



Guide de l'utilisateur des microdonnées

Enquête longitudinale nationale sur les enfants et les jeunes

Cycle 6

Septembre 2004 à juin 2005



Statistique
Canada

Statistics
Canada

Canada

Table des matières

1.0	Introduction	9
2.0	Contexte	11
3.0	Objectifs	13
4.0	Principaux changements – Du cycle 5 au cycle 6	15
4.1	Changements apportés au contenu	15
4.2	Changement relatif à la période de collecte	15
4.3	Changements méthodologiques	15
4.3.1	Échantillon	15
4.3.2	Poids	16
4.3.3	Variance	16
4.4	Changements concernant le traitement des données	16
4.5	Changements concernant les fichiers diffusés	16
5.0	Méthodologie de l'enquête – Échantillon	19
5.1	Aperçu du plan d'échantillonnage – Cycles 1 à 6	19
5.2	L'Enquête sur la population active	20
5.2.1	Population cible	20
5.2.2	Stratification	21
5.2.3	Délimitation et sélection des grappes	22
5.2.4	Sélection des logements	22
5.2.5	Renouvellement de l'échantillon	23
5.2.6	Membres des ménages admissibles aux fins de l'Enquête sur la population active	23
5.3	Précisions concernant l'échantillon de l'Enquête longitudinale nationale sur les enfants et les jeunes	24
5.3.1	Cohorte initiale – Cycles 1 à 6	24
5.3.2	Cohortes du Développement de la petite enfance (DPE) présentes au cycle 6	26
5.4	Populations de référence longitudinale et transversales pour l'établissement des poids de l'Enquête longitudinale nationale sur les enfants et les jeunes	30
5.4.1	Populations longitudinales des cohortes	31
5.4.2	Les cohortes et leurs populations transversales	32
6.0	Collecte des données	35
6.1	Composante « Ménage »	35
6.2	Composante « Enfant »	35
6.3	Composante « Adulte »	36
6.4	Composante du jeune	36
6.5	Tailles d'échantillon au cycle 6	36
6.6	Mesures directes	38
6.7	Questionnaires auto administrés de 10 à 17 ans	38
6.7.1	Personnel de collecte (formation, supervision et contrôle)	39
7.0	Traitement des données	41
7.1	La 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 de liens	43
7.6	Vérifications préliminaires	43
7.7	Vérification de cheminement	43
7.8	Codage des questions ouvertes	44

7.9	Vérification de la cohérence	44
7.10	Indicateurs d'imputation	44
7.11	Établissement de variables dérivées et de 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	Variabes démographiques	54
8.3	Questionnaire de l'adulte	54
8.4	Questionnaire de l'enfant	56
8.5	Questionnaire du jeune (16 à 21 ans)	61
8.6	Questionnaires auto administrés (10 à 17ans)	65
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 exprimées selon une échelle	69
9.2	Analyse factorielle	70
9.2.1	Analyse factorielle pour les échelles	70
9.2.2	Transformation des données au moyen de l'échelonnement optimal	71
9.2.3	Analyse factorielle axée sur les données pondérées	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 de facteurs	71
9.3.3	Poids factoriel négatif	72
9.3.4	Code de non-réponse	72
9.3.5	Items bruts	72
9.4	Mesures de fiabilité des échelles	73
9.4.1	Coefficient alpha de Cronbach	73
9.4.2	Interprétation du coefficient alpha de Cronbach	73
9.4.3	Qu'est-ce qu'un niveau satisfaisant de fiabilité?	73
9.5	Échelles relatives au questionnaire du parent	74
9.5.1	Échelle de dépression	74
9.5.2	Échelle des responsabilités familiales	74
9.5.3	Échelle de fonctionnement de la famille	75
9.5.4	Échelle de sécurité du quartier	76
9.5.5	Échelle du soutien social	77
9.5.6	Échelle de comportement	78
9.5.7	Échelle du développement moteur et social	81
9.5.8	Échelles du rôle parental	83
9.5.9	Étapes du développement	85
9.6	Échelles relatives au questionnaire des jeunes	86
9.6.1	Échelle de dépression	86
9.6.2	Échelle de structure du quartier	86
9.6.3	L'échelle relative à l'estime globale de soi	87
9.6.4	Quotient émotionnel	87
9.6.5	Échelle du soutien social	92
9.6.6	Échelle des amis	92
9.6.7	L'échelle mes parents et moi	93
9.6.8	Échelles de résolution des conflits	93
9.6.9	Échelle de comportement	94

9.7	Résumé – Échelles du cycle 6-----	95
9.7.1	Échelles relatives au questionnaire du parent -----	95
9.7.2	Échelles relatives au questionnaire des enfants/jeunes (auto administré)---	96
9.7.3	Échelles relatives au questionnaire des jeunes (auto administré) -----	96
10.0	Imputation -----	97
10.1	Imputation du revenu du ménage -----	97
10.2	Imputation du revenu des jeunes – 16 ou 17 ans -----	100
10.3	Imputation du revenu des jeunes – 18 ans et plus-----	102
10.4	Imputation de l'échelle de Développement moteur et social (DMS) -----	103
11.0	Pondération et traitement de la non-réponse -----	105
11.1	Poids disponibles au cycle 6 -----	105
11.1.1	Poids entonnoir (variable FWTCdW1L) -----	105
11.1.2	Poids longitudinaux (variable FWTCW01L) -----	106
11.1.3	Poids transversaux (variable FWTCW01C) -----	107
11.2	Méthode de pondération -----	107
11.2.1	Pondération de l'Enquête longitudinale nationale sur les enfants et des jeunes -----	107
11.2.2	Premier ajustement : ajustement pour la non-réponse -----	108
11.2.3	Second ajustement : stratification a posteriori -----	108
11.2.4	Différences entre la méthode de pondération du cycle 6 et celle des cycles précédents -----	109
11.3	Application de la méthode de pondération-----	109
11.3.1	Pondération longitudinale -----	109
11.3.2	Pondération entonnoir -----	110
11.3.3	Pondération transversale-----	111
12.0	Qualité des données, taux de réponse et couverture-----	113
12.1	Erreur d'échantillonnage -----	113
12.2	Erreurs non dues à l'échantillonnage -----	113
12.3	Non-réponse totale et biais de non-réponse -----	113
12.3.1	Définitions de la réponse-----	114
12.3.2	Taux de réponse transversaux -----	116
12.3.3	Taux d'attrition longitudinaux pour la cohorte du cycle 1 -----	118
12.3.4	Taux de réponse longitudinal pour les enfants sélectionnés au cycle 4-----	121
12.3.5	Taux de réponse longitudinal pour les enfants sélectionnés au cycle 5-----	121
12.4	Non-réponse partielle -----	121
12.4.1	Composante de l'enfant -----	122
12.4.2	Composante de la personne la mieux renseignée-----	123
12.4.3	Composante du conjoint-----	123
12.4.4	Composante du jeune -----	124
12.4.5	EVIP-----	124
12.4.6	Connaissance des nombres-----	124
12.4.7	Qui suis-je? -----	125
12.4.8	Tests de mathématiques -----	125
12.4.9	Résolution de problèmes-----	125
12.4.10	Évaluation de la littératie -----	125
12.4.11	Évaluation de la numératie -----	126
12.4.12	Questionnaires auto administrés -----	126
12.5	Non-réponse de cycle-----	126
12.6	Incidence des erreurs de réponse pour les caractéristiques rares-----	127
12.7	Erreurs de réponse liées à des comportements déviants ou à des questions de nature délicate -----	127
12.8	Erreurs de réponse dues à des approximations-----	128
12.9	Erreurs de réponse dues à des erreurs de mémoire -----	128
12.10	Erreurs de réponse dues à la collecte par personne interposée -----	128

12.11	Schémas de réponse : catégories de réponse mal définies	129
12.12	Langue de l'interview	129
12.13	Renseignements contradictoires	129
12.14	Qualité des données de l'indice de masse corporelle	130
12.14.1	Indice de masse corporelle	130
12.14.2	Indice de masse corporelle – Centers for Disease Control	130
12.14.3	Indice de masse corporelle – Tim Cole et coll.	131
12.14.4	Indice de masse corporelle – Qualité des données	131
12.15	Biais d'accoutumance	132
12.16	Personne la mieux renseignée	132
12.17	Couverture	133
12.17.1	Couverture de l'Enquête longitudinale nationale sur les enfants et les jeunes	133
12.17.2	Questions liées à la couverture et à la base de sondage	134
12.18	Conclusion	138
13.0	Estimation de la variance	139
13.1	Quelques emplois des termes erreur d'échantillonnage et variance	140
13.2	Coefficient de variation	141
13.3	À propos de l'importance de rapporter la variance due à l'échantillonnage	144
13.4	Calcul de la variance due à l'échantillonnage	144
13.4.1	Chiffrier de variances approximatives pour des proportions	146
13.4.2	Macros SAS et SPSS pour calculer la variance due à l'échantillonnage à partir des poids bootstrap : Bootvar	146
13.4.3	Autres outils informatiques	147
13.4.4	Linéarisation de Taylor et autres techniques	147
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)	158
15.0	Enquête sur les enfants du Nord, cycle 6	163
15.1	Introduction	163
15.2	Contexte	163
15.3	Établissement de l'Enquête auprès des enfants du Nord	164
16.0	Enjeux analytiques	169
16.1	Incidence de l'utilisation d'un plan de sondage complexe sur l'analyse	169
16.2	L'unité d'analyse	170
16.3	Analyse longitudinale et analyse transversale	170
16.4	Estimations pondérées simples (totaux, moyennes et proportions)	171
16.4.1	Estimation d'un total de la population	171
16.4.2	Estimation d'une moyenne de population	172
16.4.3	Estimation d'une proportion de la population	172
16.4.4	Estimation de domaines de la population	172
16.5	Poids normalisés	173
16.6	Mesures répétées	176
16.7	Regroupement des données	177
16.8	Non-réponse	179
16.9	Autres sources d'erreurs non dues à l'échantillonnage	180
16.10	Calcul de la variance avec certains logiciels	180
16.11	Coefficients de variation pour des proportions	181
16.12	Écart-type et erreur-type	181

16.13	Comprendre la différence entre « Non déclaré », « Ne sait pas », « Refus » et « Enchaînement valide »-----	181
17.0	Lignes directrices concernant la totalisation, l'analyse et la diffusion des données -----	183
17.1	Lignes directrices concernant l'arrondissement -----	183
17.2	Lignes directrices concernant la pondération de l'échantillon en vue de la totalisation -----	183
17.3	Lignes directrices concernant la modélisation statistique -----	184
17.4	Lignes directrices concernant la diffusion -----	184
Annexe I	-----	185
Annexe II	-----	191
Annexe III	-----	195
Annexe IV	-----	205

1.0 Introduction

Le cycle 6 de l'Enquête longitudinale nationale sur les enfants et les jeunes (ELNEJ) a été mené entre septembre 2004 et juin 2005 par Statistique Canada en collaboration avec Développement social Canada (maintenant sous le nom de Ressources humaines et Développement social Canada).

Le présent manuel a été produit pour faciliter la manipulation des fichiers de microdonnées sur les résultats de l'enquête et pour décrire les questions de qualité de données et autres questions analytiques concernant l'ELNEJ.

Toute question au sujet de l'ensemble de données ou de son utilisation devrait être adressée à :

Statistique Canada

Services à la clientèle

Division des enquêtes spéciales

Téléphone : (613) 951-3321 ou appelez sans frais : 1 800 461-9050

Télécopieur : (613) 951-4527

Courriel : des@statcan.ca

Division des enquêtes spéciales, Statistique Canada

150, promenade du Pré Tunney

Ottawa (Ontario) K1A 0T6

Ressources humaines et Développement social Canada

Pierre Turcotte

Directeur adjoint par intérim

Direction de l'analyse stratégique, des partenariats et de la diffusion

Ressources humaines et Développement social Canada

Téléphone : (613) 957-7483

Courriel : pierre.c.turcotte@sdsc.gc.ca

2.0 Contexte

L'Enquête longitudinale nationale sur les enfants et les jeunes (ELNEJ) est une étude à long terme qui 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 a débuté en 1994 et elle est menée par Statistique Canada et est parrainée par Ressources humaines et Développement social Canada.

L'enquête a été conçue pour 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 les conséquences de ces facteurs sur leur développement au fil du temps.

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

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

Population de l'enquête

Au cycle 6, un échantillon représentatif d'enfants canadiens âgés de 0 à 5 ans de chacune des provinces a été suivi afin de produire des estimations longitudinales et transversales. La cohorte longitudinale de jeunes âgés de 10 à 21 ans a été suivie afin de produire des estimations longitudinales.

La cohorte du nord des enfants est un recensement des enfants âgés de 5 ans qui étaient inscrits à la maternelle au Yukon et au Nunavut. Veuillez consulter le chapitre 15.0 du présent guide pour de plus amples renseignements au sujet du contexte de l'Enquête auprès des enfants du Nord ainsi qu'une description des enfants qui ont été ciblés au cycle 6.

Population cible

Les objectifs de l'ELNEJ sont de produire à la fois des estimations longitudinales et transversales. Par conséquent, plusieurs populations cibles sont visées par l'échantillon du cycle 6. Veuillez consulter le chapitre 5.0 du présent guide pour de plus amples informations au sujet de l'échantillon.

- Transversalement, l'échantillon du cycle 6 représente les enfants de 0 à 5 ans au 31 décembre 2004.
- Longitudinalement, nous avons trois cohortes (plus d'un cycle de données) :
 - 1) Une première cohorte représente les enfants de 0 à 11 ans au 31 décembre 1994 et qui habitaient dans une des 10 provinces pendant la collecte pour le cycle 1 en 1994-1995; ces enfants sont maintenant âgés de 10 à 21 ans au cycle 6.
 - 2) Une deuxième cohorte représente les enfants âgés de 0 à 1 an au 31 décembre 2000 et qui habitaient dans une des 10 provinces pendant la collecte pour le cycle 4 en 2000-2001; ces enfants sont maintenant âgés de 4 à 5 ans au cycle 6.
 - 3) Une troisième cohorte représente les enfants âgés de 0 à 1 an au 31 décembre 2002 et qui habitaient dans une des 10 provinces pendant la collecte pour le cycle 5 en 2002-2003; ces enfants sont maintenant âgés de 2 à 3 ans au cycle 6.
- Transversalement, le recensement des enfants qui demeurent dans les territoires est limité, au cycle 6, aux enfants âgés de 5 ans qui étaient inscrits à la maternelle au Yukon et au Nunavut.

Cycles de collecte

La collecte de données se fait 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

Stratégie de diffusion des données

Les données du cycle 6 sont diffusées en quatre fichiers : la cohorte longitudinale – enfants (10 à 17 ans), la cohorte longitudinale – jeunes (18 à 21 ans), la cohorte du développement de la petite enfance (de la naissance à 5 ans) et la cohorte du Nord (5 ans).

- Cohorte longitudinale – enfants : il s'agit des enfants de la cohorte longitudinale initiale. Ces enfants étaient âgés de 0 à 7 ans au premier cycle de l'ELNEJ. Au cycle 6, ils étaient âgés de 10 à 17 ans.
- Cohorte longitudinale – jeunes : ce sont les enfants de la cohorte longitudinale initiale qui avaient 6 à 11 ans au cycle 1 et 16 à 21 ans au cycle 6. Le fichier contient des données concernant les jeunes et les ménages dans lesquels ils vivent ainsi que des variables reportées des cycles précédents et ne changeant pas au fil du temps (le poids à la naissance, par exemple). Le chapitre 8.0 fournit de plus amples renseignements sur le contenu des divers questionnaires.
- Le fichier de la cohorte du développement de la petite enfance contient les données recueillies auprès de la personne la mieux renseignée au sujet de l'enfant âgé de 0 à 5 ans au cycle 6.
- La cohorte du Nord comprend les enfants de 5 ans du Yukon et du Nunavut. Initialement, les enfants des trois territoires devaient prendre part à l'enquête, mais des contraintes opérationnelles ont empêché les Territoires du Nord-Ouest de participer au cycle 6. Pour de plus amples renseignements au sujet de l'Enquête auprès des enfants du nord, veuillez consulter le chapitre 17.0.

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;
- mettre cette information au service de 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), tests de calcul, etc.).

4.0 Principaux changements – Du cycle 5 au cycle 6

Le présent chapitre expose les principaux changements apportés à l'enquête depuis le cycle 5. Des explications plus précises se trouvent dans les différents chapitres.

4.1 Changements apportés au contenu

Des changements sont apportés au contenu de l'Enquête longitudinale nationale sur les enfants et les jeunes (ELNEJ) à chaque cycle. Le cinquième caractère du nom d'une nouvelle variable ou d'une variable modifiée (libellé, catégories de réponse, population admissible, etc.) est un « f ». 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 6.

- Le livret 24, qui contenait le questionnaire à remplir par les jeunes de 18 à 19 ans eux-mêmes, a été retiré. Les questions du livret figurent maintenant dans le questionnaire de l'application d'interviews assistées par ordinateur (IAO). Le tableau de concordance à l'annexe IV indique les noms des variables des cycles 5 et 6.
- Les répondants les plus âgés de la cohorte longitudinale ont maintenant 20 et 21 ans. Bon nombre de questions s'adressant à ce groupe d'âge sont les mêmes que celles posées au cycle 5, mais il y a aussi de nouvelles questions sur le financement des études postsecondaires (FEDYf165, FEDYf166, FEDYf167, FEDYf168), sur la qualité du travail (FLYYf14A, FLYYf14B, FLYYf14C), sur les enfants à charge (FDMCfD22) et sur l'exercice du droit de vote (FACYfQ16).
- Deux évaluations directes ont été ajoutées au cycle 6. Une évaluation de la littératie chez les jeunes de 18 à 19 ans (FLIYfS01) et une évaluation de la numératie chez les jeunes de 20 à 21 ans (FNUYfS01). Ces tests sont décrits au chapitre 14.0.
- La composante de l'éducation (questionnaire de l'enseignant de la maternelle) a été retirée.
- Des questions ont été ajoutées pour permettre au répondant de déclarer volontairement qu'il est autochtone (FSDCfQ3A, FSDCfQ3BA, FSDCfQ3BB, FSDCfQ3BC).

4.2 Changement relatif à la période de collecte

La répartition de l'échantillon entre les différentes vagues de la collecte a une incidence sur l'âge en mois des enfants au moment de l'interview, ce qui peut avoir des répercussions sur les scores obtenus aux tests et à autres mesures directes. Certains enfants ont été interviewés au cycle 6 à des périodes différentes qu'au cycle 5. Par exemple, les enfants de 10 ans ont été interviewés à la première vague au cycle 5 et à la deuxième vague au cycle 6. Les périodes de collecte des données du cycle 6 sont décrites au chapitre 6.0. Les variables FMMCbQ1A et FMCCdQ1B donnent l'âge de l'enfant au moment de la collecte.

4.3 Changements méthodologiques

4.3.1 Échantillon

Il n'y a eu aucun échantillon complémentaire d'enfants de 2 à 5 ans au cycle 5. En revanche, un échantillon complémentaire d'enfants de 2 à 5 ans a été prélevé au cycle 6 dans les provinces autre que le Québec et l'Ontario. Les répondants

de 2 à 5 ans faisant partie de l'échantillon complémentaire ont un poids transversal mais aucun poids longitudinal.

Il y a eu un déplacement dans la répartition de l'échantillon selon l'âge du cycle 5 au cycle 6 en ce qui concerne la cohorte du développement de la petite enfance, puisque l'importante cohorte introduite au cycle 3 a été retirée au cycle 6. À noter que ces enfants seront suivis de nouveau au cycle 7, alors qu'ils auront 8 et 9 ans. Voir le chapitre 5 pour plus de précisions à ce sujet.

4.3.2 Poids

L'ajustement pour la non-réponse diffère des cycles précédents. Pour plus de précisions, voir le chapitre 11.0, Pondération et traitement de la non-réponse.

4.3.3 Variance

Une partie de l'échantillon de l'ELNEJ a été prélevée après le remaniement de l'Enquête sur la population active (nouvelle formation de strates et de grappes), mis en œuvre graduellement en commençant par le groupe de renouvellement de novembre 2004. Cela influe sur la façon dont les poids bootstrap ont été créés, mais ne devrait pas avoir de répercussions pour les utilisateurs des données.

Les méthodes utilisées pour calculer les poids bootstrap ont été modifiées pour respecter la stratégie de pondération au cycle 6. Des renseignements sur l'estimation de la variance se trouvent au chapitre 13.0.

4.4 Changements concernant le traitement des données

Le système de traitement généralisé a été adopté au cycle 6, de l'étape de l'épuration jusqu'à celle de la production des fichiers principaux. Le système n'a été utilisé qu'après l'étape de l'épuration au cycle 5, et des vérifications des liens de parenté ont été effectuées. De plus, pour des raisons liées à l'efficacité et à la qualité des données, les vérifications longitudinales ont été effectuées à l'étape du calcul des variables dérivées.

Il y avait 26 jeunes de 16 et 17 ans vivant de manière autonome au cycle 6 (FDMCfD03=82). Toutes les variables présentes sur le fichier longitudinal ont été codées « enchaînement valide » pour ces répondants, sauf pour quelques variables dans la section Données démographiques, et certaines variables dans les sections Socio-démographique et Garde légale de l'enfant, pour lesquelles certaines valeurs ont été importées des cycles précédents.

Des systèmes de classification à jour ont été utilisés pour coder les données sur les branches d'activité et les professions. La Classification nationale des professions pour statistiques (CNP-S) de 2001 a été utilisée pour coder les données concernant la ou les professions des répondants. Les données sur les branches d'activité ont été codées à l'aide du Système de classification des industries de l'Amérique du Nord de 2002 (SCIAN 2002).

4.5 Changements concernant les fichiers diffusés

Les variables concernant les jeunes sont diffusées dans un fichier distinct au cycle 6. Il y a maintenant deux fichiers pour la cohorte longitudinale : un pour les enfants (10 à

17 ans) et un autre pour les jeunes (16 à 21 ans). Les questions concernant les jeunes de 16 et 17 ans seront dans le fichier des jeunes, tandis que les autres données pour ces répondants se trouveront dans le fichier des enfants. Seules les questions sur les jeunes ont été posées aux jeunes de 16 et 17 ans qui ne vivent plus avec un parent ou un tuteur. Toutes les données relatives aux jeunes de 18 à 21 ans se trouvent dans le fichier du jeune.

Les variables statiques des cycles précédents (pays de naissance, poids à la naissance, etc.) ont été ajoutées aux fichiers du cycle 6.

Les variables PPERSRUK et SPERSRUK ont été ajoutées au fichier principal pour identifier la personne la mieux renseignée (PMR) et le conjoint. Ces variables seront utiles dans les cycles à venir puisqu'elles permettront de repérer facilement les changements de PMR et/ou de conjoint d'un cycle à l'autre.

5.0 Méthodologie de l'enquête – Échantillon

Le présent chapitre fournit des précisions sur l'échantillon de l'Enquête longitudinale nationale sur les enfants et les jeunes (ELNEJ). La section 5.1 présente un aperçu du plan d'échantillonnage, tandis que la section 5.2 décrit l'Enquête sur la population active (EPA) étant donné que tous les enfants qui ont pris part au cycle 6 de l'ELNEJ ont été échantillonnés à partir de l'EPA. On trouvera à la section 5.3 des précisions sur le procédé d'échantillonnage à chaque cycle ainsi que sur les taux de réponse. Puis, à la section 5.4, le lecteur trouvera des précisions sur la façon dont les différents échantillons de chaque cycle de l'ELNEJ peuvent être utilisés pour faire des inférences concernant des populations de référence longitudinales et transversales précises (des populations pour lesquelles l'ELNEJ produit des poids d'enquête).

5.1 Aperçu du plan d'échantillonnage – Cycles 1 à 6

L'ELNEJ est une enquête probabiliste conçue pour recueillir des renseignements détaillés tous les deux ans 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 avec une importante cohorte d'enfants de 0 à 11 ans vivant dans les dix provinces. Voilà ce dont nous nommons la cohorte initiale. Cette cohorte est purement longitudinale, en ce sens qu'aucune personne n'y est rajoutée 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 réunir des renseignements sur les jeunes enfants et d'établir certains indicateurs (développement moteur et social, troubles émotifs, hyperactivité, agressivité physique, comportement prosocial, compétences linguistiques, etc.). Une nouvelle cohorte d'enfants de 0 à 1 an est donc sélectionnée à chaque cycle depuis le deuxième cycle, et cette cohorte est suivie pendant trois cycles (c'est-à-dire jusqu'à ce que les enfants aient entre 4 et 5 ans)¹. Ces échantillons servent à des fins longitudinales et transversales et sont désignés « cohortes du DPE ».

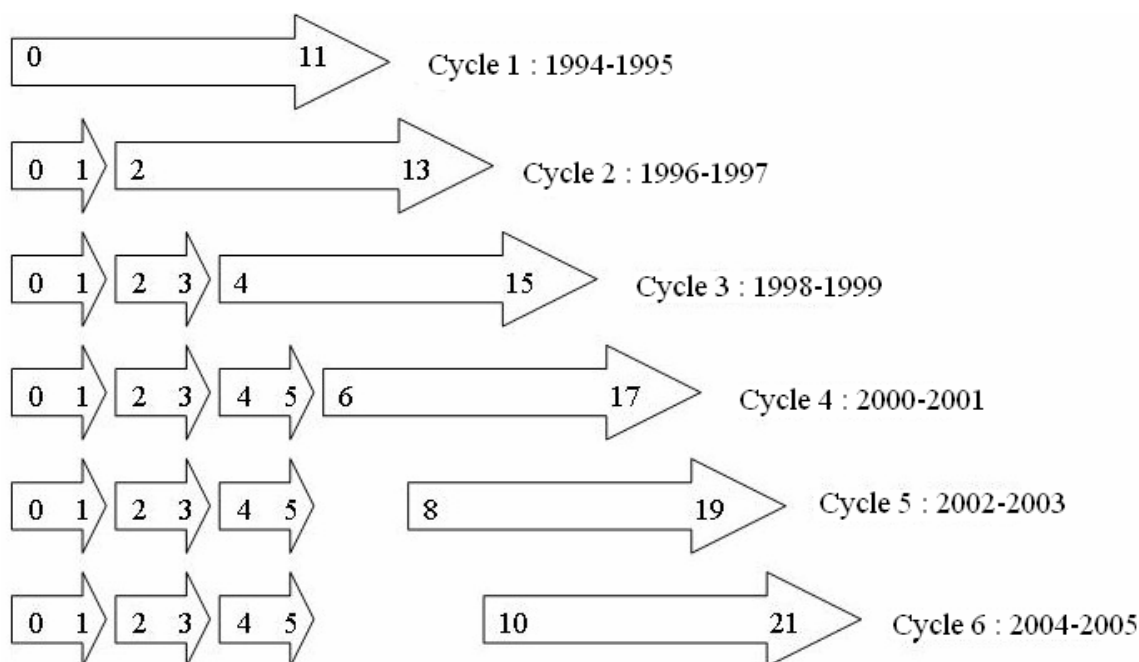
En règle générale, les enfants de l'ELNEJ sont sélectionnés dans les ménages de l'échantillon de l'EPA de Statistique Canada. Font exception les enfants de 1 an du cycle 3 et certains enfants de 5 ans des cycles 3 et 4, qui ont été sélectionnés à partir du Registre des naissances. Aucun de ces enfants ne se trouve dans l'échantillon du cycle 6 puisqu'il aurait plus que 5 ans.

Au cycle 6, l'échantillon de l'enquête comprend des enfants de 0 à 5 ans (cohorte du DPE) et de 10 à 21 ans (cohorte initiale). L'âge effectif de l'enfant au cycle 6 est établi en fonction du 31 décembre 2004. Les enfants de 0 année sont donc nés en 2004 et ceux de 1 an, en 2003.

Le diagramme 1 ci-dessous illustre bien l'échantillon de l'ELNEJ. L'année renvoie à la collecte. Les plus longues flèches représentent la cohorte initiale et les plus petites, les cohortes du DPE.

¹ Au cycle 7, les enfants de 6 à 9 ans seront inclus.

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



Voir la section 5.3 pour des précisions sur le procédé d'échantillonnage utilisé à chaque cycle.

5.2 L'Enquête sur la population active

L'EPA est une enquête mensuelle qui vise à recueillir des données relatives au marché du travail auprès d'un échantillon national d'environ 54 000 logements. 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. Il y a eu deux remaniements de l'EPA pendant toute la durée de vie de l'ELNEJ, le premier en 1994 et le second en 2004-2005 (complété en avril 2005). Il s'ensuit que l'échantillon de la cohorte initiale est formé d'une combinaison de personnes d'avant le remaniement de l'EPA de 1994 et du remaniement de 1994. La vaste majorité des enfants de la cohorte du DPE du cycle 6 provenaient du remaniement de l'EPA de 1994, quelques-uns provenant du remaniement de 2004-2005. Les parties suivantes donnent des renseignements détaillés sur le remaniement de l'EPA de 1994.

5.2.1 Population cible

L'échantillon de l'EPA est représentatif de la population civile, âgée de 15 ans ou plus, ne vivant pas dans un établissement et résidant dans les 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 plein temps des Forces armées canadiennes et les pensionnaires d'institutions. Ces groupes représentent en tout environ 2 % de la population âgée de 15 ans et plus.

5.2.2 Stratification

L'échantillon 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.

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 dans le temps. Les régions économiques de l'assurance-emploi sont également des régions géographiques qui correspondent à peu de choses près aux régions économiques aux plans de la taille et du nombre, mais elles diffèrent au chapitre de la définition. Des estimations de la population active sont produites pour les RÉAE utilisées par Ressources humaines et Développement social 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 on procède à une stratification plus poussée à l'intérieur de chacune d'elles (voir la section 5.2.3). Nota : 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. Les régions urbaines englobent les plus grandes RMR jusqu'aux plus petits villages classés comme régions urbaines (1 000 personnes ou plus) dans le cadre du Recensement de 1991, 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, et celles qui n'ont recours qu'à une base aréolaire.

Environ 1 % de la population de l'EPA demeure dans des régions éloignées des provinces, que 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 grand, 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 dans les grappes, attribuable à la construction de nouveaux immeubles d'appartements. Dans les principaux centres, les strates d'appartements sont, en outre, subdivisées en strates de faibles revenus et en strates ordinaires.

Lorsqu'il est possible et / ou 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, ce qui entraîne la formation de strates tenant compte de ces variations. Les diverses stratégies de stratification des régions rurales étaient 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. Diverses méthodes sont utilisées pour définir les grappes, selon le type de strate.

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, ou un multiple de six grappes) 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 SD sont utilisés comme grappes. Dans les strates urbaines à faible densité, on applique un plan à trois étapes en vertu duquel deux villes faisant partie d'une strate sont échantillonnées, puis six 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.

Dans chacune des strates secondaires dans les régions rurales, d'autres procédures de stratification sont effectuées, le cas échéant, afin de faire ressortir les différences parmi les caractéristiques socio-économiques de chaque strate. Dans chacune des strates rurales, six SD ou deux ou trois groupes de SD sont échantillonnés en tant que grappes.

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 on dresse 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 Renouvellement de l'échantillon

L'EPA utilise un plan de sondage avec renouvellement de panel, suivant lequel les ménages sélectionnés restent dans l'échantillon pendant six mois consécutifs. L'échantillon complet est formé de six sous échantillons ou panels représentatifs. 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. Il en résulte un chevauchement des cinq sixièmes de l'échantillon d'un mois à l'autre, ce qui procure un plan de sondage efficace pour estimer les variations d'un mois à l'autre. La rotation après six mois évite d'imposer un fardeau trop lourd aux enquêtés des ménages sélectionnés aux fins de l'enquête.

Sur le plan statistique, ce mécanisme de renouvellement a pour avantage de fournir une base d'échantillonnage commune qui permet de faire des comparaisons, d'un mois à l'autre, des caractéristiques de l'EPA. Il permet en outre de s'assurer que l'échantillon de ménages reflète adéquatement l'état actuel du parc immobilier et contribue à atténuer les problèmes de non-réponse ou de fardeau de 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 mettre à profit la caractéristique liée aux groupes de renouvellement pour utiliser des tailles d'échantillon plus grandes ou plus petites que celle de l'EPA.

5.2.6 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 fournissant de l'information sur la composition du ménage. On recueille des renseignements démographiques tels le nom, le sexe, la date de naissance et le niveau d'instruction concernant 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 ou 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 6, l'échantillon de l'ELNEJ comprend une cohorte du DPE formée d'enfants de 0 à 1 an, un échantillon complémentaire de nouveaux enfants de 2 à 5 ans ainsi que des enfants de 2 à 5 ans et des jeunes de 10 à 21 ans des cycles précédents. Tous les enfants proviennent des ménages de l'EPA. C'était la première fois qu'on utilisait l'EPA pour compléter l'échantillon existant de l'ELNEJ.

Outre qu'elle explique la façon dont l'échantillonnage a été effectué, la présente section donne des précisions sur les enfants 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 (décédés ou à l'extérieur du pays, par exemple) ou en raison de trop nombreux cycles de non-réponse.

