans.
- Longitudinalement, on a quatre cohortes représentant plus d'un cycle de données :
 - 1) Une première cohorte représentant les enfants de 0 à 11 ans au 31 décembre 1994, qui habitaient dans l'une des provinces pendant la collecte du cycle 1 en 1994-1995; ces enfants sont maintenant âgés de 14 à 25 ans au cycle 8.
 - 2) Une deuxième cohorte représentant les enfants de 0 et 1 an au 31 décembre 2002, qui habitaient dans l'une des provinces pendant la collecte du cycle 5 en 2002-2003; ces enfants sont maintenant âgés de 6 et 7 ans au cycle 8.
 - 3) Une troisième cohorte représentant les enfants de 0 et 1 an au 31 décembre 2004, qui habitaient dans l'une des provinces pendant la collecte du cycle 6 en 2004-2005; ces enfants sont maintenant âgés de 4 et 5 ans au cycle 8.
 - 4) Une quatrième cohorte représentant les enfants de 0 et 1 an au 31 décembre 2006, qui habitaient dans l'une des provinces pendant la collecte du cycle 7 en 2006-2007; ces enfants sont maintenant âgés de 2 et 3 ans au cycle 8.

Cycles de collecte

La collecte des données a été faite à intervalles de deux ans.

Cycle	Début de la collecte	Fin de la collecte
1	décembre 1994	avril 1995
2	décembre 1996	avril 1997
3	octobre 1998	juin 1999
4	septembre 2000	mai 2001
5	septembre 2002	juin 2003
6	septembre 2004	juin 2005
7	septembre 2006	juillet 2007
8	septembre 2008	juillet 2009

Stratégie de diffusion des données

Les données du cycle 8 sont diffusées en quatre fichiers : la cohorte longitudinale — enfants (14 à 17 ans); la cohorte longitudinale — jeunes (16 à 25 ans); la cohorte du développement de la petite enfance (0 à 7 ans), et le fichier des questionnaires autoadministrés (14 à 17 ans).

- Cohorte longitudinale — enfants : il s'agit des enfants de la cohorte longitudinale initiale. Ces enfants avaient de 0 à 3 ans au premier cycle de l'ELNEJ. Au cycle 8, ils sont âgés de 14 à 17 ans.
- Cohorte longitudinale — jeunes : ce sont les enfants de la cohorte longitudinale initiale, qui avaient de 2 à 11 ans au cycle 1. Au cycle 8, ces enfants sont âgés de 16 à 25 ans. Le fichier contient des données concernant les jeunes et les ménages dans lesquels ils vivent ainsi que des variables reportées des cycles précédents et qui ne changent pas au fil du temps (le poids à la naissance, par exemple). Veuillez consulter le chapitre 8.0 du présent guide pour obtenir plus de renseignements sur le contenu des divers questionnaires.
- Cohorte du développement de la petite enfance : ce fichier contient les données recueillies auprès de la personne la mieux renseignée (PMR) au sujet de l'enfant, pour les enfants âgés de 0 à 7 ans au cycle 8.
- Le fichier des questionnaires autoadministrés : contient les données recueillies auprès des enfants âgés de 14 à 17 ans au moyen de questionnaires sur format papier. Veuillez consulter le chapitre 8.0 du présent guide pour obtenir de plus amples renseignements au sujet des divers questionnaires.

3.0 Objectifs

Les objectifs de l'Enquête longitudinale nationale sur les enfants et les jeunes (ELNEJ) sont les suivants :

- déterminer la fréquence de divers facteurs de risque et de protection chez les enfants et les jeunes;
- comprendre comment ces facteurs, tout comme les événements de la vie, influent sur le développement de l'enfance;
- rendre cette information disponible afin de permettre l'élaboration de politiques et de programmes destinés à aider les enfants et les jeunes;
- recueillir des renseignements sur une grande diversité de sujets d'ordre biologique, social et économique;
- recueillir des renseignements sur le milieu où grandit l'enfant, qu'il s'agisse de la famille, des pairs, de l'école ou de la collectivité.

Les renseignements proviennent de diverses sources (parents et enfants) ainsi que de mesures directes (Échelle de vocabulaire en images de Peabody [EVIP] et de tests de calcul, entre autres).

4.0 Principaux changements apportés au cycle 8 depuis le cycle 7

Dans le présent chapitre, il est question des principaux changements apportés à l'enquête depuis le cycle 7. Des explications plus précises se trouvent dans les différents chapitres tels que cité ci-dessous.

4.1 Changements apportés au contenu

À chaque cycle, des changements sont apportés au contenu de l'Enquête longitudinale nationale sur les enfants et les jeunes (ELNEJ). Le cinquième caractère du nom d'une nouvelle variable ou d'une variable modifiée (libellé, catégories de réponse et population admissible, entre autres) est un « h ». Les règles d'affectation des noms des variables sont énoncées au chapitre 7.0, et le contenu de l'enquête est décrit en détail au chapitre 8.0.

Voici la liste des principaux changements apportés au contenu de l'enquête au cycle 8.

Livret autoadministré

On a retiré le livret 21, lequel contenait le questionnaire autoadministré des jeunes de 12 et 13 ans. Les plus jeunes répondants de la cohorte longitudinale du cycle 8 sont les 14 et 15 ans.

Engagement politique (20 à 25 ans)

En plus de demander aux jeunes s'ils ont voté lors de précédentes élections, on leur demande s'ils croient que c'est leur devoir d'aller voter et à quelle fréquence ils discutent de l'actualité politique avec les autres (HPOLhQ1 à HPOLhQ7).

Capacités (16 à 25 ans)

Les questions sur les capacités ont été ajoutées au cycle 8, mais apparaissaient précédemment dans les questionnaires papiers autoadministrés du cycle 5 (HABYhQ1 à HABYhQ6).

Formation des familles et fertilité (18 à 25 ans)

Une nouvelle section de questions a été ajoutée pour recueillir l'information sur les facteurs influençant la transition des jeunes adultes vers le mariage et la parentalité (HFFYhQ1 à HFFYhQ16).

4.2 Changements méthodologiques

Au cycle 7, l'âge des enfants de la cohorte du DPE allait de 0 à 9 ans, tandis qu'au cycle 8, l'âge des enfants de la cohorte du DPE s'étend de 0 à 7 ans.

5.0 Méthodologie de l'enquête — Échantillon

Dans le présent chapitre, on fournit des précisions sur l'échantillon de l'Enquête longitudinale nationale sur les enfants et les jeunes (ELNEJ). Dans la section 5.1, on présente un aperçu du plan d'échantillonnage, tandis que dans la section 5.2, on traite de l'Enquête sur la population active (EPA). Dans la section 5.3, on apporte des précisions sur le procédé d'échantillonnage à chaque cycle ainsi que sur les taux de réponse. Puis, dans la section 5.4, le lecteur trouvera de l'information sur la façon dont les différents échantillons de chaque cycle de l'ELNEJ peuvent être utilisés pour faire des inférences sur des populations de référence longitudinales et transversales spécifiques (l'ELNEJ produit des poids d'enquête pour ces populations).

5.1 Aperçu du plan d'échantillonnage — Cycles 1 à 8

L'ELNEJ est une enquête probabiliste conçue dans le but de recueillir, tous les deux ans, des renseignements détaillés sur les facteurs influant sur le développement cognitif, affectif et physique de l'enfant ainsi que pour évaluer l'incidence de ces facteurs au fil du temps.

La collecte des données pour le premier cycle de l'enquête a commencé en 1994; elle comprenait une importante cohorte d'enfants de 0 à 11 ans vivant dans l'une des provinces. Cet échantillon constitue la cohorte initiale, qui est purement longitudinale, en ce sens qu'on ne rajoute aucun complément pour tenir compte des changements qui surviennent dans la population en raison de l'immigration.

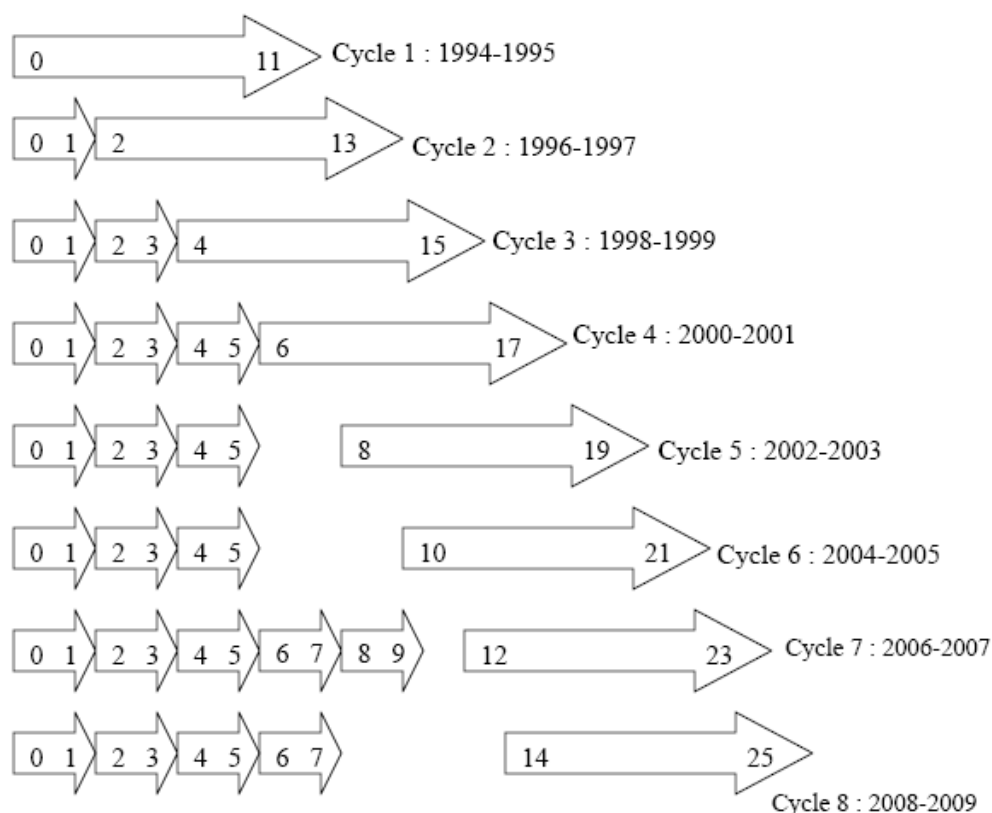
Au deuxième cycle, le champ de l'ELNEJ s'est élargi afin de mettre l'accent sur le développement de la petite enfance (DPE). L'objectif de la composante du DPE est de recueillir des renseignements sur les jeunes enfants et d'établir certains indicateurs (développement moteur et social, troubles émotifs, hyperactivité, agressivité physique, comportement prosocial, compétences linguistiques, entre autres). Une nouvelle cohorte d'enfants de 0 et 1 an est donc sélectionnée à chaque cycle depuis le deuxième cycle et elle est suivie pendant au moins trois cycles (antérieurement au cycle 7, les enfants de la cohorte du DPE étaient suivis seulement jusqu'à l'âge de 4 et 5 ans, mais au cycle 8, l'âge des enfants du DPE variait de 0 à 7 ans). Ces échantillons servent à des fins longitudinales et transversales et sont désignés « cohortes du DPE ».

En règle générale, les enfants de l'ELNEJ sont sélectionnés dans les ménages de l'échantillon de l'EPA de Statistique Canada. Font exception les enfants ayant 0 et 1 an au cycle 3 et certains enfants de 5 ans aux cycles 3 et 4, qui ont été sélectionnés à partir des données du registre des naissances.

Au cycle 8, l'échantillon de l'enquête comprend des enfants de 0 à 7 ans (cohorte du DPE) et de 14 à 25 ans (cohorte initiale). L'âge effectif des enfants au cycle 8 est établi par rapport au 31 décembre 2008. Les enfants de 0 année sont donc nés en 2008 et ceux de 1 an, en 2007.

Le diagramme ci-dessous illustre bien l'échantillon de l'ELNEJ. L'année renvoie à la collecte. Les flèches plus longues représentent la cohorte initiale et les plus petites, les cohortes du DPE.

Diagramme 1
Âge des enfants à chaque cycle, cohorte initiale et cohortes du DPE



Notes : L'âge des enfants exprimé en années est indiqué par les flèches.

Les flèches plus longues représentent la cohorte initiale et les plus petites, les cohortes du Développement de la petite enfance (DPE).

Veuillez consulter la section 5.3 du présent guide pour obtenir plus de renseignements sur le procédé d'échantillonnage utilisé à chaque cycle.

Source : Statistique Canada, Enquête longitudinale nationale sur les enfants et les jeunes.

5.2 Enquête sur la population active

L'Enquête sur la population active (EPA) est une enquête mensuelle qui vise à recueillir des données sur le marché du travail auprès d'un échantillon national de plus de 52 000 ménages (à partir du remaniement de 2004). L'EPA est remaniée après chaque recensement décennal de la population afin de tenir compte des changements survenus dans la population du Canada et de répondre aux besoins changeants en information de l'EPA. Depuis le tout début de l'ELNEJ, il y a eu deux remaniements de l'EPA, le premier a eu lieu en 1994 et le deuxième, en 2004.

Il s'ensuit que l'échantillon de la cohorte initiale est formé de différentes personnes faisant partie autant de la cohorte qui précède le remaniement de l'EPA de 1994 que de celle qui le suit. Une part des enfants de la cohorte du DPE enquêtés au cycle 8 proviennent du plan de l'EPA de 1994. Les cohortes du DPE échantillonnées au cycle 5 étaient basées sur le plan de 1994 et la cohorte du cycle 6 était formée de différentes personnes provenant du plan de 1994 et de 2004. Les cohortes du DPE des cycles 7 et 8 sont basées sur le plan de 2004. Dans les sections suivantes, on trouvera des renseignements détaillés sur les remaniements de l'EPA de 1994 et de 2004.

5.2.1 Population cible

L'échantillon de l'EPA (1994) est représentatif de la population civile de 15 ans et plus ne vivant pas dans un établissement et résidant dans l'une des dix provinces du Canada. Sont spécifiquement exclus du champ de l'enquête les résidents du Yukon, des Territoires du Nord-Ouest et du Nunavut, les personnes vivant sur les réserves indiennes, les membres à temps plein des Forces armées canadiennes et les pensionnaires d'établissements. Ces groupes représentent en tout environ 2 % de la population exclue de 15 ans et plus.

5.2.2 Stratification

Le plan de sondage de l'EPA est basé sur un plan stratifié à plusieurs degrés faisant appel à un échantillonnage probabiliste à toutes les étapes du plan. Les principes du plan de l'EPA sont les mêmes pour toutes les provinces. La stratification du remaniement de 1994 est comme suit :

Stratification primaire

Les provinces sont divisées en régions économiques (RÉ) et en régions économiques de l'assurance-emploi (RÉAE). Les régions économiques sont des régions géographiques de structure économique plus ou moins homogène constituées en vertu d'ententes fédérales-provinciales et qui sont relativement stables au fil du temps. Les régions économiques de l'assurance-emploi sont également des régions géographiques qui correspondent à peu de choses près aux régions économiques quant à la taille et au nombre, mais elles diffèrent pour ce qui est de la définition. Des estimations de la population active sont produites pour les RÉAE, lesquelles sont utilisées par Ressources humaines et Développement des compétences Canada.

Les intersections de ces deux types de régions constituent le premier niveau de stratification de l'EPA. Ces intersections sont considérées comme des strates primaires et, à l'intérieur de chacune d'elles, on procède à une stratification plus poussée (voir la section 5.2.3). Il est bon de noter qu'une troisième série de régions, les régions métropolitaines de recensement (RMR), est également présentée en strates dans le plan actuel de l'EPA, car chaque RMR est aussi une RÉAE.

Types d'unités

Les intersections de strates primaires (RÉ et RÉAE) sont ensuite subdivisées en trois types d'unités : les régions rurales, les régions urbaines et les régions éloignées. Les régions urbaines et rurales se fondent en général sur la définition des régions urbaines et rurales aux fins du recensement. Il y a quelques exceptions afin de permettre la formation de strates dans certaines régions. Pour le Recensement de 1991, les régions urbaines englobaient les plus grandes RMR jusqu'aux plus petits villages classés comme régions urbaines (1 000 personnes et plus), tandis que les régions rurales se composent de régions non désignées urbaines ou éloignées.

Toutes les régions urbaines sont ensuite divisées en deux catégories : celles qui utilisent une liste d'appartements et une base aréolaire, puis celles qui n'ont recours qu'à une base aréolaire.

Environ 1 % de la population de l'EPA demeure dans des régions éloignées dont les intervieweurs de l'EPA ne peuvent pas atteindre facilement. À des fins administratives, ce segment de la population est échantillonné séparément à l'aide de la base des régions éloignées. Les localités de moins de 10 ménages ou de 25 personnes ainsi que les secteurs de dénombrement du recensement (SD) de moins de 25 ménages ne sont pas considérés par le plan d'échantillonnage.

Stratification secondaire

Dans les régions urbaines où le nombre d'immeubles d'appartements est suffisamment important, les strates sont subdivisées en listes d'appartements et en bases aréolaires. Une

liste d'appartements est un registre tenu à jour dans les 18 plus grands centres du Canada. Cela vise à assurer une meilleure représentation des résidents d'appartements à l'intérieur de l'échantillon et à réduire l'effet de croissance attribuable à la construction de nouveaux immeubles d'appartements dans les grappes. Dans les principaux centres, les strates d'appartements sont, en outre, subdivisées en strates de faibles revenus et en strates ordinaires.

Lorsque cela est possible et nécessaire, la base aréolaire urbaine est subdivisée en strates ordinaires, en strates de revenus élevés et en strates de faible densité de population. La plupart des régions urbaines sont classées parmi les strates urbaines ordinaires qui, en fait, couvrent la majorité de la population du Canada. Les strates de revenus élevés sont concentrées dans les grandes régions urbaines, tandis que les strates urbaines de faible densité se composent de petites localités éparpillées.

Dans les régions rurales, la densité de la population peut varier fortement de relativement élevée à faible, entraînant ainsi la formation de strates qui reflètent ces variations. Les diverses stratégies de stratification des régions rurales sont fondées non seulement sur la concentration de la population, mais aussi sur les coûts et les contraintes imposées aux intervieweurs.

La base aréolaire des régions éloignées est stratifiée uniquement par province.

5.2.3 Délimitation et sélection des grappes

Les ménages des strates finales ne sont pas sélectionnés automatiquement. Chaque strate est plutôt divisée en grappes, puis un échantillon de grappes est sélectionné dans la strate. Les logements sont ensuite extraits à partir des grappes sélectionnées. Selon le type de strate, on utilise diverses méthodes pour définir les grappes.

Dans chaque strate urbaine se trouvant dans une base aréolaire urbaine, un certain nombre de groupes de logements géographiquement contigus, ou grappes, sont constitués à partir d'une combinaison de dénombrements du recensement. Ces grappes coïncident généralement avec des îlots urbains ou des côtés d'îlots. La sélection d'un échantillon de grappes (toujours six grappes ou un multiple de six) à partir de chacune de ces strates secondaires représente le premier degré de l'échantillonnage dans la plupart des régions urbaines. Dans certaines autres régions urbaines, les secteurs de dénombrements (SD) sont utilisés comme grappes. Dans les strates urbaines de faible densité, on applique un plan à trois étapes en vertu duquel deux villes faisant partie d'une strate sont échantillonnées, puis 6 ou 24 grappes de chaque ville sont échantillonnées.

Dans les strates d'appartements urbains, au lieu de définir des grappes, on utilise l'immeuble d'appartements comme unité primaire d'échantillonnage. Les immeubles d'appartements sont échantillonnés à partir de la liste, avec une probabilité proportionnelle au nombre d'appartements dans chaque immeuble.

On effectue d'autres procédures dans les régions rurales et éloignées. Dans chacune des strates rurales, six SD, ou deux ou trois groupes de SD, sont échantillonnés en tant que grappes, tandis que dans les strates éloignées le nombre d'agglomérations échantillonnées est proportionnel au nombre de logements de l'agglomération.

5.2.4 Sélection des logements

Dans les trois types de régions (régions urbaines, rurales et éloignées), les intervieweurs commencent par visiter les grappes sélectionnées, puis ils dressent une liste de tous les logements privés de chaque grappe. Un échantillon de logements est ensuite sélectionné à partir de cette liste. La taille de l'échantillon dépend du type de strate. Par exemple, dans la base aréolaire urbaine, les tailles d'échantillons sont de six ou huit logements, dépendant de la taille de la ville. Dans la liste d'appartements urbains, cinq logements par grappe sont sélectionnés alors que dans les régions rurales et les parties des SD des villes, dix logements

sont sélectionnés par grappe. Dans toutes les grappes, les logements sont sélectionnés selon un plan systématique, ce qui représente le dernier degré d'échantillonnage.

5.2.5 Changements introduits dans le cadre du remaniement de 2004

Le remaniement de 2004 a été préparé dans le contexte d'un cadre budgétaire restreint. Contrairement au remaniement de 1994, le questionnaire et l'application de collecte n'ont pas été modifiés. Le nouveau plan d'échantillonnage a été progressivement introduit à partir de novembre 2004.

Afin de réduire les coûts de fonctionnement de l'enquête, deux principaux changements ont été apportés à la méthodologie. Avant novembre 2004, la première interview sur six était menée en personne. Dans le but de réduire les coûts de collecte, la première interview est désormais effectuée par téléphone pour certains logements dans les régions urbaines. Le second changement porte sur la réduction des coûts de listage (le listage consiste à compiler une liste d'adresses résidentielles dans une certaine UPE).

Afin de réduire le coût du listage et d'améliorer la couverture de la base de sondage, le registre des adresses (RA) a été utilisé. Le RA est une base de données contenant les adresses des logements dans les centres urbains. En 2005, le RA comportait environ 13 millions d'adresses. Le Recensement de 2001 a permis d'indiquer que la majorité de ces adresses étaient des logements résidentiels valides. D'autres adresses du RA sont obtenues grâce aux mises à jour des fichiers administratifs.

Outre ces deux principaux changements, d'autres améliorations ont été apportées au plan d'échantillonnage. Dans le passé, la base de données géographiques de Statistique Canada ne couvrait pas entièrement le territoire des dix provinces. Cette base de données est nécessaire pour établir les limites des UPE. La qualité de la base de données géographiques s'est considérablement améliorée depuis le remaniement de 1994. Pour la première fois, en 2004, l'EPA a été en mesure de définir les limites des UPE pour le territoire entier des dix provinces.

Parmi les autres améliorations, il convient de mentionner que, pour mieux contrôler la répartition de l'échantillon et, par conséquent, les coûts de collecte, une nouvelle stratégie a été mise en œuvre dans les régions où les coûts de collecte sont élevés. L'EPA a également introduit des méthodes pour cibler la population des immigrants dans les grands centres et la population autochtone dans les quatre provinces de l'Ouest. Afin de réduire les coûts associés à la tenue à jour de la base de sondage, le nouveau plan d'échantillonnage ne contient plus de base de sondage des appartements. Enfin, l'échantillon des petites régions rurales est désormais sélectionné au moyen d'un plan à deux degrés au lieu du plan à trois degrés utilisé auparavant.

5.2.6 Renouvellement de l'échantillon

L'EPA utilise un plan de sondage avec un renouvellement de panel dans lequel les ménages sélectionnés restent dans l'échantillon pendant six mois consécutifs. L'échantillon mensuel complet est formé de six sous-échantillons ou panels de taille semblable, chacun étant représentatif de la population totale de l'EPA. Chaque mois, on remplace le panel qui fait partie de l'échantillon depuis six mois. Les ménages qui cessent de faire partie de l'échantillon sont remplacés par d'autres ménages du même secteur ou d'un secteur comparable.

Sur le plan statistique, ce mécanisme de renouvellement présente l'avantage de fournir une base d'échantillonnage commune qui permet de faire des comparaisons des caractéristiques de l'EPA d'un mois à l'autre. Il permet en outre de s'assurer que l'échantillon de ménages reflète en tout temps l'état actuel du parc immobilier et contribue à atténuer les problèmes de non-réponse ou de fardeau du répondant qui se produiraient si les ménages devaient

demeurer plus de six mois dans l'échantillon. Les enquêtes qui font usage de la base de sondage ou de l'échantillon de l'EPA peuvent profiter des caractéristiques liées aux groupes de renouvellement en utilisant des tailles d'échantillon plus grandes ou plus petites que les échantillons de l'EPA.

5.2.7 Membres des ménages admissibles aux fins de l'Enquête sur la population active

Le premier mois où un logement participe à l'EPA, on dresse une liste qui contient de l'information sur la composition du ménage. On recueille des renseignements démographiques comme le nom, le sexe, la date de naissance et le niveau de scolarité de toutes les personnes qui ont leur lieu de résidence habituel dans le logement sélectionné. La collecte des renseignements aux fins de la population active vise tous les membres civils du ménage âgés de 15 ans et plus.

Quand on communique avec les occupants du logement les mois suivants, on met la liste à jour en fonction des changements survenus dans la composition du ménage par rapport au mois précédent.

5.3 Précisions concernant l'échantillon de l'Enquête longitudinale nationale sur les enfants et les jeunes

Au cycle 8, l'échantillon de l'ELNEJ se compose comme suit :

- une nouvelle cohorte du DPE formée d'enfants de 0 et 1 an sélectionnés au cycle 8;
- un échantillon complémentaire de nouveaux enfants de 2 à 5 ans sélectionnés au cycle 8.

Ainsi que le retour des enfants des échantillons précédents suivants :

- le retour des enfants de 2 et 3 ans faisant partie de la cohorte du DPE d'enfants de 0 et 1 an sélectionnés au cycle 7;
- le retour des enfants de 4 et 5 ans faisant partie de la cohorte du DPE d'enfants de 0 et 1 an sélectionnés au cycle 6;
- le retour des enfants de 4 à 7 ans faisant partie de l'échantillon complémentaire d'enfants de 2 à 5 ans sélectionnés au cycle 7;
- le retour des enfants de 6 à 7 ans faisant partie de l'échantillon complémentaire d'enfants de 2 à 5 ans sélectionnés au cycle 6;
- le retour des enfants de 6 et 7 ans faisant partie de la cohorte du DPE d'enfants de 0 et 1 an sélectionnés au cycle 5;
- le retour des enfants de 14 à 25 ans faisant partie de la cohorte initiale d'enfants de 0 à 11 ans sélectionnés au cycle 1.

Tous les enfants étaient échantillonnés de l'EPA. Il est à noter qu'il n'y avait aucun enfant de 8 à 13 ans au cycle 8.

Outre l'explication sur la façon dont l'échantillonnage a été sélectionné, la présente section donne des précisions sur les enfants qui ont été interviewés. Certains enfants sont échantillonnés, mais ne sont pas interviewés parce qu'ils étaient hors du champ de l'enquête sur le plan transversal au cycle précédent c'est-à-dire qu'ils étaient décédés, qu'ils étaient à l'extérieur du pays ou qu'ils étaient non-répondants pour de trop nombreux cycles.

Pour obtenir plus de détails sur la façon dont l'échantillonnage a été sélectionné à chaque cycle, veuillez consulter le Guide de l'utilisateur du cycle en question.

5.3.1 Cohorte initiale — Cycles 1 à 8

Les paragraphes suivants décrivent la composition de la cohorte initiale à chaque cycle (voir aussi le diagramme 1). La cohorte initiale comprend tout au plus deux enfants par ménage.

Cycle 1

L'échantillonnage d'enfants prélevé au cycle 1 a été sélectionné dans le but de produire des estimations provinciales fiables (mais avec précision variant par province) pour les enfants de 0 à 11 ans, par groupe d'âge de deux ans (0 et 1 an, 2 et 3 ans, 4 et 5 ans, 6 et 7 ans, 8 et 9 ans et 10 et 11 ans). On a sélectionné tout au plus quatre enfants par ménage, et ces ménages eux-mêmes ont été sélectionnés à même les sources suivantes :

- l'ancienne EPA (avant 1994);
- l'EPA remaniée de 1994;
- l'Enquête nationale sur la santé de la population (ENSP) (menée par Statistique Canada).

À la fin du cycle 1, on comptait 22 831 enfants répondants dans l'ELNEJ, et le taux de réponse des enfants se situait à 86,5 % (voir le tableau 1).

Cycle 2

Pour des raisons d'ordre budgétaire, on a retiré certains enfants de l'échantillon au deuxième cycle. Tous les ménages prélevés dans l'échantillon de l'ENSP ont été supprimés de l'échantillon de l'ELNEJ et, pour réduire le fardeau de réponse des ménages, le nombre maximum d'enfants sélectionnés par ménage a été ramené de quatre à deux. Cette diminution donnait un échantillon de 16 903 enfants au début du cycle 2. Le taux de réponse des enfants s'établissait à 91,5 % (voir le tableau 1). Le taux de réponse longitudinal cumulé des enfants de la cohorte initiale était de 79,1 % (voir le tableau 2).

Cycle 3

Au cycle 3, on a retiré 185 enfants de l'échantillon parce qu'ils étaient hors du champ de l'enquête sur le plan transversal (71) ou qu'ils refusaient catégoriquement de répondre (114) à la fin du cycle 2. Les enfants qui étaient hors du champ de l'enquête sur le plan transversal sont ceux qui sont décédés, ceux qui n'avaient pas l'âge requis, ceux qui avaient quitté le pays de façon permanente ou ceux qui s'étaient installés dans une réserve indienne. Ainsi, sur les 16 903 enfants échantillonnés pour la cohorte initiale, 16 718 ont fourni des renseignements. Le taux de réponse des enfants s'élevait à 89,2 % (voir le tableau 1), et le taux de réponse longitudinal cumulé des enfants de la cohorte initiale se situait à 76,0 % (voir le tableau 2).

Cycle 4

Au cycle 4, afin d'accroître l'efficacité du processus de collecte, on a décidé de retirer de l'échantillon les ménages qui n'avaient pas répondu pendant deux cycles consécutifs ou plus (ainsi que les ménages comptant un cycle de non-réponse puis ayant le code « déménagé temporairement »). Donc, on a retiré 1 086 enfants de l'échantillon au cycle 4 parce que, à la fin du cycle 3, ils étaient hors du champ de l'enquête sur le plan transversal (106) ou ils n'avaient pas répondu pendant deux cycles ou plus (980). Au total, 15 632 enfants ont été interviewés. Le taux de réponse des enfants se situait à 84,5 % (voir le tableau 1), et le taux de réponse longitudinal cumulé des enfants de la cohorte initiale s'établissait à 67,8 % (voir le tableau 2).

Cycle 5

Au cycle 5, on a décidé que les jeunes de 18 et 19 ans ne seraient retirés de l'échantillon qu'après trois cycles consécutifs de non-réponse (par rapport à deux pour les plus jeunes). En effet, à 18 ans, le jeune devient le seul répondant, tandis que, avant cet âge, le principal répondant est la personne la mieux renseignée (PMR), habituellement la mère.

Ainsi, 469 jeunes ont été retirés de l'échantillon au cycle 5 parce que, à la fin du cycle 4, ils étaient hors du champ de l'enquête sur le plan transversal (32) ou ils n'avaient pas répondu pendant de trop nombreux cycles consécutifs (437). Au total, on a suivi 15 163 jeunes au cycle 5. Le taux de réponse des enfants se situait à 81,3 % (voir le tableau 1), et le taux de réponse longitudinal cumulatif des enfants de la cohorte initiale s'établissait à 63,1 % (voir le tableau 2).

Cycle 6

Au cycle 6, on a retiré 1 506 enfants de l'échantillon parce que, à la fin du cycle 5, ils n'avaient pas répondu pendant de trop nombreux cycles consécutifs. Au total, 13 657 enfants ont été interviewés au cycle 6. Le taux de réponse des enfants se situait à 82,4 % (voir le tableau 1), et le taux de réponse longitudinal cumulatif des enfants de la cohorte initiale s'établissait à 57,6 % (voir le tableau 2).

Cycle 7

Au cycle 7, 613 jeunes ont été retirés de l'échantillon parce que, à la fin du cycle 6, ils étaient hors du champ de l'enquête sur le plan transversal (11) ou ils n'avaient pas répondu pendant de trop nombreux cycles consécutifs (602). Au total, on a suivi 13 709 enfants au cycle 7. Le taux de réponse des enfants se situait à 80,5 % (voir le tableau 1), et le taux de réponse longitudinal cumulatif des enfants de la cohorte initiale s'établissait à 56,6 % (voir le tableau 2).

Il importe de noter qu'au cycle 7 une nouvelle consigne s'appliquait aux jeunes âgés de 18 ans et plus provenant des cycles précédents : on ignorait les antécédents de non-réponse de la PMR lors de la prise de décision si le jeune était interviewé ou pas. Ce qui veut dire que certains enfants qui avaient été retirés de l'échantillon dans les cycles précédents, ont été interviewés lors du cycle 7.

Cycle 8

Au cycle 8, 56 enfants ont été retirés de l'échantillon parce que, à la fin du cycle 7, ils étaient soit hors du champ de l'enquête sur le plan transversal (17) ou n'avaient pas répondu pendant de trop nombreux cycles consécutifs (39). Au total, 15 056 enfants ont été interviewés au cycle 8. Le taux de réponse des enfants se situait à 68,0 % (voir le tableau 1). Le taux de réponse longitudinal cumulatif des enfants de la cohorte initiale s'établissait à 52,7 % (voir le tableau 2).

Il importe de noter que la consigne qui s'appliquait au cycle 7 aux jeunes âgés de 18 ans et plus provenant des cycles précédents, s'appliquait également au cycle 8. Dans le cas des jeunes âgés de 14 à 17 ans, ceux qui avaient répondu pour la dernière fois au cycle 5, 6 ou 7 ont été interviewés au cycle 8.

5.3.2 Participation des cohortes du Développement de la petite enfance au cycle 8

Les enfants de la cohorte du DPE du cycle 8, avaient été échantillonnés pour la première fois aux cycles 5, 6, 7 et 8. Lorsque la première cohorte d'enfants de 0 et 1 an a été sélectionnée au cycle 2, la règle imposait un maximum d'un enfant par ménage, sauf pour ce qui est des jumeaux, auquel cas les deux étaient échantillonnés¹. Au cycle 5, un seul enfant par ménage sans exception était échantillonné². Au cycle 8, il n'y avait pas de jumeaux dans les cohortes du DPE provenant de cycles précédents.

-
1. La cohorte du DPE échantillonnée au cycle 2 incluait des enfants de 0 et 1 an qui étaient les jeunes frères et sœurs des enfants de la cohorte initiale. Il s'agissait du seul cycle où des frères et sœurs de la cohorte initiale avaient été sélectionnés. Aucun enfant de la cohorte du DPE du cycle 2 n'est présent dans l'échantillon du cycle 8.
 2. Pour l'échantillon du cycle 7, il a été décidé que l'un des jumeaux déjà présents serait retiré de l'échantillon pour les cohortes existantes du DPE. La cohorte initiale continue de compter tout au plus deux enfants par ménage.

Avant le cycle 7, pour les échantillons du DPE, seuls les répondants faisant partie du cycle précédent étaient interviewés aux cycles subséquents. Au cycle 7, cette consigne a été modifiée afin que les non-répondants faisant partie des cycles précédents soient interviewés, à condition qu'ils n'aient pas été des non-répondants pendant au moins 2 cycles consécutifs. Au cycle 8, cette condition a été abandonnée afin que tous les non-répondants faisant partie des cycles précédents soient interviewés, peu importe le nombre de cycles consécutifs auxquels ils n'ont pas répondu.

Cohorte du DPE au cycle 5

Au cycle 5, on a sélectionné un échantillon d'enfants de 0 et 1 an dans l'EPA. La taille totale de l'échantillon était de 4 492 enfants et ménages. À la fin de la collecte du cycle 5, on comptait 3 252 enfants répondants, et le taux de réponse se situait à 74,0 % (voir le tableau 1).

À la fin du cycle 5, 98 enfants étaient inadmissibles sur le plan transversal et 1 142 étaient des non-répondants. Par conséquent, un total de 3 252 répondants du cycle 5 ont été interviewés au cycle 6 alors qu'ils étaient âgés de 2 et 3 ans. Le taux de réponse était de 88,6 % (voir le tableau 1), et le taux de réponse longitudinal cumulatif se situait à 65,3% (voir le tableau 2).

À la fin du cycle 6, 4 enfants étaient inadmissibles sur le plan transversal et 33 étaient des non-répondants. Par conséquent, un total de 3 215 répondants d'un cycle antérieur ont été interviewés au cycle 7 alors qu'ils étaient âgés de 4 et 5 ans. Le taux de réponse était de 85,9 % (voir le tableau 1), et le taux de réponse longitudinal cumulatif se situait à 62,4 % (voir le tableau 2).

Finalement, un total de 3 214 répondants d'un cycle antérieur ont été interviewés au cycle 8 alors qu'ils étaient âgés de 6 et 7 ans. Le taux de réponse était de 80,5 % (voir le tableau 1), et le taux de réponse longitudinal cumulatif se situait à 58,8% (voir le tableau 2).

Cohorte du DPE au cycle 6

Au cycle 6, un échantillon d'enfants de 0 à 5 ans a été sélectionné dans l'EPA. La taille totale de l'échantillon était de 5 795 enfants et ménages. À la fin de la collecte du cycle 6, on comptait 4 684 enfants répondants, et le taux de réponse se situait à 81,3 % (voir le tableau 1).

À la fin du cycle 6, 21 enfants étaient inadmissibles sur le plan transversal et 143 étaient des non-répondants. Par conséquent, un total de 5 631 enfants âgés de 2 à 7 ans ont été interviewés au cycle 7. Le taux de réponse était de 83,0 % (voir le tableau 1), et le taux de réponse longitudinal cumulatif se situait à 79,7 % (voir le tableau 2).

Finalement, un total de 5 039 enfants âgés de 4 à 7 ans ont été interviewés au cycle 8. Le taux de réponse était de 76,9 % (voir le tableau 1), et le taux de réponse longitudinal cumulatif se situait à 75,1 % (voir le tableau 2).

Cohorte du DPE au cycle 7

Au cycle 7, un échantillon d'enfants de 0 à 5 ans a été sélectionné dans l'EPA. La taille totale de l'échantillon était de 5 843 enfants et ménages. À la fin de la collecte du cycle 7, on comptait 4 691 enfants répondants, et le taux de réponse se situait à 80,8 % (voir le tableau 1).

À la fin du cycle 7, 35 enfants étaient inadmissibles sur le plan transversal et 11 étaient des non-répondants. Par conséquent, un total de 5 797 enfants âgés de 2 à 7 ans ont été interviewés au cycle 8. Le taux de réponse était de 79,1 % (voir le tableau 1), et le taux de réponse longitudinal cumulatif se situait à 78,2 % (voir le tableau 2).

Cohorte du DPE du cycle 8

Au cycle 8, un échantillon d'enfants de 0 à 5 ans a été sélectionné dans l'EPA. La taille totale de l'échantillon était de 6 685 enfants et ménages. À la fin de la collecte du cycle 8, on comptait 5 065 enfants répondants, et le taux de réponse se situait à 76,0 % (voir le tableau 1).

Tableau 1 Réponse des enfants lors de la collecte

Type d'échantillon	Cohorte d'échantillonnage	Cycle de l'enquête	Groupe d'âge en années	Taille de l'échantillon		Réduction de l'échantillon		Hors du champ (retiré du cycle précédent)		Non-répondant (retiré du cycle précédent)		Admissible		Répondant		Taux d'admissibilité (%)	Taux de réponse (%)
				ménage	enfant	ménage	enfant	ménage	enfant	ménage	enfant	ménage	enfant	ménage	enfant	ménage	enfant
Cohorte initiale de 1994	1	1	0 à 11	43 751	15 502	26 409	13 439	22 831	35,4	86,5
		2	2 à 13	11 188	16 903	25 588	..	5 345	..	1 677	..	11 140	16 816	10 216	15 391	99,6	91,5
		3	4 à 15	11 032	16 718	0	0	38	71	73	114	10 937	16 563	9 801	14 777	99,1	89,2
		4	6 à 17	10 449	15 632	0	0	65	106	618	980	10 418	15 588	8 834	13 176	99,7	84,5
		5	8 à 19	10 355	15 163	0	0	24	32	286	437	10 320	15 113	8 582	12 280	99,7	81,3
		6	10 à 21	9 881	13 657	0	0	0	0	878	1 506	9 816	13 572	8 201	11 178	99,3	82,4
		7	12 à 23	10 522	13 709	0	0	7	11	406	602	10 454	13 616	8 561	10 966	99,4	80,5
		8	14 à 25	12 021	15 056	0	0	13	17	28	39	11 981	15 007	8 468	10 208	99,7	68,0
DPE de 1996 plus échan. additionnel du N.-B.	2	2	0 à 1	5 592	4 929	5 087	4 496	4 634	88,1	91,1
		3	2 à 3	3 992	4 046	558	..	598	..	444	..	3 950	4 004	3 592	3 640	98,9	90,9
		4	4 à 5	3 577	3 610	520	540	25	25	34	35	3 552	3 585	3 023	3 052	99,3	85,1
DPE de 1998 plus échan. complémentaire de 5	3	3	0 à 1 et 5	16 812	15 929	16 263	13 256	13 546	94,7	83,3
		4	2 à 3	7 941	8 118	6 935	..	516	..	1 420	..	7 896	8 070	6 956	7 111	99,4	88,1
		5	4 à 5	6 960	7 115	0	22	41	41	940	940	6 919	7 073	6 208	6 340	99,4	89,6
		7	8 à 9	6 016	6 016	5 988	5 988	5 321	5 321	99,5	88,9
DPE de 2000 plus échan. complémentaire de 5	4	4	0 à 1 et 5	9 439	9 116	9 192	6 908	6 961	96,6	75,7
		5	2 à 3	3 788	3 841	4 405	..	125	..	1 121	..	3 776	3 829	3 281	3 324	99,7	86,8
		6	4 à 5	3 280	3 323	0	10	11	11	497	497	3 270	3 313	2 931	2 964	99,7	89,5
		7	6 à 7	3 231	3 231	0	43	3	3	46	46	3 217	3 217	2 882	2 882	99,5	89,6
DPE de 2002	5	5	0 à 1	4 492	4 492	4 394	4 394	3 252	3 252	97,8	74,0
		6	2 à 3	3 252	3 252	0	0	98	98	1 142	1 142	3 233	3 233	2 866	2 866	99,4	88,6
		7	4 à 5	3 215	3 215	0	0	4	4	33	33	3 189	3 189	2 740	2 740	99,2	85,9
		8	6 à 7	3 214	3 214	0	0	26	26	3	3	3 205	3 205	2 580	2 580	99,7	80,5
DPE de 2004 plus échan. complémentaire de 2 à 5	6	6	0 à 5	5 795	5 795	5 763	5 763	4 684	4 684	99,4	81,3
		7	2 à 7	5 631	5 631	0	0	21	21	143	143	5 600	5 600	4 650	4 650	99,4	83,0
		8	4 à 7	5 039	5 039	0	0	31	31	688	688	5 006	5 006	3 852	3 852	99,3	76,9
DPE de 2006 plus échan. complémentaire de 2 à 5	7	7	0 à 5	5 843	5 843	5 808	5 808	4 691	4 691	99,3	80,8
		8	2 à 7	5 797	5 797	0	0	35	35	11	11	5 769	5 769	4 561	4 561	99,5	79,1
DPE de 2008 plus échan. complémentaire de 2 à 5	8	8	0 à 5	6 685	6 685	6 666	6 666	5 065	5 065	99,7	76,0

Notes :

.. indisponible

Enfant réfère à « enfants répondants ».

Source : Statistique Canada, Enquête longitudinale nationale sur les enfants et les jeunes.

Tableau 2 Réponse longitudinale pour ce qui est des enfants

Longitudinale											
Type d'échantillon	Cohorte d'échantillonnage	Cycle de l'enquête	Groupe d'âge en années	Taille de l'échantillon		Admissible		Répondant		Taux d'admissibilité (%)	Taux de réponse (%)
				ménage	enfant	ménage	enfant	ménage	enfant	ménage	enfant
Cohorte initiale de 1994	1	1	0 à 11	18 163	..	12 818	19 487	11 141	16 903	70,6	86,7
		2	2 à 13	18 210	..	12 863	19 481	10 220	15 403	70,6	79,1
		3	4 à 15	18 165	..	12 818	19 481	9 810	14 796	70,6	76,0
		4	6 à 17	18 265	..	12 912	19 435	8 839	13 168	70,7	67,8
		5	8 à 19	18 481	..	13 134	19 481	8 592	12 300	71,1	63,1
		6	10 à 21	18 885	..	13 532	19 474	8 222	11 210	71,7	57,6
		7	12 à 23	19 402	..	14 409	19 474	8 597	11 016	72,4	56,6
		8	14 à 25	20 043	..	14 690	19 474	8 510	10 268	73,3	52,7
Cohorte du DPE de 1996	2	2	0 à 1	5 271	..	4 673	4 733	4 100	4 154	88,7	87,8
		3	2 à 3	5 271	..	4 671	4 731	3 595	3 643	88,6	77,0
		4	4 à 5	4 620	..	3 846	3 880	2 899	2 928	83,2	75,5
Cohorte du DPE de 1998	3	3	0 à 1	9 877	..	9 361	9 559	7 949	8 126	94,8	85,0
		4	2 à 3	9 877	..	9 141	9 334	6 794	6 946	92,5	74,4
		5	4 à 5	9 877	..	9 141	9 334	6 060	6 189	92,5	66,3
		7	8 à 9	9 877	..	9 141	9 141	5 325	5 325	92,5	58,3
Cohorte du DPE de 2000	4	4	0 à 1	5 034	..	4 909	4 985	3 788	3 841	97,5	77,1
		5	2 à 3	5 034	..	4 907	4 983	3 279	3 322	97,5	66,7
		6	4 à 5	5 034	..	4 911	4 987	2 932	2 965	97,6	59,5
		7	6 à 7	5 034	..	4 911	4 911	2 885	2 885	97,6	58,7
Cohorte du DPE de 2002	5	5	0 à 1	4 492	4 492	4 394	4 394	3 252	3 252	97,8	74,0
		6	2 à 3	4 492	4 492	4 392	4 392	2 867	2 867	97,8	65,3
		7	4 à 5	4 492	4 492	4 392	4 392	2 741	2 741	97,8	62,4
		8	6 à 7	4 492	4 492	4 392	4 392	2 582	2 582	97,8	58,8
DPE de 2004 plus échan. complémentaire de 2 à 5	6	6	0 à 1	4 356	4 356	4 343	4 343	3 521	3 521	99,7	81,1
		7	2 à 3	4 356	4 356	4 343	4 343	3 463	3 463	99,7	79,7
		8	4 à 5	4 356	4 356	4 343	4 343	3 263	3 263	99,7	75,1
DPE de 2006 plus échan. complémentaire de 2 à 5	7	7	0 à 1	4 997	4 997	4 975	4 975	4 015	4 015	99,6	80,7
		8	2 à 3	4 997	4 997	4 973	4 973	3 888	3 888	99,5	78,2
DPE de 2008 plus échan. complémentaire de 2 à 5	8	8	0 à 1	5 482	5 482	5 464	5 464	4 107	4 107	99,7	75,2

Notes :

Ce tableau comprend des données pour seulement les enfants qui font partie de l'étude longitudinale (c.-à-d. les enfants qui font l'objet d'un suivi dans tous les cycles).

.. indisponible

Enfant réfère à « enfants répondants ».

Source : Statistique Canada, Enquête longitudinale nationale sur les enfants et les jeunes.

5.4 Populations de référence longitudinale et transversale pour l'établissement des poids de l'Enquête longitudinale nationale sur les enfants et les jeunes

Dans une enquête probabiliste, les individus sont sélectionnés de façon aléatoire à même une population bien définie, de sorte qu'une probabilité non nulle de sélection s'applique à chaque individu (c'est-à-dire que tout le monde peut être sélectionné et que personne n'est exclu) et que cette probabilité peut être calculée. Par exemple, s'il y a 100 enfants dans la population et que 10 d'entre eux sont sélectionnés selon une technique d'échantillonnage aléatoire simple, la probabilité que chaque enfant soit sélectionné s'établit ainsi : $10/100=1/10$.

Le poids qu'a l'enfant dans l'enquête correspond au nombre moyen d'enfants dans la population qu'il représente. Ce poids est calculé comme l'inverse de la probabilité de sélection (des ajustements ultérieurs sont habituellement apportés, et ce, pour tenir compte de la non-réponse et pour assurer la correspondance avec le dénombrement démographique selon l'âge et le sexe, par exemple). Donc, si la probabilité de sélection d'un enfant est de 1 sur 10, le poids (initial) qu'a l'enfant dans l'enquête est 10, ce qui signifie que l'enfant représente dix enfants dans la population. (Veuillez consulter le chapitre 11.0 pour obtenir plus de renseignements sur les poids de l'ELNEJ.)

Les poids de l'enquête renvoient à une population particulière et ils devraient être utilisés dans les analyses quand on doit faire des inférences sur cette population. Dans une enquête longitudinale comme l'ELNEJ, où les enfants sont suivis au fil des ans, il peut y avoir différents poids pour différentes populations. La raison en est que les populations changent avec le temps (décès, immigration et émigration).

Une enquête longitudinale peut avoir deux types de populations : la population longitudinale et les diverses populations transversales. La population longitudinale correspond à la population initiale lorsque l'échantillon a été prélevé pour la première fois (il n'y a qu'une seule population longitudinale), et une population transversale renvoie à une période ultérieure (il peut y en avoir plusieurs). Par exemple, la population longitudinale de la cohorte initiale est composée de tous les enfants dont l'âge effectif se situait entre 0 et 11 ans le 31 décembre 1994 et qui vivaient dans l'une des provinces pendant la collecte des données du cycle 1, soit en 1994-1995. Une population transversale au cycle 8 pourrait inclure les enfants qui avaient entre 0 et 11 ans le 31 décembre 1994 et qui vivaient dans l'une des provinces en 2008 (cette population englobe donc les immigrants arrivés depuis 1994-1995).

La cohorte initiale peut être utilisée pour faire des inférences concernant la première population, mais non la seconde, puisque la cohorte initiale n'a jamais été mise à niveau afin de tenir compte des immigrants arrivés après 1994-1995. Pour la dernière fois au cycle 4, des poids transversaux ont été produits pour la cohorte initiale. Au cycle 5, on estimait que l'absence de nouveaux immigrants était tellement importante que la cohorte initiale ne devrait pas être utilisée pour des inférences concernant les populations transversales après le cycle 4. Cependant, comme les cohortes du DPE sont conçues à des fins transversales et longitudinales, des poids transversaux et longitudinaux sont produits à chaque cycle.

La sous-section qui suit définit les diverses populations longitudinales et transversales pour lesquelles des inférences peuvent être faites à l'aide des poids de l'ELNEJ. Il est bon de noter que ce ne sont pas les seules populations à propos desquelles des inférences peuvent être faites. Plusieurs cohortes du DPE pourraient, par exemple, être regroupées pour représenter une population qui ne figure pas dans la liste ci-dessous. (Veuillez consulter le chapitre 15.0 pour obtenir plus de renseignements sur la façon de regrouper les échantillons de l'ELNEJ.)

Il convient également de mentionner que tous les poids finaux sont ajustés pour tenir compte de la non-réponse et des dénombrements démographiques selon l'âge, le sexe et la province. (Veuillez consulter le chapitre 11.0 pour obtenir plus de renseignements sur la façon dont les poids de l'ELNEJ sont établis.)

5.4.1 Populations longitudinales des cohortes

La cohorte initiale et les diverses cohortes du DPE représentent les populations longitudinales suivantes. Il est bon de noter que pour le premier cycle d'une cohorte, la population longitudinale est définie par les poids dits « transversaux ».

La cohorte initiale sélectionnée au cycle 1

- Population longitudinale : enfants âgés de 0 à 11 ans au 31 décembre 1994, qui vivaient dans l'une des provinces lors de la collecte des données du cycle 1 (1994-1995).

Un échantillon longitudinal d'enfants âgés de 0 à 11 ans a été prélevé à même l'EPA au cycle 1. Ces enfants avaient entre 14 et 25 ans au cycle 8 (au 31 décembre 2008). Cet échantillon a été réduit au cycle 2. Les enfants enlevés de l'échantillon entre les cycles 1 et 2 peuvent être considérés comme des enfants transversaux du cycle 1.

Des poids sont produits à chaque cycle pour cette population longitudinale.

Cohortes du Développement de la petite enfance sélectionnées aux cycles 2 à 8

- La population longitudinale de la cohorte du DPE sélectionnée au cycle 2 comprend les enfants âgés de 0 et 1 an au 31 décembre 1996, qui vivaient dans l'une des provinces lors de la collecte des données du cycle 2 (1996-1997).

Au cycle 2, on a prélevé à même l'EPA un échantillon longitudinal d'enfants de 0 et 1 an. Cette cohorte n'a été suivie que pendant trois cycles, jusqu'à ce que les enfants aient entre 4 et 5 ans. Elle ne fait pas partie de l'échantillon du cycle 8.

Des poids ont été produits aux cycles 2, 3 et 4 pour cette population longitudinale.

- La population longitudinale de la cohorte du DPE sélectionnée au cycle 3 comprend les enfants âgés de 0 et 1 an au 31 décembre 1998, qui vivaient dans l'une des provinces lors de la collecte des données du cycle 3 (1998-1999).

Au cycle 3, on a sélectionné un échantillon longitudinal d'enfants de 0 et 1 an de l'EPA et des données du registre des naissances. Cette cohorte a été suivie pendant quatre cycles. Au cycle 4, les enfants avaient 2 et 3 ans, au cycle 5 ils avaient 4 et 5 ans, et au cycle 7, les enfants avaient 8 et 9 ans. Ils ne font pas partie de l'échantillon du cycle 8.

Des poids ont été produits aux cycles 3, 4, 5 et 7 pour cette population longitudinale.

- La population longitudinale de la cohorte du DPE sélectionnée au cycle 4 comprend les enfants âgés de 0 et 1 an au 31 décembre 2000, qui vivaient dans l'une des provinces lors de la collecte des données du cycle 4 (2000-2001).

Au cycle 4, on a prélevé à même l'EPA un échantillon longitudinal d'enfants âgés de 0 et 1 an. Cette cohorte a été suivie pour quatre cycles. Elle ne fait pas partie de l'échantillon du cycle 8.

Des poids ont été produits aux cycles 4, 5, 6 et 7 pour cette population longitudinale.

- La population longitudinale de la cohorte du DPE sélectionnée au cycle 5 comprend les enfants âgés de 0 et 1 an au 31 décembre 2002, qui vivaient dans l'une des provinces lors de la collecte des données du cycle 5 (2002-2003).

Au cycle 5, un échantillon longitudinal d'enfants âgés de 0 et 1 an a été prélevé à même l'EPA. Au cycle 8, ces enfants avaient 6 et 7 ans.

Des poids ont été produits aux cycles 5, 6, 7 et 8 pour cette population longitudinale.

- La population longitudinale de la cohorte du DPE sélectionnée au cycle 6 comprend les enfants âgés de 0 et 1 an au 31 décembre 2004, qui vivaient dans l'une des provinces lors de la collecte des données du cycle 6 (2004-2005).

Au cycle 6, on a prélevé à même l'EPA un échantillon longitudinal d'enfants âgés de 0 et 1 an. Les enfants de 4 et 5 ans seront de nouveau interviewés au cycle 8.

Des poids ont été produits aux cycles 6, 7, et 8 pour cette population longitudinale.

- La population longitudinale de la cohorte du DPE sélectionnée au cycle 7 comprend les enfants âgés de 0 et 1 an au 31 décembre 2006, qui vivaient dans l'une des provinces lors de la collecte des données du cycle 7 (2006-2007).

Au cycle 7, un échantillon longitudinal d'enfants âgés de 0 et 1 an a été prélevé à même l'EPA. Les enfants de 2 et 3 ans provenant des cycles précédents ont été interviewés au cycle 8.

On a appliqué des facteurs de pondération à cette population longitudinale aux cycles 7 et 8

5.4.2 Cohortes et leurs populations transversales

La cohorte initiale (au cycle 8, les enfants étaient âgés de 14 à 25 ans au 31 décembre 2008)

Comme aucun échantillon complémentaire n'a été ajouté à la cohorte initiale afin de tenir compte des immigrants, il n'est pas recommandé d'utiliser cette cohorte pour représenter des populations transversales après le cycle 4. La cohorte initiale peut néanmoins être utilisée pour faire des inférences sur la population longitudinale définie à la sous-section 5.4.1 ainsi que les populations transversales suivantes.

- Population transversale du cycle 2 : enfants âgés de 2 à 13 ans au 31 décembre 1996, qui vivaient dans l'une des provinces lors de la collecte des données du cycle 2 (1996-1997).

Des poids transversaux ont été produits au cycle 2 pour cette population.

- Population transversale du cycle 3 : enfants âgés de 4 à 15 ans au 31 décembre 1998, qui vivaient dans l'une des provinces lors de la collecte des données du cycle 3 (1998-1999).

Des poids transversaux ont été produits au cycle 3 pour cette population.

- Population transversale du cycle 4 : enfants âgés de 6 à 17 ans au 31 décembre 2000, qui vivaient dans l'une des provinces lors de la collecte des données du cycle 4 (2000-2001).

Des poids transversaux ont été produits au cycle 4 pour cette population.

La cohorte initiale et les enfants des cohortes du DPE

Les enfants de la cohorte initiale et des différentes cohortes du DPE peuvent être utilisés pour faire des inférences sur les populations suivantes.

- Population transversale du cycle 2 : enfants âgés de 0 à 13 ans au 31 décembre 1996, qui vivaient dans l'une des provinces lors de la collecte des données du cycle 2 (1996-1997).

L'échantillon transversal comprend :

- les enfants âgés de 0 et 1 an appartenant à la cohorte du DPE et sélectionnés au cycle 2;
- retour des enfants de 2 à 13 ans appartenant à la cohorte initiale.

Des poids transversaux ont été produits au cycle 2 pour cette population.

- Population transversale du cycle 3 : enfants âgés de 0 à 15 ans au 31 décembre 1998, qui vivaient dans l'une des provinces lors de la collecte des données du cycle 3 (1998-1999).

L'échantillon transversal comprend :

- les enfants âgés de 0 et 1 an appartenant à la cohorte du DPE et sélectionnés au cycle 3;
- retour des enfants de 2 et 3 ans appartenant à la cohorte du DPE d'enfants de 0 et 1 an sélectionnés au cycle 2;
- l'échantillon complémentaire d'enfants de 5 ans (sélectionnés à partir des données du registre des naissances);
- retour des enfants de 4 à 15 ans appartenant à la cohorte initiale.

Des poids transversaux ont été produits au cycle 3 pour cette population.

- Population transversale du cycle 4 : enfants âgés de 0 à 17 ans au 31 décembre 2000, qui vivaient dans l'une des provinces lors de la collecte des données du cycle 4 (2000-2001).

L'échantillon transversal comprend :

- les enfants âgés de 0 et 1 an appartenant à la cohorte du DPE et sélectionnés au cycle 4;
- retour des enfants de 2 et 3 ans appartenant à la cohorte du DPE d'enfants de 0 et 1 an sélectionnés au cycle 3;
- retour des enfants de 4 et 5 ans appartenant à la cohorte du DPE d'enfants de 0 et 1 an sélectionnés au cycle 2;
- retour des enfants de 6 à 17 ans appartenant à la cohorte initiale.

Des poids transversaux ont été produits au cycle 4 pour cette population.

Après le cycle 4, les inférences sur les populations transversales ne doivent être faites qu'avec les enfants des cohortes du DPE.

- Population transversale du cycle 5 : enfants âgés de 0 à 5 ans au 31 décembre 2002, qui vivaient dans l'une des provinces lors de la collecte des données du cycle 5 (2002-2003).

L'échantillon transversal comprend :

- les enfants âgés de 0 et 1 an appartenant à la cohorte du DPE et sélectionnés au cycle 5;

- retour des enfants de 2 et 3 ans appartenant à la cohorte du DPE d'enfants de 0 et 1 an sélectionnés au cycle 4;
- retour des enfants de 4 et 5 ans appartenant à la cohorte du DPE d'enfants de 0 à 1 an sélectionnés au cycle 3.

Des poids transversaux ont été produits au cycle 5 pour cette population.

- Population transversale du cycle 6 : enfants âgés de 0 à 5 ans au 31 décembre 2004, qui vivaient dans l'une des provinces lors de la collecte des données du cycle 6 (2004-2005).

L'échantillon transversal comprend :

- les enfants âgés de 0 et 1 an appartenant à la cohorte du DPE et sélectionnés au cycle 6;
- retour des enfants de 2 et 3 ans de la cohorte du DPE d'enfants de 0 et 1 an sélectionnés au cycle 5;
- retour des enfants de 4 et 5 ans de la cohorte du DPE d'enfants de 0 et 1 an sélectionnés au cycle 4;
- un nouvel échantillon complémentaire d'enfants de 2 à 5 ans sélectionnés au cycle 6 (dans l'EPA).

Des poids transversaux ont été produits au cycle 6 pour cette population.

- Population transversale du cycle 7 : enfants âgés de 0 à 9 ans au 31 décembre 2006, qui vivaient dans l'une des provinces lors de la collecte des données du cycle 7 (2006-2007).
- les enfants âgés de 0 et 1 an appartenant à la cohorte du DPE et sélectionnés au cycle 7;
 - retour des enfants de 2 et 3 ans de la cohorte du DPE d'enfants de 0 et 1 an sélectionnés au cycle 6;
 - retour des enfants de 4 et 5 ans de la cohorte du DPE d'enfants de 0 et 1 an sélectionnés au cycle 5;
 - un nouvel échantillon complémentaire d'enfants de 2 à 5 ans sélectionnés au cycle 7 (dans l'EPA);
 - retour des enfants de 4 à 7 ans de l'échantillon complémentaire d'enfants de 2 à 5 ans sélectionnés au cycle 6;
 - retour des enfants de 6 et 7 ans de la cohorte du DPE d'enfants de 0 et 1 an sélectionnés au cycle 4;
 - retour des enfants de 8 et 9 ans de la cohorte du DPE d'enfants de 0 et 1 an sélectionnés au cycle 3.

Des poids transversaux ont été produits au cycle 7 pour cette population.

- Population transversale du cycle 8 : enfants âgés de 0 à 7 ans au 31 décembre 2008, qui vivaient dans l'une des provinces lors de la collecte des données du cycle 8 (2008-2009).
- la cohorte du DPE d'enfants de 0 et 1 an sélectionnés au cycle 8;
 - retour des enfants de 2 et 3 ans faisant partie de la cohorte du DPE d'enfants de 0 et 1 an sélectionnés au cycle 7;
 - retour des enfants de 4 et 5 ans faisant partie de la cohorte du DPE d'enfants de 0 et 1 an sélectionnés au cycle 6;
 - un nouvel échantillon complémentaire d'enfants de 2 à 5 ans sélectionnés au cycle 8 (dans l'EPA);
 - retour des enfants de 4 à 7 ans faisant partie de l'échantillon complémentaire d'enfants de 2 à 5 ans sélectionnés au cycle 7;

- retour des enfants de 6 et 7 ans faisant partie de l'échantillon complémentaire d'enfants de 2 à 5 ans sélectionnés au cycle 6;
- retour des enfants de 6 et 7 ans faisant partie de la cohorte du DPE d'enfants de 0 et 1 an sélectionnés au cycle 5.

Des poids transversaux ont été produits au cycle 8 pour cette population.

6.0 Collecte des données

La collecte des données pour le huitième cycle de l'Enquête longitudinale nationale sur les enfants et les jeunes (ELNEJ) s'est déroulée de l'automne 2008 à l'été 2009.

La collecte a été effectuée en trois vagues, lesquelles sont décrites ci-dessous. La date de l'interview est consignée dans la variable (HMMHgQ06).

Période de collecte	Groupe d'âge
Septembre à janvier, vague 1	Enfants de 0 à 6 ans et jeunes de 18 à 25 ans dont aucun frère ou sœur de 14 à 17 ans n'a été sélectionné
Janvier à avril, vague 2	Enfants de 7 ans, jeunes de 14 à 17 ans et 18 à 25 ans pour qui aucune donnée n'a été recueillie au cours de la première période de collecte
Mars à juillet, vague 3	Enfants de 0 à 5 ans

L'enquête combine l'utilisation de méthodes d'interview assistée par ordinateur (IAO) et de questionnaires sur format papier. L'ELNEJ a recours à deux types d'applications d'IAO : l'interview sur place assistée par ordinateur (IPAO) et l'interview téléphonique assistée par ordinateur (ITAO). Dans les deux cas, l'intervieweur lit les questions qui figurent à l'écran et inscrit dans l'ordinateur les réponses données par le répondant. Pour l'IPAO, le répondant et l'intervieweur remplissent le questionnaire en personne, tandis que pour l'ITAO, le répondant remplit le questionnaire au téléphone. Les applications d'IAO permettent de créer des cheminements complexes et d'intégrer des vérifications au questionnaire, ce qui aide à produire des données de meilleure qualité et à faire en sorte que le répondant ne réponde qu'aux questions qui s'appliquent à sa situation. Les questions sont les mêmes, que l'interview se fasse au moyen de la technique de l'IPAO ou de celle de l'ITAO. Selon la composition du ménage et la nature des composantes requises, l'interview est effectuée en tout ou en partie par téléphone et/ou par visite sur place. Vous trouverez ci-après une brève description des composantes assistées par ordinateur et des questionnaires sur format papier « outils de collecte » ou « instruments d'enquête » utilisés pour la collecte de l'ELNEJ. Veuillez consulter le chapitre 8.0 du présent guide pour obtenir une description plus détaillée du contenu des questionnaires.