5.3.1 Cohorte initiale – Cycles 1 à 6

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

Cycle 1

L'échantillon d'enfants prélevé au cycle 1 a été conçu pour produire des estimations provinciales fiables (mais non fiables d'une façon égale) pour les enfants de 0 à 11 ans, par groupe d'âge de deux ans (0 à 1 an, 2 à 3 ans, 4 à 5 ans, 6 à 7 ans, 8 à 9 ans et 10 à 11 ans). Quatre enfants par ménage au maximum ont été sélectionnés, et les 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, il y avait 22 831 enfants répondants dans l'ELNEJ, et le taux de réponse des ménages se situait à 86,7 % (voir le tableau 1).

Consulter le Guide de l'utilisateur des microdonnées du cycle 1 pour obtenir plus de renseignements.

Cycle 2

Certains enfants ont été retirés de l'échantillon au deuxième cycle pour des raisons d'ordre budgétaire. 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. Ce qui donnait un échantillon de 16 903 enfants au début du cycle 2. Le taux de réponse des ménages s'établissait à 91,7 % (voir le tableau 1). Le taux de réponse longitudinal cumulatif des ménages de la cohorte initiale était de 79,8 % (voir le tableau 2).

Consulter le Guide de l'utilisateur des microdonnées du cycle 2 pour obtenir plus de renseignements.

Cycle 3

Au cycle 3, 185 enfants ont été retirés 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 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 ménages s'élevait à 89,6 % (voir le tableau 1), et le taux de réponse longitudinal cumulatif des ménages de la cohorte initiale se situait à 77,5 % (voir le tableau 2).

Consulter le Guide de l'utilisateur des microdonnées du cycle 3 pour obtenir plus de renseignements.

Cycle 4

Au cycle 4, afin d'accroître l'efficacité du processus de collecte, il a été décidé que les ménages n'ayant pas répondu pendant deux cycles consécutifs ou plus seraient retirés de l'échantillon (ainsi que les ménages comptant un cycle de non-réponse puis ayant le code «déménagé temporairement»). Donc, 1 086 enfants ont été retirés 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 ménages se situait à 84,8 % (voir le tableau 1), et le taux de réponse longitudinal cumulatif des ménages de la cohorte initiale s'établissait à 69,4 % (voir le tableau 2).

Consulter le Guide de l'utilisateur des microdonnées du cycle 4 pour obtenir plus de renseignements.

Cycle 5

Au cycle 5, il a été décidé que les jeunes de 18 ou 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 enfants 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 (34) ou ils n'avaient pas répondu pendant de trop nombreux cycles consécutifs (435). Au total, 15 163 enfants ont été suivis au cycle 5. Le taux de réponse des ménages se situait à 83,2 % (voir le tableau 1), et le taux de réponse longitudinal cumulatif des ménages de la cohorte initiale s'établissait à 66,6 % (voir le tableau 2).

Consulter le Guide de l'utilisateur des microdonnées du cycle 5 pour obtenir plus de renseignements.

Cycle 6

Au cycle 6, 1 506 enfants ont été retirés de l'échantillon parce que, à la fin du cycle 5, ils étaient hors du champ de l'enquête sur le plan transversal (26) ou ils n'avaient pas répondu pendant de trop nombreux cycles consécutifs (1 480). Au total, 13 657 enfants ont été interviewés au cycle 6. Le taux de réponse des ménages se situait à 83,5 % (voir le tableau 1), et le taux de réponse longitudinal cumulatif des ménages de la cohorte initiale s'établissait à 62,1 % (voir le tableau 2).

5.3.2 Cohortes du Développement de la petite enfance (DPE) présentes au cycle 6

Au cycle 6, les seuls enfants de la cohorte du DPE présents étaient les bébés introduits aux cycles 4, 5 et 6 ainsi que l'échantillon complémentaire de nouveaux enfants de 2 à 5 ans prélevés dans l'échantillon de l'EPA. Lorsque la première cohorte de bébés du DPE 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 échantillonnés². Au cycle 5, un seul enfant par ménage sans exception était échantillonné. Les jumeaux déjà échantillonnés continuaient toutefois à être interviewés jusqu'au cycle 6³.

Pour ce qui est des échantillons de la cohorte du DPE, seuls les répondants du cycle précédent sont interviewés au cycle suivant, contrairement à la cohorte initiale, où il doit y avoir deux ou trois cycles consécutifs de non-réponse avant que l'enfant ne soit retiré de l'échantillon.

Pour des renseignements sur les cohortes du DPE introduites aux cycles 2 et 3, prière de consulter les Guides de l'utilisateur des microdonnées concernant ces cycles. Voir le diagramme 1 pour une illustration des cohortes du DPE présentes au cycle 6.

Cohorte du DPE au cycle 4

Au cycle 4, un échantillon d'enfants de 0 à 1 an a été sélectionné dans l'EPA ainsi qu'un échantillon complémentaire d'enfants de 5 ans provenant du Registre des naissances. La taille totale de l'échantillon était de 9 439 ménages. À la fin du cycle 4, il y avait 6 960 enfants répondants, et le taux de réponse se situait à 75,8 % (voir le tableau 1).

Au cycle 5, les enfants de 5 ans présents au cycle 4 ont été retirés de l'échantillon (4 405 ménages). À la fin du cycle 4, 125 ménages étaient hors du champ de l'enquête sur le plan transversal, et 1 121 étaient des non-répondants. Par conséquent, 3 788 des ménages de retour avec des enfants de 2 à 3 ans de la cohorte du DPE du cycle 4 ont été suivis au cycle 5. Le taux de réponse se situait à 86,9 % (voir le tableau 1), et le taux de réponse longitudinal cumulatif s'établissait à 67,1 % (voir le tableau 2).

À la fin du cycle 5, sur les enfants de retour âgés de 4 à 5 ans, 11 étaient inadmissibles sur le plan transversal et 507 étaient des non-répondants. Ces enfants ont été enlevés de l'échantillon, et 3 323 ont été interviewés au cycle 6. Le taux de réponse était de 89,6 % (voir le tableau 1), et le taux de réponse longitudinal cumulatif se situait à 60,1 % (voir le tableau 2).

Cohorte du DPE au cycle 5

Au cycle 5, un échantillon d'enfants de 0 à 1 an a été sélectionné dans l'EPA. La taille totale de l'échantillon était de 4 492 enfants et ménages. À la fin du cycle 5, il y avait 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, seulement 3 252 enfants parmi les répondants de 2 à 3 ans de retour de la cohorte du DPE du cycle 5 ont été interviewés au cycle 6. Le taux de réponse était de 88,6 % (voir le tableau 1), et le taux de réponse longitudinal cumulatif se situait à 65,6 % (voir le tableau 2).

2 La cohorte du DPE échantillonnée au cycle 2 incluait des enfants de 0 à 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 étaient sélectionnés. Aucun enfant de la cohorte du DPE du cycle 2 n'est présent dans l'échantillon du cycle 6.

3 Pour le 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.

Cohorte du DPE du cycle 6

Au cycle 6, un échantillon d'enfants de 0 à 1 an ainsi qu'un échantillon complémentaire de nouveaux enfants de 2 à 5 ans ont été sélectionnés dans l'EPA. La taille totale de l'échantillon était de 5 795 enfants et ménages. À la fin du cycle 6, il y avait 4 684 enfants répondants, et le taux de réponse se situait à 81,3 % (voir le tableau 1).

Tableau 1 : Réponse au niveau ménage lors de la collecte (nombre d'enfants entre parenthèses)

Cohorte d'échantillonnage	Cycle de l'enquête	Groupe d'âge (en années)	Collecte									
			Taille de l'échantillon	Réduction de l'échantillon	Hors du champ (retirés du cycle précédent)	Non répondants (retirés du cycle précédent)	Éligible	Répondants	Taux en portée (%)	Taux d'éligible (%)		
Cohorte initiale	1	0-11	43 751 (s/o)	s/o	s/o	s/o	15 502 (s/o)	13 439 (22 831)	35,4	86,7		
	2	2-13	11 188 (16 903)	25 588 (s/o)	5 345 (s/o)	1 677 s/o	11 140 (16 816)	10 216 (15 391)	99,6	91,7		
	3	4-15	11 032 (16 718)	0 (0)	38 (71)	73 (114)	10 937 (16 563)	9 801 (14 777)	99,1	89,6		
	4	6-17	10 449 (15 632)	0 (0)	65 (106)	618 (980)	10 417 (15 586)	8 833 (13 174)	99,7	84,8		
	5	8-19	10 355 (15 163)	0 (0)	25 (34)	285 (435)	10 320 (15 113)	8 582 (12 280)	99,7	83,2		
	6	10-21	9 881 (13 657)	0 (0)	17 (26)	861 (1 480)	9 816 (13 572)	8 201 (11 178)	99,3	83,5		
2	2	0-1	5 528 (s/o)	s/o	s/o	s/o	4 865 (s/o)	4 496 (4 634)	88,0	92,4		
	3	2-3	3 935 (3 989)	551 (s/o)	598 (s/o)	444 (s/o)	3 893 (3 947)	3 592 (3 640)	98,9	92,3		
	4	4-5	3 577 (3 610)	464 (484)	25 (25)	33 (34)	3 552 (3 585)	3 023 (3 052)	99,3	85,1		
3	3	0-1 & 5	16 812 (s/o)	s/o	s/o	s/o	15 929 (s/o)	13 256 (13 546)	94,7	83,2		
	4	2-3	7 941 (8 118)	6 935 (s/o)	516 (s/o)	1 420 (s/o)	7 896 (8 070)	6 956 (7 111)	99,4	88,1		
	5	4-5	6 960 (7 115)	0 (0)	41 (44)	940 (959)	6 919 (7 073)	6 208 (6 340)	99,4	89,7		
4	4	0-1 & 5	9 439 (s/o)	s/o	s/o	s/o	9 115 (s/o)	6 907 (6 960)	96,6	75,8		
	5	2-3	3 788 (3 841)	4 405 (s/o)	125 (s/o)	1 121 (s/o)	3 776 (3 829)	3 281 (3 324)	99,7	86,9		
	6	4-5	3 280 (3 323)	0 (0)	11 (11)	497 (507)	3 270 (3 313)	2 931 (2 964)	99,7	89,6		
5	5	0-1	4 492 (4 492)	s/o	s/o	s/o	4 394 (4 394)	3 252 (3 252)	97,8	74,0		
	6	2-3	3 252 (3 252)	0 (0)	98 (98)	1 142 (1 142)	3 233 (3 233)	2 866 (2 866)	99,4	88,6		
6	6	0-5	5 795 (5 795)	s/o	s/o	s/o	5 763 (5 763)	4 684 (4 684)	99,4	81,3		

Tableau 2 : Réponse longitudinale au niveau ménage (nombre d'enfants entre parenthèses)

Cohorte d'échantillonnage	Cycle de l'enquête	Groupe d'âge (en années)	Longitudinale				
			Taille de l'échantillon	Éligible	Répondants	Taux d'éligible (%)	Taux de réponse (%)
Cohorte initiale	1	0-11	18 163 (s/o)	12 818 (s/o)	11 141 (16 903)	70,6	86,9
	2	2-13	18 210 (s/o)	12 863 (s/o)	10 262 (15 472)	70,6	79,8
	3	4-15	18 165 (s/o)	12 818 (s/o)	9 933 (14 997)	70,6	77,5
	4	6-17	18 265 (s/o)	12 914 (s/o)	8 962 (13 361)	70,7	69,4
	5	8-19	18 481 (s/o)	13 134 (s/o)	8 744 (12 535)	71,1	66,6
	6	10-21	18 885 (s/o)	13 532 (s/o)	8 409 (11 483)	71,7	62,1
2	2	0-1	4 977 (s/o)	4 379 (s/o)	4 100 (4 154)	88,0	93,6
	3	2-3	4 977 (s/o)	4 377 (s/o)	3 632 (3 680)	87,9	83,0
	4	4-5	4 513 (s/o)	3 784 (s/o)	2 942 (2 971)	83,8	77,7
3	3	0-1	9 877 (s/o)	9 361 (s/o)	7 949 (8 126)	94,8	84,9
	4	2-3	9 877 (s/o)	9 198 (s/o)	6 838 (6 992)	93,1	74,3
	5	4-5	9 877 (s/o)	9 198 (s/o)	6 141 (6 273)	93,1	66,8
4	4	0-1	5 034 (s/o)	4 909 (s/o)	3 788 (3 841)	97,5	77,2
	5	2-3	5 034 (s/o)	4 907 (s/o)	3 291 (3 334)	97,5	67,1
	6	4-5	5 034 (s/o)	4 911 (s/o)	2 952 (2 985)	97,6	60,1
5	5	0-1	4 492 (4 492)	4 394 (4 394)	3 252 (3 252)	97,8	74,0
	6	2-3	4 492 (4 492)	4 392 (4 392)	2 883 (2 883)	97,8	65,6
6	6	0-1	4 356 (4 356)	4 343 (4 343)	3 521 (3 521)	99,7	81,1

Nota : Le tableau 2 comprend seulement les enfants qui font partis de l'étude longitudinale (c.-à-d., les enfants qui sont suivis à travers les cycles).

5.4 Populations de référence longitudinale et transversales 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 chacun (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 dix d'entre eux sont sélectionnés à l'aide d'une technique d'échantillonnage aléatoire simple, alors 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. C'est l'inverse de la probabilité de sélection (des ajustements ultérieurs sont habituellement apportés, notamment pour tenir compte de la non-réponse et pour assurer la correspondance avec le dénombrement démographique selon l'âge et le sexe). 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. (Voir le chapitre 11.0 pour plus de détails 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 vient le moment de faire des inférences concernant 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 au fil des ans. La raison en est que les populations changent avec le temps (décès, immigration et émigration).

Deux types de populations sont possibles dans une enquête longitudinale : la population longitudinale et 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 dix provinces pendant la collecte des données du cycle 1, en 1994-1995. Une population transversale au cycle 6 pourrait comprendre les enfants qui avaient entre 0 et 11 ans le 31 décembre 1994 et qui vivaient dans l'une des dix provinces en 2004 (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 complétée pour tenir compte des immigrants arrivés après 1994-1995. Des poids transversaux ont été produits pour la cohorte initiale pour la dernière fois au cycle 4. 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. À 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. (Voir le chapitre 16.0 pour obtenir plus de renseignements sur la façon de regrouper les échantillons de l'ELNEJ.). Il convient également de noter 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. (Voir 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. À noter que pour le premier cycle d'une cohorte, la population longitudinale est définie par les poids dits « transversaux » pour ce premier cycle.

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 dix provinces lors de la collecte des données du cycle 1 (1994-1995)

Un échantillon longitudinal d'enfants de 0 à 11 ans a été prélevé à même l'EPA au cycle 1. Ces enfants avaient entre 10 et 21 ans au cycle 6 (au 31 décembre 2004). 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 6

- Population longitudinale de la cohorte du DPE sélectionnée au cycle 2 : enfants âgés de 0 à 1 an au 31 décembre 1996 qui vivaient dans l'une des dix provinces lors de la collecte des données du cycle 2 (1996-1997)

Au cycle 2, un échantillon longitudinal d'enfants de 0 à 1 an a été prélevé à même l'EPA. 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 6.

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

- Population longitudinale de la cohorte du DPE sélectionnée au cycle 3 : enfants âgés de 0 à 1 an au 1^{er} décembre 1998 qui vivaient dans l'une des dix provinces lors de la collecte des données du cycle 3 (1998-1999)

Au cycle 3, un échantillon longitudinal d'enfants de 0 à 1 an a été sélectionné. Les enfants de moins de 1 an ont été sélectionnés dans l'EPA, tandis que ceux de 1 an provenaient du Registre des naissances. Cette cohorte n'a été suivie que pendant trois cycles, et aucun de ces enfants n'est présent au cycle 6. Ils seront toutefois de retour au cycle 7 (ils auront de 8 à 9 ans).

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

- Population longitudinale de la cohorte du DPE sélectionnée au cycle 4 : enfants âgés de 0 à 1 an au 31 décembre 2000 qui vivaient dans l'une des dix provinces lors de la collecte des données du cycle 4 (2000-2001)

Au cycle 4, un échantillon longitudinal d'enfants de 0 à 1 an a été prélevé à même l'EPA. Ces enfants avaient de 4 à 5 ans au cycle 6 (en décembre 2004). Ils seront interviewés au cycle 7 (où ils auront de 6 à 7 ans).

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

- Population longitudinale de la cohorte du DPE sélectionnée au cycle 5 : enfants âgés de 0 à 1 an au 31 décembre 2002 qui vivaient dans l'une des dix provinces lors de la collecte des données du cycle 5 (2002-2003)

Au cycle 5, un échantillon longitudinal d'enfants de 0 à 1 an a été prélevé à même l'EPA. Ces enfants avaient 2 et 3 ans au cycle 6. Ils seront interviewés au cycle 7 (où ils auront de 4 à 5 ans).

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

- Population longitudinale de la cohorte du DPE sélectionnée au cycle 6 : enfants âgés de 0 à 1 an au 31 décembre 2004 qui vivaient dans l'une des dix provinces lors de la collecte des données du cycle 6 (2004-2005)

Au cycle 6, un échantillon longitudinal d'enfants de 0 à 1 an a été prélevé à même l'EPA. Ils seront interviewés au cycle 7.

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

5.4.2 Les cohortes et leurs populations transversales

La cohorte initiale (au cycle 6, 10 à 21 ans au 31 décembre 2004)

Comme aucun échantillon complémentaire n'a été ajouté à la cohorte initiale pour 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 donc être utilisée pour faire des inférences concernant 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 dix 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 dix 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 dix 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 concernant 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 dix provinces lors de la collecte des données du cycle 2 (1996-1997)

L'échantillon transversal comprend :

- les enfants de la cohorte du DPE âgés de 0 à 1 an et sélectionnés au cycle 2;
- les enfants de 2 à 13 ans de retour et 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 dix provinces lors de la collecte des données du cycle 3 (1998-1999)

L'échantillon transversal comprend :

- les enfants de la cohorte du DPE âgés de 0 à 1 an et sélectionnés au cycle 3;
- les enfants de 2 à 3 ans de retour de la cohorte du DPE d'enfants de 0 à 1 an sélectionnés au cycle 2;
- l'échantillon complémentaire d'enfants de 5 ans (sélectionnés à partir du Registre des naissances);
- les enfants de 4 à 15 ans de retour 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 dix provinces lors de la collecte des données du cycle 4 (2000-2001)

L'échantillon transversal comprend :

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

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

Après le cycle 4, les inférences concernant les populations transversales ne doivent être faites qu'à l'aide des 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 dix provinces lors de la collecte des données du cycle 5 (2002-2003)

L'échantillon transversal comprend :

- les enfants de la cohorte du DPE âgés de 0 à 1 an et sélectionnés au cycle 5;
- les enfants de 2 à 3 ans de retour de la cohorte du DPE d'enfants de 0 à 1 an sélectionnés au cycle 4;
- les enfants de 4 à 5 ans de retour de 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 dix provinces lors de la collecte des données du cycle 6 (2004-2005)

L'échantillon transversal comprend :

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

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

6.0 Collecte des données

La collecte des données pour le sixième cycle de l'Enquête longitudinale nationale sur les enfants et les jeunes (ELNEJ) s'est déroulée de l'automne 2004 au printemps 2005.

La collecte a été effectuée en trois périodes, décrites ci-dessous. La date de l'interview est consignée dans la variable (FMMHfQ06).

Période	Groupes d'âge
Septembre à janvier Vague 1	Enfants de 0 à 5 ans et jeunes de 18 à 21 ans dont aucun frère ou sœur de 10 à 17 ans n'a été sélectionné
Janvier à avril Vague 2	Jeunes de 10 à 17 ans et de 18 à 21 ans qui n'ont pas été visés dans la première période de collecte
Mars à juin Vague 3	Enfants de 0 à 5 ans

L'enquête combine l'utilisation de méthodes d'interview assistée par ordinateur (IAO) et de questionnaires 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 les réponses données par le répondant dans l'ordinateur. 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 à l'aide 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 papier utilisés pour la collecte de l'ELNEJ. Pour une description plus détaillée du contenu des questionnaires, veuillez consulter le chapitre 8.0.

6.1 Composante « Ménage »

Cette première partie de l'interview sert à dresser la liste de tous les membres du ménage, déterminer les relations entre eux, recueillir des renseignements de dépistage et obtenir les caractéristiques démographiques de base, notamment : le sexe, la date de naissance, l'état matrimonial et les liens entre les membres du ménage.

Cette composante permet en outre d'identifier la Personne la mieux renseignée sur l'enfant (PMR). Celle-ci fournit des renseignements sur tous les enfants sélectionnés au sein du ménage ainsi que sur elle-même et sur son conjoint. Dans certains cas, il aurait été indiqué 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 les renseignements sur les liens entre les membres du ménage recueillis.

6.2 Composante « Enfant »

Une composante « Enfant » est créée pour chaque enfant choisi qui est âgé de 0 à 17 ans. La PMR sur les enfants et les jeunes 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 choisi appartient au sous-groupe de 16 et 17 ans et qu'il réside au domicile des parents une version plus courte de la composante « enfant » a été demandée. Si il ne réside plus au domicile des parents la composante ne sera pas créée.

6.3 Composante « Adulte »

Une composante « Adulte » est créée pour la Personne la mieux Renseignée et son (ou sa) conjoint(e) ou partenaire si l'enfant sélectionné est âgé de 17 ans ou moins. Seule la PMR ou son conjoint ou partenaire 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 choisi appartient au sous-groupe de 16 à 17 ans et qu'il réside au domicile des parents, une version plus courte de la composante « adulte » est demandée. Si il ne réside plus au domicile des parents la composante ne sera pas créée.

6.4 Composante du jeune

Cette composante est réservée aux jeunes âgés de 16 ans et plus. 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 d'échantillon au cycle 6

Les tableaux suivants présentent le nombre d'enfants et jeunes échantillonnés, selon l'âge et la province, pour le cycle 6.

Tableau 1 : Nombre d'enfants dans l'échantillon et taux de réponse selon l'âge, cycle 6

Âge au 1 ^{er} janvier 2005	Échantillonné	Admissible	Répondants	Taux de réponse au cycle 6 (%)
0	1 777	1 770	1 457	82,3
1	2 579	2 573	2 064	80,2
2	1 845	1 834	1 627	88,7
3	2 140	2 121	1 834	86,5
4	1 900	1 889	1 662	88,0
5	2 127	2 122	1 870	88,1
6	1	0	0	s/o
7	0	0	0	s/o
8	0	0	0	s/o
9	0	0	0	s/o
10	1 595	1 587	1 378	86,8
11	1 689	1 683	1 447	86,0
12	1 189	1 182	1 015	85,9
13	1 147	1 146	980	85,5
14	1 085	1 077	907	84,2
15	1 026	1 022	884	86,5
16	924	920	790	85,9
17	932	929	795	85,6
18	1 119	1 106	815	73,7
19	1 052	1 044	758	72,6
20	992	978	746	76,3
21	908	898	663	73,8
Total	26 027	25 881	21 692	83,8

Tableau 2 : Nombre d'enfants dans l'échantillon et taux de réponse, selon la province, cycle 6

Province	Échantillonné	Admissible	Répondants	Taux de réponse au cycle 6 (%)
Terre-Neuve-et-Labrador	1 408	1 405	1 215	86,5
Île-du-Prince-Édouard	959	955	815	85,3
Nouvelle-Écosse	1 749	1 746	1 493	85,5
Nouveau-Brunswick	1 659	1 658	1 364	82,3
Québec	4 385	4 379	3 667	83,7
Ontario	6 827	6 819	5 562	81,6
Manitoba	1 980	1 974	1 696	85,9
Saskatchewan	1 970	1 968	1 697	86,2
Alberta	2 683	2 677	2 263	84,5
Colombie-Britannique	2 304	2 300	1 920	83,5
En dehors des 10 provinces	103	0	0	n/a
Total	26 027	25 881	21 692	83,8

6.6 Mesures directes

Différentes mesures directes sont effectuées auprès du répondant sélectionné. Elles sont résumées ci-dessous. Pour obtenir plus de renseignements à ce sujet, voir le chapitre 14.0.

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

6.7 Questionnaires auto administrés de 10 à 17 ans

Les répondants âgés de 10 à 17 ans remplissent un questionnaire sur 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 sa confidentialité.

Les questionnaires auto administrés consistent en une série de quatre livrets correspondants à chaque cinq groupes d'âge visé. Vous trouverez ci-après une grille indiquant les sujets traités selon le groupe d'âge et la section à l'intérieur du livret. Un même sujet comportera 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 2004-2005, cycle 6 ».

Sujet	Section			
	10 - 11 Livret n° 20 A/F	12 - 13 Livret n° 21 A/F	14 - 15 Livret n° 22 A/F	16 - 17 Livret n° 23 A/F
Amis et famille	A	A	A	A
École	B	B	B	-
À propos de moi	C	C	C	B
Sentiments et comportements	D	D	D	C
Mes parents	E	G	G	G
Tabac, alcool et drogues	G	F	F	D
Puberté	F	H	H	-
Activités	H	E	E	-
Relations amoureuses	-	H	H	F
Santé	-	H	H	E
Travail	-	I	I	-
Remerciement	J	J	J	H

Au cycle 5, les jeunes de 18 et 19 ans complétaient un questionnaire papier. Au cycle 6, toutes les questions posées à ces jeunes ont été admistrées dans la composante du jeune elle-même.

6.7.1 Personnel de collecte (formation, supervision et contrôle)

L'ELNEJ a été effectuée par des intervieweurs de Statistique Canada. Plusieurs avaient déjà eu l'occasion de travailler à un ou 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 non-réponse. Au besoin, les cas de non-réponse sont confiés à l'intervieweur principal et réaffectés. Les intervieweurs principaux, quant à eux, relèvent des gestionnaires de programmes, qui se trouvent dans chacun des cinq bureaux régionaux de Statistique Canada.

Dans le cas de l'ELNEJ, on a combiné la formation en salle de classe et l'auto apprentissage pour s'assurer que les intervieweurs et superviseurs comprenaient bien les concepts de l'enquête. L'auto apprentissage a nécessité des gestionnaires de programme, intervieweurs principaux et intervieweurs qu'ils prennent connaissance du Guide de l'intervieweur préparé pour l'enquête et qu'ils fassent un cas pratique à compléter avant la formation. La formation en classe a été donnée par un gestionnaire de programme ou un intervieweur principal. Il y avait deux modules de formation : un pour les intervieweurs ITAO et un autre pour les intervieweurs IPAO.

7.0 Traitement des données

Les principaux produits de l'Enquête longitudinale nationale sur les enfants et les jeunes (ELNEJ), cycle 6 sont des fichiers de données « épurés ». Ce chapitre présente un bref résumé de certaines des phases de traitement inhérentes à la production de ces fichiers.

Le traitement des données du cycle 6 de l'ELNEJ a été effectué à l'aide d'un environnement de traitement généralisé. 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
- Pré édition
- Vérifications de l'enchaînement des questions
- Codage
- Vérifications de la cohérence
- Variables dérivées
- Fichier de traitement final
- Création du fichier principal

7.1 La vérification informatique

Tel qu'indiqué ci haut, tous les renseignements auprès des ménages, sauf pour les questionnaires auto administrés qui sont destinés aux enfants de 10 à 17 ans, ont été obtenus en personne, ou au moyen d'une interview téléphonique à l'aide d'une application d'interview assistée par ordinateur (IAO). On a pu ainsi inclure diverses fonctions de vérification dans le questionnaire pour les diverses composantes IAO des ménages de sorte que les renseignements recueillis soient 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 :

Écran 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'appuyait sur la liste des membres du ménage. La page-écran correspondante présentait 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 exigeant d'indiquer des valeurs numériques. Si les chiffres indiqués ne s'inscrivaient 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 de cheminement de questions

Tous les cheminements de questions sont en 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 afin de pouvoir travailler ou étudier (elle ou son conjoint/partenaire). Le cheminement des questions peut varier selon la réponse. S'il y a utilisation de garderie ou gardiennage, 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 cohérence générale

On a prévu un certain nombre de vérifications de cohérence dans le système IAO, et les intervieweurs étaient en mesure de revenir à des questions déjà posées pour rectifier les incohérences. Les intervieweurs recevaient aussi des instructions à l'écran pour traiter ou régler des problèmes de réponse incomplète ou erronée. Par exemple, dans la collecte des données de la section sur l'activité professionnelle des parents, le nombre de semaines passées à travailler, à ne pas travailler et à chercher du travail ne doit pas être de plus de 52. En cas de dépassement, le système signalait l'erreur et demandait 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 d'interview papier et crayon

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

- les questionnaires auto administrés destinés aux enfants de 10 à 17 ans
- « Qui suis-je » destiné aux enfants de 4 à 5 ans
- les exercices de calcul destinés aux enfants de 10 à 15 ans qui sont de la 4^e année scolaire ou plus haut
- 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 remplie par le répondant 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 qualité qui marquent les entrées inusitées afin de mettre les opérateurs en garde contre les risques d'entrées erronées.

Si le répondant avait coché plusieurs réponses, il devait accepter la première. Les erreurs restantes devaient être corrigées à une étape ultérieure.

7.3 Nettoyage

Exigences définitives

L'objet de cette étape est de retirer les enregistrements en doubles complets et de séparer les enregistrements d'identification en double pour examen. Ensuite, les données sont divisées en réponse et non-réponse selon les critères donnés.

On a examiné les questionnaires des répondants et des non répondants et on a établi des spécifications fondées sur cette analyse pour déterminer les enregistrements qui seraient éliminés en raison de la non-réponse. Essentiellement, s'il manquait des données essentielles ou s'il n'y avait pas de réponse à plus de la moitié des questions sur un enregistrement, il était éliminé du fichier.

À la fin de cette étape, les enregistrements ont été traités par 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 auto administrés.

Variables manquantes

Pour toutes les variables manquantes des ménages, on indiquait « non déclaré ». Si on n'avait pas assez de renseignements, le ménage était retiré de l'échantillon des répondants et on en faisait un cas de non-réponse.

Le fichier longitudinal contient aussi 253 enregistrements créés à l'égard des enfants de l'échantillon longitudinal pour qui on n'avait pas recueilli de données dans ce cycle. Il s'agit

d'enfants qui sont décédés ou qui ont déménagé à l'étranger, mais qu'on gardera au fichier longitudinal à des fins de pondération. Dans le cas de ces enregistrements, on a fixé à « non déclaré » toutes les variables sauf quelques valeurs reliées à la pondération longitudinale (FWTCW01L).

7.4 Vérification de l'âge et du genre

À cette étape, une vérification est menée pour toutes les variables relatives à l'âge. Une comparaison avec le cycle antérieur est faite entre la date de naissance et l'âge déclaré. L'âge de tous les enfants et les jeunes déclarants est comparé au cycle précédent. De plus, l'âge est vérifié afin d'être cohérent à l'âge de la cohorte. Pour ce cycle, la date de l'interview (INTVDATE) est utilisée pour vérifier l'âge effectif de tous les membres du ménage. INTVDATE figure maintenant dans les fichiers des données diffusées. Le sexe de l'enfant est vérifié afin d'être cohérent au cycle précédent.

7.5 Vérification de liens

L'étape de vérification de liens établit les liens entre les membres du ménage et crée les variables dérivées familiales. 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 par une application utilisant des règles établies, tandis que d'autres incohérences sont identifiées pour une révision manuelle et pour un recodage. Une série de variables dérivées connexes est produite à partir des vérifications de liens.

7.6 Vérifications préliminaires

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

7.7 Vérification de cheminement

Cette vérification consiste à reproduire le cheminement des questions du questionnaire. Les variables qui sont sautées en suivant l'enchaînement des questions sont changées de *Non déclaré* à *Enchaînement valide* (6, 96, 996...).

Lorsque l'instruction « passez à » est basé sur l'âge ou suit une question à laquelle on a répondu, toutes les questions sautées sont indiquées comme *Enchaînement valide*. Lorsque l'instruction « passez à » suit une réponse *Ne sait pas* ou *Refus*, toutes les questions sautées sont indiquées comme *Non déclaré*.

Il y avait 26 jeunes de 16 et 17 ans vivant de manière autonome au cycle 6 (FDMCfD03=82). Toutes les variables présentes sur le fichier longitudinal ont été codées "enchaînement valide" pour ces répondants, sauf pour quelques variables dans la section Données démographiques, et certaines variables dans les sections Socio-démographique et Garde légale de l'enfant, pour lesquelles certaines valeurs ont été importées des cycles précédents.

7.8 Codage des questions ouvertes

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

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

Des questions relatives à l'aspiration de carrière furent demandées aux 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 inscrivait littéralement les réponses données par le répondant. Au bureau central, les énoncés écrits ont été convertis en codes d'industrie et de profession pour décrire la nature du travail du répondant.

Mode de codage

Ces questions ouvertes ont été codées au moyen de plusieurs classifications types. Pour le cycle 6, les questions portant sur les professions ont été codées à l'aide de la Classification Nationale des Professions - 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)⁴. On peut consulter les listes groupées de ces codes au fichier de données (FLFPfD7A et FLFPfD8A pour la PMR, et FLFSfD7A et FLFSfD8A pour le conjoint/partenaire).