6.1 Composante « Ménage »

Cette première partie de l'interview sert à dresser la liste de tous les membres du ménage, à déterminer quels sont les liens de parenté entre eux, à recueillir des renseignements de dépistage et à prendre connaissance des caractéristiques démographiques de base, notamment celles sur le sexe, la date de naissance, l'état matrimonial et les liens de parenté entre les membres du ménage. Le sexe et la date de naissance ne sont mis à jour que pour les nouveaux répondants.

Cette composante permet en outre d'identifier la personne la mieux renseignée (PMR) sur l'enfant. Celle-ci fournit des renseignements sur tous les enfants sélectionnés au sein du ménage ainsi que sur elle-même et sur son conjoint ou son partenaire. Dans certains cas, il aurait été approprié de désigner deux PMR différentes dans un même ménage. Par exemple, dans le cas d'une famille reconstituée, il aurait peut-être été plus approprié de désigner la mère comme PMR pour un enfant et le père pour un autre enfant. Cependant, pour simplifier le déroulement de l'interview, une seule PMR est choisie par ménage.

Cette identification a lieu une fois que sont recueillis les renseignements sur les liens de parenté entre les membres du ménage.

6.2 Composante « Enfant »

Une composante « Enfant » est créée pour chaque enfant de 0 à 17 ans qui a été sélectionné sauf ceux qui sont âgés de 16 ou 17 ans qui vivent de manière autonome. La PMR répond aux questions de la composante « Enfant ». Il s'agit normalement de la mère de l'enfant, mais il peut également s'agir du père, d'un beau-parent ou d'un parent adoptif qui habite dans le même logement. Seule la PMR ou son conjoint peut répondre aux questions de cette composante.

Pour les ménages où le seul enfant sélectionné appartient au sous-groupe des 16 et 17 ans et qui réside au domicile de ses parents, on a demandé une version plus courte de la composante « Enfant ». Si l'enfant ne réside plus au domicile de ses parents, la composante ne sera pas créée.

6.3 Composante « Adulte »

Une composante « Adulte » est créée pour la personne la mieux renseignée et son conjoint ou son partenaire si l'enfant sélectionné est âgé de 17 ans et moins, sauf ceux qui sont âgés de 16 ou 17 ans qui vivent de manière autonome. Seule la PMR ou son conjoint peut répondre aux questions de cette composante. Les questions de la composante « Adulte » ne sont posées qu'une seule fois par ménage, même si plus d'un enfant a été sélectionné dans le ménage.

Pour les ménages où le seul enfant sélectionné appartient au sous-groupe des 16 et 17 ans et qui réside au domicile de ses parents, on demande une version plus courte de la composante « Adulte ». S'il ne réside plus au domicile de ses parents, la composante ne sera pas créée.

6.4 Composante « Jeune »

Cette composante est réservée aux jeunes de 16 ans et plus qui sont sélectionnés. Le jeune est la seule personne pouvant répondre aux questions de cette composante et cela, qu'il réside ou non dans le logement familial.

6.5 Tailles de l'échantillon au cycle 8

Les tableaux suivants présentent le nombre d'enfants et de jeunes qui font partie de l'échantillon, selon l'âge et la province, pour le cycle 8.

Nombre d'enfants dans l'échantillon et taux de réponse selon l'âge, cycle 8

Âge au 31 décembre 2008 (Année)	Échantillonné	Admissible	Répondant	Taux de réponse au cycle 8 (%)
	Nombre			
0	2 400	2 389	1 808	75,7
1 an	3 082	3 074	2 298	74,8
2 ans	2 064	2 060	1 643	79,8
3 ans	3 516	3 495	2 729	78,1
4 ans	2 291	2 278	1 761	77,3
5 ans	3 113	3 094	2 369	76,6
6 ans	1 992	1 984	1 636	82,5
7 ans	2 279	2 272	1 814	79,8
8 ans	1	0	0	s/o
11 ans	1	0	0	s/o
14 ans	1 528	1 527	1 227	80,4
15 ans	1 606	1 602	1 274	79,5
16 ans	1 136	1 134	890	78,5
17 ans	1 102	1 101	880	79,9
18 ans	1 304	1 302	850	65,3
19 ans	1 219	1 213	785	64,7
20 ans	1 178	1 172	689	58,8
21 ans	1 183	1 182	677	57,3
22 ans	1 232	1 221	757	62,0
23 ans	1 186	1 185	713	60,2
24 ans	1 250	1 246	755	60,6
25 ans	1 132	1 122	711	63,4
Total	35 795	35 653	26 266	73,7

Source : Statistique Canada, Enquête longitudinale nationale sur les enfants et les jeunes.

Nombre d'enfants dans l'échantillon et taux de réponse selon la province de résidence, cycle 8

Province	Échantillonné	Admissible	Répondant	Taux de réponse au cycle 8 (%)
	Nombre			
Terre-Neuve-et-Labrador	2 104	2 098	1 704	81,2
Île-du-Prince-Édouard	1 513	1 511	1 167	77,2
Nouvelle-Écosse	2 347	2 339	1 744	74,6
Nouveau-Brunswick	2 318	2 312	1 778	76,9
Québec	5 658	5 642	4 344	77,0
Ontario	8 579	8 559	6 059	70,8
Manitoba	2 710	2 697	1 940	71,9
Saskatchewan	2 868	2 857	2 219	77,7
Alberta	4 244	4 227	2 988	70,7
Colombie-Britannique	3 375	3 360	2 272	67,6
En dehors des 10 provinces	79	51	51	100,0
Total	35 795	35 653	26 266	73,7

Source : Statistique Canada, Enquête longitudinale nationale sur les enfants et les jeunes.

6.6 Mesures directes

Différentes mesures directes sont administrées aux répondants sélectionnés. Elles sont résumées dans le tableau ci-dessous. Veuillez consulter le chapitre 14.0 du présent guide pour obtenir plus de renseignements à ce sujet.

Nom de la mesure	Groupe d'âge	Méthode utilisée
Échelle de vocabulaire en images de Peabody (EVIP)	4 et 5 ans	Interview assistée par ordinateur
Qui suis-je?	4 et 5 ans	Questionnaire sur format papier
Connaissance des nombres	4 et 5 ans	Interview assistée par ordinateur
Exercice de calculs mathématiques	7 à 15 ans, de la 2 ^e à la 10 ^e année	Questionnaire sur format papier
Résolution de problèmes	16 et 17 ans	Questionnaire sur format papier
Évaluation des capacités de lecture	18 et 19 ans	Questionnaire sur format papier
Évaluation des habiletés au calcul	20 et 21 ans	Questionnaire sur format papier

6.7 Questionnaires autoadministrés (14 à 17 ans)

Les répondants de 14 à 17 ans remplissent un questionnaire sur format papier portant sur différents aspects de leur vie. On remet le questionnaire au jeune pendant l'interview et on lui demande de le remplir lui-même. Une fois le questionnaire rempli, le jeune doit le glisser dans une enveloppe, cacheter celle-ci et la remettre à l'intervieweur afin d'assurer la confidentialité.

Les questionnaires autoadministrés consistent de deux livrets correspondants à chacun des groupes d'âge visés. Vous trouverez ci-après une grille indiquant les sujets traités dans le livret selon le groupe d'âge. Un même sujet peut comporter des questions différentes selon l'âge de l'enfant. Vous trouverez un exemplaire de chacun de ces livrets en vous référant au Livre 2 du document « Enquête longitudinale nationale sur les enfants et les jeunes — Matériel d'enquête pour la collecte des données de 2008-2009, cycle 8 ».

Sujet	Section autoadministrée selon l'âge et le livret	
	14 et 15 ans Livret n° 22	16 et 17 ans Livret n° 23
Amis et famille	A	A
École	B	...
À propos de moi	C	B
Sentiments et comportements	D	C
Activités	E	...
Tabac, alcool et drogues	F	D
Mes parents	G	G
Puberté	H	...
Relations amoureuses	H	F
Santé	H	E
Travail	I	...
Remerciement	J	H

... n'ayant pas lieu de figurer

Source : Statistique Canada, Enquête longitudinale nationale sur les enfants et les jeunes.

6.8 Personnel de collecte (formation, supervision et contrôle)

Des intervieweurs de Statistique Canada ont effectué les interviews de l'ELNEJ. Plusieurs avaient déjà eu l'occasion de travailler dans le cadre d'un ou de plusieurs cycles antérieurs de l'ELNEJ.

Tous ces intervieweurs relèvent d'un groupe d'intervieweurs principaux chargés de s'assurer qu'ils sont familiers avec les concepts et les procédures de l'enquête. Les intervieweurs principaux s'assurent que des mesures de suivi sont prises rapidement dans les cas de refus et de toutes autres non-réponses quelconques. Au besoin, les cas de non-réponse sont confiés à l'intervieweur principal, puis réaffectés. Les intervieweurs principaux, quant à eux, relèvent des gestionnaires de programme, qui se trouvent dans chacun des bureaux régionaux de Statistique Canada.

Dans le cas de l'ELNEJ, on a combiné la formation en salle de classe et l'autoapprentissage pour s'assurer que les intervieweurs et les superviseurs comprenaient bien les concepts de l'enquête. L'autoapprentissage qui a précédé la formation a nécessité que les gestionnaires de programme, les intervieweurs principaux et les intervieweurs prennent connaissance du Guide de l'intervieweur préparé pour l'enquête. Les intervieweurs principaux reçoivent la formation en salle de classe en premier lieu par les employés du bureau régional pour ensuite donner la formation à l'intention des intervieweurs sous leur responsabilité dans leurs secteurs respectifs.

6.8.1 Interview effectuée dans une langue autre qu'une des langues officielles

Les questionnaires de l'ELNEJ ne sont disponibles qu'en français et en anglais. Si un répondant souhaite être interviewé dans une autre langue, celui-ci sera donc passé à un intervieweur qui parle sa langue, si cela est possible. Au cycle 8, moins de 80 cas n'ont été complétés en raison de problèmes d'ordre linguistique.

7.0 **Traitement des données**

Les principaux produits de l'Enquête longitudinale nationale sur les enfants et les jeunes (ELNEJ), cycle 8, sont des fichiers de données épurées qui peuvent être utilisés par les chercheurs. Le présent chapitre donne un bref résumé de certaines des phases de traitement inhérentes à la production de ces fichiers.

Le traitement des données du cycle 8 de l'ELNEJ a été effectué à l'aide d'un environnement de traitement généralisé de la division. Il s'agit d'un système générique qui suit les étapes suivantes pour « nettoyer » un fichier du début à la fin. Les étapes principales sont :

- Nettoyage
- Vérifications de l'âge et du genre
- Vérifications des liens familiaux
- Prédiction
- Vérifications de l'enchaînement des questions
- Codage
- Vérifications de la cohérence
- Variables dérivées
- Fichier de traitement final
- Création du fichier principal

7.1 **Vérification informatique**

Tel qu'indiqué ci-dessus, tous les renseignements de chaque ménage, sauf pour les questionnaires autoadministrés et les mesures directes qui sont destinés aux enfants de 14 à 17 ans, sont recueillis en personne ou au moyen d'une interview téléphonique, à l'aide d'une application d'interview assistée par ordinateur (IAO). On peut ainsi inclure diverses fonctions de vérification dans les questionnaires pour les diverses composantes IAO des ménages, de sorte que les renseignements recueillis sont de grande qualité. Voici quelques exemples illustrant le genre de vérifications faites dans le cadre du processus d'interviews assistées par ordinateur exécutées pour l'ELNEJ :

Écrans de révision

On a prévu des écrans de révision pour les données importantes et complexes. Par exemple, la procédure de sélection de la personne la mieux renseignée (PMR), élément clé de l'enquête, s'appuie sur la liste des membres du ménage. La page-écran correspondante présente des données démographiques sur chaque membre et son lien avec chacun des autres membres du ménage. Les données recueillies s'affichent à l'écran pour que l'intervieweur les confirme auprès du répondant avant de continuer l'interview.

Vérifications d'intervalles

On a également intégré des vérifications d'intervalles dans le système IAO pour les questions où il faut indiquer des valeurs numériques. Si les chiffres indiqués ne s'inscrivent pas dans l'intervalle, une fenêtre apparaît instantanément pour indiquer l'erreur et demander à l'intervieweur de corriger la réponse erronée. Par exemple, si la valeur entrée dans l'ordinateur pour le poids de l'enfant à la naissance est trop élevée ou trop basse, un message s'affiche et demande à l'intervieweur de vérifier la réponse auprès du répondant.

Vérifications du cheminement de questions

Tous les cheminements de questions sont en mode intégration automatique au système IAO. Par exemple, à la section sur la garde de l'enfant, on demande à la PMR si elle place l'enfant en garderie ou en gardiennage. Le cheminement des questions peut varier selon la réponse. Si l'on fait appel à une garderie, le système IAO pose une suite de questions sur les modes de garde particuliers qui sont utilisés pour l'enfant. Dans le cas contraire, il saute automatiquement ces questions.

Vérifications de la cohérence générale

On a prévu un certain nombre de vérifications de la cohérence dans le système IAO, et les intervieweurs sont en mesure de revenir à des questions déjà posées pour rectifier les incohérences. Les intervieweurs reçoivent des instructions à l'écran pour traiter ou régler des problèmes de réponse incomplète ou erronée. Le système génère une fenêtre éclair laquelle signale l'erreur et demande à l'intervieweur de revenir à la question pour vérifier les données et apporter les corrections nécessaires.

7.2 Saisie des données

Saisie des données pour les questionnaires sur format papier

On a fait la saisie des questionnaires suivants au bureau central de Statistique Canada :

- les questionnaires autoadministrés destinés aux enfants de 14 à 17 ans
- « Qui suis-je ? » destiné aux enfants de 4 et 5 ans
- les exercices de calcul destinés aux enfants de 7 ans et ceux de 14 à 15 ans qui sont en 2^e année ou plus
- les exercices de résolution de problèmes pour les jeunes de 16 et 17 ans
- le test d'évaluation de la littératie pour les jeunes de 18 et 19 ans
- le test d'évaluation de la numératie pour les jeunes de 20 et 21 ans

On fait la saisie de tout document contenant au moins une question à laquelle le répondant a donné une réponse et on remet un fichier contenant tous les enregistrements au bureau central pour la suite du traitement. De plus, on a intégré dans ce processus de saisie certaines vérifications de la qualité qui marquent les entrées inusitées afin de mettre les opérateurs en garde contre les risques d'entrées erronées.

Dans les cas où plus d'une réponse avait été cochée par le répondant, on a donné pour instruction aux opérateurs d'entrer la valeur « Non déclaré ». Il s'agit d'un changement par rapport au cycle 7, lorsque les opérateurs avaient pour instruction d'accepter la première réponse. Les erreurs qui étaient encore dans le questionnaire, étaient corrigées à une étape ultérieure.

7.3 Nettoyage

Exigences définitives

L'objet de cette étape est de retirer les enregistrements complets en doubles et de séparer les enregistrements d'identification en double afin de les examiner. Ensuite, les données sont divisées en réponses et en non-réponses, selon les critères déterminés à l'avance.

On examine les questionnaires des répondants et des non-répondants et on établit des spécifications fondées sur cette analyse afin de déterminer quels enregistrements seront éliminés en raison de la non-réponse. Essentiellement, s'il manque des données importantes sur un enregistrement ou s'il y manque plus de la moitié des questions, on élimine cet enregistrement du fichier.

À la fin de cette étape, les enregistrements sont traités selon le type de questionnaire, c'est-à-dire : le questionnaire de l'adulte, le questionnaire de l'enfant, le questionnaire du jeune, le questionnaire du ménage et les questionnaires autoadministrés.

Variables manquantes

Pour toutes les variables manquantes des ménages, on indique « Non déclaré ». Si les renseignements sont insuffisants, on retire le ménage de l'échantillon des répondants et on en fait un cas de non-réponse.

Chaque fichier contient des enregistrements d'enfants qui sont maintenant décédés ou qui ont déménagé à l'étranger, mais qu'on gardera aux fichiers à des fins de pondération. Il y a 114 enregistrements de la cohorte du DPE, 143 enregistrements longitudinaux et 189 enregistrements concernant les jeunes qui ont un poids longitudinal (HWTCW01L) et toutes les autres variables sont

fixées à « Non déclaré ». Pour les jeunes de ce groupe qui sont âgés de 16 et 17 ans, ceux-ci apparaîtront sur le fichier longitudinal ainsi que sur le fichier concernant les jeunes (53 cas).

7.4 Vérification de l'âge et du genre

À cette étape, on effectue une vérification pour les variables relatives à l'âge concernant les membres faisant partie du nouvel échantillon ainsi que les nouveaux membres de ménages de cycles précédents. Il s'agit d'un changement par rapport au cycle 6 auquel on effectuait une vérification pour tous les membres du ménage. À compter du cycle 7, l'application ne permettait pas d'effectuer des mises à jour pour les membres d'un échantillon ainsi que les membres de ménages de cycles précédents. De plus, on vérifie l'âge afin d'être cohérent avec l'âge de la cohorte. On vérifie le sexe du répondant afin d'être cohérent avec le cycle précédent.

7.5 Vérification des liens

L'étape de la vérification des liens sert précisément à authentifier les liens entre les membres du ménage et permet de créer des variables familiales dérivées. Cette étape exécute une série de vérifications standards afin de comparer les données des liens déclarés pour tous les membres d'un ménage en particulier; certaines incohérences sont réglées automatiquement grâce à une application qui utilise des règles établies, tandis qu'on indique d'autres incohérences afin de faire une révision manuelle et un recodage. On produit une série de variables dérivées connexes à partir des vérifications de liens.

7.6 Vérification préliminaire

Pour tous les enregistrements où il manque des valeurs (blancs) provenant de la collecte, on insère la valeur « 9 », « 99 », « 999 » etc. pour indiquer qu'aucune information n'a été recueillie. Les entrées « Ne sait pas » retournées par l'application IAO comme « 9 » sont modifiées à « 7 », à l'étape de la vérification préliminaire. De même, on a modifié les questions pour lesquelles on peut lire « Inscrivez toutes les réponses qui s'appliquent » afin de transformer leurs valeurs en réponses « Oui » (« 1 ») ou « Non » (« 2 »). Enfin, toutes les réponses écrites par les répondants sont enlevées du fichier de traitement et sont mises de côté afin d'être traitées séparément.

7.7 Vérification de cheminement

Cette vérification consiste à reproduire le cheminement des questions incluses dans le questionnaire. Les variables qui manquent, en suivant l'enchaînement des questions, sont changées de « Non déclaré » à « Enchaînement valide » (« 6 », « 96 », « 996 » etc.).

Lorsque l'instruction « Passez à » est basée sur l'âge ou suit une question pour laquelle on a une réponse, toutes les questions auxquelles on n'a pas répondu sont indiquées comme « Enchaînement valide ». Lorsque l'instruction « Passez à » suit une réponse « Ne sait pas » ou « Refus », toutes les questions auxquelles on n'a pas répondu sont indiquées comme « Non déclaré ».

Dans le cycle 8, on comptait 24 jeunes de 16 et 17 ans vivant de manière autonome (HDMCfD03=82). Toutes les variables présentes sur le fichier longitudinal ont été codées à « Enchaînement valide » pour ces répondants, sauf pour quelques variables dans la section sur les données démographiques.

7.8 Codage des questions ouvertes

Les intervieweurs enregistrent sur le questionnaire quelques éléments d'information sous forme de réponses à des questions ouvertes. Ainsi, à la section sur l'activité professionnelle, on pose à la PMR qui a travaillé au cours des 12 derniers mois une série de questions ouvertes sur l'emploi actuel ou l'emploi le plus récent :

- De quel genre d'entreprise, de service ou d'industrie s'agit-il/s'agissait-il?
- Quel genre de travail est-ce que vous faites/faisiez?
- Dans ce travail, quelles sont/étaient vos fonctions ou activités les plus importantes ?

On a posé des questions sur l'aspiration professionnelle dans les questionnaires du jeune pour les 18 à 21 ans :

- Quel genre de carrière ou travail espères-tu obtenir lorsque tu auras environ 30 ans?
- Spécifier le genre de carrière ou travail.
- Spécifier le genre d'entreprise.

Mode d'enregistrement

L'intervieweur inscrit littéralement les réponses données par le répondant. Au bureau central, ces énoncés écrits sont convertis en codes d'industrie et de profession pour décrire la nature du travail du répondant.

Mode de codage

Ces questions ouvertes sont codées selon plusieurs classifications types. Pour le cycle 8, les questions portant sur les professions ont été codées à l'aide de la Classification nationale des professions pour statistiques (CNP-S) 2001, tandis que celles portant sur les industries l'ont été à l'aide du Système de classification des industries de l'Amérique du Nord de 2002 (SCIAN 2002)³. On peut consulter les listes groupées de ces codes dans les fichiers de données suivants : (HLFPgD7A et HLFPfD8A pour la PMR, et HLFSgD7A et HLFSfD8A pour son conjoint ou son partenaire).

7.9 Vérification de la cohérence

Après la vérification de l'enchaînement, on procède à une vérification de la cohérence qui consiste à vérifier les liens entre deux variables ou plus. On utilise des tables de décision pour préciser les vérifications de la cohérence. Le logiciel *LogiPlus* a servi pour saisir l'information dans les tables de décision et pour générer le code SAS. On produit un rapport comprenant les comptes « avant » et « après » des variables, de même qu'un rapport précisant le nombre de décisions prises pour chaque table de décision.

Par exemple, à la section sur les activités, une question s'adressant aux enfants âgés de 3 à 7 ans ayant déclaré utiliser un ordinateur à leur maison (HACCeQ4E), il y avait une vérification de la cohérence qui permettait de comparer ces enfants à ceux ayant déclaré avoir un ordinateur à la maison (HACCeQ4D). Si l'enfant a déclaré utiliser un ordinateur à la maison alors la réponse à la question « Y a-t-il un ordinateur à la maison? » (HACCeQ4D) a été établie à « Oui » dans la vérification.

Au cycle 7, un changement a été apporté au processus de vérification de la cohérence pour ce qui est des questions portant sur le développement moteur et social. Aux cycles précédents, on établissait des corrélations entre les réponses aux questions de la section des grandes étapes du développement et celles de la section du développement moteur et social. Par exemple, si l'enfant n'était pas capable de s'asseoir seul (MILQ1), on présumait qu'il ne pouvait rester assis pendant dix minutes sans soutien (HMSCQ16). Au cycle 8, aucune corrélation n'a été établie. Cela a touché plusieurs questions de la section du développement moteur et social.

3. Des renseignements sur la classification ainsi que la concordance avec les classifications antérieures se trouvent à l'adresse suivante : www.statcan.gc.ca/concepts/index-fra.htm.

7.10 Indicateurs d'imputation

Variables manquantes

Pour diverses raisons, quelques variables sont absentes pour certains ménages répondants compris dans le fichier de l'ELNEJ. On leur attribue généralement un code de non-réponse partielle ou de réponse partielle.

Imputation

Cependant, pour certaines variables du fichier de l'ELNEJ, on procède à une imputation au lieu d'attribuer un code spécial de non-réponse. L'imputation consiste à remplacer des réponses manquantes ou incohérentes par des valeurs vraisemblables. Dans le cas de l'ELNEJ, l'imputation porte sur le revenu du ménage, sur celui de la PMR, sur celui du jeune et sur le développement moteur et social. Veuillez consulter le chapitre 10.0 du présent guide pour obtenir plus de renseignements à ce sujet.

On a ajouté des indicateurs d'imputation au fichier de l'ELNEJ ainsi que les questions précises de divers enregistrements ayant été imputées pour que l'utilisateur soit informé par rapport à l'ampleur de cette imputation.

Tous les indicateurs d'imputation du fichier de données de l'ELNEJ comportent un « I » en cinquième ou en sixième position dans le nom de la variable. Ainsi, la variable de l'indicateur d'imputation du revenu de la PMR est HINPI1AA.

7.11 Établissement de variables dérivées et vérifications longitudinales

Établissement de variables à partir de plus d'un élément d'information

Un certain nombre de variables sont dérivées en combinant des questions du questionnaire ou d'autres variables dérivées pour faciliter l'analyse des données. Par exemple, à la section sur l'activité professionnelle, on est informé de la situation actuelle du répondant quant à son activité à partir des questions sur son travail actuel. À la section sur le niveau de scolarité des adultes, on demande s'il fréquente présentement l'école, le collège ou l'université. En combinant ces deux variables, on forme une variable dérivée qui permet d'identifier la situation actuelle, pour ce qui est du travail et du niveau de scolarité du répondant (HLFPD51, HLFSD51).

Variables longitudinales dérivées

Dans les cycles précédents, une variable dérivée était créée pour indiquer les changements entre les données déclarées aux cycles précédents et celles du cycle actuel par rapport à la structure familiale, la PMR et le conjoint. À compter du cycle 7, cette variable dérivée n'existe plus. Au lieu, les identificateurs de la PMR et du conjoint ont été ajoutés aux fichiers de données des cycles 4 à 8.

Les incohérences entre les cycles quant au poids et à la taille de l'enfant sont signalées au moyen d'un indicateur. Les variables indiquant ces incohérences contiennent un « Z » en cinquième position du nom de la variable.

C'est à cette étape que sont également reportées les données des cycles précédents concernant des variables pour lesquelles des questions ne sont posées qu'une seule fois au répondant (poids à la naissance, renseignements sur l'allaitement naturel, données sociodémographiques, notamment). Des données sont aussi reportées pour ce qui est des variables dérivées des cycles précédents, obtenues à partir de données ne changeant pas au fil du temps.

Nom des variables dérivées

Toutes les variables dérivées dans le fichier de données de l'ELNEJ reçoivent un « D » en cinquième ou en sixième position dans leur nom. Le nom de la variable dérivée portant sur le principal mode de garde est HCRCgD2A.

7.12 Structure de codage des variables

On a conçu des normes pour la structure de codage des variables de l'ELNEJ afin de décrire certaines situations de façon cohérente pour toutes les variables. Voici ces diverses situations avec leurs propres codes descriptifs.

Refus

Au cours d'une interview assistée par ordinateur, le répondant peut choisir de refuser de répondre à une question particulière. Le système IAO comporte une touche de fonction sur laquelle l'intervieweur appuie pour indiquer un refus. L'information est enregistrée pour la question et transmise au bureau central.

Sur le fichier des données de l'ELNEJ, une question refusée comporte un code « 8 ». Dans le cas d'une variable à un chiffre, le code est « 8 », pour une variable à deux chiffres « 98 », pour une variable à trois chiffres « 998 », etc.

Ne sait pas

Il se peut que le répondant ne connaisse pas la réponse à une question particulière. Là encore, le système IAO comporte une touche de fonction sur laquelle l'intervieweur appuie pour expliquer cette situation.

Dans le fichier des données de l'ELNEJ, le code utilisé pour indiquer que le répondant ne connaissait pas la réponse à une question particulière est « 7 », « 97 », « 997 », etc.

Enchaînement valide

Dans certains cas, une question ne s'applique pas au répondant. Dans le fichier de données, on se sert des codes « 6 », « 96 », « 996 », etc., pour indiquer qu'une question ou une variable dérivée est un enchaînement valide.

Dans certains cas, une question ou une série de questions est sans objet. Par exemple, la question sur le nombre d'heures hebdomadaires où l'enfant est en garderie (HCRCg1GA) ne s'applique que pour les enfants qui vont à la garderie (HCRCgQ1G = 1). Dans les autres cas, on attribue le code « 996 » à cette question.

Occasionnellement, c'est une section entière du questionnaire, voire tout le questionnaire qui est sans objet. Ainsi, la section sur le développement moteur et social ne vise que les enfants de 0 à 3 ans. Pour tous les autres enfants de quatre ans et plus, le système indique que les variables en question sont des « Enchaînements valides ».

Dans le cas où la PMR n'est pas mariée ou n'a pas de conjoint de fait vivant dans le ménage, toutes les variables du conjoint (sections sur l'activité professionnelle et le niveau de scolarité du conjoint) deviennent des « Enchaînements valides » et sont codées ainsi.

Non déclaré

Dans le cas du traitement qui se fait au bureau central, on code parfois comme « Non déclaré » la réponse à une question. On indique ainsi que la réponse à la question est inconnue. On attribue les codes « Non déclaré » pour les raisons suivantes :

- Dans le cadre de l'interview assistée par ordinateur (IAO), l'intervieweur pouvait entrer un code « Refus » ou « Ne sait pas », ainsi qu'on l'a expliqué plus haut. Le système IAO est souvent programmé en pareil cas pour ne pas tenir compte de cette section particulière du questionnaire. En cas de « Refus », on suppose que les questions posées sont délicates et que le répondant ne veut probablement pas répondre à d'autres questions à ce sujet. Dans les cas « Ne sait pas », on part du principe que le répondant n'est pas suffisamment informé pour répondre à d'autres questions. Dans le traitement des données de l'ELNEJ, on a décidé que toutes les questions subséquentes se verraient attribuer un code « Non déclaré ». On entend par là que la

question n'est pas posée au répondant. On ignore même parfois si la question s'appliquait à ce dernier.

- Dans certains cas, il est impossible de remplir au complet un questionnaire et même de le commencer. Par exemple, il peut se produire une interruption ou le répondant dit ne pas vouloir continuer. Si l'on obtient suffisamment de renseignements pour que le ménage soit considéré comme répondant, on attribue le code « Non déclaré » à toutes les questions restantes (et aux questionnaires que l'on n'a pas encore commencés). Comme seule exception, si on sait qu'une section ou qu'un questionnaire ne s'applique pas, on code ces éléments comme « Enchaînement valide ».
- Pour les questionnaires sur format papier, quelques questionnaires partiellement remplis sont parfois retournés. Dans ces cas, si l'on a recueilli suffisamment de données sur le répondant, on attribue le code « Non déclaré » à toutes les autres questions sans réponse du questionnaire. Comme seule exception, si l'on sait qu'une section ou qu'un questionnaire ne s'applique pas, on code ces éléments comme « Enchaînement valide ».
- Occasionnellement, on peut avoir recours à des codes « Non déclaré » à la suite des vérifications de la cohérence. Si l'on décèle une erreur de cohérence entre le lien des groupes de variables, on attribue le code « Non déclaré » à une variable ou à plusieurs.

Dans le cas des variables dérivées, si une ou plusieurs des variables d'entrée (d'une variable dérivées) reçoivent le code « Refus », « Ne sait pas » ou « Non déclaré », on attribue donc le code « Non déclaré ».

On attribue un code « 9 », « 99 », « 999 », etc. aux cas « Non déclaré ».

7.13 Nomenclature

Le système de la documentation du fichier de microdonnées de l'ELNEJ utilise certaines normes de désignation aux noms de variables et de valeurs. Le but étant de rendre l'interprétation des données plus facile à l'utilisateur.

On a conçu une nomenclature pour chaque variable du fichier de données de l'ELNEJ afin de renseigner plus précisément l'utilisateur sur ces dernières. Tous les noms de variables sont d'au plus huit caractères, ce qui permet d'utiliser facilement ces désignations avec des progiciels d'analyse comme SAS ou SPSS. Les indicateurs « Persruk » et « Fieldruk » constituent des exceptions à cette règle d'affectation des noms.

Nomenclature des variables

Le **premier** caractère du nom de la variable désigne le cycle de l'ELNEJ :

- « **A** » indique le premier cycle;
- « **B** » indique le deuxième cycle;
- « **C** » indique le troisième cycle;
- « **D** » indique le quatrième cycle;
- « **E** » indique le cinquième cycle;
- « **F** » indique le sixième cycle
- « **G** » indique le septième cycle
- « **H** » indique le huitième cycle

Les **deuxième et troisième caractères**, lesquels forment un acronyme, désignent la section du questionnaire où la question a été posée ou dont la variable a été tirée. Veuillez consulter la section 7.15 du présent guide pour une liste des acronymes utilisés pour désigner les sections du questionnaire.

Le **quatrième** caractère du nom de la variable désigne l'unité de collecte ou l'unité à laquelle se rapporte la variable. Il y a sept possibilités⁴ :

- « **C** » si la variable indique l'enfant;
- « **P** » si la variable indique la PMR;
- « **S** » si la variable indique le conjoint ou son partenaire;
- « **H** » si la variable indique le ménage;
- « **Y** » si la variable indique le jeune;
- « **M** » si la variable indique la mère
- « **O** » si la variable indique les autres membres du ménage âgés de 15 ans et plus.

Les **cinquième, sixième, septième et huitième** caractères du nom de la variable désignent le cycle pendant lequel la variable est apparue la première fois (si ce n'est pas le cycle 1), le type de variable et un numéro séquentiel attribué à la variable. Par exemple :

- « **h** » La lettre minuscule désigne le cycle de l'ELNEJ où la variable a figuré pour la première fois au fichier ou le cycle (par exemple « h » désigne le cycle 8) dans lequel des changements ont été apportés à une question qui avait déjà été posée.
- « **Q** » Désigne la variable qui est liée à une question directement posée dans un des questionnaires de l'ELNEJ.
- « **S** » Désigne un résultat calculé pour une des échelles utilisées dans le questionnaire.
- « **D** » Signifie que la variable a été calculée ou dérivée à partir des données d'autres questions posées dans le questionnaire ou de variables codées.
- « **I** » Signifie que la variable est un indicateur d'imputation.
- « **Z** » Signifie que la variable est un indicateur d'incohérence des données déclarées entre le cycle en cours et les cycles antérieurs.
- « **nnx** » Désigne la question ou la variable d'identification; en général, « nn » est un numéro séquentiel attribué à une variable et « x », un indicateur alphabétique séquentiel pour une suite de variables du même ordre.

Entre les cycles, on a pu réviser le contenu du questionnaire. Si la révision a changé le sens ou les valeurs d'une question au cycle 8, la variable est considérée comme nouvelle et se voit accorder la lettre « h ». Par exemple, au cycle 8, la variable HCRChQ6B a été demandée à propos des enfants de 6 et 7 ans uniquement, ce qui a donné lieu à un changement de couverture. En ce qui concerne les enfants de 5 ans et moins, une nouvelle question a été posée pour l'âge en mois (HCRChQ6A).

4. Il convient de noter que, bien qu'il y ait des variables pour différentes unités d'analyse (PMR, conjoint/partenaire et ménage par exemple), on ne pourra produire que des « estimations pour les enfants » à partir du fichier de microdonnées de l'ELNEJ. On peut se reporter aux caractéristiques de la PMR, du conjoint/partenaire et du ménage pour décrire les caractéristiques de l'enfant. Ainsi, il sera possible d'estimer le nombre d'enfants vivant dans un ménage à faible revenu ou dont la PMR a reçu une note élevée à l'échelle de la dépression. Il sera toutefois impossible de produire des estimations du nombre de ménages à faible revenu ou de PMR déprimées.

7.14 Exemples de noms de variables

Voici des exemples qui illustrent la nomenclature des variables au fichier des données de l'ELNEJ.

Nom de la variable	Objet
HLFSQ02	Q2 de la section sur l'activité professionnelle du conjoint ou partenaire
H	Variable du cycle 8
LF	Section sur l'activité professionnelle
S	Conjoint ou partenaire
Q	Question directement posée dans le questionnaire
02	Identificateur de la question

Nom de la variable	Objet
HPRCS03	Note d'interaction positive à l'échelle du rôle parental pour un enfant de 2 à 15 ans
H	Variable du cycle 8
PR	Section sur le rôle parental
C	Enfant
S	Note ou résultat
03	Identificateur de la question

7.15 Siglaison des sections du questionnaire

Le tableau qui suit indique les acronymes utilisés pour les diverses sections des questionnaires de l'ELNEJ. Ces acronymes font partie du nom de toutes les variables du fichier des données de l'ELNEJ. Ils occupent les deuxième et troisième positions du nom de la variable. Veuillez consulter le chapitre 8.0 du présent guide pour plus de renseignements concernant les sections du questionnaire.

Acronyme	Variable	Recueillie ou tirée de
GE	Renseignements géographiques	Données d'échantillon
HH	Ménage	Caractéristiques du logement
MM	Variables recueillies pour établir la liste du ménage	Variables démographiques de base pour chaque membre du ménage; celles-ci figurent au fichier des données de l'ELNEJ pour l'enfant, le jeune, la PMR et son conjoint ou son partenaire.
DM	Variables démographiques calculées qui décrivent les conditions de vie de l'enfant ou du jeune	Données de la liste du ménage et de la grille des liens
SD	Renseignements sociodémographiques	Questionnaire de l'enfant La PMR et son conjoint ou son partenaire du questionnaire de l'adulte
HL	Santé	La PMR et son conjoint ou son partenaire du questionnaire de l'adulte Questionnaire de l'enfant Questionnaire du jeune
CH	États chroniques des adultes	La PMR et son conjoint ou son partenaire à la section sur la santé du questionnaire de l'adulte
RS	Limitations d'activités	La PMR et son conjoint ou son partenaire à la section sur la santé du questionnaire de l'adulte
DP	Échelle de dépression	Questionnaire des parents (échelle administrée à la PMR)
ED	Niveau de scolarité	Questionnaire de l'enfant Questionnaire de l'adulte (section pour la PMR et son conjoint ou son partenaire) Questionnaire du jeune
LF	Activité professionnelle	Questionnaire de l'adulte (section pour la PMR et son conjoint ou son partenaire) Questionnaire du jeune
IN	Revenu	Questionnaire de l'adulte (Revenu du ménage et revenu personnel de la PMR) Questionnaire du jeune
FN	Fonctionnement de la famille	Questionnaire de l'adulte (administré à la PMR)
MD	Renseignements médicaux et biologiques	Questionnaire de l'enfant
TM	Tempérament	Questionnaire de l'enfant
LT	Alphabétisation	Questionnaire de l'enfant
AC	Activités	Questionnaire de l'enfant Questionnaire du jeune
BE	Comportement	Questionnaire de l'enfant
MS	Développement moteur et social	Questionnaire de l'enfant
WB	Travail après la naissance	Questionnaire de l'enfant
RL	Rapports sociaux	Questionnaire de l'enfant Questionnaire du jeune

Acronyme	Variable	Recueillie ou tirée de
PR	Style parental	Questionnaire de l'enfant
CR	Garde des enfants	Questionnaire de l'enfant
SL	Sommeil	Questionnaire de l'enfant
PB	Comportement positif	Questionnaire de l'enfant
CM	Échelle de communication	Questionnaire de l'enfant
ML	Grandes étapes du développement	Questionnaire de l'enfant
CS	Garde légale	Questionnaire de l'enfant
AG	Étapes du développement	Questionnaire de l'enfant
SF	Sécurité du voisinage	Questionnaire de l'adulte (section administrée à la PMR ou à son conjoint ou à son partenaire)
SP	Soutien social	Questionnaire de l'adulte (section administrée à la PMR ou à son conjoint ou à son partenaire) Questionnaire du jeune
AS	Aspirations	Questionnaire de l'enfant Questionnaire du jeune
EQ	Quotient émotionnel	Questionnaires autoadministrés Questionnaire du jeune
AM	Ce que je suis, moi	Questionnaires autoadministrés Questionnaire du jeune
FB	Sentiments et comportements	Questionnaires autoadministrés Questionnaire du jeune
MO	Départ du foyer familial	Questionnaire du jeune
FF	Formation des familles et fertilité	Questionnaire du jeune
AB	Capacités	Questionnaire du jeune
PO	Engagement politique	Questionnaire du jeune
FF	Amis et famille	Questionnaires autoadministrés
SC	École	Questionnaires autoadministrés
PM	Mes parents et moi	Questionnaires autoadministrés
PU	Puberté	Questionnaires autoadministrés
DR	Usage du tabac, de l'alcool et des drogues	Questionnaires autoadministrés
AT	Activités	Questionnaires autoadministrés
HT	Santé	Questionnaires autoadministrés
WK	Travail et sources d'argent	Questionnaires autoadministrés
DA	Fréquentations	Questionnaires autoadministrés
PP	Test EVIP	Échelle de vocabulaire en images de Peabody
WI	Qui suis-je?	Mesure directe (questionnaire Qui suis-je?)
KN	Connaissance des nombres	Mesure directe (questionnaires Connaissance des nombres)
MA	Tests de mathématiques	Tests de mathématiques et Exercice de résolution de problèmes
LT	Évaluation de la littératie	Questionnaire du test d'évaluation de la littératie
NU	Évaluation de la numératie	Questionnaire du test d'évaluation de la numératie
WT	Poids	Poids faisant partie du plan d'échantillonnage

7.16 Fichiers de traitement final et fichiers principaux

Les fichiers suivants ont été créés pour le cycle 8 de l'ELNEJ :

- la *cohorte du Développement de la petite enfance* (0 à 7 ans) (questionnaires de l'adulte, de l'enfant et du ménage)
- la *cohorte longitudinale – enfant* (14 à 17 ans) (questionnaires de l'adulte, de l'enfant et du ménage)
- la *cohorte longitudinale – jeunes* (16 à 25 ans) (questionnaires du jeune et du ménage; variables statiques reportées des cycles précédents)
- les *questionnaires autoadministrés* (14 à 17 ans)

8.0 Contenu de l'enquête

L'Enquête longitudinale nationale sur les enfants et les jeunes (ELNEJ) a été conçue en fonction d'une approche écologique ou holistique permettant de mesurer le développement de l'enfant. L'enquête permet de saisir toute la diversité et la dynamique des facteurs ayant une incidence sur les enfants et les jeunes. Afin d'assurer une couverture adéquate de tous les domaines pertinents qui influent sur le développement de l'enfant, une consultation multidisciplinaire a eu lieu dès la conception de l'enquête. Le choix des priorités, des questions d'enquête et des domaines particuliers ont résulté en grande partie d'un effort concerté et à la contribution et aux conseils :

- du groupe consultatif d'experts de l'ELNEJ, composé de spécialistes du développement de l'enfant et des sciences sociales;
- de représentants de ministères fédéraux;
- de représentants provinciaux et territoriaux responsables des programmes liés au développement de l'enfant.

On a recommandé que l'ELNEJ porte sur une large gamme de caractéristiques et de facteurs influant sur la croissance et le développement de l'enfant. On a recueilli une foule de renseignements au sujet de l'enfant, de son ou ses parents, des caractéristiques de la famille et du quartier ainsi que de l'école de l'enfant et de sa vie à l'école. Dans la présente section, on donne un aperçu du contenu de chacune des sections du questionnaire ayant servi à la collecte des données de l'ELNEJ. L'ELNEJ comprend plusieurs échelles dont il sera brièvement question dans le présent chapitre. Les différentes échelles de l'ELNEJ et la validation des résultats obtenus sur celles-ci sont expliquées dans le détail au chapitre 9.0.

8.1 Composantes de l'enquête

L'ELNEJ est subdivisée en plusieurs composantes, lesquelles sont décrites au chapitre 6.0, Collecte des données. Un sommaire de chaque composante est présenté ci-dessous.

Ménage	La liste des membres du ménage est remplie au début de l'interview en recueillant des renseignements démographiques de base sur chacun des membres ainsi que des données sur les liens qui les unissent à tous les membres du ménage.
Adulte	Questions relatives à la personne la mieux renseignée (PMR) et son conjoint. Dans le cas des enfants de 16 et 17 ans, on ne pose pas les questions comprises dans toutes les sections de la composante des adultes. La composante des adultes est remplie une seule fois, même lorsque le ménage comprend deux enfants sélectionnés. Aucune composante « Adulte » n'est produite pour les jeunes de 18 ans et plus ou ceux qui sont âgés de 16 ou 17 ans qui vivent de manière autonome.
Enfant	Les questions relatives à l'enfant sélectionné sont posées à la PMR. La composante des enfants est remplie pour chacun des enfants sélectionnés âgés de 0 à 7 ans et de 14 à 17 ans. Les seules sections du questionnaire de l'enfant qui s'applique aux jeunes de 16 et 17 ans sont les suivantes : aspirations et attentes, garde légale des enfants et caractéristiques sociodémographiques.
Jeune	Dans cette section, les jeunes âgés de 16 à 25 ans répondent aux questions qui les concernent dans une interview assistée par ordinateur (IAO).
Questionnaires autoadministrés	Les jeunes de 14 à 17 ans répondent aux questions qui les concernent dans le questionnaire sur format papier.

Évaluations directes

Plusieurs évaluations directes sont remplies par les enfants et les jeunes; celles-ci sont décrites au chapitre 14.0.

8.2 Variables démographiques

Les variables démographiques sont recueillies dans la composante du ménage. Certains renseignements démographiques de base (p. ex. l'âge, le sexe, l'état matrimonial) sont recueillis pour tous les membres du ménage de l'enfant. Ce questionnaire vise en outre à recueillir des données sur la grille des liens, c'est-à-dire les liens entre tous les membres du ménage. Ces données ont permis de créer un vaste ensemble de variables pour décrire la situation familiale de l'enfant.

8.3 Questionnaire de l'adulte**Niveau de scolarité (parent)**

La section sur le niveau de scolarité est remplie pour la PMR et son conjoint ou son partenaire. L'objectif est de recueillir des renseignements sur le nombre d'années d'études terminées, le plus haut niveau de scolarité atteint ainsi que la fréquentation actuelle d'un établissement d'enseignement.

Certaines recherches ont démontré qu'il existe un lien entre le niveau de scolarité de la mère, le milieu familial et le développement de l'enfant. Les questions sur la fréquentation scolaire à temps plein et à temps partiel constituent un indicateur des principales activités de la PMR et de son conjoint ou son partenaire.

Activité sur le marché du travail

La stabilité d'emploi a un effet sur le milieu familial, tant sur le plan du revenu que sur celui du niveau de stress. Certaines recherches démontrent que le chômage des parents peut avoir un effet néfaste sur les résultats des enfants.

La section traitant de l'activité sur le marché du travail est remplie à la fois pour la PMR et son conjoint ou son partenaire. L'objectif principal de cette section est de déterminer la stabilité d'emploi, puisque celle-ci constitue un indicateur de la constance du revenu d'emploi. Les questions portent sur les périodes d'absence du travail, les raisons de l'absence la plus récente, le nombre d'heures de travail et l'organisation du travail (p. ex. le travail par quarts) au cours de l'année précédente. On recueille des renseignements sur l'emploi actuel ou l'emploi le plus récent de la PMR et de son conjoint ou son partenaire.

Une description complète de l'emploi actuel ou de l'emploi le plus récent est consignée. L'industrie et la profession sont codées selon le Système de classification des industries de l'Amérique du Nord (SCIAN) de 2002 et la Classification nationale des professions pour la statistique de 2001 (CNP-S).

Variables dérivées relatives à l'activité sur le marché du travail

Plusieurs variables dérivées relatives à l'activité sur le marché du travail ont été créées pour la PMR et son conjoint ou son partenaire. Parmi ces variables, on retrouve les suivantes :

HLFPfD5A / HLFSfD5A :	code du SCIAN 2002 pour l'emploi actuel de la PMR et de son conjoint
HLFPfD6A / HLFSfD6A :	code de la CNP-S 2001 pour l'emploi principal de la PMR et de son conjoint
HLFPgD7A / HLFSgD7A :	code de la classification type des industries pour l'emploi actuel (SCIAN 2002) — regroupé
HLFPfD8A / HLFSfD8A :	code de la classification type des professions pour l'emploi actuel (CNP-S 2001) — regroupé

Revenu

Dans la section de l'enquête portant sur le revenu, on recueille des renseignements sur les sources et le montant du revenu pour chacun des ménages. On y trouve également quelques questions sur la perception qu'a la PMR ou son conjoint quant à leur situation financière. Ces renseignements constituent un indicateur de la situation économique de la famille, une composante essentielle du cadre de vie de l'enfant.

Puisque le revenu familial est une variable importante de bon nombre d'études sur le développement de l'enfant, on a imputé une valeur au revenu du ménage en l'absence de réponses à ces questions. Le chapitre 10.0 présente une explication détaillée de la procédure d'imputation du revenu.

On a créé une variable dérivée (HINHGD4A) afin de comparer le revenu du ménage au seuil de faible revenu (SFR)⁵ avant impôts. Le seuil de faible revenu sert à distinguer les unités familiales qui sont à « faible revenu » des « autres » groupes. Une unité familiale est considérée à faible revenu lorsque son revenu est inférieur à la valeur du seuil correspondant à sa taille d'unité familiale et à la taille de sa communauté. La variable HINHGD03A donne la valeur du SFR selon la zone géographique.

Santé (adulte)

Cette section comporte des questions adressées aux PMR et à leur conjoint sur leur état de santé général, des états chroniques et des limitations d'activités, ainsi que des questions portant sur le tabagisme et sur la consommation d'alcool. On a inclus les questions sur le tabagisme parce que des recherches ont indiqué que le comportement des parents à cet égard pourrait permettre de prévoir le tabagisme chez les enfants. On aborde la consommation d'alcool en raison de son incidence potentielle sur la santé physique et mentale de l'adulte, la situation économique de la famille et les relations familiales.

États chroniques

On demande aux PMR et à leur conjoint s'ils souffrent d'états chroniques à long terme (p. ex. allergies, asthme, hypertension). Les variables dérivées (HCHPgD01 et HCHSgD01) indiquent une réponse affirmative, c'est-à-dire que les répondants ont déclaré au moins un état chronique à long terme.

Limitations d'activités

On pose aux PMR et à leur conjoint une série de questions sur leurs limitations d'activités notamment à la maison, au travail, à l'école. Les variables dérivées (HRSPdD01 et HRSSdD01) ont été créées afin de déterminer si la PMR ou son conjoint a déclaré des limitations d'activités.

Antécédents maternels

Ces questions visent principalement les antécédents gravidiques de la mère. Elles sont posées seulement à celles qui sont interviewées pour la première fois. Les questions portant sur la grossesse et l'accouchement ont été formulées par J.-F. Saucier de l'Hôpital Sainte-Justine de Montréal, puis adaptées par l'équipe de projet.

Échelle de dépression

Les PMR répondent aux questions se rapportant à l'échelle de dépression (HDPPS01) dans le questionnaire de l'adulte.

Fonctionnement de la famille

Le but de cette section est de fournir une évaluation globale du fonctionnement de la famille et une indication de la qualité des relations familiales. Ces questions sont posées à la PMR ou à son conjoint si l'enfant est âgé de 0 à 15 ans.

5. Pour obtenir plus de renseignements, veuillez consulter *Seuils de faible revenu de 2006* et *Mesures du faible revenu de 2005*, produit n° 75F0002MIF2007004 au catalogue de Statistique Canada.

Sécurité du quartier

Cette section permet de recueillir des renseignements sur la satisfaction des répondants relativement à leur quartier comme milieu de vie pour les enfants, y compris leurs perceptions quant à l'importance des dangers et des problèmes, à la cohésion sociale ou au sentiment de « voisinage ». Deux échelles ont été créées dans cette section : celle de la sécurité du quartier (HSFHhS5), qui mesure le degré de sécurité du quartier tel que le perçoivent les répondants, et l'échelle du voisinage (HSFHhS6), qui mesure le degré de cohésion sociale dans le quartier.

Soutien social

L'objet de cette section est de recueillir des renseignements sur le niveau de soutien que les PMR pensent avoir de la part de leurs amis, des membres de leur famille et de leur collectivité. Ces questions sont posées à la PMR ou à son conjoint si l'enfant est âgé de 0 à 15 ans.

Caractéristiques sociodémographiques

La section sur les caractéristiques sociodémographiques vise à recueillir des renseignements sur l'immigration, l'origine ethnique, le profil linguistique ainsi que l'appartenance religieuse de la PMR et de son conjoint. Ces données permettent d'analyser les diverses composantes de la population canadienne et d'identifier les personnes appartenant à une minorité visible.

8.4 Questionnaire de l'enfant

Les questions du questionnaire de l'enfant sont posées en fonction de l'âge effectif de l'enfant. Dans l'ELNEJ, plutôt que d'utiliser l'âge réel de l'enfant, on utilise un âge calculé appelé « âge effectif » (HMMCQ01) afin de s'assurer que l'enfant reste dans le groupe d'âge auquel il a été associé, peu importe que la collecte ait lieu avant ou après son anniversaire. Pour le cycle 8, l'âge effectif est calculé en fonction de l'année 2008, c'est-à-dire 2008 moins l'année de naissance de l'enfant. Par exemple, un enfant né en 2004 aurait un âge effectif de 4 ans (2008 moins 2004). Il est bon de noter que l'âge réel de l'enfant au moment de l'interview est parfois différent de l'âge effectif.

Niveau de scolarité (enfant)

Cette section vise à obtenir des renseignements de base au sujet des expériences scolaires de l'enfant. La quantité et le genre de renseignements recueillis varient selon l'âge de l'enfant, davantage de données étant recueillies au sujet des enfants plus âgés qui ont une expérience scolaire plus vaste.

Pour tous les groupes d'âge, on recueille les renseignements de base sur les sujets suivants : le niveau de scolarité de l'enfant, le genre d'école fréquentée et la langue d'enseignement, l'attrait qu'exerce l'école sur l'enfant, l'absentéisme, le nombre de changements d'école et de déménagements.

Dans le cas des enfants de première année ou ceux plus avancés, on pose des questions supplémentaires sur d'autres aspects de l'expérience éducative, comme le redoublement scolaire, le rendement scolaire et l'éducation spécialisée.

Mesures directes

L'objet de cette section est de préparer le terrain pour les tests de mesures directes qu'on fera passer aux enfants de 4 à 5 ans. Si les enfants ne peuvent pas y répondre (p. ex. s'il ne parlent ni l'anglais ni le français ou s'ils sont daltoniens), on ne leur fera pas passer les tests.

Santé (enfant)

Cette section vise à recueillir des renseignements sur l'état de santé physique de l'enfant — état de santé général, blessures, limitations et états chroniques — de même que sur le recours aux services de santé et aux médicaments.

Dans le cas d'un enfant de 4 à 5 ans, on recueille également des renseignements sur l'état de santé couvrant divers sujets tels que l'ouïe, la vue, la parole et le bien-être mental général. À partir de ces renseignements, on calcule l'Indice de l'état de santé (IES3) (HHLCCD2A). L'IES3 est un indice

général de l'état de santé permettant de synthétiser les aspects tant quantitatifs que qualitatifs de la santé. Cet indice, mis au point au *Centre for Health Economics and Policy Analysis* de la McMaster University, est fondé sur le *Comprehensive Health Status Measurement System* (CHSMS). Il donne une description de la santé fonctionnelle générale de la personne selon neuf attributs : la vue, l'ouïe, la parole, la mobilité (capacité de se déplacer), la dextérité (usage des mains et des doigts), la cognition (mémoire et raisonnement), l'émotion (sentiments) ainsi que la douleur et l'inconfort.

Les scores de l'IES3 reflètent les perceptions sociales de l'état de santé. Les préférences de chaque personne sont représentées par une valeur numérique (normalement entre 0 et 1) pour un état de santé donné. (Certains des pires états de santé reçoivent souvent des valeurs inférieures à 0, ce qui indique que la personne les considère comme pires que la mort.) Cet indice est également utilisé dans le cadre de l'Enquête nationale sur la santé de la population.

Renseignements médicaux et biologiques

La section des renseignements médicaux et biologiques vise les enfants de 0 à 5 ans. L'objectif principal consiste à recueillir des renseignements sur certains facteurs, comme l'âge gestationnel et le poids à la naissance, dont il a été démontré qu'ils ont un effet direct sur la croissance et le développement de l'enfant. Par exemple, à long terme, les bébés ayant un poids insuffisant courent des risques plus élevés d'être en mauvaise santé et d'éprouver des problèmes persistants liés au développement comparativement à la moyenne.

Pour chaque enfant de moins de 2 ans, on recueille des données sur la nature de l'accouchement, l'état de santé général du bébé à la naissance et l'utilisation de services spécialisés après la naissance. L'ELNEJ permet aussi de recueillir des renseignements sur la grossesse de la mère biologique et sur l'accouchement, y compris sur des sujets d'intérêt stratégique comme les expériences d'allaitement de la mère et son style de vie avant la naissance de l'enfant.

Il y a des variables dérivées qui sont créées dans cette section qui méritent d'être soulignées. Deux variables ont été dérivées pour indiquer l'âge gestationnel de l'enfant. HMDCD06 exprime cet âge en jours et HMDCD07 indique si l'enfant est né avant terme (âge gestationnel de 258 jours ou moins), à terme (âge gestationnel de 259 jours à 293 jours) ou après terme (âge gestationnel de 294 jours et plus).

On a également créé une variable (HMDCD08) permettant d'indiquer si le poids de l'enfant à la naissance est normal (2 500 grammes ou plus), modérément faible (de 1 500 à 2 499 grammes) ou très faible (moins de 1 500 grammes).

Travail de la mère après la naissance d'un enfant

On a posé ces questions pour déterminer quelle était la période écoulée entre la naissance d'un enfant et le retour au travail de la mère ainsi que le niveau de participation des mères au marché du travail après leur retour.

Questionnaires sur les étapes du développement

Les questionnaires Étapes du développement (QED) sont des outils de déclaration des parents, mis au point par Jane Squires, LaWanda Potter et Diane Bricker de la University of Oregon, et servant à repérer les bébés et les jeunes enfants susceptibles d'éprouver des problèmes de développement. Il existe 19 questionnaires visant les enfants de 4 à 60 mois. Chaque questionnaire comporte environ une trentaine de questions couvrant cinq aspects du développement :

- 1) Communication : p. ex. babillage, vocalisation, écoute et compréhension
- 2) Motricité globale : coordination des bras, du corps et des jambes
- 3) Motricité fine : coordination des mains et des doigts
- 4) Résolution de problèmes : diverses activités avec des objets, dessin
- 5) Facultés personnelles et sociales : jeux solitaires et sociaux, habillement et alimentation sans assistance.

L'aspect « motricité globale » des QED n'est pas inclus parce qu'il est déjà traité dans d'autres sections de l'enquête. Les questionnaires comprennent aussi une section générale se rapportant aux préoccupations courantes des parents. Cependant, cette section n'est pas utilisée dans le cadre de l'ELNEJ, les questions étant semblables à celles déjà incluses dans l'enquête.

L'ELNEJ fait appel aux QED pour les enfants de 3 à 47 mois inclusivement. Après avoir consulté l'éditeur, Statistique Canada a converti les questionnaires pour en faciliter l'administration dans les interviews assistées par ordinateur (IAO).

Principales étapes du développement

Ces questions ont été ajoutées pour obtenir une meilleure mesure du développement des jeunes enfants. Ensemble, les principales étapes du développement (p. ex. les premiers mots ou les premiers pas) donnent une indication générale du développement de l'enfant. Des experts du Dunedin en Nouvelle-Zélande ont recommandé à l'équipe de projet d'utiliser les étapes marquantes comme mesures du développement. Les questions sont tirées des questionnaires provisoires du *Early Childhood Longitudinal Study Program* (cohorte des naissances) du *National Center for Education Statistics* des États-Unis.

Tempérament

Cette section permet de mesurer le tempérament des jeunes enfants selon la perception qu'ont les parents du degré de difficulté que présente leur enfant. Cette mesure est fondée sur l'hypothèse que le tempérament de l'enfant est conditionné non seulement par des facteurs biologiques, mais aussi par les perceptions des parents quant au degré de difficulté de l'enfant.

Alphabétisation

Cette section permet de mesurer l'expérience des enfants avec les livres, leur intérêt pour la lecture et les activités d'apprentissage auxquelles les parents s'adonnent avec eux. La section porte principalement sur la stimulation que les jeunes enfants reçoivent à la maison.

Pour les enfants de 0 à 2 ans, on pose plusieurs questions visant à mesurer la fréquence de certaines activités auxquelles s'adonnent les parents avec leurs enfants, par exemple, leur raconter des histoires, leur chanter des chansons et leur montrer de nouveaux mots. Ces questions sont inspirées de la *Early Childhood Longitudinal Study* des États-Unis. Des questions semblables sont posées au sujet des enfants de 3 à 7 ans; elles ont été adaptées aux activités correspondant à ce groupe d'âge.

Communication

Ces questions sont adaptées de la *Competent Children Study* de la Nouvelle-Zélande. Elles portent sur les capacités de l'enfant de comprendre les messages oraux, de transmettre un message à autrui et de communiquer verbalement. La dernière question, portant sur la bonne compréhension des messages verbaux, n'est posée qu'au sujet des enfants de 3 ans. On pose une question semblable au sujet des enfants de 4 et 5 ans dans le cadre de l'Indice de l'état de santé, dans la section sur la santé.

Activités

Cette section permet de mesurer la participation des enfants à diverses activités non scolaires. On recueille des renseignements quant à l'emploi du temps de l'enfant, quels sont ses intérêts personnels et quel degré d'interaction qu'il a avec ses pairs.

On a ajouté plusieurs questions destinées aux enfants de 4 à 7 ans afin de déterminer la fréquence de certaines activités qu'ils partagent avec leurs parents, par exemple les repas, les jeux et les travaux ménagers. Lorsqu'un ménage compte deux conjoints ou partenaires, les questions sont posées à la PMR et à son conjoint ou son partenaire.

Comportement

L'objectif de cette section est d'évaluer certains aspects du comportement des enfants de 2 ans et plus ainsi que les habitudes alimentaires des enfants de 1 an à 3 ans.

Les questions que contient cette section servent à mesurer la fréquence de comportements tels que l'hyperactivité et l'agressivité physique.

Comportement positif

Cette section vise à évaluer le comportement positif des enfants de 3 à 5 ans, notamment la persévérance et l'autonomie. Selon la *Competent Children Study* de la Nouvelle-Zélande, la persévérance et l'autonomie figurent parmi l'ensemble des compétences constituant de bons indicateurs du rendement général de l'enfant.

Les questions sont inspirées de l'étude néo-zélandaise et des questions se rapportant au comportement des enfants ont été utilisées pour d'autres groupes d'âge de l'ELNEJ.

Sommeil

Les recherches révèlent que les troubles du sommeil permettent de prévoir des difficultés éventuelles chez les enfants. À l'inverse, l'absence de tels troubles est corrélée avec un tempérament facile et un développement positif.

Ces questions portent sur les heures de sommeil, les heures de sommeil ininterrompu par nuit, combien de fois le sommeil des parents a été interrompu par l'enfant, entre autres.

Développement moteur et social

L'échelle du développement moteur et social (DMS) mesure divers aspects du développement moteur, social et cognitif des enfants de 0 à 3 ans; les questions varient selon l'âge des enfants. Trois scores (HMSCS01, HMSCS02 et HMSCdS03) sont établis à partir de ces questions.

Relations

L'objectif de cette section consiste à recueillir des renseignements sur les relations que les enfants entretiennent avec les autres. Les relations positives avec d'autres enfants et adultes peuvent contribuer à neutraliser les effets d'autres facteurs qui entraînent des risques pour l'enfant.

Les questions touchant les activités avec les amis et la façon dont l'enfant s'accorde avec les parents, le personnel enseignant ainsi que les amis sont fondées sur les questions de l'Étude sur la santé des jeunes ontariens.

Rôle parental

On estime que le rôle parental exerce une influence importante sur le comportement et le développement de l'enfant. L'objectif de cette section est de mesurer certains comportements parentaux. On a créé des échelles à partir des questions de cette section.

On a posé certaines questions aux PMR ayant un conjoint ou partenaire à la maison afin de déterminer jusqu'à quel point ils s'entendent au sujet des décisions parentales. Cette question a été élaborée par l'équipe de projet et s'apparente à des questions liées à l'échelle de Strayhorn et Weidham dont sont inspirées les autres questions sur le rôle parental.

Garde légale des enfants

Cette section vise à fournir des renseignements sur la situation familiale de l'enfant : parents mariés ou non, séparés ou divorcés, et l'âge de l'enfant au moment de la séparation ou du divorce.

Garde d'enfants

Cette section fournit des renseignements de base sur les services de garde actuels des enfants ainsi que des données sur les services de garde antérieurs. Les données permettent notamment de mesurer le temps passé par l'enfant en garderie et le genre de service de garde utilisé pour chaque enfant. On recueille également des renseignements sur le nombre de changements de services de garde qu'a vécus l'enfant ainsi que sur les raisons de ces changements au cours des 12 derniers mois. Ces questions permettent aussi de déterminer s'il s'agit d'une garderie à but lucratif ou non lucratif et s'il s'agit de services de garde en milieu familial agréés ou non agréés ainsi que le ratio « travailleurs de services de garde : enfants ».