7.9 Vérification de la cohérence

Après la vérification de l'enchaînement, nous avons procédé à une vérification de cohérence qui consiste à vérifier les liens entre deux variables ou plus. On utilise des tables de décision pour préciser les vérifications. Le logiciel LogiPlus est utilisé pour générer le code SAS en utilisant l'information fournie par les tables de décision. Un rapport comprenant les comptes « avant » et « après » des variables est produit, de même qu'un rapport comprenant le nombre de décisions prises.

Par exemple, à la section sociodémographique, une question portait sur la date de la première immigration au Canada (FSDCQ2B) dans le cas des enfants nés à l'étranger. Une vérification de cohérence a permis de comparer cette réponse à l'année de naissance de l'enfant. Si l'année d'immigration précédait l'année de naissance, on indiquant « Non déclaré » pour l'année d'immigration dans le cadre de la vérification.

7.10 Indicateurs d'imputation

Variables manquantes

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

Imputation

Cependant, pour certaines variables du fichier de l'ELNEJ, on a procédé à une imputation au lieu d'attribuer un code de non-réponse. L'imputation consiste à remplacer des réponses

⁴ Des renseignements sur la classification ainsi que la concordance aux classifications antérieures se trouvent à l'adresse suivante : www.statcan.ca/francais/concepts/index_f.htm.

manquantes ou incohérentes pour des valeurs vraisemblables. Dans le cas de l'ELNEJ, l'imputation a porté sur le revenu du ménage, sur celui de la PMR, sur celui du jeune et de l'Échelle du développement moteur et social. Pour de plus amples détails, consultez le chapitre 10.0.

On a ajouté des indicateurs d'imputation au fichier de l'ELNEJ pour que l'utilisateur connaisse l'ampleur de cette imputation et les questions ayant été imputées.

Tous les indicateurs d'imputation du fichier de données de l'ELNEJ comportent un « I » en sixième position dans le nom de la variable. Ainsi, l'indicateur d'imputation du revenu de la PMR est FINPcl1a.

7.11 Établissement de variables dérivées et de vérifications longitudinales

Pour des raisons d'efficacité et de qualité des données, ces étapes ont été regroupées au cycle 6 et elles le seront aussi dans les cycles à venir.

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

Un certain nombre de variables ont été dérivées en combinant des questions et/ou des variables dérivées afin de faciliter l'analyse des données. Par exemple, à la section sur l'activité professionnelle, la situation actuelle du répondant vis-à-vis de l'activité est obtenue à partir des questions sur son travail actuel. À la section sur l'éducation des adultes, on lui demande s'il va actuellement à l'école, au collège ou à l'université. En combinant ces deux variables, on forme une variable dérivée identifiant la situation courante travail et études du répondant (FLFPD51, FLFSD51).

Variables longitudinales dérivées

On a créé des variables longitudinales dérivées pour indiquer les changements de données entre le cycle en cours et les cycles antérieurs pour ce qui est de la structure familiale, de la PMR et du conjoint/partenaire.

De plus, les données incohérentes entre les cycles ont été indiquées, et le cinquième caractère du nom des variables qui indiquent ces incohérences est un « Z ». Par exemple, si un répondant a déclaré dans un cycle précédent qu'il avait eu recours à des services de garde d'enfants et qu'il a déclaré au cycle 6 n'avoir jamais eu recours à ces services, une valeur de « 1 » est attribuée à la variable FCRCeZQ6. Les incohérences entre les cycles relativement aux modalités de la garde des enfants ainsi qu'à la taille et au poids des enfants sont aussi indiquées.

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

Au cycle 6, les données reportées proviennent des cycles 1 à 5. Au cycle 5, elles ne provenaient que du cycle 4.

Nom des variables dérivées

Toutes les variables dérivées dans le fichier de données de l'ELNEJ reçoivent un « D » en cinquième position dans leur nom. Le nom de la variable dérivée portant sur le principal mode de garde est FLFPD51.

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 en toute cohérence pour toutes les variables. Voici ces diverses situations avec leurs codes descriptifs respectifs.

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 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 de l'ELNEJ, le code utilisé pour indiquer que le répondant ne connaissait pas la réponse à une question particulière est « 7 ». Dans le cas d'une variable à un chiffre, le code est « 7 », pour une variable à deux chiffres « 97 », pour une variable à trois chiffres « 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 était sans objet. Par exemple, la question sur le nombre d'heures hebdomadaires où l'enfant est en garderie (FCRCQ1G1) ne s'applique que pour les enfants qui vont à la garderie (FCRCQ1G = 1). Dans les autres cas, on attribue le code 996 à cette question.

Dans d'autres cas, 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. Si un enfant appartient à une autre tranche d'âge (enfants de quatre ans et plus), le système indique que les variables en question sont des enchaînements valides (« 6 », « 96 », « 996 »).

Là où la PMR n'a pas de conjoint de droit ou de fait qui vit dans le ménage, toutes les variables du conjoint/partenaire (sections sur l'activité professionnelle et la scolarité du conjoint/partenaire) 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 quelques raisons :

- Dans le cadre de l'interview assistée par ordinateur, l'intervieweur pouvait entrer un code « Refus » ou « Ne sait pas », ainsi que nous l'avons expliqué plus haut. Le système IAO était souvent programmé en pareil cas pour sauter cette section particulière du questionnaire. En cas de refus, on supposait que les questions posées étaient délicates et qu'il était probable que le répondant ne veuille pas répondre à d'autres questions à ce sujet. Dans les cas « Ne sait pas », on parlait du principe que le répondant n'était 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 ». Nous entendons par-là que le répondant n'a pas répondu à la question. Nous ignorons même parfois si la question s'appliquait à ce dernier.

- Dans certains cas, il était impossible de remplir un questionnaire ni même parfois de le commencer. Par exemple, il a pu se produire une interruption ou le répondant a dit ne pas vouloir continuer. Si nous avons obtenu suffisamment de renseignements pour que le ménage soit considéré comme répondant, nous attribuerions le code « Non déclaré » à toutes les questions restantes (et aux composantes dont nous n'avions pas encore commencé). Comme seule exception, si nous savions qu'une section ou qu'un questionnaire ne s'appliquait pas, nous avons codé ces éléments comme « Sans objet ».
- Pour les questionnaires papier, quelques questionnaires partiellement remplis ont été retournés par la poste. Si nous avons recueilli suffisamment de données sur le répondant, nous attribuons le code « Non déclaré » à toutes les autres questions sans réponse du questionnaire. Comme seule exception, si nous savions qu'une section ou qu'un questionnaire ne s'appliquait pas, nous avons codé ces éléments comme « Sans objet ».
- Dans une autre situation, nous avons recouru à des codes « Non déclaré » à la suite de vérifications de cohérence. Si nous décelions une erreur de cohérence entre des groupes de variables, une ou plusieurs des variables en question se voyaient attribuer le code « Non déclaré ».

Dans le cas des variables dérivées, si une ou plusieurs des variables d'entrée avaient reçu le code « Refus », « Ne sait pas » ou « Non déclaré », nous avons fait de même pour la variable dérivée.

On a attribué un code « 9 » aux cas « Non déclaré ». Dans le cas d'une variable à un chiffre, le code est « 9 », pour une variable à deux chiffres « 99 », pour une variable à trois chiffres « 999 », etc.

7.13 Nomenclature

La documentation du fichier de microdonnées de l'ELNEJ utilise certaines normes de désignation et de codage des valeurs des variables, le but étant de rendre l'interprétation des données plus facile à l'utilisateur.

Nous avons conçu une nomenclature pour les diverses variables du fichier de données de l'ELNEJ pour 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 » sont les 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 et
- « **F** » indique le sixième cycle.

Les **deuxième et troisième** caractères désignent la section du questionnaire où la question a été posée dont la variable a été tirée. Veuillez consulter la section 7.15 pour les 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 neuf possibilités⁵:

- « **C** » si la variable indique l'enfant,
- « **P** » si la variable indique la PMR,
- « **S** » si la variable indique le conjoint/partenaire,
- « **H** » si la variable indique le ménage
- « **Y** » si la variable indique le jeune
- « **W** » si la variable indique la variable pondérée
- « **M** » si la variable indique la mère

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 :

- f** la lettre minuscule désigne le cycle de l'ELNEJ où la variable a figuré pour la première fois au dossier ou le cycle dans lequel des changements ont été apportés à une question qui avait déjà été posée.

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 6 la variable est considérée comme nouvelle et se voit accorder la lettre « f ».

Par exemple, pour la variable FLTCfQ1E, la couverture comprend maintenant les enfants de 0 à 5 ans au cycle 6 alors qu'elle ne comprenait que les enfants de 0 à 2 ans au cycle précédent. La variable FLTCfQ4B correspond à une nouvelle question posée au cycle 6.

- 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.
- 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; 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.

5 Il convient de noter que, bien qu'il y ait des variables pour une diversité d'unités d'analyse (PMR, conjoint/partenaire et ménage), on ne pourra produire que des « estimations pour les enfants » à partir du fichier de micro données 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 de données de l'ELNEJ.

Nom de la variable se rapporte à :

Nom de la variable	Objet :
FLFSQ2	Q2 de la section sur l'activité professionnelle pour le conjoint/partenaire
F	Variable du cycle 6
LF	Section sur l'activité professionnelle
S	Conjoint/partenaire
Q	Question directement posée dans le questionnaire
2	Deuxième question de la section sur l'activité professionnelle au questionnaire de l'adulte

Nom de la variable	Objet :
FPRCS03	Note d'interaction positive à l'échelle du rôle parental pour un enfant de 2 à 15 ans
F	Variable du cycle 6
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 employés pour les diverses sections des questionnaires de l'ELNEJ. Ces acronymes font partie du nom des variables du fichier de données de l'ELNEJ. Ils occupent les deuxième et troisième positions du nom de variable.

Acronyme	Variable	Recueillie ou générée à partir de :
GE	Renseignements géographiques	Données d'échantillon
HH	Ménage	Caractéristiques du logement
MM	Variables d'observation de la liste du ménage	Variables démographiques de base pour chaque membre du ménage; celles-ci figurent au fichier de données de l'ELNEJ pour l'enfant, le jeune, la PMR et le conjoint/partenaire.
DM	Variables démographiques calculées qui décrivent les conditions de vie de l'enfant	Données de la liste du ménage et de la grille des liens
SD	Renseignements sociodémographiques	L'enfant au questionnaire de l'enfant et de la PMR et conjoint/partenaire au questionnaire de l'adulte
HL	Santé	La PMR et conjoint/partenaire au questionnaire de l'adulte; de l'enfant au questionnaire de l'enfant; et du jeune de 16 ans et plus au questionnaire du jeune.
CH	États chroniques des adultes	La PMR et conjoint/partenaire à la section sur la santé du questionnaire de l'adulte
RS	Limitations d'activités	La PMR et conjoint/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	Scolarité	L'enfant de 4 et 5 et 10 à 15 ans au questionnaire de l'enfant; la PMR et conjoint/partenaire au questionnaire de l'adulte; et du jeune de 16 ans et plus au questionnaire du jeune
LF	Activité professionnelle	La PMR et conjoint/partenaire au questionnaire de l'adulte; et du jeune de 16 ans et plus au questionnaire du jeune.
IN	Revenu	Revenu du ménage et revenu personnel de la PMR au questionnaire de l'adulte; et celui du jeune de 16 ans et plus au questionnaire du jeune
FN	Fonctionnement de la famille	Questionnaire de l'adulte (administré à la PMR ou au conjoint/partenaire)
MD	Renseignements médicaux et biologiques	Questionnaire de l'enfant (de 0 à 3 ans)
TM	Tempérament	Questionnaire de l'enfant (3 à 35 mois)
LT	Alphabétisation	Questionnaire de l'enfant (0 à 5 ans)
AC	Activités	Questionnaire de l'enfant (0 à 5 et 10 à 15 ans) et du jeune de 16 ans et plus au questionnaire du jeune
BE	Comportement	Questionnaire de l'enfant (0 à 5 et 10 et 11 ans)
MS	Développement moteur et social	Questionnaire de l'enfant (0 à 47 mois)
WB	Travail après la naissance	Questionnaire de l'enfant (0 à 5 et 10 à 15 ans)
RL	Rapports sociaux	Questionnaire de l'enfant (4 et 5 ans) et du jeune de 18 ans et plus au questionnaire du jeune
PR	Style parental	Questionnaire de l'enfant (0 à 5 et 10 à 15 ans)

Acronyme	Variable	Recueillie ou générée à partir de :
CR	Garde des enfants	Questionnaire de l'enfant (0 à 5 et 10 à 13 ans)
SL	Sommeil	Questionnaire de l'enfant (0 à 13 ans)
PB	Comportement positif	Questionnaire de l'enfant (3 à 5 ans)
CM	Échelle de communication	Questionnaire de l'enfant (3 à 5 ans)
ML	Grandes étapes du développement	Questionnaire de l'enfant (9 à 47 mois)
CS	Garde légale	Questionnaire de l'enfant (0 à 5 et 10 à 17 ans)
AG	Étapes du développement	Questionnaire de l'enfant (3 à 47 mois)
SF	Sécurité du voisinage	Questionnaire de l'adulte (section administrée à la PMR ou au conjoint(e)/partenaire)
SP	Soutien social	Questionnaire de l'adulte (section administrée à la PMR ou au conjoint(e)/partenaire) et le jeune de 18 ans et plus dans le questionnaire du jeune
AS	Aspirations	Questionnaire de l'enfant (16 et 17 ans) et du jeune 18 ans et plus dans le questionnaire du jeune
EQ	Quotient émotionnel	Questionnaires auto administrés des 10 à 17 ans et du jeune de 20 et 21 ans dans le questionnaire du jeune
AM	Ce que je suis, moi	Questionnaires auto administrés des 10 à 17 ans et du jeune de 18 et 19 ans dans le questionnaire du jeune
FB	Sentiments et comportements	Questionnaires auto administrés des 10 à 17 ans et du jeune de 18 ans et plus dans le questionnaire du jeune
MO	Départ du foyer familial	Questionnaire du jeune (18 ans et plus)
FF	Amis et famille	Questionnaires auto administrés des 10 à 17 ans
SC	École	Questionnaires auto administrés des 10 à 15 ans
PM	Mes parents et moi	Questionnaires auto administrés des 10 à 17 ans
PU	Puberté	Questionnaires auto administrés des 10 à 17 ans
DR	Usage du tabac, de l'alcool et des drogues	Questionnaires auto administrés des 10 à 17 ans
AT	Activités	Questionnaires auto administrés des 10 à 15 ans
HT	Santé	Questionnaires auto administrés des 12 à 17 ans
WK	Travail et sources d'argent	Questionnaires auto administrés des 12 à 15 ans
DA	Fréquentations	Questionnaires auto administrés des 12 à 17 ans
PP	Test EVIP	Enfants de 4 et 5 ans
WM	Qui suis-je?	Mesure directe (4 à 5 ans)
KN	Connaissance des nombres	Mesure directe (4 à 5 ans)
MA	Test de calcul	Enfants de 8 à 15 ans étant de la 2 ^e à la 10 ^e année scolaire et exercices de résolution de problèmes aux 16 et 17 ans
LI	Évaluation de la littératie	Jeunes de 18 à 19 ans
NU	Évaluation de la numératie	Jeunes de 20 à 21 ans
WT	Poids	Poids faisant partie de la méthodologie de l'échantillon

7.16 Fichiers de traitement final et fichiers principaux

Les fichiers suivants ont été créés pour le cycle 6 de l'ELNEJ :

- la cohorte du développement de la petite enfance : 0 à 5 ans (adulte, enfant, ménage)
- la cohorte longitudinale : 10 à 17 ans (adulte, enfant, ménage)
- la cohorte longitudinale : 16 à 21 ans (jeune et ménage; variables statiques reportées des cycles précédents)
- les questionnaires auto administrés
- nord (enfants de 5 ans provenant du Yukon et du Nunavut)

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 pour 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. Pour 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 a résulté en grande partie d'un effort concerté soutenu par la contribution et les conseils :

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

Le groupe 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 sa vie à l'école. La présente section donne un aperçu du contenu de chacune des sections du questionnaire ayant servi à la collecte des données de l'ELNEJ. Les différentes échelles utilisées dans le cadre de l'ELNEJ seront brièvement examinées dans ce chapitre, mais pour en savoir davantage sur ce sujet ou sur la validation des résultats obtenus sur les échelles, veuillez consulter le 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 | Cette composante constitue la première partie de l'interview. On recueille, pour la liste des membres du ménage, des renseignements démographiques de base sur chacun des membres et des données sur les liens qui l'unissent à tous les autres membres du ménage. |
| Adulte | Questions relatives à la personne la mieux renseignée (PMR) et son conjoint. Dans le cas des enfants âgés de 16 et 17 ans, on ne pose pas les questions de toutes les sections de la composante des adultes. La composante des adultes est remplie une seule fois, même lorsque le ménage comprend deux enfants. Aucune composante adulte n'est produite pour les jeunes de 18 ans et plus. |
| 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 5 ans et de 10 à 17 ans. Les seules sections du questionnaire de l'enfant qui s'applique aux jeunes âgés de 16 et 17 ans sont les suivantes : aspirations et attentes, garde légale des enfants, et les caractéristiques sociodémographiques. |
| Jeune | Questions sur le jeune sélectionné âgé de 16 à 21 ans. Dans cette section, les jeunes répondent eux-mêmes aux questions qui les concernent dans une interview assistée par ordinateur (IAO). |
| Questionnaire auto administré | Les jeunes âgés de 10 à 17 ans répondent eux-mêmes aux questions qui les concernent sur un questionnaire en format imprimé. |

Évaluations directes

On effectue plusieurs évaluations directes relativement aux enfants et aux jeunes; celles-ci sont décrites au chapitre 14.0.

8.2 Variables démographiques

Les variables démographiques sont recueillies dans le questionnaire du ménage. À partir de ce questionnaire, 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.

Il a fallu procéder à toute une série de vérifications des données recueillies. Voici des exemples des vérifications effectuées :

- un parent biologique doit avoir au moins 12 ans de plus (et au maximum 55 ans de plus) que l'enfant biologique;
- la différence d'âge entre mari et femme doit être inférieure à 29 ans.

8.3 Questionnaire de l'adulte

Éducation (parent)

La section sur le niveau de scolarité est remplie pour la PMR et son conjoint/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 montré 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/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 montrent que le chômage des parents peut avoir un effet néfaste sur les résultats des enfants.

La section relative à l'activité sur le marché du travail est remplie à la fois pour la PMR et son conjoint/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 (par exemple, travail par quarts) au cours de l'année précédente. On recueille des renseignements portant sur l'emploi actuel ou l'emploi le plus récent de la PMR et de son conjoint/partenaire.

Une description complète de l'emploi actuel ou de l'emploi le plus récent est consignée. La branche d'activité et la profession sont codées selon le Système de classification des industries de l'Amérique du Nord (SCIAN) de 2002, la Classification type des professions de 1991, 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/partenaire. Parmi ces variables figurent les suivantes :

- FLFPfD5A / FLFSfD5A : code du SCIAN 2002 pour l'emploi actuel de la PMR / de son conjoint
- FLFPfD6A / FLFSfD6A : code de la CNP-S 2001 pour l'emploi principal de la PMR / de son conjoint
- FLFPfD7A / FLFSfD7A : code de la classification type des industries pour l'emploi actuel (SCIAN 2002) – regroupé
- FLFPfD8A / FLFSfD8A : 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. Il y a également quelques questions portant sur la perception qu'ont la PMR et son conjoint de 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 (FINHD04A) afin de comparer le revenu du ménage au seuil de faible revenu (SFR)⁶. 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 FINHD03A donne la valeur du SFR selon la zone géographique.

Santé (adulte)

Cette section comporte des questions adressées aux PMR et à leur conjoint au sujet de leur état général de santé, les états chroniques et les limitations d'activités, le tabagisme et 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. La consommation d'alcool est abordée 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 (p. ex., allergies, asthme, hypertension). Une variable dérivée (FCHPD01 ou FCHSD01) indique une réponse affirmative, c'est-à-dire que les répondants ont déclaré au moins un état chronique.

Limitations d'activités

On pose aux PMR et à leur conjoint une série de questions sur les limitations de leurs activités à la maison, au travail, à l'école, etc. Une variable dérivée (FRSPdD01 ou FRSSdD01) permet de déterminer si la PMR ou son conjoint ont déclaré des limitations d'activités.

⁶ Pour de plus amples renseignements sur les mesures utilisées par Statistique Canada en matière de faible revenu, veuillez consulter *Seuils de faible revenu de 2005* et *Mesures du faible revenu de 2004*, N° 75F0002MIF2006004 au catalogue.

Antécédents maternels

Cette section permet de cerner les antécédents gravidiques de la mère. Ces questions sont posées seulement à celles qui sont interviewées pour la première fois. Les questions portant sur la grossesse et l'accouchement ont été élaborées par le D^r 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 (FDPPS01) dans le cadre du 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.

Sécurité du quartier

Cette section permet de recueillir des renseignements sur la satisfaction des répondants quant à leur quartier en tant que milieu de vie pour les enfants, y compris leurs perceptions relativement à l'importance des dangers et des problèmes, à la cohésion sociale ou sentiment de « voisinage ». Deux échelles ont été créées dans cette section : celle de la sécurité du quartier (FSFHS5), qui mesure le degré de sécurité du quartier tel que le perçoivent les répondants, et l'échelle du voisinage (FSFHS6), qui mesure le degré de cohésion sociale dans le quartier.

Nota : Cette section est omise dans l'Enquête auprès des enfants du Nord.

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 et le profil linguistique des membres du ménage. Ces données permettent l'analyse des diverses composantes de la population canadienne et l'identification des personnes appartenant à une minorité visible. La section comporte en outre une question sur l'appartenance religieuse. Ces questions se rapportent à la PMR, à son conjoint et à l'enfant.

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 » (FMMCQ01) 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 6, l'âge effectif est calculé en fonction de l'année 2004, c'est-à-dire 2004 moins l'année de naissance de l'enfant. Par exemple, un enfant né en 2000 aurait un âge effectif de 4 ans (2004-2000). À noter que l'âge réel de l'enfant au moment de l'interview est parfois différent de l'âge effectif.

Éducation (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 scolaire 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 plus avancés, on pose des questions supplémentaires sur d'autres aspects de l'expérience éducative, comme le saut d'une classe et 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 et 5 ans. Si les enfants ne peuvent pas y répondre (i.e. s'il ne parlent ni l'anglais ni le français ou s'ils sont daltoniens), on ne leur fera pas passer les tests.

Nota : Cette section est omise dans l'Enquête auprès des enfants du Nord.

Santé (enfant)

Cette section vise à recueillir des renseignements sur l'état de santé physique de l'enfant – état général de santé, 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 ou 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. On calcule, à partir de ces renseignements, l'Indice de l'état de santé (IES3) (FHLCcD2A). 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 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 huit 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 3 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 de développement persistants.

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.

Deux variables dérivées créées dans cette section méritent d'être soulignées. Ces deux variables servent à indiquer l'âge gestationnel de l'enfant. FMDCD06 exprime cet âge en jours et FMDCD07 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 (FMDCD08) permettant d'indiquer si le poids de l'enfant à la naissance est normal (2 500 grammes), 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 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 (QAS) sont des outils de déclaration des parents, mis au point par Jane Squires, LaWanda Potter et Diane Bricker de 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, par 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.

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 QAS pour les enfants âgés de 3 à 47 mois inclusivement. Après avoir consulté l'éditeur, Statistique Canada a converti les questionnaires pour en faciliter l'administration de l'IAO (interviews assistées par ordinateur). L'aspect « motricité globale » des QAS n'est pas inclus parce qu'il se trouve déjà dans d'autres sections de l'enquête.

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 le contact qu'on les enfants avec les livres, leur intérêt pour la lecture et les activités d'apprentissage auxquelles les parents s'adonnent avec leurs enfants. La section porte principalement sur la stimulation des jeunes enfants reçue à la maison.

Pour les enfants âgés 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 âgés de 3 à 5 ans, ces questions étant adaptées aux activités correspondant à ce groupe d'âge. On a ajouté une question sur les activités de calcul, conformément aux recommandations de collègues de D^r Robbie Case, l'auteur du test sur la connaissance des nombres, selon qui les capacités de calcul constituent un facteur déterminant en ce qui a trait à l'apprentissage et à l'alphabétisation. On a ajouté cette question pour obtenir des renseignements sur les activités de calcul des enfants susceptibles de confirmer les résultats du test de connaissance des nombres.

Communication

Ces questions sont adaptées à partir 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 âgés de 4 et 5 ans dans le cadre de l'Indice de l'état de santé de la section relative à la santé.

Activités

Cette section permet de mesurer la participation des enfants à diverses activités non scolaires ainsi que les responsabilités qu'assument les enfants âgés de 10 à 13 ans à la maison. Ces questions servent à établir l'échelle des responsabilités familiales (FACCS06) qui mesure le degré de responsabilités assumées à la maison. Cette section donne certaines indications quant à l'emploi du temps de l'enfant, ses intérêts personnels et son degré d'interaction avec ses pairs.

On a ajouté plusieurs questions destinées aux enfants de 4 et 5 ans afin de déterminer la fréquence de certaines activités que font les parents avec leurs enfants, 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.

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 âgés de 1 an à 3 ans.

Les questions de cette section servent à mesurer la fréquence de comportements tels que l'hyperactivité et l'agressivité physique. Les échelles établies à partir de ces questions sont décrites de façon détaillée au chapitre 9.0.

Comportement positif

Cette section vise à évaluer le comportement positif des enfants âgés 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 de l'ELNEJ se rapportant au comportement des enfants d'autres groupes d'âge.

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, etc.

Développement moteur et social

L'échelle du développement moteur et social (DMS) mesure divers aspects du développement moteur, social et cognitif des enfants âgés de 0 à 3 ans; les questions varient selon l'âge des

enfants. Trois scores (FMSCS01, FMSCS02 et FMSCdS03) 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 des 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 les relations avec les parents, les enseignants et les amis sont fondées sur celles 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 demandé certaines questions aux PMR ayant un conjoint/partenaire à la maison pour 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, âge de l'enfant au moment de la séparation ou du divorce, etc.

Garde d'enfants

Cette section fournit des renseignements de base sur les services de garde actuels des enfants pendant que les parents travaillent ou étudient, 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 sans but 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 ».

Caractéristiques sociodémographiques

Ces questions visent à recueillir des renseignements sociodémographiques sur les enfants sélectionnés, par exemple leur origine ethnique et leur pays d'origine. Au Cycle 6, deux questions ont été ajoutées au sujet de l'identité Autochtone.

Aspirations et attentes (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. L'aide aux devoirs, la discussion de l'expérience scolaire et des 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

Ces questions visent à recueillir des renseignements sociodémographiques sur les enfants sélectionnés, par exemple leur origine ethnique et leur pays d'origine. Au cycle 6, deux questions ont été ajoutées au sujet de l'identité Autochtone.

8.5 Questionnaire du jeune (16 à 21 ans)

Départ du foyer parental (18 à 21 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.

Cette section vise à recueillir des renseignements sur le nombre de fois que les répondants de 18 à 21 ans ont quitté leur foyer. 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 l'information qui aura été recueillie durant la composante « Ménage ».

Les questions ont été élaborées en consultation avec D^{re} Dianne Looker (directrice du département de sociologie d'Acadia University). D^{re} Looker 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*.

Éducation des jeunes (16 à 21 ans)

Cette section recueille de l'information sur l'expérience scolaire des jeunes. Les premières questions servent à établir le niveau d'instruction actuel. Les questions qui s'appliquent à sa 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) et
- 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.

Au cycle 5, on a élaboré une section intégrée sur les études. 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 en fonction non de leur âge mais de leur situation par rapport aux études, passée et présente.

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 auto administrés des 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.

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 le jeune (16 et 17 ans). Ce sont notamment les suivantes : FLFYfD5A, FLFYfD6A, FLFYfD7A, FLFYfD8A, et FLFYeD02.

Activité des jeunes sur le marché du travail (18 à 21 ans)

Les questions dans cette section sont similaires à celles posées aux jeunes de 16 et 17 ans, mais il y en a encore plus prises de la section auprès l'activité des adultes sur le marché du travail, et cela, parce que les répondants sont maintenant plus vieux. Pour quelques-uns parmi les jeunes, le travail peut être l'activité principale et les questions doivent tenir compte de cette possibilité.

Ces questions permettent de recueillir des renseignements qui aideront à se faire une idée d'ensemble de 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, une description complète de l'emploi actuel ou de l'emploi le plus récent est consignée. La branche d'activité et la profession sont codées selon le SCIAN de 2002 et la CNP-S de 2001.

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 le jeune (18 à 21 ans). Ce sont notamment les suivantes : FLYYfD5A, FLYYfD6A, FLYYfD7A et FLYYfD8A.

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 dépendant de 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 D^r Dianne Looker d'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, de ces renseignements, on a créé la variable FASYfD03.

Revenu des jeunes (16 et 17 ans)

La section sur le revenu des jeunes se rapporte au revenu que les jeunes ont tiré de diverses sources au cours des 12 derniers mois. Ces questions ressemblent à celles posées aux parents. Les renseignements sur le revenu permettent de mesurer l'argent de poche dont disposent les jeunes pour prendre leurs propres décisions.

Revenu des jeunes (18 à 21 ans)

Il s'agit d'une nouvelle section introduite au cycle 6 pour ce groupe d'âge. Les questions ressemblent 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.

Il y a une série de questions concernant le paiement des dépenses d'habitation. À cet âge, il se peut que bien des jeunes quittent le foyer parental pour la première fois. Déterminer s'ils paient leurs frais de logement est un renseignement important qui permet d'évaluer comment ils s'adaptent à leurs responsabilités financières.

Santé des jeunes (16 à 21 ans)

Cette section porte sur l'état général de santé, 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. De nouvelles questions sur le sommeil ont été ajoutées. 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, etc.

Santé des jeunes (18 à 21 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. Au cycle 6, ces questions sont passées des questionnaires auto administrés à l'application d'IAO.

En outre, l'échelle de dépression des questionnaires auto administrés est comprise dans cette section (FHTYfS01). 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)

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* de 1992 de la Colombie-Britannique et ressemblent aux questions auxquelles doivent répondre les plus jeunes répondants dans les questionnaires auto administrés.

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 des gangs.

Activités des jeunes (16 à 21 ans)

À l'âge jeune adulte, les jeunes 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 et 17 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 sur ce sujet.

D'autres questions portent sur la mesure dans laquelle le jeune fait du bénévolat dans sa collectivité. Le cycle 6 comprend aussi une nouvelle question qui demande au jeune s'il a voté aux dernières élections municipales, provinciales ou fédérales. On cherche ainsi à déterminer dans quelle mesure le jeune participe dans la société.

Relations (18 à 21 ans)

Au fur et à mesure que le jeune fait son entrée dans la vie adulte, la nature de ses relations avec un partenaire/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 21 ans selon leur âge et leur état matrimonial.

Quartier (16 et 17 ans)

Les facteurs relatifs au quartier ont une incidence démontrée sur la situation des enfants et des jeunes à plusieurs égards (p. ex., rendement scolaire, comportement, fonctionnement émotif et social, développement moteur et social). Cette incidence s'accroît à mesure que les enfants franchissent les étapes de la vie et intensifient leurs interactions et leurs contacts avec le milieu extrafamilial. C'est ce qu'ont démontré des ouvrages universitaires ainsi que des recherches fondées sur les données de l'ELNEJ (Boyle et Lipman, Kohen et al., Offord et Lipman).

À propos de moi (18 et 19 ans)

L'objectif des questions au début de la section est de déterminer le niveau d'estime de soi du jeune. Les réponses permettent d'établir un score global d'estime de soi (FAMYfS01).

De plus, des questions supplémentaires portent sur les expériences douloureuses que le jeune peut avoir vécues pendant les deux dernières années, comme la rupture avec son ami(e), 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 ressemblent à celles des questionnaires que doivent remplir les enfants plus jeunes.

Quotient émotionnel (20 et 21 ans)

L'échelle du quotient émotionnel a été mise au point par Reuven BarOn, Ph.D et James D.A. Parker, Ph.D., pour les jeunes qui avaient 20 et 21 ans au cycle 6. 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 et ceux des autres. On a posé 20 questions aux répondants portant sur leurs sentiments, leurs émotions et leurs perceptions. Cette version s'adressant aux jeunes adultes ressemble aux 15 questions du questionnaire que devaient remplir eux-mêmes les jeunes de 10 à 17 ans.

Soutien social des jeunes (18 à 21 ans)

Les questions de cette section sont posées aux jeunes de 18 ans et plus dans le cadre de l'application IAO. Variant selon l'âge du jeune, elles visent à recueillir des renseignements sur le réseau de soutien social des jeunes. Cette mesure a été ajoutée pour garantir l'adéquation avec l'âge du répondant. Les questions servent à déterminer le soutien social reçu de la part de la famille et des amis tel qu'il est perçu par le jeune.

Cette section contient aussi quelques questions s'adressant aux personnes de 20 et 21 ans et portant sur le nombre d'enfants à charge qu'ils ont ainsi que sur leurs relations avec leur mère/père. Des renseignements sont ainsi recueillis sur le réseau familial et sur le rôle qu'il joue dans la vie du jeune.