Les questions concernant les interactions parentales avec les fournisseurs de services de garde des enfants (HCRCgQ20, HCRCg22A, HCRCg22B) ont été adaptées à partir de l'Enquête sociale générale de Statistique Canada tout comme celles traitant la forme préférée de garde des enfants (HCRCg27, HCRCg28A - HCRCg28K). Des questions supplémentaires concernant la forme préférée de garde des enfants (HCRCgQ29, HCRCg30A - HCRCg30I) ont été développées par l'équipe de projet. Les questions concernant les critères de sélection des services de garde des enfants (HCRCg25A - HCRCg25K, HCRCg26A - HCRCg26K) ainsi que les coûts des services de garde des enfants (HCRCg40A, HCRCg40B) ont été adaptées à partir de la *Early Childhood Longitudinal Study* du *National Center for Education Statistics* des États-Unis. Les questions traitant la disponibilité de l'aide financière ou des subventions (HCRCgQ36, HCRCgQ38, HCRCgQ39) ont été tirées de l'enquête intitulée *Sensibilisation et attitudes concernant l'apprentissage précoce et la garde des enfants* (EKOS).

À partir du cycle 7, on a demandé aux répondants qui ont affirmé ne pas avoir recours aux services de garde d'enfants pendant leurs heures de travail ou d'études s'ils les utilisaient pour d'autres raisons. Puisqu'il s'agissait d'un important changement dans la manière dont les données sont recueillies pour la section sur la garde d'enfants, l'équipe de projet de l'ELNEJ a pris la décision de créer deux ensembles de variables dérivées pour cette section. Une série a été créée pour ceux qui ont indiqué utiliser les services de garde pendant leurs heures de travail ou d'études, et une autre série de variables dérivées a été créée pour inclure tous les répondants utilisateurs de services de garde. Par exemple, la variable dérivée « Nombre total de modes de garde utilisés » (HCRCgD03) comprend les répondants qui utilisent les services de garde pour permettre à la PMR et à son conjoint de travailler ou d'étudier. Elle a aussi été créée en tant que HCRCgD08, mais celle-ci comprend tous les répondants qui ont recours à la garde d'enfants. Deux ensembles de variables dérivées ont également été créés pour chacune des catégories suivantes : Nombre d'heures par semaine pour tous les modes de garde, Mode principal de garde, Nombre d'heures par semaine d'utilisation du mode de garde principal, et Durée d'utilisation du mode de garde principal, en mois.

Attentes et aspirations (16 et 17 ans)

On a ajouté des questions pour évaluer les aspirations et les attentes des parents à l'égard de leurs enfants ainsi que l'opinion des parents quant à l'expérience scolaire de leurs enfants. La discussion de l'expérience scolaire et les projets d'études sont associés au succès scolaire.

Ces questions ont été élaborées par le Centre de la statistique de l'éducation de Statistique Canada, à partir des questions de l'ELNEJ et d'autres enquêtes sur l'éducation comme l'Enquête auprès des jeunes en transition et l'Enquête auprès des sortants.

Caractéristiques sociodémographiques

Les questions de cette section visent à recueillir des renseignements sociodémographiques sur l'origine ethnique, le pays d'origine, les langues parlées et l'identité autochtone.

8.5 Questionnaire du jeune (16 à 25 ans)

Départ du foyer parental (18 à 25 ans)

De l'adolescence à l'âge adulte, les jeunes passent par de nombreuses transitions. L'une des plus importantes est sans aucun doute quand les jeunes quittent le foyer parental pour la première fois pour vivre de façon autonome dans leur propre logement.

Certains de ces jeunes quittent le foyer de façon permanente ou temporaire pour aller étudier ou travailler. On a estimé que des renseignements devaient être recueillis sur cette transition en raison de son importance dans le passage de l'enfance à l'âge adulte.

Les questions dans cette section varieront selon l'information déjà recueillie et celle qui l'a été durant la composante « Ménage ».

Les questions ont été élaborées en consultation avec Dianne Looker, directrice du département de sociologie de la Acadia University, qui nous a également fourni des questions qu'elle a utilisées

dans son étude longitudinale intitulée *The Transition from Education to Employment: A Longitudinal and Cohort Analysis of Canadian Youth*.

Niveau de scolarité des jeunes (16 à 25 ans)

Cette section recueille de l'information sur l'expérience scolaire des jeunes. Les premières questions servent à établir le niveau de scolarité actuel. Les questions qui s'appliquent à leur situation sont ensuite posées au répondant. Il y a quatre possibilités :

- 1) les sortants (ceux qui ont abandonné leurs études avant l'obtention du diplôme d'études secondaires);
- 2) les finissants (ceux qui ne fréquentent plus l'école, mais qui ont obtenu leur diplôme);
- 3) les élèves actuellement à l'école (ceux qui fréquentent encore une école secondaire);
- 4) les étudiants actuellement au postsecondaire (ceux qui poursuivent des études postsecondaires).

Ces questions ont été élaborées par le Centre de la statistique de l'éducation de Statistique Canada, à partir des questions de l'ELNEJ et d'autres enquêtes sur l'éducation comme l'Enquête auprès des jeunes en transition et l'Enquête auprès des sortants.

Cette section vise entre autres à cerner les facteurs qui influent sur la décision des jeunes de poursuivre leurs études ou de les abandonner.

Étant donné qu'il y a de nombreux répondants d'âges différents, mais dans des situations similaires par rapport aux études, les jeunes sont dirigés vers les bonnes questions non en fonction de leur âge mais de leur situation par rapport aux études, passées et présentes.

Activité des jeunes sur le marché du travail (16 et 17 ans)

La section relative à l'activité des jeunes a pour objectif de mesurer l'expérience des jeunes sur le marché du travail. Certains jeunes travaillent à temps partiel alors qu'ils poursuivent des études, d'autres sont passés des études au marché du travail. Cette section comprend une combinaison de questions de l'ELNEJ tirées des questionnaires autoadministrés des enfants de 14 et 15 ans et de questions sur l'activité des adultes sur le marché du travail.

On recueille auprès des jeunes des renseignements sur leur emploi actuel, sur l'emploi pendant l'année scolaire courante et sur leur travail au cours de l'été précédent.

Activité des jeunes sur le marché du travail (18 à 25 ans)

Les questions dans la présente section sont semblables à celles posées aux jeunes de 16 et 17 ans, mais plusieurs questions sont prises dans la section sur l'activité des adultes sur le marché du travail, et cela, parce que les répondants sont maintenant plus âgés. Pour quelques jeunes, le travail peut constituer l'activité principale, et les questions doivent tenir compte de cette possibilité.

Ces questions permettent de recueillir des renseignements qui permettront de se faire une idée d'ensemble quant à la participation des jeunes sur le marché du travail, notamment en ce qui a trait à la situation d'emploi, aux caractéristiques de l'emploi, au nombre d'heures travaillées, à la stabilité d'emploi et aux liens entre les objectifs en matière de travail et d'études et les réalisations.

Comme c'était le cas dans la section sur les activités des adultes relatives au marché du travail, on a consigné une description complète de l'emploi actuel ou de l'emploi le plus récent. L'industrie et la profession sont codées selon le Système de classification des industries de l'Amérique du Nord (SCIAN) de 2002 et la Classification nationale des professions pour la statistique de 2001 (CNP-S).

Variables dérivées relatives à l'activité sur le marché du travail

Plusieurs variables dérivées relatives à l'activité sur le marché du travail ont été créées pour les jeunes de 18 à 25 ans. Ce sont notamment les suivantes : HLYYeD5A, HLYYeD6A, HLYYgD7A et HLYYeD8A.

Aspirations professionnelles des jeunes (18 à 21 ans)

Cette section vise à recueillir des données sur le type de renseignements que les jeunes se procurent sur différents cheminements de carrière. Elle indique également si le jeune a pris une décision au sujet de sa carrière. Les questions varient selon l'âge du répondant.

On pense que les aspirations professionnelles chez les jeunes leur donnent une orientation réaliste et leur permettent de trouver des emplois convenables et satisfaisants. Il est important de recueillir des renseignements sur leurs attentes en matière de travail afin de comprendre dans quelle mesure ils planifient leur future carrière.

Certaines des questions figurant dans cette section ont été élaborées en consultation avec Dianne Looker de la Acadia University. D'autres questions ont été tirées de son étude intitulée *The Transition from Education to Employment: A Longitudinal and Cohort Analysis of Canadian Youth*.

Dans le cas des jeunes ayant indiqué vers quelle carrière ils désiraient se diriger, on a effectué le codage des professions choisies à l'aide des codes de la CNP-S (2001) et, à partir de ces renseignements, on a créé la variable HASYfD03.

Revenu des jeunes (16 et 17 ans)

La section sur le revenu des jeunes se rapporte aux revenus que les jeunes ont tirés de diverses sources au cours des 12 derniers mois.

Revenu des jeunes (18 à 25 ans)

Les questions sont semblables à celles posées aux adultes. Le jeune donne des renseignements sur son revenu personnel, s'il est célibataire ou sur le revenu du ménage s'il est marié ou s'il vit en union libre.

À cet âge, nombre de jeunes quittent la maison familiale pour la première fois. Cette section comprend une série de questions sur le paiement des dépenses de logement et des dettes et les économies financières. Il est important de recueillir des données visant à déterminer s'ils paient un logement, s'ils remboursent leurs dettes et s'ils ont des économies ou des placements afin d'évaluer comment les jeunes s'adaptent à leurs responsabilités financières.

Santé des jeunes (16 à 25 ans)

Cette section porte sur l'état de santé général, les blessures, les états chroniques et les limitations d'activités des jeunes. Les questions sont semblables aux questions sur la santé des enfants et des adultes. Des questions sur la structure du sommeil des jeunes y sont aussi posées. Le sommeil constitue un important indicateur de l'attitude du jeune face à son corps et de sa façon de prendre soin de sa personne. La durée des périodes de sommeil déclarées peut servir à déterminer si les jeunes parviennent à concilier le travail, les études, les activités bénévoles, les sports, entre autres.

Santé des jeunes (18 à 25 ans)

Cette section comprend des questions portant sur la taille, le poids et les comportements à risque, comme la consommation de tabac, d'alcool et de drogues.

L'échelle de dépression est comprise dans cette section (HHTYfS01). Compte tenu du nombre de transitions que doivent vivre les jeunes adultes, cette période peut se révéler extrêmement stressante et causer certaines émotions. Des recherches laissent entendre que de nombreuses personnes souffrent de dépression. Comme la santé mentale est aussi essentielle que la santé physique, il importe de recueillir des renseignements sur ces deux aspects.

Sentiments et comportements (18 et 19 ans et 22 et 23 ans)

L'objectif de cette section est de déterminer si le jeune connaît une personne qui s'est suicidée et s'il a lui-même envisagé sérieusement le suicide ou fait une tentative de suicide. Ces questions sont inspirées de l'*Adolescent Health Survey* menée en 1992 en Colombie-Britannique et sont semblables aux questions auxquelles doivent répondre les plus jeunes répondants dans les questionnaires autoadministrés.

D'autre part, certaines questions portent sur les comportements à risque, comme le vol, les bagarres, le fait de conduire avec facultés affaiblies et l'appartenance à un gang.

Relations (18 à 25 ans)

Au fur et à mesure que le jeune fait son entrée dans la vie adulte, la nature de ses relations avec un partenaire ou un conjoint constitue un facteur déterminant du bonheur et de la qualité de vie en général. Ces renseignements sont importants afin d'établir l'incidence de ses relations sur le jeune.

Les questions sur la santé sexuelle, les grossesses et le nombre d'amoureux sont posées aux jeunes de 18 à 25 ans, selon leur âge et leur état matrimonial.

Activités des jeunes (16 à 25 ans)

Alors qu'ils atteignent l'âge de jeunes adultes, ils se lancent parfois dans une multitude d'activités qui ne sont pas liées aux études. Il importe de mesurer ces activités pour comprendre comment elles peuvent favoriser un bon développement. La mesure dans laquelle les jeunes participent à des activités d'apprentissage continu et accordent de l'importance à l'autonomie intéresse particulièrement les chercheurs.

Cette section comprend des questions sur les activités physiques, les activités de lecture, l'écoute de la télévision, l'utilisation de l'ordinateur, l'engagement communautaire et la spiritualité. Ces questions sont des adaptations de celles posées aux adolescents plus jeunes. Le jeune répond à des questions différentes selon son âge.

On a également ajouté des questions pour les jeunes de 16 à 18 ans sur l'accès qu'ont les jeunes à un véhicule automobile et sur l'obtention ou non du permis de conduire. La conduite automobile représente un important « rite de passage » pour les jeunes de ce groupe d'âge, et il est important de recueillir des données à ce sujet.

Engagement politique (20 à 25 ans)

Une nouvelle série de questions sur l'engagement politique a été ajoutée au cycle 8. En plus de demander aux jeunes s'ils ont voté lors de précédentes élections, on leur demande s'ils croient que c'est leur devoir d'aller voter et à quelle fréquence ils discutent de l'actualité politique avec les autres. Le contenu a été suggéré par André Blais de l'Université de Montréal et adapté par l'équipe de projet de l'ELNEJ.

À propos de moi (18 et 19 ans et 22 et 23 ans)

L'objectif des questions au début de la section est de déterminer quel est le niveau d'estime de soi du jeune. Les réponses permettent d'établir un score global d'estime de soi (HAMYS01).

De plus, des questions supplémentaires portent sur les expériences douloureuses que le jeune peut avoir vécues pendant les deux dernières années, comme une rupture douloureuse avec un copain ou une copine, un problème grave à l'école ou au travail, le décès ou la maladie d'une personne proche, le divorce ou la séparation de ses parents, un important problème d'argent ou toute autre expérience difficile.

Ces questions sont semblables à celles des questionnaires que doivent remplir eux-mêmes les enfants plus jeunes.

Capacités (16 à 25 ans)

Les questions sur les capacités ont été ajoutées au cycle 8, mais apparaissaient précédemment dans les questionnaires papier autoadministrés du cycle 5. Ces questions aident à déterminer les capacités qu'ont les jeunes, selon eux, en matière d'utilisation d'un ordinateur, d'écriture, de lecture, de communication orale, de résolution de problèmes et de mathématiques.

Quotient émotionnel (20 et 21 ans et 24 et 25 ans)

L'échelle du quotient émotionnel a été mise au point par Reuven BarOn et James D.A. Parker. L'échelle permet d'évaluer dans quelle mesure les jeunes établissent des relations avec les autres à la maison, à l'école et au travail. L'intelligence émotionnelle englobe la capacité de déceler et de différencier ses propres sentiments et émotions de ceux des autres. On a posé 20 questions aux répondants sur leurs sentiments, leurs émotions et leurs perceptions. Cette version est semblable aux 15 questions du questionnaire que devaient remplir eux-mêmes les jeunes de 14 à 17 ans.

Soutien social des jeunes (18 à 25 ans)

Les questions de cette section varient selon l'âge du jeune et comportent l'échelle du Soutien social de la composante « Adulte » et elles visent à recueillir des renseignements sur le réseau de soutien social des jeunes. Les questions servent à déterminer quel est le soutien social reçu de la part de la famille et des amis, selon la perception du jeune. Cette section contient aussi des questions sur leur relation avec leur mère et leur père.

Formation des familles et fertilité (18 à 25 ans)

La cohorte originale de l'ELNEJ passe de l'adolescence à l'âge adulte, et des questions sur le mariage et les unions libres et les projets d'avoir des enfants deviennent plus pertinentes pour ce groupe. Dans le cycle 8, une nouvelle section de questions a été ajoutée pour recueillir l'information sur les facteurs influençant la transition des jeunes adultes vers le mariage et la parentalité. On demande aux jeunes célibataires s'ils désirent se marier ou vivre en union libre. Des questions sont également posées sur les facteurs pouvant influencer les jeunes à avoir un ou plusieurs enfants. On demande à ceux qui indiquent qu'il est peu probable qu'ils aient des enfants les raisons qui les poussent à donner cette réponse.

8.6 Questionnaires autoadministrés (14 à 17 ans)

Avec ces questionnaires, on vise à recueillir des renseignements directement auprès des jeunes sur divers aspects de leur vie afin de compléter les renseignements obtenus auprès des parents. Les questionnaires recueillent également auprès des jeunes de l'information que seuls ceux-ci peuvent fournir de façon fiable. Dans le cas des jeunes de 16 et 17 ans, certaines données sont toujours recueillies au moyen des questionnaires autoadministrés, même si ces jeunes fournissent leurs propres renseignements dans le cadre de l'interview assistée par ordinateur. On était d'avis que les jeunes se sentaient plus à l'aise de répondre à des questions délicates sur un questionnaire sur format papier que dans le cadre d'une rencontre avec un intervieweur.

Parents et amis (14 à 17 ans)

Cette section vise à déterminer dans quelle mesure les jeunes s'entendent avec les autres.

La section permet de recueillir des renseignements sur la portée et la qualité du réseau de soutien social du jeune, par exemple, le nombre d'amis intimes, le temps passé avec les amis et la présence d'une personne à qui le jeune peut se confier. Les questions varient selon l'âge des jeunes. Elles sont inspirées de l'Étude sur la santé des jeunes ontariens et du questionnaire de l'enfant de l'ELNEJ.

L'échelle des amis (HFFcS01) est établie à partir de ces questions.

Cette section contient également une mesure de l'intimité des jeunes de 14 et 15 ans. La question sur le nombre de fois que le jeune partage ses secrets et ses sentiments privés avec des amis proches a été adaptée de *Network of Relationships Inventory* de Furman et Buhrmester.

École (14 et 15 ans)

Cette section porte sur l'attitude des jeunes à l'égard de l'école, le rendement scolaire, l'importance de bons résultats, le sentiment de sécurité et d'acceptation à l'école, la perception qu'ont les jeunes du personnel enseignant pour ce qui est de l'équité et de l'assistance supplémentaire. On pose aux jeunes de 14 et 15 ans un ensemble de questions sur les activités parascolaires organisées par l'école, comme les sports et le théâtre. L'équipe de projet a adapté ces questions à partir de la *Child Health Survey* menée en l'Australie-Occidentale, la *Health Attitudes, Knowledge and Behaviours Study* réalisée dans les Territoires du Nord-Ouest, le *Marsh Self-Description Questionnaire*, et l'Enquête sur le comportement et la santé des élèves d'âge scolaire de l'Organisation mondiale de la santé (OMS).

L'attitude face à l'école peut exercer une influence importante sur les réalisations scolaires du jeune. Des recherches révèlent qu'une attitude négative à l'égard de l'école peut être associée à un rendement scolaire médiocre.

À propos de moi (14 à 17 ans)

Ces questions servent à déterminer l'estime de soi générale des jeunes. On calcule un score (HAMCS02) à partir des réponses à ces questions.

On pose aux jeunes des questions supplémentaires sur leurs sentiments quant à leur vie actuelle et future. Ces questions sont tirées de la *Child Health Survey* menée en l'Australie-Occidentale.

Il y a également une série de questions visant à mesurer « l'intelligence émotionnelle ». Ces 15 questions ont été conçues par Reuven BarOn et James D.A. Parker. Il s'agit de la version pour les jeunes des questions posées aux personnes de 20 et 21 ans et de 24 et 25 ans. Cette mesure a été choisie parce qu'elle permet d'évaluer les aptitudes sociales, personnelles et émotionnelles des répondants plutôt que leurs comportements.

On pose également aux jeunes des questions sur leurs expériences douloureuses, comme une rupture avec un copain ou une copine ou la mort d'un proche.

Sentiments et comportements (14 à 17 ans)

Liste de comportements (14 et 15 ans)

Cette section reprend la liste de comportements utilisée dans le questionnaire de l'enfant rempli par le parent dans le cadre de l'IAO. On y trouve des indicateurs des comportements suivants : troubles du comportement, hyperactivité, inattention, agressivité physique, agressivité indirecte, troubles émotifs, anxiété et comportement prosocial. On a calculé des scores pour ces comportements.

Comportements à risque (14 à 17 ans)

Ces questions sur les comportements à risque, comme le fait de passer la nuit hors du foyer sans permission, sont également tirées du questionnaire de l'enfant. On en a élargi la portée pour les jeunes plus âgés de manière à cerner les comportements qui deviennent plus fréquents à mesure que les jeunes vieillissent. L'équipe de projet a adapté ces questions à partir de la *National Longitudinal Survey of Youth* réalisée par la University of Ohio, la *Child Health Survey* menée en l'Australie-Occidentale et de questions fournies par Richard Tremblay de l'Université de Montréal.

Suicide (14 à 17 ans)

Cette section comporte des questions sur le suicide, notamment pour déterminer si le jeune connaît une personne qui s'est suicidée et s'il a lui-même envisagé sérieusement le suicide ou fait une tentative de suicide. Ces questions sont inspirées de l'*Adolescent Health Survey* menée en Colombie-Britannique en 1992.

Dépression (16 et 17 ans)

On interroge ces jeunes au sujet des sentiments de dépression au moyen de questions identiques à celles posées aux PMR et aux jeunes adultes de 18 à 25 ans. Un score (HHTCbS1b) est calculé à partir de ces questions.

Mes parents (14 à 17 ans)

Cette section a pour but d'évaluer la relation entre les jeunes et leurs parents ou tuteurs et ce, sous différents angles. Les questions sont conçues pour déterminer dans quelle mesure ils sont traités avec compréhension, équité et affection par chaque parent ou tuteur, de même que pour connaître les méthodes de résolution des conflits et le niveau de supervision des parents. On leur demande aussi quelle est leur impression au sujet de la relation de leurs parents ou tuteurs et de leurs aptitudes à résoudre les conflits.

Mes parents et moi (14 et 15 ans)

Trois échelles sont établies à partir de ces questions :

- 1) les soins prodigués par les parents (HPMCcS1);
- 2) le rejet par les parents (HPMCbS2b);
- 3) la surveillance par les parents (HPMCcS3).

Échelle de résolution des conflits (16 et 17 ans)

Ces questions sont identiques à celles posées aux parents de jeunes âgés de 12 à 15 ans. Dans le cas des jeunes de 16 et 17 ans, ces questions sont posées séparément au sujet de la mère et du père des jeunes. On a calculé deux scores à partir de ces questions :

- 1) l'échelle de résolution des conflits — mère (HPMCdS4) et
- 2) l'échelle de résolution des conflits — père (HPMCdS5).

Tabac, alcool et drogues (14 à 17 ans)

Cette section comporte des questions permettant de déterminer si le jeune fume ou consomme de l'alcool ou des drogues, ainsi que l'importance de sa consommation. La consommation de tabac, d'alcool ou de drogues est corrélée avec des comportements et un développement négatifs, notamment la délinquance et un rendement scolaire médiocre. Les questions varient selon l'âge.

Les questions sur le tabagisme sont adaptées de l'Enquête sur le tabagisme chez les jeunes, de l'Enquête sur le comportement et la santé des élèves d'âge scolaire réalisée par l'OMS et de la *Child Health Survey* menée en l'Australie-Occidentale.

Les questions sur la consommation d'alcool sont adaptées de la *Child Health Survey* réalisée en l'Australie-Occidentale et de questions fournies par Richard Tremblay de l'Université de Montréal.

Les questions sur la consommation de drogues et de substances toxicomano-gènes sont adaptées de la *Health Attitudes, Knowledge and Behaviours Study* réalisée dans les Territoires du Nord-Ouest.

Les questions se rapportant à la conduite avec facultés affaiblies et à la présence du jeune dans un véhicule conduit par un conducteur ayant pris de l'alcool ont été ajoutées pour les jeunes de 16 et 17 ans. Il s'agit d'importants comportements à risque dans ce groupe d'âge. Les questions sont adaptées de l'*Evaluation of School-Based Health Centers* réalisée en Caroline du Nord.

Activités (14 et 15 ans)

Cette section vise à déterminer l'importance de la participation des jeunes à des activités en dehors des heures de classe et l'emploi de leur temps libre. Parmi ces activités figurent les sports, les arts, la danse ou la musique, le scoutisme, l'emploi et le bénévolat. La lecture extrascolaire, l'utilisation d'un ordinateur et l'écoute de la télévision sont également prises en considération. Les activités sont également couvertes dans le questionnaire de l'enfant de l'IAO rempli par le parent pour les plus jeunes enfants.

Activités de lecture (14 et 15 ans)

Ces questions visent à mesurer à quelle fréquence les jeunes s'adonnent à des activités extrascolaires liées à la littérature, par exemple la fréquentation d'une bibliothèque ou la lecture. Ces questions sont semblables à celles posées à la PMR pour les plus jeunes enfants et aux jeunes de 16 à 25 ans au moyen du questionnaire de l'IAO.

Santé (14 à 17 ans)

On demande aux jeunes de fournir des renseignements sur leur taille et leur poids, les symptômes de stress, l'alimentation saine et les relations amoureuses. Les questions varient selon l'âge.

Les questions relatives aux indicateurs physiques du stress sont adaptées de l'Enquête sur le comportement et la santé des élèves d'âge scolaire réalisée par l'OMS.

Les questions sur le port de la ceinture de sécurité et d'un casque protecteur sont adaptées de la *Youth Risk Behaviour Survey* menée aux États-Unis. Ces questions ont aussi été utilisées dans l'*Adolescent Health Survey* réalisée en Colombie-Britannique en 1992.

On pose également des questions sur la puberté puisqu'elle est un indicateur important du développement physique. Ces questions portent sur des indicateurs physiologiques clés et sur la perception qu'ont les jeunes de leur propre puberté. Ces questions ont été fournies par Richard Tremblay de l'Université de Montréal et fondées d'après le *Puberty Development Scale (PDS)* développé par A. C. Petersen et L. Crockett.

L'équipe de projet a adapté les questions sur les relations amoureuses et l'activité sexuelle de divers questionnaires destinés aux jeunes, notamment ceux de la *Minnesota Adolescent Health Survey* ainsi que de l'*Adolescent Health Survey* réalisée en Colombie-Britannique en 1992.

Travail (14 et 15 ans)

On a posé aux jeunes des questions sur le travail pendant l'année scolaire et sur leur travail l'été précédent. On leur pose des questions sur leurs emplois, par exemple sur les heures de travail et si le travail réduit le temps qu'ils consacrent à leurs études.

Relations amoureuses (14 à 17 ans)

Dans cette section, on pose aux jeunes des questions sur leurs expériences amoureuses et sur leurs activités sexuelles. On pose également des questions sur l'utilisation de moyens de contraception et, pour les jeunes de 16 et 17 ans, sur les raisons de leur abstinence sexuelle ou de leur refus d'utiliser des moyens pour contrôler les naissances. Ces questions ont été développées par l'équipe de projet, en collaboration avec des experts d'enquêtes sur les jeunes comme l'*Adolescent Health Survey* réalisée en Colombie-Britannique en 1992 et la *Minnesota Adolescent Health Survey*.

9.0 Validation des échelles d'enquête

9.1 Validation des données d'échelle

9.1.1 Définition de l'échelle

Pour certains des concepts dont la mesure a été jugée importante pour l'Enquête longitudinale nationale sur les enfants et les jeunes (ELNEJ), on a décidé qu'il conviendrait mieux d'utiliser une échelle. Une échelle est constituée simplement d'un groupe de questions ou items dont les réponses regroupées permettent de mesurer certains concepts.

Par exemple, à partir du questionnaire de l'enfant, on a déterminé qu'il serait important d'évaluer trois comportements parentaux. L'échelle vise à mesurer l'interaction positive, l'inefficacité parentale et la cohérence des pratiques parentales.

9.1.2 Échelles et calculs

On calcule un score pour chaque facteur mesuré au moyen d'une échelle. Le score attribué à un facteur particulier peut servir à classer les personnes. Par exemple, aux échelles du rôle parental pour les enfants ayant les scores les plus élevés pour le facteur d'interaction positive, la personne la mieux renseignée (PMR) a déclaré avoir plus de rapports positifs avec l'enfant (ils riaient ensemble davantage, l'enfant était félicité plus souvent, entre autres). Le score lié à un facteur particulier est généralement fondé sur une série de questions, étant donné qu'une seule question ne peut mesurer le concept ou le facteur avec suffisamment de précision.

Au moment d'élaborer l'ELNEJ, lorsqu'on a pensé aux échelles particulières qui devraient servir à mesurer un concept donné, on a choisi des échelles utilisées dans le cadre d'autres études. On connaissait ainsi les propriétés psychométriques des mesures produites pour chaque échelle et on disposait de références complètes.

9.1.3 Évaluation des données de l'échelle

Dans plusieurs cas, on a modifié le libellé de certaines questions de l'échelle initiale et, dans certains cas, on a ajouté de nouvelles questions. Parfois, l'échelle utilisée n'avait jamais servi à des enquêtes sur les enfants au Canada ou n'avait été appliquée qu'à de très petits échantillons. Compte tenu de ces préoccupations et d'autres relatives aux conditions dans lesquelles se dérouleraient les interviews, on était d'avis que les structures des facteurs des échelles utilisées dans le cadre de l'ELNEJ pourraient différer de celles décrites dans les ouvrages spécialisés. Par conséquent, l'équipe de projet a décidé qu'il fallait procéder à une évaluation approfondie des données de l'échelle afin de s'assurer que les propriétés psychométriques présentées dans les autres études s'appliquaient aussi à l'ELNEJ.

L'analyse des données d'une échelle s'est faite en trois grandes étapes. Tout d'abord, une nouvelle analyse factorielle a été effectuée pour toutes les échelles afin de déterminer quels étaient les concepts ou facteurs inhérents à chacune. Puis, les scores des échelles ont été calculés selon cette structure factorielle. Enfin, des mesures de fiabilité ont été produites. Les procédures générales utilisées pour chacune de ces étapes sont décrites en détail dans les pages suivantes.

Note : Bon nombre des échelles ont été élaborées et validées au cycle 1. Au cours des cycles subséquents, on a imposé la structure factorielle émanant de l'analyse du cycle 1. Cette façon de procéder assure la cohérence des échelles au fil du temps et permet de faire des analyses longitudinales ainsi que les comparaisons transversales. Pour chaque échelle, une note indique à quel cycle la validation a été effectuée.

9.2 Analyse factorielle

9.2.1 Analyse factorielle des échelles

On a déterminé la structure factorielle de chacune des échelles à l'aide des données du premier cycle. La structure factorielle imposée aux échelles utilisées au cycle 1, puis de nouveau aux cycles ultérieurs de l'enquête, est le résultat d'analyses des données du cycle 1. Les résultats de l'analyse factorielle du premier cycle sont présentés dans le Guide de l'utilisateur des microdonnées du cycle 1.

1. Pour chaque échelle (et groupe d'âge, si l'échelle utilisait des questions différentes pour les divers groupes d'âge), l'échantillon de répondants a été divisé au hasard en deux sous-échantillons. On a procédé ainsi pour déterminer si des échantillons différents produiraient les mêmes résultats.
2. Pour chacun des sous-échantillons, on a effectué une analyse des composantes principales afin de déterminer le nombre de facteurs à extraire dans le cadre de l'analyse factorielle subséquente. En principe, on s'attendait à arriver au même nombre de facteurs que ceux décrits dans les ouvrages spécialisés. En pratique, toutefois, certaines échelles ont fait ressortir un nombre différent de facteurs parce que, dans certains cas, des facteurs se sont combinés, tandis que dans d'autres, de nouveaux facteurs sont apparus.
3. L'analyse factorielle a été effectuée pour chaque sous-échantillon, et la structure des facteurs ainsi que les items de chaque facteur ont été comparés.
4. Pour effectuer l'analyse factorielle, on a utilisé la valeur des items attribuée à chaque enfant du groupe d'âge approprié, multipliée par le poids normalisé d'enquête de l'enfant. On normalise le poids statistique d'un individu en divisant son poids (_WTCW01C)⁶ par la moyenne des poids de tous les individus. Ainsi, la somme des poids normalisés est égale à la taille de l'échantillon.
5. Une fois les structures factorielles analysées et les items inclus dans chaque facteur déterminés, on procède au calcul des scores. Pour produire les scores, on a retranché le chiffre un de la valeur de chacun des items afin que la plus faible valeur possible pour le score soit zéro. Un score de zéro indique l'absence de problème chez l'enfant pour tous les facteurs de l'échelle de comportement à l'exception du facteur « prosocial », pour lequel un score de zéro indique l'absence de comportement prosocial. Certains items ont été imputés. Les valeurs à imputer ont été déterminées grâce à une procédure (la procédure PRINQUAL du progiciel SAS) qui indique parmi les valeurs possibles d'un item celle qui semble la plus plausible pour un individu, étant donné son profil de réponse, le profil de réponse des autres individus dans l'échantillon ainsi que le nombre de facteurs considérés dans l'analyse.
6. On calcule le score pour chacun des facteurs de l'échelle en faisant la somme des valeurs des items qui composent les facteurs (en tenant compte des valeurs imputées). Le score est désigné « manquant » si trop de valeurs pour un item compris dans un facteur ne sont pas déclarées. Une valeur peut être manquante si le parent n'a pas voulu répondre ou s'il ne connaissait pas la réponse à la question.

6. Dans ce chapitre, on utilise un trait de soulignement « _ » au début du nom de chaque variable au lieu d'une lettre indiquant un certain cycle d'enquête. Par exemple, le nom de variable « HPRCS01 » au cycle 8 débute avec la lettre « H » au fichier de microdonnées et ici, on en fait référence comme suit : « _PRCS01 ».

9.2.2 Transformation des données au moyen de l'échelonnage optimal

L'analyse factorielle nécessite que les données aient des propriétés d'intervalle ou de ratio, c'est-à-dire que la distance entre chaque catégorie de réponse à la question soit la même. Par exemple, dans les échelles dont les choix de réponse sont : « Jamais », « Parfois », « Souvent » et « Toujours », on doit partir du principe que l'écart entre « Jamais » et « Parfois » est le même que celui entre « Parfois » et « Souvent », du point de vue du répondant. On estime que cela n'est pas toujours le cas pour les échelles utilisées dans l'ELNEJ.

Par conséquent, avant de procéder à l'analyse factorielle des échelles de l'ELNEJ, les données ont été transformées au moyen de l'échelonnage optimal. La méthode utilisée était l'une de celles proposées par Young et autres (Young, 1981), qui constitue une variante de la méthode de l'échelonnage optimal de Fisher. La méthode est présentée comme un moyen de transformer des données qui sont essentiellement nominales ou ordinales, en intervalles ou en ratios, afin que les techniques statistiques s'appliquant seulement à ces derniers puissent être utilisées.

9.3 Calcul des scores et imputation des réponses

9.3.1 Calcul des scores pour chaque facteur

Les résultats des analyses factorielles ont été utilisés pour déterminer quels sont les items « rattachés » à chaque facteur. L'étape suivante a consisté à calculer un score pour chaque facteur. À cette fin, on a additionné les valeurs de chaque question faisant partie du facteur. Dans certains cas, on a réechelonné des valeurs avant de calculer le score final. Voici un exemple du calcul des scores des facteurs.

9.3.2 Exemple de calcul des scores factoriels

Dans le cas de l'échelle relative au rôle parental dans le questionnaire de l'enfant, un des concepts qui est ressorti des analyses factorielles est celui de l'inefficacité parentale. L'analyse factorielle des données du cycle 1 a cerné sept questions s'y rattachant.

_PRCQ04	À quelle fréquence vous arrive-t-il d'être contrarié par une parole ou un geste que [votre enfant] n'est pas censé dire ou faire?
_PRCQ08	Lorsque vous parlez à [votre enfant] de son comportement, dans quelle proportion du temps le félicitez-vous?
_PRCQ09	Lorsque vous lui parlez de sa conduite, dans quelle proportion du temps le désapprouvez-vous?
_PRCgQ13	À quelle fréquence vous mettez-vous en colère lorsque vous punissez [votre enfant]?
_PRCgQ14	À quelle fréquence croyez-vous que le genre de punition que vous lui donnez dépend de votre humeur?
_PRCQ15	À quelle fréquence croyez-vous avoir de la difficulté à vous y prendre avec lui en général?
_PRCQ18	À quelle fréquence devez-vous le punir à maintes reprises pour la même chose?

Ces questions sont associées à des catégories de réponse de deux genres :

1. Jamais
2. Environ une fois par semaine ou moins
3. Quelques fois par semaine
4. Une ou deux fois par jour
5. Plusieurs fois par jour

1. Jamais
2. Moins de la moitié du temps
3. Environ la moitié du temps
4. Plus de la moitié du temps
5. Tout le temps

Au moment du calcul du score de ce facteur d'inefficacité parentale, les catégories ont été rééchelonnées de 0 à 4 (c'est-à-dire la catégorie « Jamais » correspondant à 0, la catégorie « Environ une fois par semaine ou moins/moins de la moitié du temps » à 1, et la catégorie « Plusieurs fois par jour/tout le temps » à 4). Pour le calcul du score, ces valeurs ont été additionnées pour les sept questions comprises dans le facteur avec un résultat de 0 à 28 pour l'inefficacité parentale. Un faible score de zéro correspond à l'absence de problème et un score élevé de 28 est l'indicateur le plus élevé d'un problème. Pour la plupart des scores calculés dans le cadre de l'ELNEJ, un score de zéro correspond à une absence de problème. Toutefois, il existe des exceptions à cette règle, lesquelles sont mentionnées dans la documentation relative à chaque échelle.

9.3.3 Poids factoriel négatif

Il convient de souligner que la deuxième question incluse dans le facteur d'inefficacité parentale, _PRCQ08 (Lorsque vous parlez à [votre enfant] de son comportement, dans quelle proportion du temps le félicitez-vous?), va dans la direction opposée de celle des autres questions. En réalité, la question a une incidence « négative » sur le facteur. Par conséquent, dans le calcul du score, les valeurs pour cette question ont été inversées, (c'est-à-dire « Tout le temps » correspondant à 0, « Plus de la moitié du temps » à 1, et « Jamais » à 4). Dans la documentation relative à chaque échelle, toutes les questions dont l'algorithme de notation a été inversé en raison d'un poids factoriel négatif comportent une indication à cet effet.

9.3.4 Codes de non-réponse

Lors du calcul du score pour chaque facteur, il est arrivé que l'une ou plusieurs des questions prises en compte dans le calcul comportent un code de non-réponse (« Ne sait pas », « Refus » ou « Non déclaré »). Si l'une des questions comportent un code de non-réponse, le score du facteur était établi à « Non déclaré ».

9.3.5 Items bruts

Il convient de souligner que tant les scores que les items bruts pour chaque échelle figurent dans le fichier de microdonnées. Cela permettra aux chercheurs d'opter pour d'autres structures factorielles s'ils le souhaitent. Dans le cas des items bruts, on a retenu les valeurs originales avant le rééchelonnage ou le renversement des valeurs.

9.4 Mesures de fiabilité des échelles

La fiabilité désigne la précision, la fidélité, la cohérence ou la capacité de reproduire une échelle donnée. En termes plus techniques, la fiabilité indique jusqu'à quel point les scores d'échelles sont exempts d'erreurs de mesure. Il existe de nombreuses façons de mesurer la fiabilité.

9.4.1 Coefficient alpha de Cronbach

Un des coefficients de fiabilité les plus courants est le coefficient alpha de Cronbach (Cronbach, 1951). Le coefficient alpha mesure la cohérence interne des items compris dans un facteur. Il est fondé sur la covariance moyenne de ceux-ci. On part du principe que les items compris dans un facteur comportent une corrélation positive les uns avec les autres, étant donné qu'ils visent à mesurer, jusqu'à un certain point, une entité ou un concept commun.

9.4.2 Interprétations du coefficient alpha de Cronbach

Le coefficient alpha de Cronbach donne lieu à plusieurs interprétations. Il peut être considéré comme la corrélation entre une échelle ou un facteur et toutes les autres échelles possibles comprenant le même nombre de questions, lesquelles pourraient être tirées d'un univers hypothétique de questions mesurant la caractéristique visée. Pour ce qui est du facteur de l'inefficacité parentale, par exemple, les sept questions incluses dans l'échelle peuvent être considérées comme un échantillon de l'univers de toutes les questions possibles. On aurait pu aussi demander aux parents : « Combien de fois élevez-vous la voix lorsque vous grondez votre enfant? » ou « Combien de fois menacez-vous de punir votre enfant sans le faire? ». Le coefficient alpha de Cronbach indique le niveau de corrélation prévu entre l'échelle utilisée et toutes les autres échelles possibles à sept questions qui mesurent le même phénomène.

Le coefficient alpha de Cronbach peut aussi être interprété comme le carré de la corrélation entre le score obtenu par une personne pour un facteur en particulier (score observé) et le score que cette personne aurait obtenu si on lui avait posé toutes les questions possibles de l'univers (score réel). Étant donné que le coefficient est interprété comme un coefficient de corrélation, sa valeur se situe entre 0 et 1.

On a démontré que le coefficient est en général la limite inférieure de fiabilité d'une échelle de n questions (Novick et Lewis, 1967). Autrement dit, dans la plupart des cas, le coefficient donne une estimation prudente de la fiabilité d'un score.

9.4.3 Qu'est-ce qu'un niveau satisfaisant de fiabilité?

Il est difficile de préciser un niveau unique pouvant s'appliquer à toutes les situations. Certains chercheurs sont d'avis que la fiabilité ne devrait pas être inférieure à 0,8 pour les échelles abondamment utilisées. À ce niveau, les corrélations sont peu touchées par l'erreur de mesure aléatoire. Cependant, il est souvent très coûteux, en argent et en temps, d'obtenir un coefficient plus fiable. Il convient de souligner que dans le cas de certains des facteurs pour lesquels des scores ont été calculés dans le cadre de l'ELNEJ, la fiabilité est inférieure à ce niveau. Le coefficient alpha de Cronbach figure dans la documentation relative à chaque score calculé. Les chercheurs peuvent déterminer eux-mêmes si le score est suffisamment fiable pour les fins auxquelles ils le destinent.

Enfin, il importe de noter que pour l'ELNEJ, le coefficient alpha de Cronbach de chaque score de facteur a été calculé au moyen de SAS. En règle générale, les coefficients calculés par SAS sont inférieurs à ceux calculés à l'aide de SPSS.

9.5 Échelles relatives dans le questionnaire du parent

Le reste du chapitre présente une description détaillée des sources des échelles de l'ELNEJ ainsi que tous les résultats de l'analyse factorielle et de l'analyse de fiabilité. Les modifications apportées aux échelles d'un cycle à l'autre sont également décrites. Les échelles sont présentées dans l'ordre dans lequel elles apparaissent dans le questionnaire.

9.5.1 Échelle de dépression

Objectifs et aperçu

L'échelle de dépression a été administrée à la PMR dans le questionnaire du parent. Les questions rattachées à cette échelle (_DPPQ12A à _DPPQ12L) constituent une version réduite de l'échelle de dépression (CES-D) contenant 20 questions et mise au point par L. S. Radloff du Centre d'étude épidémiologique du *National Institute of Mental Health* des États-Unis. Cette échelle sert à mesurer la fréquence des symptômes dans le grand public. On mesure la fréquence et la gravité des symptômes associés à la dépression pendant la semaine qui a précédé l'interview. L'échelle a été réduite à 12 questions par M. Boyle, de Chedoke-McMaster Hospital de la McMaster University.

L'objectif de cette échelle est de recueillir des renseignements sur la santé mentale des répondants, surtout en ce qui a trait aux symptômes de dépression. Plusieurs membres du groupe consultatif d'experts de l'ELNEJ ont laissé entendre que la meilleure façon de procéder est de mesurer un aspect particulier de la santé mentale de la PMR, au lieu de tenter de mesurer son état de santé général mentale. On a proposé que cette section soit centrée sur la dépression pour les raisons suivantes : il s'agit d'un état prévalent. En effet, il a été démontré que la dépression d'un parent a des répercussions sur les enfants; les recherches actuelles sur ce sujet sont généralement fondées sur des groupes de démonstration et non sur des échantillons de population, et on croit que l'adoption de politiques dans ce domaine peut améliorer la situation.

Aperçu des résultats de l'échelle de dépression de la PMR des enfants de 0 à 15 ans

Score	Items inclus	Taille de l'échantillon	Univers (âge en années)	Coefficient alpha de Cronbach (standardisé)
Dépression de la PMR (_DPPS01)	_DPPQ12A,			
	_DPPQ12B,			
	_DPPQ12C,			
	_DPPQ12D,	3 893	0 et 1	0,818
	_DPPQ12E,	4 153	2 et 3	0,803
	_DPPQ12F*,	3 888	4 et 5	0,821
	_DPPQ12G,			
	_DPPQ12H*,	3 332	6 et 7	0,850
	_DPPQ12I,	2 430	14 et 15	0,853
	_DPPQ12J*,			
	_DPPQ12K,			
	_DPPQ12L			

* Indique que les valeurs ont été inversées.

Source : Statistique Canada, Enquête longitudinale nationale sur les enfants et les jeunes.

9.5.2 Échelle de fonctionnement de la famille

Objectifs et aperçu

Les questions sur le fonctionnement de la famille (_FNHhQ1A à _FNHhQ1L) ont été développées par les chercheurs de Chedoke-McMaster Hospital de la McMaster University et sont grandement utilisées tant au Canada qu'à l'étranger. Il s'agit d'une échelle qui sert à mesurer différents aspects du fonctionnement de la famille (la résolution de problèmes, la communication, les rôles, l'engagement affectif, la sensibilité affective et la maîtrise du comportement).

La question _FNHhQ1M, tirée du suivi de l'Étude sur la santé des jeunes ontariens, a été ajoutée à l'échelle originale pour permettre de déterminer si la consommation d'alcool avait

des répercussions sur la dynamique globale de la famille. Elle n'a cependant pas été retenue dans l'analyse de l'échelle.

L'objectif de l'échelle est d'obtenir une évaluation globale du fonctionnement de la famille et une indication de la qualité des relations entre les parents. Pour cette raison et étant donné le nombre restreint de questions, on n'a pas tenté de mesurer les différents aspects du fonctionnement de la famille.

D'autres enquêtes ont montré que les relations entre les membres de la famille ont des répercussions considérables sur les enfants. Par exemple, les résultats de l'Étude sur la santé des jeunes ontariens indiquent un lien important entre le dysfonctionnement familial et certains problèmes de santé mentale chez l'enfant.

Administration de l'échelle de fonctionnement de la famille

L'échelle de fonctionnement de la famille a été administrée à la PMR dans le questionnaire du parent. L'échelle comporte 12 questions et chacune d'elles est associée à quatre catégories de réponse. Pour que la plus faible valeur du score soit 0, on a réduit la valeur des catégories de réponse de 1 lors du calcul du score. On a également inversé les catégories de réponse pour les questions ayant un poids factoriel négatif (_FNHhQ1A, _FNHhQ1C, _FNHhQ1E, _FNHhQ1G, _FNHhQ1I, et _FNHhQ1K). Le score total (_FNHhS01) peut donc varier de 0 à 36, un score élevé dénotant un dysfonctionnement de la famille.

Aperçu des résultats de l'échelle du fonctionnement de la famille des enfants de 0 à 15 ans

Score	Items inclus	Taille de l'échantillon	Univers (âge en années)	Coefficient alpha de Cronbach (standardisé)
Fonctionnement de la famille (_FNHhS01)	_FNHhQ1A*,			
	_FNHhQ1B,			
	_FNHhQ1C*,	3 853	0 et 1	0,914
	_FNHhQ1D,	4 127	2 et 3	0,910
	_FNHhQ1E*, _FNHhQ1F,	3 879	4 et 5	0,916
	_FNHhQ1G*,	3 314	6 et 7	0,907
	_FNHhQ1H, _FNHhQ1I*,	2 410	14 et 15	0,917
	_FNHhQ1J, _FNHhQ1K*, _FNHhQ1L			

* Indique que les valeurs ont été inversées.

Source : Statistique Canada, Enquête longitudinale nationale sur les enfants et les jeunes.

9.5.3 Échelle de sécurité du quartier

Objectifs et aperçu

L'échelle de sécurité du quartier vise à réunir des renseignements sur la satisfaction du répondant à l'égard de son quartier en tant que milieu de vie pour les enfants, y compris sur les perceptions par rapport au danger et aux problèmes ainsi qu'à la cohésion sociale ou « sentiment de voisinage ». Selon des recherches menées par Jacqueline Barnes du *Judge Baker Children's Centre* de la Harvard University de Boston, la peur du danger ressentie par les parents et les perceptions par rapport au désordre social dans le quartier ont une incidence sur leur attachement à ce quartier et sur leurs stratégies en matière de discipline.

Les questions _SFHhQ01, _SFHhQ02 et _SFHhQ5A à _SFHhQ6E couvrent la durée de résidence dans le quartier, la satisfaction à l'égard du quartier comme milieu de vie pour les enfants, la sécurité, la cohésion sociale et les problèmes du quartier. Il s'agit d'une version révisée des sections particulières du *Simcha-Fagan Neighbourhood Questionnaire* dont s'est servie Jacqueline Barnes dans ses travaux sur certains quartiers de Boston et de Chicago. Ce questionnaire a été modifié, en consultation avec Jacqueline Barnes, à la lumière de l'analyse

factorielle des sections. La question _SFHhQ03 sur l'engagement bénévole s'inspire d'une question de l'Enquête nationale sur la santé de la population.

Modifications apportées à la section relative au quartier d'un cycle à l'autre

Ces échelles ont été utilisées par intermittence durant les huit cycles de l'ELNEJ. Au cycle 1, trois échelles ont été créées : sécurité du quartier (_SFHhQ5A à _SFHhQ5C), voisinage (_SFHhQ6A à _SFHhQ6E) et problèmes du quartier. Au cycle 2, aucune question de la section relative au quartier n'a été posée aux répondants. Au cycle 3, la section se rapportant au quartier a été réintroduite, mais sans les questions _SFHhQ5A à _SFHhQ5C et ASFHQ07A à ASFHQ07F. Par ailleurs, les questions rattachées à l'échelle des problèmes du quartier dans le cycle 1 (ASFHQ07A à ASFHQ07F) n'ont pas été reprises après le cycle 4.

Aperçu des résultats des échelles de sécurité du quartier des enfants de 0 à 15 ans

Score	Items inclus	Taille de l'échantillon	Univers (âge en années)	Coefficient alpha de Cronbach (standardisé)
Sécurité du quartier (_SFHhS5)	_SFHhQ5A*,	3 885	0 et 1	0,732
	_SFHhQ5B*,	4 151	2 et 3	0,712
	_SFHhQ5C*	3 865	4 et 5	0,701
		3 325	6 et 7	0,735
		2 423	14 et 15	0,743
Voisinage (_SFHhS6)	_SFHhQ6A*,	3 225	0 et 1	0,905
	_SFHhQ6B*,	3 540	2 et 3	0,905
	_SFHhQ6C*,	3 367	4 et 5	0,913
	_SFHhQ6D*,	2 978	6 et 7	0,913
	_SFHhQ6E*	2 153	14 et 15	0,899

* Indique que les valeurs ont été inversées.

Source : Statistique Canada, Enquête longitudinale nationale sur les enfants et les jeunes.

9.5.4 Échelle du soutien social

Objectifs et aperçu

On a posé les questions sur le soutien social aux PMR ayant des enfants ou des jeunes de moins de 16 ans. L'échelle initiale comportait 24 items tirés du *Social Provisions Model* de Robert Weiss, lequel décrit six fonctions ou « dispositions » sociales distinctes qu'il est possible d'acquérir grâce aux relations avec autrui. Compte tenu de la taille de l'échelle et suivant les recommandations de M. Boyle de Chedoke-McMaster Hospital de la McMaster University, on a utilisé dans le cadre de l'enquête une version abrégée (à six items) inspirée du projet *Partir d'un bon pas pour un avenir meilleur* du gouvernement de l'Ontario. Cette échelle permet de mesurer l'encadrement (deux questions), les alliances fiables (deux questions) et l'attachement (deux questions). Par ailleurs, au cycle 1, quatre questions supplémentaires portant sur différents genres de soutien social (p. ex. le soutien religieux et les services communautaires) ont été ajoutées comme le suggérait Tom Hay. Cependant, ces questions ont été retirées au cycle 3 en raison du manque de variabilité des réponses. Des questions s'apparentant à celles proposées par Tom Hay ont été tirées de *Family Crisis Oriented Personal Evaluation Scales* (F-COPES) et intégrées aux cycles 4 et 5. Les F-COPES se fondent sur les dimensions de l'adaptation du *Resiliency Model of Family Adjustment and Adaptation* (McCubbin, Olson et Larsen, 1981). La mesure globale du soutien social comporte huit questions et couvre non seulement la portée du soutien social mais également la qualité de ce soutien.

Au cycle 2, on a abandonné toute la section se rapportant au soutien social, car on s'attendait à obtenir peu de variations temporelles pour ce qui est du soutien social dont jouissent les personnes, et aussi en raison d'inquiétudes quant au fardeau de réponse.

Modifications apportées à la section relative au soutien social parmi les cycles

Au cycle 4, les modifications suivantes ont été apportées à la section sur le soutien social et elles ont été conservées dans les cycles subséquents :

- Les six items initiaux utilisés au cycle 1 et au cycle 3 ont été conservés; cependant, les items _SPHQ02A à _SPHQ02D utilisés au cycle 1 ont été remplacés par des items tirés des F-COPES.
- Deux items additionnels tirés de l'échelle secondaire de l'intégration sociale susmentionnée (items _SPHh01H et _SPHh01I) ont également été ajoutés. Les questions sur l'intégration sociale sont importantes parce qu'elles permettent de mesurer le sentiment d'appartenance à un groupe ayant des intérêts, des préoccupations et des activités semblables, et donc, de mesurer un autre facteur du soutien social.
- On a ajouté quatre autres questions tirées des F-COPES ainsi qu'une question fondée sur le cadre des F-COPES, toutes centrées sur le raisonnement sous-jacent aux questions du cycle 1. Toutefois, les questions proposées diffèrent des simples réponses affirmatives et négatives, lesquelles n'indiquent pas de variabilité et sont plutôt associées aux catégories de réponse « Entièrement en désaccord », « En désaccord », « D'accord » et « Entièrement d'accord ».

Aperçu des résultats de l'échelle du soutien social des enfants de 0 à 15 ans

Score	Items inclus	Taille de l'échantillon	Univers (âge en années)	Coefficient alpha de Cronbach (standardisé)
Soutien social (_SPHhS01)	_SPHh01A*,	3 913	0 et 1	0,895
	_SPHh01B,	4 164	2 et 3	0,898
	_SPHh01C,	3 893	4 et 5	0,906
	_SPHh01D*,	3 337	6 et 7	0,909
	_SPHh01E*,	2 426	14 et 15	0,906
	_SPHh01F,			
	_SPHh01H, _SPHh01I*			

* Indique que les valeurs ont été inversées.

Source : Statistique Canada, Enquête longitudinale nationale sur les enfants et les jeunes.

9.5.5 Échelles de comportement

Objectif

L'objectif des échelles de comportement est d'évaluer des aspects du comportement des enfants de 2 à 11 ans. Cependant, l'échantillon ne comprenait pas d'enfants âgés de 8 à 11 ans au cycle 8, donc ces questions ont été posées aux enfants âgés de 2 à 7 ans.

Angoisse de séparation (enfants de 2 et 3 ans)

Questions _BECQ6CC, _BEC6DD1, _BEC8LL1, _BEC8PP1 et _BEC8TT1 de la *Child Behavior Checklist* (CBCL) d'Achenbach.

Opposition (enfants de 2 et 3 ans)

Questions _BECQ6G, _BECQ6R1, _BECQ8E1, _BECQ8T1, _BECQ8Z1 et _BECQ8J1 également tirées de la CBCL d'Achenbach.

Troubles de conduite (enfants de 4 à 7 ans)

Questions _BECQ6AA, _BECQ6FF, _BECd6JJ et _BECQ6NN tirées de l'Étude sur la santé des jeunes Ontariens (ESJO).

Hyperactivité (enfants de 2 à 7 ans)

Questions _BECQ6B, _BECQ6I, _BECQ6P et _BECQ6W tirées de l'ESJO, et questions _BECQ6QQ et _BEDQ8HH tirées de l'Enquête longitudinale de Montréal. Dans les cycles antérieurs, la question _BECQ6N était rattachée à ce concept. On a décidé de retirer cette question du cycle 4 et des cycles ultérieurs, les répondants la trouvant trop répétitive.

Troubles émotifs et anxiété (enfants de 2 à 7 ans)

Questions _BECdQ6F, _BECQ6K, _BECQ6Q, _BECQ6V, _BECQ6MM et _BECQ6RR tirées de l'ESJO. L'anxiété est mesurée par les questions de l'ELNEJ tirées des questions de l'ESJO se rapportant aux troubles émotifs (_BECdQ6F, _BECQ6Q, _BECQ6V et _BECQ6CC). Les cycles antérieurs comprenaient les questions _BECQ6Y et _BECQ6II. On a décidé de retirer ces deux questions du cycle 4 et des cycles ultérieurs.

Agressivité indirecte (enfants de 4 à 7 ans)

Questions _BECQ6J, _BECQ6R, _BECQ6Z, _BECQ6LL et _BECQ6TT tirées des travaux des chercheurs finlandais Lagerspetz, Bjornqvist et Peltonen.

Agressivité physique (enfants de 2 à 7 ans)

Question _BECQ6X tirée de l'Enquête longitudinale de Montréal et questions _BECQ6G, _BECQ6AA et _BECQ6NN tirées de l'ESJO.

Inattention (enfants de 2 à 7 ans)

Question _BECQ6P tirée de l'ESJO et question _BECQ6QQ tirée de l'Enquête longitudinale de Montréal.

Comportement prosocial (enfants de 6 et 7 ans)

Questions _BECQ6A, _BECQ6H, _BECQ6M, _BECQ6GG et _BECQ6OO tirées de l'ESJO, et questions _BECQ6D, _BECQ6U, _BECQ6BB, _BECQ6SS et _BECQ6UU tirées de l'Enquête longitudinale de Montréal; les quatre dernières questions proviennent d'une échelle mise au point par K. Weir et G. Duveen. Dans les cycles 1 à 3, ces questions ont été posées au sujet de tous les enfants de 4 à 11 ans. Dans le cycle 4, tous les enfants de 4 et 5 ans ont été exclus de cette échelle, on a plutôt posé ces questions dans la section se rapportant au comportement positif.

Aperçu des résultats des échelles de comportement des enfants de 2 et 3 ans

Score	Items inclus	Taille de l'échantillon	Univers (âge en années)	Coefficient alpha de Cronbach (standardisé)
Hyperactivité – inattention (_BECeS01)	_BECQ6B, _BECQ6I, _BECQ6P, _BECQ6W, _BECQ6QQ, _BEDQ8HH	4 292	2 et 3 ans	0,727
Troubles émotifs – anxiété (_BECdS03)	_BECdQ6F, _BECQ6K, _BECQ6Q, _BECQ6V, _BECQ6MM, _BECQ6RR	4 316	2 et 3 ans	0,638
Agressivité physique – opposition (_BECS04)	_BECQ6G, _BECQ6X, _BECQ6NN, _BECQ6R1, _BECQ8E1, _BECQ8T1, _BECQ8Z1, _BECQ8J1	4 277	2 et 3 ans	0,723
Angoisse de séparation (_BECS05)	_BECQ6CC, _BEC6DD1, _BEC8LL1, _BEC8PP1, _BEC8TT1	4 317	2 et 3 ans	0,579

Source : Statistique Canada, Enquête longitudinale nationale sur les enfants et les jeunes.

Aperçu des résultats des échelles de comportement des enfants de 4 à 7 ans

Score	Items inclus	Taille de l'échantillon	Univers (âge en années)	Coefficient alpha de Cronbach (standardisé)
Hyperactivité – inattention (_BECdS06)	_BECQ6B, _BECQ6I, _BECQ6P, _BECQ6S, _BECQ6W, _BECQ6QQ, _BEDQ8HH	4 059 3 412	4 et 5 ans 6 et 7 ans	0,809 0,817
Troubles émotifs – anxiété (_BECdS08)	_BECdQ6F, _BECQ6K, _BECQ6Q, _BECQ6V, _BECQ6MM, _BECQ6RR, _BECQ6CC	4 072 3 422	4 et 5 ans 6 et 7 ans	0,691 0,737
Agressivité physique – problèmes de conduite (_BECdS09)	_BECQ6G, _BECQ6X, _BECQ6AA, _BECQ6FF, _BECd6JJ, _BECQ6NN	4 067 3 425	4 et 5 ans 6 et 7 ans	0,774 0,735
Agressivité indirecte (_BECS10)	_BECQ6J, _BECQ6R, _BECQ6Z, _BECQ6LL, _BECQ6TT	3 973 3 217	4 et 5 ans 6 et 7 ans	0,632 0,696
Comportement prosocial (_BECdS07)	_BECQ6A, _BECQ6D, _BECQ6H, _BECQ6M, _BECQ6U, _BECQ6BB, _BECQ6GG, _BECQ6OO, _BECQ6SS, _BECQ6UU	3 138	6 et 7 ans	0,831

Source : Statistique Canada, Enquête longitudinale nationale sur les enfants et les jeunes.

9.5.6 Échelle du développement moteur et social

Section sur le développement moteur et social

La section sur le développement moteur et social du questionnaire de l'enfant s'adresse aux enfants de 0 à 3 ans. L'objectif de cette section est de mesurer le développement moteur, social et cognitif des jeunes enfants. Une échelle composée de 48 questions au total (_MSCQ01 à _MSCQ48) a été utilisée à cette fin. Chaque enfant, selon son âge en mois, se fait poser 15 de ces 48 questions.

L'échelle relative au développement moteur et social

L'échelle relative au développement moteur et social (DMS) a été mise au point par Gail Poe du *National Center for Health Statistics* des États-Unis. Tous les jeunes enfants se voient poser 15 questions qui varient selon l'âge de l'enfant et qui mesurent certains aspects de leur développement moteur, social et cognitif, allant de leur naissance à l'âge de 3 ans. Chaque question vise à déterminer si un enfant peut ou non effectuer une tâche donnée. L'échelle a été utilisée pour la *National Longitudinal Survey of Youth* réalisée aux États-Unis et pour la *National Child Development Survey* menée en Angleterre.

Le tableau suivant indique les questions qui ont été posées à chaque groupe d'âge.

Âge (en mois)	Questions du DMS
0 à 3	_MSCQ01 à _MSCQ15
4 à 6	_MSCQ08 à _MSCQ22
7 à 9	_MSCQ12 à _MSCQ26
10 à 12	_MSCQ18 à _MSCQ32
13 à 15	_MSCQ22 à _MSCQ36
16 à 18	_MSCQ26 à _MSCQ40
19 à 21	_MSCQ29 à _MSCQ43
22 à 47	_MSCQ34 à _MSCQ48

Scores bruts

Un score brut a été calculé pour chaque enfant en additionnant le nombre de réponses « Oui » à chaque question de l'échelle (_MSCS01). Même s'il existait divers ensembles de questions appropriées à l'âge de l'enfant exprimé en mois, des différences ont été observées au moment de la comparaison des scores à l'intérieur des groupes d'âge. Par exemple, un ensemble précis de questions s'adressaient aux enfants de quatre à six mois. On a découvert que les enfants âgés de six mois avaient en moyenne des scores plus élevés que ceux de quatre mois. C'est pourquoi on a décidé de produire des scores normalisés. Ces scores, calculés pour chaque âge en mois, vont permettre de faire des comparaisons d'un âge à l'autre. Tous les enfants de 3 ans et moins ont obtenu un score normalisé basé sur les données du cycle 1 et un score normalisé basé sur les données du cycle 8.

Scores normalisés basés sur les normes du cycle 8

Un score normalisé est calculé pour chaque enfant de 4 à 47 mois. La normalisation est faite pour chaque groupe d'âge exprimé en mois. Pour chaque groupe d'âge en mois, la moyenne et l'écart type du score brut ont été calculés afin de produire un score normalisé avec une moyenne de 0 et un écart type de 1. Ce score a ensuite été ajusté de façon à ce que la moyenne des scores du DMS soit de 100 et que l'écart type soit de 15. Par conséquent, les enfants de 4 mois ont un score du DMS moyen de 100, les enfants de 5 mois ont un score du DMS moyen de 100 et ainsi de suite pour les enfants de 47 mois.

Une fois ces scores calculés, on a identifié les enfants dont le score était éloigné de plus de 3 fois de l'écart type par rapport à la moyenne (scores inférieurs à 55 ou supérieurs à 145) et les scores ont été recalculés en enlevant ces enfants. On a considéré ces enfants comme ayant des valeurs aberrantes et donc, ils ne sont pas représentatifs des enfants dans leur groupe d'âge. Par conséquent, la moyenne des scores du DMS sur le fichier de données selon l'âge exprimé en mois, peut ne pas correspondre exactement à 100. L'utilisation du score normalisé (_MSCdS03) permet de comparer les scores des enfants du groupe d'âge de 4 à 47 mois sans devoir procéder à un ajustement selon l'âge.

Ce score n'a pas été calculé pour les enfants de 0 à 3 mois, car il n'y a pas suffisamment d'enfants répondants (dans ce groupe d'âge) pour permettre d'établir une norme.

Scores normalisés basés sur les normes du cycle 1

Un deuxième score normalisé (_MSCS02) a été calculé pour les enfants de 0 à 47 mois. Ce score a été calculé de la même façon que le score mentionné ci-dessus, sauf que les normes ont été calculées d'après les données du cycle 1 et ensuite appliquées aux données du cycle 8.

De façon générale, il n'y a pas de différences majeures entre les scores basés sur les normes établies au cycle 1 par rapport à ceux obtenus en utilisant les normes établies au cycle 8. Le score calculé en utilisant les normes établies au cycle 1 devrait être utilisé pour comparer les scores d'un cycle à l'autre. Ce score est disponible pour tous les cycles de données.

Les questions de l'échelle du développement moteur et social n'ont pas changé tout au long des huit cycles de l'ELNEJ, mais on a fait certains changements dans le calcul des scores finaux. Pour obtenir plus de précisions concernant ces changements, veuillez consulter l'annexe portant sur les changements apportés aux cycles précédents du Guide de l'utilisateur des microdonnées du cycle 4.

9.5.7 Échelles du rôle parental

Objectifs et aperçu

L'objectif de cette série de questions est de mesurer certaines pratiques parentales. Plus précisément, deux échelles ont été utilisées. La première a été divisée en trois sous-échelles conçue afin de mesurer l'interaction positive, l'inefficacité et la cohérence des parents envers les enfants. La deuxième échelle vise à mesurer les pratiques de la rationalité parentale.

Les items du questionnaire de l'enfant qui ont servi à mesurer ces aspects des pratiques parentales sont indiqués dans les paragraphes qui suivent. On a effectué des analyses factorielles complètes des questions portant sur le rôle parental afin d'en évaluer les propriétés psychométriques pour chacune des sous-échelles pour la population de l'ELNEJ.

Les questions _PRCQ01 à _PRCQ18 et _PRCQ21 à _PRCQ24 sur l'interaction positive, sur l'inefficacité et sur la cohérence ont été fournies par M. Boyle de Chedoke-McMaster Hospital de la McMaster University à partir des travaux de Ken Dodge (Vanderbilt University) et d'une adaptation de la *Parent Practices Scale* de Strayhorn et Weidman.

Aperçu des résultats des échelles du rôle parental des enfants de 0 et 1 an

Score	Items inclus	Taille de l'échantillon	Univers (âge en années)	Coefficient alpha de Cronbach (standardisé)
Interaction positive (_PRCS01)	_PRCQ01, _PRCQ02, _PRCQ03, _PRCQ06, _PRCQ07	3 945	0 et 1 an	0,622
Inefficacité parentale (_PRCS02)	_PRCQ04, _PRCQ05	3 954	0 et 1 an	0,343

Source : Statistique Canada, Enquête longitudinale nationale sur les enfants et les jeunes.

Aperçu des résultats des échelles du rôle parental des enfants de 2 à 7 ans

Score	Items inclus	Taille de l'échantillon	Univers (âge en années)	Coefficient alpha de Cronbach (standardisé)
Interaction positive (_PRCS03)	_PRCQ01, _PRCQ02, _PRCQ03, _PRCQ06, _PRCQ07	4 267	2 et 3	0,679
		4 012	4 et 5	0,677
		3 396	6 et 7	0,682
Inefficacité parentale (_PRCgS04)	_PRCQ04, _PRCQ08*, _PRCQ09, _PRCgQ13, _PRCgQ14, _PRCQ15, _PRCQ18	4 225	2 et 3	0,607
		3 965	4 et 5	0,630
		3 350	6 et 7	0,672
Cohérence des pratiques parentales (_PRCgS05)	_PRCe10, _PRCe11, _PRCgQ12*, _PRCgQ16*, _PRCgQ17*	4 240	2 et 3	0,723
		3 989	4 et 5	0,650
		3 378	6 et 7	0,520
Rationalité parentale (_PRCS06)	_PRCQ21, _PRCQ22*, _PRCQ23, _PRCQ24*	4 243	2 et 3	0,562
		3 989	4 et 5	0,528
		3 383	6 et 7	0,568

* Indique que les valeurs ont été inversées.

Source : Statistique Canada, Enquête longitudinale nationale sur les enfants et les jeunes.

Aperçu des résultats de l'échelle du rôle parental des enfants de 14 et 15 ans

Score	Items inclus	Taille de l'échantillon	Univers (âge en années)	Coefficient alpha de Cronbach (standardisé)
Résolution des conflits (_PRCbS09)	_PRCb30A*, _PRCb30B, _PRCb30C, _PRCb30D, _PRCb30E, _PRCb30F, _PRCb30G, _PRCb30H*	2 434	14 et 15	0,746

* Indique que les valeurs ont été inversées.

Source : Statistique Canada, Enquête longitudinale nationale sur les enfants et les jeunes.

9.5.8 Échelle relative dans les étapes du développement

La composante Étapes du développement a été produite pour tous les enfants de 3 à 47 mois ayant un âge effectif de 3 ans et moins. Les questions ont été regroupées en quatre catégories énumérées ci-dessous, chaque répondant obtenant un score dans l'intervalle de 0 à 60. Pour cette mesure, un score élevé indique que l'enfant se situe dans un intervalle normal pour son groupe d'âge ou supérieur à son groupe d'âge. Pour obtenir plus de renseignements sur cette mesure, veuillez consulter le chapitre 8.0 du présent guide ou communiquer avec Brookes Publishing Co. and Health Professions Press afin d'obtenir une copie de chaque question.

Facteur	Score	Intervalle des scores
Score en résolution de problèmes	_AGCdS01	0 à 60
Score personnel	_AGCdS02	0 à 60
Score en communication	_AGCdS03	0 à 60
Score en motricité fine	_AGCdS04	0 à 60

Source : Statistique Canada, Enquête longitudinale nationale sur les enfants et les jeunes.

9.6 Échelles relatives dans le questionnaire des jeunes

9.6.1 Échelle de dépression

Aperçu

L'échelle de dépression utilisée pour mesurer la dépression chez la PMR a également été appliquée aux jeunes de 16 ans et plus. En ce qui concerne les jeunes de 16 et 17 ans, les questions figuraient dans le questionnaire sur format papier qu'ils devaient eux-mêmes remplir. Pour les jeunes de 18 ans et plus, les questions ont été posées dans le contexte de l'application d'Interview assistée par ordinateur (IAO). La structure factorielle de l'échelle appliquée aux PMR a également été imposée à l'échelle des jeunes.

Le score total (_HTCbS1B et _HTYfS01) peut donc varier de 0 à 36, un score élevé dénotant la présence de symptômes de dépression.

Aperçu des résultats de l'échelle autoadministrée de dépression relative aux jeunes de 16 à 25 ans

Score	Items inclus	Taille de l'échantillon	Univers (âge en années)	Coefficient alpha de Cronbach (standardisé)
Dépression (_HTCbS1B)	_FBCd10A, _FBCd10B, _FBCd10C, _FBCd10D, _FBCd10E, _FBCd10F*, _FBCd10G, _FBCd10H*, _FBCd10I, _FBCd10J*, _FBCd10K, _FBCd10L	1 358	16 et 17	0,833
Dépression (_HTYfS01)	_HTYf14A, _HTYf14B, _HTYf14C, _HTYf14D, _HTYf14E, _HTYf14F*, _HTYf14G, _HTYf14H*, _HTYf14J, _HTYf14K*, _HTYf14M, _HTYf14N	1 622 1 354 1 464 1 460	18 et 19 20 et 21 22 et 23 24 et 25	0,799 0,805 0,816 0,832

* Indique que les valeurs ont été inversées.

Source : Statistique Canada, Enquête longitudinale nationale sur les enfants et les jeunes.

9.6.2 Échelle de structure du quartier

Objectifs et aperçu

L'échelle de structure du quartier vise à réunir des renseignements sur la satisfaction du répondant à l'égard de son quartier, y compris sur les perceptions du niveau de risque et de problèmes ainsi que de cohésion sociale ou « sentiment de voisinage ». Ces questions sont posées aux jeunes de 16 et 17 ans dans le questionnaire des jeunes.

Il s'agit d'une version révisée des sections particulières du *Simcha-Fagan Neighbourhood Questionnaire* dont s'est servie Jacqueline Barnes dans ses travaux sur des quartiers de Boston et de Chicago.

Aperçu des résultats de l'échelle de structure du quartier des jeunes de 16 et 17 ans

Score	Items inclus	Taille de l'échantillon	Univers (âge en années)	Coefficient alpha de Cronbach (standardisé)
Structure du quartier (_ACYdS01)	_ACYd13A*, _ACYd13B*, _ACYd13C*, _ACYd13D*, _ACYd13F*, _ACYd13G*	1 495	16 et 17	0,809

* Indique que les valeurs ont été inversées.

Source : Statistique Canada, Enquête longitudinale nationale sur les enfants et les jeunes.

9.6.3 Échelle de l'estime globale de soi

Le but de l'échelle de l'estime globale de soi est de mesurer l'estime globale de soi du jeune. Les questions formant l'échelle de l'estime de soi sont posées aux jeunes de 14 à 17 ans dans le questionnaire sur format papier qu'ils doivent eux-mêmes remplir et aux jeunes de 18 et 19 ans et de 22 et 23 ans dans le questionnaire de l'application d'Interview sur place assistée par ordinateur (IPAO). Les questions sur l'estime globale de soi sont tirées de l'échelle générale d'autoévaluation du questionnaire d'autodescription de Marsh, mises au point par H. W. Marsh.

Aperçu des résultats de l'échelle de l'estime globale de soi des jeunes de 14 à 23 ans

Score	Items inclus	Taille de l'échantillon	Univers (âge en années)	Coefficient alpha de Cronbach (standardisé)
Estime globale de soi (_AMcS02)	_AMCQ01A, _AMCQ01B, _AMCQ01C, _AMCQ01D	2 091	14 et 15	0,830
		1 373	16 et 17	0,817
Estime globale de soi (_AMYfS01)	_AMYfQ01, _AMYfQ02, _AMYfQ03, _AMYfQ04	1 624	18 et 19	0,764
Estime globale de soi (_AMYfS01)	_AMYfQ01, _AMYfQ02, _AMYfQ03, _AMYfQ04	1 465	22 et 23	0,778

Source : Statistique Canada, Enquête longitudinale nationale sur les enfants et les jeunes.

9.6.4 Échelle relative au quotient émotionnel

Mis au point par Reuven BarOn et James D. A. Parker, le test de quotient émotionnel version adultes (EQ-i) et le test de quotient émotionnel version jeunes (EQ-i : YV) sont des mesures de l'intelligence émotionnelle. Ces mesures comprennent cinq grandes dimensions : intrapersonnelle, interpersonnelle, adaptabilité, gestion du stress et humeur générale. Après l'avoir comparé à d'autres mesures possibles, on a finalement choisi ce test pour différentes raisons. D'abord, les questions sont généralement très positives et elles sont courtes et simples. Elles permettent aussi d'évaluer les capacités sociales, personnelles et émotionnelles du répondant plutôt que ses comportements.

Les questions servant à mesurer l'EQ-i : YV ont été posées aux jeunes de 14 à 17 ans et celles servant à mesurer l'EQ-i, aux jeunes de 20 et 21 et de 24 et 25 ans. Les premières se trouvaient dans le questionnaire sur format papier que les jeunes de 14 à 17 ans devaient remplir eux-mêmes, tandis que les secondes, pour les 20 et 21 et les 24 et 25 ans, se trouvaient dans le questionnaire de l'application d'IPAQ.

Avant de calculer le score de quotient émotionnel (EQ-i), les valeurs des catégories de réponse ont été inversées pour la dimension intrapersonnelle et la gestion du stress. Pour calculer l'EQ-i : YV, seules les valeurs de la gestion du stress ont été inversées. On a ensuite soustrait le chiffre un de chacun des items afin d'obtenir un score de zéro. Une fois ces deux étapes réalisées, les valeurs ont été additionnées pour chacune des dimensions, et cinq scores ont été créés.

Les scores les plus faibles pour une échelle particulière représentent la partie négative du continuum de l'intelligence émotionnelle, tandis que les scores plus élevés représentent la partie positive. Par exemple, un score total de 33 signifie que la personne a une intelligence émotionnelle bien supérieure à celle qui obtient un score de 12 pour la même échelle. La ligne de séparation entre un score (a) extrêmement élevé, (b) élevé, (c) moyen, (d) faible et (e) très faible est essentiellement un écart type de +/- 1 de la valeur moyenne pour une échelle particulière.

Les scores normalisés ne sont pas fournis dans le fichier principal. Cependant, le tableau ci-dessous présente un guide d'interprétation des scores normalisés pour les utilisateurs de données qui désirent créer leurs propres scores. Les scores normalisés EQ-i et EQ-i : YV placent les valeurs moyennes à 100 et chaque écart type à 15. Cependant, on remarquera qu'il y a 10 points liées aux valeurs moyennes pour permettre de différencier les descripteurs dans le tableau ci-dessous.

Guide d'interprétation des scores normalisés des échelles EQ-i et EQ-i : YV	
130 et plus	Extrêmement élevé (capacité émotionnelle/sociale exceptionnellement développée)
120 à 129	Très élevé (capacité émotionnelle/sociale extrêmement développée)
110 à 119	Élevé (capacité émotionnelle/sociale bien développée)
90 à 109	Moyen (capacité émotionnelle/sociale suffisante)
80 à 89	Faible (capacité émotionnelle/sociale sous-développée)
70 à 79	Très faible (capacité émotionnelle/sociale extrêmement sous-développée)
Moins de 70	Extrêmement faible (capacité émotionnelle/sociale exceptionnellement déficiente)

Aperçu des résultats des échelles relatives au EQ-i : YV des jeunes de 14 à 17 ans

Score	Items inclus	Taille de l'échantillon	Univers (âge en années)	Coefficient alpha de Cronbach (standardisé)
Aptitudes intrapersonnelles (_EQYeS06)	_AMCe25A, _AMCe25F, _AMCe25K	2 089 1 372	14 et 15 16 et 17	0,847 0,837
Aptitudes interpersonnelles (_EQYeS07)	_AMCe25B, _AMCe25G, _AMCe25L	2 082 1 370	14 et 15 16 et 17	0,643 0,663
Aptitudes de gestion du stress (_EQYeS08)	_AMCe25C*, _AMCe25H*, _AMCe25M*	2 078 1 366	14 et 15 16 et 17	0,670 0,708
Aptitudes d'adaptabilité (_EQYeS09)	_AMCe25D, _AMCe25I, _AMCe25N	2 085 1 371	14 et 15 16 et 17	0,773 0,782
Humeur générale (_EQYeS10)	_AMCe25E, _AMCe25J, _AMCe25O	2 082 1 370	14 et 15 16 et 17	0,654 0,646
Quotient émotionnel (EQ4) (_EQYeS04)	_AMCe25A, _AMCe25F, _AMCe25K, _AMCe25B, _AMCe25G, _AMCe25L, _AMCe25C*, _AMCe25H*, _AMCe25M*, _AMCe25D, _AMCe25I, _AMCe25N	2 052 1 349	14 et 15 16 et 17	0,697 0,716
Quotient émotionnel (EQ5) (_EQYeS05)	_AMCe25A, _AMCe25F, _AMCe25K, _AMCe25B, _AMCe25G, _AMCe25L, _AMCe25C*, _AMCe25H*, _AMCe25M*, _AMCe25D, _AMCe25I, _AMCe25N, _AMCe25E, _AMCe25J, _AMCe25O	2 046 1 343	14 et 15 16 et 17	0,774 0,776

* Indique que les valeurs ont été inversées.

Source : Statistique Canada, Enquête longitudinale nationale sur les enfants et les jeunes.

Aperçu des résultats des échelles relatives au EQ-i des jeunes de 20 et 21 et de 24 et 25 ans

Score	Items inclus	Taille de l'échantillon	Univers (âge en années)	Coefficient alpha de Cronbach (standardisé)
Aptitudes interpersonnelles (_EQYfs11)	_EQYfQ01, _EQYfQ06, _EQYfQ11, _EQYfQ16	1 358 1 465	20 et 21 24 et 25	0,691 0,712
Aptitudes intrapersonnelles (_EQYfs12)	_EQYfQ02*, _EQYfQ07*, _EQYfQ12*, _EQYfQ17*	1 341 1 442	20 et 21 24 et 25	0,590 0,589
Aptitudes de gestion du stress (_EQYfs13)	_EQYfQ03*, _EQYfQ08*, _EQYfQ13*, _EQYfQ18*	1 358 1 460	20 et 21 24 et 25	0,700 0,702
Aptitudes d'adaptabilité (_EQYfs14)	_EQYfQ04, _EQYfQ09, _EQYfQ14, _EQYfQ19	1 352 1 459	20 et 21 24 et 25	0,639 0,703
Humeur générale (_EQYfs15)	_EQYfQ05, _EQYfQ10, _EQYfQ15, _EQYfQ20	1 355 1 461	20 et 21 24 et 25	0,650 0,707
Quotient émotionnel (EQ4) (_EQYfs16)	_EQYfQ02*, _EQYfQ07*, _EQYfQ12*, _EQYfQ17*, _EQYfQ01, _EQYfQ06, _EQYfQ11, _EQYfQ16, _EQYfQ03*, _EQYfQ08*, _EQYfQ13*, _EQYfQ18*, _EQYfQ04, _EQYfQ09, _EQYfQ14, _EQYfQ19,	1 334 1 435	20 et 21 24 et 25	0,792 0,787
Quotient émotionnel (EQ5) (_EQYfs17)	_EQYfQ02*, _EQYfQ07*, _EQYfQ12*, _EQYfQ17*, _EQYfQ01, _EQYfQ06, _EQYfQ11, _EQYfQ16, _EQYfQ03*, _EQYfQ08*, _EQYfQ13*, _EQYfQ18*, _EQYfQ04, _EQYfQ09, _EQYfQ14, _EQYfQ19, _EQYfQ05, _EQYfQ10, _EQYfQ15, _EQYfQ20	1 333 1 434	20 et 21 24 et 25	0,838 0,843

* Indique que les valeurs ont été inversées.

Source : Statistique Canada, Enquête longitudinale nationale sur les enfants et les jeunes.

Voici de brèves définitions de ce qui est mesuré à l'aide des 5 échelles composites et des 15 sous-échelles (le EQ-i : YV ne comprend que les 5 échelles composites, tandis que le EQ-i comprend les 5 échelles et les 15 sous-échelles qui figurent ci-après sous chacune des échelles composites).

1) **Aptitudes intrapersonnelles (conscience de soi et extériorisation)**

Les aptitudes intrapersonnelles comprennent les sous-composantes suivantes, lesquelles déterminent la capacité de prendre conscience de soi, de comprendre ses forces et ses faiblesses et d'exprimer ses pensées et ses sentiments de façon non destructive :

- Estime de soi : capacité de se connaître, de se comprendre et de s'accepter.

- Conscience de ses propres émotions : capacité de prendre conscience de ses émotions et de les comprendre.
- Affirmation de soi : capacité d'exprimer ses sentiments et de s'extérioriser de façon non destructive.
- Indépendance : capacité d'être autonome et sans dépendance émotive à l'égard des autres.
- Réalisation de soi : capacité de se fixer des buts et de se motiver pour les atteindre.

2) **Aptitudes interpersonnelles (conscience sociale et relations interpersonnelles)**

Les aptitudes interpersonnelles comprennent les sous-composantes suivantes, lesquelles déterminent la capacité de prendre conscience des émotions, des sentiments et des besoins des autres et d'établir et d'entretenir des relations coopératives, constructives et mutuellement satisfaisantes :

- Empathie : capacité de savoir et de comprendre comment les autres se sentent.
- Responsabilité sociale : capacité de s'identifier à son groupe social et d'avoir un sentiment d'appartenance à ce groupe.
- Relations interpersonnelles : capacité d'établir et d'entretenir des relations mutuellement satisfaisantes avec les autres.

3) **Aptitudes de gestion du stress (gestion et contrôle des émotions)**

Les aptitudes de gestion du stress comprennent les sous-composantes suivantes, lesquelles déterminent la capacité de gérer ses émotions de façon efficace et constructive et non le contraire :

- Tolérance au stress : capacité de gérer ses émotions de façon efficace et constructive.
- Contrôle des impulsions : capacité de contrôler ses émotions de façon efficace et constructive.

4) **Aptitudes d'adaptabilité (gestion du changement)**

Les aptitudes d'adaptabilité comprennent les sous-composantes suivantes, lesquelles déterminent la capacité de gérer le changement, c'est-à-dire de faire face avec réalisme et souplesse à la situation immédiate et de résoudre efficacement les problèmes à mesure qu'ils surviennent :

- Épreuve de réalité : capacité de valider ses sentiments et ses pensées en les confrontant avec la réalité extérieure.
- Souplesse : capacité de composer et de s'adapter avec le changement dans sa vie quotidienne et de s'y adapter.
- Résolution des problèmes : capacité de trouver des solutions efficaces aux problèmes d'ordre personnel et social.

5) **Humeur générale (motivation personnelle)**

L'humeur générale est un facteur qui favorise un comportement intelligent sur le plan émotif et social et comprend les sous-composantes suivantes, lesquelles permettent de déterminer la capacité d'être optimiste, positif et suffisamment motivé pour se fixer des buts et les atteindre :

- Optimisme : capacité d'avoir un point de vue positif et de prendre les choses du bon côté.
- Joie de vivre : capacité d'être bien avec soi-même et avec les autres et d'être satisfait de la vie en général.

Pour obtenir plus de renseignements à ce sujet, veuillez consulter les ouvrages suivants :

- 1) BarOn, Reuven., 2004. The BarOn Emotional Quotient Inventory (EQ-i): Rationale, description, and summary of psychometric properties, *Measurement of emotional*

intelligence: Common ground and controversy, Glenn Geher (Éd.), Hauppauge (New York), Nova Science Publishers, p. 111 à 142.

- 2) BarOn, Reuven., et Parker, James D. A., 2000. *Emotional Quotient Inventory: Youth Version (EQ-i :YV): Technical manual*. Multi-Health Systems, Toronto.

9.6.5 Échelle du soutien social

L'échelle initiale comportait 24 items tirés du *Social Provisions Model* de Robert Weiss, qui décrit six fonctions ou « dispositions » sociales distinctes qu'il est possible d'acquérir grâce aux relations avec autrui. Compte tenu de la taille de l'échelle et suivant les recommandations de M. Boyle de Chedoke-McMaster Hospital de la McMaster University, on a utilisé dans le cadre de l'enquête une version abrégée (à six items) inspirée du projet *Partir d'un bon pas pour un avenir meilleur* du gouvernement de l'Ontario. Cette échelle permet de mesurer l'encadrement (deux questions), les alliances fiables (deux questions) et l'attachement (deux questions). Quatre questions supplémentaires portant sur différents genres de soutien social (p. ex. le soutien religieux et les services communautaires) ont été ajoutées comme le suggérait Tom Hay. Des questions s'apparentant à celles proposées par Tom Hay ont été tirées des *Family Crisis Oriented Personal Evaluation Scales* (F-COPES). Les F-COPES se fondent sur les dimensions de l'adaptation du *Resiliency Model of Family Adjustment and Adaptation* (McCubbin, Olson et Larsen, 1981). La mesure globale du soutien social comprend huit questions et porte non seulement sur l'importance du soutien social mais également sur la qualité de ce soutien. Les questions sont posées aux jeunes de 18 et 19 ans dans le questionnaire des jeunes.

Aperçu des résultats de l'échelle du soutien social des jeunes de 18 et 19 ans

Score	Items inclus	Taille de l'échantillon	Univers (âge en années)	Coefficient alpha de Cronbach (standardisé)
Soutien social (_SPYeS01)	_SPYeQ1A*, _SPYeQ1B, _SPYeQ1C, _SPYeQ1D*, _SPYeQ1E*, _SPYeQ1F, _SPYeQ1G, _SPYeQ1H*	1 627	18 et 19	0,840

*Indique que les valeurs ont été inversées.

Source : Statistique Canada, Enquête longitudinale nationale sur les enfants et les jeunes.

9.6.6 Échelle des amis

L'échelle des amis sert à mesurer comment le jeune se sent par rapport à ses relations avec ses pairs. Cette information est importante lorsqu'il s'agit de déterminer l'ampleur et la qualité du réseau de soutien social de l'enfant. Cet ensemble de questions sur l'entente avec les pairs provient de la *Peer Relations Subscale* tirée du *Marsh Self-description Questionnaire*, mis au point par H. W. Marsh.

Aperçu des résultats de l'échelle des amis des jeunes de 14 à 17 ans

Score	Items inclus	Taille de l'échantillon	Univers (âge en années)	Coefficient alpha de Cronbach (standardisé)
Amis (_FFcS01)	_FFCQ01, _FFCQ02, _FFCQ03, _FFCQ04	2 059 1 365	14 et 15 16 et 17	0,845 0,865

Source : Statistique Canada, Enquête longitudinale nationale sur les enfants et les jeunes.

9.6.7 Échelles « Mes parents et moi »

Les échelles « Mes parents et moi » visent à compléter la section sur le rôle parental du questionnaire de l'enfant remplie par le parent, grâce à la collecte de renseignements auprès de l'enfant lui-même sur sa perception quant à sa relation avec ses parents. On jugeait important que le questionnaire rempli par l'enfant permette aussi de mesurer la supervision parentale (c'est-à-dire la surveillance) puisqu'on sait qu'elle est liée au développement de l'enfant — on observe en effet une corrélation entre le manque de supervision et les situations négatives, par exemple, la délinquance juvénile et d'autres comportements à risque.

Cette échelle a également été utilisée dans le cadre de la *Child Health Survey* menée en Australie-Occidentale. Elle a été élaborée par Lempers et autres (1989), qui se sont fondés sur les travaux de Schaefer (1965) et de Roberts et autres (1984). Cette échelle mesure les soins prodigués par les parents, le rejet et la surveillance.

Aperçu des résultats des échelles « Mes parents et moi » des jeunes de 14 et 15 ans

Score	Items inclus	Taille de l'échantillon	Univers (âge en années)	Coefficient alpha de Cronbach (standardisé)
Soins prodigués par les parents (_PMCcS1)	_PMCcQ1A, _PMCcQ1D, _PMCcQ1K, _PMCcQ1H, _PMCcQ1I, _PMCcQ1M, _PMCcQ1Q	2 025	14 et 15	0,911
Rejet par les parents (_PMCbS2B)	_PMCcQ1C, _PMCcQ1G, _PMCcQ1J, _PMCcQ1L, _PMCcQ1O, _PMCcQ1P, _PMCcQ1R	1 997	14 et 15	0,767
Surveillance par les parents (_PMCcS3)	_PMCcQ1B, _PMCcQ1F, _PMCcQ1N, _PMCcQ1E, _PMCdQ1T	2 049	14 et 15	0,415

Source : Statistique Canada, Enquête longitudinale nationale sur les enfants et les jeunes.

9.6.8 Échelle de résolution des conflits

Deux scores ont été créés pour évaluer la résolution des conflits chez les jeunes de 16 et 17 ans, en s'inspirant de questions posées dans le questionnaire autoadministré. Un score porte sur la relation entre le jeune et sa mère et l'autre, sur la relation entre le jeune et son père. Un score élevé indique un nombre élevé de conflits entre le jeune et ses parents.

Aperçu des résultats des échelles de résolution des conflits des jeunes de 16 et 17 ans

Score	Items inclus	Taille de l'échantillon	Univers (âge en années)	Coefficient alpha de Cronbach (standardisé)
Résolution des conflits – mère (_PMCdS4)	_PMCdQ6C*, _PMCdQ6D, _PMCdQ6E, _PMCdQ6F, _PMCdQ6G, _PMCdQ6H, _PMCdQ6I, _PMCdQ6J*, _PMCdQ6K, _PMCdQ6L	1 326	16 et 17	0,840
Résolution des conflits – père (_PMCdS5)	_PMCdQ9C*, _PMCdQ9D, _PMCdQ9E, _PMCdQ9F, _PMCdQ9G, _PMCdQ9H, _PMCdQ9I, _PMCdQ9J*, _PMCdQ9K, _PMCdQ9L	1 282	16 et 17	0,837

*Indique que les valeurs ont été inversées.

Source : Statistique Canada, Enquête longitudinale nationale sur les enfants et les jeunes.

9.6.9 Échelles de comportement

Cette section reprend la liste de comportements qui a été utilisée dans le questionnaire de l'enfant rempli par la PMR au cycle 8. Tous les jeunes de 14 et 15 ans répondent à ces questions dans la section de l'enquête à remplir soi-même. La section vise à fournir des indicateurs des comportements suivants : troubles de la conduite, hyperactivité, inattention, agressivité physique, agressivité indirecte, troubles émotifs, anxiété et comportement prosocial ainsi que des comportements liés à des atteintes à la propriété.

Aperçu des résultats des échelles de comportement des jeunes de 14 et 15 ans

Score	Items inclus	Taille de l'échantillon	Univers (âge en années)	Coefficient alpha de Cronbach (standardisé)
Agressivité indirecte (_FBcS01)	_FBCQ01J, _FBCQ01R, _FBCQ01Z, _FBCQ1LL, _FBCQ1TT	2 070	14 et 15	0,702
Troubles émotifs – anxiété (_FBCdS02)	_FBCQ01F, _FBCQ01K, _FBCQ01Q, _FBCc01V, _FBCQ1CC, _FBCQ1MM, _FBCQ1RR	2 069	14 et 15	0,810
Agressivité physique – problèmes de conduite (_FBcS03)	_FBCQ01G, _FBCd01X, _FBCQ1AA, _FBCQ1FF, _FBCQ1JJ, _FBCQ1NN	2 068	14 et 15	0,763
Hyperactivité – inattention (_FBCdS04)	_FBCc01B, _FBCQ01I, _FBCQ01P, _FBCQ01S, _FBCQ01W, _FBCQ1HH, _FBCQ1QQ	2 051	14 et 15	0,783
Comportement prosocial (_FBcS05)	_FBCQ01A, _FBCQ01D, _FBCQ01H, _FBCQ01M, _FBCQ01U, _FBCQ1BB, _FBCQ1GG, _FBCQ1OO, _FBCQ1SS, _FBCc1UU	2 047	14 et 15	0,871
Atteintes à la propriété (_FBcS07)	_FBCQ01C, _FBCQ01E, _FBCQ01L, _FBCQ01T, _FBCQ1DD, _FBCQ1PP	2 060	14 et 15	0,701

Source : Statistique Canada, Enquête longitudinale nationale sur les enfants et les jeunes.

9.7 Résumé des échelles du cycle 8

9.7.1 Échelles relatives dans le questionnaire du parent

Variable de score	Nom de l'échelle	Univers
_DPSS01	Dépression	PMR des enfants de 0 à 15 ans
_FNHhS01	Fonctionnement de la famille	PMR des enfants de 0 à 15 ans
_SFHhS5	Sécurité du quartier	PMR des enfants de 0 à 15 ans
_SFHhS6	Voisinage	PMR des enfants de 0 à 15 ans
_SPHhS01	Soutien social	PMR des enfants de 0 à 15 ans
_BECeS01	Hyperactivité – inattention	PMR des enfants de 2 et 3 ans
_BECdS03	Troubles émotifs – anxiété	PMR des enfants de 2 et 3 ans
_BECS04	Agressivité physique – opposition	PMR des enfants de 2 et 3 ans
_BECS05	Angoisse de séparation	PMR des enfants de 2 et 3 ans
_BECdS06	Hyperactivité – inattention	PMR des enfants de 4 à 7 ans
_BECdS07	Comportement prosocial	PMR des enfants de 6 et 7 ans
_BECdS08	Troubles émotifs – anxiété	PMR des enfants de 4 à 7 ans
_BECdS09	Agressivité physique – problèmes de conduite	PMR des enfants de 4 à 7 ans
_BECS10	Agressivité indirecte	PMR des enfants de 4 à 7 ans
_MSCS01	Score brut – DMS	PMR des enfants de 0 à 47 mois
_MSCS02	Score normalisé DMS – basé sur les normes du cycle 1	PMR des enfants de 0 à 47 mois
_MSCdS03	Score normalisé DMS – basé sur les normes du cycle 8	PMR des enfants de 4 à 47 mois
_PRCS01	Interaction positive	PMR des enfants de 0 à 23 mois
_PRCS02	Inefficacité parentale	PMR des enfants de 0 à 23 mois
_PRCS03	Interaction positive	PMR des enfants de 2 à 7 ans
_PRCgS04	Inefficacité parentale	PMR des enfants de 2 à 7 ans
_PRCgS05	Cohérence parentale	PMR des enfants de 2 à 7 ans
_PRCS06	Rationalité parentale	PMR des enfants de 2 à 7 ans
_PRCbS09	Résolution des conflits	PMR des enfants de 14 et 15 ans

Source : Statistique Canada, Enquête longitudinale nationale sur les enfants et les jeunes.

9.7.2 Échelles relatives dans le questionnaire des enfants et des jeunes (autoadministré)

Variable de score	Nom de l'échelle	Univers
_FFcS01	Amis	Enfants et jeunes de 14 à 17 ans
_AMcS02	Estime globale de soi	Enfants et jeunes de 14 à 17 ans
_FBcS01	Agressivité indirecte	Enfants et jeunes de 14 et 15 ans
_FBCdS02	Troubles émotifs – anxiété	Enfants et jeunes de 14 et 15 ans
_FBcS03	Agressivité physique – troubles de la conduite	Enfants et jeunes de 14 et 15 ans
_FBCdS04	Hyperactivité – inattention	Enfants et jeunes de 14 et 15 ans
_FBcS05	Comportement prosocial	Enfants et jeunes de 14 et 15 ans
_FBcS07	Atteintes à la propriété	Enfants et jeunes de 14 et 15 ans
_PMCcS1	Soins prodigués par les parents	Enfants et jeunes de 14 et 15 ans
_PMCbS2b	Rejet par les parents	Enfants et jeunes de 14 et 15 ans
_PMCcS3	Surveillance par les parents	Enfants et jeunes de 14 et 15 ans
_EQYeS04	Quotient émotionnel (4 facteurs)	Enfants et jeunes de 14 à 17 ans
_EQYeS05	Quotient émotionnel (5 facteurs)	Enfants et jeunes de 14 à 17 ans
_EQYeS06	Aptitudes intrapersonnelles	Enfants et jeunes de 14 à 17 ans
_EQYeS07	Aptitudes interpersonnelles	Enfants et jeunes de 14 à 17 ans
_EQYeS08	Aptitudes de gestion du stress	Enfants et jeunes de 14 à 17 ans
_EQYeS09	Aptitudes d'adaptabilité	Enfants et jeunes de 14 à 17 ans
_EQYeS10	Humeur générale	Enfants et jeunes de 14 à 17 ans
_PMCdS4	Résolution des conflits – mère	Jeunes de 16 et 17 ans
_PMCdS5	Résolution des conflits – père	Jeunes de 16 et 17 ans
_HTCbS1B	Dépression	Jeunes de 16 et 17 ans

Source : Statistique Canada, Enquête longitudinale nationale sur les enfants et les jeunes.

9.7.3 Échelles relatives au questionnaire des jeunes (autoadministré)

Variable de score	Nom de l'échelle	Univers
_ACYdS01	Structure du quartier	Jeunes de 16 et 17 ans
_HTYfS01	Dépression	Jeunes de 18 à 25 ans
_SPYeS01	Soutien social	Jeunes de 18 et 19 ans
_AMYfS01	Estime globale de soi	Jeunes de 18 et 23 ans
_EQYfS11	Interpersonnelle	Jeunes de 20 et 21 et 24 et 25 ans
_EQYfS12	Intrapersonnelle	Jeunes de 20 et 21 et 24 et 25 ans
_EQYfS13	Gestion du stress	Jeunes de 20 et 21 et 24 et 25 ans
_EQYfS14	Adaptabilité	Jeunes de 20 et 21 et 24 et 25 ans
_EQYfS15	Humeur générale	Jeunes de 20 et 21 et 24 et 25 ans
_EQYfS16	Quotient émotionnel (EQ4)	Jeunes de 20 et 21 et 24 et 25 ans
_EQYfS17	Quotient émotionnel (EQ5)	Jeunes de 20 et 21 et 24 et 25 ans

Source : Statistique Canada, Enquête longitudinale nationale sur les enfants et les jeunes.

10.0 Imputation

Un répondant est défini comme étant un enfant ou un jeune pour lequel au moins une des composantes destinées à l'enfant, au jeune ou à l'adulte a été remplie. On observe de nombreux cas de non-réponse partielle parmi ces répondants. Il peut s'agir d'une non-réponse à toute une composante ou à certaines questions seulement. L'imputation est le procédé qui consiste à remplacer les valeurs manquantes ou incohérentes par des valeurs acceptables. Dans le cadre de l'Enquête longitudinale nationale sur les enfants et les jeunes (ELNEJ), ce procédé est utilisé pour certaines variables des sections sur le revenu des adultes et des jeunes ainsi que de la section sur le développement moteur et social de l'enfant.

On a intégré des indicateurs d'imputation au fichier de l'ELNEJ afin d'informer les utilisateurs de l'ampleur de l'imputation et des éléments particuliers ayant fait l'objet d'une imputation dans tel ou tel enregistrement. Tous les indicateurs d'imputation du fichier de données de l'ELNEJ comporte un « I » au cinquième ou sixième caractère du nom de la variable. Ainsi, HINPeI03 représente l'indicateur d'imputation pour le revenu de la PMR (HINPeD04).

10.1 Imputation du revenu du ménage

L'interview du ménage dans le cadre de l'ELNEJ comporte plusieurs questions relatives au revenu. On recueille des renseignements sur le revenu de la personne la mieux renseignée (PMR) et de son conjoint selon les trois sources suivantes : salaires et traitements, revenu net tiré d'un travail autonome et prestations d'assurance-emploi. On a demandé les revenus des trois mêmes sources pour tous les autres membres du ménage ayant au moins 15 ans; ces revenus n'ont pas été déclarés pour chaque membre respectif mais pour l'ensemble de ces derniers. On recueille aussi des renseignements sur le revenu du ménage selon les quatre sources suivantes : prestation fiscale pour enfants/prestation nationale pour enfants, assistance sociale, allocation de soutien des enfants et du conjoint et toutes autres sources. La somme de ces 13 sources de revenu constitue le revenu total du ménage.

Le revenu est un sujet délicat. Certains répondants refusent donc de donner des réponses aux questions détaillées se rapportant au revenu. Parmi ceux-ci, certains fournissent une estimation du revenu total de leur ménage ou une estimation de leur revenu au moyen des intervalles de revenu. Enfin, parmi ceux qui répondent aux questions détaillées se rapportant au revenu, il arrive que les montants indiqués à la section sur le revenu soient incompatibles avec les réponses données à la section relative à l'activité sur le marché du travail (par exemple, un répondant qui, selon les réponses fournies à la section relative à l'activité sur le marché du travail, a travaillé au cours des 12 derniers mois mais qui ne déclare pas de salaire ou de revenu d'un emploi autonome à la section sur le revenu). On procède à une imputation du revenu pour combler les valeurs manquantes attribuables à la non-réponse partielle et pour corriger les données incohérentes lorsque cela est possible. L'imputation a aussi été faite pour les ménages qui ont déclaré un revenu total de moins de 6 000 \$.

L'imputation du revenu du ménage a été effectuée uniquement pour les ménages admissibles à une composante s'adressant aux adultes. Cela englobe tous les ménages, à l'exception de ceux où seuls des jeunes de 18 et plus sont sélectionnés et de ceux composés uniquement de jeunes de 16 et 17 ans vivant de façon autonome. Au moins une variable du revenu a été imputée pour 4 488 des 19 619 ménages admissibles. Ces 19 619 ménages admissibles correspondent en fait à 20 305 enfants ou jeunes répartis dans deux fichiers selon leur âge effectif : i) le fichier longitudinal pour les enfants ou les jeunes de la cohorte originale sélectionnée au cycle 1, qui ont entre 14 et 17 ans au cycle 8 et au cycle 2) le fichier du développement de la petite enfance (DPE) pour les enfants ou les jeunes faisant partie des cohortes de naissances sélectionnées aux cycles 5, 6, 7 et 8, et qui sont âgés de 0 à 7 ans au cycle 8. Le fichier longitudinal contient 4 247 enfants ou jeunes. De ce nombre, on a imputé des valeurs à 801 d'entre eux, tandis que le fichier du DPE contient 16 058 enfants, et on a imputé des valeurs à 3 822 d'entre eux. Dans l'ensemble, des valeurs ont été imputées à 4 623 des 20 305 enfants ou jeunes. Au cycle 8, les réponses par personne interposée ne sont plus acceptées pour la section sur le revenu, ce qui signifie que si quelqu'un autre que la PMR répond au questionnaire, la section sur le revenu est ignorée. Cette décision a suscité 254 non-réponses à la section sur le revenu. Nous avons réussi à trouver 160 de ces non-réponses

dans le cycle précédent, mais pour 94 d'entre eux nous n'avions pas de renseignements sur le revenu pour aider à l'imputation.

Le revenu total du ménage constitue le renseignement le plus crucial de la section sur le revenu des adultes. Notre stratégie d'imputation pour le cycle 8 visait à déterminer la meilleure valeur possible concernant le revenu total du ménage, parfois au détriment des sources de revenu déclarées. L'imputation a été effectuée à l'aide de diverses formes de la méthode du plus proche voisin. Cette méthode consiste tout d'abord à retracer un répondant ayant fourni une réponse à la section sur le revenu (un donneur) et dont les caractéristiques sont semblables à celles du répondant n'ayant pas fourni de renseignements complets sur le revenu (le receveur). L'enregistrement du donneur est alors utilisé pour calculer les valeurs à imputer pour l'enregistrement du receveur. L'imputation a été effectuée en quatre étapes : a) les ménages qui ont indiqué un revenu estimatif; b) les ménages qui ont répondu au cycle 7; c) les ménages ayant déclaré un revenu à l'aide d'intervalles de revenu; d) les autres ménages (incluant les ménages ayant seulement des sources de revenu manquantes pour les membres du ménage âgés de 15 ans et plus n'étant pas une PMR ou un conjoint).

1) Imputation pour les ménages qui ont indiqué un revenu estimatif

Pour ces ménages, on considère le revenu estimatif comme étant le revenu total du ménage. S'il manque une seule source de revenu, elle est imputée de façon déterministe. S'il manque plus d'une source de revenu, les 13 sources de revenu sont imputées simultanément à l'aide de la répartition des sources de revenu établie à partir d'un ménage donneur (imputation par le ratio du donneur). En fait, chaque source de revenu est imputée si elle n'est pas déclarée ou si la valeur calculée fondée sur le donneur diffère de plus de 10 % de la valeur déclarée. En fin de compte, le revenu total réel du ménage obtenu en additionnant les 13 sources de revenu peut être légèrement différent du revenu estimatif fourni.

2) Imputation pour les ménages qui ont répondu au cycle 7

Afin de conserver la cohérence longitudinale au fil du temps, l'imputation du revenu total des ménages qui ont répondu au cycle 7 a été réalisée au moyen de la méthode de la tendance du plus proche voisin, sauf en ce qui concerne les ménages donneurs qui affichaient des tendances extrêmes sur le plan du revenu du cycle 7 au cycle 8. Lorsqu'un receveur déclare son revenu total à l'aide des intervalles de revenu, on s'assure que le revenu total imputé respecte l'intervalle précisé. Tout comme dans la situation 1), la répartition des sources de revenu du ménage donneur est utilisée pour imputer les 13 sources de revenu du ménage receveur (imputation par le ratio du donneur).

3) Imputation pour les ménages ayant déclaré un revenu à l'aide d'intervalles de revenu

Une imputation est effectuée pour chaque source de revenu manquante à l'aide de la méthode traditionnelle du plus proche voisin. Les sources de revenu déclarées par le donneur sont directement utilisées pour imputer les sources de revenu manquantes du receveur. Les sources de revenu manquantes sont tous imputées en même temps en utilisant le même ménage donneur et en s'assurant que le revenu total du ménage respecte l'intervalle retenu.

4) Imputation pour les autres ménages

On ne dispose pas de beaucoup de renseignements sur le revenu total des autres ménages. Comme dans la situation précédente, une imputation est effectuée pour chaque source de revenu manquante à l'aide de la méthode traditionnelle du plus proche voisin. Les sources de revenu déclarées par le donneur sont directement utilisées pour imputer les sources de revenu manquantes du receveur. L'imputation est toutefois réalisée en trois étapes cette fois-ci : i) imputation des sources de revenu de la PMR; ii) imputation des sources de revenu du conjoint; iii) imputation des sources de revenu des autres membres du ménage et les sources de revenu au niveau du ménage.

Veuillez noter qu'au cycle 8, une partie des ménages de cette étape regroupait ceux qui, lorsqu'on leur a demandé une estimation du revenu total du ménage, toutes sources confondues, ont donné une estimation à 1 000 \$ près de la somme déclarée par la PMR, le conjoint et le ménage (10 sources au total). Il semble que les répondants dans ces ménages ne connaissaient probablement pas les revenus des autres membres du ménage et qu'ils ne se sont donc pas hasardés à fournir une estimation ou qu'ils ne les ont pas considérés comme contribuant au revenu du ménage. Par

conséquent, leur revenu estimé n'a pas été pris en compte. Comme il a été mentionné précédemment, les sources de revenu manquantes ont été imputées simultanément à l'aide de la méthode traditionnelle du plus proche voisin. Au total, 51 ménages des 903 de l'étape 4 étaient concernés.

Répartition des ménages pour lesquels on a effectué une imputation selon l'étape du procédé d'imputation pour le cycle 8

Étape du procédé d'imputation	Ménages	
	Nombre	%
1) Ménages indiquant un revenu estimatif	1 253	27,9
2) Ménages ayant répondu au cycle 7	1 734	38,6
3) Ménages déclarant un revenu à l'aide d'intervalles de revenu	598	13,3
4) Autres ménages	903	20,1
Total	4 488	100

Source : Statistique Canada, Enquête longitudinale nationale sur les enfants et les jeunes.

Les indicateurs d'imputation fournissent des renseignements sur la façon dont l'imputation a été effectuée. Voici la valeur des indicateurs.

Indicateur du revenu total du ménage (HINHgl03) :

Indicateur d'imputation = 0	Aucune imputation
Indicateur d'imputation = 1	Imputation du revenu estimatif
Indicateur d'imputation = 2	Imputation par la tendance du donneur selon l'intervalle de revenu
Indicateur d'imputation = 3	Imputation par la tendance du donneur
Indicateur d'imputation = 4	Imputation par donneur selon l'intervalle de revenu
Indicateur d'imputation = 5	Imputation par donneur

Indicateurs du revenu partiel (HINPeI03 et HINSeI03) :

Indicateur d'imputation = 0	Aucune imputation
Indicateur d'imputation = 1	Au moins une source de revenu imputée
Indicateur d'imputation = 6	Enchaînement valide (pas de conjoint dans le ménage)

Indicateurs des sources de revenu (HINPI1AA, HINPI1AB, HINPI1AC, HINSI1AA, HINSI1AB, HINSI1AC, HINHI1AD, HINHI1AE, HINHI1AF et HINHI1AG) :

Indicateur d'imputation = 0	Aucune imputation
Indicateur d'imputation = 1	Imputation déterministe
Indicateur d'imputation = 2	Imputation par le ratio du donneur
Indicateur d'imputation = 3	Imputation par donneur
Indicateur d'imputation = 6	Enchaînement valide (pas de conjoint dans le ménage)

Les taux d'imputation au niveau des enfants pour les variables du revenu, cycle 8

Variable	Indicateur d'imputation	Taux d'imputation global	Taux d'imputation pour le fichier longitudinal (14 à 17 ans)	Taux d'imputation pour le fichier du DPE (0 à 7 ans)
		(%)		
HINPc1AA Revenu de la PMR (salaires et traitements)	HINPI1AA	14,4	12,8	14,8
HINPc1AB Revenu de la PMR (emploi autonome)	HINPI1AB	11,5	10,2	11,9
HINPc1AC Revenu de la PMR (prestations d'assurance-emploi)	HINPI1AC	12,9	10,4	13,5
HINPeD04 Revenu total personnel de la PMR	HINPeI03	17,6	14,5	18,4
HINSc1AA Revenu du conjoint (salaires et traitements)	HINSI1AA	18,3 ¹	15,6 ¹	18,9 ¹
HINSc1AB Revenu du conjoint (emploi autonome)	HINSI1AB	14,3 ¹	13,2 ¹	14,5 ¹
HINSc1AC Revenu du conjoint (prestations d'assurance-emploi)	HINSI1AC	13,0 ¹	11,0 ¹	13,5 ¹
HINSeD04 Revenu total personnel du conjoint	HINSeI03	20,2 ¹	17,0 ¹	20,9 ¹
HINOg1AA Revenu des autres membres 15 ans et plus du ménage (salaires et traitements)	HINOI1AA	20,3 ²	15,3 ²	29,2 ²
HINOg1AB Revenu des autres membres 15 ans et plus du ménage (emploi autonome)	HINOI1AB	12,3 ²	10,2 ²	15,8 ²
HINOg1AC Revenu des autres membres 15 ans et plus du ménage (prestations d'assurance-emploi)	HINOI1AC	12,1 ²	9,7 ²	16,3 ²
HINHeD3P Revenu total personnel des autres membres 15 ans et plus du ménage	HINHeI3P	21,1 ²	15,9 ²	30,4 ²
HINHe1AD Revenu du ménage (prestation fiscale pour enfants)	HINHI1AD	18,1	15,0	18,9
HINHe1AE Revenu du ménage (assistance sociale)	HINHI1AE	11,4	9,5	11,9
HINHe1AF Revenu du ménage (allocation de soutien des enfants ou au conjoint)	HINHI1AF	11,7	9,9	12,2
HINHe1AG Revenu du ménage (autres sources)	HINHI1AG	12,0	10,3	12,4
HINHgQ03 Revenu total du ménage	HINHgl03	22,8	18,9	23,8

1. Les ménages où il n'y avait pas de conjoint ne sont pas inclus dans le calcul du taux d'imputation.

2. Les ménages où les seuls membres âgés d'au moins 15 ans sont la PMR et le conjoint ne sont pas inclus dans le calcul du taux d'imputation.

Source : Statistique Canada, Enquête longitudinale nationale sur les enfants et les jeunes.

10.2 Imputation du revenu des jeunes — 16 et 17 ans

Dans le cadre de la composante sur les jeunes, on recueille des renseignements sur le revenu selon cinq sources auprès des jeunes âgés de 16 et 17 ans au cycle 8. On demande aux jeunes de déclarer les revenus provenant de travaux divers, d'un employeur, de bourses d'études et de subventions, des parents et toutes autres sources. La somme de ces cinq sources de revenu représente le revenu total des jeunes.

Le revenu est un sujet délicat. En effet, certains répondants refusent de donner des réponses aux questions détaillées se rapportant au revenu. Parmi ceux-ci, certains fournissent une estimation de leur revenu au moyen des intervalles de revenu. On procède à une imputation du revenu pour combler les valeurs manquantes attribuables à la non-réponse partielle.

On a envisagé une imputation du revenu pour tous les répondants de 16 et 17 ans qui ont fourni des renseignements dans la composante sur les jeunes. Sur les 1 746 jeunes répondants de 16 et 17 ans au cycle 8, 1 614 ont rempli la composante sur les jeunes. Les 132 autres jeunes n'ont rempli que la composante de l'enfant et/ou de l'adulte. On ne disposait pas de renseignements suffisants sur les jeunes n'ayant pas rempli la composante du jeune pour procéder à l'imputation de leur revenu. Parmi les répondants qui ont rempli la composante du jeune, 218 jeunes ont fait l'objet d'une imputation d'au moins une source de revenu.

On a procédé à l'imputation pour chacune des cinq sources de revenu, dans la plupart des cas, par la méthode du plus proche voisin. Cette méthode consiste tout d'abord à retracer un répondant ayant fourni une réponse à la section sur le revenu (un donneur) et dont les caractéristiques sont semblables à celles de la personne n'ayant pas fourni de renseignements complets sur le revenu (le receveur). Une fois qu'on a identifié le plus proche voisin, on copie les sources de revenu manquantes dans l'enregistrement du receveur. Lorsqu'il est fourni, l'intervalle de revenu total est pris en considération dans la sélection du donneur, de sorte que le revenu total imputé se situe dans l'intervalle précisé. Lorsqu'une seule des cinq sources de revenu doit être imputée et que l'intervalle de revenu total est indiqué en intervalles, on impute une valeur plausible choisie au hasard à même une répartition uniforme de valeurs possibles.

Les indicateurs d'imputation indiquent de quelle façon l'imputation a été faite. Voici la valeur des indicateurs.

Indicateur d'imputation = 0	Aucune imputation
Indicateur d'imputation = 1	Imputation par donneur selon l'intervalle de revenu
Indicateur d'imputation = 2	Imputation d'une valeur plausible
Indicateur d'imputation = 3	Imputation par donneur

Répartition des jeunes de 16 et 17 ans pour lesquels on a effectué une imputation selon l'étape du procédé d'imputation pour le cycle 8

Méthode d'imputation ¹	Nombre de jeunes ²	%
Imputation par donneur selon l'intervalle de revenu	74	33,9
Imputation d'une valeur plausible	103	47,3
Imputation par donneur	41	18,8
Total	218	100,0

1. Indicateur d'imputation H1NYel02.

2. Répondants pour lesquels au moins une des sources de revenu a été imputée.

Source : Statistique Canada, Enquête longitudinale nationale sur les enfants et les jeunes.

Taux d'imputation pour les variables du revenu pour les jeunes de 16 et 17 ans, cycle 8

Variable	Indicateur d'imputation	Taux d'imputation (%)
HINYeQ1A (Revenu provenant de travaux divers)	HINYeI1A	4,4
HINYeQ1B (Revenu provenant d'employeurs)	HINYeI1B	5,0
HINYeQ1E (Revenu provenant des bourses d'études et de subventions)	HINYeI1E	1,6
HINYeQ1C (Revenu provenant des parents)	HINYeI1C	8,3
HINYeQ1D (Revenu provenant d'autres sources)	HINYeI1D	3,9
HINYeD01 (Revenu total du jeune)	HINYeI02	13,5

Source : Statistique Canada, Enquête longitudinale nationale sur les enfants et les jeunes.

10.3 Imputation du revenu des jeunes — 18 ans et plus

Dans le cadre de la composante sur les jeunes, des renseignements sur le revenu sont recueillis auprès de ceux qui ont 18 ans et plus. Les jeunes sont priés d'indiquer leur revenu total ainsi que leurs sources de revenu, mais pas le montant attribué à chaque source. Voici la liste des sources de revenu qui leur est présentée.

1. Salaires et traitements (y compris les commissions, les pourboires et les gratifications)
2. Revenu d'un travail autonome
3. Bourses d'études
4. Prêt d'études consenti par le gouvernement
5. Régime enregistré d'épargne-études (REEEE)
6. Intérêts, dividendes, gains en capital ou autres revenus de placement
7. Assurance-emploi (prestations)
8. Indemnité pour accident du travail
9. Prestations du Régime de pensions du Canada ou du Régime de rentes du Québec, supplément de revenu garanti ou allocation au conjoint
10. Prestation fiscale pour enfants
11. Assistance ou aide sociale de l'administration provinciale ou municipale
12. Pension alimentaire pour enfant
13. Pension alimentaire
14. Autre
15. Aucun revenu

Le revenu est un sujet délicat. En effet, certains répondants refusent de donner des réponses aux questions détaillées se rapportant au revenu. Parmi ceux-ci, certains fournissent une estimation de leur revenu au moyen des intervalles de revenu. Enfin, parmi ceux qui répondent aux questions détaillées se rapportant au revenu, il arrive que les montants indiqués à la section sur le revenu soient incompatibles avec les réponses données à la section relative à l'activité sur le marché du travail (par exemple, un jeune qui, selon les réponses fournies à la section relative à l'activité sur le marché du travail, fait des travaux contre rémunération, mais qui ne déclare pas de revenu à la section sur le revenu). Une imputation du revenu est effectuée pour combler les valeurs manquantes attribuables à la non-réponse partielle ou pour corriger les données incohérentes lorsque cela est possible.

On a envisagé effectuer une imputation du revenu pour tous les répondants de 18 ans et plus. Parmi les 5 937 jeunes répondants qui avaient 18 ans et plus au cycle 8, on a imputé le revenu total

personnel pour 694 d'entre eux et la liste des sources de revenu a été imputée pour 20 jeunes (dont le revenu de un ne nécessitait pas d'imputation). En outre, comme au cycle 7, la variable du revenu du ménage applicable aux jeunes (HIYYgQ2B), qui a été demandée à tous les jeunes ayant un conjoint, a été diffusée. Le revenu du ménage de 191 des 1 103 jeunes ayant un conjoint a fait l'objet d'une imputation.

L'imputation a été réalisée à l'aide de la méthode du plus proche voisin. Cette méthode consiste tout d'abord à retracer un répondant ayant fourni une réponse à la section sur le revenu (un donneur) et dont les caractéristiques sont semblables à celles de la personne n'ayant pas fourni de renseignements complets sur le revenu (le receveur). Une fois qu'on a identifié le plus proche voisin, les données manquantes sont copiées dans l'enregistrement du receveur. Lorsqu'il est fourni, l'intervalle du revenu total est pris en considération dans la sélection du donneur de sorte que le revenu total imputé se situe dans l'intervalle précisé.

Les indicateurs d'imputation montrent de quelle façon l'imputation a été faite. Voici la valeur des indicateurs.

Sources de revenu des jeunes (HIYYfQ1A à HIYYfQ1O)

Indicateur d'imputation = 0 Aucune imputation
Indicateur d'imputation = 1 Imputation par donneur

Revenu personnel des jeunes (HIYYfD01)

Indicateur d'imputation = 0 Aucune imputation
Indicateur d'imputation = 1 Imputation par donneur selon l'intervalle de revenu
Indicateur d'imputation = 2 Imputation par donneur

Revenu du ménage des jeunes (HIYYgQ2B)

Indicateur d'imputation = 0 Aucune imputation
Indicateur d'imputation = 1 Imputation par donneur

Répartition des jeunes de 18 ans et plus pour lesquels on a effectué une imputation selon l'étape du procédé d'imputation pour le cycle 8¹

Méthode d'imputation ²	Nombre de jeunes ³	%
Imputation par donneur selon l'intervalle de revenu	528	76,1
Imputation par donneur	166	23,9
Total	694	100,0

1. Ce tableau présente le revenu personnel, qui est la seule variable imputée à l'aide d'intervalles fournies par le répondant. Les taux d'imputation pour les sources de revenu et le revenu du ménage sont présentés dans le tableau suivant.

2. Indicateur d'imputation HIYYfI2A.

3. Répondants pour lesquels leur revenu total personnel a été imputé.

Source : Statistique Canada, Enquête longitudinale nationale sur les enfants et les jeunes.

Taux d'imputation pour les variables du revenu pour les jeunes de 18 ans et plus, cycle 8

Variable	Indicateur d'imputation	Taux d'imputation (%)
HIYYfQ1A à HIYYfQ1O (Sources de revenu)	HIYYf11	0,3
HIYYfD01 (Revenu total personnel)	HIYYf12A	11,7
HIYYgQ2B (Revenu total du ménage)	HIYYg12B	17,3

Source : Statistique Canada, Enquête longitudinale nationale sur les enfants et les jeunes.

10.4 Imputation de l'échelle de Développement moteur et social (DMS)

La méthodologie retenue au cycle 8 pour l'imputation de l'échelle de Développement moteur et social (DMS) est la même que celle retenue aux cycles 5, 6 et 7. Antérieurement au cycle 5, pour qu'un enfant obtienne un score brut pour l'échelle de DMS (variable HMSCS01), on devait répondre par « Oui » ou « Non » aux 15 questions que comporte l'échelle de DMS. Cependant, pour la vaste majorité des cas où l'on ne pouvait pas calculer un score brut, il a été noté que seulement une ou deux des 15 questions de l'échelle contenaient une réponse manquante. On a donc décidé, depuis le cycle 5, qu'on pourrait obtenir un score très valable en faisant usage des 13 ou 14 questions pour lesquelles on avait obtenu une réponse, et en imputant les données manquantes. Nous avons donc procédé de la façon suivante, depuis lors.

Plus précisément, si un enfant a 13 ou 14 réponses valides, un enregistrement donneur est choisi aléatoirement parmi les enfants qui ont répondu à toutes les questions et qui ont des réponses identiques aux questions communes. Lorsqu'une question a été imputée, la valeur manquante initiale est remplacée par un « Oui » ou un « Non » provenant de l'enregistrement donneur sélectionné. L'imputation est faite de façon indépendante lorsque les réponses à deux questions sont imputées. Ainsi, les réponses de deux donneurs différents peuvent être utilisées pour deux valeurs manquantes.

Il n'y a pas toujours de donneur présentant le même schéma de réponse aux questions communes. Dans ces cas, un donneur est choisi parmi les enfants qui ont répondu aux 15 questions et qui affichent un score partiel identique aux questions communes.

Naturellement, afin d'avoir 13 ou 14 questions en commun, tous les donneurs potentiels doivent se situer dans le même groupe d'âge (en mois) que celui de l'enfant dont il faut imputer des réponses. Par exemple, un enfant de 8 mois ayant une valeur manquante pour la variable HMSCQ21 a comme donneurs potentiels les enfants de 7 à 9 mois à qui les mêmes 15 questions (HMSCQ12 à HMSCQ26) ont été posées à la PMR et qui présentent le même schéma de réponse « Oui » et « Non » pour HMSCQ12 à HMSCQ20 et HMSCQ22 à HMSCQ26.

Dans ce contexte, une réponse valide ne passe jamais de « Oui » à « Non » et vice versa. Seulement les valeurs manquantes sont remplacées par un « Oui » ou un « Non ».

Au total, 439 scores additionnels du DMS ont été obtenus en imputant au moins une réponse. En fait, une réponse a été imputée dans 386 cas, et deux réponses ont été imputées dans 53 cas. Cela représente 5,2 % de tous les enfants admissibles.

Les indicateurs d'imputation pour les variables HMSCIS1A et HMSCIS1B désignent les questions de l'échelle de DMS pour lesquelles des réponses sont imputées. Une valeur de zéro pour ces deux indicateurs d'imputation signifie qu'aucune imputation n'a été effectuée pour les questions sur le DMS.

11.0 Pondération et traitement de la non-réponse

L'Enquête longitudinale nationale sur les enfants et les jeunes (ELNEJ) est une enquête probabiliste. Comme c'est le cas pour toutes les enquêtes probabilistes, l'échantillon est sélectionné de façon à pouvoir produire des estimations pour une population de référence. Pour ce faire, chaque unité de l'échantillon représente un certain nombre d'unités dans la population. Dans l'ELNEJ, plusieurs populations sont représentées. L'échantillon complet du cycle 8 est une combinaison d'échantillons sélectionnés aux cycles 1, 5, 6, 7 et 8. Pour obtenir des renseignements au sujet du plan d'échantillonnage et de la composition de l'échantillon de chaque cycle de l'ELNEJ, veuillez consulter le chapitre 5.0 du présent guide.

On a vu au chapitre 5.0 que, dans une enquête longitudinale comme l'ELNEJ, deux types de populations sont possibles : la population longitudinale et les populations transversales. La première est formée par la population initiale au moment où l'échantillon a été prélevé la première fois et elle ne change pas au fil du temps, tandis que la seconde pourrait faire référence à d'autres périodes ultérieures. Les différences entre la population longitudinale et les populations transversales sont attribuables aux naissances, aux décès, à l'immigration et à l'émigration.

L'ELNEJ permet de produire trois ensembles de poids à chaque cycle : deux ensembles de poids longitudinaux (entonnoir et non-entonnoir) et un ensemble de poids transversaux. Les poids entonnoir sont attribués aux enfants longitudinaux qui ont répondu à chaque cycle, tandis que les poids non-entonnoir sont attribués aux enfants longitudinaux qui ont répondu au cycle le plus récent, mais pas nécessairement à tous les cycles précédents.

On calcule les poids d'enquête à partir du poids déterminé par le plan d'échantillonnage de l'enfant, que l'on rajuste pour tenir compte de la non-réponse et de la stratification a posteriori, afin que les poids d'enquête définitifs correspondent aux nombres connus d'enfants, selon l'âge, le sexe et la province. Le poids déterminé par le plan d'échantillonnage est l'inverse de la probabilité de sélection, c'est-à-dire la probabilité qu'un enfant dans la population soit sélectionné dans l'échantillon de l'ELNEJ.

Le présent chapitre porte sur la façon dont les poids d'enquête sont calculés et sur les populations auxquelles ils se rapportent.

11.1 Poids disponibles au cycle 8

Les paragraphes qui suivent portent sur les populations de référence pour les différents poids produits au cycle 8. Pour plus de renseignements sur les poids d'enquête et leur population de référence pour les cycles précédents, veuillez consulter le chapitre 5.0.

11.1.1 Poids longitudinaux : poids entonnoir (variable HWTCWd1L) et poids non-entonnoir (variable HWTCW01L)

Des poids entonnoir et des poids non-entonnoir longitudinaux ont été attribués aux répondants du cycle 8 qui appartenaient à :

- la cohorte initiale formée d'enfants de 0 à 11 ans sélectionnés au cycle 1;
- la cohorte du DPE formée d'enfants de 0 à 1 an sélectionnés au cycle 5;
- la cohorte du DPE formée d'enfants de 0 à 1 an sélectionnés au cycle 6;
- la cohorte du DPE formée d'enfants de 0 à 1 an sélectionnés au cycle 7.