8.6 Questionnaires auto administrés (10 à 17ans)

On vise avec ces questionnaires à recueillir des renseignements directement auprès des jeunes au sujet de 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 ces jeunes peuvent fournir de façon fiable. Dans le cas des jeunes de 16 et 17 ans, certaines données sont toujours recueillies au moyen des questionnaires auto administrés, même si ces jeunes fournissent leurs propres renseignements dans le cadre de l'interview assistée par ordinateur. Nous étions d'avis que les jeunes se sentent plus à l'aise de répondre à des questions délicates sur un questionnaire imprimé que dans le cadre d'une rencontre avec un intervieweur.

Parents et amis (10 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 l'enfant 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 (FFCS01) est établie à partir de ces questions.

Cette section contient également une mesure de l'intimité des 14 et 15 ans. La question concernant combien de fois 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 (10 à 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 de l'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* de l'Australie-Occidentale, la *Health Attitudes, Knowledge and Behaviours Study* des 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 (10 à 17 ans)

Ces questions servent à déterminer l'estime de soi générale des jeunes. On calcule un score (FAMcS02) à partir des réponses à ces questions.

On pose aux jeunes de 12 à 17 ans des questions supplémentaires sur leurs sentiments face à la vie, maintenant et à l'avenir. Ces questions sont tirées de la *Child Health Survey* de 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, Ph.D et James D.A. Parker, Ph.D. Il s'agit de la version pour les jeunes des questions posées aux personnes de 20 et 21 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 âgés de 14 à 17 ans des questions sur des expériences douloureuses, comme la rupture avec leur ami(e) ou la mort d'un proche.

Sentiments et comportements (10 à 17 ans)

Liste de comportements (10 à 15 ans)

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

Comportements à risque (10 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 des 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* d'University of Ohio, la *Child Health Survey* de l'Australie-Occidentale et de questions fournies par D^r Richard Tremblay de l'Université de Montréal.

Suicide (12 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* de 1992 de la Colombie-Britannique.

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 21 ans. Un score (FHTCbS1b) est calculé à partir de ces questions.

Mes parents (10 à 17 ans)

Cette section a pour but de saisir la relation des jeunes avec leurs parents/tuteurs sous différents angles. Les questions sont conçues pour savoir dans quelle mesure ils sont traités avec compréhension, équité et affection par chaque parent/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 des relations de leurs parents/tuteurs et de leurs aptitudes à résoudre les conflits.

Mes parents et moi (10 à 15 ans)

Trois échelles sont établies à partir de ces questions :

- 1) les soins prodigués par les parents (FPMCcS1),
- 2) le rejet par les parents (FPMCbS2b) et
- 3) la surveillance par les parents (FPMCcS3).

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 (FPMCdS4) et
- 2) l'échelle de résolution des conflits – Père (FPMCdS5).

Puberté (10 à 17 ans)

La puberté est un marqueur important du développement physique. Dans cette section, on pose aux jeunes des questions concernant leurs indicateurs physiologiques clés et leur propre perception de leur puberté. Ces questions ont été élaborées par D' R. Tremblay de l'Université de Montréal.

Nota : Pour les jeunes de 12 à 17 ans, ces questions se trouvent à la section sur la santé.

Tabac, alcool et drogues (10 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 de l'OMS et de la *Child Health Survey* de l'Australie-Occidentale.

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

Les questions sur la consommation de drogues et de substances toxicomanogènes sont adaptées de la *Health Attitudes, Knowledge and Behaviours Study* des 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 en état d'ébriété ont été ajoutées pour les jeunes âgés de 16 et 17 ans. Il s'agit là d'importants comportements à risque dans ce groupe d'âge. Les questions sont adaptées de l'*Evaluation of School-Based Health Centers* de la Caroline du Nord.

Activités (10 à 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. En règle générale, les activités sont également couvertes dans le questionnaire de l'enfant de l'IAO rempli par le parent pour les enfants de moins de 10 ans.

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 lecture, 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 21 ans au moyen du questionnaire de l'IAO.

Santé (12 à 17 ans)

On demande aux jeunes de fournir des renseignements sur leur taille et leur poids, les symptômes de stress, le port de la ceinture de sécurité et d'un casque protecteur, 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 de 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* des États-Unis qui ont été aussi employées dans l'*Adolescent Health Survey* de 1992 de la Colombie-Britannique.

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* et de l'*Adolescent Health Survey* de 1992 de la Colombie-Britannique.

Travail et argent (12 à 15 ans)

On a posé aux jeunes des questions sur le travail pendant l'année scolaire et, dans le cas des jeunes âgés de 14 et 15 ans, sur le travail l'été précédent. Les jeunes de 14 et 15 ans doivent répondre à des questions plus détaillées sur leurs emplois, par exemple sur les heures de travail et la rémunération. On leur demande également si le travail réduit le temps qu'ils consacrent à leurs études.

On demande aussi aux jeunes d'indiquer le montant qu'ils reçoivent de diverses sources, comme les parents et le travail.

Ces questions ont été élaborées par l'équipe de projet à la suite de l'examen de plusieurs enquêtes.

Relations amoureuses (12 à 17 ans)

Dans cette section, on pose aux jeunes des questions sur leurs expériences avec un copain ou une copine et sur leurs activités sexuelles. La question sur le comportement sexuel dans le questionnaire des 12 et 13 ans a été adaptée à partir de l'Enquête sur les jeunes et le sida. On pose également des questions sur l'utilisation des moyens de contraception et, pour les 16 à 17 ans, sur les raisons de leur abstinence du sexe ou de leur refus d'utiliser des moyens contrôle des naissances. Ces questions ont été élaborées par l'équipe de projet en collaboration avec des experts d'enquêtes sur les jeunes comme l'*Adolescent Health Survey* de 1992 de la Colombie-Britannique 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 dont les réponses mises ensemble permettent de mesurer certains concepts.

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

9.1.2 Échelles et calculs

Pour chaque facteur mesuré au moyen d'une échelle, on calcule un score (note). 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, etc.). Le score pour 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 autant que possible 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 exprimées selon une 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 exprimées selon une échelle, afin de s'assurer que les propriétés psychométriques présentées dans les autres études s'appliquaient à l'ELNEJ.

L'analyse des données exprimées selon 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 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 employées pour chacune de ces étapes sont décrites en détail ci-dessous.

Nota : 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 les 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 pour les échelles

La détermination de la structure factorielle de chacune des échelles s'est faite à partir des données du premier cycle. La structure factorielle imposée aux échelles employées au cycle 1, puis réutilisées 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, une analyse des composantes principales a été effectuée 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 de l'enfant. On normalise le poids statistique d'un individu en divisant son poids (`_WTCW01C`)⁷ 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 déterminé, on procède au calcul des scores. Pour produire les scores, on a retranché 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 problèmes chez l'enfant pour tout facteur 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

⁷ 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 « FPRCS01 » au cycle 6 débute avec la lettre « F » au fichier de microdonnées et ici on en fait référence comme « _PRCS01 ».

si le parent n'a pas voulu répondre ou s'il ne connaissait pas la réponse à la question.

9.2.2 Transformation des données au moyen de l'échelonnement optimal

Avant de procéder à l'analyse factorielle des échelles de l'ELNEJ, les données ont été transformées au moyen de l'échelonnement optimal. La méthode utilisée était l'une de celles proposées par Young et plusieurs collaborateurs (Young, 1981), qui constitue une variante de la méthode d'échelonnement 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.

9.2.3 Analyse factorielle axée sur les données pondérées

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.

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 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 de facteurs

Dans le cas de l'échelle relative au rôle parental dans le questionnaire de l'enfant, un des concepts qui est ressorti des analyses de facteurs est celui de l'inefficacité parentale (pour les enfants âgés de 2 à 11 ans). 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/ée dire ou faire?
_PRCQ08	Lorsque vous parlez à [votre enfant] de son comportement, dans quelle proportion du temps le/la félicitez-vous?
_PRCQ09	Lorsque vous lui parlez de sa conduite, dans quelle proportion du temps le/la désapprouvez-vous?
_PRCQ13	À quelle fréquence vous mettez-vous en colère lorsque vous punissez [votre enfant]?
_PRCQ14	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/elle en général?

_PRCQ18 À quelle fréquence devez-vous le/la 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 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 signalées 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/la 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, les réponses (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 Code de non-réponse

Le score pour le facteur d'inefficacité parentale est étiqueté _PRCS04 dans le cliché d'article du fichier de microdonnées. Le « S » qui figure au cinquième caractère du nom de la variable indique qu'il s'agit d'un score.

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 le nombre de questions comportant un code de non-réponse se situait au-dessus d'un certain seuil, le score du facteur était établi à « non déclaré ».

9.3.5 Items bruts

Il convient de souligner qu'outre les scores, les items bruts pour chaque échelle figurent dans le fichier de microdonnées. Cela permettra aux chercheurs d'opter pour d'autres structures de facteurs s'ils le souhaitent. Dans le cas des items bruts, les valeurs

originales (variant de 1 à 5 pour l'échelle relative au rôle parental), avant le rééchelonnement ou le renversement des valeurs, ont été retenues.

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'échelle 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 Interprétation 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 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 qui pourrait 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 largement 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 au 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 des facteurs et de l'analyse de la fiabilité. Les modifications apportées aux échelles d'un cycle à l'autre sont également décrites. Pour des raisons de commodités, les échelles sont présentées en ordre alphabétique. Le tableau ci-dessous donne un bref résumé des échelles de l'ELNEJ de même qu'une description de chacune des échelles.

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) de 20 questions 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 précédente. L'échelle a été réduite à 12 questions par D' M. Boyle, de Chedoke-McMaster Hospital de 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 expert 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 général de santé mentale. On a proposé que cette section soit centrée sur la dépression pour les raisons suivantes : il s'agit d'un état prévalent; 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.

Score	Items inclus	Taille de l'échantillon	Univers (âge en ans)	Non-réponse à l'item	Alpha de Cronbach (standardisé)
Dépression - PMR (_DPPS01)	_DPPQ12A,	3 178	0 à 1	7,2 à 8,0 %	0,812
	_DPPQ12B,				
	_DPPQ12C,	3 312	2 à 3	3,8 à 4,0 %	0,835
	_DPPQ12D,				
	_DPPQ12E,	3 364	4 à 5	4,0 à 4,2 %	0,827
	_DPPQ12F*,				
	_DPPQ12G,	2 764	10 à 11	1,7 à 1,8 %	0,857
	_DPPQ12H*,				
	_DPPQ12I,	1 950	12 à 13	1,7 à 2,0 %	0,859
	_DPPQ12J*,				
	_DPPQ12K,	1 756	14 à 15	1,5 à 1,6 %	0,862
	_DPPQ12L				

*Indique que les valeurs ont été renversées.

9.5.2 Échelle des responsabilités familiales

Objectifs

L'échelle des activités a pour but de mesurer la participation de l'enfant aux responsabilités familiales. Cet ensemble de questions sur les responsabilités est tiré du questionnaire abrégé du *Home Observation for Measurement of the Environment* de la *National Longitudinal Survey of Youth* menée par l'University of Ohio.

Au cycle 1, on n'a posé ces questions qu'aux enfants de 10 et 11 ans parce qu'ils constituaient le plus vieux groupe d'âge. Aux cycles ultérieurs, on a posé ces questions à tous les enfants de 10 à 13 ans.

Score	Items inclus	Taille de l'échantillon	Univers (âge en ans)	Non-réponse à l'item	Alpha de Cronbach (standardisé)
Responsabilités familiales (_ACCS6)	_ACCQ6A, _ACCQ6B,	2 811	10 à 11	0,4 %	0,794
	_ACCQ6C, _ACCQ6D, _ACCQ6E	1 988	12 à 13	0,4 %	0,794

9.5.3 Échelle de fonctionnement de la famille

Objectifs

Les questions sur le fonctionnement de la famille (_FNHQ01A à _FNHQ01L) ont été élaborées par les chercheurs de Chedoke-McMaster Hospital de McMaster University et sont largement 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 _FNHQ01M, tirée du suivi de l'Étude sur la santé des jeunes ontariens, a été ajoutée à l'échelle originale pour 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 cette échelle est d'obtenir une évaluation globale du fonctionnement de la famille et une indication de la qualité des relations entre les parents ou partenaires. 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 ou à son conjoint/partenaire dans le cadre du 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 (_FNHQ01A, _FNHQ01C, _FNHQ01E, _FNHQ01G, _FNHQ01I et _FNHQ01K). Le score total (_FNHS01) peut donc varier de 0 à 36, un score élevé dénotant un dysfonctionnement de la famille.

Score	Items inclus	Taille de l'échantillon	Univers (âge en ans)	Non-réponse à l'item	Alpha de Cronbach (standardisé)
Fonctionnement de la famille (_FNHS01)	_FNHQ01A*, _FNHQ01B,	3 232	0 à 1	4,7 à 5,8 %	0,915
	_FNHQ01C*, _FNHQ01D,	3 353	2 à 3	2,1 à 2,3 %	0,909
	_FNHQ01E*, _FNHQ01F,	3 439	4 à 5	1,5 à 1,8 %	0,910
	_FNHQ01G*, _FNHQ01H,	2 775	10 à 11	1,0 à 1,1 %	0,916
	_FNHQ01I*, _FNHQ01J,	1 968	12 à 13	0,3 à 0,6 %	0,907
	_FNHQ01K*	1 764	14 à 15	0,6 à 0,7 %	0,906

9.5.4 Échelle de sécurité du quartier

Objectifs et aperçu

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

Les questions _SFHQ01, _SFHQ02, et _SFHQ05A à _SFHQ06E couvrent la durée de résidence dans le quartier, la satisfaction à l'égard du quartier en tant que 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 D^{re} Jacqueline Barnes dans ses travaux sur des quartiers de Boston et de Chicago. Ce questionnaire a été modifié, en consultation avec D^{re} Jacqueline Barnes, à la lumière de l'analyse des facteurs des sections. La question _SFHQ03 sur l'engagement bénévole s'inspire d'une question de l'Enquête nationale sur la santé de la population.

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

Ces échelles ont été utilisées par intermittence au cours des quatre cycles de l'ELNEJ. Au cycle 1, trois échelles ont été créées : sécurité du quartier (_SFHQ05A à _SFHQ05C), voisinage (_SFHQ06A à _SFHQ06E) et problèmes du quartier. Au cycle 2, aucune question de la section relative au quartier n'a été posée aux répondants, de sorte qu'aucune échelle n'a été créée. Au cycle 3, la section se rapportant au quartier a été réintroduite, mais sans les questions _SFHQ05A à _SFHQ05C et _SFHQ07A à _SFHQ07F.

Les questions du cycle 6 sont identiques à celles du cycle 1, à l'exception de la question _SFHQ05C, dont on a légèrement modifié le libellé. Par ailleurs, les questions rattachées aux problèmes du quartier dans le cycle 1 (_SFHQ07A à _SFHQ07F) n'ont pas été reprises après le cycle 4.

Score	Items inclus	Taille de l'échantillon	Univers (âge en ans)	Non-réponse à l'item	Alpha de Cronbach (standardisé)
Sécurité du quartier (_SFHS5)	_SFHQ05A,	3 251	0 à 1	4,8 à 5,0 %	0,673
	_SFHQ05B,	3 354	2 à 3	2,0 à 2,1 %	0,672
	_SFHQ05C	3 436	4 à 5	1,5 à 1,5 %	0,682
		2 759	10 à 11	0,8 à 0,9 %	0,704
		1 973	12 à 13	0,3 à 0,3 %	0,702
		1 765	14 à 15	0,5 à 0,6 %	0,719
Voisinage (_SFHS6)	_SFHQ06A,	2 674	0 à 1	5,2 à 5,6 %	0,889
	_SFHQ06B,	2 918	2 à 3	2,1 à 2,5 %	0,897
	_SFHQ06C,	3 021	4 à 5	1,6 à 1,9 %	0,901
	_SFHQ06D,	2 498	10 à 11	0,9 à 1,4 %	0,910
	_SFHQ06E	1 768	12 à 13	0,3 à 0,8 %	0,906
		1 586	14 à 15	0,6 à 0,9 %	0,900

9.5.5 Échelle du soutien social

Objectifs et aperçu

On a posé ces questions à tous les PMR 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 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., soutien religieux, 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 des *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 six 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, parce qu'on s'attendait à obtenir peu de variations temporelles au chapitre du soutien social dont jouissent les personnes, et en raison d'inquiétudes quant au fardeau de réponse.

Modifications à la section relative au soutien social

Au cycle 4, les modifications suivantes ont été apportées à la section relative au 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 tirés de l'échelle secondaire de l'intégration sociale susmentionnée (items _SPHQ01H et _SPHQ01I) 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.
- 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, ont été ajoutées. Toutefois, les questions proposées rompent avec les simples réponses affirmatives et négatives qui n'indiquent pas de variabilité et sont plutôt associées aux catégories de réponse « entièrement en désaccord », « en désaccord », « d'accord » et « entièrement d'accord ».

Score	Items inclus	Taille de l'échantillon	Univers (âge en ans)	Non-réponse à l'item	Alpha de Cronbach (standardisé)
Soutien social (_SPHS01)	_SPHQ01A*	3 270	0 à 1	4,7 à 5,7 %	0,906
	_SPHQ01B,	3 362	2 à 3	2,1 à 2,2 %	0,898
	_SPHQ01C,	3 450	4 à 5	1,6 à 1,8 %	0,901
	_SPHQ01D*,	2 785	10 à 11	0,8 à 0,9 %	0,909
	_SPHQ01E*,	1 979	12 à 13	0,3 à 0,4 %	0,912
	_SPHQ01F,	1 777	14 à 15	0,5 à 0,5 %	0,915
	_SPHQ01H,				
	_SPHQ01I*.				

*Indique que les valeurs ont été renversées.

9.5.6 Échelle de comportement

Objectif

L'objectif de cette échelle est d'évaluer certains aspects du comportement des enfants de 2 à 11 ans.

Angoisse de séparation (enfants âgés de 2 et 3 ans)

Questions _BECQ6CC, _BEC6QDD1, _BECQ8LL1, _BECQ8PP1 et _BEC8TT1 de la *Child Behavior Checklist* (CBCL) d'Achenbach.

Opposition (enfants âgés de 2 et 3 ans)

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

Troubles des conduites (enfants âgés de 10 à 11 ans)

Questions _BECQ6AA, _BECQ6FF, _BECQ6JJ et _BECQ6NN tirées de l'Étude sur la santé des jeunes ontariens (ESJO).

Hyperactivité (enfants âgés de 2 à 5 et 10 à 11 ans)

Questions _BECQ6B, _BECQ6I, _BECQ6P et _BECQ6W tirées de l'ESJO, et questions _BECQ6QQ et _BECQ8HH 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 (enfants âgés de 2 à 5 et 10 à 11 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 âgés de 10 à 11 ans)

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

Agressivité physique (enfants âgés de 2 à 5 ans et de 10 à 11 ans)

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

Inattention (enfants âgés de 2 à 5 et de 10 à 11 ans)

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

Comportement prosocial (enfants âgés de 10 à 11 ans)

Questions _BECQ6A, _BECQ6H, _BECQ6M, _BECQ6GG et _BECQ6OO tirées de l'ESJO, et questions _BECQ6D, _BECQ6U, _BECQ6BB, _BECQ6SS et _BECc6UU 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 âgés de 4 à 11 ans. Dans le cycle 4, tous les enfants âgés 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 échelles sur le comportement concernant les jeunes de 2 et 3 ans

Score	Items inclus	Taille de l'échantillon	Univers (âge en ans)	Non-réponse à l'item	Alpha de Cronbach (standardisé)
Hyperactivité – inattention (_BECdS01)	_BECQ6B, _BECQ6I, _BECQ6P, _BECQ6W, _BECQ6QQ, _BECQ8HH	3 391	2 à 3 ans	0,55 % à 0,84 %	0,730
Troubles émotifs – anxiété (_BECdS03)	_BECdQ6F, _BECQ6K, _BECQ6Q, _BECQ6V, _BECQ6MM, _BECQ6RR	3 407	2 à 3 ans	0,61 % à 0,87 %	0,602
Agressivité physique – opposition (_BECS04)	_BECQ6G, _BECQ6X, _BECQ6NN, _BECQ6R1, _BECQ8E1, _BECQ8T1, _BECQ8Z1, _BECQ8J1	3 376	2 à 3 ans	0,61 % à 0,81 %	0,727
Angoisse de séparation (_BECS05)	_BECQ6CC, _BECQ6DD1, _BECQ8LL1, _BECQ8PP1, _BEC8TT1	3 421	2 à 3 ans	0,78 % à 0,89 %	0,566

Aperçu des échelles sur le comportement concernant les jeunes de 4 et 5 ans

Score	Items inclus	Taille de l'échantillon	Univers (âge en ans)	Non-réponse à l'item	Alpha de Cronbach (standardisé)
Hyperactivité – inattention (_BECDS06)	_BECQ6B, _BECQ6I, _BECQ6P, _BECQ6S, _BECQ6W, _BECQ6QQ, _BECQ8HH	3 471	4 à 5 ans	0,96 % à 1,19 %	0,787
Troubles émotifs – anxiété (_BECDS08)	_BECdQ6F, _BECQ6K, _BECQ6Q, _BECQ6V, _BECQ6MM, _BECQ6RR, _BECQ6CC	3 475	4 à 5 ans	0,88 % à 1,28 %	0,665
Agressivité physique – troubles des conduites (_BECDS09)	_BECQ6G, _BECQ6X, _BECQ6AA, _BECQ6FF, _BECQ6JJ, _BECQ6NN	3 483	4 à 5 ans	0,88 % à 0,93 %	0,764
Agressivité indirecte (_BECs10)	_BECQ6J, _BECQ6R, _BECQ6Z, _BECQ6LL, _BECQ6TT	3 409	4 à 5 ans	1,56 % à 2,04 %	0,640

Aperçu des échelles sur le comportement concernant les jeunes de 10 et 11 ans

Score	Items inclus	Taille de l'échantillon	Univers (âge en ans)	Non-réponse à l'item	Alpha de Cronbach (standardisé)
Comportement prosocial (_BECdS07)	_BECQ6A, _BECQ6D, _BECQ6H, _BECQ6M, _BECQ6U, _BECQ6BB, _BECQ6GG, _BECQ6OO, _BECQ6SS, _BECc6UU	2 647	10 à 11 ans	0,9 % à 2,9 %	0,841
L'atteinte à la propriété (_BECdS11)	_BECQ6C, _BECQ6E, _BECQ6L, _BECQ6T, _BECQ6DD, _BECQ6PP	2 784	10 à 11 ans	0,6 % à 1,1 %	0,608
Hyperactivité – inattention (_BECdS06)	_BECQ6B, _BECQ6I, _BECQ6P, _BECQ6S, _BECQ6W, _BECQ6QQ, _BECQ8HH	2 782	10 à 11 ans	0,6 % à 1,1 %	0,847
Troubles émotifs – anxiété (_BECdS08)	_BECdQ6F, _BECQ6K, _BECQ6Q, _BECQ6V, _BECQ6MM, _BECQ6RR, _BECQ6CC	2 787	10 à 11 ans	0,6 % à 1,0 %	0,776
Agressivité physique – troubles des conduites (_BECDS09)	_BECQ6G, _BECQ6X, _BECQ6AA, _BECQ6FF, _BECQ6JJ, _BECQ6NN	2 779	10 à 11 ans	0,7 % à 1,2 %	0,795
Agressivité indirecte (_BECs10)	_BECQ6J, _BECQ6R, _BECQ6Z, _BECQ6LL, _BECQ6TT	2 565	10 à 11 ans	2,0 % à 6,1 %	0,798

9.5.7 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 D^{re} Gail Poe du National Centre 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, de leur naissance à leur troisième année. Les questions varient selon l'âge de l'enfant. 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* effectuée aux États-Unis et pour la *National Child Development Survey* effectuée 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). À cause d'un problème avec l'application, la question 26 (_MSCQ26) n'a pas été posée aux enfants de 7 et 8 mois. Par conséquent, le score brut pour ces enfants est de 14. En regardant l'information des cycles antérieurs, nous avons vu qu'au moins 93 % des enfants de cet âge répondent « non » à cette question. De plus, les enfants qui auraient répondu « oui » à cette question auraient tout de même obtenu les scores les plus élevés pour ce groupe d'âge, même sans tenir compte de cette question. Par conséquent, aucun ajustement n'a été fait pour tenir compte de cette erreur.

Même s'il existait divers ensembles de questions selon l'âge de l'enfant exprimé en mois, des différences ont été observées au moment de la comparaison des notes 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 notes plus élevées que ceux de quatre mois. C'est pourquoi on a décidé de produire des notes normalisées. 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 6.

Scores normalisés basés sur les normes du cycle 6

Un score normalisé est calculé pour chaque enfant de 3 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 soit de 100 et que l'écart type soit de 15. Par conséquent, les enfants de 3 mois ont un score moyen de 100, les enfants de 4 mois ont un score moyen de 100 et ainsi de suite pour les enfants de 47 mois.

Une fois ces scores calculés, les enfants dont le score était éloigné de plus de 3 fois l'écart type par rapport à la moyenne (score inférieur à 55 ou supérieur à 145) ont été identifiés et les scores ont été recalculés en enlevant ces enfants. Ces enfants ont été considérés comme ayant des valeurs aberrantes et par conséquent, non représentatifs des enfants de leur âge. Par conséquent, la moyenne des scores normalisés, sur le fichier de données selon l'âge en mois, peut ne pas être exactement 100. L'utilisation de ce score normalisé (_MSCdS03) permet de comparer les scores des enfants du groupe d'âge de 3 à 47 mois sans devoir faire d'ajustement pour l'âge.

Ce score n'a pas été calculé pour les enfants de 0 à 2 mois car il n'y a pas suffisamment d'enfants de ces âges dans l'échantillon pour 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-haut, sauf que les données du cycle 1 ont été utilisées pour ensuite être appliquées aux données du cycle 6.

De façon générale, il n'y a pas de différences majeures entre les scores basés sur les normes établies au cycle 1 versus ceux obtenus en utilisant les normes établies au cycle 6. 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.

Les questions de l'échelle du développement moteur et social n'ont pas changé à travers les six cycles de l'ELNEJ, mais il y a eu certains changements dans le calcul des scores finaux. Pour de plus amples renseignements 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.

9.5.8 Échelles du rôle parental

Objectifs et aperçu

L'objectif de cette échelle est de mesurer certaines pratiques parentales. Plus précisément, deux échelles ont été utilisées. La première a été conçue afin de mesurer l'interaction positive, l'hostilité ou l'inefficacité et la cohérence des parents envers les enfants. La deuxième échelle vise à mesurer les pratiques parentales provoquant ou non l'aversion.

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 échelles du rôle parental afin d'en évaluer les propriétés psychométriques pour la population de l'ELNEJ. La composition de chaque facteur obtenue lors de ces analyses a été comparée à celle indiquée dans la documentation. Les résultats de ces analyses sont présentés plus loin dans cette section.

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

Résultats de l'analyse visant les enfants de 0 à 1 an concernant les échelles du rôle parental

Score	Items inclus	Taille de l'échantillon	Univers (âge en ans)	Non-réponse à l'item	Alpha de Cronbach (standardisé)
Interaction positive (_PRCS01)	_PRCQ01, _PRCQ02, _PRCQ03, _PRCQ06, _PRCQ07	3,393	0 à 1 an	2,1 à 3,4 %	0,638
Inefficacité (_PRCS02)	_PRCQ04, _PRCQ05	3 409	0 à 1 an	2,5 à 2,9 %	0,367

Résultats de l'analyse visant les enfants de 2 à 5 ans concernant les échelles du rôle parental

Score	Items inclus	Taille de l'échantillon	Univers (âge en ans)	Non-réponse à l'item	Alpha de Cronbach (standardisé)
Interaction positive (_PRCS03)	_PRCQ01, _PRCQ02, _PRCQ03, _PRCQ06, _PRCQ07	3 375	2 à 3	2,4 à 2,5 %	0,703
		3 454	4 à 5	2,1 à 2,1 %	0,703
Inefficacité (_PRCS04)	PRCQ04, _PRCQ08*, _PRCQ09, _PRCQ13, _PRCQ14, _PRCQ15, _PRCQ18	3 323	2 à 3	2,5 à 3,2 %	0,680
		3 416	4 à 5	2,2 à 2,6 %	0,678
Cohérence (_PRCS05)	_PRCQ10, _PRCQ11, _PRCQ12*, _PRCQ16*, _PRCQ17*	3 260	2 à 3	2,7 à 3,3 %	0,668
		3 374	4 à 5	2,3 à 2,6 %	0,665
Rationalité (_PRCS06)	_PRCQ21, _PRCQ22*, _PRCQ23, _PRCQ24*	3 354	2 à 3	2,8 à 3,1 %	0,545
		3 442	4 à 5	2,4 à 2,4 %	0,566

*Indique que les items ont été renversés.

Résultats de l'analyse visant les jeunes de 10 à 11 ans concernant les échelles du rôle parental

Score	Items inclus	Taille de l'échantillon	Univers (âge en ans)	Non-réponse à l'item	Alpha de Cronbach (standardisé)
Interaction positive (_PRCS03)	_PRCQ01, _PRCQ02, _PRCQ03, _PRCQ06, _PRCQ07	2 769	10 à 11	1,8 à 1,9 %	0,721
Inefficacité (_PRCS04)	_PRCQ04, _PRCQ08*, _PRCQ09, _PRCQ13, _PRCQ14, _PRCQ15, _PRCQ18	2 750	10 à 11	1,8 à 2,1 %	0,733
Cohérence (_PRCS05)	_PRCQ10, _PRCQ11, _PRCQ12, _PRCQ16, _PRCQ17	2 658	10 à 11	2,2 à 5,2 %	0,686
Rationalité (_PRCS06)	_PRCQ21, _PRCQ22*, _PRCQ23, _PRCQ24*	2 765	10 à 11	2,1 à 2,1 %	0,536
Résolution des conflits (_PRCbS09)	_PRCB30A*, _PRCB30B, _PRCB30C, _PRCB30D, _PRCB30E, _PRCB30F, _PRCB30G, _PRCB30H*	1 957	12 à 13	1,5 à 1,9 %	0,747
		1 751	14 à 15	1,6 à 1,8 %	0,730

*Indique que les items ont été renversés.

9.5.9 Étapes du développement

La composante Étapes du développement a été produite pour tous les enfants de 3 à 47 mois. Les questions comprises dans les questionnaires Étapes du développement ont été groupées 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 de plus amples renseignements sur cette mesure, veuillez consulter le chapitre 8.0 ou communiquer avec Brookes Publishing Co. And Health Professions Press.

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

9.6 Échelles relatives au 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 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'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.

Score	Items inclus	Taille de l'échantillon	Univers (âge en ans)	Non-réponse à l'item	Alpha de Cronbach (standardisé)
Dépression (_HTCbS1B)	_FBCd10A, _FBCd10B, _FBCd10C, _FBCd10D, _FBCd10E, _FBCd10F*, _FBCd10G, _FBCd10H*, _FBCd10I, _FBCd10J*, _FBCd10K, _FBCd10L	1 176	16 à 17	0,5 à 0,5 %	0,602
Dépression (_HTYfS01)	_HTYf14A, _HTYf14B, _HTYf14C, _HTYf14D, _HTYf14E, _HTYf14F*, _HTYf14G, _HTYf14H*, _HTYf14J, _HTYf14K*, _HTYf14M, _HTYf14N	1 567 1 400	18 à 19 20 à 21	0,0 à 0,3 % 0,0 à 0,3 %	0,811 0,811

*Indique que les items ont été renversés

9.6.2 Échelle de structure du quartier

Objectifs et aperçu

Cette échelle 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.

Score	Items inclus	Taille de l'échantillon	Univers (âge en ans)	Non-réponse à l'item	Alpha de Cronbach (standardisé)
Structure du quartier (_ACYDS01)	_ACYD13A, _ACYD13B, _ACYD13C, _ACYD13D, _ACYD13F, _ACYD13G	1 359	16 à 17	9,0 à 13,4 %	0,829

9.6.3 L'échelle relative à l'estime globale de soi

Le but de l'échelle de l'estime globale de soi est de mesurer l'estime de soi de l'enfant. Les questions formant l'échelle de l'estime de soi sont posées aux jeunes de 10 à 17 ans dans le questionnaire papier qu'ils doivent eux-mêmes remplir et aux jeunes de 18 à 19 ans dans le questionnaire de l'application d'IPAO. Les questions sur l'estime globale de soi sont tirées de l'échelle générale d'auto-évaluation du questionnaire d'auto-description de Marsh, mises au point par H.W. Marsh.

Score	Items inclus	Taille de l'échantillon	Univers (âge en ans)	Non-réponse à l'item	Alpha de Cronbach (standardisé)
L'estime globale de soi (_AMCS02)	_AMCQ01A, _AMCQ01B, _AMCQ01C, _AMCQ01D	2 459	10 à 11	0,2 à 0,2 %	0,792
		1 707	12 à 13	0,2 à 0,3 %	0,823
		1 507	14 à 15	0,3 à 0,3 %	0,844
		1 199	16 à 17	0,4 à 0,4 %	0,862
L'estime globale de soi (_AMYfS01)	_FAMyfQ01, _FAMyfQ02, _FAMyfQ03, _FAMyfQ04	1 568	18 à 19	0,3 à 0,3%	0,746

9.6.4 Quotient émotionnel

Mis au point par D^r Reuven BarOn et D^r James D.A. Parker, le test de quotient émotionnel version adultes (EQ-i) et le test de quotient émotionnel version jeunes (Emotional Quotient Inventory Youth Version - 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 10 à 17 ans et celles servant à mesurer l'EQ-i, aux jeunes de 20 à 21 ans. Les premières se trouvaient sur le questionnaire papier que les jeunes devaient remplir eux-mêmes, tandis que les secondes se trouvaient dans le questionnaire de l'application d'IPAO.