Dans le cas des cohortes du DPE, les poids entonnoir, `_wtc wd1l`, ont été utilisés pour la première fois au cycle 7. Avant le cycle 7, les enfants répondaient à l'enquête uniquement s'ils avaient répondu à tous les cycles précédents. Par conséquent, les poids longitudinaux, `_wtc w01l`, reçus par les enfants des cohortes du DPE avant le cycle 7 pourraient être

interprétés comme des poids entonnoir. Toutefois, à partir du cycle 7, les enfants des cohortes du DPE peuvent répondre à l'enquête, peu importe s'ils ont répondu aux cycles précédents. Ce qui fait, des poids entonnoir et des poids non-entonnoir, `_wtc wd11` et `_wtc w011`, respectivement, ont pu être calculés pour les enfants des cohortes du DPE.

Au moment du cycle d'introduction d'un enfant, seuls des poids transversaux sont attribués. C'est au moment des cycles subséquents que les enfants qui retournent dans l'enquête reçoivent des poids longitudinaux. Ainsi, au cycle 1, les enfants de la cohorte originale ont reçu un poids transversal, et ce n'est qu'aux cycles subséquents qu'ils ont reçu des poids longitudinaux. De même, les enfants des cohortes du DPE qui ont répondu pour la première fois à l'enquête entre 0 et 1 an ont reçu un poids transversal au moment de leur cycle d'introduction, et des poids transversaux et longitudinaux aux cycles subséquents.

Par ailleurs, dans le cas des enfants de la cohorte originale, seuls ceux qui ont répondu au cycle 1 ont reçu des poids longitudinaux. Dans le cas des cohortes du DPE, des poids longitudinaux sont attribués uniquement aux enfants qui ont été échantillonnés pour la première fois entre 0 et 1 an, mais ces enfants doivent ne pas avoir répondu à l'enquête à ce cycle d'introduction. Tous les enfants des cohortes du DPE qui sont échantillonnés pour la première fois entre 2 et 5 ans (c'est-à-dire les échantillons complémentaires des cycles 6, 7 et 8) recevront uniquement un poids transversal, parce que ces échantillons complémentaires sont utilisés à des fins transversales seulement.

11.1.2 Populations longitudinales

Ci-dessous, vous trouverez la description des populations longitudinales faisant partie de la cohorte initiale et des cohortes du DPE. Les enfants appartenant à une cohorte particulière peuvent être identifiés au moyen de la variable MEMCYCLE laquelle indique le cycle auquel l'enfant a commencé à participer à l'enquête.

La cohorte initiale

La population longitudinale de la cohorte initiale est définie par les enfants âgés de 0 à 11 ans au 31 décembre 1994 qui vivaient dans l'une des provinces lors de la collecte du cycle 1 (1994-1995).

Dans les fichiers de données, il est possible d'identifier ces personnes grâce à la mention « MEMCYCLE = 01 ».

La cohorte du DPE introduite au cycle 5

La population longitudinale de la cohorte du DPE sélectionnée au cycle 5 est définie par les enfants âgés de 0 à 1 an au 31 décembre 2002 qui vivaient dans l'une des provinces lors de la collecte du cycle 5 (2002-2003).

Dans les fichiers de données, il est possible d'identifier ces personnes grâce à la mention « MEMCYCLE = 05 ».

La cohorte du DPE introduite au cycle 6

La population longitudinale de la cohorte du DPE sélectionnée au cycle 6 est définie par les enfants âgés de 0 à 1 an au 31 décembre 2004 qui vivaient dans l'une des provinces lors de la collecte du cycle 6 (2004-2005).

Dans les fichiers de données, il est possible d'identifier ces personnes grâce à la mention « MEMCYCLE = 06 ».

La cohorte du DPE introduite au cycle 7

La population longitudinale de la cohorte du DPE sélectionnée au cycle 7 est définie par les enfants âgés de 0 à 1 an au 31 décembre 2006 qui vivaient dans l'une des provinces lors de la collecte du cycle 7 (2006-2007).

Dans les fichiers de données, il est possible d'identifier ces personnes grâce à la mention « MEMCYCLE = 07 ».

11.1.3 Poids transversaux (variable HWTCW01C)

Au cycle 8, des poids transversaux ont été attribués aux enfants répondants des cohortes du DPE qui pouvaient servir à représenter la population transversale des 0 à 7 ans du cycle 8. Étant donné que la cohorte originale n'a jamais été complétée pour tenir compte des immigrants, elle ne devrait pas servir aux analyses transversales.

Au cycle 8, l'échantillon transversal des enfants âgés de 0 à 7 ans de la cohorte du DPE comprenait :

- les enfants de 0 et 1 an faisant partie de la cohorte du DPE sélectionnés au cycle 8;
- les enfants de 2 et 3 ans ayant fait partie de la cohorte du DPE d'enfants de 0 et 1 an sélectionnés au cycle 7;
- les enfants de 4 et 5 ans ayant fait partie de la cohorte du DPE d'enfants de 0 à 1 an sélectionnés au cycle 6;
- un nouvel échantillon complémentaire d'enfants de 2 à 5 ans (sélectionné dans l'EPA);
- les enfants de 4 à 7 ans ayant fait partie de l'échantillon complémentaire d'enfants de 2 à 5 ans sélectionnés au cycle 7;

La population transversale au cycle 8 est définie par les enfants âgés de 0 à 7 ans au 31 décembre 2008, qui vivaient dans l'une des provinces lors de la collecte du cycle 8 (2008-2009).

11.2 Méthode de pondération

Un aperçu de la méthode de pondération utilisée est présenté ci-dessous. Les utilisateurs de l'ELNEJ qui veulent en savoir davantage peuvent contacter Statistique Canada.

11.2.1 Poids déterminé par le plan de l'échantillonnage de l'Enquête longitudinale nationale sur les enfants et les jeunes

La stratégie de pondération de l'ELNEJ repose sur une série d'ajustements appliqués au poids déterminé par le plan d'échantillonnage de l'enfant le quel est égal à l'inverse de la probabilité de sélection de l'enfant. Dans le présent chapitre, on renvoie au poids déterminé par le plan de l'ELNEJ ainsi : $W_{\text{plan d'échantillonnage de l'ELNEJ}}$.

Tous les enfants encore présents au cycle 8 de l'ELNEJ ont été sélectionnés, à un moment ou à un autre, à partir de l'EPA. Le sous poids de l'EPA tient compte de tous les renseignements sur la conception d'échantillon de l'échantillon de l'EPA. Pour l'ELNEJ, nous sélectionnons en outre des ménages à partir de l'EPA, puis des enfants au sein de ces ménages. Afin de rendre compte de ces étapes supplémentaires du plan d'échantillon, le sous poids de l'EPA est multiplié par plusieurs autres facteurs. Après avoir multiplié le sous

poids de l'EPA par ces divers facteurs, nous obtenons le poids déterminé par le plan d'échantillonnage de l'ELNEJ.

Un poids d'enquête définitif d'un enfant de l'ELNEJ est obtenu en appliquant les ajustements pour la non-réponse et la stratification a posteriori au poids déterminé par le plan d'échantillonnage de l'ELNEJ.

11.2.2 Premier ajustement : ajustement pour la non-réponse

En vérité, dans la plupart des enquêtes, les personnes qui font partie de l'échantillon ne répondent pas toutes, et l'ELNEJ ne fait pas exception. En raison de cette non-réponse, on doit ajuster les poids de façon telle que les répondants représentent les non-répondants. Sinon, on sous-estimerait les totaux, entre autres choses.

L'objectif de l'ajustement pour la non-réponse est plus précisément de gonfler les poids déterminés par le plan d'échantillonnage de l'ELNEJ des répondants de sorte que les poids ajustés de la non-réponse donnent la somme des poids déterminés par le plan d'échantillonnage de l'ELNEJ de tous les membres de l'échantillon initial. En d'autres mots, l'ajustement pour la non-réponse vise à donner les poids déterminés par le plan d'échantillonnage de l'ELNEJ des non-répondants aux répondants d'une façon avisée.

Les enfants de l'échantillon initial tirés au cycle 1 ont eu huit cycles de non-réponse au moment du cycle 8. La cohorte du DPE du cycle 5 a été affectée par quatre cycles de non-réponse (cycles 5, 6, 7 et 8), la cohorte du DPE du cycle 6 a connue trois cycles de non-réponse, et la cohorte du DPE du cycle 7 a perdu des répondants pendant deux cycles. L'échantillon de la cohorte du DPE du cycle 8 n'a eu qu'un seul cycle de non-réponse. L'ajustement pour la non-réponse vise à corriger tous ces épisodes de non-réponse.

Afin de décider de quelle façon le poids des non-répondants est attribué vers les répondants, on a recours à la méthode des groupes de réponses homogènes (GRH). La méthode des GRH vise à regrouper les individus ayant la même propension à répondre. Par la suite, un facteur d'ajustement est calculé pour chaque GRH. Ce facteur est défini comme suit :

$$\text{Ajustement de non - réponse dans un GRH} = \frac{\sum_{\substack{\text{Répondants} \\ \text{et} \\ \text{non-répondants}}} w_{\text{plan d'échantillonnage de l'ELNEJ}}}{\sum_{\text{Répondants}} w_{\text{plan d'échantillonnage de l'ELNEJ}}}$$

Au cycle 8, la propension à répondre est déterminée par des modèles de régression logistique à l'aide d'une méthode semblable à celle appliquée au cycle 7. Le modèle permet de prédire des probabilités de réponse (dont la valeur se situe entre 0 et 1) pour chacun. Ces probabilités sont divisées en plusieurs GRH à l'aide d'une analyse par grappes. Le nombre de GRH est déterminé de manière à ce que l'homogénéité globale des grappes soit maximisée et que certaines contraintes (taille du facteur d'ajustement, nombre minimal de répondants dans chaque GRH) soient respectées afin que des facteurs d'ajustement raisonnables et fiables soient établis.

Pour la pondération au cycle 8, l'ajustement pour la non-réponse a été effectué en deux phases à l'aide de deux modèles logistiques différents. Les non-répondants ont été divisés en deux groupes : les cas de non-contact et les autres (des refus pour la plupart). Dans le cas du premier ajustement, les poids des cas avec lesquels nous n'avons pas pu établir de contact ont été attribués aux cas pour lesquels un contact avait été établi (répondants et autres non-répondants) tandis que dans le cas du deuxième ajustement, les poids des autres non-répondants ont été attribués aux répondants.

Dix modèles de régression logistique ont été conçus afin de modéliser la non-réponse. Les modèles sont calculés pour la réponse en entonnoir et non-entonnoir longitudinale de la cohorte initiale; la réponse en entonnoir et non-entonnoir longitudinale pour les enfants de la cohorte du DPE des cycles 5, 6 et 7, et la réponse transversale pour les enfants de la cohorte du DPE des cycles 5, 6, 7 et 8. Des modèles distincts ont été construits parce que la cohorte originale avait connu beaucoup plus de cycles de non-réponse que les cohortes du DPE, et que tout porte à croire que le comportement au regard de la non-réponse varie en fonction de l'augmentation du nombre de fois où une personne est interviewée. Pour chacun des ajustements pour la non-réponse susmentionnés, deux modèles différents ont été construits pour les deux groupes de non-répondants.

11.2.3 Second ajustement : stratification a posteriori

Le second facteur d'ajustement vise à assurer la cohérence entre les estimations produites à partir de l'enquête et des estimations démographiques selon l'âge, le sexe et la province que produit Statistique Canada. Cette méthode porte le nom de stratification a posteriori. Le but de cet ajustement est de s'assurer que la somme des poids corresponde aux totaux connus de la population. Les totaux de la stratification a posteriori dépendent de la population de référence.

Tous les poids d'enquête finals sont stratifiés a posteriori. En ce qui concerne la pondération transversale au cycle 8, l'année de référence pour calculer l'âge effectif de l'enfant est 2008. Les chiffres utilisés dans la stratification a posteriori renvoient à janvier 2009, de sorte qu'on puisse disposer d'un dénombrement fiable des enfants d'un âge donné au 31 décembre 2008. Dans le même ordre d'idées, les poids entonnoir et non-entonnoir longitudinaux sont stratifiés a posteriori aux chiffres de janvier suivant l'année de référence pour l'année d'intérêt (énumérés à la section 11.1.2).

Pour savoir dans quelle strate de second niveau une personne en particulier se situe, voir les variables PSTRATC (strate de second niveau aux fins de la pondération transversale) et PSTRATL (strate de second niveau aux fins de la pondération longitudinale).

11.2.4 Différences entre la méthode de pondération des cycles 6, 7 et 8 et celle des cycles précédents

La méthode d'ajustement pour la non-réponse utilisée aux cycles 6, 7 et 8 diffère de celle utilisée aux cycles précédents. Au cycle 6, nous sommes passés d'un modèle de segmentation à un modèle de régression logistique. Nous obtenons ainsi des GRH moins nombreux, plus robustes et davantage distincts. En outre, au cycle 6, nous avons conçu un modèle de non-réponse véritablement longitudinal pour les poids longitudinaux. Au cycle 8, l'analyse par grappes a été introduite pour créer des GRH et les ajustements pour la non-réponse ont été divisés en deux phases afin de tenir compte des deux types de non-répondants.

Pour plus de renseignements concernant la pondération du cycle 6, veuillez consulter le Guide de l'utilisateur du cycle 6.

11.3 Application de la méthode de pondération

11.3.1 Pondération longitudinale non-entonnoir

Définition d'un répondant longitudinal

Un répondant longitudinal est un enfant introduit à un cycle précédent et pour lequel la composante adulte ou la composante enfant ou jeune a été remplie. Pour les jeunes de 18

ans et plus, la composante du jeune doit être remplie pour que le jeune puisse être considéré comme un répondant. Les enfants introduits à un cycle précédent qui sont décédés ou qui ont déménagé à l'extérieur des dix provinces canadiennes sont également des répondants longitudinaux. Ils représentent des enfants de la population de référence qui sont dans la même situation.

Premier ajustement : ajustement pour la non-réponse

Deux ensembles de modèles de non-réponse ont été créés : un pour la cohorte initiale et un autre pour les enfants de la cohorte du DPE. Les deux modèles utilisaient des variables de l'EPA (propriétaire ou locataire du logement, niveau de scolarité le plus élevé, entre autres) et, dans la mesure du possible, un score supplémentaire de collaboration était attribué à la cohorte initiale comme variable indépendante. (La collaboration d'un individu est calculée en pourcentage des questions auxquelles il aurait répondu à un cycle précédent de l'ELNEJ.) L'ajustement du poids pour la non-réponse est calculé pour chaque GRH à l'aide de la formule présentée plus haut.

En ce qui concerne la cohorte initiale et les cohortes du DPE, le nombre de GRH utilisé se situait entre 6 et 13. Même si les cohortes du DPE ont été modélisées ensemble, l'ajustement pour la non-réponse a été calculé séparément pour la simple et bonne raison qu'elles représentent leur population de référence distincte.

Second ajustement : ajustement pour la stratification a posteriori

En ce qui concerne la cohorte initiale, la population de référence est constituée de la population de tous les enfants âgés de 0 à 11 ans au 31 décembre 1994. Pour ce qui est des cohortes du DPE, les populations de référence sont énumérées à la section 11.1.2. Chaque groupe a été stratifié a posteriori aux chiffres de population applicables selon l'âge, le sexe et la province pour le mois de janvier suivant l'année de référence.

11.3.2 Pondération longitudinale entonnoir

Définition d'un répondant entonnoir

Un répondant entonnoir est un répondant longitudinal au cycle 8 (voir la définition ci-dessus) qui a aussi répondu à tous les cycles précédents. On décrit ci-dessous à quel cycle une cohorte en particulier aurait dû répondre.

<u>Cohorte</u>	<u>Cycles auxquels un répondant aurait dû répondre</u>
Cohorte initiale	1, 2, 3, 4, 5, 6, 7, 8
Cohorte du DPE du cycle 5	5, 6, 7, 8
Cohorte du DPE du cycle 6	6, 7, 8
Cohorte du DPE du cycle 7	7, 8

Il est bon de noter que les personnes qui ont répondu à tous les cycles (par exemple, les répondants entonnoir) constituent un sous-ensemble des répondants au cycle 8.

Premier ajustement : ajustement pour la non-réponse

L'ajustement pour la non-réponse repose encore une fois sur la création de GRH. Les groupes de non-réponse pour la cohorte initiale sont établis à l'aide de variables de l'EPA et d'un score de collaboration. Les groupes de non-réponse pour les cohortes du DPE sont établis à l'aide des variables de l'EPA. Les GRH sont produits avec les extrants du modèle. Il est bon de noter que ces GRH sont différents des GRH établis pour la pondération longitudinale puisque les répondants au cycle 8 qui n'ont pas répondu à tous les cycles précédents affichent un mécanisme de non-réponse différent de celui des personnes qui ont répondu à tous les cycles. Le facteur d'ajustement est calculé pour chaque GRH.

Second ajustement : ajustement pour la stratification a posteriori

En ce qui concerne la cohorte initiale, la population de référence est constituée de la population de tous les enfants âgés de 0 à 11 ans au 31 décembre 1994. Pour ce qui est des cohortes du DPE, les populations de référence sont énumérées à la section 11.1.2. Chaque groupe a été stratifié a posteriori aux chiffres de population applicables selon l'âge, le sexe et la province pour le mois de janvier suivant l'année de référence.

11.3.3 Pondération transversale

Définition d'un répondant transversal

Un répondant transversal est un enfant pour lequel la composante adulte ou la composante enfant est remplie. À la différence du répondant longitudinal, les enfants introduits à un cycle précédent qui sont décédés ou qui ont déménagé à l'extérieur des dix provinces canadiennes sont des enfants inadmissibles. Ils ne font pas partie de la population cible.

Premier ajustement : ajustement pour la non-réponse

Un poids transversal est attribué aux enfants de l'ELNEJ qui avaient entre 0 et 7 ans en 2008. Des modèles de non-réponse ont été établis à partir des variables de l'EPA. Les probabilités estimées de réponse à partir des modèles ont par la suite été combinées, en vue de constituer les GRH. À l'intérieur de chaque GRH, on a procédé à un ajustement pour la non-réponse.

Second ajustement : stratification a posteriori

La population de référence est les enfants qui avaient entre 0 et 7 ans au 31 décembre 2008. L'ajustement est calculé pour chaque ensemble âge-sexe-province pour les chiffres de janvier 2009.

12.0 Qualité des données, taux de réponse et couverture

Le présent chapitre permet à l'utilisateur de prendre connaissance des divers éléments qui influent sur la qualité des données de l'enquête. On distingue deux principaux types d'erreurs : les erreurs d'échantillonnage et les erreurs non dues à l'échantillonnage. On accorde ici une attention particulière aux erreurs non dues à l'échantillonnage.

Pour obtenir plus de renseignements d'ordre général sur la qualité des données de l'enquête et l'assurance de la qualité veuillez consulter le site www.statcan.gc.ca.

12.1 Erreur d'échantillonnage

Les estimations établies à partir des données de l'enquête sont fondées sur un échantillon d'enfants. On aurait pu obtenir des valeurs légèrement différentes si l'on avait procédé à un recensement complet de la population ciblée en faisant appel aux mêmes questionnaires et aux mêmes intervieweurs, superviseurs, méthodes de traitement, entre autres. L'écart entre les estimations reposant sur un échantillon et celles obtenues à partir d'un dénombrement complet effectué dans des conditions similaires est appelé erreur d'échantillonnage des estimations.

L'erreur d'échantillonnage peut être estimée à l'aide de la variance d'échantillonnage. Le chapitre 13.0 fournit plus de renseignements sur le calcul de cette estimation.

12.2 Erreurs non dues à l'échantillonnage

Dans toute enquête, les sources d'erreurs non dues à l'échantillonnage sont nombreuses. Il arrive que des intervieweurs comprennent mal les instructions, que des répondants ne répondent pas adéquatement aux questions, que des réponses soient mal inscrites sur le questionnaire ou que des erreurs se produisent au moment du traitement des données. Ce sont là des erreurs non dues à l'échantillonnage pouvant difficilement être quantifiées. D'autres types d'erreurs, particulièrement les erreurs de non-réponse et les erreurs de couverture de la population cible, sont plus facilement quantifiables.

Une erreur non due à l'échantillonnage peut entraîner un biais, c'est-à-dire un écart entre la valeur estimative prévue et la valeur vraie de la population. Comme les valeurs vraies de la population ne sont pas connues, il est très difficile de mesurer les biais.

12.3 Non-réponse totale et biais de non-réponse

Dans les enquêtes, la non-réponse découle de l'impossibilité d'obtenir un ensemble de mesures pour des unités sélectionnées dans l'échantillon. Il y a deux types de non-réponse : la non-réponse totale, où l'unité sélectionnée ne répond à aucune question donc libellée ainsi comme non-répondante et la non-réponse partielle, où l'unité sélectionnée répond à suffisamment de questions pour être considérée comme répondante, sans toutefois répondre à toutes les questions. La présente section porte sur la non-réponse totale, et la section 12.4, sur la non-réponse partielle.

La non-réponse peut biaiser les estimations de l'enquête. Les estimations obtenues peuvent être biaisées si les non-répondants présentent des caractéristiques très différentes de celles des répondants. Tant l'importance de la non-réponse que la mesure dans laquelle les non-répondants auraient déclaré des réponses différentes de celles des répondants influent sur l'ampleur du biais dans les estimations. Nous sommes dans l'impossibilité de mesurer exactement ce que les non-répondants « auraient déclaré », mais nous pouvons mesurer le niveau de non-réponse. Les taux de réponse transversaux et les taux d'attrition longitudinaux sont indiqués plus loin dans la section. De plus, on trouvera davantage de précisions sur le procédé de pondération et sur la façon dont elle est utilisée pour l'ajustement de non-réponse totale au chapitre 11.0. Dans le chapitre 13.0, on donne un aperçu du biais de non-réponse.

Nos procédés de pondération ajustent les poids d'échantillonnage pour essayer de réduire le biais potentiel dû à la non-réponse. Cette pratique se fonde toutefois sur certaines hypothèses, et rien ne garantit qu'il n'y a pas de biais en raison de non-réponse.

Dans le cas de l'Enquête longitudinale nationale sur les enfants et les jeunes (ELNEJ), on a créé des groupes de réponses homogènes (GRH), de sorte que les poids des répondants compensent pour les non-répondants ayant une propension prédite à fournir des réponses semblables, là où cette propension prédite à répondre repose sur des caractéristiques recueillies précédemment (niveau de scolarité, type de logement, entre autres). Néanmoins, quel que soit le GRH, les non-répondants peuvent différer des répondants de diverses façons importantes encore non observées ou inconnues.

La non-réponse augmente au fil du temps. Comme on a de moins en moins de participants, l'erreur d'échantillonnage estimée s'accroît, tout comme d'ailleurs le risque de biais. Après de nombreux cycles, il serait fort improbable que les participants qui continuent de collaborer constituent un sous-échantillon aléatoire des répondants du cycle 1.

En fait, dans des cas extrêmes, certains sous-ensembles de la population peuvent ne plus être représentés dans ce qui reste de l'échantillon. Voici un exemple purement hypothétique. Supposons que l'échantillon initial contenait 20 jeunes filles autistes dans une province donnée et qu'aucune de celles-ci ne répondent au cycle 8. Peu importe le procédé de pondération, il serait désormais impossible de produire des estimations concernant les jeunes filles autistes dans cette province.

12.3.1 Définitions de la réponse

Il y a deux différents types de taux de réponse : le **taux de réponse à la phase de la collecte**, qui mesure l'efficacité du processus de collecte des données et est établi en fonction des unités auprès desquelles on recueille des données, et le **taux de réponse à la phase de l'estimation**, qui constitue un indicateur de la qualité des estimations produites. Le taux de réponse à la phase de l'estimation est établi en fonction de l'unité statistique (l'enfant, dans le cas de l'ELNEJ), et il vise à montrer la mesure dans laquelle des données sont manquantes.

Dans le contexte de l'ELNEJ, les enfants qui demeurent admissibles mais qui ne faisaient pas partie de l'échantillon pour la collecte du cycle 8 constituent la principale différence entre les deux types de taux de réponse. Par exemple, les unités avec des refus catégoriques dans un cycle précédent sont retirées de l'échantillon du cycle 8, et leur absence influe sur la qualité des données. De telles unités sont comprises dans les taux de réponse à la phase de l'estimation mais ne le sont pas dans les taux de réponse à la phase de la collecte. Des taux de réponse à la phase de la collecte au niveau de l'enfant sont donnés aux chapitres 5.0 et 6.0. Les taux de réponse à la phase de l'estimation indiqués dans la présente section ne contredisent pas les taux à la phase de la collecte. Ils visent à les compléter et à présenter des renseignements légèrement différents.

Taux pondérés et taux non pondérés

Il arrive souvent que les taux de réponse à la phase de l'estimation soient pondérés de façon à tenir compte du fait que la non-réponse de certaines unités (c'est-à-dire les unités dont les poids d'échantillonnage sont importants dans les enquêtes sociales ainsi que les unités ayant une grande incidence sur les estimations dans les enquêtes auprès des entreprises) cause plus de dommage aux estimations de l'enquête que la non-réponse d'unités moins importantes. Ce raisonnement se défend.

L'utilisation des taux de réponse non pondérés présente également des avantages. Il est plus facile de les définir et de les produire (et peut-être de les interpréter), ils ne sont pas touchés par les révisions des poids de l'enquête et ils ont été utilisés dans les cycles précédents de l'ELNEJ. En outre, on observe très peu de différence dans les provinces entre les taux pondérés et les taux non pondérés pour l'ELNEJ.

Les taux de réponse dont il est question dans le présent chapitre sont **non pondérés**.

Définitions

Les définitions suivantes sont essentielles à la compréhension des tableaux du présent chapitre.

Les enfants et les jeunes de différents groupes d'âge doivent répondre à différentes composantes de l'enquête. Par ailleurs, les critères à respecter pour être considéré comme un répondant varient selon l'âge de la personne sélectionnée.

La composante de l'enfant prend la forme d'une interview assistée par ordinateur (IAO) dans le cadre de laquelle les questions concernant l'enfant sont posées à la personne la mieux renseignée (PMR). La composante de la PMR est aussi une IAO où les questions concernant la PMR sont posées à la PMR. La composante du conjoint est une IAO où les questions sur le conjoint sont posées directement au conjoint de la PMR. Les composantes de la PMR et de son conjoint sont parfois appelées composantes de l'adulte. Enfin, la composante du jeune est une IAO où les questions concernant le jeune sont posées au jeune. Une composante pour laquelle on a répondu à un ensemble de questions clés est considérée comme **remplie**. Il est bon de noter qu'une composante considérée comme remplie peut présenter une importante non-réponse partielle (voir la section 12.4).

L'enfant dont l'âge effectif se situe entre 0 et 15 ans est considéré comme un **répondant** si la composante de l'enfant ou une composante de l'adulte (PMR ou son conjoint) est remplie. Un jeune dont l'âge effectif est de 16 ou 17 ans est considéré comme un **répondant** si la composante de l'enfant ou du jeune ou si une composante de l'adulte (PMR ou son conjoint) est remplie. Le jeune dont l'âge effectif est de 18 ans et plus est considéré comme un **répondant** si la composante du jeune est remplie. Aucune composante de l'adulte n'est administrée pour les jeunes de 18 ans et plus.

Un **enfant inadmissible** est un enfant qui a été sélectionné dans l'échantillon mais qui ne fait pas partie de la population observée. Un enfant peut être inadmissible parce qu'il est décédé, qu'il réside à l'extérieur du Canada ou qu'il réside en établissement. À l'opposé, un **enfant admissible** est un enfant sélectionné dans l'échantillon **et** fait partie de la population cible. La somme du nombre d'enfants admissibles et inadmissibles égale la taille de l'échantillon. Il est bon de noter que certains enfants peuvent être inadmissibles du point de vue transversal mais admissibles sur le plan longitudinal. Cette situation survient, par exemple, avec les enfants décédés ou déménagés hors du pays qui ne représentent plus personne sur le plan transversal dans la population cible. En revanche, du point de vue longitudinal, ces enfants représentent d'autres enfants dans la même situation et présents dans la population cible longitudinale au moment de leur introduction dans l'enquête. D'autre part, il est également possible d'avoir des enfants qui sont admissibles sur le plan transversal mais inadmissibles sur le plan longitudinal (les enfants de l'échantillon complémentaire, par exemple).

Un **non-répondant** est un enfant ou un jeune admissible qui ne répond pas à nos critères relatifs à la réponse. Il peut y avoir non-réponse parce que les participants ciblés refusent de participer à l'enquête (ou ne donnent pas suffisamment de réponses), parce que l'enfant ou le jeune n'a pu être retracé ou parce que l'intervieweur n'a pu terminer l'interview pour d'autres raisons.

Le **taux de réponse à la phase de l'estimation** correspond au nombre d'enfants ou de jeunes répondants par rapport au nombre d'enfants ou de jeunes admissibles dans l'échantillon initial.

Le **taux de réponse à la phase de la collecte** correspond au nombre d'enfants ou de jeunes répondants par rapport au nombre d'enfants ou de jeunes admissibles parmi les unités envoyées à la collecte.

12.3.2 Taux de réponse transversaux

L'échantillon transversal du cycle 8 est formé de l'agrégation des enfants sélectionnés dans quatre cohortes différentes, celles des cycles 5, 6, 7 et 8.

Taux de réponse transversaux non pondérés selon la cohorte échantillonnée, cycle 8

Âge effectif (ans)	Échantillon initial		Envoyés à la collecte		Répondants	Taux de réponse à la phase de la collecte (%)	Taux de réponse à la phase de l'estimation (%)
	Taille de l'échantillon	Admissible	Taille de l'échantillon	Admissible			
0 et 1	5 479	5 463	5 479	5 463	4 106	75,2	75,2
2 et 3 (de retour)	4 991	4 946	4 964	4 940	3 886	78,7	78,6
2 et 3 Échant. comp.	615	615	615	615	486	79,0	79,0
4 et 5 (de retour)	4 857	4 788	4 816	4 785	3 658	76,4	76,4
4 et 5 Échant. comp.	588	587	588	587	472	80,4	80,4
6 et 7 (de retour)	5 568	5 406	4 271	4 256	3 450	81,1	63,8
Total	22 098	21 805	20 733	20 646	16 058	77,8	73,6

Définitions du titre des colonnes

Âge effectif — Établit d'après l'année de naissance. Ceux dont l'âge effectif est 0 année sont nés en 2008; ceux dont l'âge effectif est de 1 an sont nés en 2007, et ainsi de suite.

Échantillon initial/Taille de l'échantillon – Nombre total d'enfants sélectionnés dans la base de sondage, y compris les enfants déclarés inadmissibles lors de la collecte initiale ou d'une collecte suivante.

Échantillon initial/Admissible – Nombre d'enfants admissibles sur le plan transversal au cycle 8. Ce nombre constitue le dénominateur du taux de réponse à la phase de l'estimation. Voir la définition complète à la section 12.3.1.

Envoyés à la collecte/Taille de l'échantillon – Nombre d'enfants qui ont été envoyés à la collecte au cycle 8.

Envoyés à la collecte/Admissibles – Nombre d'enfants admissibles sur le plan transversal parmi les enfants qui ont été envoyés à la collecte au cycle 8. Ce nombre constitue le dénominateur du taux de réponse à la phase de la collecte.

Répondants – Numérateur des taux de réponse. Ces enfants respectent nos critères relatifs à la réponse énoncés à la section précédente.

Taux de réponse à la phase de la collecte – Ratio du nombre de répondants et du nombre d'enfants admissibles envoyés à la collecte.

Taux de réponse à la phase de l'estimation – Ratio du nombre de répondants et du nombre d'enfants admissibles dans l'échantillon initial.

Source : Statistique Canada, Enquête longitudinale nationale sur les enfants et les jeunes.

Le tableau ci-dessus illustre la différence entre le taux de réponse à la phase de la collecte et le taux de réponse à la phase de l'estimation. Par exemple, en ce qui concerne les enfants de retour dont l'âge effectif était de 6 ou 7 ans, le taux de réponse à la phase de la collecte se situait à 81,1 %. Cependant, si l'on prend en compte le fait que plusieurs

centaines d'enfants admissibles ont été retirés de l'échantillon du cycle 8, tout particulièrement en raison de la non-réponse au cycle 5, le taux de réponse à la phase de l'estimation est considérablement plus faible (63,8 %).

Au cycle 7, bien que nous ayons commencé à réintégrer des non-répondants des cycles précédents, nous ne l'avons pas fait pour ceux qui étaient des non-répondants à leur premier cycle. Au cycle 5, il y avait 1 142 non-répondants initiaux. Ceux-ci étaient tous admissibles sur le plan transversal au cycle 7, mais n'ont pas été envoyés au terrain pour la collecte. Cela a donné lieu à un important écart entre l'échantillon initial et les cas envoyés à la collecte, contribuant au taux de réponse faible à l'estimation.

En revanche, pour les 4 et 5 ans de retour au cycle 8, on a renvoyé à la collecte ceux qui au départ n'avaient pas répondu au cycle 6. En tout, 852 non-répondants initiaux du cycle 6 ont été envoyés à la collecte au cycle 8.

La stratégie de collecte du cycle 7 a entraîné des taux de réponse faibles à l'estimation, bien que les taux de réponse à la collecte étaient élevés. Au cycle 8, les deux taux sont essentiellement les mêmes, exception faite des taux pour les 6 et 7 ans de retour, les non-répondants initiaux du cycle 5 n'étant toujours pas envoyés au terrain pour la collecte.

Le prochain tableau montre ces taux à la phase de l'estimation selon la province.

Taux de réponse à la phase de l'estimation, transversaux, non pondérés selon la province, cycle 8

Province	Échantillon admissible	Répondants	Taux de réponse à la phase de l'estimation (%)
Terre-Neuve-et-Labrador	1 435	1 192	83,1
Île-du-Prince-Édouard	1 171	893	76,3
Nouvelle-Écosse	1 380	1 063	77,0
Nouveau-Brunswick	1 517	1 154	76,1
Québec	2 957	2 280	77,1
Ontario	5 064	3 511	69,3
Manitoba	1 787	1 274	71,3
Saskatchewan	1 824	1 405	77,0
Alberta	2 516	1 811	72,0
Colombie-Britannique	2 154	1 475	68,5
Canada	21 805	16 058	73,6

Source : Statistique Canada, Enquête longitudinale nationale sur les enfants et les jeunes.

Les chiffres de la colonne *Échantillon admissible* sont fondés sur la province de sélection, c'est-à-dire la province où l'enfant a été choisi. Les chiffres de la colonne *Répondants* sont fondés sur la province de résidence au cycle 8 (connue uniquement dans le cas des répondants).

12.3.3 Taux d'attrition longitudinaux pour la cohorte du cycle 1

Dans une enquête longitudinale, le taux de réponse longitudinal indique la proportion de répondants participant encore à l'enquête. Normalement, ce taux représente le ratio du nombre d'enfants longitudinaux qui répondent au cycle actuel sur le nombre d'enfants

sélectionnés au premier cycle. Cependant, comme la méthode de sélection utilisée lors de la collecte des deux premiers cycles diffère de celle utilisée dans les cycles suivants, il est impossible de calculer de façon précise un taux de réponse longitudinal qui tient compte de la non-réponse au cycle 1 pour les enfants sélectionnés à ce même cycle. Plus précisément, le nombre réel d'enfants admissibles dans les ménages non-répondants n'est pas connu. Par conséquent, le dénominateur requis pour déterminer le taux de réponse est aussi inconnu.

On définit plutôt le **taux d'attrition** comme la proportion de répondants participant toujours à l'enquête par rapport au nombre de **répondants** au cycle 1.

Taux d'attrition longitudinaux non pondérés pour les enfants du cycle 1 selon l'âge effectif

Âge effectif au cycle 8 (ans)	Nombre de répondants au cycle 1	Données recueillies pour le cycle 8		Avec poids longitudinal au cycle 8 ¹		Avec poids entonnoir au cycle 8 ²	
		Nombre	% du cycle 1	Nombre	% du cycle 1	Nombre	% du cycle 1
14 à 17 ³	6 967	4 271	61,3	4 414	63,4	3 624	52,0
18 à 25	9 936	5 937	59,8	6 073	61,1	3 848	38,7
Total	16 903	10 208	60,4	10 487	62,0	7 472	44,2

Notes : Le dénominateur pour tous les pourcentages présentés dans ce tableau est le nombre d'enfants répondants au cycle 1 qui ont été suivis au cycle 2. Comme la taille de l'échantillon a été réduite du cycle 1 au cycle 2, les enfants répondants au cycle 1 ne sont pas tous considérés comme longitudinaux.

Afin de réduire au minimum l'attrition des enfants de la cohorte initiale, des efforts sont faits afin d'obtenir une réponse pour les enfants n'ayant pas répondu à un cycle précédent. Il est donc possible d'avoir des enfants qui n'ont pas répondu lors d'un cycle donné mais qui ont répondu dans le cycle courant.

1. L'enfant à qui un poids longitudinal a été attribué au cycle 8 peut être classé dans deux catégories : ceux pour lesquels des données ont été recueillies et ceux pour lesquels aucune donnée n'a été recueillie. Les enfants qui sont décédés ou qui n'habitent plus dans l'une des dix provinces se trouvent dans la seconde catégorie. Aucune donnée n'a été recueillie concernant ces enfants, mais ceux-ci ont tout de même encore un poids longitudinal car ils représentent d'autres enfants dans la population longitudinale qui sont dans la même situation. D'une façon plus courante, l'enfant ou le jeune qui a un poids longitudinal a répondu aux questions de l'enquête. Ceux pour lesquels des données ont été recueillies sont dans la colonne *Données recueillies pour le cycle 8* et constituent un sous-ensemble de ceux auxquels un poids longitudinal a été attribué.
2. La colonne *Avec poids entonnoir au cycle 8* montre le nombre d'enfants qui n'ont jamais été considérés comme des non-répondants. En d'autres termes, un poids longitudinal leur a été attribué à chaque cycle dans le fichier principal. Précisons de nouveau que des données ont été recueillies à chaque cycle pour la majorité de ces enfants. Un nombre beaucoup moins élevé d'enfants, soit ceux qui sont décédés ou qui sont déménagés à l'extérieur du Canada, conservent toujours un poids même si aucune donnée les concernant n'a été recueillie. Voir le chapitre 11.0 pour plus de renseignements sur les poids entonnoir (variable HWTCWd1L). À signaler que le léger écart entre le nombre total d'unités qui possèdent un poids entonnoir et le nombre total d'unités qui sont des « répondants à tous les cycles » dans le tableau ci-après est attribuable aux changements dans le concept de l'« âge » qui s'est produit dans le cycle 4. Quelques unités qui ont un poids entonnoir ont été considérés longitudinalement inadmissibles pour le cycle 4 seulement de sorte qu'on ne leur a pas attribué un poids longitudinal pour ce cycle.
3. À la première rangée du tableau, le répondant principal est la PMR. Les jeunes de 18 ans et plus répondent eux-mêmes.

Source : Statistique Canada, Enquête longitudinale nationale sur les enfants et les jeunes.

Le tableau suivant montre l'attrition survenue au cours des différents cycles. Dans ce tableau, le mot « répondants » renvoie aux enfants auxquels un poids longitudinal a été attribué, y compris ceux pour lesquels aucune donnée n'a été recueillie. La proportion des répondants du cycle 1 qui avaient un poids longitudinal a diminué de presque 5 %, soit de

66,9 % à 62,0 %, entre le cycle 7 et le cycle 8. Bien que cette diminution n'était que de 1 % seulement entre le cycle 6 et le cycle 7, la diminution entre le cycle 7 et le cycle 8 est semblable à la diminution observée à chacun des cycles précédant le cycle 7.

Taux d'attrition longitudinaux non pondérés pour les enfants du cycle 1, (cycles 2 à 8) selon la province

Province au cycle 1	Répondants au cycle 1	Répondants au cycle 2		Répondants au cycle 3		Répondants au cycle 4		Répondants au cycle 5		Répondants au cycle 6		Répondants au cycle 7		Répondants au cycle 8		Répondants à tous les cycles	
	Nombre	Nombre	% du cycle 1	Nombre	% du cycle 1	Nombre	% du cycle 1	Nombre	% du cycle 1	Nombre	% du cycle 1	Nombre	% du cycle 1	Nombre	% du cycle 1	Nombre	% du cycle 1
Terre-Neuve-et-Labrador	950	892	93,9	845	88,9	777	81,8	755	79,5	689	72,5	679	71,5	636	66,9	491	51,7
Île-du-Prince-Édouard	467	443	94,9	434	92,9	392	83,9	364	77,9	339	72,6	340	72,8	305	65,3	227	48,6
Nouvelle-Écosse	1 191	1 068	89,7	1 085	91,1	988	83,0	903	75,8	839	70,4	809	67,9	748	62,8	536	45,0
Nouveau-Brunswick	1 070	958	89,5	958	89,5	836	78,1	792	74,0	710	66,4	685	64,0	678	63,4	449	42,0
Québec	3 182	2 944	92,5	2 844	89,4	2 522	79,3	2 361	74,2	2 108	66,2	2 147	67,5	2 093	65,8	1 453	45,7
Ontario	4 342	3 899	89,8	3 760	86,6	3 318	76,4	3 104	71,5	2 834	65,3	2 804	64,6	2 598	59,8	1 809	41,7
Manitoba	1 232	1 161	94,2	1 112	90,3	1 019	82,7	1 004	81,5	905	73,5	868	70,5	743	60,3	568	46,1
Saskatchewan	1 413	1 305	92,4	1 257	89,0	1 073	75,9	1 002	70,9	958	67,8	971	68,7	906	64,1	640	45,3
Alberta	1 599	1 465	91,6	1 420	88,8	1 242	77,7	1 162	72,7	1 109	69,4	1 083	67,7	968	60,5	686	42,9
Colombie-Britannique	1 457	1 333	91,5	1 282	88,0	1 143	78,4	1 076	73,9	992	68,1	923	63,3	812	55,7	596	40,9
Canada	16 903	15 468	91,5	14 997	88,7	13 310	78,7	12 523	74,1	11 483	67,9	11 309	66,9	10 487	62,0	7 455	44,1

Source : Statistique Canada, Enquête longitudinale nationale sur les enfants et les jeunes.

12.3.4 Taux de réponse longitudinaux pour les enfants sélectionnés aux cycles 5 et 6

Les taux de réponse à la phase de l'estimation indiqués aux rangées des 2 et 3 ans (*de retour*), des 4 et 5 ans (*de retour*), des 6 et 7 ans (*de retour*) du tableau dans la section 12.3.2, donnent une bonne idée des taux de réponse longitudinaux pour les enfants sélectionnés aux cycles 7, 6, et 5 respectivement.

12.4 Non-réponse partielle

La section précédente sur la non-réponse totale a porté sur les problèmes de représentation de l'échantillon relativement à l'estimation transversale ou longitudinale. Ces erreurs non dues à l'échantillonnage peuvent habituellement être rectifiées efficacement par l'ajustement des poids de l'enquête afin de refléter la population observée. D'autres types de non-réponse sont également mesurés dans cette enquête et ne sont habituellement pas corrigés par un ajustement des poids de l'enquête.

Même si une personne donne suffisamment d'information pour être considérée comme un répondant, certaines questions (variables) peuvent rester sans réponse, entraînant une non-réponse partielle. Plusieurs raisons peuvent expliquer cela, notamment les suivantes (sans ordre particulier) : collaboration de certains participants ciblés du ménage (mais non de tous les participants), refus de répondre à des questions de nature délicate, lassitude des répondants, omission accidentelle de certaines parties du questionnaire ou difficultés opérationnelles.

La non-réponse à une composante est une situation qui se produit lorsqu'une personne participe à l'enquête et que d'autres n'y participent pas. Par exemple, dans le ménage d'un enfant de 17 ans sélectionné, la PMR peut collaborer et répondre aux composantes de l'enfant et de la PMR, alors que son conjoint peut refuser de répondre à la composante du conjoint et que le jeune sélectionné peut refuser de répondre à la composante du jeune. Conformément à nos définitions, ce jeune est considéré comme un répondant et fait l'objet d'un enregistrement dans le fichier principal, mais on est devant un cas de non-réponse partielle. Il peut aussi y avoir non-réponse partielle dans les cas où les données du volet téléphonique sont recueillies mais où celles des autres composantes sont manquantes.

La nature de la non-réponse partielle dépend habituellement du sujet. Par exemple, la section sur le développement moteur et social est rigoureusement remplie pour les enfants de 0 à 3 ans, car les parents s'intéressent davantage au sujet. En revanche, les questions portant sur le revenu peuvent être jugées trop personnelles par certains répondants, entraînant des cas de non-réponse partielle.

La **non-réponse partielle** est mesurée pour chaque variable et représente l'information non recueillie auprès du répondant au moment de l'interview. Ce type de non-réponse n'est pas corrigé, sauf lorsqu'un indicateur d'imputation indique expressément le contraire. La non-réponse partielle est détaillée dans le livre des codes de l'enquête, qui comprend des catégories de réponse telles que « Refus » ou « Non déclaré ». La catégorie « Ne sait pas » est considérée comme une non-réponse pendant l'analyse, mais certains analystes peuvent considérer cette catégorie comme une réponse acceptable selon l'information recherchée et l'interprétation de variables précises. À des fins analytiques, les chercheurs doivent se rappeler que les catégories « Refus » et « Ne sait pas » sont utilisées lorsque le répondant a été interrogé au sujet de cette information particulière, alors que la catégorie « Non déclaré » indique habituellement que cette information n'a pas été demandée au répondant. Cela est vrai pour la saisie des réponses assistée par ordinateur. Cependant, les questionnaires autoadministrés sur format papier font exception. Pour ces questionnaires, les réponses laissées en blanc sont classées dans la catégorie « Non déclaré », bien qu'il soit possible que le répondant ait vu la question.

Il est bon de noter que la catégorie « Enchaînement valide » n'est pas une non-réponse, mais bel et bien un enchaînement valide d'une information particulière pour un répondant donné. Par exemple,

de nombreuses questions s'adressent à un groupe d'âge précis, et les enfants qui ne sont pas dans ce groupe d'âge ont un code « Enchaînement valide » pour ces variables.

Pour avoir des précisions sur la non-réponse à des questions précises, il faut consulter le livre des codes qui accompagne le fichier de microdonnées. Certains participants considérés comme des « répondants à une composante » peuvent n'avoir répondu qu'à une partie de la composante.

Les analystes qui ont recours aux données de l'ELNEJ doivent être conscients de l'incidence de la non-réponse partielle sur les données qu'ils tentent d'analyser. À l'instar de la non-réponse totale, la non-réponse partielle peut être plus fréquente dans le cas des répondants qui présentent une caractéristique particulière (les adolescents qui ont de piètres résultats scolaires peuvent se montrer plus réticents à remplir le questionnaire autoadministré, par exemple). Cela entraîne un biais qui, s'il est marqué, peut semer le doute quant aux résultats de l'analyse. Des techniques telles que la repondération et l'imputation permettent de remédier à la non-réponse partielle. On encourage vivement les utilisateurs de données à évaluer le degré de non-réponse dans le cadre de leur analyse et à prendre, s'il y a lieu, des mesures correctives pour compenser l'effet de la non-réponse. Ils doivent au moins exposer en détail l'incidence de la non-réponse à des composantes ou à des questions dans leurs résultats. Cet aspect est examiné plus tard dans le chapitre 15.0.

Les sections qui suivent étudieront les questions de la non-réponse aux composantes de l'ELNEJ. Cette étude a pour but d'indiquer aux chercheurs qui utilisent ces variables dans leurs analyses les sources d'erreurs possibles qui ne sont pas corrigées par les poids de l'enquête. Tous les taux présentés dans les tableaux suivants sont non pondérés, et le dénominateur correspond au nombre d'enfants admissibles pour la composante parmi ceux qui sont considérés comme des répondants. Ceci est différent de la non-réponse totale pour la composante. Par exemple, le taux de réponse pour la composante de l'EVIP est de 83,9%. Cependant, lorsque la non-réponse partielle et la non-réponse totale sont considérées ensemble, on obtient un taux de réponse de 64,5 % pour les 5 375 enfants admissibles pour lesquels un test de l'EVIP était souhaitable et lesquels l'ont complété.

En outre, comme le présent document a été rédigé avant la diffusion du fichier principal, il peut y avoir de légères différences entre les deux.

12.4.1 Composante de l'enfant

La composante de l'enfant est une IAO pour laquelle des questions sur l'enfant sont posées à la PMR. Tous les enfants dont l'âge effectif est de 0 à 17 ans sont admissibles, à l'exception de ceux qui ont 16 et 17 ans et qui vivent de façon autonome.

Il est bon de noter que le chiffre dans la colonne *Nombre admissible* est fondé sur le nombre d'enfants pour lesquels des données ont été recueillies. Ne sont pas compris, les enfants longitudinaux admissibles pour lesquels aucune donnée n'a été recueillie au cycle 8, mais qui figurent néanmoins dans le fichier principal avec un poids longitudinal.

Dans la colonne *Nombre de personnes ayant répondu*, le chiffre englobe les personnes ayant répondu ou non à toutes les questions.

Composante de l'enfant — taux de réponse

Nombre admissible	Nombre de personnes ayant répondu	Taux de réponse à la composante (%)
20 305	20 212	99,5

Source : Statistique Canada, Enquête longitudinale nationale sur les enfants et les jeunes.

Il est extrêmement rare que des participants ne répondent à aucune question de la composante de l'enfant. Si l'on n'a pas répondu à la composante de l'enfant, on a néanmoins répondu à la composante de l'adulte ou du jeune.

12.4.2 Composante de la personne la mieux renseignée

La composante de la PMR est une IAO dans le cadre de laquelle des questions sur la PMR sont posées à la PMR. Tous les enfants dont l'âge effectif est de 0 à 17 ans sont admissibles, à l'exception de ceux qui ont 16 et 17 ans et qui vivent de façon autonome.

Le tableau qui suit est établi en fonction des enfants et non des adultes. Dans les ménages comptant deux enfants sélectionnés, il n'y a qu'une seule PMR. On souhaite recueillir des renseignements auprès de 19 619 PMR. Pour ce qui est des enfants, on souhaite en avoir 20 305 pour obtenir des renseignements sur les PMR.

Composante de la personne la mieux renseignée — taux de réponse

Nombre admissible	Nombre de personnes ayant répondu	Taux de réponse à la composante (%)
20 305	19 817	97,6

Source : Statistique Canada, Enquête longitudinale nationale sur les enfants et les jeunes.

Il est rare que des participants ne répondent à aucune question de la composante de la PMR.

12.4.3 Composante du conjoint

La composante du conjoint est une IAO dans le cadre de laquelle des questions sur le conjoint sont posées au conjoint. Tous les enfants dont l'âge effectif est de 0 à 17 ans et dont la PMR a un conjoint sont admissibles. Comme il n'y a pas de PMR pour les jeunes de 16 et 17 ans qui vivent de façon autonome, il n'y a évidemment pas de conjoint de la PMR.

Le tableau qui suit est établi en fonction des enfants et non des adultes. Dans les ménages comptant deux enfants sélectionnés, un seul conjoint répond aux questions. On souhaite recueillir des renseignements auprès de 16 594 personnes. Pour ce qui est des enfants, on souhaite en avoir 17 141 pour obtenir des renseignements sur les conjoints. Les enfants vivant avec un parent seul ne sont pas admissibles pour cette composante.

Composante du conjoint — taux de réponse

Nombre admissible	Nombre de personnes ayant répondu	Taux de réponse à la composante (%)
17 141	16 560	96,6

Source : Statistique Canada, Enquête longitudinale nationale sur les enfants et les jeunes.

Il est rare que des participants ne répondent à aucune question de la composante du conjoint. Il est bon de noter que les réponses par personne interposée étaient permises pour cette composante lorsque la PMR pouvait répondre au nom du conjoint. Veuillez consulter la section 12.10 du présent guide pour obtenir plus de renseignements sur les réponses par personne interposée.

12.4.4 Composante du jeune

La composante du jeune est une IAO dans le cadre de laquelle des questions sur le jeune sont posées au jeune.

Le taux de réponse de la composante n'est significatif que pour les jeunes de 16 et 17 ans. Les jeunes de 18 ans et plus doivent remplir la composante du jeune pour être considérés comme des répondants.

Composante du jeune — taux de réponse

Nombre admissible	Nombre de personnes ayant répondu	Taux de réponse à la composante (%)
1 746	1 590	91,1

Source : Statistique Canada, Enquête longitudinale nationale sur les enfants et les jeunes.

Il est relativement courant qu'un adulte participe alors que le jeune de 16 ou 17 ans ne participe pas.

12.4.5 Échelle de vocabulaire en images de Peabody

Les enfants de 4 et 5 ans ont subi le test de l'Échelle de vocabulaire en images de Peabody (EVIP). L'Échelle de vocabulaire en images de Peabody (EVIP) est conçue pour mesurer le vocabulaire compris.

Échelle de vocabulaire en images de Peabody — taux de réponse

Nombre admissible	Nombre de personnes ayant répondu	Taux de réponse à la composante (%)
4 130	3 467	83,9

Source : Statistique Canada, Enquête longitudinale nationale sur les enfants et les jeunes.

12.4.6 Connaissance des nombres

La connaissance des nombres a été évaluée chez les enfants de 4 et 5 ans. L'évaluation de la connaissance des nombres vise à évaluer comment la compréhension des nombres se développe chez les enfants, grâce à l'examen de la compréhension du système des nombres entiers.

Connaissance des nombres — taux de réponse

Nombre admissible	Nombre de personnes ayant répondu suffisamment pour obtenir un score	Taux de réponse à la composante (%)
4 130	3 485	84,4

Source : Statistique Canada, Enquête longitudinale nationale sur les enfants et les jeunes.

12.4.7 Qui suis-je?

Les enfants de 4 et 5 ans ont subi le test « Qui suis-je? » Le module Qui suis-je? vise à évaluer le niveau de développement des jeunes enfants.

Qui suis-je? — taux de réponse

Nombre admissible	Nombre de personnes ayant répondu suffisamment pour obtenir un score	Taux de réponse à la composante (%)
4 130	3 461	83,8

Source : Statistique Canada, Enquête longitudinale nationale sur les enfants et les jeunes.

12.4.8 Tests de mathématiques

Les tests de mathématiques de l'ELNEJ sont composés de 20 questions de calcul auxquelles les répondants de 7 à 15 ans doivent répondre à la maison. Le niveau du test (de 2 à 10) est déterminé par le niveau scolaire de l'enfant. Si l'on ne connaît pas ce niveau de scolarité, l'âge effectif de l'enfant déterminera le niveau du test.

Tests de mathématiques — taux de réponse

Nombre admissible	Nombre de personnes ayant répondu suffisamment pour obtenir un score	Taux de réponse à la composante (%)
4 315	3 470	80,4

Source : Statistique Canada, Enquête longitudinale nationale sur les enfants et les jeunes.

12.4.9 Résolution de problèmes

Le test de résolution de problèmes s'adresse aux jeunes de 16 et 17 ans faisant partie de l'ELNEJ et comprend 18 questions qui visent à évaluer les points forts dans les domaines de la compréhension de la lecture, de la résolution de problèmes et de la prise de décisions. Le test porte également sur certaines habiletés mathématiques.

Résolution de problèmes — taux de réponse

Nombre admissible	Nombre de personnes ayant répondu suffisamment pour obtenir un score	Taux de réponse de la composante (%)
1 770	1 413	79,8

Source : Statistique Canada, Enquête longitudinale nationale sur les enfants et les jeunes.

12.4.10 Évaluation de la littératie

Le test d'évaluation de la littératie, que subissent les jeunes de 18 et 19 ans, comprend 36 questions mettant l'accent sur l'extraction de renseignements dans des textes, des tableaux et des graphiques. Une visite en personne s'impose pour ce test, tandis que la composante du jeune peut être remplie au téléphone.

Évaluation de la littératie — taux de réponse

Nombre admissible	Nombre de personnes ayant répondu suffisamment pour obtenir un score	Taux de réponse à la composante (%)
1 635	1 240	75,8

Source : Statistique Canada, Enquête longitudinale nationale sur les enfants et les jeunes.

12.4.11 Évaluation de la numératie

Le test d'évaluation de la numératie s'adresse aux jeunes de 20 et 21 ans et comprend 32 questions. Il vise à mettre à l'épreuve la capacité des jeunes adultes de fonctionner dans la société et de se débrouiller dans diverses situations d'ordre mathématique. Une visite en personne s'impose pour ce test, tandis que la composante du jeune peut être remplie au téléphone.

Évaluation de la numératie — taux de réponse

Nombre admissible	Nombre de personnes ayant répondu suffisamment pour obtenir un score	Taux de réponse à la composante (%)
1 366	1 011	74,0

Source : Statistique Canada, Enquête longitudinale nationale sur les enfants et les jeunes.

12.4.12 Questionnaires autoadministrés

La composante des questionnaires autoadministrés de l'ELNEJ consiste en un court livret comprenant des questions sur des sujets de nature intime, comme la mauvaise conduite, les sentiments, les parents et la puberté. Les sujets varient selon le groupe d'âge. Il s'agit de questionnaires que l'enfant remplit lui-même en privé, loin des parents et de l'intervieweur. Les questionnaires sont placés dans une enveloppe scellée remise personnellement à l'intervieweur lors de la visite.

Questionnaires autoadministrés — taux de réponse

Âge effectif (ans)	Nombre admissible	Nombre de personnes ayant répondu	Taux de réponse à la composante (%)
14 et 15	2 501	2 239	89,5
16 et 17	1 746	1 502	86,0
Total	4 247	3 741	88,1

Source : Statistique Canada, Enquête longitudinale nationale sur les enfants et les jeunes.

12.5 Non-réponse de cycle

Certains répondants longitudinaux ne participent pas à chaque cycle. C'est ce qu'on appelle la non-réponse de cycle. On ne dispose pas nécessairement pour chaque répondant de données longitudinales pour tous les cycles. Il peut arriver, par exemple, qu'un enfant ait répondu aux cycles 1, 5, 7 et 8 mais non aux cycles 2, 3, 4 et 6.

Si les données de tous les cycles sont essentielles, l'analyste peut restreindre le champ de son analyse aux enfants sans non-réponse de cycle et avoir recours aux poids entonnoir longitudinaux correspondant à ce groupe (variable HWTCWd1L).

12.6 Incidence des erreurs de réponse pour les caractéristiques rares

Les enquêtes auprès de la population en général se prêtent mal à la mesure de caractéristiques rares.

Des erreurs de réponse ou d'inscription se produisent au cours de la collecte des données. Par exemple, on prévoit qu'un certain pourcentage de répondants aux milliers d'interviews menées ne répondront pas honnêtement à toutes les questions. Il est aussi possible que l'intervieweur se trompe tout simplement de bouton. Dans la plupart des cas, ce type d'erreurs de déclaration n'a pas d'incidence importante. Pour beaucoup de variables, les erreurs s'équilibrent et leur incidence globale est minime. Cependant, si l'enquête sert à tirer des conclusions au sujet de caractéristiques, d'événements ou de comportements rares, ces erreurs de réponse peuvent devenir relativement plus importantes et lourdes de conséquences. On ne s'attend alors plus à ce que les erreurs s'équilibrent. Si les erreurs de réponse sont aléatoires, la caractéristique rare est systématiquement surestimée. Imaginons une enquête générale dans le cadre de laquelle on demande à 1 000 adultes (995 personnes n'étaient pas titulaires d'un doctorat et 5 personnes étaient titulaires d'un doctorat) quel est le niveau de scolarité le plus élevé qu'ils ont atteint. Il est beaucoup plus probable qu'un non-titulaire de doctorat déclare faussement qu'il possède un doctorat que l'inverse. Supposons que l'erreur de réponse associée à cette question soit de 0,2 % (0,2 % de 995 donne approximativement 2, et 0,2 % de 5 équivaut presque à zéro). L'estimation de l'enquête correspondant à la proportion de titulaires d'un doctorat serait de 7/1 000 plutôt que de 5/1 000. L'écart n'est pas prononcé, mais en termes relatifs, la surestimation de 40 % est importante et inquiétante. Il existe des techniques (comme celle qui consiste à poser plusieurs questions plutôt qu'une seule) qui permettent d'atténuer ce problème, mais elles ont pour effet d'allonger le questionnaire et de rendre l'enquête plus complexe. Vu le caractère général de l'ELNEJ, il n'était ni pratique ni possible d'accorder un tel degré d'attention à chaque élément d'information recueilli.

L'utilisateur des données de l'ELNEJ désireux d'étudier des comportements rares (forte consommation de drogue, comportement violent, ou autres) doit garder à l'esprit cette limite.

En outre, dans le cas de beaucoup de variables, il se peut que l'hypothèse de l'erreur de réponse aléatoire ne tienne pas, particulièrement pour les réponses considérées comme socialement indésirables. Cette question est analysée à la section 12.7. Par exemple, la probabilité qu'un jeune qui n'est pas violent déclare faussement qu'il manifeste un comportement violent peut différer de celle d'un jeune qui est violent et qui affirme à tort qu'il ne présente pas de comportement violent.

12.7 Erreurs de réponse liées à des comportements déviants ou à des questions de nature délicate

Dans une interview, les répondants ne disent pas toujours la vérité au sujet de leurs comportements considérés comme répréhensibles ou anormaux. C'est ce qu'on appelle le biais de désirabilité sociale. Par exemple, les parents qui infligent souvent des punitions corporelles à leurs enfants déformeront peut-être la vérité lorsqu'on les interrogera à ce sujet. De même, il se peut que certains

répondants mentent et donnent une image positive mais non réaliste d'eux-mêmes et de leurs enfants. Par exemple, certains parents ne répondent pas de façon honnête aux questions portant sur les lectures qu'ils font à leur enfant, reconnaissant qu'ils **devraient** en faire davantage.

Comme une grande partie des données d'enquête sont fournies par les répondants plutôt que d'être le résultat d'observations ou de mesures concrètes, il y a lieu de faire cette distinction lorsqu'on présente les conclusions d'une enquête. Par exemple, on ne peut conclure, à partir des résultats de l'ELNEJ, que X % des enfants au Canada reçoivent parfois une punition corporelle. En fait, l'enquête permet simplement d'affirmer que, selon les données **déclarées**, X % des enfants au Canada reçoivent parfois une punition corporelle.

12.8 Erreurs de réponse dues à des approximations

Même si cela peut sembler évident, il convient de mentionner que certaines valeurs recueillies font souvent l'objet d'approximations de la part du répondant. L'utilisateur de données doit être conscient du fait que les variables mesurant des concepts (le revenu, la taille, ou autres) et pouvant être considérés à juste titre comme continus dans la population ne conservent pas nécessairement ces traits caractéristiques dans le fichier d'enquête. Ainsi, on constate que le revenu déclaré correspond exactement, dans de nombreux cas, à un multiple de 10 000 \$ et que la taille est souvent arrondie au pouce exact (voir le graphique à la section 12.14.4). Au sein de la population, le nombre de ménages dont le revenu se situe entre 19 501 \$ et 20 500 \$ est probablement comparable au nombre de ménages dont le revenu oscille entre 20 501 \$ et 21 500 \$. Selon les résultats de l'enquête, la situation serait toute autre, le premier groupe étant beaucoup plus nombreux que le deuxième du fait que les répondants fournissent une approximation de leur revenu.

On observe aussi ce phénomène lorsqu'on demande l'âge qu'un enfant avait au moment d'un événement particulier. Par exemple, on demande l'âge de l'enfant en années et en mois au moment de la séparation de leurs parents. Pour ce qui est de la composante mois, le chiffre zéro est de loin celui qu'on mentionne le plus souvent.

12.9 Erreurs de réponse dues à des erreurs de mémoire

Un autre type d'erreur de réponse survient lorsqu'un répondant ne peut se souvenir de quelque chose de façon précise. Ce genre d'erreur est particulièrement associé à de longues périodes de référence. Par exemple, il arrive parfois que le répondant ne sache pas exactement combien de fois l'enfant a été examiné par un médecin au cours des 12 derniers mois. Souvent, le répondant oublie des maladies bénignes survenues plusieurs mois auparavant. En revanche, il arrive que le répondant se focalise sur des événements importants et indique que ceux-ci sont survenus au cours de la période de référence, alors qu'ils sont en fait antérieurs à celle-ci.

12.10 Erreurs de réponse dues aux réponses de la personne interposée

Dans le contexte de l'ELNEJ, il est permis de répondre par personne interposée aux composantes de l'adulte. Cela signifie que les renseignements concernant une personne sont communiqués par une autre personne (la mère de l'enfant qui répond à sa propre composante de la PMR et à la composante du conjoint au nom de son mari, par exemple). Normalement, un des membres du couple connaît suffisamment bien son conjoint pour répondre convenablement aux questions. Il est toutefois possible que la personne ciblée ait donné des réponses différentes de celles fournies par le répondant substitut.

Les taux de réponse par personne interposée sont suivis par les responsables de l'ELNEJ, mais aucune étude détaillée à ce sujet n'a été entreprise. Il est bien entendu qu'il est toujours impossible de savoir ce que le répondant non substitut aurait répondu.

Il est bon de noter que le tableau ci-dessous est établi en fonction des enfants et non des adultes.

Taux de personnes interposées, composantes de la personne la mieux renseignée et son conjoint

Composante	Enfants admissibles	Enfants ayant des réponses par personne interposée	Taux de réponse par personne interposée (%)
	Nombre		
PMR	20 305	257	1,3
Conjoint	17 141	11 221	65,5

Source : Statistique Canada, Enquête longitudinale nationale sur les enfants et les jeunes.

Le conjoint répond rarement à la composante de la PMR, mais il arrive très souvent que la PMR réponde aux deux composantes de l'adulte.

12.11 Schémas de réponse dont les catégories de réponse sont mal définies

Pour un grand nombre d'éléments d'information recueillies au moyen du questionnaire de l'ELNEJ, les catégories de réponse possibles sont mal définies, c'est-à-dire qu'elles ne sont pas définies concrètement et de façon précise (« Jamais », « Parfois » ou « Souvent », par exemple). Le sens qu'on donne aux mots « Parfois » et « Souvent » peut varier énormément d'une personne à une autre. Il en est de même pour « Entièrement d'accord » et « D'accord ». C'est ce qui explique la conséquence indésirable qu'on observe, alors que des répondants ayant les mêmes types de comportement ne présentent pas nécessairement des résultats d'enquête identiques. En règle générale, cela ne signifie pas que les données fondées sur des catégories de réponse mal définies soient inexactes ou peu fiables. Cependant, il faut être prudent lorsqu'on compare des groupes différents. Il faut savoir que les différences entre les régions ou les groupes ethniques quant aux schémas de réponse ne sont pas nécessairement attribuables à des différences réelles entre les enfants. Il se peut, par exemple, que la propension à répondre « Souvent » plutôt que « Parfois » s'explique par des facteurs culturels.

12.12 Langue de l'interview

En raison des nuances de la langue, il est impossible de traduire exactement certaines expressions ou questions. Cela contribue à introduire des différences artificielles dans les résultats d'enquête, malgré le fait qu'il n'existe pas de différences réelles entre les populations. Il convient aussi de noter que l'intervieweur peut passer du français à l'anglais au cours de l'interview. La variable « Langue » indique la principale langue dans laquelle l'interview s'est déroulée, mais il se peut que certaines questions aient été posées dans l'autre langue.

Par ailleurs, certaines interviews sont menées dans d'autres langues que le français ou l'anglais, l'intervieweur traduisant les questions dans la langue que le répondant préfère.

12.13 Renseignements contradictoires

Il arrive parfois qu'un répondant fournisse des renseignements contradictoires. Dans certains cas, la contradiction peut être résolue grâce à des règles de contrôle déterministes. Par exemple, si un répondant déclare ayant atteint un plus haut niveau d'études moins élevé que ce qu'il a déclaré au cycle précédent, le plus haut niveau d'études est établi à la valeur déjà déclarée.

Dans d'autres cas, la contradiction ne peut être résolue facilement. Par exemple, le répondant peut répondre « Oui » à la question « Votre enfant est-il capable de dire huit mots ou davantage en plus de « maman » et « papa » ? » dans le module Étapes du développement, et répondre « Non » à la question « Est-ce qu'il ou elle a prononcé deux mots reconnaissables à part « maman » et « papa » ? » dans le module Développement moteur et social. De toute évidence, ces réponses sont contradictoires, mais on ne change rien dans ces cas-là.

Il est frustrant que les données recueillies ne soient pas cohérentes. Cependant, étant donné qu'on ne peut pas en toute confiance corriger les incohérences, il en subsiste dans les fichiers définitifs de l'enquête.

Il se peut également que les données recueillies dans le cycle en cours soient contradictoires par rapport à celles recueillies au cours des cycles précédents. Dans le cas de certains enfants, il arrive par exemple qu'on ait indiqué au cycle 1 que les parents étaient séparés et que la PMR affirme, à un cycle suivant, que les parents ont continuellement vécu ensemble depuis la naissance de l'enfant. Il arrive aussi, au cours de l'enquête, que plus d'une personne se disent la mère ou le père biologique de l'enfant. Le cas échéant, on accepte la réponse fournie dans le cycle en cours.

Par ailleurs, les résultats de l'ELNEJ peuvent aussi être contradictoires par rapport à ceux provenant d'autres sources. Il se peut que certaines définitions ou concepts ne soient pas entièrement compatibles ou qu'on ait eu recours à des pratiques différentes dans la collecte des données. Il se peut également qu'une erreur se soit glissée à l'étape du traitement du fichier de microdonnées.

12.14 Qualité des données de l'indice de masse corporelle

12.14.1 Indice de masse corporelle

L'indice de masse corporelle (IMC) est une échelle normalisée servant à mesurer la masse corporelle. Le score de l'IMC se calcule en divisant le poids par le carré de la taille :

$$\text{IMC} = \frac{\text{poids en kilogrammes}}{(\text{taille en mètres}) \times (\text{taille en mètres})}$$

Pour les enfants de moins de 2 ans, l'utilisation de l'IMC aux fins d'analyse n'est pas recommandée. Par conséquent, l'IMC n'est pas calculé pour ce groupe d'âge. Pour les enfants de 2 ans et plus, les variables suivantes de taille et de poids sont utilisées pour calculer l'IMC à partir des données de l'ELNEJ.

- HHLCQ03B (taille) et HHLCQ04A (poids) (données déclarées par la PMR pour les enfants de 2 à 7 ans) : score de l'IMC du répondant (HHLCE01);
- HHTCbQ01 (taille) et HHTCbQ02 (poids) (données déclarées par les répondants de 14 à 17 ans dans les questionnaires qu'ils remplissent eux-mêmes) : score de l'IMC du répondant (HHLCE01);
- HHTYeD01 (taille) et HHTYeD02 (poids) (données déclarées par les répondants de 16 à 25 ans dans les questionnaires qu'ils remplissent eux-mêmes de la composante des jeunes) : score de l'IMC du répondant (HHTYe03). Il n'est pas recommandé d'effectuer des analyses de l'IMC pour les femmes de ce groupe car l'ELNEJ ne précise pas si elles sont enceintes ou si elles allaitent, facteurs qui peuvent influencer grandement sur l'IMC.

En calculant le score de l'IMC, on peut ensuite le comparer aux autres pour voir dans quel percentile il se situe. Différents seuils ou intervalles percentiles ont été proposés pour permettre de déterminer si le score de l'IMC d'une personne est classé dans la catégorie « poids insuffisant », « poids normal », « à risque d'excédent de poids », « excédent de poids » ou « obésité ». Les *Centers for Disease Control* (CDC) des États-Unis ont proposé des seuils pour les enfants, les jeunes et les adultes. De même, Tim Cole et ses collaborateurs ont proposé des seuils internationaux pour les enfants et les jeunes en se servant d'une méthode différente.

12.14.2 Indice de masse corporelle — Centers for Disease Control

Les CDC ont proposé un ensemble d'intervalles percentiles pour classer les scores d'IMC dans les catégories « poids insuffisant », « poids normal », « à risque d'excédent de poids » et « excédent de poids ». Ces intervalles sont établis selon l'âge et le sexe et sont fondés sur des données américaines de taille et de poids. Les seuils des CDC sont fondés sur l'âge de la personne réparti en intervalles d'un mois. Par conséquent, lors du traitement des données de l'ELNEJ, on a utilisé la variable de l'âge en mois (HMMCdQ1B) pour calculer les seuils. Les intervalles percentiles proposés par les CDC pourraient être utilisés pour les personnes de 0 à 19 ans. Le nom de diffusion de cette variable est HHLCeD03 pour les jeunes de 17 ans et moins et HHTYeD05 pour les personnes de 18 ans et plus.

On peut obtenir plus de renseignements sur les seuils de l'IMC des CDC pour les enfants et les jeunes à l'adresse suivante : <http://www.cdc.gov/growthcharts>.

12.14.3 Indice de masse corporelle : seuils internationaux

Des seuils internationaux d'IMC pour les jeunes de 2 à 18 ans ont été proposés par Tim Cole, Mary Bellizzi, Katherine Flegal et William Dietz dans le *British Medical Journal* (« Establishing a standard definition for child overweight and obesity worldwide: international survey », *BMJ* 2000, vol. 320, p. 1 à 6). Selon ces seuils, les scores de l'IMC sont classés en trois catégories : « poids normal », « excédent de poids » et « obésité ». Comme les seuils des CDC, ces catégories sont établies selon l'âge et le sexe. Mais contrairement aux seuils des CDC, les seuils internationaux sont fondés sur l'âge réparti en intervalles de six mois. Ils ont été établis d'après des études effectuées sur des ensembles de données représentatives de six pays sur les IMC pendant l'enfance (États-Unis, Brésil, Grande-Bretagne, Hong-Kong, Pays-Bas et Singapour). Le nom de diffusion de cette variable est HHLCeD02 pour les jeunes de 17 ans et moins et HHTYeD04 pour les jeunes de 18 ans et plus. Après l'âge de 17 ans, les seuils pour les adultes de l'Organisation mondiale de la santé sont utilisés.

On peut obtenir plus de renseignements sur les seuils de l'IMC proposés par Tim Cole et ses collaborateurs à l'adresse suivante : <http://www.bmj.com>.

12.14.4 Indice de masse corporelle — Qualité des données

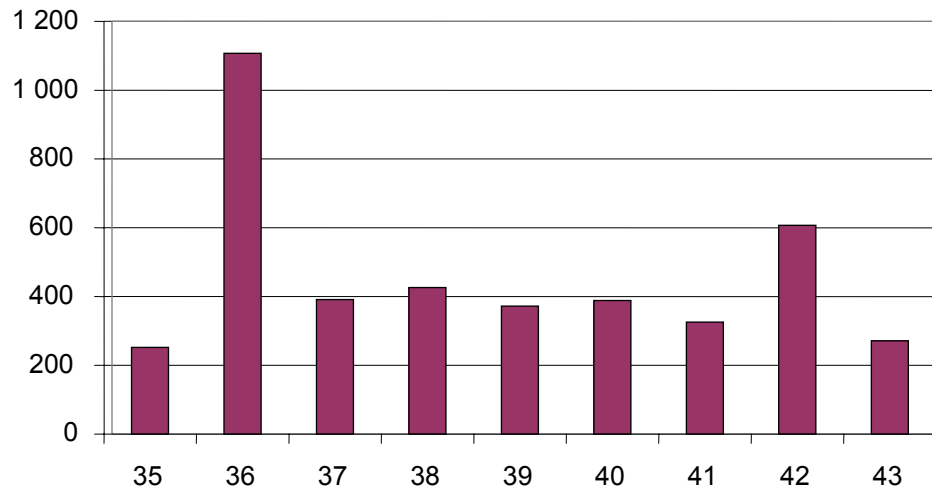
Plusieurs facteurs influent sur la qualité des scores de l'IMC. D'abord, le taux de non-réponse est plus élevé pour les variables de l'IMC, car il faut que les variables de taille et de poids contiennent toutes deux des réponses valides pour pouvoir calculer un score. Il manque notamment une grande quantité de données sur la taille des enfants. On s'est penché sur le problème et des mesures seront prises à l'avenir pour améliorer les taux de réponse ainsi que l'exactitude des mesures.

Ensuite, les données recueillies sur la taille et le poids ne sont que des estimations fournies par le parent ou le jeune et non des mesures cliniques exactes. Les résultats obtenus avec cette méthode de collecte, particulièrement auprès des parents, sont des valeurs moins précises de taille et de poids et, par conséquent, des valeurs moins précises de l'IMC. De façon générale, le répondant arrondit le chiffre de la taille ou du poids qu'il déclare, ce qui donne des valeurs différentes d'IMC que celles qu'on aurait obtenues à partir de mesures cliniques. Par exemple, la PMR peut déclarer que l'enfant mesure 5 pieds plutôt que 5 pieds 2 pouces ou 4 pieds 10 pouces ou encore, que l'enfant pèse 110 livres plutôt que 113 ou 108 livres. Ces petites erreurs d'estimation peuvent se traduire en un degré d'erreur beaucoup plus élevé dans l'IMC, ce qui pourrait donner lieu à un changement de catégorie, par exemple de « Excédent de poids » à

« Poids normal » ou à « Obésité », selon la façon dont la taille et le poids ont été arrondis.

Le graphique ci-dessous montre que l'on a déclaré que beaucoup plus d'enfants mesureraient exactement 3 pieds que près de 3 pieds. Bien que les résultats ci-après s'appliquent au cycle 6, des comportements semblables ont été remarqués pour d'autres variables et à d'autres cycles, y compris au cycle 8.

Cycle 6 — Chiffres non pondérés de certaines valeurs de taille déclarées en pouces



Source : Statistique Canada, Enquête longitudinale nationale sur les enfants et les jeunes.

12.15 Biais d'accoutumance

Il arrive que des participants à une enquête longitudinale modifient leur comportement du fait qu'ils sont conscients de participer à une telle enquête. En outre, le fait de répondre au questionnaire est susceptible d'influer sur le comportement des répondants. Par exemple, après avoir été interrogé sur la fréquence des lectures faites à l'enfant au cycle 1, le parent peut décider de s'adonner plus souvent à cette activité. Ce parent n'est plus représentatif des autres parents canadiens qui n'ont pas participé à l'enquête, puisque sa participation à celle-ci a modifié son comportement.

Il se peut également que les répondants répondent de façon à réduire la durée de l'interview. Ils sont parfois conscients du fait qu'en répondant par l'affirmative à certaines questions, ils déclenchent un ensemble de questions complémentaires détaillées. Cela peut les inciter à ne pas répondre sincèrement à ces questions.

Même si ces biais sont probablement négligeables, il est impossible de les mesurer de façon précise.

12.16 Personne la mieux renseignée

À chaque cycle, un adulte au sein du ménage est désigné comme étant la personne la mieux renseignée (PMR) au sujet de l'enfant. La PMR répond aux questions de la composante de l'enfant, fournissant des renseignements à propos de la santé de l'enfant, de son niveau de scolarité, de son comportement, entre autres. Les caractéristiques de l'enfant sont mesurées indirectement en fonction de ce que déclare la PMR. La PMR n'est pas toujours la même d'un cycle à l'autre. Pour un enfant particulier, il se peut que la mère ait été la PMR au cycle 1, que le père l'ait été aux cycles 2

et 3, et que la mère soit redevenue la PMR aux cycles 4, 5, 6, 7 et 8. Les changements observés en ce qui touche les caractéristiques de l'enfant peuvent être attribuables, en partie, au fait que différentes personnes ont répondu à ces questions. Le milieu dans lequel se trouve l'enfant n'est pas vraiment différent, mais la personne qui répond aux questions n'est pas la même et a, naturellement, un point de vue différent. La variable HDMCD06 donne le lien de la PMR à l'enfant. Les variables PPERSC4, PPERSC5, PPERSC6, PPERSC7 et PPERSRUK identifient la PMR aux cycles 4, 5, 6, 7 et 8, respectivement.

12.17 Couverture des enfants canadiens par l'échantillon de l'ELNEJ

La population observée est la population représentée par l'échantillon et à propos de laquelle des inférences peuvent être faites lorsque les poids de l'enquête sont utilisés à l'analyse. Par exemple, la population observée pour la cohorte originale est composée d'enfants âgés de 0 à 11 ans au 31 décembre 1994 qui vivaient dans une province ou l'autre au moment de la collecte du cycle 1 (1994-1995).

Il faut se rappeler que tous les enfants de l'ELNEJ reçoivent un poids longitudinal (sauf les échantillons de remise à niveau), mais que seulement les enfants du fichier DPE reçoivent un poids sur le plan transversal. Les diverses populations observées couvertes par les poids sur les plans longitudinal et transversal de l'ELNEJ sont décrites au chapitre 5.0.

Les utilisateurs des données de l'ELNEJ doivent savoir qu'il y a certaines exclusions et que la couverture de certains enfants canadiens, comme les immigrants, n'est pas égale, puisque ces facteurs pourraient avoir un effet sur leurs analyses.

12.17.1 Exclusions de l'EPA

Des enfants sont exclus de l'EPA et par conséquent des échantillons de l'ELNEJ qui ont été tirés de l'EPA. L'EPA porte seulement sur la population civile excluant les pensionnaires d'établissements institutionnels dans les 10 provinces du Canada. Elle exclut le Yukon, le Nunavut et les Territoires du Nord-Ouest, les personnes qui vivent dans les réserves indiennes, les membres à plein temps des Forces armées canadiennes et les détenus dans les prisons.

12.17.2 Couverture des immigrants récents

Comme il n'y a jamais eu de remise à niveau de la cohorte originale pour les immigrants qui sont arrivés après 1994-1995, il ne faut pas l'utiliser pour les analyses sur le plan transversal. Le cycle 4 a été le dernier cycle pour lequel des poids sur le plan transversal ont été produits pour la cohorte originale. Au cycle 5, nous avons pensé que l'absence d'immigrants récents était si importante qu'il ne faudrait plus utiliser la cohorte originale pour faire des inférences au sujet des populations sur le plan transversal après le cycle 4.

Dans le cas des cohortes du DPE, il a fallu remettre à niveau l'échantillon dans certaines provinces, pour certains âges, de sorte que certains nouveaux immigrants sont présents, mais la base de sondage ne garantit pas expressément un échantillon suffisant de nouveaux immigrants. Par conséquent, il faut être prudent lorsqu'on utilise l'échantillon du fichier DPE pour faire des inférences au sujet des immigrants.

Plus expressément, avant le cycle 7, les cohortes du DPE n'ont été suivies que pour trois cycles (jusqu'à l'âge de 4 et 5 ans) de sorte que, même en l'absence de remises à niveau, nous avons pensé qu'elles pouvaient être utilisées pour l'analyse sur le plan transversal. Aux cycles 6, 7 et 8, on a fait une remise à niveau pour les 2 à 5 ans dans

l'ELNEJ dans certaines provinces (mais non l'Ontario ou le Québec) de sorte que certains immigrants récents ont été pris en compte avec ces remises à niveau. Toutefois, la taille des échantillons de remise à niveau a été généralement modeste.

Les enfants du DPE du cycle 8 comprennent donc certains nouveaux immigrants en raison de la remise à niveau des 2 à 5 ans du cycle 8, de la remise à niveau des enfants de 4 à 7 ans qui sont de retour du cycle 7, et de la remise à niveau des enfants de 6 à 7 ans qui sont de retour du cycle 6.

Par conséquent pour les enfants du DPE présents au cycle 8, bien que certains enfants immigrants récents soient présents, la base de sondage ne garantit pas que l'échantillon est représentatif de tous les enfants immigrants âgés de 0 à 7 ans.

12.17.3 Couverture selon le mois de naissance

Étant donné la façon dont l'ELNEJ échantillonne les bébés à partir de l'EPA, les bébés nés à la fin de l'année civile ont généralement une probabilité moindre de sélection par rapport à ceux qui sont nés au début de l'année. Cette distribution inégale de l'échantillon selon le mois de naissance est devenue plus prononcée aux cycles 6 et 7 et des rajustements de poids ont été effectués : au cycle 6, un rajustement uniforme a été ajouté aux poids de l'enquête pour les 0 et 1 an, et au cycle 7, le rajustement du poids selon le mois de naissance pour les 0 et 1 an a été précisé. Au cycle 8, la méthodologie du cycle 7 a été utilisée.

12.17.4 Couverture selon l'ordre de naissance dans la cohorte originale

Au cycle 2, l'échantillon de la cohorte originale a été réduit pour des raisons budgétaires. Certains ménages ont été éliminés et, dans d'autres ménages, le nombre d'enfants retenus a été réduit à un maximum de deux enfants comparativement au maximum de quatre au cycle 1. Pour la grande majorité des ménages, le choix des enfants retenus a été aléatoire et les enfants retenus correspondent exactement à la population observée.

Pour une fraction de l'échantillon, c'est-à-dire les ménages dont au moins un enfant de 0 ou 1 an a été choisi dans l'EPA après le remaniement de 1994, plus tous les ménages au Nouveau-Brunswick, le choix des enfants à retenir pour respecter le nouveau maximum de deux enfants par ménage n'a pas été aléatoire. Plutôt, pour cette fraction de l'échantillon, les deux plus jeunes enfants ont été retenus.

Il y a donc une légère distorsion selon l'ordre de naissance. Par exemple, dans la cohorte originale, nous avons une proportion légèrement élevée d'enfants de 0 et 1 an (âge au cycle 1) avec deux frères ou sœurs plus âgés comparativement à l'ensemble de la population. Il y a une proportion légèrement inférieure correspondante d'enfants âgés de 2 à 11 ans (âge au cycle 1) ayant au moins deux frères ou sœurs cadets.

12.18 Validation des données

Pour les cycles 7 et 8, une étape supplémentaire de validation a été utilisée dans le processus d'enquête en vue de vérifier la cohérence des estimations sur plusieurs cycles. L'outil, appelé outil de validation préalable à la diffusion, identifie automatiquement les écarts importants dans les estimations pour les cycles consécutifs pour toutes les variables communes à ces cycles. Cet outil s'est révélé très utile pour repérer et corriger les erreurs avant la diffusion des données et devrait améliorer la qualité globale des données des fichiers diffusés.

12.19 Conclusion

Divers types d'erreurs influent sur la qualité des données. On s'efforce de réduire les risques d'erreurs à toutes les étapes (formation des intervieweurs, contrôle de la collecte des données, traitement et pondération, notamment).

On encourage les utilisateurs des données à déterminer en quoi les erreurs d'échantillonnage et les erreurs non dues à l'échantillonnage peuvent influencer sur les variables qu'ils tentent d'analyser.

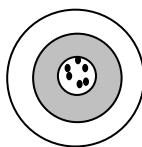
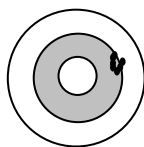
13.0 Estimation de la variance

L'Enquête longitudinale nationale sur les enfants et les jeunes (ELNEJ) est une enquête probabiliste pour laquelle des échantillons ont été sélectionnés afin de représenter diverses populations cibles. Une quantité d'intérêt à propos d'une population cible donnée s'appelle un **paramètre**. La taille moyenne des enfants d'un même âge est un exemple de paramètre. La valeur d'un paramètre est inconnue et nous menons une enquête telle que l'ELNEJ afin d'obtenir de l'information utile auprès d'un échantillon de tous les individus appartenant à la population cible d'intérêt. L'information pertinente contenue dans cet échantillon observé au sujet du paramètre est extraite des données à l'aide d'un outil mathématique qu'on appelle un **estimateur**. La valeur qui représente une approximation raisonnable du paramètre que l'on puisse faire à partir de l'information observée s'appelle une **estimation**, qui est tout simplement la valeur de sortie de l'estimateur après lui avoir soumis l'échantillon observé. Soumettre différents échantillons à l'estimateur donne différentes approximations numériques, c'est-à-dire les estimations qui sont faites au sujet du paramètre. La mesure dans laquelle ces estimations diffèrent globalement de la valeur du paramètre est l'erreur d'échantillonnage. L'un des éléments clés de l'échantillonnage consiste à avoir la capacité à mesurer de façon mathématique, l'ampleur de l'erreur d'échantillonnage. Par définition, un recensement ne comporte pas d'erreur d'échantillonnage puisque le seul échantillon possible est la population entière (et par conséquent, la seule estimation raisonnable que nous obtenons correspond exactement à la valeur du paramètre).

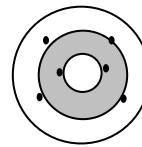
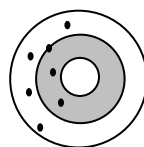
Même si, dans un contexte donné, on ne parvient pas à mesurer l'erreur d'échantillonnage, il serait possible de décrire certaines de ses caractéristiques. Par exemple, plus l'échantillon est grand, plus, en général, l'erreur d'échantillonnage est petite; cela vient du fait qu'un plus grand échantillon renferme de l'information utile à propos d'une plus grande partie de la population. Parmi les autres facteurs qui influent sur la magnitude de l'erreur d'échantillonnage figurent la taille de la population, la façon de sélectionner l'échantillon (précisée dans le **plan d'échantillonnage**) et la variabilité dans la population cible des caractéristiques sur lesquelles reposeront notre estimation.

Il y a deux composantes à l'erreur d'échantillonnage : le biais et la variance d'échantillonnage. Un estimateur ne présente pas de biais d'échantillonnage lorsque la moyenne de toutes les estimations qu'il produit à partir de tous les échantillons possibles, correspond en gros à la valeur du paramètre. Ainsi, les estimations prises individuellement peuvent toutes être différentes de la valeur du paramètre mais lui correspondre en moyenne. On dit alors de l'estimateur qu'il est non biaisé (autrement, on le dit biaisé). L'autre composante de l'erreur d'échantillonnage est la variance d'échantillonnage, qui indique le degré de différence entre les estimations.

L'analogie bien connue que voici aide à comprendre ces concepts. On demande à un lanceur de dards (l'estimateur) de lancer une série de dards vers une cible; chacun des lancers sur la cible correspond à une estimation. On ne s'attend pas à ce que tous les lancers atteignent le centre de la cible. Chacun des lancers contribue à accroître l'erreur d'échantillonnage. Il y a quatre scénarios possibles pour l'ensemble des lancers sur la cible selon la magnitude des deux composantes de l'erreur d'échantillonnage, soit le biais et la variance :



Biais et faible variance Absence de biais et faible variance



Biais et forte variance Absence de biais et forte variance

Dans la pratique, il est généralement impossible d'obtenir un estimateur qui donne de bons résultats pour les deux composantes, c'est-à-dire un estimateur à la fois à faible biais et à faible variance. Habituellement, un estimateur à faible variance sera fortement biaisé, tandis qu'un estimateur peu biaisé aura une forte variance. Les statisticiens d'enquête ont généralement recours à des estimateurs qu'on sait n'être pas ou que peu biaisés; ils préfèrent réduire au minimum la composante de l'erreur d'échantillonnage résultant du biais et faire ensuite ce qu'ils peuvent au sujet de la variance. C'est aussi ce que nous préférons et nous ne traiterons par conséquent des erreurs d'échantillonnage que dans l'optique de la variance d'échantillonnage.

Afin d'évaluer la variance d'échantillonnage d'un estimateur, nous avons besoin d'estimations provenant de tous les échantillons possibles. En pratique, cependant, nous ne disposons que d'un seul échantillon, celui pour lequel la collecte a été réalisée. Bien que nous ne puissions pas mesurer la variance d'échantillonnage d'un estimateur, nous pouvons habituellement l'estimer à partir de l'échantillon dont nous disposons.

L'estimation de la variance nous permet de prédire avec exactitude la variabilité des résultats provenant de tous les échantillons en n'utilisant qu'un seul échantillon. Si la situation générale est plutôt compliquée, l'estimation de la variance repose fondamentalement sur l'observation suivante. On peut démontrer que dans le cas de l'emploi d'estimateurs simples (comme celui de la moyenne) avec un mécanisme très simple de sélection de l'échantillon, la variance d'échantillonnage est fonction directe de la dispersion dans la population. Alors, dans un tel cas simple, la variance d'échantillonnage nous est inconnue parce qu'il s'agit d'une fonction (connue) d'une quantité inconnue, en l'occurrence la dispersion dans la population. Si nous connaissions la variance d'échantillonnage, nous connaîtrions la dispersion dans la population. Est-il possible d'établir une bonne approximation de cette dispersion dans la population à partir de l'échantillon des valeurs? Encore une fois, dans ce cas des plus simples, la dispersion des valeurs de l'échantillon donne une mesure directe de la dispersion dans la population.

En effet, dans ce cadre particulier, l'on obtient habituellement un échantillon bien équilibré puisqu'il contient à peu près les mêmes proportions de petites, de moyennes et de grandes valeurs qu'au sein de la population même, ce qui fait de la dispersion observée un miroir de la dispersion dans la population. Dans les cadres plus complexes, il n'existe pas de telle relation directe entre la dispersion dans la population et la dispersion dans l'échantillon, qui pourrait, à son tour, être utilisée pour estimer la variance d'échantillonnage. Il y a néanmoins des méthodes astucieuses comme le bootstrap, dont nous parlerons plus loin, qui permettent de déterminer la variance d'échantillonnage à partir d'un seul échantillon observé, comme nous l'avons fait ci-dessus.

Dans le présent chapitre, nous expliquerons pourquoi il est important de calculer la variance d'échantillonnage et nous présenterons différents outils pour le faire dans le cadre de l'ELNEJ.

13.1 Termes liés à l'erreur et à la variance d'échantillonnage

On se méprend parfois sur le sens des termes « variance de la population », « variance d'échantillonnage », « écart type » et « erreur type ». L'objectif de la présente section est de le clarifier.

Malheureusement, le mot « variance » a deux sens très distincts selon qu'il s'agisse du terme « variance de la population » ou du terme « variance d'échantillonnage ». La variabilité observée dans les valeurs d'une caractéristique de la population cible est souvent appelée (malheureusement) « variance dans la population » (de la caractéristique). Par exemple, la taille en centimètres de la population canadienne de tous les garçons de 10 ans affiche une certaine variabilité.

Mais le terme « variance » ne devrait être utilisé qu'en présence d'un processus aléatoire sous-jacent (comme le processus aléatoire servant à choisir des échantillons dans le cadre d'enquêtes). Puisque les valeurs d'une caractéristique de la population sont fixes (n'étant donc pas le résultat d'un processus aléatoire), nous proposons plutôt d'appeler leur variabilité inhérente « dispersion dans la population » (de cette caractéristique) plutôt que « variance dans la population ». Dans

l'exemple ci-dessus, nous dirions qu'il y a une certaine dispersion dans les tailles mesurées en centimètres de tous les garçons de 10 ans au Canada.

La définition mathématique suivante de la dispersion dans la population est souvent utilisée :

$$\sigma^2 = \frac{\sum_{i=1}^N (y_i - \bar{Y})^2}{N - 1}$$

où y_i est la valeur de la caractéristique Y pour la personne i ;

\bar{Y} est la moyenne des y_i dans la population; et

N est la taille de la population.

Présentée ainsi, la dispersion dans la population n'est qu'un des nombreux paramètres de la population, comme l'est par exemple \bar{Y} , la valeur moyenne de la population pour Y . Le lecteur notera que le symbole employé pour désigner la dispersion dans la population est $\sigma^2 = \sigma \times \sigma$ et non simplement σ . C'est là une façon de se rappeler que la mesure de dispersion ne s'exprime pas dans les mêmes unités de mesure que la variable elle-même, mais plutôt selon son carré. Par conséquent, pour obtenir une mesure de dispersion qui soit exprimée dans les mêmes unités que la variable elle-même, il semble raisonnable de prendre la racine carrée de la dispersion dans la population. Et c'est précisément ce que nous faisons; cela donne ce qui s'appelle l'**écart type**.

Nous avons déjà parlé de la **variance d'échantillonnage**, qui est la variabilité qu'on observerait dans les estimations d'un même paramètre si l'on traitait tous les échantillons possibles. Puisque la sélection d'échantillons est un processus aléatoire, il est logique ici de parler de « variance » plutôt que de dispersion dans la population, comme nous le faisons ci-dessus. Tout comme la mesure de dispersion ne s'exprime pas dans les mêmes unités que la variable dont on estime la variabilité, mais plutôt selon son carré, la dispersion associée à d'autres valeurs s'exprime selon son carré. Par exemple, si l'estimation d'un revenu personnel s'exprime en dollars (\$), alors la variance d'échantillonnage s'exprime aussi en \$². Une fois de plus, il est logique de prendre la racine carrée de la valeur afin de rétablir la comparabilité s'exprimant dans des unités semblables à celles de l'estimation; la racine carrée de la valeur s'appelle l'**erreur type**.

13.2 Coefficient de variation

Une mesure de la variance d'échantillonnage nous donne une indication précieuse de la fiabilité d'une estimation. Comme nous l'avons vu, l'erreur type est un meilleur outil que la variance d'échantillonnage parce qu'elle s'exprime dans les mêmes unités de mesure que l'estimation elle-même. Mais l'erreur type, tout comme la variance, présente un inconvénient majeur, car les erreurs types rattachées à des estimations différentes ne sont pas comparables dans des situations semblables. Par exemple, l'estimation obtenue pour la province A est-elle plus ou moins fiable que l'estimation correspondante obtenue pour la province B?

Afin d'illustrer cet inconvénient, supposons qu'on tire un échantillon pour recueillir de l'information afin d'estimer le total d'une caractéristique Y . Avec la même information, nous pouvons également estimer la **moyenne** de Y en divisant le total estimé par la taille (présumée connue) de la population N . Néanmoins, les deux estimations sont obtenues à partir des mêmes données d'échantillon, et ce, de la même façon, par l'entremise du total estimé. Par conséquent, si les erreurs types tant du total que de la moyenne étaient des instruments comparables de mesure de la précision relative des estimations, nous nous attendrions à ce qu'elles aient la même valeur. Leurs valeurs sont toutefois différentes, car l'erreur type du total sera N fois plus grande que celle de la moyenne; les estimations de la moyenne et du total tirées du même échantillon observé diffèrent toutes par le même facteur, soit N .

Le **coefficient de variation** (CV) est une façon d'obtenir une mesure relative de l'erreur d'échantillonnage qui fournit la même valeur dans des situations d'échantillonnage comparables. Le CV s'obtient en divisant l'erreur type de l'estimation $\hat{\theta}$ par l'estimation elle-même :

$$CV(\hat{\theta}) = \frac{e.t.(\hat{\theta})}{\hat{\theta}}$$

Maintenant, contrairement à la variance d'échantillonnage associée à une estimation, le CV permet à l'analyste de comparer des estimations de magnitude différente ou mesurées à l'aide d'unités différentes sur une échelle commune pour des variables du même type que le revenu⁷. L'exemple, ci-dessus, de la comparaison des estimations de la moyenne et des totaux nous montre que dans un tel cas, les CV seraient égaux. En effet, bien que, comme nous l'avons dit, l'erreur type du total est N fois plus grande que celle de la moyenne, le dénominateur correspondant pour le CV du total est aussi N fois plus grand que celui utilisé pour calculer le CV de la moyenne. Par conséquent, le même facteur N apparaît à la fois dans le numérateur et le dénominateur et s'annule donc de lui-même.

Bien que le CV soit utile pour mesurer la fiabilité d'estimations telles que les totaux, il présente des inconvénients dont l'utilisateur devrait être conscient lorsqu'il l'utilise pour examiner la qualité des proportions. Quelques-uns de ces problèmes potentiels sont décrits ci-dessous.

Problème 1 : Coefficients de variation pour de très petites ou de très grandes proportions

Comme l'erreur type pour une proportion p est la même que pour l'erreur type pour une proportion $(1-p)$, les CV pour p et $(1-p)$ peuvent être très différents parce que les dénominateurs sont p et $(1-p)$ respectivement. Il est possible d'imaginer un scénario où p est très petit, ce qui lui donne un très grand CV pour p , mais où le CV de $(1-p)$ est excellent. Utilisons l'exemple suivant pour en faire la démonstration. Supposons que nous ayons calculé que l'erreur type des estimations p et $(1-p)$ est 0,0475, ce qui donne une valeur p de 0,95. Le CV pour la valeur 0,95 de p serait :

$$0,0475 / 0,95 = 0,05 \text{ ou } 5 \%, \text{ ce qui est un très bon CV.}$$

D'autre part, le CV pour la proportion $(1-p)$ est :

$$0,0475 / 0,05 = 0,95 \text{ ou } 95 \%, \text{ ce qui est un très mauvais CV.}$$

Problème 2 : Appliquer les règles d'évaluation de l'erreur d'échantillonnage dans le cas des proportions

Des règles ont été établies pour aider les utilisateurs à évaluer la validité d'une estimation à partir de la magnitude de l'erreur d'échantillonnage telle qu'elle est mesurée par le CV. D'ordinaire, ces règles stipulent qu'une estimation qui s'accompagne d'un CV de moins de 16,5 % est de **bonne qualité** (les conclusions tirées de ces résultats sont fiables). Un CV d'entre 16,5 % et 33 % représente une estimation de **qualité acceptable** (les conclusions tirées de ces résultats doivent être considérées avec prudence), alors qu'un CV dépassant 33 % représente une estimation de **piètre qualité** (les conclusions tirées de ces résultats sont purement exploratoires et ne sont pas fiables). Ces règles peuvent aider l'utilisateur de données à comprendre les résultats d'une enquête, mais elles ont leurs limites. Consulter un statisticien d'enquête est sans nul doute le meilleur moyen d'interpréter les résultats puisque

7. Nous entendons par là une variable d'intérêt qui est continue et positive, par opposition notamment à une variable dichotomique.

d'autres facteurs spécifiques à l'analyse que l'erreur d'échantillonnage peuvent influencer sur la validité des conclusions. Dans le cas de proportions, ces règles doivent être employées avec davantage de discernement, comme l'illustrent les exemples qui suivent.

Exemple 1 : Une proportion estimée de 0,50 avec un intervalle de confiance de 99 % de 0,10 % à 0,90 % se situe dans la catégorie **marginale** compte tenu des lignes directrices relatives à la qualité publiées précédemment (le CV est de 31 %). L'intervalle de confiance est tellement grand que l'estimation nous renseigne peu.

Exemple 2 : Supposons que nous ayons un échantillon de 27 000 individus, dont 44 possèdent une caractéristique que nous sommes intéressés à étudier. Avec les poids de sondage, nous calculons que 0,16 % de la population possède cette caractéristique (le CV est de 34 %). Un CV de 34 % est considéré **inacceptable** selon les lignes directrices relatives à la qualité publiées précédemment. Bien que 0,16 % soit une très petite proportion, si nous construisons son intervalle de confiance de 99 %, nous obtenons (0,02 %, 0,30 %). Cela signifie que nous pouvons affirmer avec beaucoup de confiance que le vrai taux est, disons, plus petit que 0,5 %. Bien entendu, les utilisateurs de données doivent se montrer prudents, car le taux véritable pourrait être de 0,05 %, de 0,25 % ou de 0,30 %. Selon le but de la recherche, peut-être serait-il utile d'affirmer que l'estimation est inférieure à 0,5 %. Il est donc inapproprié d'appliquer aveuglément les lignes directrices relatives à la qualité et de rejeter une estimation à cause de son CV élevé.

Par conséquent, les utilisateurs qui voudraient utiliser le CV comme mesure de l'erreur d'échantillonnage avec des proportions sont fortement encouragés à calculer les CV tant pour la proportion p que pour la proportion $(1-p)$. Les CV associés à des proportions, particulièrement à des proportions extrêmes, peuvent être trompeuses, comme l'illustre les exemples précédents. Nous ne recommandons pas vraiment de comparer les CV de différentes proportions, mais plutôt de construire des intervalles de confiance et de s'y fier afin de tirer des conclusions adéquates au sujet des proportions.

Des recherches sont en cours pour trouver des solutions de rechange au CV pour les proportions extrêmes (près de 0 et près de 1). Cependant, aucune solution qui répondra à tous les besoins n'a encore été trouvée. Par conséquent, les utilisateurs doivent faire preuve de prudence s'ils veulent utiliser les CV pour déterminer des proportions.

13.3 Importance de déclarer la variance d'échantillonnage

Comme nous l'avons vu dans les sections précédentes, une estimation sera, selon toute vraisemblance, différente de la vraie valeur, à supposer que nous connaissions celle-ci; cette différence numérique entre l'estimation et le paramètre constitue l'erreur. Il faut donc supposer que chaque estimation s'accompagne d'une erreur. **Par conséquent, on doit donner une indication de la magnitude ou de l'étendue de cette erreur dans la conclusion à ceux qui consulteront les estimations.** En effet, il peut être très trompeur de faire état d'estimations d'enquête sans les mesures correspondantes de l'erreur qu'elles comportent. L'une des deux composantes de cette erreur est l'erreur d'échantillonnage et l'autre, l'erreur non due à l'échantillonnage. Cette dernière comprend des choses telles que les erreurs de réponse et les erreurs de traitement. D'une part, la magnitude de l'erreur non due à l'échantillonnage est très difficile, voire impossible, à quantifier en pratique. (Lorsque l'impact des erreurs ne peut pas être mesuré, comme c'est le cas pour la plupart des erreurs non dues à l'échantillonnage, elles doivent alors à tout le moins être clairement rapportées et décrites.) D'autre part, l'erreur d'échantillonnage **peut** être quantifiée si nous parvenons à suivre mathématiquement l'effet de l'aléatoire sur le rendement des estimations. Nous mesurons l'erreur d'échantillonnage d'une estimation en estimant la variance d'échantillonnage du

processus qui l'a engendrée. Il faut donc calculer la variance d'échantillonnage et la fournir aux utilisateurs pour leur décrire la qualité des estimations fournies.

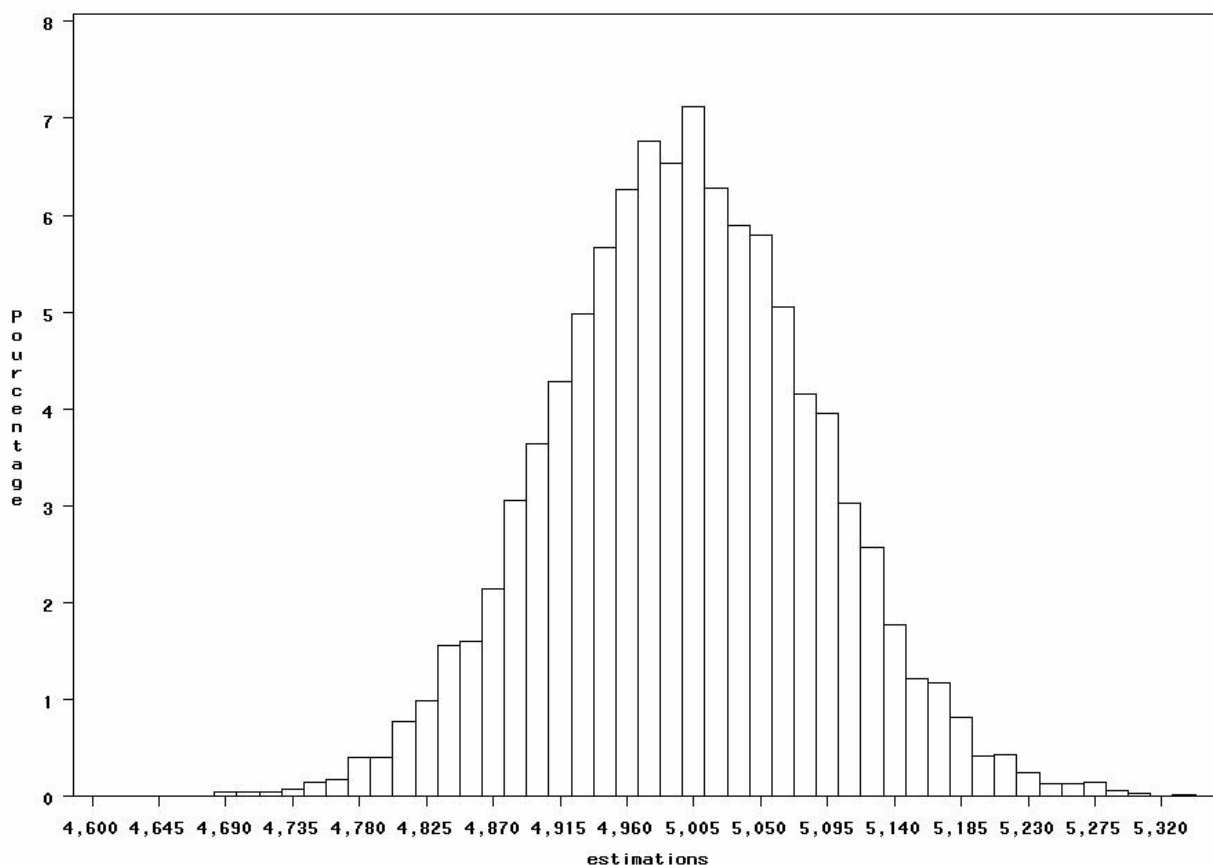
13.4 Calcul de la variance d'échantillonnage

Il serait difficile, pour ne pas dire impossible, d'obtenir une formule exacte qui permettrait de calculer la variance d'échantillonnage pour l'ELNEJ en raison de son plan de sondage complexe, des ajustements pour la non-réponse, du traitement des unités hors champ et de la stratification a posteriori. En fait, cela ne pourrait se faire qu'en ayant recours à des hypothèses si fortes que nous obtiendrions un cadre trop simpliste pour être d'une quelconque utilité dans la pratique. Une très bonne façon de produire une approximation de la variance d'échantillonnage est de recourir à la méthode du bootstrap⁸.

La méthode bootstrap consiste à faire une sélection aléatoire de sous-échantillons à partir d'un échantillon principal, et ce, de telle façon que chaque sous-échantillon (ou réplique) suive le même plan de sondage que l'échantillon principal. Les poids définitifs des unités dans chacun des sous-échantillons sont calculés en suivant les mêmes étapes de pondération que celles utilisées pour l'échantillon principal (voir le chapitre 11.0). Ces poids bootstrap sont utilisés pour calculer une estimation de la population pour chaque réplique. La variance des estimations obtenue à partir de ces répliques pour une caractéristique donnée est une estimation de la variance d'échantillonnage de l'estimation de la population provenant de l'échantillon principal.

Afin d'aider à comprendre ce qu'est le bootstrap, il faut introduire le concept de distribution d'échantillonnage d'un estimateur. Nous avons déjà noté que si tous les échantillons possibles pouvaient être soumis à l'extracteur d'information de l'échantillon qu'est l'estimateur, nous obtiendrions alors toutes les estimations possibles. Supposons que nous ayons toutes les estimations possibles. Nous pourrions alors représenter graphiquement la fréquence à laquelle nous observons chacune de ces estimations; cela s'appelle la distribution d'échantillonnage de l'estimateur. Voici un exemple d'une telle représentation graphique, la distribution d'échantillonnage d'un estimateur.

8. Le défi ici est de parvenir à adapter le bootstrap avec succès à l'échantillonnage, bien que cette méthode ait d'abord été conçue à d'autres fins que des enquêtes. Notre compréhension du bootstrap dans le cadre d'une enquête et de sa bonne mise en œuvre ne cesse d'évoluer depuis une dizaine d'années, soit depuis que l'ELNEJ s'est mise à l'utiliser.



Source : Statistique Canada, Enquête longitudinale nationale sur les enfants et les jeunes.

Nous remarquons, par exemple, que de très petites estimations sont rares, tout comme les très grandes estimations, et que les estimations les plus fréquentes sont quelque part au centre. Dans notre exemple, les estimations sont rassemblées en grappe autour de la valeur moyenne de 5 (qui serait la valeur inconnue si l'estimateur était sans biais) et les estimations supérieures à 5,275, par exemple, sont rares (elles surviennent moins de 0,5 % du temps). Le fait que les estimations ne sont pas toutes identiques, mais qu'elles s'étendent plutôt en une plage de valeurs, est la représentation graphique de la variance d'échantillonnage. En effet, par définition, la variance d'échantillonnage est une mesure de la variabilité observée dans la distribution des estimations de l'histogramme. La variance d'échantillonnage est donc l'une des nombreuses caractéristiques de la distribution d'échantillonnage de l'estimateur.

La distribution d'échantillonnage réelle d'un estimateur donné se fait à partir de tous les échantillons possibles. La représentation bootstrap de cette distribution d'échantillonnage est basée sur le rééchantillonnage effectué à partir de l'échantillon observé, c'est-à-dire celui pour lequel nous recueillons des données sur le terrain. Quand cette correspondance existe, on peut alors estimer un paramètre donné de la distribution d'échantillonnage (comme sa variance) en le calculant pour la distribution bootstrap. Tout comme l'échantillonnage de départ a donné lieu à des poids de sondage, le rééchantillonnage produit un ensemble de poids bootstrap.

Pour l'ELNEJ, un ensemble de 1 000 poids bootstrap est disponible. Afin d'obtenir une estimation de la variance d'échantillonnage d'un estimateur par le bootstrap, il faut produire les 1 000 estimations qui correspondent aux 1 000 poids bootstrap. En effet, on remplace l'ensemble des poids de sondage utilisé par chacun des 1 000 ensembles de poids bootstrap. Tout comme l'ensemble des poids d'enquête a produit une estimation, les 1 000 ensembles de poids bootstrap produiront 1 000

estimations. La variance calculée de ces 1 000 estimations (bootstrap) correspond précisément à l'estimation de la variance d'échantillonnage de l'estimateur que nous obtenons du bootstrap. Algébriquement, l'estimation de la variance bootstrap \hat{v} est calculée comme suit :

$$\hat{v} = \frac{1}{1000} \sum_{k=1}^{1000} (\theta_k - \theta^*)^2$$

où θ_k est la k-ième estimation bootstrap et θ^* est l'estimation originale pour θ obtenue à partir de l'échantillon. (**Nota** : En pratique, on peut la plupart du temps substituer la valeur moyenne des estimations bootstrap θ_k à θ^* , l'estimation obtenue à partir de l'échantillon, dans le calcul de la variance bootstrap. Cette simplification est utilisée chaque fois qu'on calcule la variance bootstrap avec PROC SUMMARY en SAS sur l'ensemble des estimations bootstrap en ayant recours à la statistique VAR.)

Deux outils, tous deux s'appuyant sur les poids bootstrap, ont été développés afin d'aider les utilisateurs à calculer la variance d'échantillonnage et le CV pour leurs estimations :

- un chiffrier Excel muni d'une interface en Visual Basic qui permet à l'utilisateur d'obtenir des variances d'échantillonnage approximatives pour des proportions dans un grand nombre de domaines, par exemple selon l'âge ou la province (n'est pas disponible pour le cycle 8);
- des macros pour calculer la variance d'échantillonnage à partir des poids bootstrap.

Le choix de l'outil approprié dépend du type d'analyse et du niveau de précision requis.

Dans les cycles précédant le cycle 5, un troisième outil d'approximation des variances était disponible: les tables de recherche de CV. L'effet de plan représentatif (l'effet de plan compare la variance des estimateurs du plan de sondage de l'ELNEJ à la variance des estimateurs d'un échantillon aléatoire simple) permettait aux utilisateurs d'obtenir des CV pour certains domaines, selon la cohorte d'âge ou la province. L'interface en Visual Basic et les macros bootstrap étant plus flexibles et plus précises, ces tables ne sont plus fournies.

13.4.1 Chiffrier de variances d'échantillonnage approximatives pour des proportions

Un ensemble de chiffriers est mis à la disposition des utilisateurs pour leur permettre de calculer la variance d'échantillonnage approximative associée aux estimations de proportions. Disponible en format Excel, l'interface en Visual Basic présente les résultats (calculés par des méthodes de rééchantillonnage) pour des milliers de domaines. Ces domaines comprennent les totalisations croisées selon l'âge, les groupes d'âge, les provinces ou les régions. La taille des échantillons de chaque domaine est aussi disponible.

Les détails sur la façon dont les chiffriers et les interfaces ont été produits, sur leur contenu et sur la façon de les utiliser se trouvent dans la documentation qui accompagne les chiffriers.

13.4.2 Macros SAS et SPSS pour calculer la variance d'échantillonnage à partir des poids bootstrap : Bootvar

Des macros ont été développées en SAS et en SPSS pour calculer la variance d'échantillonnage avec les poids bootstrap; elles forment ce que l'on appelle Bootvar. Les versions les plus récentes de Bootvar sont 3.2 en SAS et 3.0 en SPSS. Bootvar est accessible par l'entremise du lien *Outils de recherche* à http://www.statcan.gc.ca/rdc-cdr/bootvar_sas-fra.htm. Bootvar peut calculer des estimations de la variance pour des choses telles que les totaux, les ratios, la différence des ratios et les coefficients de régression linéaire et logistique, et ce, quel que soit le domaine. En plus du plan d'échantillonnage, cette façon de calculer la variance d'échantillonnage tient également compte des caractéristiques spécifiques à la variable d'intérêt. Enfin, en utilisant Bootvar plutôt que le chiffrier, l'utilisateur n'est pas limité qu'à des domaines prédéfinis⁹.

Cette méthode présente beaucoup d'avantages, mais elle exige un peu plus de travail de la part du chercheur. Il faut plus de temps pour calculer la variance d'échantillonnage avec ces macros qu'avec le chiffrier. L'utilisateur doit se familiariser avec les macros avant de les utiliser. Toutefois, ces macros ont été conçues de façon à être faciles à utiliser. L'utilisateur doit avoir accès aux macros, aux fichiers de données et aux fichiers de poids bootstrap. Il peut avoir accès à ces outils dans les Centres de données de recherche (CDR) de Statistique Canada. De la documentation détaillée sur l'utilisation de ces macros SAS ou SPSS est également disponible dans les CDR.

Malgré le temps requis pour faire les calculs, il est fortement recommandé d'utiliser Bootvar au lieu de Visual Basic, qui utilise des feuilles de calcul Excel, pour obtenir une estimation de la variance d'échantillonnage de toute estimation devant être publiée. Bootvar permet d'obtenir une mesure beaucoup plus précise et adéquate de la variance d'échantillonnage.

Les détails sur l'utilisation de ces programmes se trouvent dans la documentation qui accompagne les programmes et les poids bootstrap.

13.4.3 Autres outils informatiques

D'autres logiciels commerciaux sont en mesure de bien calculer la variance d'échantillonnage avec les poids bootstrap fournis. Par exemple, SUDAAN (avec l'option DESIGN = BRR), WesVar et STATA9.

S'il veut calculer la variance d'échantillonnage pour d'autres estimations que celles comprises dans Bootvar, l'analyste voudra peut-être écrire ses propres programmes de mise en œuvre de la méthode bootstrap. Il s'agit toutefois d'un exercice fastidieux.

9. Il convient de souligner qu'à chaque cycle, l'attrition va croissant par rapport à l'échantillon longitudinal d'origine de l'ELNEJ. L'attrition peut être attribuable à différentes circonstances, par exemple l'incapacité de retrouver les enfants et leurs parents, le refus d'enfants plus âgés ou de leurs parents de participer à l'enquête, ou encore lorsque les enfants ont déménagé à l'étranger. Ainsi, certaines strates a posteriori, qui étaient au départ calibrées selon les chiffres du recensement d'après l'âge des enfants répartis selon le sexe et la province pour l'année 1994 (p. ex. filles de l'Île-du-Prince-Édouard âgées de 5 ans en 1994), risquent désormais de comporter des valeurs manquantes/nulles pour certaines répliques bootstrap. Cette situation aura certains effets sur les estimations de la variance dans le cas des données étudiées à d'autres niveaux d'agrégation pour les données du cycle 8 de l'échantillon longitudinal d'origine (p. ex. selon l'âge seulement; selon le sexe seulement; selon la province seulement; à l'échelle nationale). Par conséquent, dans le cas des chiffres que l'on retrouve à ces niveaux d'agrégation supérieurs, les estimations de l'ELNEJ peuvent être accompagnées d'estimations de la variance qui sont différentes de zéro même si elles concordent avec les chiffres du recensement.

13.4.4 Linéarisation de Taylor et autres techniques

Les fichiers de poids bootstrap contiennent des variables indiquant l'unité primaire d'échantillonnage (UPE) et la strate à partir desquelles l'individu a été sélectionné. Certains logiciels existants (comme Stata, SUDAAN et SAS) ont des procédures qui calculent les estimations de la variance d'échantillonnage en utilisant l'information du plan (la strate et l'UPE) et les poids de sondage. La technique est connue sous plusieurs noms, dont la linéarisation de Taylor ou l'estimation de la variance robuste (ou de Binder). Le problème à utiliser ces procédures avec les données de l'ELNEJ tient au fait qu'elles requièrent au moins deux UPE par strate, une condition qui n'est très souvent pas satisfaite par l'ELNEJ. Bien que les strates puissent être combinées, on ne sait pas encore avec certitude si cela est efficace, car aucune comparaison n'a encore été faite des variances d'échantillonnage obtenues ainsi aux variances d'échantillonnage bootstrap. De plus, plusieurs ajustements doivent être apportés afin de transformer les poids déterminés par le plan d'échantillonnage en poids pour diffusion, et la linéarisation de Taylor ne permet pas de tenir compte de l'effet qu'ont ces ajustements sur la variance. Par conséquent, nous recommandons d'utiliser un des outils déjà décrits dans le présent chapitre (l'interface en Visual Basic ou les poids bootstrap) pour obtenir des estimations de variance d'échantillonnage concordant avec le plan.

Finalement, les logiciels comme SAS ou SPSS calculent une variance pour les estimations produites avec leurs propres procédures intégrées (par exemple PROC UNIVARIATE en SAS). Cependant, bon nombre de ces procédures ne tiennent pas compte du plan de sondage (par exemple, la stratification) même si l'on a recours à l'énoncé WEIGHT, ce qui signifie que la variance obtenue n'est pas la variance d'échantillonnage recherchée (laquelle est très souvent sous-estimée de cette façon). Ces procédures ne sont donc pas recommandées, car elles peuvent mener à des conclusions erronées.

Mais avec SAS 9.2, les procédures PROC SURVEYFREQ, PROC SURVEYLOGISTIC, PROC SURVEYMEANS et PROC SURVEYREG permettent de calculer correctement la variance d'échantillonnage à partir des poids bootstrap fournis. Cela peut être fait avec l'énoncé REPWEIGHTS en spécifiant VARMETHOD=BRR dans l'énoncé PROC.

14.0 Mesures directes

Contexte

La recherche sur le développement des enfants et des jeunes joue un rôle important dans l'élaboration des politiques concernant ceux-ci. L'utilisation de divers outils d'évaluation dans le cadre de l'Enquête longitudinale nationale sur les enfants et les jeunes (ELNEJ) permettra de mieux comprendre les processus du développement au cours de la petite enfance et de l'adolescence et fournira des données pertinentes qui serviront de fondement aux orientations stratégiques visant ces populations.

Les outils d'évaluation retenus aux fins de l'ELNEJ ont été choisis à partir d'une recension détaillée des écrits, de l'élaboration d'un cadre de recherche sur le développement et l'apprentissage de l'enfant, de consultations auprès de nombreux experts au Canada et à l'étranger, de l'examen des documents traitant des différents outils possibles et de l'essai sur le terrain des outils les plus prometteurs. Ces outils ont également été examinés à la lumière de divers critères, notamment leur fiabilité et leur validité, la couverture des domaines visés par le cadre de recherche, la capacité de l'outil de distinguer le développement normal et les retards de développement, la facilité d'utilisation par des intervieweurs et la disponibilité de l'outil en anglais ou en français (ou la facilité de traduction en français ou en anglais). La décision finale a été largement conditionnée par l'opinion des principaux experts qui conseillent l'équipe de l'ELNEJ depuis un certain temps.

L'ELNEJ effectue des mesures directes pour les enfants de 4 à 21 ans. Ces évaluations sont décrites dans ce chapitre.

14.1 Échelle de vocabulaire en images de Peabody

L'Échelle de vocabulaire en images de Peabody (EVIP) est conçue pour mesurer le vocabulaire compris et peut s'appliquer à tous les groupes d'âge jusqu'à l'âge adulte. Élaboré par Lloyd et Leota Dunn, de la University of Hawaii, le test est largement utilisé dans des collectes de données d'envergure ainsi que dans le cadre d'évaluations. Son adaptation française a été mise au point par ces chercheurs et par Claudia M. Thériault de l'Université St. Thomas de Fredericton au Nouveau-Brunswick.

Pour les fins de l'ELNEJ, l'EVIP a servi à mesurer la maturité scolaire des enfants de 4 et 5 ans. Le consentement verbal des parents était requis avant l'administration du test. Une fois ce consentement obtenu, l'intervieweur faisait passer le test à l'enfant à la maison. Il s'agissait pour l'enfant de regarder des images sur un tableau et d'indiquer celle qui correspondait au mot lu à haute voix par l'intervieweur.

Chaque enfant soumis à l'EVIP s'est vu attribuer un score brut basé sur le nombre de réponses correctes. On a également attribué un score normalisé à chacun des enfants. Les scores normalisés permettent d'établir des comparaisons entre les scores des différents groupes d'âge. De toute évidence, un enfant de 5 ans devrait obtenir de meilleurs résultats sur l'EVIP qu'un enfant de 4 ans, et il affichera donc un score plus élevé. Le score normalisé, cependant, tient compte de l'âge de l'enfant.

Les scores normalisés pour un examen sont habituellement basés sur la distribution des scores obtenus à partir de la population entière. En l'absence de scores pour la population entière, on utilise plus souvent une distribution à partir d'échantillon représentatif, appelé l'échantillon standard. À chaque cycle, l'ELNEJ comporte un échantillon représentatif d'enfants. N'importe lequel de ces échantillons serait une option viable pour estimer la distribution des scores mesurés dans la population. Même si que chaque échantillon est choisi de façon probabiliste selon différentes périodes de référence, les experts dans le domaine des mesures cognitives pourraient être en désaccord pour justifier que les différences entre les distributions estimées d'un échantillon à l'autre reflète une véritable différence de population au fil du temps, ou résulte simplement de l'erreur d'échantillonnage.

Pour la robustesse, on avait choisi de ne pas utiliser qu'un seul échantillon normatif mais plutôt tous les échantillons normatifs pour chacun des cycles de collecte de l'ELNEJ, jusqu'à ce que la

répartition globale devienne relativement statique. Ce point a été atteint et les scores normalisés du cycle 8 ont donc été calculés à l'aide des mêmes normes que celles utilisées pour le cycle 4 (v2), le cycle 5, le cycle 6 et le cycle 7¹⁰. Pour le cycle 8, les normes sont basées sur les 28 214 enregistrements provenant des cycles 1 à 5 de l'enquête principale de l'ELNEJ avec les scores bruts de l'EVIP. Quelques enregistrements avec un score brut de zéro à l'EVIP ont été exclus des données utilisées pour créer les normes : 1 enregistrement du cycle 2, 49 enregistrements du cycle 3, et 6 enregistrements du cycle 4. Ces scores de zéro étaient probablement des tests inachevés, ainsi ils ne sont pas fiables et sous-estimeraient la mesure réelle des capacités (en particulier au cycle 3). Pour obtenir les normes, chaque enregistrement a été pondéré par son poids transversal divisé par le poids transversal moyen de tous les enregistrements du même cycle. Dans le cadre de l'EVIP, on a attribué aux individus de l'échantillon de référence des scores normalisés de sorte que la moyenne des scores normalisés corresponde à 100 et que l'écart type s'établisse à 15 pour tous les groupes d'âge en mois. Le lissage « loess » a été appliqué aux données pour s'assurer que les normes de l'EVIP augmentent avec l'âge.

Les mesures de fiabilité de l'EVIP ont été calculées en fonction de l'échantillon de référence américain (Dunn et Dunn, 1981).

14.1.1 Propriétés psychométriques des scores

La présente section traite de la qualité du test dans son application à la population ciblée par l'enquête par opposition à la population initiale pour laquelle le test a été conçu. On estime que le test fournit toujours une évaluation raisonnable des habiletés des enfants et ce, pour les raisons évoquées ci-dessous.

Le score brut

L'un des principaux avantages du test dans le cadre d'une enquête est que celui-ci est adapté à l'âge et au rendement de l'enfant, de sorte qu'il n'est pas nécessaire de poser toutes les questions pour déterminer le niveau de compétence. On choisit une question de départ en fonction de l'âge du répondant et on continue de poser des questions de plus en plus difficiles. Lorsque le répondant semble répondre de façon aléatoire — c'est-à-dire qu'au moins six des huit dernières réponses sont erronées — on met un terme au test et on calcule le score en fonction du rang de la dernière question et du nombre de réponses erronées.

Les questions sont classées en ordre croissant de difficulté et sont conçues de manière à se répartir également sur l'« échelle de difficulté ». Initialement, l'EVIP était calibrée au moyen d'un échantillon représentatif de quelque 5 000 enfants d'expression anglaise. Une démarche semblable a été entreprise pour en calibrer la version française. Le degré de difficulté des items a été calibré à l'aide du modèle de Rasch. Dans le langage de la théorie de la réponse d'item (TRI), on désigne cela par le terme « modèle logistique à un paramètre ».

Étant donné que la calibration du test a été effectuée il y a plusieurs années, on s'attend à observer un certain glissement du niveau de difficulté des items. En effet, la langue elle-même évolue et certains mots deviennent plus ou moins courants. Pour vérifier si c'est le cas, au cycle 4, on a effectué une analyse des items fondée sur la théorie de la réponse d'item et on a calculé les scores en fonction du nouveau degré de difficulté des items. Dans certains cas, on a observé un écart par rapport aux résultats attendus selon le test original.

Cependant, on n'a observé aucun écart systématique dans les différences des résultats mesurés. De façon constante, aucune surestimation ou sous-estimation des capacités de l'enfant n'a été mesurée dans aucune partie du test, et les scores calculés selon la TRI concordaient avec les scores bruts. Le coefficient de corrélation entre les deux scores des

10. Notez qu'aux cycles 1, 2 et 3, on a utilisé différents échantillons standards. Quoique de petites variations soient attendues, celles-ci sont bien à l'intérieur de l'erreur d'échantillonnage provenant des mesures de différents échantillons de norme.

tests s'est établi à 94 % pour la version anglaise et à 96 % pour la version française, ce qui est élevé, peu importe la norme retenue. On estime donc que les scores bruts peuvent être utilisés tels quels.

Le score normalisé

Comme on l'a expliqué plus tôt, on établit le score normalisé à partir de la répartition de la population de chaque âge. En fait, on ne peut pas réellement déterminer la répartition de la population, puisqu'il n'est pas possible d'appliquer le test à l'ensemble de la population. Pour résoudre cette difficulté, on utilise l'échantillon d'un groupe d'âge considéré comme représentatif de la population de ce groupe d'âge pour calculer les percentiles requis.