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, seule la valeur de la gestion du stress a été inversée. On a ensuite soustrait 1 de chacun des items afin de permettre un score de zéro. Une fois ces deux étapes accomplies, 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 cependant pas fournis dans le fichier principal. Le tableau ci-dessous présente un guide d'interprétation des scores normalisés. 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 autour des 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)

Résultats de l'analyse concernant l'EQ-i :YV

Score	Items inclus	Taille de l'échantillon	Univers (âge en ans)	Non-réponse à l'item	Alpha de Cronbach (standardisé)
Aptitudes intra personnelles (_EQYES06)	_AMCE25A, _AMCE25F, _AMCE25K	2 427	10 à 11	2,0 à 2,9 %	0,810
		1 703	12 à 13	2,2 à 2,7 %	0,825
		1 499	14 à 15	3,1 à 3,3 %	0,855
		1 197	16 à 17	4,2 à 4,5 %	0,855
Aptitudes interpersonnelles (_EQYES07)	_AMCE25B, _AMCE25G, _AMCE25L	2 415	10 à 11	2,1 à 3,0 %	0,631
		1 699	12 à 13	2,3 à 2,8 %	0,635
		1 504	14 à 15	3,0 à 3,2 %	0,666
		1 190	16 à 17	4,4 à 4,7 %	0,618
Aptitudes gestion du stress (_EQYES08)	_AMCE25C*, _AMCE25H*, _AMCE25M*	2 403	10 à 11	2,9 à 3,3 %	0,631
		1 686	12 à 13	2,8 à 2,9 %	0,649
		1 494	14 à 15	3,2 à 3,5 %	0,658
		1 198	16 à 17	4,2 à 4,4 %	0,678
Aptitudes d'adaptabilité (_EQYES09)	_AMCE25D, _AMCE25I, _AMCE25N	2 439	10 à 11	2,3 à 2,6 %	0,679
		1 710	12 à 13	2,2 à 2,4 %	0,734
		1 503	14 à 15	3,1 à 3,2 %	0,778
		1 197	16 à 17	4,2 à 4,4 %	0,774
Humeur générale (_EQYES10)	_AMCE25E, _AMCE25J, _AMCE25O	2 437	10 à 11	2,2 à 2,9 %	0,689
		1 705	12 à 13	2,3 à 2,6 %	0,704
		1 501	14 à 15	3,0 à 3,5 %	0,676
		1 197	16 à 17	4,3 à 4,4 %	0,673
Quotient émotionnel (EQ4) (_EQYES04)	_AMCE25A, _AMCE25F, _AMCE25K, _AMCE25B, _AMCE25G, _AMCE25L _AMCE25C*, _AMCE25H*, _AMCE25M* _AMCE25D, _AMCE25I, _AMCE25N	2 297	10 à 11	2,0 à 3,3 %	0,563
		1 655	12 à 13	2,2 à 2,9 %	0,526
		1 477	14 à 15	3,1 à 3,5 %	0,481
		1 182	16 à 17	4,2 à 4,7 %	0,42
Quotient émotionnel (EQ5) (_EQYES05)	_AMCE25A, _AMCE25F, _AMCE25K, _AMCE25B, _AMCE25G, _AMCE25L _AMCE25C*, _AMCE25H*, _AMCE25M* _AMCE25D, _AMCE25I, _AMCE25N _AMCE25E, _AMCE25J, _AMCE25O	2 270	10 à 11	2,0 à 3,3 %	0,691
		1 645	12 à 13	2,2 à 2,9 %	0,667
		1 470	14 à 15	3,0 à 3,5 %	0,636
		1 181	16 à 17	4,2 à 4,7 %	0,585

Résultats de l'analyse concernant l'EQ-i

Score	Items inclus	Taille de l'échantillon	Univers (âge en ans)	Non-réponse à l'item	Alpha de Cronbach (standardisé)
Aptitudes intra personnelles (_EQYFS11)	_EQYfQ02*, _EQYfQ07*, _EQYfQ12*, _EQYfQ17*	1 395	20 à 21	0,0 à 0,1 %	0,627
Aptitudes interpersonnelles (_EQYFS12)	_EQYfQ01, _EQYfQ06, _EQYfQ11, _EQYfQ16	1 407	20 à 21	0,1 à 0,9 %	0,648
Aptitudes gestion du stress (_EQYFS13)	_EQYfQ03*, _EQYfQ08*, _EQYfQ13*, _EQYfQ18*	1 403	20 à 21	0,0 à 0,4 %	0,642
Aptitudes d'adaptabilité (_EQYFS14)	_EQYfQ04, _EQYfQ09, _EQYfQ14, _EQYfQ19	1 403	20 à 21	0,1 à 0,2 %	0,678
Humeur générale (_EQYFS15)	_EQYfQ05, _EQYfQ10, _EQYfQ15, _EQYfQ20	1 404	20 à 21	0,0 à 0,2 %	0,689
Quotient émotionnel (EQ4) (_EQYFS16)	_EQYfQ02*, _EQYfQ07*, _EQYfQ12*, _EQYfQ17*, _EQYfQ01, _EQYfQ06, _EQYfQ11, _EQYfQ16, _EQYfQ03*, _EQYfQ08*, _EQYfQ13*, _EQYfQ18*, _EQYfQ04, _EQYfQ09, _EQYfQ14, _EQYfQ19	1 388	20 à 21	0,0 à 0,9 %	0,598
Quotient émotionnel (EQ5) (_EQYFS17)	_EQYfQ02*, _EQYfQ07*, _EQYfQ12*, _EQYfQ17*, _EQYfQ01, _EQYfQ06, _EQYfQ11, _EQYfQ16, _EQYfQ03*, _EQYfQ08*, _EQYfQ13*, _EQYfQ18*, _EQYfQ04, _EQYfQ09, _EQYfQ14, _EQYfQ19, _EQYfQ05, _EQYfQ10, _EQYfQ15, _EQYfQ20	1 386	20 à 21	0,0 à 0,9 %	0,709

Voici de brèves définitions de ce qui est mesuré par les 5 échelles composites et les 15 sous-échelles [le EQ-i :YV (test de QE – version jeunes) 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) :**

Ces aptitudes comprennent les composantes suivantes qui 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.
- Émotions.
- 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) :**

Ces aptitudes comprennent les composantes suivantes qui 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 gestion du stress (gestion et contrôle des émotions) :**

Ces aptitudes comprennent les composantes suivantes qui déterminent la capacité de gérer ses émotions pour qu'elles travaillent pour soi et non contre soi :

- 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) :**

Ces aptitudes comprennent les composantes suivantes qui 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 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 facilite un comportement intelligent sur le plan émotif et social et comprend les composantes suivantes qui déterminent 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 plus de renseignements, voir les suivants :

- 1) BarOn, R., 2004, The BarOn Emotional Quotient Inventory (EQ-i): Rationale, description, and summary of psychometric properties, dans Glenn Geher (Éd.), *Measurement of emotional intelligence: Common ground and controversy*, Hauppauge (New-York), Nova Science Publishers, p. 111 à 142.

- 2) BarOn, R., et Parker, J.D.A., 2000. *Emotional Quotient Inventory: Youth Version (EQ-I : YV): Technical manual*. Toronto, Canada, Multi-Health Systems.

9.6.5 Échelle du soutien social

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 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., soutien religieux, services communautaires) ont été ajoutées comme le suggérait D^r Tom Hay. Des questions s'apparentant à celles proposées par D^r 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 comporte six questions et couvre non seulement la portée du soutien social mais également la qualité de ce soutien. La mesure globale du soutien social comporte 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.

Score	Items inclus	Taille de l'échantillon	Univers (âge en ans)	Non-réponse à l'item	Alpha de Cronbach (standardisé)
Soutien social (_SPYES01)	_SPYEQ1A*, _SPYEQ1B, _SPYEQ1C, _SPYEQ1D*, _SPYEQ1E*, _SPYEQ1F, _SPYEQ1G, _SPYEQ1H*	1 562	18 à 19	0,5 à 1,6 %	0,847

*Indique que les items ont été renversés.

9.6.6 Échelle des amis

Cette section de l'enquête sert à recueillir des renseignements sur le nombre d'amis intimes, le temps passé avec les amis, la présence d'une personne à qui l'enfant peut se confier et la qualité des relations avec d'autres personnes, comme les parents, les pairs et les enseignants. Ces renseignements sont importants 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.

Score	Items inclus	Taille de l'échantillon	Univers (âge en ans)	Non-réponse à l'item	Alpha de Cronbach (standardisé)
Amis (_FFCS01)	_FFCQ01, _FFCQ02, _FFCQ03, _FFCQ04	2 461	10 à 11	1,5 à 2,1 %	0,770
		1 702	12 à 13	1,9 à 2,8 %	0,809
		1 473	14 à 15	4,3 à 5,1 %	0,835
		1 191	16 à 17	4,3 à 5,0 %	0,874

9.6.7 L'échelle mes parents et moi

L'échelle mes parents et moi vise à compléter la section sur le rôle parental du questionnaire de l'enfant remplie par la PMR, grâce à la collecte de renseignements directement auprès de l'enfant sur la perception que celui-ci a de 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.

L'échelle employée a également été utilisée dans le cadre de la *Child Health Survey* de l'Australie-Occidentale. Elle a été élaborée par Lempers et al. (1989), qui se sont fondés sur les travaux de Schaefer (1965) et de Roberts et al. (1984). Cette échelle mesure les soins prodigués par les parents, le rejet et la surveillance. Elle permet de compléter les concepts mesurés par le questionnaire de l'enfant rempli par le parent (interaction positive entre parents et enfants, interaction inefficace, interaction cohérente et pratiques parentales provoquant ou non de l'aversion).

Score	Items inclus	Taille de l'échantillon	Univers (âge en ans)	Non-réponse à l'item	Alpha de Cronbach (standardisé)
Soins prodigués par les parents (_PMCCS1)	_PMCCQ1A, _PMCCQ1D,	2 269	10 à 11	3,2 à 5,9 %	0,860
	_PMCCQ1K, _PMCCQ1H,	1 581	12 à 13	3,1 à 6,6 %	0,904
	_PMCCQ1L, _PMCCQ1M, _PMCCQ1Q	1 466	14 à 15	4,2 à 4,6 %	0,920
Rejet par les parents (_PMCBS2B)	_PMCCQ1C, _PMCCQ1G,	2 197	10 à 11	3,9 à 8,0 %	0,664
	_PMCCQ1J, _PMCCQ1L,	1 583	12 à 13	3,6 à 5,8 %	0,739
	_PMCCQ1O, _PMCCQ1P, _PMCCQ1R	1 456	14 à 15	4,3 à 5,1 %	0,771
Surveillance par les parents (_PMCCS3)	_PMCCQ1B, _PMCCQ1F,	2 321	10 à 11	3,4 à 6,0 %	0,482
	_PMCCQ1N, _PMCCQ1E,	1 648	12 à 13	3,6 à 4,5 %	0,505
	_PMCCQ1T	1 469	14 à 15	4,2 à 4,9 %	0,481

9.6.8 Échelles de résolution des conflits

Deux scores ont été créés pour évaluer la résolution des conflits chez les jeunes de 16 et 17 ans en fonction des questions posées dans le questionnaire auto administré. 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 jeune et parent.

Score	Items inclus	Taille de l'échantillon	Univers (âge en ans)	Non-réponse à l'item	Alpha de Cronbach (standardisé)
Résolution des conflits - Mère (_PMCDS4)	_PMCdQ6C, _PMCdQ6D, _PMCdQ6E, _PMCdQ6F, _PMCdQ6G, _PMCdQ6H, _PMCdQ6I, _PMCdQ6J, _PMCdQ6K, _PMCdQ6L	1 162	16 à 17	4,0 à 4,4 %	0,617
Résolution des conflits - Père (_PMCDS5)	_PMCdQ9C, _PMCdQ9D, _PMCdQ9E, _PMCdQ9F, _PMCdQ9G, _PMCdQ9H, _PMCdQ9I, _PMCdQ9J, _PMCdQ9K, _PMCdQ9L	1 110	16 à 17	4,9 à 5,3 %	0,669

9.6.9 Échelle de comportement

Cette section reprend la liste de comportements utilisée dans le questionnaire de l'enfant rempli par la PMR. Tous les enfants âgés de 10 à 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 des conduites, hyperactivité, inattention, agressivité physique, agressivité indirecte, troubles émotifs, anxiété et comportement prosocial.

Score	Items inclus	Taille de l'échantillon	Univers (âge en ans)	Non-réponse à l'item	Alpha de Cronbach (standardisé)
Agressivité indirecte (_FBCS01)	_FBCQ01J, FBCQ01R, _FBCQ01Z, _FBCQ1LL, _FBCQ1TT	2 346	10 à 11	3,7 à 4,3 %	0,730
		1 677	12 à 13	2,4 à 3,1 %	0,747
		1 495	14 à 15	3,3 à 3,7 %	0,714
Troubles émotifs – anxiété (_FBCdS02)	_FBCQ01F, _FBCQ01K, _FBCQ01Q, _FBCQ01V, _FBCQ1CC, _FBCQ1MM, _FBCQ1RR	2 295	10 à 11	3,9 à 4,8 %	0,730
		1 665	12 à 13	2,4 à 3,3 %	0,775
		1 485	14 à 15	3,3 à 3,9 %	0,784
Agressivité physique - trouble des conduites (_FBCS03)	_FBCQ01G, _FBCQ01X, _FBCQ1AA, _FBCQ1FF, FBCQ1JJ, _FBCQ1NN	2 334	10 à 11	3,8 à 4,3 %	0,760
		1 681	12 à 13	2,4 à 3,0 %	0,803
		1 489	14 à 15	3,3 à 3,7 %	0,801
Hyperactivité - inattention (_FBCdS04)	_FBCQ01B, _FBCQ01I, _FBCQ01P, _FBCQ01S, _FBCQ01W, _FBCQ1HH, _FBCQ1QQ	2 294	10 à 11	3,8 à 5,1 %	0,755
		1 664	12 à 13	2,5 à 4,1 %	0,769
		1 474	14 à 15	3,3 à 4,0 %	0,796
Comportement prosocial (_FBCS05)	_FBCQ01A, _FBCQ01D, _FBCQ01H, _FBCQ01M, _FBCQ01U, _FBCQ1BB, _FBCQ1GG, _FBCQ1OO, _FBCQ1SS, _FBCQ1UU	2 296	10 à 11	3,5 à 5,0 %	0,817
		1 658	12 à 13	2,4 à 3,3 %	0,857
		1 490	14 à 15	3,1 à 3,7 %	0,874
L'atteinte à la propriété (_FBCS07)	_FBCQ01C, _FBCQ01E, _FBCQ01L, _FBCQ01T, _FBCQ1DD, _FBCQ1PP	2 314	10 à 11	3,6 à 5,7 %	0,633
		1 680	12 à 13	2,4 à 3,3 %	0,704
		1 483	14 à 15	3,1 à 4,0 %	0,712

9.7 Résumé – Échelles du cycle 6

9.7.1 Échelles relatives au questionnaire du parent

Variable de score	Nom de l'échelle	Univers
_DPPS01	Échelle de dépression	PMR des enfants de 0 à 15 ans
_FNHS01	Échelle de fonctionnement de la famille	PMR ou conjoint(e) de la PMR des enfants de 0 à 15 ans
_SFHS5	Échelle de sécurité du quartier	PMR ou conjoint(e) de la PMR des enfants de 0 à 15 ans
_SFHS6	Échelle des voisins	PMR ou conjoint(e) de la PMR des enfants de 0 à 15 ans
_SPHS01	Échelle du soutien social	PMR ou conjoint(e) de la PMR des enfants de 0 à 15 ans
_ACCS6	Échelle des responsabilités familiales	PMR des enfants de 10 à 13 ans
_BECdS01	Hyperactivité - inattention	PMR des enfants de 2 à 3 ans
_BECdS03	Troubles émotifs - anxiété	PMR des enfants de 2 à 3 ans
_BECS04	Agressivité physique - opposition	PMR des enfants de 2 à 3 ans
_BECS05	Angoisse de séparation	PMR des enfants de 2 à 3 ans
_BECdS06	Hyperactivité - inattention	PMR des enfants de 4 à 11 ans
_BECdS07	Comportement prosocial	PMR des enfants de 8 à 11 ans
_BECdS08	Troubles émotifs - anxiété	PMR des enfants de 4 à 11 ans
_BECdS09	Agressivité physique - troubles des conduites	PMR des enfants de 4 à 11 ans
_BECS10	Agressivité indirecte	PMR des enfants de 4 à 11 ans
_BECdS11	L'atteinte à la propriété	PMR des enfants de 8 à 11 ans
_MSCS01	Score brut - DMS	PMR des enfants de 0 à 47 mois
_MSCS02	Score normalisé DMS – basé sur les données du cycle 1	PMR des enfants de 0 à 47 mois
_MSCdS03	Score normalisé DMS – basé sur les données du cycle 5	PMR des enfants de 3 à 47 mois
_PRCS01	Interaction positive	PMR des enfants de 0 à 23 mois
_PRCS02	Inefficacité	PMR des enfants de 0 à 23 mois
_PRCS03	Interaction positive	PMR des enfants de 2 à 11 ans
_PRCS04	Inefficacité	PMR des enfants de 2 à 11 ans
_PRCS05	Cohérence	PMR des enfants de 2 à 11 ans
_PRCS06	Rationalité	PMR des enfants de 2 à 11 ans
_PRCbS09	Échelle de résolution des conflits	PMR des enfants de 12 à 15 ans

9.7.2 Échelles relatives au questionnaire des enfants/jeunes (auto administré)

Variable de score	Nom de l'échelle	Univers
_FFCS01	Échelle des amis	Enfants/Jeunes de 10 à 17 ans
_AMCS02	L'estime globale de soi	Enfants/Jeunes de 10 à 17 ans
_FBCS01	Agressivité indirecte	Enfants/Jeunes de 10 à 15 ans
_FBCdS02	Troubles émotifs - anxiété	Enfants/Jeunes de 10 à 15 ans
_FBCS03	Agressivité physique - troubles des conduites	Enfants/Jeunes de 10 à 15 ans
_FBCdS04	Hyperactivité - inattention	Enfants/Jeunes de 10 à 15 ans
_FBCS05	Comportement prosocial	Enfants/Jeunes de 10 à 15 ans
_FBCS07	L'atteinte à la propriété	Enfants/Jeunes de 10 à 15 ans
_PMcC_S1	Soins prodigués par les parents	Enfants/Jeunes de 10 à 15 ans
_PMcC_S2b	Rejet par les parents	Enfants/Jeunes de 10 à 15 ans
_PMcC_S3	Surveillance par les parents	Enfants/Jeunes de 10 à 15 ans
_EQYES04	Quotient émotionnel (4 facteurs)	Enfants/Jeunes de 10 à 17 ans
_EQYES05	Quotient émotionnel (5 facteurs)	Enfants/Jeunes de 10 à 17 ans
_EQYES06	Aptitudes intrapersonnelles	Enfants/Jeunes de 10 à 17 ans
_EQYES07	Aptitudes interpersonnelles	Enfants/Jeunes de 10 à 17 ans
_EQYES08	Aptitudes gestion du stress	Enfants/Jeunes de 10 à 17 ans
_EQYES09	Aptitudes d'adaptabilité	Enfants/Jeunes de 10 à 17 ans
_EQYES10	Humeur générale	Enfants/Jeunes de 10 à 17 ans
_PMcC_S4	Résolution des conflits - Mère	Jeunes de 16 à 17 ans
_PMcC_S5	Résolution des conflits - Père	Jeunes de 16 à 17 ans
_HTCbS1B	Échelle de dépression	Jeunes de 16 à 17 ans

9.7.3 Échelles relatives au questionnaire des jeunes (auto administré)

Variable de score	Nom de l'échelle	Univers
_ACYdS01	Échelle de structure du quartier	Jeunes de 16 à 17 ans
_HTYfS01	Échelle de dépression	Jeunes de 18 ans et plus
_SPYES01	Échelle de soutien social	Jeunes de 18 à 19 ans
_AMYfS01	L'estime globale de soi	Jeunes de 18 à 19 ans

10.0 Imputation

La définition d'un répondant est 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 chez les 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'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 pour 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 sixième caractère du nom de la variable. Ainsi, FINHel02 représente l'indicateur d'imputation pour le revenu du ménage (FINHeQ03).

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 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 autres sources. La somme de ces dix 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, 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 net 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 ou 17 ans vivant de façon autonome. Au moins une variable du revenu a été imputée pour 3 160 des 16 491 ménages admissibles. Ces 16 491 ménages admissibles correspondent en fait à 18 684 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 10 et 21 ans au cycle 6, et ii) le fichier du développement de la petite enfance (DPE) pour les enfants ou les jeunes faisant partie des cohortes de naissance sélectionnées aux cycles 4, 5 et 6, âgés de 0 à 5 ans au cycle 6. Le fichier longitudinal contient 10 514 enfants ou jeunes, pour 2 291 desquels des valeurs ont été imputées, tandis que le fichier du DPE contient 8 170 enfants, pour 1 159 desquels des valeurs ont été imputées. Dans l'ensemble, des valeurs ont été imputées pour 3 450 des 18 684 enfants ou jeunes.

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 6 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 indiquent un revenu estimatif, b) les ménages qui ont répondu au cycle 5, c) les ménages ayant déclaré un revenu à l'aide d'intervalles de revenu et d) les autres ménages.

a) Imputation pour les ménages qui indiquent un revenu estimatif

Pour ces ménages, nous tenons le revenu estimatif pour 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 dix sources de revenu sont imputées 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 dix sources de revenu peut être légèrement différent du revenu estimatif fourni.

b) Imputation pour les ménages qui ont répondu au cycle 5

Afin de conserver la cohérence longitudinale au fil du temps, l'imputation du revenu total des ménages qui ont répondu au cycle 5 a été réalisée au moyen de la méthode de la tendance du plus proche voisin, sauf en ce qui concerne les ménages donneurs affichant des tendances extrêmes sur le plan du revenu du cycle 5 au cycle 6. Lorsqu'un receveur déclare son revenu total à l'aide des intervalles de revenu, nous nous assurons que le revenu total imputé respecte l'intervalle précisé. Tout comme dans la situation a), la répartition des sources de revenu du ménage donneur est utilisée pour imputer les dix sources de revenu du ménage receveur (imputation par le ratio du donneur).

c) Imputation pour les ménages qui déclarent 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 imputées en utilisant le même ménage donneur et en s'assurant que le revenu total du ménage respecte l'intervalle retenu.

d) Imputation pour les autres ménages

Nous ne disposons 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 et iii) imputation des sources de revenu du ménage.

Le tableau ci-dessous montre la ventilation des ménages imputés selon l'étape du procédé d'imputation.

Étape du procédé d'imputation	Nombre de ménages	Pourcentage de ménages imputés
a) Ménages indiquant un revenu estimatif	698	22,1 %
b) Ménages ayant répondu au cycle 5	1 314	41,6 %
c) Ménages déclarant un revenu à l'aide d'intervalles de revenu	388	12,3 %
d) Autres ménages	760	24,0 %
Total	3 160	100 %

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 (FINHel02) :

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 (FINPel03 et FINSel03) :

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 (FINP11AA, FINP11AB, FINP11AC, FINS11AA, FINS11AB, FINS11AC, FINHI1AD, FINHI1AF, FINHI1AE et FINHI1AG) :

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 donner
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 se trouvent dans le tableau suivant.

Variable	Indicateur d'imputation	Taux d'imputation global	Taux d'imputation pour le fichier longitudinal (âges 10 à 21)	Taux d'imputation pour le fichier du DPE (0 à 5 ans)
FINPc1AA Revenu de la PMR (salaires et traitements)	FINPI1AA	11,1 %	9,6 %	12,2 %
FINPc1AB Revenu de la PMR (emploi autonome)	FINPI1AB	9,2 %	7,8 %	10,2 %
FINPc1AC Revenu de la PMR (prestations d'assurance-emploi)	FINPI1AC	9,8 %	7,6 %	11,5 %
FINPeD04 Revenu personnel total de la PMR	FINPeI03	14,0 %	11,6 %	15,9 %
FINS1AA Revenu du conjoint (salaires et traitements)	FINSI1AA	14,3 % *	12,0 % *	16,0 % *
FINS1AB Revenu du conjoint (emploi autonome)	FINSI1AB	11,9 % *	10,0 % *	13,3 % *
FINS1AC Revenu du conjoint (prestations d'assurance-emploi)	FINSI1AC	10,5 % *	8,4 % *	11,9 % *
FINSeD04 Revenu personnel total du conjoint	FINSeI03	16,4 % *	13,2 % *	18,8 % *
FINHe1AD Revenu du ménage (prestation fiscale pour enfants)	FINHI1AD	14,5 %	12,4 %	16,1 %
FINHe1AE Revenu du ménage (assistance sociale)	FINHI1AE	8,8 %	7,1 %	10,1 %
FINHe1AF Revenu du ménage (allocation de soutien des enfants ou au conjoint)	FINHI1AF	8,9 %	7,5 %	10,0 %
FINHe1AG Revenu du ménage (autres sources)	FINHI1AG	9,4 %	7,8 %	10,6 %
FINHeQ03 Revenu total du ménage	FINHeI02	18,5 %	14,2 %	21,8 %

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

10.2 Imputation du revenu des jeunes – 16 ou 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 ou 17 ans au cycle 6. 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 d'autres sources. La somme de ces cinq sources de revenu représente le revenu total des jeunes.

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 de leur revenu au moyen des intervalles de revenu. Enfin, parmi ceux qui répondent aux questions, il arrive que les montants indiqués à la section touchant 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 divers contre rémunération, mais qui ne déclare pas de revenu à ce chapitre dans la section sur le revenu). On procède à une imputation du revenu pour combler les valeurs manquantes attribuables à la non-réponse partielle ou pour corriger les données incohérentes lorsque cela est possible.

L'imputation du revenu a été envisagée pour tous les répondants de 16 ou 17 ans qui ont fourni des renseignements dans la composante sur les jeunes. Sur les 1 585 jeunes répondants de 16 ou 17 ans au cycle 6, 1 436 ont rempli la composante sur les jeunes. Les 149 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, 358 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 quatre sources de revenu doit être imputée et que l'intervalle de revenu total était indiqué, une valeur plausible choisie au hasard à même une répartition uniforme de valeurs possibles a été imputée.

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

Le tableau suivant indique le nombre et le pourcentage de jeunes répondants pour lesquels au moins une des sources de revenu a été imputée, par méthode d'imputation.

Méthode d'imputation (indicateur d'imputation FINYe102)	Nombre de jeunes	Pourcentage de jeunes ayant une valeur imputée
Imputation d'une valeur plausible	61	17,0 %
Imputation par donneur selon l'intervalle de revenu	79	22,1 %
Imputation par donneur	218	60,9 %
Total	358	100 %

Les taux d'imputation pour les variables du revenu se trouvent dans le tableau suivant.

Variable	Indicateur d'imputation	Taux d'imputation (%)
FINYeQ1A Revenu provenant de travaux divers	FINYeI1A	12,4 %
FINYdQ1B Revenu provenant d'employeurs	FINYeI1B	12,3 %
FINYeQ1E Revenu provenant des bourses d'études et de subventions	FINYeI1E	2,9 %
FINYeQ1C Revenu provenant des parents	FINYeI1C	8,6 %
FINYdQ1D Revenu provenant d'autres sources	FINYeI1D	5,4 %
FINYeD01 Revenu total du jeune	FINYeI02	24,9 %

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 (sans montants détaillés). 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
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. Certains répondants refusent donc de donner des réponses aux questions sur le revenu. Parmi ceux-ci, certains fournissent une estimation de leur revenu au moyen des intervalles de revenu. Enfin, parmi ceux qui répondent aux questions, 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.

L'imputation du revenu a été envisagée pour tous les répondants de 18 ans et plus. Le revenu total de 300 jeunes répondants sur les 2 982 qui avaient 18 ans et plus au cycle 6 a été imputé, tandis que la liste des sources de revenu a été imputée pour 11 de ces 300 répondants.

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 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 par donneur

Le tableau suivant indique le nombre et le pourcentage de jeunes répondants pour lesquels le revenu total a été imputé, par méthode d'imputation.

Méthode d'imputation (indicateur d'imputation FIYYfI2A)	Nombre de jeunes	Pourcentage de jeunes ayant une valeur imputée
Imputation par donneur selon l'intervalle de revenu	236	78,7 %
Imputation par donneur	64	21,3 %
Total	300	100 %

Les taux d'imputation pour les variables du revenu se trouvent dans le tableau suivant.

Variable	Indicateur d'imputation	Taux d'imputation
FIYYfQ1A à FIYYfQ1O Sources de revenu	FIYYfI1	0,4 %
FIYYfD01 Revenu total des jeunes	FIYYfI2A	10,1 %

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

La méthodologie retenue au cycle 6 pour l'imputation de l'échelle de Développement moteur et social (DMS) est la même que celle retenue au cycle 5. Antérieurement au cycle 5, pour qu'un enfant obtienne un score brut pour l'échelle de DMS (variable **FMSCS01**), les 15 questions que comporte l'échelle de DMS devaient être répondues par « Oui » ou « Non ». Cependant, pour la vaste majorité des cas où on ne pouvait pas calculer un score brut, il a été noté que seulement une ou deux des 15 questions de l'échelle avaient une réponse manquante. Nous avons donc décidé, depuis le cycle 5, que nous pourrions obtenir un score très valable en faisant usage des 13 ou 14 questions répondues et en imputant les données manquantes.

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. Ainsi, la valeur manquante initiale est remplacée par un « Oui » ou un « Non » provenant de l'enregistrement donneur sélectionné.

Lorsque les réponses à deux questions sont imputées, l'imputation est faite de façon indépendante. Ainsi, deux donneurs différents peuvent être utilisés 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 âgé de 8 mois et ayant une valeur manquante pour la variable FMSCQ21 a comme donneurs potentiels les enfants âgés de 7 à 9 mois à qui les mêmes 15 questions (FMSCQ12 à FMSCQ26) ont été posées et qui présentent le même schéma de réponse « Oui » et « Non » pour FMSCQ12 à FMSCQ20 et FMSCQ22 à FMSCQ26.

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, 525 scores additionnels du DMS ont été obtenus en imputant au moins une réponse. En fait, une réponse a été imputée dans 472 cas, et deux réponses ont été imputées dans 53 cas. Cela représente 7,5 % de tous les enfants admissibles.

Les indicateurs d'imputation FMSCIS1A et FMSCIS1B 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'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 dans l'échantillon représente un certain nombre d'unités dans la population. Dans l'ELNEJ, plusieurs populations sont représentées. L'échantillon complet du cycle 6 est une combinaison d'échantillons sélectionnés aux cycles 1, 4 et 5 de l'enquête (1994, 2000 et 2002) et d'un nouvel échantillon sélectionné au cycle 6 en 2004. Pour des renseignements détaillés 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.

Nous avons vu au chapitre 5.0 que, dans une enquête longitudinale comme l'ELNEJ, deux types de populations sont possibles, longitudinale et transversale. La première est formée par la population initiale au moment où l'échantillon a été prélevé la première fois, tandis que la seconde renvoie à des périodes ultérieures. Les différences entre la population longitudinale et les populations transversales ultérieures 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 et un ensemble de poids transversaux. La différence entre les deux ensembles de poids longitudinaux est la suivante : les poids « entonnoir » ne concernent que les membres de la cohorte initiale qui ont répondu à tous les cycles, tandis que les autres poids longitudinaux concernent les répondants longitudinaux qui n'ont peut-être pas répondu à tous les cycles (cohorte initiale ou cohortes du Développement de la petite enfance (DPE)).

Les poids longitudinaux ne sont attribués qu'aux personnes de retour qui avaient de 0 à 1 an quand elles ont commencé à prendre part à l'ELNEJ. Le participant qui en est à son premier cycle ne reçoit qu'un poids transversal.

11.1 Poids disponibles au cycle 6

Les paragraphes qui suivent portent sur les populations de référence pour les différents poids produits au cycle 6.

11.1.1 Poids entonnoir (variable FWTCdW1L)

Les poids entonnoir ne sont attribués qu'aux membres longitudinaux de la cohorte initiale qui ont répondu à chaque cycle.

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

Un échantillon longitudinal d'enfants de 0 à 11 ans a été prélevé au cycle 1 à partir de l'EPA. Ces enfants avaient entre 10 et 21 ans au cycle 6 (au 31 décembre 2004). L'échantillon a été réduit au cycle 2. Les enfants retirés entre les cycles 1 et 2 peuvent être considérés comme les enfants transversaux du cycle 1.

Depuis le cycle 4, des poids entonnoir sont produits à tous les cycles pour cette population longitudinale.