Cependant, on doit traiter certaines contraintes de l'échantillon avant de normaliser les scores. Lorsqu'on examine les percentiles pour différents âges, on s'attend à obtenir une courbe ascendante, en fonction de l'âge, pour la mesure des aptitudes. Si l'on observe effectivement une tendance globale à la hausse, on relève également un renversement de cette courbe pour bon nombre d'âges. Cela est attribuable à la taille insuffisante de l'échantillon pour chaque groupe d'âge, de sorte que la faible représentation contribue à introduire beaucoup de bruit.

Il est préférable d'utiliser les percentiles de l'échantillon comme point de départ et de lisser la progression en fonction de l'âge jusqu'à ce que l'on obtienne une progression jugée « naturelle ». On a utilisé la progression des percentiles initiaux tirée du manuel de l'EVIP pour déterminer le degré de lissage prévu. On a ensuite utilisé les points obtenus comme percentiles aux fins de la normalisation. Il convient de souligner que même si les caractéristiques des normes étaient semblables, les percentiles tendaient à augmenter au fil des ans, ce qui, selon les experts, était à prévoir.

Le test s'applique habituellement aux enfants pour qui l'âge effectif est 4 ou 5 ans (il est bon de noter que ce groupe peut comprendre des enfants d'âge réel de 3 et 6 ans s'ils passent le test avant leur quatrième anniversaire ou après leur sixième anniversaire au cours de l'année civile suivante).

Note finale

Les scores de l'EVIP utilisés dans le cadre de l'ELNEJ sont une mesure valide des aptitudes. Cependant, pour réduire le plus possible le biais potentiel des estimations dans les analyses, la non-réponse doit être traitée au cas par cas. Des renseignements plus détaillés sur la non-réponse sont présentés au chapitre 12.0.

14.2 Évaluation de la connaissance des nombres

L'évaluation de la connaissance des nombres vise à évaluer comment la compréhension des nombres se développe chez les enfants, grâce à l'examen de la compréhension du système des nombres entiers. Dans le cadre de l'ELNEJ, cette évaluation vise les enfants âgés de 4 et 5 ans.

Cette évaluation a été mise au point par Robbie Case de l'Université de Toronto en collaboration avec ses collègues, notamment Yukari Okamoto de l'University of California à Santa Barbara. Elle se fonde sur la théorie des structures conceptuelles centrales élaborée par Robbie Case pour expliquer le développement de la pensée de l'enfant. Avant sa mort en mai 2002, Robbie Case travaillait à l'adaptation du test aux fins de l'ELNEJ. Après le décès de Robbie Case, Yukari Okamoto a aidé l'équipe de l'ELNEJ à terminer les travaux d'adaptation de l'évaluation pour l'enquête.

Cadre théorique

Selon la théorie de Robbie Case, on peut distinguer quatre niveaux de développement de la compréhension des nombres par les enfants : le niveau prédimensionnel, le niveau unidimensionnel, le niveau bidimensionnel et le niveau bidimensionnel intégré. L'enfant doit acquérir une certaine maîtrise de chaque niveau avant de passer au suivant. Normalement, les enfants atteignent ces

niveaux à l'âge de 4, 6, 8 et 10 ans. Le niveau prédimensionnel évalue la capacité de compter par cœur et de quantifier de petits ensembles à l'aide d'objets concrets. Ces connaissances sont nécessaires pour le niveau unidimensionnel, où les enfants traitent les changements de quantité sans recourir à des objets tangibles ou visibles. Le niveau unidimensionnel évalue les connaissances des enfants relativement à la séquence des nombres ainsi que leur capacité de résoudre des problèmes arithmétiques simples. Pour répondre aux questions, les enfants doivent faire appel au calcul mental. Cette faculté leur permet d'intégrer leur compréhension des nombres et des quantités. L'évaluation mesure les connaissances requises pour assurer l'apprentissage scolaire.

Description de l'évaluation

En consultation avec Robbie Case et Yukari Okamoto, on a adapté le test au contexte de l'ELNEJ. On l'a transformé en une évaluation continue à trois niveaux; certains items ont été révisés ou éliminés pour raccourcir un peu le test. À l'origine, il s'agissait d'un test discontinu (c.-à-d. que l'enfant devait répondre correctement à un nombre suffisant de questions à un niveau donné avant de passer au niveau suivant). Puisqu'on ne peut pas s'attendre à ce qu'un enfant obtienne de bons résultats à un niveau donné sans avoir déjà obtenu de bons résultats au niveau précédent, il semble raisonnable d'interrompre l'évaluation après un certain nombre de réponses erronées. Le test a également été programmé dans l'application informatique de sorte que la règle d'interruption s'applique automatiquement. L'intervieweur pose la question à l'enfant et inscrit la réponse. C'est l'application qui détermine si l'enfant y a répondu correctement ou non.

Le test comporte 22 items. Certaines d'entre elles ont deux volets – a) et b). Les enfants doivent répondre correctement aux deux volets de la question pour réussir cet item. On a adopté cette convention parce que tous les items à deux volets proposent deux options de réponse à l'enfant, de sorte que celui-ci a 50 % de chances de répondre correctement en ne se fiant qu'au hasard¹¹. Le fait d'exiger que les enfants répondent correctement aux deux volets de ces items pour obtenir un point accroît l'assurance que ceux-ci ont effectivement les connaissances requises dans le cadre de ces items.

Les enfants n'ont pas le droit d'utiliser un crayon et du papier pour répondre aux questions, puisqu'elles sont posées oralement. Ils doivent plutôt faire appel au calcul mental, qui intègre la compréhension qu'ont les enfants des nombres et des quantités. Les enfants ont cependant accès à divers instruments comme des jetons et une carte de nombres pour les aider à résoudre les problèmes.

L'évaluation de la connaissance des nombres devrait prendre de 10 à 15 minutes environ.

Calcul des scores

Trois types de scores sont mis à la disposition des chercheurs au cycle 8 : le « score d'âge » (HKNCdS01), le « score brut de 30 points » (HKNCfS03) et le « score brut de 30 points normalisé selon l'âge » (HKNCfS02).

On a calculé le « score d'âge » en fonction des réponses de l'enfant aux items. Pour déterminer le « score d'âge », on attribue un point pour chacun des trois niveaux réussis, et on additionne ensuite les points obtenus (on peut attribuer un maximum de un point par niveau complété). Pour réussir un niveau, l'enfant doit répondre correctement à un certain nombre d'items associés à ce niveau, par exemple au niveau prédimensionnel, il doit donner trois réponses correctes sur cinq. Un enfant qui n'a pas pu répondre correctement aux questions du premier niveau recevra le score minimum (zéro) alors qu'un enfant ayant répondu correctement à toutes les questions des trois niveaux recevra le score maximum (trois).

Le niveau 1 représente la proportion de réponses correctes pour le niveau prédimensionnel, qui comporte cinq items. Pour avoir le score d'âge de ce niveau, l'enfant doit obtenir une proportion d'au moins 0,6 (c.-à-d. trois réponses correctes sur cinq). Le niveau 2 représente la proportion de

11. Par exemple, on demande dans le volet a) laquelle des deux piles de jetons est la plus grande et, dans le volet b), laquelle est la plus petite.

réponses correctes pour le niveau unidimensionnel, qui comporte huit items. Pour avoir le score d'âge de ce niveau, l'enfant doit obtenir une proportion d'au moins 0,6 (c.-à-d. cinq réponses correctes sur huit). Le niveau 3 représente la proportion de réponses correctes pour le niveau bidimensionnel, qui comporte neuf items. Pour avoir le score d'âge de ce niveau, l'enfant doit obtenir une proportion d'au moins 0,6 (c.-à-d. six réponses correctes sur neuf).

Même si le test d'évaluation de la connaissance des nombres comporte 22 questions, l'enfant qui se rend jusqu'au bout doit répondre en réalité à 30 questions, puisque certaines d'entre elles ont deux volets, soit a) et b). Le score brut de 30 points correspond simplement au nombre total de réponses exactes à ces 30 questions.

Un score brut de 30 points normalisé selon l'âge a également été attribué à chaque enfant. Les scores normalisés permettent d'établir des comparaisons entre les scores des groupes d'âge différents. De toute évidence, un enfant de 5 ans devrait obtenir de meilleurs résultats qu'un enfant de 4 ans au test d'évaluation de la connaissance des nombres. Le score normalisé tient compte de l'âge de l'enfant. Les normes utilisées dans le procédé de normalisation ont été établies à l'aide des scores bruts de 30 points pour la connaissance des nombres des cycles 4 et 5. Pour obtenir les normes, chaque enregistrement a été pondéré par son poids transversal divisé par le poids transversal moyen de tous les enregistrements du même cycle. Les enfants de l'échantillon de référence ont reçu des scores normalisés de sorte que la moyenne des scores normalisés correspondait à 100 et que l'écart type s'établissait à 15 pour tous les groupes d'âge. Cette procédure de normalisation a été effectuée pour tous les groupes d'âge en mois. Le lissage « loess » a été appliqué aux données pour s'assurer que les normes augmentent avec l'âge.

Validité de l'évaluation

On a effectué une analyse des données relatives à la connaissance des nombres afin de valider l'évaluation. Dans le cadre de cette analyse, on a établi des comparaisons entre le « score d'âge » et l'âge de l'enfant ainsi qu'avec le module Qui suis-je? et on a examiné les items et la non-réponse. Lors de la mise au point des procédures de calcul des scores, l'équipe de l'ELNEJ a consulté Yukari Okamoto pour s'assurer que ces procédures concordaient avec la théorie élaborée par Robbie Case.

L'ensemble des preuves ont validé le test et il devrait fournir aux utilisateurs de données des renseignements sur l'acquisition par l'enfant des compétences requises pour réussir le programme scolaire de mathématiques. Cependant, cette évaluation n'est pas exempte du biais de non-réponse. Des renseignements détaillés sur la non-réponse sont présentés au chapitre 12.0.

14.3 Qui suis-je?

Le module Qui suis-je?¹² vise à évaluer le niveau de développement des jeunes enfants âgés de 3 à 7 ans. Dans l'ELNEJ, cette évaluation s'applique aux enfants de 4 et 5 ans.

L'évaluation a été mise au point par Molly de Lemos et ses collègues de l'*Australian Council for Educational Research (ACER)*. L'équipe de l'ELNEJ a collaboré étroitement avec Molly de Lemos pour apporter des modifications à l'évaluation aux fins de l'ELNEJ (principalement l'élimination de la tâche de dessin) ainsi que pour améliorer l'administration et les procédures de calcul des scores en fonction de l'ELNEJ.

Cadre théorique

Le module Qui suis-je? évalue le niveau de développement des jeunes enfants de 3 à 7 ans. Le Qui suis-je? comporte des tâches de reproduction et d'écriture. Les tâches de reproduction visent à évaluer la capacité de l'enfant à conceptualiser et à reconstruire une forme géométrique. Les tâches

12. Des renseignements plus détaillés sur Qui suis-je? sont présentés dans le document « Patterns of Young Children's Development: An International Comparison of Development as Assessed by *Who Am I?* » de Molly de Lemos (R-02-5E). Ce document de recherche a été publié par Ressources humaines et Développement des compétences Canada (www.rhdcc.gc.ca).

d'écriture évaluent sa capacité à comprendre et à utiliser des représentations symboliques telles que les nombres, les lettres et les mots. La capacité de l'enfant à effectuer les tâches dépend de plusieurs facteurs comme la maturité, la culture, l'expérience et les compétences linguistiques.

Le recours aux tâches de reproduction de formes géométriques pour évaluer le niveau de développement des enfants est bien établi. Ce genre d'évaluation fait partie des mesures de l'intelligence et du développement depuis longtemps. Les recherches de Piaget sur le développement des concepts spatiaux chez les jeunes enfants confirment également la validité des tâches de reproduction pour mesurer le niveau de développement.

Puisque le module Qui suis-je? évalue le langage non verbal, on peut l'utiliser pour évaluer les enfants ayant une connaissance limitée de l'anglais ou du français. Ces enfants peuvent effectuer les tâches dans leur langue maternelle de même qu'en anglais ou en français. Leurs scores dans leur langue maternelle donneront des indications quant au niveau de développement, tandis que les scores en anglais ou en français donneront une idée du développement des enfants dans cette langue. Dans le cadre de l'ELNEJ, on a convenu d'évaluer les enfants uniquement en anglais ou en français et ce, pour deux raisons. Premièrement, on a jugé que l'évaluation du développement de l'enfant dans l'une des langues officielles constitue un indicateur important de la capacité de l'enfant à fonctionner dans le système scolaire canadien. Deuxièmement, il aurait été difficile sur le plan opérationnel d'attribuer des scores pour les questionnaires dans toutes les langues parlées au Canada.

Les tâches ont été mises au point à la lumière de recherches indiquant que les compétences de reproduction sont fortement associées au rendement scolaire ultérieur, qu'elles sont valides pour différents groupes culturels et qu'elles fournissent une mesure fiable du développement au moment de l'évaluation. Par ailleurs, les tentatives précoces d'écriture par les enfants sont liées à une compréhension croissante de la représentation des sons parlés par les caractères écrits.

Description de l'évaluation

L'évaluation Qui suis-je? comporte trois échelles : l'échelle de reproduction, l'échelle des symboles et l'échelle du dessin. L'échelle de reproduction présente des formes (un cercle, une croix, un carré, un triangle et un losange) que l'enfant doit tenter de reproduire. L'échelle des symboles est composée de tâches d'écriture (écrire leur nom, quelques lettres, des nombres, des mots et une phrase) que l'enfant tente d'accomplir. Les enfants ne sont tenus de faire que ce dont ils se sentent capables, mais on les encourage à essayer au moins chacun des exercices. Sur l'échelle du dessin, on demande aux enfants de dessiner une image d'il- ou d'elle-même. L'échelle du dessin n'est pas utilisée dans l'ELNEJ en raison de contraintes de temps. On a consulté Molly de Lemos avant d'éliminer l'échelle du dessin de l'ELNEJ.

L'évaluation se présente sous la forme d'un livret attrayant dans lequel l'enfant effectue les tâches à mesure que l'évaluateur tourne les pages et donne ses instructions. L'évaluation prend une dizaine de minutes et le calcul des scores est fait au bureau central. L'enfant en fait autant qu'il peut, mais on l'encourage à produire au moins un griffonnage pour chacune des tâches.

Calcul des scores

Dans le cadre de l'ELNEJ, le module Qui suis-je? est noté manuellement par des employés qualifiés de Statistique Canada. Ces employés ont été formés pour reconnaître, dans les réponses de l'enfant, certains signes propres à chaque niveau. Les correcteurs qui ne sont pas en mesure de déterminer quel est le niveau de l'enfant parce que les travaux de celui-ci ne correspondent pas clairement à un niveau donné doivent trancher la question du niveau de l'enfant à la lumière des scores obtenus à d'autres items. La notation a été effectuée par un nombre restreint de personnes et a été soumise à des procédures de contrôle de la qualité.

Toutes les questions sont notées par les correcteurs sur une échelle de 1 à 4. Si l'enfant n'a pas essayé de répondre, un score initial de 0 est attribué. Les réponses seront imputées plus tard. En fin de compte, un score de 1 à 4 sera attribué à toutes les questions. Le score de l'échelle de reproduction (HWICdS02) correspond à la somme des scores attribués aux tâches liées à la reproduction d'un symbole. Le score de l'échelle des symboles (HWICdS03) correspond pour sa

part à la somme des scores attribués aux tâches liées à la rédaction. Comme il y a cinq tâches pour chacune de ces échelles, le score de l'échelle de reproduction et le score de l'échelle des symboles oscillent tous les deux entre 5 et 20.

Un score total combiné, le score total de l'échelle Qui suis-je? (HWICdS01), complète les deux échelles retenues aux fins de l'ELNEJ. Ce score n'est en fait que le total du score de l'échelle de reproduction et du score de l'échelle des symboles. Il se situe donc entre 10 et 40, et il donne une indication générale du niveau de développement de l'enfant.

Il convient de noter encore une fois que les réponses aux questions sont imputées avant de calculer les scores.

Des scores d'âge normalisés étaient disponibles pour la première fois au cycle 6. Le score normalisé total de l'échelle Qui suis-je? (HWICfS04), le score normalisé de l'échelle de reproduction (HWICfS05) et le score normalisé de l'échelle des symboles (HWICfS06) constituent les variables du cycle 8. Les scores normalisés permettent d'établir des comparaisons entre les scores des groupes d'âge différents. De toute évidence, un enfant de 5 ans devrait obtenir de meilleurs résultats qu'un enfant de 4 ans au test Qui suis-je? Les scores normalisés tiennent compte de l'âge de l'enfant. Les normes utilisées dans le procédé de normalisation ont été établies à l'aide des scores bruts du test Qui suis-je? des cycles 4 et 5. Pour obtenir les normes, chaque enregistrement a été pondéré par son poids transversal divisé par le poids transversal moyen de tous les enregistrements du même cycle. Les enfants de l'échantillon de référence ont reçu des scores normalisés, de sorte que la moyenne des scores normalisés correspondait à 100 et que l'écart type s'établissait à 15 pour tous les groupes d'âge. Cette procédure de normalisation a été effectuée pour tous les groupes d'âge (en mois). Le lissage « loess » a été appliqué aux données pour s'assurer que les normes augmentent avec l'âge.

Imputation

Selon Molly de Lemos, il est nécessaire d'attribuer un score dans le cas des réponses codées « 0 » (pas d'essai) lors de la sommation des scores des tâches de Qui suis-je? afin d'obtenir le score total des échelles de reproduction et des symboles.

Dans la plupart des cas, on présume que cette forme de non-réponse signifie que l'enfant est incapable d'accomplir la tâche. Du point de vue du développement, cela équivaut à un griffonnage. Pour l'établissement des normes, les réponses « pas d'essai » ont été considérées comme des griffonnages et ont été associées à un score de 1.

Toutefois, on a observé dans certains cas que des enfants capables de réponses plus avancées à des items antérieurs n'ont pas tenté d'effectuer certaines des tâches plus difficiles, notamment la reproduction du losange et la rédaction de la phrase. Dans ces cas, l'attribution d'un score de 1 aurait sous-estimé le niveau de développement de l'enfant. On a donc adopté une procédure du test pour traiter les cas où l'enfant n'essaie pas d'accomplir une tâche donnée. Cette procédure consiste à attribuer un score fondé sur celui obtenu à d'autres items. Par exemple, l'enfant qui a obtenu un score de 4 pour la reproduction du carré et qui n'a pas tenté de reproduire le losange obtient un score de 3 pour la tâche du losange.

Selon Molly de Lemos, il importe de procéder à une imputation pour assurer une meilleure compatibilité entre les données de l'ELNEJ et les données recueillies par le module Qui suis-je? dans d'autres études. Dans la plupart des cas, le module Qui suis-je? est administré par l'enseignant de l'enfant ou un chercheur de l'ACER ayant une formation en développement de l'enfant. L'ELNEJ s'appuie sur des intervieweurs profanes qui ne disposent que de peu de temps, dans un contexte d'interview, pour établir une relation avec l'enfant. Il est donc plus difficile pour l'intervieweur de convaincre l'enfant de tenter d'accomplir les tâches plus ardues. Les règles d'imputation visent à redresser les scores de manière à mieux refléter le niveau de développement de l'enfant.

Validité de l'évaluation

On a effectué une analyse des données issues du module Qui suis-je? pour déterminer le degré de validité de l'évaluation. Dans le cadre de cette analyse, on a établi des comparaisons entre le score d'âge et l'âge de l'enfant, des comparaisons avec le test de connaissance des nombres, des comparaisons avec le Qui suis-je? Lors de la mise au point des procédures de calcul des scores, l'équipe de l'ELNEJ a consulté Molly de Lemos.

Toutes les preuves indiquent que le test était valide et qu'il devrait fournir aux utilisateurs des données des renseignements sur le niveau de développement de l'enfant. Cette évaluation n'est pas exempte du biais de non-réponse. Des renseignements détaillés sur la non-réponse sont présentés au chapitre 12.0.

14.4 Exercice de calculs mathématiques

L'exercice de calculs mathématiques administré à l'enfant est une version abrégée de l'épreuve de calculs mathématiques du test canadien normalisé de connaissances, deuxième édition (Canadian Achievement Test – CAT/2). Il s'agit d'un ensemble de tests destinés à mesurer l'acquisition des compétences scolaires de base.

Le CAT/2 mesure les capacités de l'élève à effectuer des opérations d'addition, de soustraction, de multiplication et de division sur les nombres entiers, les décimales, des fractions, les négatifs et les exposants. La résolution de problèmes comportant des pourcentages et l'ordre des opérations sont également évalués. Depuis le cycle 5, la version abrégée du test mise au point pour les fins de l'ELNEJ se compose de 20 questions à chaque niveau.

À chaque niveau, le test subi au cycle 8 est le même qu'au cycle 7. Aucune modification n'a été apportée au nombre de questions ni aux questions elles-mêmes.

Calcul des scores

Tous les enfants qui ont subi le test de mathématiques se sont vu attribuer un score brut (HMACS01) et un score gradué (que l'on appelle « score gradué classique ») (HMACfS02) et une variable (HMACgD03) qui fournit de l'information sur le nombre total de questions répondues par le répondant et exclut les questions comportant des valeurs manquantes.

Le score gradué classique est calculé selon les normes établies en 1992 par le *Canadian Test Centre* (CTC). Le CTC a mis au point ces normes à partir d'un échantillon d'enfants des dix provinces du Canada, appelé « échantillon normatif » (par contre, le test a été élaboré en anglais seulement, de sorte que l'échantillon ne représente que les élèves des écoles anglaises). Les enfants de l'échantillon normatif ont subi le test complet. Les scores gradués sont des unités d'une seule échelle ayant des intervalles équidistants couvrant tous les niveaux de scolarité. L'échelle a été construite selon la procédure de Thurstone, qui découle de la théorie classique des tests.

L'emploi d'une version abrégée du test pour les enfants dans l'échantillon de l'ELNEJ fait en sorte qu'il n'est pas possible d'associer directement les scores gradués du CTC aux scores bruts obtenus dans l'enquête. Par conséquent, on a utilisé l'échantillon normatif du CTC pour calculer le rang-centile de chaque score brut pour notre version abrégée du test. Par exemple, au niveau 6, dans le test abrégé, un rang-centile de 0,94 % correspond à un score brut de 1. Dans le test complet, les rangs-centiles de 0,55 % et 0,99 % correspondent respectivement à des scores bruts de 3 et 4 et à des scores gradués de 315 et 319. Après une interpolation linéaire, on obtient un score gradué de 318 correspondant à un score brut de 1 dans la version abrégée du test.

Le tableau ci-dessous montre le lien entre les scores bruts et les scores gradués selon le niveau pour le test de mathématiques de l'ELNEJ.

Lien entre les scores bruts et les scores gradués (classiques) selon le niveau pour le test de mathématiques

Score brut	Les scores gradués (classiques)						
	Niveau 4	Niveau 5	Niveau 6	Niveau 7	Niveau 8	Niveau 9	Niveau 10
0	267	294	311	330	361	397	406
1	285	306	318	338	376	423	419
2	301	324	332	359	401	449	430
3	314	339	347	381	425	477	443
4	327	355	365	405	443	504	475
5	339	370	383	426	464	530	495
6	350	382	397	444	480	554	518
7	361	392	409	461	494	574	536
8	371	403	421	477	506	589	565
9	380	414	433	492	517	605	581
10	388	425	445	506	529	623	597
11	396	434	456	518	540	641	619
12	405	443	468	529	557	659	636
13	416	453	480	541	570	678	662
14	425	464	495	550	583	696	681
15	434	478	510	559	597	717	703
16	445	489	527	574	614	739	724
17	458	503	544	594	637	760	751
18	475	522	564	611	664	781	791
19	497	540	584	636	684	803	830
20	524	568	622	674	729	825	871

Source : Statistique Canada, Enquête longitudinale nationale sur les enfants et les jeunes.

Au cycle 5, les évaluations de mathématiques des enfants des niveaux 2 à 10 (chez les moins de 16 ans) ont été administrées beaucoup plus tôt que dans les cycles précédents. L'effet de ce changement est mentionné dans le Guide de l'utilisateur du cycle 5 comme suit.

« Les scores bruts mesurés pendant la collecte sont influencés par la période de collecte différente dans le cas des enfants et des jeunes qui ont passé le test. Par conséquent, le score équivalent du CTC subira le même effet. C'est une évaluation injuste des enfants, car leurs capacités sont directement liées au programme d'études suivi. On a utilisé un modèle estimatif de l'apprentissage progressif, fondé sur tous les enfants évalués au fil du temps, pour compenser ces différences et on a produit un score corrigé (EMACfS02). »

Les utilisateurs doivent noter qu'à l'exception du cycle 5, une méthodologie semblable a été appliquée à tous les autres scores restants (p. ex. AMACs02-DMACs02, FMACfS02, GMACfS02 et HMAcS02).

14.5 Exercice de résolution de problèmes (16 et 17 ans)

L'exercice de résolution de problèmes a été mis en place au cycle 6 pour les jeunes de 16 et 17 ans. Cette évaluation est en fait une combinaison des mesures cognitives utilisées dans les cycles précédents pour ce groupe d'âge. Afin de continuer à mesurer le développement des enfants, on a estimé qu'une mesure plus complète des capacités était nécessaire pour voir comment ceux-ci se préparaient à relever les défis à venir. Cette période de la transition est très importante, car certaines décisions relatives aux études commencent à influencer sur le cheminement de carrière des enfants. On sait que certaines habiletés cognitives, comme la compréhension de la lecture, la résolution de problèmes et la prise de décisions jouent un rôle clé dans les choix et les possibilités qui se présentent alors aux jeunes.

Stratégie et révision

L'exercice de résolution de problèmes (livret 32) a été révisé, passant de 20 questions au cycle 6 à 18 questions au cycle 7. Il s'agit de la combinaison des livrets 30 et 31 utilisés dans le cycle 5. Les questions 1 et 2 ont été éliminées de l'évaluation du cycle 7, et les questions restantes ont été renumérotées. Les deux questions (questions 1 et 2) avaient été choisies parmi les tests de mathématiques de niveaux 7 et 9, qui ne permettent pas l'emploi d'une calculatrice. Toutefois, les répondants peuvent utiliser une calculatrice pour effectuer l'exercice de résolution de problèmes. Au cycle 6, les deux questions sur le calcul mathématique ne pouvaient donc pas être utilisées pour faciliter la correspondance entre l'évaluation de l'exercice de résolution de problèmes avec le test de mathématiques. Par conséquent, on a décidé de supprimer ces questions du cycle 7. Le test administré au cycle 8 est le même que le test utilisé au cycle 7.

Les questions de cet instrument de mesure proviennent d'un autre instrument déjà élaboré et mis à l'essai dans le contexte de l'Enquête auprès des jeunes en transition (EJET). Compte tenu des restrictions relatives au droit d'auteur concernant les questions et au fait que l'ELNEJ utilise des questions similaires auprès de populations qui pouvaient se chevaucher, l'ELNEJ a obtenu la permission d'utiliser des items mis à l'essai mais qui étaient exclus du questionnaire principal de l'EJET. La population cible utilisée était différente de celle évaluée dans le cadre de l'ELNEJ, mais on a estimé que le concept cognitif était encore approprié pour cette cohorte.

Méthode de calcul des scores

Neuf des dix-huit questions de l'exercice sont notées manuellement par des employés qualifiés de Statistique Canada. Ces personnes ont été formées pour noter les questions selon le guide de notation qui leur a été remis. La notation a été effectuée par un nombre restreint de personnes et a été soumise à des procédures de contrôle de la qualité. Le chef de l'équipe de notation a également consulté un membre de l'équipe de spécialistes au sujet des procédures de notation.

Les correcteurs attribuent une valeur de « 0 » à « 3 » ou de « 9 » aux questions complexes. Un score de « 0 » est attribué pour une mauvaise réponse, un score de « 1 » est attribué lorsqu'il y a une réponse partiellement bonne et un score de « 2 » est attribué lorsqu'il y a une bonne réponse. L'exception à la règle est le cas d'une question pour laquelle on attribue un « 2 » pour une réponse partiellement bonne de niveau supérieur et un « 3 » pour une bonne réponse. Un score de « 9 » est attribué pour une valeur manquante. Lorsqu'une réponse a été proposée puis rayée, un score de « 0 » (aucune réponse) est attribué par opposition à un « 9 » (valeur manquante). Cela englobe les réponses effacées ou rayées, sauf s'il est clair que la réponse effacée ou rayée est bonne ou partiellement bonne, auquel cas un score de « 1 », « 2 » ou « 3 » est attribué. Si le répondant a écrit quelque chose d'illisible pour le correcteur (du gribouillage, par exemple), un score de « 0 » est attribué. Les questions complexes proviennent du PISA (Programme international de l'OCDE pour le suivi des acquis des élèves), qui permet des notes partielles dans son guide de notation. Par conséquent, cette approche de notation a été adoptée pour le cycle 7 et le cycle 8. Veuillez noter qu'au cycle 6, les questions complexes notées par les correcteurs ont été associées aux valeurs 1, 2 ou 9. Un score de « 1 » a été donné pour une bonne réponse, un score de « 2 » pour une mauvaise réponse, et un score de « 9 » pour les valeurs manquantes. Aucun score n'est attribué pour les réponses partiellement bonnes.

Les items restants dans le livret sont des données saisies par du personnel compétent de Statistique Canada et analysées par des méthodologistes de l'ELNEJ. Les scores des items notés et les données constituant les réponses saisies permettent de faire des analyses et de calculer un score global.

Scores

(HMAygS04)

Ce score est parfois nommé le « score brut » et représente la somme des scores associés aux 18 questions, sans y inclure les valeurs manquantes.

(HMAygD02)

Au cycle 7, une nouvelle variable a été introduite. Cette variable fournit de l'information sur le nombre total de questions auxquelles a répondu le répondant et exclut les questions comportant des valeurs manquantes.

14.6 Évaluation de la littératie (18 et 19 ans)

Cette évaluation directe qui mesure le degré de littératie des jeunes de 18 et 19 ans a été ajoutée à l'ELNEJ au cycle 6. Elle comporte 36 questions extraites d'une version abrégée et adaptée de l'Enquête internationale sur l'alphabétisation et les compétences des adultes (EIAA) et de l'Enquête sur la littératie et les compétences des adultes (ELCA). Le principal objectif de l'évaluation est d'établir dans quelle mesure ces jeunes utilisent l'information imprimée pour fonctionner dans la société. Les questions ont été retenues parce qu'elles mesurent des capacités cognitives déjà établies. L'évaluation porte sur deux types de littératie.

- **Compréhension de textes suivis** qui s'agit des connaissances et compétences requises pour comprendre et utiliser l'information contenue dans des textes comme des éditoriaux, des reportages, des poèmes et des œuvres de fiction.
- **Compréhension de textes schématiques** qui s'agit des connaissances et compétences requises pour repérer et utiliser l'information contenue dans différents documents comme des tableaux, des formulaires, des graphiques et des diagrammes.

Il était important d'intégrer les mesures de la littératie dans l'ELNEJ, compte tenu de l'évolution de la signification de ce concept. Les définitions de la lecture et de la littératie ont évolué au fil du temps, parallèlement à l'évolution de la société, de l'économie et de la culture. En accordant de plus en plus d'importance à l'acquisition continue du savoir, on a étendu les notions de lecture et de littératie et accru les exigences à leur égard. La littératie est considérée comme un ensemble évolutif de compétences, de connaissances et de stratégies qu'une personne met en œuvre tout au long de sa vie dans divers contextes ainsi que dans ses relations avec ses pairs et sa collectivité.

Méthode de calcul des scores

Cet exercice est noté manuellement par des employés qualifiés de Statistique Canada. Ces personnes ont été formées pour noter les questions selon le guide de notation qui leur a été remis. La notation a été effectuée par un nombre restreint de personnes et a été soumise à des procédures de contrôle de la qualité. Le chef de l'équipe de notation a également consulté un membre de l'équipe de spécialistes au sujet des procédures de notation.

Pour l'évaluation de la littératie, il est possible d'attribuer trois codes à chaque question. Le code « 1 » est attribué pour une bonne réponse, le code « 7 » est attribué pour une mauvaise réponse et le code « 0 » est attribué lorsqu'il y a une valeur manquante. Ces codes sont inscrits sur la feuille de scores à la fin de chaque livret. Une fois l'évaluation terminée, les données sur les feuilles de scores sont saisies et transmises à l'équipe de l'ELNEJ à des fins d'analyse.

Le score attribué à la littératie (HLTYfS01) correspond tout simplement au nombre de bonnes réponses données aux 36 questions. Une variable additionnelle est dérivée (HLYgD02) laquelle fournit de l'information sur le nombre total de questions auxquelles a répondu le répondant et exclut les questions comportant des valeurs manquantes.

Même si les questions du test d'évaluation de la littératie de l'ELNEJ ont été choisies à même celles de l'EIAA de 2003, les scores de l'évaluation de la littératie de l'ELNEJ et celui de l'évaluation de la compréhension de textes schématiques de l'EIAA ne sont pas comparables. Les différences relatives à la façon dont l'évaluation a été réalisée, au processus de notation et à la méthode de calcul des scores sont si importantes que les deux scores ne doivent pas être comparés.

Note : Les questions du test d'évaluation de la littératie sont posées à la maison par l'intervieweur, et le jeune n'est pas limité dans le temps, c'est-à-dire qu'il peut prendre le temps qu'il juge nécessaire pour effectuer le test.

14.7 Évaluation de la numératie (20 et 21 ans)

Ce test d'évaluation a été ajouté au cycle 6 pour mesurer les compétences en numératie des répondants les plus âgés. Comme dans le cas de l'évaluation de la littératie, les 32 questions du test proviennent d'une version abrégée et adaptée de l'EIACA et de l'ELCA. Et comme c'était le cas aussi pour l'évaluation de la littératie, l'objectif principal du test est d'établir la mesure dans laquelle ces jeunes utilisent l'information imprimée pour fonctionner dans la société.

La numératie renvoie aux connaissances et aux compétences requises pour gérer efficacement diverses situations dans lesquelles on fait appel aux mathématiques. Certains chercheurs ont mis au jour des liens entre la littératie et la numératie, et ces deux concepts constituent d'importants déterminants de la réussite en milieu de travail.

Comme dans le cas des tâches concernant la compréhension de textes suivis et schématiques, les individus doivent, pour accomplir les tâches quantitatives, jumeler de l'information extraite d'une question ou d'une directive à de l'information énoncée dans un ou plusieurs textes, qu'ils soient continus ou non continus. En outre, les répondants peuvent avoir à composer avec des distracteurs plausibles dans le cadre de l'extraction d'information pour une opération arithmétique. Ils doivent aussi traiter certains types de renseignements. Tandis que le type de renseignements varie pour ce qui est de la compréhension de textes suivis et schématiques, l'information requise est toujours exprimée en tant que tâches quantitatives.

Méthode de calcul des scores

Tout comme pour le test d'évaluation de la littératie, cet exercice est noté manuellement par des employés qualifiés de Statistique Canada. Ces personnes ont été formées pour noter les questions selon le guide de notation qui leur a été remis. La notation a été effectuée par un nombre restreint de personnes et a été soumise à des procédures de contrôle de la qualité. Le chef de l'équipe de notation a également consulté un membre de l'équipe de spécialistes au sujet des procédures de notation.

Pour l'évaluation de la numératie, il est possible d'attribuer trois codes à chaque question. Le code « 1 » est attribué pour une bonne réponse, le code « 7 » est attribué pour une mauvaise réponse et le code « 0 » est attribué lorsqu'il y a une valeur manquante. Ces codes sont inscrits sur la feuille de scores à la fin de chaque livret. Une fois l'évaluation terminée, les données sur les feuilles de scores sont saisies et transmises à l'équipe de l'ELNEJ à des fins d'analyse.

Le score attribué à la numératie (HNUYfS01) correspond tout simplement au nombre de bonnes réponses données aux 32 questions. Une variable additionnelle est dérivée (HNUYgD02) laquelle fournit de l'information sur le nombre total de questions auxquelles a répondu le répondant et exclut les questions comportant des valeurs manquantes.

Même si les questions du test d'évaluation de la numératie de l'ELNEJ ont été choisies à même celles de l'EIACA de 2003, le score de l'évaluation de la numératie de l'ELNEJ et celui de l'évaluation de la numératie de l'EIACA ne sont pas comparables. Les différences relatives à la façon dont l'évaluation a été réalisée, au processus de notation et à la méthode de calcul des scores sont si importantes que les deux scores ne doivent pas être comparés.

Note : Les questions du test d'évaluation de la numératie sont posées à la maison par l'intervieweur, et le jeune n'est pas limité dans le temps, c'est-à-dire qu'il peut prendre le temps qu'il juge nécessaire pour effectuer le test.

15.0 Enjeux analytiques

Dans le présent chapitre, on présente aux utilisateurs un aperçu des différents enjeux analytiques qu'ils doivent considérer lors de tout projet d'analyse avec les données de l'Enquête longitudinale nationale sur les enfants et les jeunes (ELNEJ). On y soulève quelques points qui ont déjà été abordés plus en détail dans d'autres chapitres de ce guide. L'objectif est de faire ressortir quelques enjeux importants sur l'analyse de données tels que :

- Statistique Canada recommande d'utiliser les poids de l'enquête dans les analyses lorsque ceci est possible;
- il y a lieu d'être vigilant en regroupant les données de plusieurs cycles puisque les mesures reprises au fil du temps concernant un individu sont susceptibles d'être corrélées;
- des poids standardisés ou normalisés peuvent se solder par des estimations incorrectes de la variance;
- une estimation de l'erreur d'échantillonnage doit accompagner les analyses;
- des suggestions concernant les données manquantes devraient être faites.

Dans le chapitre 5.0 « Méthodologie de l'enquête – Échantillon », on donne une description détaillée de l'échantillon de l'ELNEJ, alors que dans le chapitre 11.0 « Pondération et traitement de la non-réponse », on fait une description de la méthode de calcul des poids de l'enquête. Une description de la méthode d'estimation de l'erreur d'échantillonnage est présentée au chapitre 13.0 « Estimation de la variance », et pour en savoir davantage sur la qualité des données, y compris une description détaillée des diverses sources d'erreurs non dues à l'échantillonnage dans une enquête (non-réponse, réponse, sous-dénombrement, erreurs de traitement, entre autres), veuillez consulter le chapitre 12.0 « Qualité des données, taux de réponse et couverture ».

À l'automne 2008, un atelier de deux jours sur l'analyse des données issues de l'ELNEJ a eu lieu au centre de données de recherche de Toronto. L'atelier a couvert plusieurs sujets traités dans ce chapitre. Les présentations sont accessibles via <http://www.utoronto.ca/rdc/events.html#NLSCY2008>. (Veuillez noter que ces présentations sont seulement disponibles en anglais.)

Pour obtenir de l'aide concernant une question analytique particulière, veuillez communiquer avec le Centre de ressources en analyse de données de Statistique Canada en acheminant votre question à : georgia.roberts@statcan.gc.ca.

15.1 Incidence de l'utilisation d'un plan de sondage complexe sur l'analyse

L'analyse des données suppose qu'on les résume et qu'on les interprète d'une façon permettant d'apporter des réponses claires aux questions à la base de l'enquête. L'analyste ne veut parfois que décrire l'échantillon, mais il arrive plus souvent qu'il désire utiliser l'échantillon pour décrire une certaine population.

En ce qui concerne les inférences concernant une population observée, Statistique Canada recommande d'utiliser les poids de l'enquête (transversaux ou longitudinaux, selon l'analyse). En effet, il est improbable que la répartition d'une certaine caractéristique d'intérêt dans l'échantillon reflète la répartition de la même caractéristique dans la population en raison du plan de sondage complexe. Ce n'est qu'en appliquant les poids de l'enquête que la répartition de la population peut être préservée.

Par exemple, la stratification et la mise en grappes (toutes deux présentes dans le plan de sondage de l'ELNEJ) conduisent à des probabilités inégales de sélection : la probabilité qu'un enfant dans la population soit sélectionné pour l'ELNEJ dépend de son âge, de sa province de résidence, notamment. (Un nombre disproportionné d'enfants dans l'échantillon proviennent des petites provinces.) Des taux de non-réponse inégaux dans la population peuvent également se solder par une représentation inégale des enfants dans l'échantillon. De plus, la mise en grappes de

l'échantillon entraîne une non indépendance statistique des unités : les caractéristiques des enfants d'un même ménage sont corrélés.

Supposons que l'analyste désire connaître la répartition des enfants de la cohorte initiale au Canada selon la province. La population d'inférence est composée des enfants qui avaient entre 0 et 11 ans au 31 décembre 1994 et qui vivaient dans l'une des provinces lors de la collecte du cycle 1 (1994-1995). Deux ensembles distincts de poids longitudinaux pourraient être utilisés : les poids « entonnoir » (pour les enfants qui ont répondu à tous les cycles) et les poids non-entonnoir (pour les enfants qui ont répondu aux cycles 1 et 8, mais pas nécessairement aux cycles intermédiaires). Le tableau ci-dessous illustre la différence entre les estimations pondérées et non pondérées du nombre et de la proportion d'enfants au Canada, en utilisant les poids entonnoir.

Répartition des enfants dans la population, pour la cohorte initiale, estimations pondérées et non pondérées utilisant les poids entonnoir du cycle 8¹

Province	Non pondéré		Pondéré	
	Nombre	%	Nombre	%
Terre-Neuve-et-Labrador	491	6,57	88 986	1,91
Île-du-Prince-Édouard	227	3,04	23 148	0,50
Nouvelle-Écosse	536	7,17	144 088	3,09
Nouveau-Brunswick	450	6,02	115 131	2,47
Québec	1 457	19,50	1 090 582	23,41
Ontario	1 814	24,28	1 773 616	38,08
Manitoba	569	7,62	182 869	3,93
Saskatchewan	643	8,61	173 611	3,73
Alberta	688	9,21	489 913	10,52
Colombie-Britannique	597	7,99	576 125	12,37
Total	7 472	100,00	4 658 069	100,00

1. HWTCWd1L. Voir la section 15.3 ci-dessous.

Source : Statistique Canada, Enquête longitudinale nationale sur les enfants et les jeunes.

S'il n'utilise pas les poids, l'analyste en viendrait à la conclusion inexacte que 22,8 % des enfants habitent dans les provinces de l'Atlantique alors que, en vérité, le pourcentage réel ne se situe qu'à 7,97 %. Les proportions non pondérées reflètent le fait que l'échantillon comprend un nombre disproportionné d'enfants provenant de petites provinces (afin de garantir une taille d'échantillon appropriée dans ces provinces).

15.2 Unité d'analyse

Pour l'ELNEJ, l'unité d'analyse est toujours l'enfant. Certains éléments d'information ont été recueillis au niveau du ménage, mais aucune estimation n'a été faite à ce niveau. Toutes les estimations doivent être faites au niveau de l'enfant. Par exemple, on peut estimer le nombre d'enfants vivant dans un ménage avec un seul parent, mais on **ne peut pas** estimer le nombre de ménages composé d'un seul parent.

15.3 Analyse longitudinale et analyse transversale

Avec l'ELNEJ, les utilisateurs ont la possibilité de faire l'analyse longitudinale ou transversale.

Analyse longitudinale

La population représentée par les poids longitudinaux est la population au moment de la sélection initiale de l'enfant. En ce qui concerne les enfants de la cohorte initiale (âgés de 14 à 25 ans au cycle 8), deux ensembles de poids longitudinaux sont disponibles, c'est-à-dire les variables HWTCW01L et HWTCWd1L. Le premier ensemble de poids comprend les enfants longitudinaux qui ont répondu (ou dont la PMR a répondu) aux cycles 1 et 8, sans égard à leur réponse aux cycles intermédiaires. Le deuxième ensemble de poids longitudinaux, appelé poids « entonnoir », est calculé uniquement pour les enfants qui ont répondu (ou dont la PMR a répondu) à tous les cycles (de 1 à 8).

Par exemple, si vous faites une analyse sur les enfants introduits au cycle 1 et voulez regarder leurs caractéristiques au cycle 8, sans avoir besoin de leur information aux cycles 2 à 7, le premier ensemble de poids longitudinaux vous conviendrait mieux. Toutefois, si vous voulez utiliser l'information de tous les cycles, le deuxième ensemble de poids longitudinaux serait préférable. (Cette analyse serait un exemple de mesures répétées, dont il est question à la section suivante.) Veuillez noter toutefois que les deux ensembles de poids longitudinaux ont été testés et qu'ils donnent des estimations tout à fait semblables pour une série de variables-clés.

Analyse transversale

L'analyse transversale renvoie à l'analyse où la population d'intérêt est une population formée à partir ou à la suite de la population initiale longitudinale. Pour le cycle 8, des poids transversaux ne sont calculés que pour les enfants âgés de 0 à 7 ans (pour obtenir une description de l'échantillon formé de ces enfants, veuillez consulter le chapitre 5.0 ou 11.0 du présent guide). Il n'est pas recommandé d'utiliser la cohorte initiale pour faire des inférences concernant la population transversale de tous les jeunes de 14 à 25 ans vivant au Canada lors de la collecte du cycle 8 puisque cet échantillon n'a jamais été complété pour tenir compte des immigrants arrivés après 1994. (Pour obtenir plus de renseignements sur l'incidence de l'immigration, veuillez consulter le chapitre 12.0 du présent guide.)

15.4 Estimations pondérées simples (totaux, moyennes et proportions)

Dans la présente section, on montre comment utiliser les poids de l'enquête afin de produire des estimations pour des statistiques descriptives simples. Des totaux, des moyennes et des proportions sont généralement estimés concernant un vaste éventail de caractéristiques recueillies auprès des unités de l'échantillon. Ces caractéristiques, souvent appelées des variables, peuvent être nominales ou qualitatives (le sexe, par exemple) ou encore quantitatives (le poids à la naissance, par exemple). Différentes statistiques et différents procédés statistiques peuvent être utilisés pour la production d'estimations, selon le type de données.

Des estimations peuvent être produites pour toute la population observée ou pour des sous-groupes ou domaines précis de la population (les provinces, par exemple). Les estimateurs suivants peuvent être appliqués à tous les plans de sondage probabiliste, simples ou complexes.

15.4.1 Estimation d'un total de la population

L'estimation du nombre total d'unités dans la population observée se calcule en additionnant les poids finaux des unités répondantes :

$$\hat{N} = \sum_{i \in S_r} w_i$$

où i est la i^{e} unité répondante de l'échantillon et w_i est son poids final de l'enquête (c.-à-d. le plan d'échantillonnage ajusté pour la non-réponse et pour la stratification a posteriori tel que décrit au chapitre 11.0) et S_r est l'ensemble des unités répondantes.

Pour les données quantitatives, l'estimation d'une valeur totale correspond au produit du poids final, w_i , et de la valeur y_i , de chaque unité répondante, additionnés pour toutes les unités répondantes :

$$\hat{Y} = \sum_{i \in S_r} w_i y_i$$

15.4.2 Estimation d'une moyenne de population

Pour les données quantitatives, l'estimation d'une valeur moyenne dans la population s'obtient en additionnant le produit de la valeur de l'échantillon et le poids de chaque unité répondante, puis en divisant ce montant par la somme des poids. En d'autres mots, l'estimation de la moyenne dans la population correspond à l'estimation de la valeur totale des données quantitatives divisée par l'estimation du nombre total d'unités dans la population :

$$\begin{aligned}\hat{\bar{Y}} &= \frac{\sum_{i \in S_r} w_i y_i}{\sum_{i \in S_r} w_i} \\ &= \frac{\hat{Y}}{\hat{N}}\end{aligned}$$

15.4.3 Estimation d'une proportion de la population

Pour les données qualitatives, l'estimation de la proportion des unités de la population observée possédant une caractéristique donnée s'obtient en additionnant les poids des unités possédant cette caractéristique, puis en divisant ce total par la somme des poids de tous les répondants. En d'autres mots, l'estimation de la proportion de la population correspond à l'estimation du nombre total d'unités possédant la caractéristique donnée divisée par l'estimation du nombre total d'unités dans la population :

$$\begin{aligned}\hat{P} &= \frac{\sum_{i \in S_r \cap C} w_i}{\sum_{i \in S_r} w_i} \\ &= \frac{\hat{N}_C}{\hat{N}}\end{aligned}$$

où C correspond aux unités de collecte possédant la caractéristique donnée.

15.4.4 Estimation de domaines de la population

Des estimations peuvent être requises pour des domaines tels que le groupe d'âge ou le sexe, entre autres.

La taille de la population concernant un domaine d'intérêt pour des données qualitatives ou quantitatives est estimée de la façon suivante :

$$\hat{N}_{\text{domaine}} = \sum_{i \in S_r \cap \text{domaine}} w_i$$

La formule d'estimation d'un total d'un domaine pour des données quantitatives est :

$$\hat{Y}_{\text{domaine}} = \sum_{i \in S_r \cap \text{domaine}} w_i y_i$$

La formule d'estimation d'une moyenne d'un domaine pour des données quantitatives est :

$$\begin{aligned} \hat{\bar{Y}}_{\text{domaine}} &= \frac{\sum_{i \in S_r \cap \text{domaine}} w_i y_i}{\sum_{i \in S_r \cap \text{domaine}} w_i} \\ &= \frac{\hat{Y}_{\text{domaine}}}{\hat{N}_{\text{domaine}}} \end{aligned}$$

La formule d'estimation d'une proportion d'un domaine pour des données qualitatives ou quantitatives est :

$$\begin{aligned} \hat{P}_{\text{domaine}} &= \frac{\sum_{i \in S_r \cap \text{domaine} \cap C} w_i}{\sum_{i \in S_r \cap \text{domaine}} w_i} \\ &= \frac{\hat{N}_{\text{domaine} \cap C}}{\hat{N}_{\text{domaine}}} \end{aligned}$$

15.5 Poids normalisés

Pour s'assurer que les estimations des caractéristiques de la population finie de l'enquête sont plus ou moins non biaisées par rapport au plan de sondage, un poids est attribué à chaque observation. Cependant, l'utilisation des poids de l'enquête avec certains logiciels non spécialisés dans l'analyse des données d'enquête (SAS et SPSS, par exemple) peut se solder par des conclusions erronées. La raison en est que le logiciel interprète la somme des poids comme le nombre d'unités observées dans l'échantillon et qu'il surestime par conséquent le pouvoir statistique du test.

En effectuant des tests statistiques à l'aide d'un logiciel basé sur un modèle, l'analyste doit au moins remettre à l'échelle les poids initiaux de l'enquête pour s'assurer que la somme des poids normalisés est égale à la taille de l'échantillon. (La plupart des procédures SAS et SPSS sont dotées de fonctions pour normaliser les poids.) Ainsi, le nombre d'unités observées est correct. Toutefois, comme le logiciel basé sur un modèle traite encore les unités comme si elles étaient réparties de façon indépendante et identique, les estimations de la variance d'échantillon produites sont probablement incorrectes.

L'utilisation de poids normalisés peut être considérée comme une mise en œuvre incomplète de la méthode fondée sur le plan d'enquête. On recommande que les poids normalisés ne soient utilisés que s'il est impossible de réaliser correctement une analyse statistique à l'aide d'un logiciel basé sur le plan de sondage, ou si les renseignements concernant le plan d'enquête sont insuffisants pour calculer correctement des variances d'échantillonnage (lorsqu'il n'y a pas de poids bootstrap, par exemple) ou si l'analyste n'exécute que des analyses préliminaires.

Procédé de normalisation

Les poids normalisés sont calculés en divisant le poids final de l'enquête de chaque unité à analyser par la moyenne (non pondérée) des poids de l'enquête de toutes les unités analysées :

$$w_i^{norme} = \frac{w_i}{\bar{w}}$$

Supposons, par exemple, qu'il y a six enfants dans l'échantillon dont les poids finaux d'enquête s'établissent à 1, 3, 4 et 6. Les poids normalisés sont présentés dans le tableau ci-dessous.

Exemple de poids normalisés

Numéro d'observation	Poids final de l'enquête	Poids normalisé
1	1	0,25
2	3	0,75
3	4	1,00
4	4	1,00
5	6	1,50
6	6	1,50
Total	24	6

On ne doit avoir recours à la normalisation que pour des répondants. Par exemple, si la caractéristique d'intérêt est manquante pour deux des unités de l'échantillon, la somme des poids normalisés devrait alors être égale à 4 et non à 6 (voir le tableau ci-dessous).

Exemple de poids normalisés en présence d'unités non-répondantes

Numéro d'observation	Statut de réponse pour la variable d'intérêt	Poids d'enquête	Poids normalisé incorrect (calculé en utilisant les non-répondants)	Poids normalisé correct (répondants seulement)
1	Répondant	1	0,25	0,33
2	Répondant	3	0,75	1,00
3	Répondant	4	1,00	1,33
4	Répondant	4	1,00	1,33
5	Non-répondant	6	Exclu de l'analyse	Exclu de l'analyse
6	Non-répondant	6	Exclu de l'analyse	Exclu de l'analyse
Total		24	3	4

En outre, la normalisation doit être effectuée séparément pour chaque domaine d'analyse afin de s'assurer que la somme des poids normalisés respecte la taille de l'échantillon selon le domaine. Supposons que l'analyse est réalisée par province et que deux répondants vivent dans la province A et que quatre vivent dans la province B. La somme des poids normalisés pour la province A doit correspondre à 2 et la somme des poids normalisés pour la province B doit correspondre à 4 (voir le tableau ci-dessous).

Exemple de poids normalisés selon le domaine d'analyse

Numéro d'observation	Statut de réponse de la variable d'intérêt	Province	Poids d'enquête	Poids normalisé (selon la province)
1	Répondant	A	1	0,50
2	Répondant	A	3	1,50
Sous-total			4	2
3	Répondant	B	4	0,80
4	Répondant	B	4	0,80
5	Répondant	B	6	1,20
6	Répondant	B	6	1,20
Sous-total			20	4

Problème lié à la normalisation

Dans la plupart des enquêtes comportant un plan de sondage complexe, le nombre effectif d'unités dans l'échantillon est plus petit que le nombre réel en raison de la corrélation des unités échantillonnées (qui est souvent le résultat de la mise en grappes). Dans de tels cas, la normalisation entraîne :

- une surestimation du nombre effectif d'unités;
- une sous-estimation de la variabilité;
- de trop nombreux résultats significatifs.

Certains utilisateurs des poids normalisés ont par conséquent recours à un niveau de test plus prudent (habituellement 1 % plutôt que 5 %) avant de déclarer qu'un résultat est significatif. Mais cette règle est imparfaite : dans un cas donné, ou elle est trop rigoureuse ou elle ne l'est pas assez.

L'exemple ci-dessous illustre comment la valeur de p estimée, produite au moyen de poids normalisés, peut être incorrecte. On a recours à :

- un test d'indépendance avec PROC FREQ dans SAS;
- une régression logistique avec PROC LOGISTIC dans SAS.

L'exemple utilise des données synthétiques d'une enquête sur la santé de Statistique Canada.

Exemple 1 — La date de naissance d'une personne mariée est-elle un prédicteur du divorce?

Dans cet exemple, l'analyste désire savoir si l'état matrimonial d'une personne et le moment de l'année où se situe la naissance de cette personne sont corrélés. Est-il nécessaire d'être né dans les trois premiers mois de l'année pour être marié? Les personnes nées dans les trois derniers mois de l'année sont-elles plus susceptibles de ne pas être mariées?

Résultats après la normalisation

SAS : une valeur de $\chi^2 = 24,33$ ($p=0,0038$)

Conclusion — L'analyste en vient à la conclusion que l'état matrimonial et la date de naissance sont corrélés.

Résultats avec un logiciel fondé sur le plan de sondage

SUDAAN : une valeur de $\chi^2 = 14,95$ ($p=0,0955$)

Conclusion — L'analyste en vient à la conclusion fondée que l'état matrimonial et la date de naissance ne sont pas corrélés.

Conclusion

Avec un logiciel basé sur un modèle, la normalisation constitue une tentative pour corriger un certain nombre de procédures, mais elle représente une mise en œuvre incomplète de la méthode fondée sur le plan de sondage puisqu'elle tient compte des poids de l'enquête mais non des autres aspects du plan (comme les unités qui ne sont pas réparties de façon indépendante et identique).

Les poids normalisés entraînent souvent une sous-estimation de la variance d'échantillonnage des estimations et, par conséquent, des énoncés de différences significatives alors qu'il n'y en a pas. Une règle empirique est souvent utilisée pour compenser cet effet, mais cet ajustement est parfois trop prudent et parfois pas assez prudent.

Le domaine d'analyse et le nombre de répondants doivent être pris en considération dans le calcul des poids normalisés.

Si l'on utilise un logiciel fondé sur le plan de sondage (comme SUDAAN ou Bootvar), la normalisation ne doit pas être faite.

15.6 Mesures répétées

Les mesures répétées sont des observations multiples des mêmes variables de la même unité échantillonnée. Les mesures répétées se produisent lorsqu'un échantillon est suivi au fil du temps — comme dans le cas d'une enquête longitudinale — et que les mêmes renseignements sont recueillis à différents moments (taille, poids, entre autres). Dans ce cas, la dépendance entre les observations répétées du même individu doit être justifiée dans l'estimation de la variance d'échantillonnage.

Lorsque l'estimation de la variance est effectuée à l'aide des poids bootstrap (comme cela est recommandé pour l'ELNEJ), il y a un moyen simple de tenir compte de la structure de corrélation des observations multiples du même enfant échantillonné : il faut établir le fichier d'analyse de sorte que chaque enregistrement constitue un événement ou une observation, et que le poids de l'enquête et les poids bootstrap associés à cet enregistrement correspondent au poids de l'enquête et aux poids bootstrap de l'enfant qui a vécu l'événement ou produit l'observation.

Supposons, par exemple, qu'un chercheur soit intéressé par les mesures répétées x et y , disons la taille et le poids, pendant les trois premiers cycles de l'ELNEJ. Disons que x_{ij} représente la taille de l'enfant « i » au cycle « j » et que y_{ij} représente le poids de l'enfant « i » au cycle « j ». Disons que w_1 est le poids d'enquête de l'enfant (poids longitudinal non-entonnoir, par exemple) et disons que $bs1_1$, $bs2_1$, et autres, correspondent aux poids bootstrap de l'enfant. Supposons que, pour le premier enfant, on disposait des données des cycles 1, 2 et 3 et que, pour le deuxième enfant, on ne disposait que des données du cycle 1. Dans le cas du troisième enfant, on a les données des cycles 1 et 2 et, pour ce qui est du quatrième enfant, on a uniquement les données des cycles 2 et 3. Le fichier d'entrée serait alors construit comme ci-dessous, et la variance d'échantillonnage estimée à l'aide des poids bootstrap serait calculée comme d'habitude.

Exemple de construction de fichier d'entrée

Enregis- tremment	Enfant	Poids d'enquête, w	Variable x	Variable y	Poids bootstrap, réplique 1	Poids bootstrap, réplique 2
1	1	w_1	x_{11}	y_{11}	$bs1_1$	$bs2_1$
2	1	w_1	x_{12}	y_{12}	$bs1_1$	$bs2_1$
3	1	w_1	x_{13}	y_{13}	$bs1_1$	$bs2_1$
4	2	w_2	x_{21}	y_{21}	$bs1_2$	$bs2_2$
5	3	w_3	x_{31}	y_{31}	$bs1_3$	$bs2_3$
6	3	w_3	x_{32}	y_{32}	$bs1_3$	$bs2_3$
7	4	w_4	x_{42}	y_{42}	$bs1_4$	$bs2_4$
8	4	w_4	x_{43}	y_{43}	$bs1_4$	$bs2_4$

Il est bon de noter que la somme des poids d'enquête serait beaucoup plus élevée que le nombre total d'enfants dans la population longitudinale. La section qui suit porte sur les questions dont l'analyste doit tenir compte pour regrouper des données.

15.7 Regroupement des données

Les analystes qui regroupent les données de différents cycles doivent être conscients que l'analyse qui en découle peut être erronée, sauf si l'exercice est réalisé soigneusement. Les points dont il faut tenir compte sont les suivants :

- Il est possible que les données relatives à un enfant figurent plus d'une fois dans les données regroupées (mesures répétées).
- Pour éviter les mesures répétées, l'analyste doit regrouper des échantillons indépendants. Il est possible que l'analyste doive choisir une mesure dans plusieurs cycles ou qu'il établisse une nouvelle mesure combinée.
- Il peut être difficile de définir la population de référence représentée par l'échantillon regroupé et à propos de laquelle on fait des inférences.
- Il est possible que l'analyste doive calculer de nouveaux poids pour l'échantillon regroupé.

Ces points sont expliqués ci-dessous.

Si des échantillons dépendants sont regroupés et que les données relatives à certains enfants figurent plus d'une fois dans l'échantillon regroupé, la structure de corrélation doit être prise en compte dans les estimations de la variance d'échantillonnage afin de tester la validité des statistiques (l'analyste doit composer ici avec des mesures répétées).

Le regroupement des données de différents cycles de la cohorte initiale de l'ELNEJ entraînera la formation d'un échantillon regroupé dépendant et, par conséquent, l'existence de mesures répétées. Il peut s'agir par exemple du regroupement des données concernant les jeunes de 0 à 11 ans du cycle 1 avec celles des jeunes de 2 à 13 ans du cycle 2 (ce sont les mêmes enfants).

Il est bon de noter que le simple fait qu'un enfant ait un poids transversal plus élevé que zéro ne signifie pas qu'il n'est pas aussi suivi du point de vue longitudinal. Pour savoir si les données concernant un enfant figurent plus d'une fois dans l'échantillon regroupé, il suffit de vérifier si l'identificateur de l'enfant (variable PERSRUK) y figure plus d'une fois.

Si des échantillons indépendants sont regroupés, les données de l'enfant ne figureront qu'une seule fois dans le fichier. Par exemple, comme un nouvel échantillon indépendant d'enfants de 1 an est sélectionné à chaque cycle de l'ELNEJ, les échantillons d'enfants de 1 an peuvent facilement être regroupés. L'analyste peut alors tout simplement utiliser le poids transversal et les poids bootstrap pour chaque bébé (du cycle approprié), et la population de référence de l'analyse serait composée de tous les enfants de 1 an nés pendant les années de référence visées par les cycles regroupés.

Dans le cadre de cette méthode, chaque année de naissance est considérée comme une strate, et l'analyste peut facilement utiliser les poids d'enquête (aucun ajustement n'est requis). De plus, il n'y aurait qu'une seule mesure par enfant, et elle serait liée au cycle pendant lequel l'enfant avait 1 an.

Souvent, il existe de nombreuses façons de combiner les données de plusieurs cycles. Supposons par exemple que l'analyste s'intéresse aux données des enfants de 0 à 4 ans des cycles 1 et 2. Il pourrait tout simplement regrouper tous les enfants de 0 à 4 ans du cycle 1 et ceux du même groupe d'âge du cycle 2, auquel cas les données de certains enfants figureraient plus d'une fois puisque les enfants de 0 à 2 ans du cycle 1 sont les enfants de 2 à 4 ans du cycle 2. L'analyste devrait alors composer avec des mesures répétées.

Si l'analyste désire que les données de l'enfant ne figurent qu'une fois dans le fichier, il pourrait alors sélectionner tous les enfants de 0 à 4 ans du cycle 1 et les regrouper avec l'échantillon indépendant d'enfants de 0 à 1 an du cycle 2. Il pourrait ensuite utiliser les poids transversaux des enfants de 0 à 4 ans du cycle 1 et les poids transversaux des enfants de 0 à 1 an du cycle 2. La population de référence de l'analyse serait composée de tous les enfants qui avaient entre 0 et 1 an aux cycles 1 et 2 et de ceux qui avaient entre 2 et 4 ans au cycle 1.

Autrement, l'analyste pourrait calculer une nouvelle mesure combinant tous les renseignements des différents cycles puis regrouper les échantillons de façon indépendante, de sorte que les données de chaque enfant ne figurent qu'une seule fois dans le fichier final, avec la nouvelle mesure calculée.

Pour illustrer ces différentes possibilités, prenons l'exemple présenté plus tôt (enfants de 0 à 4 ans des cycles 1 et 2) et supposons que l'analyste s'intéresse à l'état de santé général de l'enfant (excellent, très bon, bon, passable, médiocre). Il s'agit de la variable AHLCQ01 au cycle 1 et de la variable BHLCCQ01 au cycle 2. En ce qui concerne les enfants de l'échantillon regroupé qui sont âgés de 0 à 2 ans au cycle 1 et de 2 à 4 ans au cycle 2, plusieurs choix s'offrent à l'analyste : les données de ces enfants pourraient figurer deux fois dans le fichier (mesures répétées), l'analyste pourrait utiliser les données du cycle 1 seulement ou du cycle 2 seulement, ou il pourrait encore calculer une mesure combinant l'information des deux cycles.