11.1.2 Poids longitudinaux (variable FWTCW01L)

Des poids longitudinaux (et non des poids entonnoir) sont attribués aux membres longitudinaux de retour de la cohorte initiale et des cohortes du DPE qui sont des répondants au cycle 6 mais qui n'ont peut-être pas répondu à tous les cycles précédents.

La cohorte initiale

Population longitudinale (même définition que pour les poids entonnoir) : enfants âgés de 0 à 11 ans au 31 décembre 1994 qui vivaient dans l'une des dix provinces lors de la collecte des données du cycle 1 (1994-1995).

Un échantillon longitudinal d'enfants de 0 à 11 ans a été prélevé au cycle 1 à partir de l'EPA. Ces enfants avaient entre 10 et 21 ans au cycle 6 (au 31 décembre 2004). L'échantillon a été réduit au cycle 2. Les enfants retirés entre les cycles 1 et 2 peuvent être considérés comme les enfants transversaux du cycle 1.

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

Dans les fichiers de données, il est possible d'identifier ces personnes grâce à la mention « MEMCYCLE = 01 ». La variable MEMCYCLE indique le cycle au cours duquel l'enfant est entré dans l'enquête.

Cohorte du DPE sélectionnée au cycle 4

Population longitudinale de la cohorte du DPE sélectionnée au cycle 4 : enfants âgés de 0 à 1 ans au 31 décembre 2000 qui vivaient dans l'une des dix provinces lors de la collecte des données du cycle 4 (2000-2001).

Un échantillon longitudinal d'enfants de 0 à 1 an a été prélevé au cycle 4 à partir de l'EPA. Ces enfants avaient de 4 à 5 ans au cycle 6 (en décembre 2004).

Des poids longitudinaux ont été produits pour cette population aux cycles 5 et 6.

Dans les fichiers de données, il est possible d'identifier ces personnes grâce à la mention « MEMCYCLE = 04 ». La variable MEMCYCLE indique le cycle au cours duquel l'enfant est entré dans l'enquête.

Cohorte du DPE sélectionnée au cycle 5

Population longitudinale de la cohorte du DPE sélectionnée au cycle 5 : enfants âgés de 0 à 1 ans au 31 décembre 2002 qui vivaient dans l'une des dix provinces lors de la collecte des données du cycle 5 (2002-2003).

Un échantillon longitudinal d'enfants de 0 à 1 an a été prélevé au cycle 5 à partir de l'EPA. Ces enfants avaient de 2 à 3 ans au cycle 6 (en décembre 2004).

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

Dans les fichiers de données, il est possible d'identifier ces personnes grâce à la mention « MEMCYCLE = 05 ». La variable MEMCYCLE indique le cycle au cours duquel l'enfant est entré dans l'enquête.

11.1.3 Poids transversaux (variable FWTCW01C)

Poids transversaux (variable FWTCW01C) : uniquement pour les membres de l'échantillon de l'ELNEJ qui, au cycle 6, peuvent représenter la population transversale suivante.

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

L'échantillon transversal comprend :

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

11.2 Méthode de pondération

La méthode de pondération est exposée en détail ci-dessous. Les utilisateurs qui veulent en savoir davantage peuvent se mettre en rapport avec Statistique Canada.

11.2.1 Pondération de l'Enquête longitudinale nationale sur les enfants et des jeunes

La stratégie de pondération de l'ELNEJ repose sur une série d'ajustements appliqués au poids d'échantillonnage de l'ELNEJ. Le poids d'échantillonnage de chaque enfant est égal à l'inverse de la probabilité de sélection de l'enfant. Tous les enfants qui participent encore à l'ELNEJ au cycle 6 ont été sélectionnés, à un moment ou à un autre, à partir de l'EPA. Le sous-poids de cette dernière enquête tient compte de toute l'information du plan de sondage pour l'échantillon de l'EPA. Dans le cas de l'ELNEJ, nous sélectionnons des ménages à même l'EPA, puis nous sélectionnons des enfants dans ces ménages. Pour tenir compte de ces étapes supplémentaires, le sous-poids de l'EPA est multiplié par plusieurs autres facteurs pour enfin arriver au poids du plan de sondage de l'ELNEJ. Le poids final est obtenu en appliquant des ajustements pour la non-réponse et la stratification a posteriori au poids du plan de l'ELNEJ. Dans le présent chapitre, nous renvoyons au poids du plan de l'ELNEJ ainsi : W poids initial de l'ELNEJ*

11.2.2 Premier ajustement : ajustement pour la non-réponse

En vérité, dans la plupart des enquêtes, les personnes échantillonnées ne répondent pas toutes, et l'ELNEJ ne fait pas exception. En raison de cette non-réponse, nous devons ajuster les poids de façon que les répondants représentent les non-répondants. Sinon, nous sous-estimerions les totaux, entre autres choses.

L'objectif de l'ajustement pour la non-réponse est plus précisément de gonfler les poids des répondants à l'ELNEJ de façon que les poids ajustés de la non-réponse donnent la somme des poids de tous les membres de l'échantillon initial de l'ELNEJ. En d'autres mots, l'ajustement pour la non-réponse vise à donner les poids d'échantillonnage de l'ELNEJ des non-répondants aux répondants d'une façon intelligente. Les enfants de l'échantillon du cycle 1 ont connu six cycles de non-réponse (cycles 1 à 6 inclusivement). Ceux de l'échantillon du cycle 4 ont perdu des répondants sur trois cycles (cycles 4, 5 et 6), tandis que les enfants de l'échantillon du cycle 5 ont été affectés par deux cycles de non-réponse (cycles 5 et 6). L'échantillon du cycle 6 n'a connu qu'un seul cycle de non-réponse (cycle 6). L'ajustement pour la non-réponse vise à corriger tous ces épisodes de non-réponse en une seule étape.

Afin de décider de quelle façon le poids des non-répondants est déplacé vers les répondants, nous avons recours à la méthode des groupes de réponse 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 :

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

Au cycle 6, la propension à répondre est déterminée par des modèles de régression logistique qui permettent de prédire des probabilités de réponse (dont la valeur se situe entre 0 et 1) pour tous. Ces probabilités sont ensuite triées et divisées en Y GRH, où Y est établi de façon qu'un taux de réponse monotonique soit atteint dans les groupes et de façon que certaines contraintes (taille du facteur d'ajustement et taille minimale de chaque GRH) soient respectées afin que des facteurs d'ajustement raisonnables et fiables soient établis.

Trois modèles de régression logistique sont construits pour la pondération au cycle 6 : un pour la réponse longitudinale de l'échantillon du cycle 1, un autre pour la réponse en entonnoir de l'échantillon du cycle 1 et un autre pour la réponse des échantillons des cycles 4, 5 et 6. Des modèles distincts sont construits parce que l'échantillon du cycle 1 a connu beaucoup plus de cycles de non-réponse que les cohortes du DPE, et tout porte à croire que le comportement au regard de la non-réponse varie en fonction de l'augmentation du nombre de fois où un individu est interviewé. À partir de ces modèles, les GRH sont construits séparément pour les poids entonnoir, longitudinaux et transversaux.

11.2.3 Second ajustement : stratification a posteriori

Le second facteur d'ajustement a pour objet d'assurer la cohérence entre les estimations produites à partir de l'enquête et les estimations démographiques produites par

Statistique Canada selon l'âge, le sexe et la province. Cette méthode porte le nom de stratification a posteriori. Les totaux de la stratification a posteriori dépendent de la population de référence.

Tant les poids transversaux que les poids longitudinaux (y compris les poids entonnoir) sont stratifiés a posteriori. En ce qui concerne la pondération transversale au cycle 6, l'année de référence pour calculer l'âge effectif de l'enfant est 2004. Les chiffres utilisés dans la stratification a posteriori renvoient à janvier 2005, de façon que nous puissions disposer d'un dénombrement fiable des enfants d'un âge donné au 31 décembre de l'année précédente. Dans le même ordre d'idées, la pondération longitudinale utilise les chiffres de janvier de l'année suivant l'année de référence pour la population d'intérêt.

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 du cycle 6 et celle des cycles précédents

La méthode de pondération exposée ci-dessus pour le cycle 6 est différente de la méthode utilisée dans les cycles précédents. Nous avons remanié le modèle de non-réponse pour le cycle 6, de sorte que le modèle de non-réponse longitudinal est maintenant véritablement longitudinal. La méthode de modélisation de la non-réponse a aussi été modifiée : elle est passé d'un modèle de segmentation utilisé dans les cycles précédents à un modèle de régression logistique au cycle 6. Nous obtenons ainsi des GRH moins nombreux, plus robustes et davantage distincts. L'utilisation d'une « variable de collaboration » dans le modèle de non-réponse est une nouveauté : un score est attribué aux individus au moyen de cette variable en fonction de leur participation à l'enquête dans le passé, et cette variable s'est révélée un prédicteur de la non-réponse.

Par exemple, le score attribué à la collaboration à l'échantillon du cycle 1 est obtenu à l'aide d'information provenant du cycle 1. Nous établissons essentiellement la mesure dans laquelle le répondant a répondu aux questions au cycle 1. S'il a répondu à toutes les questions, le score de collaboration est élevé. En revanche, s'il n'a pas répondu à un grand nombre de questions, le score de collaboration est faible. Le score lui-même est une variable nominale. Dans le cas des non-répondants au cycle 1, le score de collaboration a été imputé.

11.3 Application de la méthode de pondération

11.3.1 Pondération longitudinale

Le poids longitudinal, FWTCW01L, s'applique aux trois échantillons distincts d'enfants de l'ELNEJ :

- la cohorte initiale 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 4;
- la cohorte du DPE formée d'enfants de 0 à 1 an sélectionnés cycle 5.

À noter que le poids longitudinal renvoie à une population de référence distincte pour chaque échantillon.

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 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, etc.) et un score supplémentaire de collaboration était attribué à la cohorte initiale (variable dépendante). L'ajustement du poids pour la non-réponse est calculé par GRH à l'aide de la formule présentée plus haut.

En ce qui concerne la cohorte initiale et l'échantillon du cycle 4, 9 GRH et 5 GRH ont été utilisés respectivement. Cinq GRH ont aussi été utilisés pour l'échantillon du cycle 5. Même si les cycles 4 et 5 ont été modélisés ensemble, l'ajustement pour la non-réponse a été calculé séparément pour la simple et bonne raison que chaque cycle représente ses propres populations de référence.

Second ajustement : ajustement pour la stratification a posteriori

En ce qui concerne l'échantillon d'enfants sélectionnés au cycle 1, 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 de l'échantillon d'enfants sélectionnés au cycle 4, la population de référence est celle de tous les enfants âgés de 0 à 1 an au 31 décembre 2000. Enfin, dans le cas de l'échantillon d'enfants sélectionnés au cycle 5, la population de référence est celle de tous les enfants de 0 à 1 an au 31 décembre 2002.

11.3.2 Pondération entonnoir

Le poids entonnoir, FWTCdW1L, n'est produit que pour :

- la cohorte initiale des enfants de 0 à 11 ans sélectionnés au cycle 1

Définition d'un répondant entonnoir

Un répondant entonnoir est un répondant longitudinal au cycle 6 (voir la définition ci-dessus) qui a répondu à tous les cycles précédents, en l'occurrence les cycles 1 à 6. À noter que les personnes qui ont répondu à tous les cycles constituent un sous-ensemble des répondants au cycle 6.

Premier ajustement : ajustement pour la non-réponse

L'ajustement pour la non-réponse repose encore une fois sur la création de GRH. Le modèle de non-réponse est établi à l'aide de variables de l'EPA et d'un score de collaboration. Les GRH sont produits avec les extraits du modèle. À noter que ces GRH sont différents des GRH établis pour la pondération longitudinale puisque les répondants au cycle 6 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. Onze GRH ont été utilisés en tout.

Second ajustement : ajustement pour la stratification a posteriori

Ici encore, la population de référence est constituée des enfants âgés 0 à 11 ans au 31 décembre 1994. L'ajustement est calculé pour chaque combinaison d'âge, de sexe et de province.

11.3.3 Pondération transversale

Le poids transversal, FWTCW01C, s'applique aux quatre échantillons distincts d'enfants de l'ELNEJ :

- la cohorte du DPE formée d'enfants de 0 à 1 an sélectionnés au cycle 6;
- les enfants de retour de 2 à 3 ans provenant de la cohorte du DPE d'enfants de 0 à 1 an sélectionnés au cycle 5;
- les enfants de retour de 4 à 5 ans provenant de la cohorte du DPE d'enfants de 0 à 1 an sélectionnés au cycle 4;
- un nouvel échantillon complémentaire d'enfants de 2 à 5 ans.

La pondération transversale des enfants de 0 à 5 ans consiste à représenter la population au moment de la collecte du cycle 6.

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 non admissibles. 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 5 ans en 2004. Le modèle de non-réponse utilise des variables de l'EPA. Dix GRH ont été utilisés en tout pour exécuter l'ajustement pour la non-réponse.

Second ajustement : stratification a posteriori

Les enfants de l'ELNEJ qui avaient entre 0 et 5 ans au 31 décembre 2004 ont été stratifiés a posteriori selon les chiffres de population par âge, sexe et province.

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. Nous accorderons ici une attention particulière aux erreurs non dues à l'échantillonnage.

De plus amples renseignements d'ordre général sur la qualité des données de l'enquête et l'assurance de la qualité sont fournis à l'adresse www.statcan.ca.

12.1 Erreur d'échantillonnage

Les estimations établies à partir des données de l'enquête sont fondées sur un échantillon d'enfants. Nous aurions pu obtenir des valeurs légèrement différentes si nous avions procédé à un recensement complet en faisant appel aux mêmes questionnaires et aux mêmes intervieweurs, superviseurs, méthodes de traitement, etc. 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. Les chapitres 13.0 et 17.0 fournissent 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 répondent aux questions de façon erronée, 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 visée, 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, et la non-réponse partielle où l'unité sélectionnée répond à suffisamment de questions pour être considérée répondant, 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. Si nous sommes dans l'impossibilité de mesurer exactement ce que les non-répondants « auraient déclaré », nous pouvons au moins 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 au chapitre 11.0. Le chapitre 13.0 présente pour sa part un examen général 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 de non-réponse.

Dans le cas de l'ELNEJ, des Groupes de réponses homogènes (GRH) ont été créés de façon que les poids des répondants compensent pour les non-répondants ayant une propension prédite à répondre semblable, là où cette propension prédite à répondre repose sur des caractéristiques recueillies précédemment (niveau de scolarité, type de logement, etc.). Néanmoins, quel que soit le GRH, les non-répondants peuvent différer des répondants de diverses façons importantes non observées ou inconnues.

La non-réponse s'amplifie au fil du temps. Comme nous avons 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 par 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 6. 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 types distincts de taux de réponse : le **taux de réponse à la phase de la collecte** mesure l'efficacité du processus de collecte des données et est établi en fonction des unités auprès desquelles des données sont recueillies.

Le **taux de réponse à la phase de l'estimation** 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 du cycle 6 constituent la principale différence entre les deux types de taux de réponse. Les refus catégoriques ou chroniques dans les cycles précédents pour la cohorte initiale ainsi que n'importe quel type de non-réponse dans les cycles précédents pour la cohorte du Développement de la petite enfance (DPE) en sont des exemples. Ces enfants sont retirés de l'échantillon du cycle 6, et leur absence influe sur la qualité des données. Comme ils ne sont pas envoyés sur le terrain au cycle 6, ces enfants ne sont pas pris en considération dans l'établissement du 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 au chapitre 6.0, et des taux au niveau du ménage sont donnés au chapitre 5.0. 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 plutôt 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 pour tenir compte du fait que la non-réponse de certaines unités (c'est-à-dire les unités ayant d'importants poids d'échantillonnage dans les enquêtes sociales et 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 de 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, nous observons très peu de différence dans les provinces entre les taux pondérés et les taux non pondérés pour l'ELNEJ.

Les taux de réponse présentés 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. Par ailleurs, les critères à respecter pour être considéré 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 PMR. La composante de la personne la mieux renseignée (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 du 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.

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 conjoint) est remplie. Une composante pour laquelle on a répondu à un ensemble de questions clés est considérée comme **remplie**. À noter qu'une composante considérée comme remplie peut présenter une importante non-réponse partielle (voir la section 11.4).

Un jeune dont l'âge effectif se situe entre 16 et 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 conjoint) est remplie.

Le jeune dont l'âge effectif est de 18 ans ou 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 ou 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 faisant partie de la population cible. La somme du nombre d'enfants admissibles et inadmissibles égale la taille de l'échantillon. À 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** correspond au nombre d'enfants ou de jeunes répondants par rapport au nombre d'enfants ou de jeunes admissibles.

12.3.2 Taux de réponse transversaux

L'échantillon transversal du cycle 6 est formé de l'agrégation des enfants sélectionnés pour la première fois dans trois cohortes différentes, celles des cycles 4, 5 et 6. Seuls les enregistrements des personnes qui ont répondu au cycle précédent sont envoyés sur le terrain au cycle 6.

Taux de réponse transversaux non pondérés selon la cohorte

Âge effectif	Échantillon initial	Inadmissibles	Admissibles	Répondants	NR C4	NR C5	NR C6	Taux de réponse à l'estimation (%)
0 et 1 an	4 356	13	4 343	3 521	s/o	s/o	822	81,1
2 et 3 ans de retour	4 492	117	4 375	2 866	s/o	1 142	367	65,5
2 et 3 ans – Échant. comp.	734	12	722	595	s/o	s/o	127	82,4
4 et 5 ans de retour	5 087	144	4 943	2 964	1 125	505	349	60,0
4 et 5 ans – Échant. comp.	703	5	698	568	s/o	s/o	130	81,4
Total	15 372	291	15 081	10 514	1 125	1 647	1 795	69,7

Définitions du titre des colonnes du tableau ci-dessus

Âge effectif – Établi d'après l'année de naissance. Ceux dont l'âge effectif est 0 année sont nés en 2004; ceux dont l'âge effectif est de 1 an sont nés en 2003; etc.

Échantillon initial – 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.

Inadmissibles – Nombre d'enfants inadmissibles sur le plan transversal au cycle 6. Il est possible que les enfants aient été déclarés inadmissibles lors de la collecte des données du cycle 6, du cycle 5 ou du cycle 4. Comme prévu, la proportion d'enfants inadmissibles est petite en ce qui concerne l'échantillon prélevé au cycle 6, et elle est de plus en plus importante au fur et à mesure que l'on s'éloigne de l'échantillon initial. Plus le temps passe, plus la possibilité d'émigrer est élevée.

Admissibles – Ce nombre représente le dénominateur du taux de réponse. Voir la définition complète à la section précédente.

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

NR C4 – Nombre de non-répondants à la collecte de données du cycle 4 parmi les enfants admissibles au cycle 6. Ces enfants ont été retirés des échantillons des cycles 5 et 6.

NR C5 – Nombre de non-répondants à la collecte de données du cycle 5 parmi les enfants admissibles au cycle 6. Ces enfants ont été retirés de l'échantillon du cycles 6.

NR C6 – Nombre de non-répondants à la collecte de données du cycle 6.

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. En ce qui concerne les enfants de retour dont l'âge effectif est de 4 ou 5 ans, le taux de réponse chez les enfants de l'échantillon du cycle 6 se situait à près de 90 %. Cependant, si on prend en compte le fait que plusieurs centaines d'enfants admissibles ont été retirés de l'échantillon du cycle 6 en raison de la non-réponse à un cycle précédent, le taux de réponse à la phase de l'estimation est considérablement plus faible (60 %). Ces faibles taux de réponse nous préoccupent davantage relativement à l'éventualité du biais de non-réponse examiné plus haut dans le présent chapitre.

Dans ce contexte, nous avons modifié notre stratégie de collecte pour le cycle 7. En effet, une nouvelle tentative sera faite au cycle 7 pour obtenir des données concernant les enfants non répondants aux cycles précédents. Cela devrait accroître les taux de réponse transversaux à la phase de l'estimation au cycle 7.

Le tableau qui suit résume le tableau précédent *Taux de réponse transversaux non pondérés selon la cohorte* en regroupant l'échantillon des enfants de retour et l'échantillon complémentaire.

Taux de réponse transversaux non pondérés selon l'âge effectif

Âge effectif	Admissibles	Répondants	Taux de réponse à l'estimation (%)
0 et 1 an	4 343	3 521	81,1
2 et 3 ans	5 097	3 461	67,9
4 et 5 ans	5 641	3 532	62,6
Total	15 081	10 514	69,7

Le prochain tableau montre les taux par province. L'Ontario affiche le taux de réponse le plus faible, comme dans les cycles précédents.

Taux de réponse transversaux non pondérés selon la province

Province	Admissibles	Répondants	Taux de réponse à l'estimation (%)
Terre-Neuve-et-Labrador	781	618	79,1
Île-du-Prince-Édouard	645	494	76,6
Nouvelle-Écosse	931	701	75,3
Nouveau-Brunswick	995	702	70,6
Québec	2 288	1 589	69,4
Ontario	4 230	2 746	64,9
Manitoba	1 178	847	71,9
Saskatchewan	1 063	819	77,0
Alberta	1 439	1 038	72,1
Colombie-Britannique	1 431	960	67,1
Canada	15 081	10 514	69,7

Les chiffres de la colonne *Admissibles* 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 6 (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 courant sur le nombre d'enfants *sélectionnés* au premier cycle. Cependant, comme la méthode de sélection utilisée lors 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 tienne compte de la non-réponse au cycle 1 pour les enfants sélectionnés à ce même cycle. Plus précisément, le nombre 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.

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

Voici quelques éléments à prendre en considération concernant le tableau ci-dessous.

- Le dénominateur pour tous les pourcentages présentés dans ces tableaux 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.
- L'enfant à qui un poids longitudinal a été attribué au cycle 6 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 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 6* et constituent un sous-ensemble de ceux auxquels un poids longitudinal a été attribué.
- 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.
- La colonne *Poids entonnoir au cycle 6* 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 des dix provinces, conservent toujours un poids même si aucune donnée n'a été recueillie les concernant. Voir le chapitre sur la pondération pour plus de précisions sur les poids entonnoir (variable FWTCWd1L).
- À la première rangée, le répondant principal est la PMR. Les jeunes de 18 ans et plus répondent pour eux-mêmes.

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

Âge effectif au cycle 6	Nombre de répondants au cycle 1	Données recueillies pour le cycle 6		Poids longitudinal au cycle 6		Poids entonnoir au cycle 6	
		Nombre	% du cycle 1	Nombre	% du cycle 1	Nombre	% du cycle 1
10 à 17 ans	12 027	8 196	68,1	8 399	69,8	7 364	61,2
18 à 21 ans	4 876	2 982	61,2	3 084	63,2	2 490	51,1
Total	16 903	11 178	66,1	11 483	67,9	9 854	58,3

Le tableau suivant montre l'attrition jusqu'au cycle 5. 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. Au cycle 5, 74,1 % des répondants du cycle 1 avaient un poids longitudinal. Ce pourcentage a chuté à 67,9 % au cycle 6, et il est encore plus faible chez les 18 ans et plus.

Taux d'attrition longitudinaux non pondérés (cycles 2 à 5) pour les enfants du cycle 1, selon la province

Province au cycle 1	Nombre de répondants au cycle 1	Répondants au cycle 2		Répondants au cycle 3		Répondants au cycle 4		Répondants au cycle 5		Répondants à tous les cycles	
		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
Île-du-Prince-Édouard	467	443	94,9	434	92,9	392	83,9	364	77,9	330	70,7
Nouvelle-Écosse	1 191	1 068	89,7	1 085	91,1	988	83,0	903	75,8	811	68,1
Nouveau-Brunswick	1 070	958	89,5	958	89,5	836	78,1	792	74,0	691	64,6
Québec	3 182	2 944	92,5	2 844	89,4	2 522	79,3	2 361	74,2	2 108	66,2
Ontario	4 342	3 899	89,8	3 760	86,6	3 318	76,4	3 104	71,5	2 714	62,5
Manitoba	1 232	1 161	94,2	1 112	90,3	1 019	82,7	1 004	81,5	891	72,3
Saskatchewan	1 413	1 305	92,4	1 257	89,0	1 073	75,9	1 002	70,9	893	63,2
Alberta	1 599	1 465	91,6	1 420	88,8	1 242	77,7	1 162	72,7	1 031	64,5
Colombie-Britannique	1 457	1 333	91,5	1 282	88,0	1 143	78,4	1 076	73,9	978	67,1
Canada	16 903	15 468	91,5	14 997	88,7	13 310	78,7	12 523	74,1	11 136	65,9

12.3.4 Taux de réponse longitudinal pour les enfants sélectionnés au cycle 4

Le taux de réponse indiqué à la rangée 4 et 5 ans de retour du tableau *Taux de réponse transversaux*, à la section 12.3.2, donne une bonne idée du taux de réponse longitudinal.

12.3.5 Taux de réponse longitudinal pour les enfants sélectionnés au cycle 5

Le taux de réponse indiqué à la rangée 2 et 3 ans de retour du tableau *Taux de réponse transversaux*, à la section 12.3.2, donne une bonne idée du taux de réponse longitudinal.

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, fatigue, omission accidentelle de certaines parties du questionnaire ou difficultés opérationnelles.

La non-réponse de 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 nous sommes 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 cherchée 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

l'information n'a pas été demandée au répondant. Cela est vrai pour la saisie des réponses assistée par ordinateur. Cependant, les questionnaires papier auto administrés font exception. Pour ces questionnaires, les réponses laissées en blanc sont classées dans la catégorie « Non déclaré », quoique le répondant ait pu avoir vu la question.

À 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 auto administré, 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. Nous encourageons 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 sur l'analyse.

Les sections qui suivent portent sur la non-réponse aux composantes pour l'ELNEJ. Cette analyse 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 indiqués 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.

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 dans le cadre de 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.

À noter que le chiffre dans la colonne *Admissibles* est fondé sur le nombre d'enfants pour lesquels des données ont été recueillies. Les enfants longitudinaux admissibles pour lesquels aucune donnée n'a été recueillie au cycle 6, mais qui figurent néanmoins dans le fichier principal avec un poids longitudinal, ne sont pas compris.

Dans la colonne *Nombre ayant répondu*, le chiffre englobe les personnes ayant répondu à toutes les questions ou non.

Admissibles	Nombre ayant répondu	Taux de réponse à la composante
18 684	18 561	99,3 %

Il est extrêmement rare que des participants ne répondent à aucune question de la composante de l'enfant. Si 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. Nous souhaitons recueillir des renseignements auprès de 16 491 PMR. Pour ce qui est des enfants, nous souhaitons en avoir 18 684 pour obtenir des renseignements sur les PMR.

Admissibles	Nombre ayant répondu	Taux de réponse à la composante
18 684	18 349	98,2 %

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. Nous souhaitons recueillir des renseignements auprès de 14 017 personnes. Pour ce qui est des enfants, nous souhaitons en avoir 15 867 pour obtenir des renseignements sur les conjoints. Les enfants vivant avec un parent seul ne sont pas admissibles.

Admissibles	Nombre ayant répondu	Taux de réponse à la composante
15 867	15 411	97,1 %

Il est rare que des participants ne répondent à aucune question de la composante du conjoint. À noter que les réponses par personne interposée étaient permises pour cette composante lorsque la PMR pouvait répondre au nom du conjoint. Voir la section 12.10 pour 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.

Admissibles	Nombre ayant répondu	Taux de réponse à la composante
1 585	1 457	91,9 %

À noter que, sur les 1 457 personnes qui ont au moins répondu en partie à la composante, 1 436 ont donné suffisamment de réponses pour que le revenu manquant soit imputé.

Il est relativement courant qu'un adulte participe alors que le jeune de 16 ou 17 ans ne participe pas.

12.4.5 EVIP

Les enfants de 4 et 5 ans ont subi le test de l'Échelle de vocabulaire en images de Peabody (EVIP).

Admissibles	Nombre ayant répondu (suffisamment pour obtenir un score)	Taux de réponse à la composante
3 532	3 184	90,1 %

À la section sur la non-réponse totale, nous avons vu qu'il y avait 5 641 enfants admissibles dans ce groupe d'âge. En tout, 3 184 enfants ont subi le test de l'EVIP. L'ensemble de la non-réponse totale et partielle signifie que 56,4 % des enfants pour lesquels un test de l'EVIP était souhaitable ont effectivement subi le test. La non-réponse totale est compensée par les poids de l'enquête, mais la non-réponse à la composante ne l'est pas.

12.4.6 Connaissance des nombres

La connaissance des nombres a été évaluée chez les enfants de 4 et 5 ans.

Admissibles	Nombre ayant répondu (suffisamment pour obtenir un score)	Taux de réponse à la composante
3 532	3 167	89,7

12.4.7 Qui suis-je?

Les enfants de 4 et 5 ans ont passé le test « Qui suis-je? »

Admissibles	Nombre ayant répondu (suffisamment pour obtenir un score)	Taux de réponse à la composante
3 532	2 798	79,2

On peut s'interroger sur la raison pour laquelle le taux de réponse à cette composante est moins élevé qu'à la composante de l'EVIP ou celle de la connaissance des nombres. Le test était peut-être trop difficile pour certains enfants admissibles. Ce taux est légèrement inférieur à celui observé au cycle 5.

12.4.8 Tests de mathématiques

Les tests de mathématiques de l'ELNEJ sont composés de 20 questions de calcul auxquelles les répondants de 10 à 15 ans doivent répondre à la maison. Le niveau du test (de 4 à 10) est déterminé par le niveau scolaire de l'enfant. Si on ne connaît pas ce niveau scolaire, l'âge effectif de l'enfant détermine le niveau du test.

Admissibles	Nombre ayant répondu (suffisamment pour obtenir un score)	Taux de réponse à la composante
6 611	5 825	88,1

12.4.9 Résolution de problèmes

Le test de résolution de problèmes s'adresse aux jeunes de 16 et 17 ans et comprend 20 questions qui visent à évaluer les points forts dans les domaines de la compréhension de l'écrit, de la résolution de problèmes et de la prise de décisions. Le test porte également sur certaines habiletés mathématiques.

Admissibles	Nombre ayant répondu (suffisamment pour obtenir un score)	Taux de réponse de la composante
1 585	1 290	81,4

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.

Admissibles	Nombre ayant répondu (suffisamment pour obtenir un score)	Taux de réponse à la composante
1 573	1 329	84,5

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.

Admissibles	Nombre ayant répondu (suffisamment pour obtenir un score)	Taux de réponse à la composante
1 409	1 145	81,3

12.4.12 Questionnaires auto administrés

La composante des questionnaires auto administrés 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.

Âge effectif	Admissibles	Nombre ayant répondu	Taux de réponse à la composante
10 à 11 ans	2 825	2 544	90,1
12 à 13 ans	1 995	1 782	89,3
14 à 15 ans	1 791	1 562	87,2
16 à 17 ans	1 585	1 281	80,8
Total	8 196	7 169	87,5

12.5 Non-réponse de cycle

Certains répondants longitudinaux ne participent pas à chaque cycle. C'est ce que l'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, 3, 4 et 6 mais non aux cycles 2 et 5.

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 longitudinaux correspondant à ce groupe (variable FWTCWd1L). Ces poids sont disponibles pour les enfants introduits au cycle 1.

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, nous prévoyons 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 un enquête générale dans le cadre de laquelle on demande à 1 000 adultes (995 non-titulaires de doctorat et 5 titulaires de 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 doctorats serait de 7/1 000 plutôt que de 5/1 000. L'écart n'est pas prononcé; cependant, 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, etc.) 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 manifester 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 d'eux-mêmes et de leurs enfants une image positive mais non réaliste. 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* leur en faire plus.

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 des données doit être conscient du fait que les variables mesurant des concepts (le revenu, la taille, etc.) et pouvant être considérées à juste titre comme continues dans la population ne conservent pas nécessairement ces traits caractéristiques dans le fichier d'enquête. Ainsi, nous constatons que le revenu déclaré correspond exactement dans de nombreux cas à un multiple de 10 000 \$ et que la taille est souvent arrondie au pied près. (Voir le graphique à la section 12.10.4.) 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, nous demandons 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 associée 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 douze 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 à la collecte par 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 aux questions de façon appropriée. Il est toutefois possible que la personne ciblée aurait 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.

À noter que le tableau ci-dessous est établi en fonction des enfants et non des adultes.

Composante	Enfants admissibles	Enfants avec des réponses par personne interposée	Taux de réponse par personne interposée
PMR	18 684	309	1,7 %
Conjoint	15 867	9 644	60,8 %

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 : catégories de réponse 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 « fortement 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 régions ou 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. Cette situation introduit 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. Ainsi, si l'année d'immigration indiquée par le répondant est antérieure à l'année de sa naissance, cette dernière est réputée être l'année d'immigration.

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 que nous ne pouvons pas en toute confiance corriger le tir, il subsiste des incohérences 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, nous acceptons 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 :

$$\text{IMC} = \frac{\text{poids en kilogrammes}}{(\text{taille en mètres}) \times (\text{taille en mètres})}$$

Les variables de taille et de poids utilisées pour calculer l'IMC à partir des données de l'ELNEJ sont les suivantes :

- FHLCQ03B (taille) et FHLCQ04A (poids) (données déclarées par la PMR pour les enfants de 2 à 11 ans) : score de l'IMC du répondant (FHLCE01);
- FHTCbQ01 (taille) et FHTCbQ02 (poids) (données déclarées par les répondants de 12 à 17 ans dans les questionnaires qu'ils remplissent eux-mêmes) : score de l'IMC du répondant (FHLCE01);
- FHTYeD01 (taille) et FHTYeD02 (poids) (données déclarées par les répondants de 18 à 21 ans dans les questionnaires qu'ils remplissent eux-mêmes) : score de l'IMC du répondant (FHTYe03).