En combinant les données de différents cycles, l'analyste pourrait devoir recalculer de nouveaux poids pour le nouvel échantillon regroupé. Supposons, par exemple, que le chercheur s'intéresse aux données des enfants de 0 à 5 ans des cycles 2 et 3. Il peut choisir de regrouper les enfants de 0 à 5 ans du cycle 2 avec ceux de 0 à 5 ans du cycle 3 qui n'étaient pas dans le cycle 2. En ce qui a trait à cet échantillon regroupé, l'analyste pourrait utiliser les poids transversaux pour les enfants de 0 à 5 ans du cycle 2. Toutefois, comme il n'a choisi qu'un sous-ensemble des enfants du cycle 3, les poids du cycle 3 devront être ajustés. Ils pourraient par exemple être remis à l'échelle de sorte que pour le sous-échantillon d'enfants de 0 à 5 ans sélectionnés, les poids correspondent aux totaux connus de la population d'enfants de 0 à 5 ans au cycle 3.

Dans l'analyse des données provenant d'échantillons regroupés, il faut aussi garder en mémoire que le nombre d'enfants dans l'échantillon peut varier considérablement d'un cycle à l'autre pour un âge donné. Ainsi, on observe un échantillon inhabituellement important d'enfants de 5 ans au cycle 3.

15.8 Non-réponse

Comme toute enquête, l'ELNEJ est soumise à de la non-réponse. Il existe deux grandes catégories de non-réponse, la non-réponse totale et la non-réponse partielle.

La non-réponse totale est l'absence complète de données pour un enfant sélectionné et admissible, en raison de facteurs tels que le refus de participer à l'enquête ou l'impossibilité de retrouver l'enfant. La non-réponse totale est prise en considération et corrigée lors de la pondération. Veuillez consulter le chapitre 11.0 du présent guide pour obtenir plus de détails à ce sujet.

Il y a non-réponse partielle lorsque des questions précises sont laissées sans réponse. Il est bon de noter qu'un enfant ou un jeune ne doit pas avoir répondu à toutes les questions au cycle 8 afin d'être considéré comme un répondant. Veuillez consulter le chapitre 11.0 du présent guide pour obtenir plus d'exemples de non-réponses partielles. Certaines réponses sont imputées, mais la plupart ne le sont pas.

La non-réponse constitue une préoccupation pour les analystes étant donné qu'elle peut se solder par des estimations biaisées si elle n'est pas corrigée correctement. Or, les conclusions fondées sur des estimations biaisées peuvent être fausses. Dans l'analyse des données de l'ELNEJ, il est courant de se retrouver face à une non-réponse partielle concernant certaines variables d'intérêt. Voici quelques moyens de composer avec cette situation.

1. Ne conserver pour l'analyse principale que les enregistrements pour lesquels des renseignements complets sont disponibles et établir le profil des non-répondants dans une analyse secondaire distincte. Il est possible, par exemple, de signaler que le groupe de répondants est plus susceptible d'être établi dans certaines provinces ou d'avoir certains niveaux de revenu et que ces caractéristiques diffèrent de celles des non-répondants. Le fait de ne rien dissimuler relativement aux différences entre les deux groupes constitue une mise en garde pour les lecteurs qui prennent connaissance des résultats que la non-réponse a posé un problème dans les données et les aide à interpréter correctement les conclusions analytiques.
2. Déclarer la non-réponse partielle comme une catégorie. Cela signifie que la catégorie de non-réponse est considérée comme une catégorie valable dans les tableaux ou les modèles. Cette méthode convient particulièrement bien aux données catégoriques (la plupart des données de l'ELNEJ sont en fait des données catégoriques).
3. Répondérer les enregistrements comportant une réponse afin de tenir compte des répondants affichant une non-réponse partielle.
4. Ignorer les enregistrements affichant une non-réponse partielle, mais augmentez les poids des répondants afin de tenir compte des non-répondants. Cette stratégie se révèle particulièrement intéressante lorsqu'une composante entière de l'enquête est manquante (le questionnaire à remplir soi-même, par exemple). Toutefois, si seulement quelques données sont manquantes ici et là, ce n'est peut-être pas la meilleure stratégie puisqu'elle élimine tous les enregistrements comportant une non-réponse partielle, quelle qu'elle soit.
5. L'ajustement des poids des répondants n'est que l'inverse de la probabilité de réponse, que l'on assimile souvent au taux de réponse pondéré.
6. Il est bon de noter que, pour l'ELNEJ, les poids sont stratifiés a posteriori aux chiffres connus selon l'âge, le sexe et la province. Au moment de la repondération visant à tenir compte de la non-réponse partielle, ces totaux de contrôle ne sont plus respectés du tout. Il est possible de stratifier a posteriori de nouveau après l'ajustement pour la non-réponse ou, plutôt que de procéder à une nouvelle pondération, d'utiliser la technique de l'imputation (point suivant).
7. Imputer la non-réponse partielle (remplacer les valeurs manquantes par d'autres valeurs). L'avantage de cette méthode par rapport à une nouvelle pondération est que tous les enregistrements sont conservés, ce qui signifie qu'aucune donnée n'est rejetée. Seules les valeurs manquantes de chaque enregistrement sont

imputées. Il est important de citer le taux d'imputation avec les résultats analytiques et de donner des explications sur la stratégie d'imputation utilisée.

L'annexe II contient un exemple de la façon d'évaluer et de signaler la non-réponse partielle. Des documents de référence sont cités ci-dessous pour ceux qui désirent davantage de renseignements sur la non-réponse. La liste n'est aucunement exhaustive.

Aperçu de la non-réponse :

- « Méthodes et pratiques d'enquête » (Statistique Canada, 2003)
 - Différents chapitres portent sur des questions liées à la non-réponse.
 - La publication peut être commandée sur le site Web de Statistique Canada (<http://www.statcan.gc.ca>).
- *Sampling: Design and Analysis* (Lohr, S., Duxbury Press, 1999)
 - Le chapitre 11.0 porte exclusivement sur la non-réponse.

Non-réponse traitée plus en détail :

- *Incomplete Data in Sample Surveys* (Madow., W. G. et autres, New York, Academic Press, 1983).
- *Nonresponse in Household Surveys* (Groves, R. et M. Cooper, New York, Wiley, 1998).
- *Statistical Analysis with Missing Data*, deuxième édition (Little, R. J. A. et D. B. Rubin, New York, Wiley, 2002).

15.9 Autres sources d'erreurs non dues à l'échantillonnage

L'analyste ne doit pas perdre de vue les effets que les autres erreurs non dues à l'échantillonnage (en plus de la non-réponse) peuvent avoir sur son analyse. Ces erreurs, qui peuvent biaiser les résultats des analyses, sont notamment dues au sous-dénombrement de la population, aux erreurs de traitement et aux erreurs de réponse, entre autres. Veuillez consulter le chapitre 12.0 du présent guide pour en connaître davantage sur les erreurs non dues à l'échantillonnage.

15.10 Calcul de la variance avec certains logiciels

Les logiciels SAS ou SPSS, utilisés couramment par les analystes, peuvent calculer correctement les estimations en utilisant les poids de sondage. À l'exception de quelques procédures en SAS, ces logiciels ne sont pas programmés pour tenir compte du plan de sondage de l'ELNEJ lors du calcul de la variance d'échantillonnage d'une estimation (y compris le plan de sondage complexe, l'ajustement des poids pour la non-réponse et la stratification a posteriori). Il en résulte que plusieurs logiciels sous-estimeraient la variance d'échantillonnage de l'ELNEJ de celle-ci, qui peut être parfois importante. On encourage donc l'utilisateur à utiliser des poids bootstrap pour estimer la variance.

Certains logiciels sont capables d'utiliser les poids bootstrap tels que SUDAAN, WesVar et STATA9 lesquels calculent la variance basée sur le plan d'échantillonnage au moyen de la méthode de rééchantillonnage à répliques équilibrées (RRE). La création de poids RRE diffère en théorie de celle des poids bootstrap, mais l'estimateur de variance est le même, alors les poids bootstrap de l'ELNEJ peuvent être utilisés dans ces logiciels.

Il existe une version autonome de SUDAAN, ainsi qu'une version qui s'intègre au SAS. La dernière version a une flexibilité qui permet de faire appel aux procédures de SUDAAN à l'intérieur d'un programme SAS. À l'aide de ses neuf procédures, on est capable de calculer des estimations de moyennes, ratios et totaux, d'effectuer des tests d'indépendance, des régressions linéaires, log-linéaires, logistiques et de survie. SUDAAN peut faire la lecture de fichiers SAS, SPSS et d'autres formats communs.

Le logiciel WesVar présente une approche à l'analyse « pointer-cliquer », ce qui le rend très facile à apprendre. Les types d'analyses sont plus limités que dans SUDAAN, mais WesVar offre quand même le calcul des estimations de variance de moyennes, ratios et totaux, des tests d'indépendance et des régressions linéaires et logistiques. La lecture de fichiers SAS, SPSS et

d'autres formats communs est possible avec WesVar, mais le logiciel requiert une étape supplémentaire : créer un fichier de données WesVar pour effectuer l'analyse. Les poids bootstrap peuvent être utilisés avec d'autres logiciels qui offrent un environnement de programmation et les techniques d'analyse désirées. Des macros SAS et SPSS ont été conçues dans le cadre de l'ELNEJ afin d'utiliser les poids bootstrap pour arriver à des estimations de variance basées sur le plan d'échantillonnage.

Le lecteur trouvera des renseignements utiles sur l'utilisation des poids bootstrap avec SUDAAN et WesVar à l'adresse <http://www.statcan.gc.ca/pub/12-002-x/12-002-x2004002-fra.pdf>. Veuillez consulter le chapitre 13.0 du présent guide pour obtenir plus de renseignements sur l'estimation de la variance d'échantillonnage, y compris des précisions sur un autre outil appelé Bootvar.

15.11 Coefficients de variation pour des proportions

Il est d'usage d'utiliser le coefficient de variation (CV) pour déterminer la qualité d'une estimation, et ce pour différents types d'estimation tels que des totaux, des proportions ou autres. Cependant, lorsque le CV est utilisé pour déterminer la qualité d'une proportion estimée, l'analyste devra faire preuve d'une grande prudence. Le CV n'est pas dans tous les cas une bonne mesure de la qualité des proportions estimées. Plus de détails à ce sujet se trouvent dans le chapitre 13.0 du présent guide.

15.12 Écart type et erreur type

Les expressions « écart type » et « erreur type » sèment parfois la confusion. Veuillez consulter le chapitre 13.0 du présent guide sur le calcul de la variance pour obtenir des définitions claires de ces expressions et des explications sur la façon de les appliquer à l'ELNEJ.

15.13 Comprendre la différence entre « Non déclaré », « Ne sait pas », « Refus » et « Enchaînement valide »

Toutes les questions de l'ELNEJ ne s'appliquent pas à tous les enfants. Dans l'ELNEJ, une question qui ne s'applique pas à un enfant donné aura la réponse « Enchaînement valide ». Lorsqu'une question s'applique à un individu, mais qu'aucune réponse n'a été donnée, les valeurs « Non déclaré », « Ne sait pas » ou « Refus » seront présentes dans le fichier de données. Lorsque vous analysez une population en particulier, assurez-vous que les questions qui vous intéressent s'appliquent à votre population. De plus, en faisant l'analyse de la non-réponse, assurez-vous **de ne pas** traiter les cas « Enchaînement valide » comme étant des non-répondants. Dans ces cas, ce n'est pas qu'on n'a pas répondu à la question, c'est qu'elle ne s'applique pas. À l'occasion, « Enchaînement valide » peut avoir une signification précise comme « 0 » ou « Pas à l'école ». Consultez le questionnaire pour connaître les détails.

16.0 Lignes directrices concernant la totalisation, l'analyse et la diffusion des données

Ce chapitre du présent document décrit les lignes directrices que doivent suivre les utilisateurs qui totalisent, analysent, publient ou autrement diffusent des données tirées des fichiers de microdonnées de l'enquête. Ces lignes directrices devraient permettre aux utilisateurs de microdonnées d'arriver aux mêmes chiffres que Statistique Canada tout en étant en mesure de produire, conformément aux lignes directrices établies, des données qui n'ont pas été publiées.

16.1 Lignes directrices concernant l'arrondissement

Afin que les estimations destinées à la publication ou à toute autre forme de diffusion qui sont tirées des fichiers de microdonnées de l'Enquête longitudinale nationale sur les enfants et les jeunes (ELNEJ) correspondent à celles produites par Statistique Canada, on conseille vivement aux utilisateurs de se conformer aux lignes directrices suivantes en ce qui concerne leur arrondissement.

- a) Les estimations dans le corps principal d'un tableau statistique doivent être arrondies **à la centaine près** en utilisant la technique d'arrondissement normale. Suivant cette technique, si le premier ou le seul chiffre à supprimer se situe entre 0 et 4, on ne modifie pas le dernier chiffre à conserver. Si le premier ou le seul chiffre à supprimer se situe entre 5 et 9, on augmente de 1 le dernier chiffre à conserver. Par exemple, selon la technique de l'arrondissement normale à la centaine près, si les deux derniers chiffres se situent entre 00 et 49, on les remplace par 00 et on laisse inchangé le chiffre précédent (celui des centaines). Si les deux derniers chiffres se situent entre 50 et 99, on les remplace par 00 et on augmente de 1 le chiffre précédent. Ainsi, un total estimatif de 21 352 sera arrondi à 21 400.
- b) Les totaux partiels marginaux et les totaux marginaux des tableaux statistiques doivent être calculés à partir de leurs composantes non arrondies correspondantes, puis être arrondis à la centaine près en utilisant la technique d'arrondissement normale.
- c) Les moyennes, les proportions, les taux et les pourcentages doivent être calculés à partir de composantes non arrondies (c'est-à-dire des numérateurs ou des dénominateurs), puis **être arrondis à leur tour à une décimale** en utilisant la technique d'arrondissement normale.
- d) Les sommes et les différences d'agrégats (ou de rapports) doivent être calculées à partir de leurs composantes non arrondies correspondantes, puis être arrondies à leur tour à la centaine près (ou à la décimale près) en utilisant la technique d'arrondissement normale.
- e) Dans le cas, où, en raison de limitations d'ordre techniques ou de toutes autres limites, une technique d'arrondissement autre que la technique normale est utilisée produisant des estimations à être publiées ou à être diffusées qui diffèrent des estimations correspondantes publiées par Statistique Canada, on conseille vivement aux utilisateurs d'indiquer la raison pour ces différences dans le ou les documents à publier ou à diffuser.
- f) Les utilisateurs ne doivent en aucun cas publier ou autrement diffuser des estimations non arrondies. Des estimations non arrondies laissent supposer qu'elles sont plus précises qu'elles le sont en réalité.

16.2 Lignes directrices concernant la pondération de l'échantillon en vue de la totalisation

Dans une estimation provenant d'une enquête, chaque unité d'échantillonnage représente non seulement cette unité d'échantillonnage, mais aussi plusieurs autres unités au sein de la population observée. Dans le cas de l'ELNEJ, le poids d'enquête attribué à chaque enfant tient compte du nombre d'enfants que représente un répondant donné.

Lorsqu'il produit des estimations simples, y compris des tableaux statistiques ordinaires, l'utilisateur doit appliquer le poids d'échantillonnage approprié. Si l'on n'utilise pas les poids appropriés, on ne peut considérer les estimations calculées à partir du fichier de microdonnées comme étant représentatives de la population ciblée par l'enquête, ces estimations ne correspondraient pas à celles produites par Statistique Canada.

16.3 Lignes directrices concernant la modélisation statistique

Plan d'échantillonnage

Tel que mentionné précédemment, l'ELNEJ est basée sur un plan d'échantillonnage complexe comportant une stratification, de multiples degrés de sélection et des probabilités inégales de sélection des répondants. Pour obtenir des détails concernant le plan d'échantillonnage, veuillez consulter le chapitre 5.0 du présent guide. L'utilisation des données provenant d'enquêtes aussi complexes pose des problèmes aux analystes parce que le plan d'enquête et les probabilités de sélection influent sur les méthodes d'estimation et de calcul de variance qui doivent être utilisées. Pour que les estimations et les analyses de l'enquête ne soient pas entachées de biais, on doit utiliser les poids de l'enquête appropriés dans la mesure du possible. Veuillez consulter le chapitre 11.0 du présent guide pour obtenir plus de détails au sujet des poids de l'enquête.

Estimations de variance

Bien que de nombreuses méthodes d'analyse incluses dans les progiciels statistiques permettent d'utiliser des poids, la définition ou le sens de ces derniers peut différer de celui qui conviendrait à une enquête par sondage, de sorte que si les estimations produites grâce à ces progiciels sont exactes dans bien des cas, les variances calculées sont peu fiables. L'utilisateur des données devrait estimer des variances conformes au plan à l'aide des poids bootstrap et des outils décrits au chapitre 13.0.

16.4 Lignes directrices concernant la diffusion

L'utilisateur des données ne doit diffuser ni publier aucune estimation permettant de découvrir l'identité d'un répondant ou de divulguer les réponses de quelqu'un. C'est la raison pour laquelle on devrait se fonder sur les réponses d'au moins cinq répondants pour établir des estimations (p. ex. les cases dans un tableau croisé).

Abstraction faite de l'exigence énoncée ci-dessus, toutes les estimations peuvent être considérées comme pouvant faire l'objet d'une diffusion. Toutefois, avant de diffuser ou de publier une estimation de l'ELNEJ, l'utilisateur devrait d'abord déterminer la qualité de celle-ci (p. ex. la fiabilité de l'estimation), ce qui implique le calcul de l'erreur type qui y est associée (veuillez consulter le chapitre 13.0 du présent guide). L'utilisateur devrait également tenir compte de l'incidence que des erreurs non dues à l'échantillonnage peuvent avoir sur l'estimation (veuillez consulter le chapitre 11.0 du présent guide).

Une unité de mesure standard utilisée pour déterminer la qualité d'une estimation pour les données diffusées de l'ELNEJ est le coefficient de variation (veuillez consulter la section 13.2 du présent guide). Le tableau suivant décrit les lignes directrices relatives au niveau de qualité de l'estimation utilisant le coefficient de variation.

Lignes directrices relatives au niveau de qualité de l'estimation

Niveau de qualité de l'estimation	Lignes directrices
1) Acceptable	Ces estimations présentent de faibles coefficients de variation, de l'ordre de 0,0 à 16,5 %. Aucune mise en garde n'est requise.
2) Médiocre	Ces estimations présentent des coefficients de variation élevés, de l'ordre de 16,6 à 33,3 %. Ces estimations sont signalées par la lettre ^E . Elles sont aussi accompagnées d'une mise en garde avertissant les utilisateurs des niveaux élevés d'erreur associés aux estimations.
3) Inacceptable	Ces estimations présentent des coefficients de variation très élevés, supérieurs à 33,3 %. Ces estimations seront supprimées et désignées par la lettre « F » car elles sont d'une qualité inacceptable. Les conclusions qui reposeront sur ces estimations ne seront pas fiables et seront très probablement invalides.

En outre, il convient de déterminer le nombre d'enfants sur lequel est fondé le calcul de l'estimation. Si ce nombre est petit, on peut prévoir que l'erreur type associée à l'estimation pondérée sera probablement grande, et il est probable que l'estimation ne soit pas fiable dans la plupart des cas. Proportionnellement parlant, on peut certes inférer, à partir des données de l'enquête, qu'une certaine caractéristique est rare, mais il est habituellement impossible de déterminer le taux véritable en s'appuyant sur ces données. Par exemple, on peut se fonder sur l'ELNEJ pour estimer à 1 sur 1 500 la proportion d'enfants ayant un problème de santé particulier, mais le taux véritable peut équivaloir au double ou à la moitié de cette estimation. Dans ces cas-là, il est préférable de diffuser une estimation moins exacte (p. ex. on estime le taux à moins de 0,5 %) et plus conforme au niveau de précision auquel l'enquête peut légitimement prétendre.

16.5 Modélisation des données de l'ELNEJ et des poids bootstrap

L'ELNEJ est fondée sur un plan d'échantillonnage complexe (voir l'aperçu à la section 5.0 du Guide de l'utilisateur de l'ELNEJ). Par conséquent, certaines des méthodes d'échantillonnage utilisées entraînent une dépendance quelconque en ce qui concerne la sélection des unités (comme la stratification et/ou la mise en grappes, c.-à-d. la création de groupes à partir desquels les unités sont sélectionnées). La complexité du plan d'échantillonnage, comme celui qui sert à l'ELNEJ, influence largement notre façon de calculer la variance associée à une estimation.

L'utilisation de poids normalisés ne suffit pas à elle seule à estimer la variance totale adéquatement; les poids normalisés ne peuvent pas tenir compte des dépendances en ce qui concerne la sélection des unités.

Pour tenir compte des conséquences du plan d'échantillonnage sur la variance, nous recommandons aux utilisateurs la méthode bootstrap, un processus d'estimation de la variance fondée sur le rééchantillonnage.

Les poids bootstrap sont indiqués dans un fichier distinct et peuvent être fusionnés avec le fichier de données au moyen de la variable PERSRUK. Tout ce qui reste à faire pour les utilisateurs consiste à répéter l'analyse qu'ils viennent d'effectuer au moyen de poids diffusés, en utilisant cette fois-ci chacun des poids bootstrap à tour de rôle. Les utilisateurs peuvent tirer profit des programmes Bootvar, offerts en SAS et en SPSS à cette fin¹³. Certains progiciels peuvent également exécuter

13. Bootvar est accessible par l'entremise du lien *Outils de recherche* à www.statcan.gc.ca/rdc-cdr/bootvar_sas-fra.htm ou à www.statcan.gc.ca/rdc-cdr/bootvar_spss-fra.htm.

cette dernière étape pour l'utilisateur, notamment SUDAAN (au moyen de l'option BRR de SUDAAN).

L'utilisation de la méthode bootstrap améliore l'exactitude de l'estimation de la variance en tenant compte des diverses composantes d'un plan d'échantillonnage complexe dans l'estimation de la variance totale. Cet élément est particulièrement important si des résultats importants se rapprochent du seuil de rejet. Le fait de ne pas tenir compte du plan d'échantillonnage entraînera des estimations de la variance inexactes, ce qui entacherait les statistiques des tests et qui pourrait donner lieu à l'établissement inopportun d'une signification statistique inexistante.

Annexe 1 — Lignes directrices pour les chercheurs et les analystes de l'Enquête longitudinale nationale sur les enfants et les jeunes

Des analystes et des chercheurs qui utilisent la base de données de l'Enquête longitudinale nationale sur les enfants et les jeunes (ELNEJ) ont exprimé le besoin d'avoir des lignes directrices qui leur permettraient de planifier leurs analyses et de présenter leurs résultats. Le but du présent guide est de répondre à ces besoins.

Le guide de l'utilisateur de l'ELNEJ comprend deux sections principales. Dans la première section, on traite des différents éléments méthodologiques dont il faut tenir compte lors de la soumission de projets de recherche qui utilisent les données de l'ELNEJ. Dans la deuxième section, on fait des recommandations sur ce qui doit être considéré lorsque les données de l'ELNEJ sont utilisées pour écrire un article ou un résumé. Bon nombre des éléments qui se trouvent dans la section sur la préparation d'une proposition de projets de recherche se retrouvent également dans la section sur les éléments à considérer lors de la rédaction d'articles. Ces deux composantes peuvent donc être utilisées comme deux entités distinctes, au besoin.

I. Avant de soumettre un projet de recherche pour révision

Considérations méthodologiques

Avant d'entreprendre toute analyse avec les données de l'ELNEJ, les chercheurs et analystes devraient tout d'abord se familiariser avec la complexité de l'enquête et les conséquences qui en découlent. Le but du présent document est de faciliter leur travail en indiquant clairement les éléments essentiels méthodologiques qui doivent être pris en compte lorsqu'ils utilisent les données de l'ELNEJ.

Ce document indique les éléments essentiels méthodologiques qui doivent être pris en compte lorsqu'ils soumettent une proposition de recherche. Un document d'accompagnement spécifie tous les éléments qui doivent être considérés lorsqu'ils soumettent un document pour être passé en revue. On encourage fortement les auteurs à utiliser ces documents afin de s'assurer qu'ils ont couvert tous les éléments appropriés avant de soumettre leur proposition de recherche ou leur document.

Les données de l'ELNEJ peuvent être utilisées à différentes fins. L'objectif principal de l'ELNEJ est de pouvoir produire, à partir d'un échantillon probabiliste, des inférences au niveau de la population. C'est avec cet objectif en tête que le présent document a été préparé. Lorsque des données de l'ELNEJ sont utilisées à d'autres fins que pour faire des inférences, certains des éléments présentés dans ce document peuvent ne pas s'appliquer. Par contre, dans ces cas, des indications claires à ce sujet doivent figurer dans l'article afin d'encadrer les énoncés de l'analyse pour le lecteur.

Les éléments du cadre analytique

Il y a six éléments principaux à considérer lors de la préparation d'une proposition de recherche ou d'un document en utilisant les données de l'ELNEJ. Ces éléments sont :

- 1) Les sources de données;
- 2) Les facteurs inhérents à l'analyse;
- 3) Les variables;
- 4) Le type d'analyse;
- 5) L'estimation de la variance; et
- 6) Les méthodes analytiques.

1) Sources de données

Toutes les sources de données utilisées dans l'analyse peuvent être spécifiées de la façon suivante :

- a Spécifiez la source principale de données à être utilisée pour l'analyse.
 - ELNEJ
 - Autre (spécifiez).
- b Indiquez quelles autres sources de données seront utilisées dans l'analyse, le cas échéant, et si ces données seront présentées sous forme de données brutes ou en forme de tableaux.
- c Si l'analyse se limite à un sous-groupe ou à un domaine, donnez une description du sous-groupe ou du domaine, p. ex. les groupes d'âge, les provinces, les variables avec certaines caractéristiques, entre autres.
- d Identifiez le ou les cycle(s) de données de l'ELNEJ qui seront utilisés.

2) Facteurs inhérents à l'analyse

La proposition de recherche devrait inclure une description des facteurs qui peuvent limiter ou affecter l'analyse :

- a Description de la population cible :
 - Fournissez une définition claire de la population cible de l'ELNEJ.
 - Si votre population cible diffère de celle de l'ELNEJ, vous devez inclure un énoncé à cet effet indiquant l'incidence possible de cette différence sur votre analyse.
 - Si d'autres sources de données sont utilisées pour fins de comparaison, vous devez inclure un énoncé décrivant comment leurs populations cibles diffèrent de la population de l'ELNEJ.
- b Traitement de la non-réponse :
 - Si des variables que vous prévoyez utiliser comprennent de la non-réponse, veuillez inclure un énoncé indiquant le niveau de non-réponse de ces variables et l'incidence possible de cette non-réponse sur votre analyse.
 - Identifiez de quelle façon la non-réponse partielle sera traitée :
 - Imputation
 - Repondération
 - Rapportée comme une valeur
 - Ignorée, analyse faite uniquement avec les répondants.
 - Une analyse des caractéristiques des non-répondants par rapport aux répondants devrait être faite afin d'identifier les biais possibles.
- c Restrictions avec les données de l'ELNEJ :
 - Fournissez les tailles d'échantillon, globalement et pour tous les sous-domaines, lorsque l'information est disponible. Les tailles d'échantillon doivent être suffisantes pour fins de confidentialité et pour produire des estimations fiables.
 - S'il y a lieu, indiquez si vous prévoyez d'autres restrictions à l'utilisation des données de l'ELNEJ dans votre projet.

3) Variables

- Fournissez une liste préliminaire des variables du fichier de l'ELNEJ qui seront utilisées dans votre analyse.
- Indiquez les variables prédictives et les variables de résultats à être considérées, si elles sont connues.

Notez que de l'information détaillée sur les variables peut être obtenue avant même d'accéder aux fichiers maîtres, en utilisant les questionnaires disponibles (sur le site Internet de Statistique Canada) ou en examinant la liste de variables dans les fichiers de données (par l'entremise de l'initiative de démocratisation des données dans les bibliothèques universitaires).

4) Type d'analyse

- a Indiquez le type d'analyse que vous prévoyez faire, c'est-à-dire longitudinal, transversal ou les deux. Notez que si les deux types d'analyse sont inclus dans votre analyse, la population cible peut différer d'un type à l'autre.
- b Spécifiez les poids de sondage — transversaux et/ou longitudinaux — que vous utiliserez dans votre analyse. Notez que si des estimations des populations transversales et longitudinales sont analysées, assurez-vous d'utiliser les bons poids selon l'analyse que vous effectuez.
- c Si les poids de sondage ne sont pas utilisés, vous devez inclure une note justificatrice à cet effet. Notez qu'il est très difficile de justifier la non utilisation des poids de sondage dans l'analyse.

5) Estimation de la variance

Différentes méthodes sont disponibles pour calculer la précision des estimations quand on doit faire des inférences comme le calcul de la variance et/ou de coefficients de variation (CV). La soumission pour le projet de recherche devrait inclure, si possible, des indications sur les méthodes qui seront utilisées pour calculer la variance et le CV résultant. Les options sont les suivantes :

- Les approximations utilisant les tables d'approximation de variabilité échantillonnale (disponibles pour les 4 premiers cycles).
- L'utilisation de l'interface Excel pour l'ELNEJ qui contient des CV pour une série de domaines d'intérêt.
- L'utilisation des poids bootstrap avec le programme Bootvar, SUDAAN ou autres programmes qui peuvent utiliser les poids bootstrap.
- La variance et le coefficient de variation non calculés. Note : ceci signifie qu'aucune inférence statistique ne peut être faite.
- L'utilisation d'un autre logiciel (précisez le logiciel : _____).

6) Méthodes analytiques

- a Présentez une description des méthodes analytiques que vous prévoyez utiliser.
- b Décrivez les méthodes statistiques que vous utiliserez pour déterminer si les estimations sont statistiquement différentes.
- c Prévoyez inclure les intervalles de confiance de vos estimations, basés sur un calcul approprié de la variance.

II. Avant de soumettre un article pour révision

Considérations méthodologiques

Avant d'entreprendre toute analyse avec les données de l'ELNEJ, les chercheurs et analystes devraient tout d'abord se familiariser avec la complexité de l'enquête et les implications qui en découlent. Le but de ce document est de faciliter leur travail en leur exposant clairement les éléments méthodologiques clés lorsqu'ils veulent utiliser les données de l'ELNEJ.

Ce document identifie les éléments méthodologiques essentiels qui doivent être considérés lors de la soumission d'un document pour révision. Un document d'accompagnement spécifie tous les éléments qui doivent être considérés lorsqu'ils soumettent une proposition de recherche. Les auteurs sont encouragés à utiliser ces documents afin de s'assurer qu'ils ont couvert tous les éléments appropriés avant de soumettre leur proposition de recherche ou leur document.

Les données de l'ELNEJ peuvent être utilisées à différentes fins. L'objectif principal de l'ELNEJ est de pouvoir produire, à partir d'un échantillon probabiliste, des inférences au niveau de la population. C'est avec cet objectif en tête que ce document a été développé. Dans les cas où les données de l'ELNEJ sont utilisées à d'autres fins que pour faire des inférences, certains des éléments présentés dans ce document peuvent ne pas s'appliquer. Par contre, dans ces cas, des indications claires à ce sujet doivent être présentées dans l'article afin d'encadrer les énoncés de l'analyse pour le lecteur.

Les éléments du cadre analytique

Il y a six éléments principaux à considérer lors de la préparation d'une proposition de recherche ou d'un document en utilisant les données de l'ELNEJ. Ces éléments sont :

- 1) Sources des données;
- 2) Facteurs inhérents à l'analyse;
- 3) Variables;
- 4) Type d'analyse;
- 5) Estimation de la variance; et
- 6) Méthodes analytiques.

1) Sources des données

Toutes les sources de données utilisées dans l'analyse peuvent être spécifiées de la façon suivante :

- a Spécifiez la source principale de données utilisée pour l'analyse.
 - ELNEJ
 - Autre (spécifiez).
- b Indiquez quelles autres sources de données, si applicable, ont été utilisées et si ces données sont sous forme de données brutes ou bien en forme de tableaux.
- c Si l'analyse se limite à un sous-groupe ou un domaine, donnez une description du sous-groupe ou du domaine, p. ex., les groupes d'âge, les provinces, les variables avec certaines caractéristiques, etc.
- d Identifiez le ou les cycle(s) de données de l'ELNEJ qui ont été utilisés.

2) Facteurs inhérents à l'analyse

L'article devrait inclure une description des facteurs qui peuvent limiter ou affecter l'analyse :

- a Description de la population cible :
 - Fournir une définition claire de la population cible de l'ELNEJ.
 - Si votre population cible diffère de celle de l'ELNEJ, inclure un énoncé à cet effet indiquant l'impact potentiel sur votre analyse de cette différence.

- Si d'autres sources de données sont utilisées pour fins de comparaison, inclure un énoncé décrivant comment leurs populations cibles diffèrent de la population de l'ELNEJ.
- b Traitement de la non-réponse (s'il y a lieu) :
 - Si des variables que vous avez utilisées présentent de la non-réponse, inclure un énoncé indiquant le niveau de non-réponse de ces variables et l'impact de cette non-réponse sur votre analyse.
 - Identifiez de quelle façon la non-réponse partielle a été traitée :
 - Imputation;
 - Repondération;
 - Déclarée comme une valeur;
 - Ignorée, analyse faite uniquement avec les répondants.
 - Analyse des caractéristiques des non-répondants par rapport aux répondants devrait être faite afin d'identifier les biais possibles.
- c Restrictions des données :
 - Fournissez les tailles d'échantillon, globalement et pour tous les sous-domaines.
 - Assurez-vous que les tailles d'échantillon sont suffisantes pour fins de confidentialité et pour produire des estimations fiables.
 - S'il y a lieu, indiquez s'il y a d'autres restrictions à l'utilisation des données de l'ELNEJ dans votre projet (p. ex. les variables utilisées).

3) Variables

- Décrivez les variables du fichier de l'ELNEJ qui ont été utilisées dans votre analyse.

4) Type d'analyse

- a Indiquez le type d'analyse que vous avez fait, c'est-à-dire longitudinal, transversal ou les deux. Notez que si les deux types d'analyse sont inclus dans votre analyse, la population cible peut différer d'un type à l'autre.
- b Spécifiez les poids de sondage, longitudinaux, transversaux ou les deux, qui ont été utilisés dans votre analyse. Notez que si des estimations pour des populations transversales et longitudinales ont été déclarées, vous devez utiliser les poids appropriés dans chaque analyse.
- c Si les poids de sondage ne sont pas utilisés, vous devez inclure une note justificatrice à cet effet. Notez qu'il est très difficile de justifier la non utilisation des poids de sondage dans l'analyse.

5) Estimation de la variance

Décrivez la méthode que vous avez utilisée pour calculer la précision de l'estimation lors de faire des inférences, ainsi que le calcul suivant des variances et/ou des coefficients de variation (CV) :

- Approximations utilisant les tables d'approximation de variabilité échantillonnale (disponibles pour les 4 premiers cycles).
- Utilisation de l'interface Excel pour l'ELNEJ qui contient des CV pour une série de domaines d'intérêt.
- Utilisation des poids bootstrap avec le programme Bootvar, SUDAAN ou autres programmes qui utilisent les poids bootstrap.
- Variance et CV non calculés. Nota : ceci signifie qu'aucune inférence statistique ne peut être faite dans cet article.
- Utilisation d'un autre logiciel (précisez le logiciel : _____).

6) Méthodes analytiques

- a Présentez une description des méthodes analytiques que vous avez utilisées.
- b Décrivez les méthodes statistiques que vous avez utilisées pour déterminer si les estimations sont statistiquement différentes.
- c Veuillez inclure les intervalles de confiance de vos estimations, basés sur un calcul approprié de la variance.

Conclusion

Un réviseur de votre article, qui a accès aux mêmes données que vous, devrait être en mesure de reproduire parfaitement vos résultats et vos conclusions, si la méthodologie que vous avez utilisée est correcte et bien expliquée dans votre article.

III. Exemples de comment citer les produits de l'ELNEJ

Comment citer le fichier principal de l'ELNEJ (utilisateurs ayant accès aux données à distance) :

STATISTIQUE CANADA. 2006-2007. *Enquête longitudinale nationale sur les enfants et les jeunes : ELNEJ2006_C7_LONG_fichier principal*.

Comment citer un fichier principal de l'ELNEJ (consultation dans un centre de données de recherche) :

STATISTIQUE CANADA. 2006-2007. *Enquête longitudinale nationale sur les enfants et les jeunes : ELNEJ2006_C7_ECD_fichier principal*, Centre de données de recherche de l'Université de l'Alberta. Date et mois de diffusion 2008. http://www.statcan.gc.ca/cgi-bin/imdb/p2SV_f.pl?Function=getSurvey&SDDS=4450&lang=en&db=imdb&dbq=f&adm=8&dis=2 (site consulté le 9 novembre, 2010).

Comment citer le Guide de l'utilisateur de l'ELNEJ :

STATISTIQUE CANADA. s.d. *Guide de l'utilisateur des microdonnées, Enquête longitudinale nationale sur les enfants et les jeunes, cycle 7, septembre 2006 à juillet 2007*. XX p. http://www.statcan.gc.ca/imdb-bmdi/document/4450_D4_T9_V7-fra.pdf (site consulté le 9 novembre, 2010).

Pour obtenir d'autres renseignements :

STATISTIQUE CANADA. 2006. *Comment citer les produits de Statistique Canada*, produit n° 12-591-XWF au catalogue de Statistique Canada, version mise à jour le 31 mars, Ottawa, Ontario, <http://www.statcan.gc.ca/pub/12-591-x/12-591-x2006001-fra.htm> (site consulté le 9 novembre, 2010).

Annexe 2 — Analyse de la non-réponse partielle

Tel qu'il est décrit à la section 12.4 du présent guide, les analystes qui utilisent les données de l'ELNEJ doivent être conscients de l'incidence de la non-réponse partielle sur les données qu'ils tentent d'analyser. La non-réponse partielle peut être plus fréquente dans le cas des répondants qui présentent une caractéristique particulière. Cela pourrait entraîner un biais qui pourrait semer le doute quant aux résultats de l'analyse. Des techniques telles que la repondération et l'imputation permettent de remédier à la non-réponse partielle. On encourage les utilisateurs de données à évaluer, au minimum, le degré de la non-réponse à des composantes ou à des questions dans leurs résultats et exposer ces résultats dans leur rapport. Un exemple d'une telle analyse est présenté ci-dessous.

L'analyse de la non-réponse partielle qui suit, laquelle est fondée sur les résultats du cycle 5 de l'Enquête longitudinale nationale sur les enfants et les jeunes (ELNEJ), a été effectuée afin d'accompagner la diffusion intitulée « Enquête longitudinale nationale sur les enfants et les jeunes : milieu familial, revenu et comportement des enfants » parue dans *Le Quotidien*, le 21 février 2005.

Comme toute enquête, l'ELNEJ doit composer avec la non-réponse, laquelle se divise en deux grandes catégories : la non-réponse totale et la non-réponse partielle. La non-réponse totale se caractérise par l'absence complète de données, ou trop peu de données pour que le cas puisse être déclaré répondant, pour une unité échantillonnée. Les poids déterminés joints aux fichiers de données ont été ajustés afin de tenir compte de la non-réponse totale.

La non-réponse partielle se caractérise par l'absence d'information pour certaines questions seulement, la personne sélectionnée ayant toutefois répondu à suffisamment de questions pour être déclarée répondante. L'objectif de ce document est d'évaluer la non-réponse partielle pour les variables de l'ELNEJ qui ont été utilisées dans le rapport noté plus haut.

Le tableau suivant présente le pourcentage de répondants âgés de 2 à 5 ans du cycle 1 en 1994-1995 et de 10 à 13 ans du cycle 5 en 2002-2003, pour lesquels les données sont disponibles pour chaque variable prédictive et de résultats à l'étude dans le rapport.

En 1994-1995, toutes les réponses ont été données par les parents déclarants. Dans l'ensemble, les taux de réponse pour ces variables étaient très élevés, se situant entre 96 % et 98 % pour les variables prédictives (fonctionnement de la famille, dépression de la mère, pratiques parentales répressives), et entre 87 % et 94 % pour les variables de résultats (comportement agressif de l'enfant, anxiété chez l'enfant, comportement prosocial de l'enfant).

En 2002-2003, les réponses relatives à deux variables prédictives ont été fournies par les parents, et les taux de réponse globaux étaient encore une fois très élevés, se situant à 96 % pour le fonctionnement de la famille et à 95 % pour la dépression de la mère. Les réponses restantes ont été données par les enfants eux-mêmes, à l'aide de questionnaires autoadministrés. Les taux de réponse étaient plus bas que pour les renseignements fournis par les parents, se situant entre 74 % et 78 % pour les variables prédictives (pratiques parentales répressives, soins prodigués par les parents, surveillance parentale) et entre 76 % et 81 % pour les variables des résultats axées sur le comportement (comportement agressif, anxiété, comportement prosocial, estime de soi).

Afin de déterminer les sources possibles de biais de non-réponse dans les données, les taux de réponse ont été comparés entre les garçons et les filles pour les ménages à faible revenu et à revenu élevé, et pour les cinq régions du Canada. Les résultats de ces analyses sont les suivants :

- Aucune différence des taux de réponse selon le sexe n'a été constatée pour l'une ou l'autre des variables.
- On a constaté une différence des taux de réponse entre les répondants à faible revenu et les répondants à revenu élevé. Le taux de réponse était plus faible pour les répondants à faible revenu que pour les répondants à revenu élevé pour ce qui est de la variable des soins

prodigués par les parents (68 % comparativement à 75 %). Quoique non déterminante, cette différence est statistiquement significative ($p < 0,05$).

- Des différences régionales du taux de réponse ont été enregistrées pour l'anxiété chez l'enfant en 1994-1995, pour les pratiques parentales répressives en 1994-1995 et pour la dépression de la mère en 2002-2003. Les répondants de la région des Prairies ont enregistré un taux de réponse beaucoup moins élevé pour l'anxiété chez l'enfant en 1994-1995 que le taux enregistré dans la région de l'Atlantique (92 % comparativement à 97 %). Aucune autre différence statistiquement significative n'a été enregistrée pour cette variable. Le taux de réponse des répondants de la région des Prairies était beaucoup plus bas pour les pratiques parentales répressives en 1994-1995 que celui des répondants de la région de l'Atlantique et de ceux du Québec (94 % comparativement à 99 % et 98 % respectivement). Aucune autre différence statistiquement significative n'a été enregistrée pour cette variable. Enfin, les répondants de l'Ontario ont eu un taux de réponse beaucoup moins élevé en ce qui a trait à la dépression de la mère en 2002-2003 que celui des répondants de la région de l'Atlantique (93 % comparativement à 98 %). Aucune autre différence statistiquement significative n'a été enregistrée pour cette variable.

Aucune modification particulière n'a été apportée aux analyses de ces variations du taux de non-réponse et les résultats devraient être interprétés en gardant ces limites à l'esprit.

Taux de réponse pour les échelles 1994-1995 (de 2 à 5 ans) et 2002-2003 (de 10 à 13 ans) selon le sexe, la situation de revenu et la région de résidence

Revenu et la région de résidence										
Mesure des résultats	Taux de réponse global	Sexe		Situation de revenu		Région de résidence en 2002				
		Fille	Garçon	Faible	Élevé	Atlantique	Québec	Ontario	Prairies	Colombie- Britannique
%										
Score pour le comportement agressif – 1994-1995	92	93	92	91	93	95	93	92	92	92
Score pour le comportement agressif – 2002-2003	77	78	76	74	78	78	78	77	75	77
Score pour l'anxiété – 1994-1995 ¹	94	94	94	93	94	97	95	93	92	94
Score pour l'anxiété – 2002-2003	76	77	76	75	77	77	78	77	74	76
Score pour le comportement prosocial – 1994-1995	87	86	88	86	88	91	88	87	88	86
Score pour le comportement prosocial – 2002-2003	76	76	76	75	76	76	78	77	73	75
Score pour l'estime de soi – 2002-2003	81	81	82	79	82	83	82	82	78	84
Score pour le fonctionnement de la famille – 1994-1995	98	98	98	97	98	99	99	98	97	98
Score pour le fonctionnement de la famille – 2002-2003	96	96	96	94	96	98	96	95	97	96
Score pour la dépression de la mère – 1994-1995	98	98	98	99	98	98	97	99	97	98
Score pour la dépression de la mère – 2002-2003 ²	95	95	95	92	96	98	96	93	97	97
Score pour les pratiques parentales répressives – 1994-1995 ³	96	96	96	95	97	99	98	96	94	95
Score pour les pratiques parentales répressives – 2002-2003	78	79	78	74	79	81	77	80	76	80
Score pour les soins prodigués par les parents – 2002-2003 ⁴	74	74	73	68	75	74	76	73	71	76
Score pour la surveillance des parents – 2002-2003	76	77	75	70	77	77	77	76	72	77

Notes

Les données en caractères gras représentent les différences statistiquement significatives au $p < 0,05$.

1. Taux de réponse global moins élevé pour la région des Prairies que pour la région de l'Atlantique.

2. Taux de réponse global moins élevé pour l'Ontario que pour la région de l'Atlantique.

3. Taux de réponse global moins élevé pour la région des Prairies que pour les régions de l'Atlantique et le Québec.

4. Taux de réponse global moins élevé pour les répondants à faible revenu que pour les répondants à revenu élevé.

Source : Statistique Canada, Enquête longitudinale nationale sur les enfants et les jeunes.

Annexe 3 — Concordance entre les noms de variables du traitement et les noms de variables de la diffusion, pour les questionnaires autoadministrés du cycle 8 de l'ELNEJ

Notes :

1. Les cellules vierges = sans objet
2. Dans la colonne intitulée « Type », C = caractère, N = numérique

Variable	Longueur	Type	Livret 22	Livret 23
PERSRUK	14	C	PERSRUK	PERSRUK
HMMCQ01	3	N	EFFAGEYR	EFFAGEYR
HMMCbQ1A	3	N	AGE	AGE
HMMCQ02	1	C	GENDER	GENDER
HFFCQ01	1	N	A1	A1
HFFCQ02	1	N	A2	A2
HFFCQ03	1	N	A3	A3
HFFCQ04	1	N	A4	A4
HFFCcQ4A	1	N	A5	A5
HFFCQ05	2	N	A6	A6
HFFCbQ13	2	N	A7B	A7B
HFFCbQ14	2	N	A8B	A8B
HFFCbQ15	1	N	A9	A9
HFFCQ07	1	N	A13	A13
HFFCQ08A	1	N	A14A	A14A
HFFCQ08B	1	N	A14B	A14B
HFFCQ08C	1	N	A14C	A14C
HFFCQ08D	1	N	A14D	A14D
HFFCQ08E	1	N	A14E	A14E
HFFCQ08F	1	N	A14F	A14F
HFFCQ08G	1	N	A14G	A14G
HFFCQ08H	1	N	A14H	A14H
HFFCc08I	1	N	A14I	A14I
HFFCQ08K	1	N	A14J	A14J
HFFCQ08L	1	N	A14K	A14K
HFFCQ08M	1	N	A14L	A14L
HFFCQ08N	1	N	A14M	A14M
HFFCd12A	1	N	A15	A15
HDRCdQ05	1	N	A10A	A10A
HDRCdQ09	1	N	A10B	A10B
HFFCd16C	1	N	A10C	A10C
HFFCc16D	1	N	A10D	A10D
HFFCc16E	1	N	A10E	A10E
HFFCc19A	1	N	A11A	A11A
HFFCc19B	1	N	A11B	A11B
HFFCc19C	1	N	A11C	A11C
HFFCc19D	1	N	A11D	A11D
HFFCc20A	1	N	A12A	A12A

Variable	Longueur	Type	Livret 22	Livret 23
HFFCc20B	1	N	A12B	A12B
HFFCc20C	1	N	A12C	A12C
HSCCQ01	1	N	B1	
HSCCbQ18	1	N	B2	
HSCCb19A	1	N	B3A	
HSCCb19B	1	N	B3B	
HSCCb19C	1	N	B3C	
HSCCb19D	1	N	B3D	
HSCCc19F	1	N	B3E	
HSCCb19E	1	N	B3F	
HSCCb20A	1	N	B4A	
HSCCb20B	1	N	B4B	
HSCCb20C	1	N	B4C	
HSCCb20E	1	N	B4D	
HSCCc20H	1	N	B4E	
HSCCc20I	1	N	B4F	
HSCCc20J	1	N	B4G	
HSCCb20G	1	N	B4H	
HSCCQ02	1	N	B5	
HSCCcQ3a	1	N	B6A	
HSCCcQ03	1	N	B6B	
HSCCcQ3b	1	N	B6C	
HSCCcQ3D	1	N	B6D	
HSCCcQ3C	1	N	B6E	
HSCCcQ3E	1	N	B6F	
HSCCcQ3F	1	N	B6G	
HSCCd3G	1	N	B6H	
HSCCc21A	1	N	B7A	
HSCCc21B	1	N	B7D	
HSCCc21C	1	N	B7B	
HSCCc21D	1	N	B7C	
HSCCc21E	1	N	B7E	
HSCCc21F	1	N	B7F	
HSCCcQ26	1	N	B8	
HSCCcQ27	1	N	B9	
HSCCcQ10	1	N	B10	B7
HSCCbQ22	1	N	B12A	
HSCCc22A	1	N	B12B	
HSCCcQ28	1	N	B13	
HSCCcQ29	1	N	B14	
HSCCQ12	1	N	B15A	
HSCCQ11	2	N	B15B	
HSCCcQ16	2	N	B15C	
HSCCQ17	2	N	B15D	
HSCCcQ30	2	N	B16	
HSCCQ13	2	N	B17A	

Variable	Longueur	Type	Livret 22	Livret 23
HSCCQ14	1	N	B17B	
HSCCQ15	1	N	B17C	
HSCCcQ24	2	N	B18	
HSCCc31A	1	N	B11A	
HSCCc31B	1	N	B11B	
HSCCc31C	1	N	B11C	
HSCCc31D	1	N	B11D	
HSCCc31E	1	N	B11E	
HAMCQ01A	1	N	C1A	B1A
HAMCQ01B	1	N	C1B	B1B
HAMCQ01C	1	N	C1C	B1C
HAMCQ01D	1	N	C1D	B1D
HAMCcQ1E	1	N	C1E	B1E
HAMCe25A	1	N	C2A	B2A
HAMCe25B	1	N	C2B	B2B
HAMCe25C	1	N	C2C	B2C
HAMCe25D	1	N	C2D	B2D
HAMCe25E	1	N	C2E	B2E
HAMCe25F	1	N	C2F	B2F
HAMCe25G	1	N	C2G	B2G
HAMCe25H	1	N	C2H	B2H
HAMCe25I	1	N	C2I	B2I
HAMCe25J	1	N	C2J	B2J
HAMCe25K	1	N	C2K	B2K
HAMCe25L	1	N	C2L	B2L
HAMCe25M	1	N	C2M	B2M
HAMCe25N	1	N	C2N	B2N
HAMCe25O	1	N	C2O	B2O
HAMCbQ02	1	N	C3	B3
HAMCcQ03	1	N	C4	B4
HAMCcQ6A	1	N	C7A	
HAMCcQ6B	1	N	C7B	
HAMCdQ6C	1	N		B8A
HAMCcQ7A	1	N	C8A	
HAMCcQ7B	1	N	C8B	
HAMCdQ7C	1	N		B8B
HAMCcQ8A	1	N	C9A	
HAMCcQ8B	1	N	C9B	
HAMCdQ8C	1	N		B8C
HAMCdQ4A	1	N	C5A	
HAMCcQ4B	1	N	C5B	
HAMCcQ4C	1	N	C5C	
HAMCcQ4D	1	N	C5D	
HAMCeQ4E	1	N	C5E	
HAMCe26A	1	N		B5A
HAMCe26B	1	N		B5B

Variable	Longueur	Type	Livret 22	Livret 23
HAMCe26C	1	N		B5C
HAMCe26D	1	N		B5D
HAMCe26E	1	N		B5E
HAMCe26F	1	N		B5F
HAMCcQ5A	1	N	C6A	B6A
HAMCcQ5B	1	N	C6B	B6B
HAMCcQ5C	1	N	C6C	B6C
HAMCcQ5D	1	N	C6D	B6D
HAMCdQ09	1	N		B9
HAMCdQ10	1	N		B10
HFBCQ01A	1	N	D1A	
HFBCc01B	1	N	D1B	
HFBCQ01C	1	N	D1C	
HFBCQ01D	1	N	D1D	
HFBCQ01E	1	N	D1E	
HFBCQ01F	1	N	D1F	
HFBCQ01G	1	N	D1G	
HFBCQ01H	1	N	D1H	
HFBCQ01I	1	N	D1I	
HFBCQ01J	1	N	D1J	
HFBCQ01K	1	N	D1K	
HFBCQ01L	1	N	D1L	
HFBCQ01M	1	N	D1M	
HFBCQ01P	1	N	D1N	
HFBCQ01Q	1	N	D1O	
HFBCQ01R	1	N	D1P	
HFBCQ01S	1	N	D1Q	
HFBCQ01T	1	N	D1R	
HFBCQ01U	1	N	D1S	
HFBCc01V	1	N	D1T	
HFBCQ01W	1	N	D1U	
HFBCd01X	1	N	D1V	
HFBCQ01Z	1	N	D1W	
HFBCQ1AA	1	N	D1X	
HFBCQ1BB	1	N	D1Y	
HFBCQ1CC	1	N	D1Z	
HFBCQ1DD	1	N	D1AA	
HFBCQ1FF	1	N	D1BB	
HFBCQ1GG	1	N	D1CC	
HFBCQ1HH	1	N	D1EE	
HFBCQ1JJ	1	N	D1DD	
HFBCQ1LL	1	N	D1FF	
HFBCQ1MM	1	N	D1GG	
HFBCQ1NN	1	N	D1HH	
HFBCQ1OO	1	N	D1II	
HFBCQ1PP	1	N	D1JJ	

Variable	Longueur	Type	Livret 22	Livret 23
HFBCQ1QQ	1	N	D1KK	
HFBCQ1RR	1	N	D1LL	
HFBCQ1SS	1	N	D1MM	
HFBCQ1TT	1	N	D1NN	
HFBCc1UU	1	N	D1OO	
HFBCd10A	1	N		C1A
HFBCd10B	1	N		C1B
HFBCd10C	1	N		C1C
HFBCd10D	1	N		C1D
HFBCd10E	1	N		C1E
HFBCd10F	1	N		C1F
HFBCd10G	1	N		C1G
HFBCd10H	1	N		C1H
HFBCd10I	1	N		C1I
HFBCd10J	1	N		C1J
HFBCd10K	1	N		C1K
HFBCd10L	1	N		C1L
HFBCcQ04	1	N	D2	C2
HFBCcQ4A	1	N	D3	C3
HFBCcQ05	1	N	D4	C4
HFBCcQ07	1	N	D5	C5
HFBCcQ08	1	N	D6	C6
HFBCbQ2B	1	N	D7A	C7A
HFBCbQ2E	1	N	D7B	C7B
HFBCbQ2F	1	N	D7C	C7C
HFBCcQ2H	1	N		C7D
HFBCeQ2O	1	N	D7D	C7E
HFBCbQ2P	1	N	D7E	C7F
HFBCdQ2Z	1	N		C7G
HFBCe2ZZ	1	N	D7F	C7H
HFBCbQ2Y	1	N	D7G	C7I
HFBCe2BB	1	N	D7H	C7J
HFBCcQ3A	1	N	D8	C8
HATCc1AA	1	N	E1A	
HATCc1BB	1	N	E1B	
HATCc1CC	1	N	E1C	
HATCc1DD	1	N	E1D	
HATCc1EE	1	N	E1E	
HATCc1II	1	N	E1F	
HATCeQ20	2	N	E2	
HATCbQ04	1	N	E3	
HATCdQ07	2	N	E7	
HATCeQ12	1	N	E8A	
HATCe12B	1	N	E8B	
HATCe12C	1	N	E8C	
HATCeQ13	1	N	E9A	

Variable	Longueur	Type	Livret 22	Livret 23
HATCe13B	1	N	E9B	
HATCe13C	1	N	E9C	
HATCeQ21	2	N	E10	
HATCeQ22	1	N	E11	
HATCdQ10	2	N	E12	
HATCdQ5A	1	N	E5A	
HATCbQ5B	1	N	E5B	
HATCbQ5C	1	N	E5C	
HATCbQ5D	1	N	E5D	
HATCbQ5E	1	N	E5E	
HATCbQ5F	1	N	E5F	
HATCbQ5G	1	N	E5G	
HATCdQ05	1	N	E6	
HATCc14A	1	N	E4A	
HATCc14B	1	N	E4B	
HATCc14C	1	N	E4C	
HATCc14D	1	N	E4D	
HDRCdQ01	2	N	F1	D1
HDRCdQ04	2	N	F2	D2
HDRCdQ03	2	N	F3B	
HDRCdQ6A	2	N	F4	D3
HDRCdQ07	2	N	F5	
HDRCdQ9A	1	N	F6	
HDRCbQ9B	2	N	F7	
HDRCdQ9C	2	N	F8	D4
HDRCdQ15	2	N	F9	D5
HDRCd14A	2	N	F11AB	
HDRCc14C	2	N	F11BB	
HDRCc14B	2	N	F11CB	
HDRCd14F	2	N	F11DB	
HDRCc14D	2	N	F11EB	
HDRCc18A	2	N	F10A	D6A
HDRCd18B	2	N	F10B	D6B
HDRCd18C	2	N	F10C	D6C
HDRCc18D	2	N	F10D	D6D
HDRCdQ19	1	N		D7
HDRCdQ20	1	N		D8
HPMCdQ04	2	N	G1	G1
HPMCcQ5A	1	N	G2A	G2A
HPMCcQ5B	1	N	G2B	G2B
HPMCcQ5C	1	N	G2C	G2C
HPMCcQ06	1	N	G3	G3
HPMCdQ6A	2	N		G4A
HPMCdQ6B	2	N		G4B
HPMCdQ6C	1	N		G5A
HPMCdQ6D	1	N		G5B

Variable	Longueur	Type	Livret 22	Livret 23
HPMCdQ6E	1	N		G5C
HPMCdQ6F	1	N		G5D
HPMCdQ6G	1	N		G5E
HPMCdQ6H	1	N		G5F
HPMCdQ6I	1	N		G5G
HPMCdQ6J	1	N		G5H
HPMCdQ6K	1	N		G5I
HPMCdQ6L	1	N		G5J
HPMCdQ07	2	N	G4	G6
HPMCcQ8A	1	N	G5A	G7A
HPMCcQ8B	1	N	G5B	G7B
HPMCcQ8C	1	N	G5C	G7C
HPMCcQ09	1	N	G6	G8
HPMCdQ9A	2	N		G9A
HPMCdQ9B	2	N		G9B
HPMCdQ9C	1	N		G10A
HPMCdQ9D	1	N		G10B
HPMCdQ9E	1	N		G10C
HPMCdQ9F	1	N		G10D
HPMCdQ9G	1	N		G10E
HPMCdQ9H	1	N		G10F
HPMCdQ9I	1	N		G10G
HPMCdQ9J	1	N		G10H
HPMCdQ9K	1	N		G10I
HPMCdQ9L	1	N		G10J
HPMCcQ10	1	N	G7	G12
HPMCd11A	2	N	G8	G13
HPMCc11B	2	N	G9	G14
HPMCcQ1A	1	N	G10A	
HPMCcQ1B	1	N	G10B	
HPMCcQ1C	1	N	G10C	
HPMCcQ1D	1	N	G10D	
HPMCcQ1E	1	N	G10E	
HPMCcQ1F	1	N	G10F	G11A
HPMCcQ1G	1	N	G10G	
HPMCcQ1H	1	N	G10H	
HPMCcQ1I	1	N	G10I	
HPMCcQ1J	1	N	G10J	
HPMCcQ1R	1	N	G10K	
HPMCcQ1K	1	N	G10L	
HPMCcQ1L	1	N	G10M	
HPMCcQ1M	1	N	G10N	
HPMCcQ1N	1	N	G10O	
HPMCcQ1O	1	N	G10P	
HPMCcQ1P	1	N	G10Q	
HPMCcQ1Q	1	N	G10R	

Variable	Longueur	Type	Livret 22	Livret 23
HPMCdQ1S	1	N	G10S	
HPMCdQ1T	1	N	G10T	G11B
HPMCdQ9U	1	N		G11C
HPMCdQ9V	1	N		G11D
HPMCd12A	1	N		G15
HPMCd12B	1	N		G16
HPMCd12C	1	N		G17A
HPMCd12D	1	N		G17B
HPMCd12E	1	N		G17C
HPMCd12F	1	N		G17D
HPMCd12G	1	N		G17E
HPMCd12H	1	N		G17F
HPMCd12I	1	N		G17G
HPMCd12J	1	N		G17H
HPMCd12K	1	N		G17I
HPMCd12L	1	N		G17J
HHTCcQ03	1	N	H1	
HHTCbQ3A	1	N	H4A	
HHTCbQ3B	1	N	H4B	
HHTCbQ3C	1	N	H4C	
HHTCbQ3G	1	N	H4D	
HHTCdQ06	1	N	H5	
HHTCeQ5A	1	N	H6	E8
HHTCeQ5B	1	N		E8AA
HHTCeQ5C	1	N		E8AB
HHTCeQ5D	1	N		E8AC
HHTCeQ5E	1	N		E8AD
HHTCdQ5F	1	N		E8AE
HHTCdQ5G	1	N		E8BA
HHTCdQ5H	1	N		E8BB
HHTCdQ5I	1	N		E8BC
HHTCdQ5J	1	N		E8BD
HHTCeQ5K	1	N		E8CA
HHTCeQ5L	1	N		E8CB
HHTCeQ5M	1	N		E8CC
HHTCeQ5N	1	N		E8CD
HHTCeQ5O	1	N		E8CE
HPUCQ01	1	N	H7	E3
HPUCQ02	1	N	H8	E4
HPUCcQ3A	3	N	H9	E5
HPUCdQ04	1	N	H10	E6
HPUCQ05	1	N	H11	E7
HDACcQ1A	2	N	H12B	F1B
HDACcQ1B	1	N	H13	F2
HDACcQ1C	1	N	H14	F3
HDACcQ02	2	N	H15	F4

Variable	Longueur	Type	Livret 22	Livret 23
HDACcQ2A	1	N	H16	F5
HDACcQ05	1	N	H17	F6
HDACdQ23	1	N		F9
HDACeQ26	1	N		F11
HDACeQ25	2	N		F12
HDACcQ06	2	N	H18	F7
HDACcQ07	2	N	H19A	F8
HDACcQ08	1	N	H20	
HDACcQ09	1	N	H21	
HDACd22A	1	N		F10A
HDACd22B	1	N		F10B
HDACd22C	1	N		F10C
HDACd22D	1	N		F10D
HDACd22E	1	N		F10E
HDACd22F	1	N		F10F
HDACd22G	1	N		F10G
HDACd22H	1	N		F10H
HWKcQ4A	1	N	I1A	
HWKcQ4B	1	N	I1B	
HWKcQ4C	1	N	I1C	
HWKcQ4D	1	N	I1D	
HWKCe06A	1	N	I2A	
HWKCe06B	1	N	I2B	
HWKCe06C	1	N	I2C	
HWKCe06D	1	N	I2D	
HWKCe06E	1	N	I2E	
HWKCe06F	1	N	I2F	
HWKcQ9A	2	N	I3AA	
HWKcQ9B	2	N	I3BB	
HWKCeQ11	1	N	I4	
HWKcd12A	1	N	I5A	
HWKcd12B	1	N	I5B	
HWKcd12C	1	N	I5C	
HWKcd12D	1	N	I5D	
HWKcd13A	1	N	I6A	
HWKcd13B	1	N	I6B	
HWKcd13C	1	N	I6C	
HWKcd13D	1	N	I6D	
HWKcd13E	1	N	I6E	
HWKcd13F	1	N	I6F	
HWKcd13G	1	N	I6G	
HWKcd13H	1	N	I6H	
HWKcd13I	1	N	I6I	
HHTCbQ01	4,2	N	HEIGHT	HEIGHT
HHTCbQ02	7,3	N	WEIGHT	WEIGHT
HBMleS01	7,3	N	SELFBMI	SELFBMI

Variable	Longueur	Type	Livret 22	Livret 23
HBMleS02	2	N	SELFBMI2	SELFBMI2
HBMleS03	1	N	SELFBMI3	SELFBMI3
HEQYeS06	2	N	RA	RA
HEQYeS07	2	N	ER	ER
HEQYeS08	2	N	SM	SM
HEQYeS09	2	N	AD	AD
HEQYeS10	2	N	GM	GM
HEQYeS04	2	N	EQ4	EQ4
HEQYeS05	2	N	EQ5	EQ5
HFFcS01	2	N	FRFS01	FRFS01
HAMcS02	2	N	ABMS01	ABMS01
HFBcS01	2	N	BEHS01	
HFBCdS02	2	N	BEHS02	
HFBcS03	2	N	BEHS03	
HFBCdS04	2	N	BEHS04	
HFBcS05	2	N	BEHS05	
HFBcS07	2	N	BEHS07	
HPMCcS1	2	N	PARS01	
HPMCbS2B	2	N	PARS02	
HPMCcS3	2	N	PARS03	
HPMCdS4	2	N		PARS04
HPMCdS5	2	N		PARS05
HHTCbS1b	2	N		DEPS01
HWTCw01L	12,4	N		
HWTCwD1L	12,4	N		

Source : Statistique Canada, Enquête longitudinale nationale sur les enfants et les jeunes.