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 (FMMCdQ1B) pour calculer les seuils. Les intervalles percentiles proposés par les CDC pourraient être utilisés pour les personnes de 0 à 20 ans. Le nom de diffusion de cette variable est FHLCeD03 pour les jeunes de 17 ans ou moins et FHTYeD05 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/nccdphp/dnpa/bmi/bmi-for-age.htm>.

12.14.3 Indice de masse corporelle – Tim Cole et coll.

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;320:1-6). 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, ces 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 FHLCEd02 pour les jeunes de 17 ans ou moins et FHTYeD04 pour les jeunes de 18 ans ou plus.

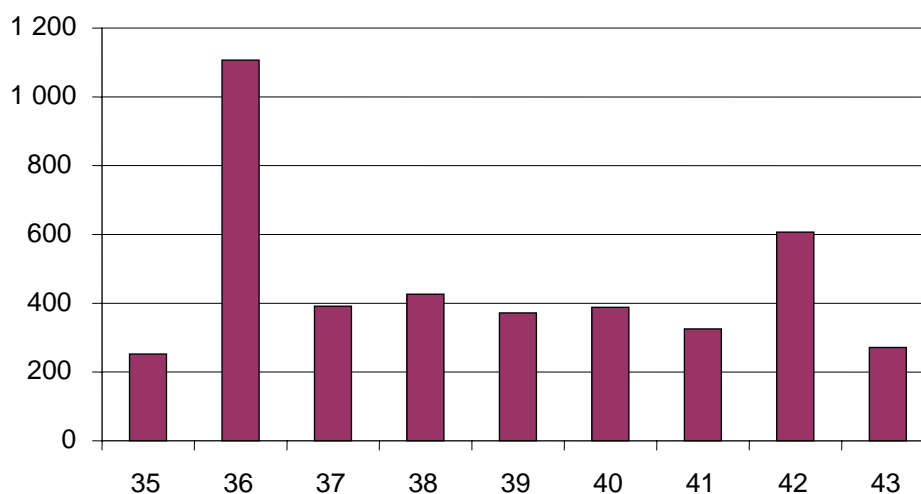
On peut obtenir plus de renseignements sur les seuils de l'IMC proposés par Tim Cole et ses collaborateurs à l'adresse suivante : <http://www.bmj.com>.

12.14.4 Indice de masse corporelle – Qualité des données

Il y a plusieurs facteurs qui influent sur la qualité des scores de l'IMC. D'abord, il y a un taux plus élevé de non-réponse 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. 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. 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 trois pieds que près de trois pieds.

Cycle 6 – Nombre non pondéré de certaines valeurs de taille déclarées en pouces



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 initiative. 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 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 éducation, de son comportement, etc. 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 et 6. 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 FDMCD06 donne le lien de la PMR à l'enfant.

12.17 Couverture

Il est essentiel de définir ici certaines notions.

La **population cible** est la population pour laquelle des renseignements sont désirés.

La **population observée** est la population réellement couverte par l'enquête. Dans le contexte de l'ELNEJ, diverses populations observées sont décrites au chapitre sur l'échantillonnage (chapitre 5.0).

La **base de sondage** contient les renseignements requis pour identifier et contacter les unités de la population observée.

Les différences entre la population observée et la population cible sont souvent attribuables aux limites de la base de sondage. Elles peuvent aussi être dues aux difficultés opérationnelles de la collecte des données. Une enquête éclair, par exemple, ne peut pas être réalisée instantanément. Le temps écoulé entre la sélection de l'échantillon et la collecte des données ainsi que le temps écoulé pendant la collecte des données peuvent influencer sur la population observée. La mesure dans laquelle la population observée correspond à la population cible équivaut à la **couverture** de l'enquête.

Imaginons par exemple que la population cible d'un sondage d'opinion est constituée de tous les électeurs admissibles du Canada. Un système d'appel aléatoire forme la base de sondage. La population observée devient alors le groupe de tous les électeurs admissibles du Canada qui vivent dans un ménage ayant le téléphone. La base de sondage impose certaines restrictions à la population cible initiale.

Les **erreurs de base de sondage** (ou erreurs de couverture) comprennent les omissions, les inclusions erronées, les doubles comptes et les erreurs de classification des unités dans la base de sondage. Après avoir défini la couverture, la base de sondage et la population observée, il est possible que la base de sondage présente certaines failles. Par exemple, la base de sondage constituée de la liste des naissances enregistrées au Canada regroupe en théorie toutes les naissances mais, dans la pratique, la liste n'est peut-être pas à jour. Les naissances manquantes constituent une erreur de base de sondage.

12.17.1 Couverture de l'Enquête longitudinale nationale sur les enfants et les jeunes

Pour la cohorte originale de l'ELNEJ, la population cible peut être précisément définie comme étant les enfants nés de 1983 à 1994 et vivant dans les dix provinces du Canada au 31 décembre 1994.

Dans le cas de la cohorte du DPE, la population cible transversale peut être définie comme étant les enfants nés entre 1999 et 2004 et vivant dans l'une des dix provinces du pays au 31 décembre 2004.

La base de sondage de l'ELNEJ est formée des ménages qui ont participé à l'Enquête sur la population active (EPA). Cette base de sondage impose certaines restrictions à notre population observée. Plus précisément, comme nous l'avons vu au chapitre 5.0, l'EPA exclut les personnes vivant dans des réserves indiennes, les membres à temps plein des Forces armées canadiennes et les pensionnaires d'établissement. Une des différences manifestes entre la population cible et la population observée pour l'ELNEJ est le fait que les enfants vivant dans des réserves ne fassent pas partie de la population observée. Cette différence ainsi que d'autres questions liées aux erreurs de base de sondage et de couverture sont examinées à la prochaine section.

12.17.2 Questions liées à la couverture et à la base de sondage

1) Exclusions de l'EPA

À noter que les enfants vivant dans les territoires sont exclus de notre population cible et de notre population observée. En outre, l'EPA ne tient pas compte des personnes qui vivent dans des réserves indiennes, des membres à temps plein des Forces armées canadiennes et des pensionnaires d'établissement. L'exclusion des personnes vivant dans des réserves est celle qui touche le plus l'ELNEJ. Même si le nombre total d'enfants vivant dans des réserves ne représente pas un important pourcentage de la population (autour de 2 %), l'incidence de cette exclusion varie considérablement selon la province. Selon les données du Recensement de 2001, près de 10 % des enfants de 0 à 5 ans en Saskatchewan et au Manitoba vivent dans des réserves.

2) L'admissibilité à l'ELNEJ dépend des valeurs des variables de l'EPA

Il y a bel et bien des erreurs de base de sondage dans l'ELNEJ. Il est possible par exemple que, dans le fichier de l'EPA, les enfants fassent l'objet d'une classification erronée selon l'année de naissance. À chaque cycle en effet, un petit nombre d'enfants qui ne sont pas véritablement admissibles à l'ELNEJ semblent l'être si on se fie à l'information de l'EPA. De même, nous supposons qu'il y a un petit nombre d'enfants qui sont admissibles à l'ELNEJ mais qui semblent ne pas l'être selon les renseignements de l'EPA.

Pour être admissible à l'ELNEJ, l'enfant doit vivre dans un ménage qui répond aux critères de l'EPA. Les enfants admissibles à l'ELNEJ vivant dans des ménages non répondants sélectionnés par l'EPA ne sont pas retenus pour l'ELNEJ, en grande partie parce que nous ne pouvons pas identifier les enfants sauf si le ménage répond à l'EPA. Les poids de l'enquête compensent pour cette non-réponse, mais un biais est toujours possible. Comme il faut s'y attendre, la non-réponse à l'EPA peut influencer sur la qualité de l'ELNEJ.

Le taux de réponse à l'EPA est toujours élevé (autour de 90 % à 93 %), et il varie légèrement d'un mois à l'autre, ce qui réduit la probabilité qu'un important biais soit associé à la non-réponse à l'EPA.

3) Immigration

En ce qui concerne une partie du fichier du DPE, entre le moment où l'échantillon a été sélectionné et celui où la collecte des données du cycle 6 a commencé, deux à quatre années se sont écoulées. Les enfants immigrants arrivés au Canada au cours de cette période ne pouvaient pas être sélectionnés.

Par exemple, au cycle 6, l'échantillon complet d'enfants de 4 et 5 ans en Ontario a été sélectionné en 2000 au cycle 4, alors que les enfants avaient 0 et 1 an. L'enfant né en 1999 ou 2000, qui s'est établi en Ontario en 2001 ou plus tard n'avait aucune chance d'être sélectionné pour l'ELNEJ.

Le tableau ci-dessous vise à montrer les différences entre les populations cibles transversales et longitudinales. La colonne *Nombre total d'enfants* indique le nombre total d'enfants âgés de 2 à 5 ans en janvier 2005. Autrement dit, il s'agit de la taille de la population cible transversale pour ce groupe d'âge. La colonne *Taille de la population cible longitudinale* indique la somme du nombre d'enfants de 0 et 1 an présents en janvier 2001 et du nombre d'enfants de 0 et 1 an présents en janvier 2003. Il s'agit essentiellement de la taille de notre population cible longitudinale.

Des projections de la population admissible à l'EPA établies d'après les données du recensement constituent la source des deux colonnes de totaux de contrôle. Des totaux provenant de cette même source, mais à un niveau différent de détail, sont utilisés pour stratifier a posteriori l'EPA.

Écart entre les estimations de la population totale pour le groupe des 2 à 5 ans selon la province, cycle 6

Province	Nombre total d'enfants	Taille de la population cible longitudinale	Écart	Écart relatif (%)
Ontario	557 463	527 435	30 028	5,39
Québec	294 803	283 481	11 322	3,84
Alberta	151 323	145 798	5 525	3,65
Colombie-Britannique	163 351	158 153	5 198	3,18
Terre-Neuve-et-Labrador	19 664	19 106	558	2,84
Manitoba	51 610	50 230	1 380	2,67
Île-du-Prince-Édouard	5 758	5 609	149	2,59
Nouveau-Brunswick	29 056	28 474	582	2,00
Saskatchewan	43 786	43 191	595	1,36
Nouvelle-Écosse	35 709	35 258	451	1,26
Canada	1 352 523	1 296 735	55 788	4,12

Les écarts sont assez marqués en Ontario (plus de 5 %) et dans l'ensemble du pays (4,1 %). Dans chaque province, le nombre d'enfants de 2 à 5 ans en janvier 2005 est plus élevé que la somme des enfants de 0 à 1 an en janvier 2001 et des enfants de 0 à 1 an en janvier 2003.

Étant donné que les immigrants récents et les enfants nés au Canada peuvent avoir des caractéristiques différentes, le risque de biais existe bel et bien. Cette lacune dans la couverture est due à la longue période entre la sélection de l'échantillon et la collecte des données pour les cohortes de retour. En ce qui a trait aux enfants de 2 à 5 ans au cycle 6, nous estimons toujours que l'échantillon est représentatif d'un point de vue transversal, mais nous encourageons vivement les utilisateurs à ne pas perdre de vue le fait que les récents immigrants ne sont pas véritablement représentés par l'échantillon.

En outre, ce tableau n'explique pas tout. Un certain nombre d'enfants âgés de 2 à 5 ans décèdent ou quittent les dix provinces entre le moment où ils sont sélectionnés et la collecte des données du cycle 6. Ces enfants ne peuvent contribuer aux estimations transversales du cycle 6 et ne sont pas pris en compte dans la première colonne. Notre taux de sous-dénombrement de la population transversale s'établit à au moins 4,1 %. En fait, il est légèrement supérieur, étant donné que le chiffre de 55 788 représente un compte net d'immigrants et que nous devrions, idéalement, utiliser un compte brut d'immigrants. L'immigration brute se calcule comme suit : la taille de la population cible transversale moins la taille de la population cible longitudinale plus le nombre de décès et le nombre d'émigrants.

Il y a également eu des échantillons complémentaires dans certaines provinces au cycle 6. Dans ces provinces, il y a couverture partielle des immigrants récents.

D'un point de vue pratique, seule une petite proportion de récents immigrants âgés de 2 à 5 ans pourraient être sélectionnés pour l'ELNEJ parce que les provinces qui attirent le plus grand nombre d'immigrants (Ontario et le Québec) n'ont pas eu d'échantillon complémentaire.

4) Migration interprovinciale

Une autre conséquence du long laps de temps séparant la sélection de l'échantillon de la collecte des données est la migration interprovinciale. L'échantillon transversal se veut représentatif de la population d'enfants de chaque province au moment de la collecte des données. Or, en cas de déménagement, la province de résidence peut être différente de la province où résidait l'enfant au moment de la sélection de l'échantillon. Le poids initial représente la province au moment de la sélection de l'échantillon, et ce poids est susceptible d'être beaucoup plus élevé ou petit que les poids des autres enfants dans la province de résidence au moment de la collecte des données. Cela peut se solder par des estimations transversales instables.

Le scénario suivant décrit bien le problème. Des enfants sélectionnés dans des provinces où la probabilité de sélection était faible (l'Ontario et le Québec, par exemple) sont déménagés dans une petite province où la probabilité de sélection est beaucoup plus élevée. L'incidence de ces enfants pourrait être trop importante dans l'estimation en raison de leur poids d'échantillonnage excessivement élevé. Pour remédier à cette situation, on a recours à des techniques de détection des valeurs aberrantes permettant de déterminer les cas où une modification du poids initial s'impose.

Dans certains cas, les poids initiaux d'une base de sondage désuète (vieille de deux ou quatre ans) ne sont fondamentalement plus appropriés du point de vue transversal. Des mesures sont prises pour corriger ce problème au moment de l'établissement des poids transversaux.

5) Couverture non uniforme du mois de naissance

Un nouveau panel ou groupe de renouvellement est introduit chaque mois dans l'EPA, tandis qu'un autre en sort. Pour répondre aux critères relatifs à la taille de l'échantillon de la population visée pour l'ELNEJ, plusieurs groupes de renouvellement sont pris en considération et tous les ménages comptant un enfant admissible à l'ELNEJ sont retenus.

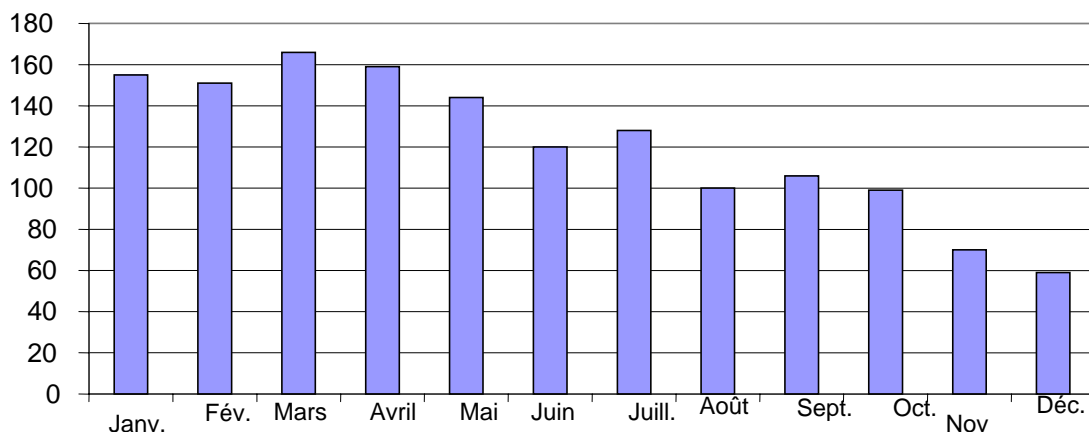
Une des difficultés rencontrées est la suivante : lorsque nous essayons de cibler des enfants de 0 an, certains groupes de renouvellement utilisés ne peuvent identifier qu'une portion des enfants nés pendant l'année de référence.

Par exemple, pour notre échantillon d'enfants de 0 an (nés en 2004), les ménages interviewés dans le contexte de l'EPA en janvier 2005 peuvent déclarer des enfants nés à n'importe quel moment de l'année 2004. Ces groupes de renouvellement couvrent entièrement les naissances survenues en 2004. Nous utilisons effectivement tous ces ménages, mais nous souhaitons inclure un plus grand nombre d'enfants de 0 an dans l'échantillon et nous sommes obligés d'avoir recours à des groupes de renouvellement qui ont quitté l'EPA en 2004 et qui ne présente qu'une couverture partielle des naissances survenues en 2004. Ainsi, les membres du groupe de renouvellement qui ont quitté l'EPA en août 2004 ne peuvent déclarer que les naissances survenues de janvier à août 2004. Nous ne savons rien des naissances qui ont eu lieu dans ces ménages de septembre à décembre 2004.

C'est pour cette raison que nous ne sommes pas en présence d'une couverture uniforme par mois de naissance pour les enfants de 0 an nés en 2000, 2002 ou 2004. Pour ces cohortes, nous avons un plus grand nombre d'enfants nés dans les

premiers mois de l'année et un plus petit nombre d'enfants nés dans les derniers mois de l'année. Les poids de l'enquête sont ajustés pour tenir compte de la couverture partielle des naissances, mais nous n'avons pas essayé de supprimer la distorsion sous l'angle du mois de naissance. L'ampleur de cet effet est illustrée dans le graphique ci-dessous.

Cycle 6 – Nombre d'enfants répondants nés en 2004, selon le mois de naissance



6) Couverture non uniforme selon l'âge en mois

Nous observons une légère surreprésentation des enfants de 24 à 28 mois au moment de l'interview du cycle 6 (variable FMCCdQ1B). Cette situation est attribuable à la méthode de collecte des données au cycle 6. Lors de la première vague de collecte des données, qui a débuté en octobre 2004, les enfants nés à l'été 2002 et sélectionnés au cycle 5, et qui avaient de 24 à 28 mois en octobre 2004, ont été suivis. Lors de la dernière vague de collecte des données, en avril 2005, les enfants sélectionnés au cycle 6 et nés à l'hiver 2003 ont été interviewés. L'âge en mois de ces enfants au moment de l'interview se situait aussi entre 24 et 28 mois.

7) Décalage entre la sélection et la collecte des données pour la cohorte originale

Au cycle 1, les logements de l'EPA comptant des enfants admissibles à l'ELNEJ formaient l'échantillon de cette dernière enquête. Il y avait un décalage de plusieurs semaines entre la sélection de l'échantillon et le début de la collecte. Une fois celle-ci commencée, certains répondants n'ont pas été contactés avant quelques autres semaines.

Si l'enfant changeait d'adresse avant que l'intervieweur de l'ELNEJ ne le joigne, l'enfant était considéré comme étant hors du champ de l'enquête. Étant donné que le décalage était relativement court, nous n'avons relevé qu'un tout petit nombre de ces cas. Néanmoins, il est possible que les enfants vivant dans des familles qui déménagent souvent soient légèrement sous-représentés chez les répondants au cycle 1.

Cette question n'entre pas en jeu en ce qui concerne les enfants introduits aux cycles 5 ou 6 puisqu'ils ont été présélectionnés et que nous avons essayé de les contacter s'ils avaient déménagé depuis la sélection de l'échantillon. Dans le cas des enfants introduits au cycle 4, des protocoles de travail sur le terrain différents ont été utilisés. Certaines personnes qui avaient déménagé ont été retracées et d'autres ne l'ont pas été.

8) Incidence sur la couverture de la diminution de la taille de l'échantillon au cycle 2

La taille de la cohorte longitudinale a été réduite au cycle 2. Certains ménages ont été retirés de l'échantillon et, au sein d'autres ménages, le nombre maximal d'enfants sélectionnés a été ramené de quatre à deux.

Pour la vaste majorité des ménages, le choix des enfants retenus était aléatoire, et les enfants maintenus dans l'échantillon reflétait fidèlement la population observée.

Dans le cas d'une partie de l'échantillon, à savoir les ménages comptant au moins un enfant de 0 ou de 1 an sélectionné à partir de l'EPA après le remaniement de 1994 plus tous les ménages du Nouveau-Brunswick, le choix des enfants à conserver dans l'échantillon afin de respecter le nouveau plafond de deux enfants par ménage n'a pas été fait de façon aléatoire. Nous avons en effet retenu les deux plus jeunes enfants pour cette partie de l'échantillon.

Il en résulte une légère distorsion par ordre de naissance. Par exemple, dans notre échantillon longitudinal du cycle 1, nous observons une proportion légèrement plus élevée d'enfants de 0 et de 1 an (âge au cycle 1) ayant au moins deux frères ou sœurs plus âgés en comparaison avec la population dans son ensemble. Dans notre échantillon longitudinal du cycle 1, il y a une proportion correspondante légèrement plus faible d'enfants de 2 à 11 ans ayant au moins deux frères ou sœurs plus jeunes.

L'incidence n'est pas négligeable au Nouveau-Brunswick, où tous les enfants de 2 à 11 ans ayant au moins deux frères ou sœurs plus jeunes ont été retranchés de l'échantillon. Au Nouveau-Brunswick en effet, l'ajustement des poids de l'enquête n'est valide que si nous estimons que les enfants qui ne sont pas les aînés d'une grosse famille peuvent représenter les aînés des grosses familles.

9) Conception et couverture de l'EPA

Le dernier remaniement de l'EPA remonte à 2004. Au fil du temps, le nombre de logements occupés dans certaines régions varie et l'efficacité du plan d'enquête se dégrade progressivement, situation qui a une légère incidence sur toutes les enquêtes qui, à l'instar de l'ELNEJ, ont l'EPA comme base de sondage.

En outre, l'EPA, bien qu'elle se révèle généralement un bon instrument pour les enquêtes sociales, a été conçue pour mesurer les caractéristiques de la population active et pas nécessairement pour évaluer les caractéristiques des enfants canadiens. Il n'y a pas de véritable problème à proprement parler. Nous ne faisons que constater que le plan d'enquête, bien que quasi optimal et très efficace en termes de coûts, n'est pas idéal dans le cas d'une enquête portant sur les enfants.

12.18 Conclusion

Divers types d'erreurs influencent la qualité des données. On s'efforce à toutes les étapes (formation des intervieweurs, contrôle de la collecte des données, traitement, pondération, etc.) de réduire les risques d'erreurs.

Nous encourageons les utilisateurs des données à déterminer en quoi les erreurs d'échantillonnage et les erreurs non dues à l'échantillonnage peuvent influencer les variables qu'ils tentent d'analyser.

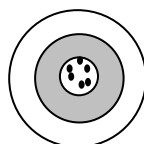
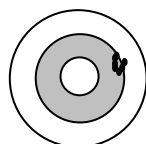
13.0 Estimation de la variance

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 visées. Une caractéristique d'intérêt à propos d'une population visée constitue un paramètre. Par exemple, la taille moyenne des enfants d'un même âge constitue un paramètre. La valeur d'un paramètre n'est connue que si on dispose des données pertinentes pour tous les individus qui composent la population visée i.e., un recensement. Nous menons une enquête telle l'ELNEJ afin d'obtenir de l'information utile se rapportant à ce paramètre auprès d'un (et un seul) échantillon (de taille restreinte) d'enfants appartenant à la population visée. L'information utile sur le paramètre contenue dans les données recueillies auprès des enfants de l'échantillon est extraite à l'aide d'un outil mathématique qu'on appelle un **estimateur**. Une valeur établie à partir des données recueillies et qui représente une approximation raisonnable de la valeur (inconnue) prise par le paramètre constitue une **estimation**; c'est tout bonnement la valeur en sortie de l'estimateur après qu'on lui ait soumis en entrée le jeu de données recueillies. En soumettant des échantillons différents en entrée à l'estimateur on obtient en sortie une série correspondante d'estimations du paramètre. L'élément clé de l'échantillonnage est le suivant : on va jauger la performance d'un estimateur en étudiant mathématiquement la qualité des approximations qui seraient produites si on lui fournissait en entrée tous les échantillons possibles. Puisque ces estimations sont obtenues à partir d'échantillons d'individus seulement, elles ne coïncideraient pas (autrement que par chance) avec la valeur du paramètre si elle nous était révélée; c'est là l'erreur liée à l'échantillonnage (on dit aussi erreur d'échantillonnage). Par définition, un recensement n'admet pas d'erreur liée à l'échantillonnage puisque le seul échantillon possible est celui dont on disposera, à savoir ici la population entière (et la seule estimation raisonnable qu'on puisse alors faire correspondra, par définition, à la valeur du paramètre).

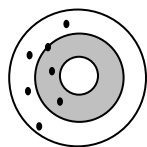
Même si dans un contexte donné on ne parvient pas à quantifier l'erreur liée à l'échantillonnage, il est généralement possible de décrire certaines de ses caractéristiques. Par exemple, plus grande est la taille de l'échantillon sélectionné, plus petite généralement est la magnitude de l'erreur liée à l'échantillonnage ; cela provient du fait qu'un grand échantillon recèle de l'information utile à propos d'une grande partie de la population visée. Parmi d'autres facteurs qui influent sur la magnitude de l'erreur liée à l'échantillonnage on compte notamment la taille de la population visée, les règles qui sous-tendent la sélection de l'échantillon (ces règles forment le plan d'échantillonnage) et la variabilité dans la population visée des variables pertinentes à l'estimation du paramètre d'intérêt sur lesquelles reposent l'estimation.

Il y a deux composantes à l'erreur liée à l'échantillonnage: le biais et la variance liée à l'échantillonnage. En termes simples, un estimateur ne présente pas de biais lorsque la moyenne de toutes les estimations qu'il produit (et qui, rappelons-le, s'obtiennent en fournissant en entrée tous les échantillons possibles) correspond précisément à la valeur du paramètre recherchée. Ainsi, les estimations prises individuellement peuvent toutes être différentes de la valeur du paramètre mais, prises collectivement, lui correspondre parfaitement. Dans un tel cas l'estimateur est dit être sans biais pour le paramètre en question (et dans l'autre cas on le dit être biaisé). L'autre composante de l'erreur liée à l'échantillonnage est la variance qui exprime dans quelle mesure toutes les estimations possibles diffèrent entre elles.

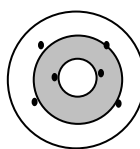
L'analogie bien connue qui suit est utile pour comprendre ces concepts. Un lanceur de dards (l'estimateur) effectue une série de lancers vers une cible; chacun des impacts de dard sur la cible correspond à une estimation. On ne s'attend assurément pas à ce que tous les lancers atterrissent en plein centre de la cible ; chacun des lancers hors du centre contribuera donc à accroître l'erreur liée à l'échantillonnage. Il y a principalement 4 scénarios possibles qui décrivent l'ensemble des impacts de dards sur la cible ; ils diffèrent entre eux par la magnitude associée à chacune des deux composantes de l'erreur liée à l'échantillonnage : le biais et la variance.



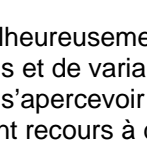
Biais présent mais faible variance



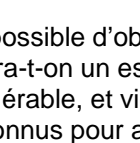
Absence de biais et faible variance



Biais présent et forte variance



Absence de biais mais forte variance



En pratique, malheureusement, il est généralement impossible d'obtenir un estimateur qui performe bien en terme de biais et de variance à la fois. Ainsi, trouvera-t-on un estimateur avec une faible variance seulement pour s'apercevoir qu'il admet un biais considérable, et vice-versa. Les statisticiens d'enquête ont généralement recours à des estimateurs qui sont connus pour avoir peu ou pas de biais; ils préfèrent donc généralement réduire la composante d'erreur qu'est le biais à son minimum et faire ensuite ce qu'ils peuvent au niveau de la variance. Cela sera l'hypothèse de travail que nous ferons ici; par conséquent, il sera question dans le reste de ce chapitre que de variance puisque nous supposerons que le biais a déjà été réduit à un minimum de part le choix de l'estimateur sur ses concurrents.

Afin de mesurer exactement la magnitude de la variance liée à l'échantillonnage d'un estimateur, il nous faudrait disposer de toutes les estimations que ce dernier puisse produire. En pratique, cependant, on ne dispose que d'un seul échantillon (et donc d'une seule estimation), celui pour lequel le travail de collecte des données sur le terrain sera mené. Bien qu'on ne puisse donc pas mesurer exactement la variance, il est souvent possible d'en établir une approximation à l'aide du seul échantillon prélevé. Le lecteur trouve sans doute suspect qu'on prétende pouvoir établir une approximation de la variabilité de toute une gamme d'estimations à partir d'une seule d'entre elles. Bien que le portrait global de la problématique d'estimation de variance soit plutôt compliqué, l'idée maîtresse se révèle dans un cas particulièrement simple. On peut en effet démontrer que pour un estimateur simple comme celui pour la moyenne, lorsqu'il est employé avec un plan d'échantillonnage lui aussi simple, sa variance est fonction d'une seule quantité inconnue: la variabilité de la variable d'intérêt dans la population. Ainsi, dans ce cas de base, établir une approximation valable de la variance de l'estimateur revient tout bonnement à obtenir une approximation valable de la variabilité de la variable d'intérêt dans la population, et à rien d'autre. Certes, la taille de la population et celle de l'échantillon influent aussi mais ces deux quantités nous sont *connues*. Or, dans ce cas des plus simples, l'échantillon sélectionné fournit bel et bien l'ingrédient nécessaire pour construire une approximation valable de la variabilité recherchée. En effet, on peut démontrer que la variabilité observée parmi les valeurs échantillonnées est un bon indicateur de la variabilité dans toute la population. Il en est essentiellement ainsi parce qu'un tel échantillon est vraisemblablement bien balancé : il contient sensiblement les mêmes proportions de petites, moyennes et grandes valeurs pour la variable d'intérêt qu'on retrouve dans la population entière. Dans les cadres plus complexes, il n'existe généralement pas de relation directe de cette sorte entre la variabilité des valeurs dans la population et l'échantillon qui puisse être exploitée pour obtenir une estimation de la variance liée à l'échantillonnage. Il existe néanmoins des méthodes comme le bootstrap⁸ qui exploitent de façon astucieuse l'information contenue dans un seul échantillon afin d'établir une estimation de la variance d'un estimateur à la façon que nous y sommes parvenus ci haut dans un cadre très simple. Dans ce chapitre, nous expliquons pourquoi il est important de quantifier la variance liée à l'échantillonnage et nous présentons quelques façons d'y arriver.

13.1 Quelques emplois des termes erreur d'échantillonnage et variance

Il y a parfois de la confusion en ce qui a trait à l'emploi des termes variance dans la population, variance d'échantillonnage, écart-type et erreur-type. Le but de cette section est de clarifier le tout.

⁸ Il appert que le terme français approprié pour désigner le bootstrap est « estimation de variance par auto-amorçage » ; nous utiliserons ici la dénomination anglaise car elle est celle d'emploi courant en français.

Malheureusement, le terme « variance » ne signifie pas une seule et même chose lorsque employé dans les expressions suivantes : « variance dans la population » et « variance liée à l'échantillonnage » ; c'est de ce double emploi qu'origine toute la confusion. **La variance dans la population** (d'une variable) est définie comme étant la variabilité observée dans les valeurs prises par la variable dans la population visée. Par exemple, dans la population des garçons au Canada qui ont présentement 10 ans, il existe une certaine variabilité quant à leur taille mesurée en centimètres; c'est donc là la variance dans la population de la taille en centimètres pour ces garçons.

Mais cet emploi du terme « variance » n'est pas approprié ici car la variable en question n'a rien à voir avec un processus aléatoire. Puisque dans le cadre statistique des enquêtes on considère que les variables d'intérêt ont des valeurs fixes (connues ou non), nous proposons plutôt de parler ici de **dispersion de la variable dans la population**. Ainsi, dans l'exemple des garçons que nous venons de considérer, nous dirions plutôt qu'il existe une certaine dispersion dans les tailles mesurées en centimètres des garçons au Canada qui ont présentement 10 ans.

La définition mathématique généralement donnée de la dispersion de la variable y dans la population est :

$$\sigma^2 = \frac{\sum_{i=1}^N (y_i - \bar{Y})^2}{N - 1}$$

où y_i est la valeur de la variable Y pour l'individu i ;

\bar{Y} est la moyenne des y_i de la population; et

N désigne la taille de la population.

Présentée ainsi, la dispersion d'une variable dans la population n'est rien d'autre qu'un paramètre de la population parmi d'autres, comme l'est la moyenne de la variable Y , \bar{Y} , par exemple. Le lecteur notera que le symbole employé pour désigner la dispersion de la population est

$\sigma^2 = \sigma \times \sigma$ et non simplement σ , par exemple. C'est là une façon de se rappeler que la mesure de dispersion ne s'exprime pas dans les mêmes unités que la variable elle-même, mais plutôt en terme de son carré. Par conséquent, pour obtenir une mesure de dispersion qui soit elle exprimée dans les mêmes unités, il apparaît naturel de prendre la racine carrée de la mesure actuelle de dispersion. Et c'est précisément là l'usage en statistique : la racine carrée de la dispersion est **l'écart-type de la variable dans la population**.

Nous avons déjà défini ce qu'est la **variance due à l'échantillonnage**. C'est la variabilité qu'on observerait dans les estimations d'un même paramètre issues si tous les échantillons possibles pouvaient être considérés. Puisque la sélection d'échantillons est un processus aléatoire, l'emploi du terme « variance » ici s'en trouve justifié (contrairement à la variance d'une variable dans la population). Tout comme pour la mesure de dispersion, la variance due à l'échantillonnage ne s'exprime pas dans les mêmes unités que les estimations elles-mêmes. Ainsi, si les estimations sont en terme de \$, alors la variance due à l'échantillonnage s'exprimera en terme de \$². Il apparaît naturel de prendre la racine carrée de la variance afin d'obtenir une mesure de variabilité qui s'exprime en terme des mêmes unités que les estimations elles-mêmes; la quantité résultante de cette opération s'appelle **l'erreur-type**.

13.2 Coefficient de variation

Une mesure de la variance due à l'échantillonnage nous renseigne sur la qualité d'une estimation. Cependant, l'erreur-type lui est généralement préférée parce qu'elle s'exprime dans

les mêmes unités que l'estimation elle-même. Mais l'erreur-type, tout comme la variance, présente un fâcheux inconvénient : les erreurs-types rattachées à des estimations différentes ne sont pas comparables bien que dans plusieurs situations elles le devraient. La comparabilité de mesure de qualité d'estimations est un atout : est-ce que l'estimation obtenue pour un paramètre pour la province *A* est plus ou moins fiable que l'estimation correspondante obtenue pour la province *B*?

Afin de mettre en relief cet inconvénient, supposons qu'un échantillon ait été tiré pour lequel une collecte subséquente de données ait fournie toute l'information utile pour l'estimation du total d'une certaine variable *Y*. Supposons que le cadre d'échantillonnage soit tel qu'une estimation de la moyenne cette fois de *Y* puisse s'obtenir tout simplement en divisant le total déjà estimé par la taille (supposée être connue, ici, pour nos fins) de la population *N*. L'idée ici est la suivante : les deux estimations, soit celle du total et celle de la moyenne, exploite la même information de l'échantillon et ce de la même façon (via le total estimé). Par conséquent, si les erreurs-types associées à l'estimation du total et de la moyenne étaient comparables (i.e., l'erreur-type serait alors un instrument adéquat de mesure de qualité *relative* d'estimations), alors elles devraient être ici *égales*. Cependant elles ne le seront pas. En effet, l'erreur-type associée à l'estimation du total sera *N* fois trop grande par rapport à celle de la moyenne puisqu'à la base les estimations correspondantes diffèrent l'une de l'autre par un facteur multiplicatif de *N*.

Une façon d'obtenir une mesure relative de précision, une qui fournit la même valeur dans des situations comparables, s'obtient en considérant le coefficient de variation (CV). Le coefficient de variation s'obtient en divisant l'erreur-type de l'estimation $s.e.(\hat{\theta})$ par l'estimation elle-même, $\hat{\theta}$:

$$CV(\hat{\theta}) = \frac{s.e.(\hat{\theta})}{\hat{\theta}}$$

Maintenant, contrairement à la variance associée à une estimation, le coefficient de variation, lui, permet à l'analyste de comparer la précision d'estimations entre elles pour des variables du même type que le revenu⁹. Retournons à notre exemple d'estimation d'un total et d'une moyenne. Dans ce cas on peut voir que les CVs correspondants seraient égaux. En effet, bien que le numérateur du CV du total (i.e., l'erreur-type qui lui est associée) soit *N* fois plus grand que le numérateur du CV de la moyenne, il en va de même des dénominateurs. Par conséquent, le facteur multiplicatif *N* est neutralisé lorsqu'il entre dans la composition du CV.

Bien que le CV est tout indiqué pour mesurer la fiabilité sous échantillonnage d'un estimateur d'un paramètre tel que le total d'un revenu, il présente des inconvénients qu'il convient de connaître et qui limitent son applicabilité dans le cas de proportions. Quelques-uns de ces inconvénients sont présentés ci-bas.

Enjeu 1 : Coefficients de variation pour des proportions très petites ou très grandes

Comme l'erreur type pour une proportion *p* est la même que l'erreur type pour une proportion $(1-p)$, les CVs pour *p* et $(1-p)$ peuvent être très différents, simplement à cause des dénominateurs *p* et $(1-p)$ respectivement. Quelqu'un peut penser à un scénario où le CV peut être très grand avec une valeur de *p* très petite, mais par contre un CV excellent pour $(1-p)$. Utilisons l'exemple suivant pour illustrer. Pour un cas donné, vous avez calculé que l'erreur type des estimations *p* et $(1-p)$ est 0,0475, pour une valeur de *p* de 0,95. Le CV pour la valeur de *p* de 0,95 serait:

⁹ Nous entendons par là une variable d'intérêt qui soit continue et positive, par opposition notamment à une variable dichotomique.

$0,0475 / 0,95 = 0,05$ ou 5 %, ce qui est un très bon CV.

D'autre part, le CV pour la proportion $(1-p)$ est:

$0,0475 / 0,05 = 0,95$ ou 95 %. Ceci est un très mauvais CV.

Enjeu 2 : Sur les règles en circulation guidant l'évaluation de la qualité d'estimations dans le cas de proportions

Des règles circulent quant à l'interprétation du CV d'une estimation sous la forme d'une proportion pour en déterminer la fiabilité. Typiquement, ces règles stipulent qu'une proportion estimée accompagnée d'un CV de moins de 16.5% est de bonne qualité (les conclusions tirées de telles estimations sont considérées fiables); un CV entre 16.5% et 33% représente une estimation de qualité acceptable (les conclusions tirées de telles estimations doivent être considérées avec prudence); un CV au-delà de 33% représente une estimation de piètre qualité (les conclusions tirées de telles estimations sont purement exploratoires et ne doivent pas autrement être considérées). Ces règles peuvent être très utiles pour guider l'analyste dans son travail à partir de données d'enquêtes mais elles requièrent de la vigilance, faute de quoi des erreurs sur l'énoncé de qualité des estimations seront commises. Une consultation auprès d'un statisticien d'enquête est sans nul doute ici la façon appropriée de tirer le maximum des résultats obtenus puisque plusieurs facteurs spécifiques à l'analyse interviennent pour déterminer la validité des conclusions tirées. Dans le cas de proportions, ces règles doivent être employées avec beaucoup de discernement, comme l'illustre les exemples qui suivent.

Exemple 1 : L'intervalle de confiance à 99 % pour une proportion estimée de 0,5 avec un CV de 31 % est (0,10 à 0,90). En appliquant les lignes directrices sur la qualité qui ont été publiées auparavant, cette estimation se retrouve dans la catégorie **marginale**. Par contre, l'intervalle de confiance est tellement large que l'estimation ne veut réellement rien dire.

Exemple 2 : Supposons que nous avons un échantillon de 27 000 individus et 44 de ceux-ci possèdent une caractéristique que nous voulons étudier. Avec les poids de sondage, nous calculons que 0,16 % de la population possède cette caractéristique, avec un CV correspondant de 34 %. Un CV de 34 % est considéré **inacceptable** en appliquant les règles énoncées ci haut. Même si 0,16 % est une très petite proportion, si nous construisons un intervalle de confiance de 99 %, nous obtenons (0,02 %, 0,30 %). Ceci signifie que nous avons une grande confiance que la vraie valeur est, disons, plus petite que 0,5 %. Évidemment, les utilisateurs doivent être conscients que la vraie valeur peut être 0,05 % ou 0,25 % or même 0,30 %. Selon le but de la recherche, peut-être qu'une affirmation que le taux est inférieur à 0,5 % est utile. Donc, appliquer aveuglément les lignes directrices de qualité et rejeter cette estimation à cause du CV élevé peut ne pas être approprié.

Par conséquent, si les utilisateurs veulent utiliser le CV comme mesure de l'erreur d'échantillonnage avec des proportions, ils sont fortement encouragés à calculer les CV tant pour la proportion p que pour $(1-p)$. Les CV pour des proportions, particulièrement pour des proportions extrêmes, peuvent être trompeuses, comme on l'a vu dans les exemples précédents. Nous ne recommandons pas la comparaison de la qualité d'estimations sous la forme de proportions en terme de CV. Nous suggérons plutôt de construire des intervalles de confiance et les utiliser afin de tirer des conclusions statistiquement valides.

De la recherche est présentement menée afin de trouver une alternative au CV comme outil d'évaluation de la fiabilité d'estimations dans le cas de proportions extrêmes (près de 0 ou de 1). Cependant, aucune solution pleinement satisfaisante n'a encore été trouvée. Par conséquent,

les analystes doivent faire preuve de prudence s'ils veulent utiliser les CV pour la détermination de la fiabilité d'estimations sous la forme de proportions.

13.3 À propos de l'importance de rapporter la variance due à l'échantillonnage

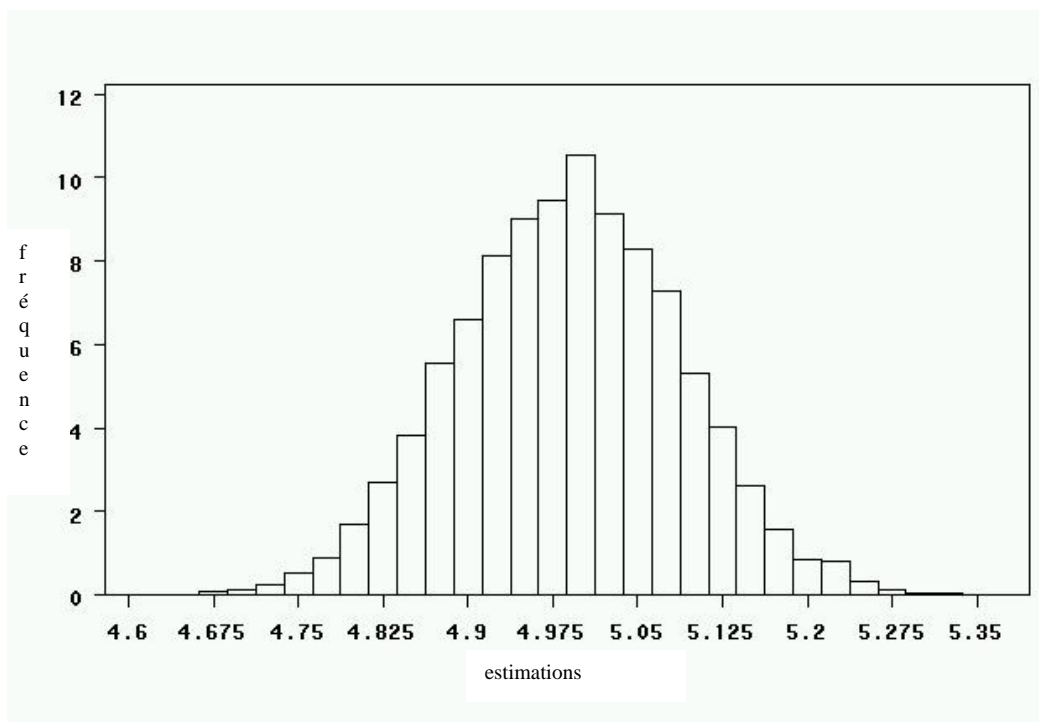
Comme nous l'avons vu dans les sections précédentes, une estimation sera selon toute vraisemblance différente de la vraie valeur qu'on cherche à estimer; la différence numérique entre ces deux valeurs constitue l'erreur. Il faut donc supposer chaque fois qu'une estimation est produite qu'elle s'accompagne d'une erreur. **Par conséquent, une indication de la magnitude ou de l'étendue de cette erreur doit être fournie à ceux qui consulteront les estimations produites.** En effet, divulguer des estimations d'enquête sans fournir d'indication sur l'erreur encourue induira en erreur le lecteur. Une composante de cette erreur est l'erreur due à l'échantillonnage et l'autre est l'erreur non-due à l'échantillonnage. Cette dernière comprend les erreurs de réponse, les erreurs liées au traitement des données. En pratique, la magnitude de l'erreur non-due à l'échantillonnage est très difficile voire même impossible à établir. (Lorsque l'impact d'erreurs ne peuvent pas être quantifiés, comme c'est le cas avec la plupart des erreurs non-dues à l'échantillonnage, alors elles doivent être clairement rapportées et décrites en détails.) En revanche, la magnitude de l'erreur due à l'échantillonnage peut généralement être quantifiée pourvu qu'on parvienne mathématiquement à suivre les effets de l'aléa dans la construction des estimations. On mesure l'erreur due à l'échantillonnage d'une estimation en calculant la variance due à l'échantillonnage du processus aléatoire qui l'a engendrée. Il faut donc faire l'effort de calculer la variance puisque cette mesure d'erreur est généralement disponible et la rapporter avec les estimations afin qu'elle puisse servir à juger de leur fiabilité.

13.4 Calcul de la variance due à l'échantillonnage

Il serait difficile, pour ne pas dire impossible, d'obtenir une formule exacte qui permettrait de calculer la variance due à l'échantillonnage pour l'ELNEJ en raison de son plan de sondage complexe, les ajustements pour contrer les effets de la non-réponse, et des unités hors-champ, et la post-stratification. En fait, une telle tâche ne pourrait être entreprise qu'en ayant recours à d'importantes hypothèses simplificatrices qui mineraient sérieusement la validité des résultats obtenus. Une bonne façon de produire une approximation de la variance due à l'échantillonnage dans un contexte tel celui de l'ELNEJ est de recourir à la méthode du bootstrap¹⁰.

Afin de comprendre ce qu'accompli le bootstrap, il faut introduire le concept de distribution sous échantillonnage d'un estimateur. Nous avons déjà noté que si tous les échantillons possibles pouvaient être soumis à l'extracteur d'information qu'est l'estimateur alors nous aurions toutes les estimations possibles. Supposons maintenant que nous ayons justement en main toutes ces estimations. Nous pourrions alors établir le diagramme de fréquences des estimations; c'est ce qu'on appelle la distribution sous échantillonnage de l'estimateur. Voici un exemple d'un tel histogramme, d'une telle distribution sous échantillonnage.

10 Le défi ici est de parvenir à adapter le bootstrap avec succès à la statistique d'enquêtes alors qu'il a avant tout été conçu pour un tout autre cadre statistique. Notre compréhension du bootstrap dans un cadre d'enquête, et comment l'implanter adéquatement, n'a cessé d'évoluer au cours des 10 dernières années, soit depuis la conception même de l'ELNEJ.



Nous remarquons en consultant l'histogramme que, par exemple, des estimations très petites, tout comme les estimations très grandes, sont rares, et que les estimations les plus fréquemment produites sont au centre de la plage de valeurs possibles. Dans cet exemple, les estimations sont distribuées de façon presque symétrique autour de la valeur 5 (qui serait alors la valeur du paramètre inconnu si l'estimateur était sans biais) et les estimations supérieures à 5.275, par exemple, sont rarissimes (elles surviennent moins de 0,5% du temps). Le fait que les estimations ne sont pas toutes concentrées en un point, mais s'étendent plutôt en une plage de valeurs, est la représentation graphique de la variance due à l'échantillonnage. En effet, par définition, la variance due à l'échantillonnage est une mesure de la variabilité qu'on observe au sein des estimations produites par un estimateur. Nous voyons aussi que la variance due à l'échantillonnage est un attribut, un parmi d'autres, de la distribution de l'estimateur sous échantillonnage.

Le bootstrap tente de recréer cette distribution en procédant à un certain ré-échantillonnage du seul échantillon observé. De cette correspondance des distributions (celle de ré-échantillonnage avec celle d'échantillonnage) on peut tirer des estimations de tous les paramètres de la distribution d'échantillonnage qui nous intéressent, un d'eux étant sa variance qui correspond à la variance due à l'échantillonnage. Tout comme l'échantillonnage qui a donné lieu à l'échantillon observé s'accompagne de poids de sondage, le ré-échantillonnage qui sous-tend le bootstrap donne lui aussi lieu à des "poids de sondage"; ce sont les poids bootstrap. Alors que les poids de sondage servent eux à la construction d'une estimation d'un paramètre d'intérêt, les poids bootstrap mènent eux à la construction d'une estimation de la variance liée à cette estimation.

Pour l'ELNEJ, un ensemble de 1,000 poids bootstrap est disponible. Afin d'obtenir en pratique une estimation de la variance d'un estimateur par le bootstrap, il faut d'abord produire les 1000 estimations qui sont induites par les 1000 poids bootstrap. Ensuite, la variabilité observée dans ces 1000 estimations constitue l'estimation de variance par le bootstrap. Algébriquement, on a :

$$\hat{v} = \frac{1}{1000} \sum_{k=1}^{1000} (\theta_k - \theta^*)^2$$

où θ_k est la k-ième estimation bootstrap et θ^* est l'estimation originelle pour θ obtenue à partir de l'échantillon. (Note: en pratique on peut la plupart du temps substituer à θ^* dans le calcul de la variance bootstrap la valeur moyenne des estimations bootstrap θ_k . Cette simplification est utilisée chaque fois qu'on calcule la variance bootstrap à partir de PROC SUMMARY en SAS en ayant recours à la statistique VAR.)

Deux outils, tous deux s'appuyant sur les poids bootstrap, ont été développés afin d'assister les utilisateurs dans le calcul de la variance due à l'échantillonnage et du CV liés à des estimations; ils sont :

- Un chiffrier Excel muni d'une interface développée en Visual Basic for Applications (VBA) qui permet à l'utilisateur d'obtenir facilement des variances approximatives pour des proportions diverses et ce pour un certain éventail de domaines.
- Des macros SAS qui permettent le calcul de la variance due à l'échantillonnage par le bootstrap.

Le choix de l'outil approprié repose sur le type d'analyse qui est menée et sur le niveau de précision requis.

Dans les cycles précédant le cycle 5, un troisième outil était disponible: des tableaux de CV approximatifs. Ces tableaux étaient construits en faisant l'hypothèse d'un effet de plan « représentatif » qui permettait une approximation grossière de la variance. Cet outil n'est plus disponible puisque les deux autres outils le supplantent en flexibilité et en fiabilité.

13.4.1 Chiffrier de variances approximatives pour des proportions

Un ensemble de chiffriers est disponible aux utilisateurs pour leurs permettre de calculer (approximativement) la variance due à l'échantillonnage pour des estimations de proportions. Disponible sous format électronique Excel, l'interface développée en Visual Basic for Applications (VBA) présente les estimations de variance due à l'échantillonnage et coefficients de variation pour des milliers de domaines. Quelques domaines sont définis selon l'âge, le groupe d'âge, la province ou la région (groupes de provinces et tout le pays), et selon les différents croisements d'âge et de province. Les tailles d'échantillon pour chaque domaine sont aussi disponibles.

Les détails sur la façon dont les chiffriers ont été produits, sur ce qu'ils contiennent 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 due à l'échantillonnage à partir des poids bootstrap : Bootvar

Des macros ont été développées, en SAS et en SPSS, pour calculer la variance due à l'échantillonnage avec les poids bootstrap; elles forment Bootvar. En SAS la version la plus récente est 3,1 alors que pour SPSS c'est 3,0. Bootvar peut être téléchargé via « outils de recherche » au www.statcan.ca/francais/rdc/whatdata_f.htm. Bootvar permet de calculer la variance due à l'échantillonnage pour une grande variété d'estimations incluant les totaux, les ratios, la différence des ratios, les coefficients de régressions linéaire et logistique, et ce, quel que soit le domaine. En plus de tenir compte du plan de sondage, cette façon de calculer une estimation de la variance due à l'échantillonnage tient également compte des caractéristiques spécifiques de la variable d'intérêt.

Contrairement à l'outil du chiffrier Excel, nous ne sommes pas limités ici qu'à certains domaines pré-identifiés.

Le programme Bootvar présente beaucoup d'avantages mais exige un peu plus d'effort de la part du chercheur qui l'utilise. Le calcul de la variance due à l'échantillonnage avec l'utilisation de ces macros demande plus de temps que le recours au chiffrier et l'utilisateur doit évidemment se familiariser avec les macros de Bootvar avant de les soumettre. Toutefois, ces macros ont été conçues de façon à ce qu'elles soient faciles à utiliser. L'utilisateur doit évidemment avoir accès aux macros, aux fichiers de données et aux fichiers de poids bootstrap. Cet accès est possible dans les Centres de données de recherche (CDR) de Statistique Canada. De plus, de la documentation détaillée existe dans les CDR qui décrit comment utiliser les macros SAS ou SPSS pour le calcul de la variance due à l'échantillonnage.

Malgré le temps requis plus long pour faire les calculs, Bootvar est à préférer au chiffrier Excel pour obtenir une estimation de la variance d'estimations qui doivent être publiées dans des articles analytiques. Bootvar permet d'obtenir une variance qui est beaucoup plus précise et davantage justifiée.

L'utilisateur trouvera toute l'information utile à l'emploi de Bootvar dans la documentation qui l'accompagne.

13.4.3 Autres outils informatiques

D'autres logiciels commerciaux ont la capacité d'incorporer les poids bootstrap dans leurs calculs de la variance due à l'échantillonnage. Deux de ces logiciels sont disponibles au CDR : SUDAAN (avec l'option DESIGN=BRR) et WesVar. Une autre possibilité est de considérer STATA9.

Si l'analyste veut calculer la variance due à l'échantillonnage pour des estimations autres que celles comprises dans Bootvar, il peut également écrire ses propres macros qui incorporent la méthode bootstrap bien que ce ne soit pas là un exercice trivial.

13.4.4 Linéarisation de Taylor et autres techniques

Les fichiers de poids bootstrap contiennent les variables indiquant l'unité primaire d'échantillonnage (UPE) et la strate à partir desquelles l'individu a été sélectionné. Certains logiciels, incluant Stata, SUDAAN et SAS, ont des procédures qui utilisent l'information du plan de sondage (la strate et UPE) avec les poids de sondage pour calculer les estimations de la variance due à l'échantillonnage. La technique est connue sous plusieurs noms: la linéarisation de Taylor ou Binder ou l'estimation robuste. Ces procédures requièrent au moins deux UPE par strate, une condition qui n'est pas toujours satisfaite par l'ELNEJ. Bien que les strates puissent être combinées, ce n'est pas clair en ce moment si le résultat est satisfaisant car une comparaison des variances dues à l'échantillonnage obtenue par bootstrap avec celles obtenues avec cette méthode n'a pas encore été faite. De plus, plusieurs ajustements aux poids de départ sont requis afin d'obtenir les poids pour diffusion et l'approche par taylorisation ne permet pas de tenir compte de l'impact qu'ont ces ajustements sur la variance. Par conséquent, il est recommandé d'utiliser un des outils déjà décrits dans ce chapitre (le chiffrier Excel ou les poids bootstrap) pour obtenir les estimations de variance due à l'échantillonnage qui tiennent compte de façon appropriée du plan de sondage.

Finalement, les logiciels comme SAS ou SPSS ont déjà des procédures pour calculer la variance due à l'échantillonnage des estimations (par exemple, PROC UNIVARIATE en SAS). Cependant, plusieurs de ces procédures ne tiennent pas compte du plan de sondage (par exemple, la stratification) **même** si on a recours à l'énoncé WEIGHT; la

variance obtenue est généralement une sous-estimation de la variance due à l'échantillonnage. Pour cette raison, ces procédures ne sont pas recommandées car elles peuvent mener à de fausses conclusions.

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 profanes 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.

On effectue, dans le cadre de l'ELNEJ, des évaluations des enfants âgés de 4 ans et plus. Ces évaluations sont décrites dans le présent chapitre.

14.1 Échelle de vocabulaire en images de Peabody

L'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, d'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 fondé sur le calcul des 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 groupes d'âge différents. 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 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 des scores pour la population entière une distribution à partir d'échantillon représentatif, appelé l'échantillon de norme, est souvent utilisée. Chaque cycle de 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. Malgré que chaque échantillon est choisi de façon probabiliste, avec différente période de référence, les experts dans le domaine des mesures cognitives pourraient être en désaccord afin de conclure que les différences entre les distributions estimées d'un échantillon à l'autre reflète une véritable différence de population à travers le temps ou résulte simplement de l'erreur d'échantillonnage.

Pour la robustesse, nous avons choisi de ne pas utiliser qu'un échantillon normatif mais 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 6 ont donc été calculés à l'aide des mêmes normes que celles utilisées pour le cycle 4 (v2) et le cycle 5¹¹. Pour le cycle 6, 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 employées pour créer les normes : 1 enregistrement au cycle 2, 49 enregistrements au cycle 3, et 6 enregistrements au cycle 4. Ces scores de zéro étaient probablement les essais 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 manière 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 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 visée 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 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

11 Notez qu'aux cycles 1, 2 et 3 des échantillons de norme différents furent utilisés. Quoique de petites variations soient attendues, celles-ci sont bien à l'intérieure de l'erreur d'échantillonnage provenant des mesures de différents échantillons de norme.

des 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.

Dans le cas de 23 enfants, on n'a pas terminé le test sur le terrain selon les règles d'utilisation de l'outil, et l'application n'a pu attribuer de score. Ces enfants ont obtenu un score de 0, même s'ils ont répondu à un nombre appréciable de questions dans certains cas. On a pu calculer un score brut approximatif pour ces enfants en utilisant le TRI. Pour les autres répondants – c.-à-d., 99 % entre eux -, on a utilisé le score produit par l'application conformément aux règles initiales de l'EVIP.

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 introduit un bruit important.

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 employé 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 à noter que ce groupe peut comprendre des enfants d'âge effectif 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 employés 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 aux chapitres 10.0 et 13.0.

14.2 Évaluation de la connaissance des nombres

L'évaluation de la connaissance des nombres vise à cerner l'évolution de la compréhension des nombres chez les enfants par 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 D^r Robbie Case d'University of Toronto en collaboration avec ses collègues, notamment Yukari Okamoto de l'University of California at Santa Barbara. Elle se fonde sur la théorie des structures conceptuelles centrales élaborée par D^r Case pour expliquer le développement de la pensée de l'enfant. Avant sa mort en mai 2002, D^r Case travaillait à l'adaptation du test aux fins de l'ELNEJ. Après le décès de D^r 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 D' Case, on peut distinguer quatre niveaux de développement de la compréhension des nombres par les enfants : niveau prédimensionnel, niveau unidimensionnel, niveau bidimensionnel et 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 et 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 D' Case et D' 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 manière 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 % des chances de répondre correctement en ne se fiant qu'au hasard¹². Le fait d'exiger que les enfants répondent correctement aux deux volets de tels items pour obtenir un point accroît l'assurance que ceux-ci ont effectivement les connaissances requises par ces items.

Les enfants n'ont pas le droit d'utiliser un crayon et du papier pour répondre aux questions, qui 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 6 : le « score d'âge » (FKNCdS01), le « score brut de 30 points » (FKNCfS01) et le « score d'âge brut normalisé de 30 points » (FKNCfS02).

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 (un maximum d'un point par niveau complété peut être attribué). 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

12 Par exemple, on demande au volet a) laquelle de deux piles de jetons est la plus grande et, au volet b), laquelle est la plus petite.

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 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, a) et b). Le score brut de 30 points correspond simplement au nombre total de réponses exactes à ces questions.

Un score d'âge brut normalisé de 30 points 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 relatifs à 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 manière que la moyenne des scores normalisés corresponde à 100 et que l'écart-type s'établisse 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 D^r Okamoto pour s'assurer que ces procédures concordaient avec la théorie élaborée par D^r 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 aux chapitres 10.0 et 13.0.

14.3 Qui suis-je?

Le module Qui suis-je ? 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 D^{re} Molly de Lemos et ses collègues de l'Australian Council for Educational Research (ACER). L'équipe de l'ELNEJ a collaboré étroitement avec D^{re} 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 qu'a l'enfant de conceptualiser et de reconstruire une forme géométrique. Les tâches d'écriture évaluent la capacité qu'a l'enfant de comprendre et d'employer 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'employer 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 D^{re} 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 dans 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 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. Le chef de l'équipe de notation a également consulté un représentant de l'ACER au sujet des procédures de notation.

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 (FWICdS02) correspond à la somme des scores attribués aux tâches liées à la reproduction d'un symbole. Le score de l'échelle des symboles (FWICdS03) 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? (FWICdS01), 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 seront également disponibles pour la première fois au cycle 6 : le score normalisé total de l'échelle Qui suis-je? (FWICfS01), le score normalisé de l'échelle de reproduction (FWICfS02) et le score normalisé de l'échelle des symboles (FWICfS03). 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? 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 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 manière que la moyenne des scores normalisés corresponde à 100 et que l'écart-type s'établisse 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 D^e de Lemos, il est nécessaire d'attribuer un score dans les cas des réponses codées « 0 » (pas d'essai) lors de la sommation des scores des tâches de Qui suis-je? pour 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 D^e 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 ? d'autres études 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 D^o de Lemos.

Toutes les preuves indiquent que le test était valide et qu'il devrait fournir aux utilisateurs 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 aux chapitres 10.0 et 11.0.

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. La version abrégée du test mise au point pour les fins de l'ELNEJ se compose de 20 questions à chaque niveau, depuis le cycle 5.

À chaque niveau, le test subi au cycle 6 est le même qu'au cycle 5. Aucune modification n'a été apportée au nombre de questions ni aux questions elles-mêmes. Au cycle 6, les tests de mathématiques ont tous été subis entre janvier et juin 2005. C'est pourquoi l'ajustement exécuté au cycle 5 pour prendre en compte l'évaluation précédente n'a pas dû être repris au cycle 6.

Calcul des scores

Tous les enfants qui ont subi le test de mathématiques se sont vu attribuer un score brut (FMACS01), un score gradué (que l'on appelle « score gradué classique ») (FMACS02) et un score gradué selon la TRI (FMACS03). On obtient le score brut en additionnant simplement le nombre de réponses correctes. Le score gradué classique et le score gradué selon la TRI sont décrits ci-dessous.

Au cycle 2, un approche TRI a fait ses preuves dans le calcul des scores des tests de compréhension de lecture. La TRI est une méthode de mesure couramment employé dans les tests psychométriques et de connaissances. Le calcul des scores de tests selon la TRI consiste à déterminer la valeur la plus probable ou la valeur prévue correspondant aux aptitudes du répondant. On présume que la probabilité d'une réponse correcte à une question suit une fonction logistique donnée des capacités du répondant. Cette probabilité prend la forme d'une courbe en S pour la gamme des capacités. La forme de la courbe dépend de la difficulté de la question et, parfois, de son pouvoir discriminatif (dans le modèle de la TRI à deux paramètres) ainsi que de la probabilité de trouver la réponse correcte de façon tout à fait aléatoire (dans le modèle à trois paramètres, pour des questions à choix multiples).

Contrairement à la méthode classique, la TRI permet de graduer les scores sans normes de population prédéfinies. À l'aide des items communs aux différents niveaux, on estime les normes à partir de la population totale des enfants qui subissent le test dans le cycle en question. Les scores sont attribués de façon à classer chaque enfant dans un niveau et ils sont échelonnés progressivement de manière à refléter la progression des scores sur tous les niveaux. Afin d'assurer la comparabilité des données d'une année à l'autre, les échantillons de chaque cycle doivent représenter des populations équivalentes.

Le modèle à trois paramètres a été choisi pour les tests de mathématiques. Le modèle à trois paramètres tient compte de la difficulté de l'item et de son pouvoir discriminatif et tient compte en outre de la composante des réponses aléatoires. Dans ce sens, la TRI prend en considération les structures de réponse. Ainsi, deux enfants ayant le même score brut n'obtiendront pas le même score gradué, sauf s'ils ont donné exactement les mêmes réponses. Par exemple, un enfant qui n'a répondu correctement qu'aux cinq questions les plus faciles aura un score gradué inférieur à celui qui a répondu correctement aux cinq questions les plus difficiles seulement.

Ce score TRI se distingue de l'autre score gradué du test de mathématiques, que l'on a décrit, en ce sens qu'il fournit des estimations plus précises du principal trait psychologique latent (qui, dans l'analyse factorielle, constitue le premier facteur devant être estimé par la pondération linéaire des éléments qui explique la plus grande partie possible de la variance). Contrairement à l'autre score gradué, ce score n'est pas établi en fonction des résultats prévus d'une population externe, mais plutôt en fonction des résultats prévus de la population à l'étude, tels qu'ils sont estimés pour tous les sujets soumis au test. Les analyses de rangs antérieures effectuées à l'aide des deux méthodes de notation n'ont pas révélé d'écart important entre les deux mesures.

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 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 scolaires. 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 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 au cycle 4

Score brut	Score gradué classique						
	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

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 nouvelle évaluation est en fait une combinaison des mesures cognitives utilisées dans les cycles précédents pour ce groupe d'âge. Dans le but de continuer à mesurer le développement chez les enfants, on a estimé qu'une mesure plus complète des capacités était nécessaire pour voir comment les enfants 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 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) comprend 20 questions. En vérité, il s'agit de la combinaison des livrets 30 et 31 utilisés dans le cycle précédent. Pour le cycle 6, il a été décidé de regrouper les deux livrets et de faire subir la même évaluation à tous les répondants. De plus, deux questions provenaient des niveaux 7 et 9 des tests de mathématiques.

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 vingt 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 un « 1 », un « 2 » ou un « 9 » aux questions complexes. Un score de « 1 » est attribué lorsqu'il y a une réponse, un score de « 2 » est attribué lorsque la question est omise et un score de « 9 » correspond à une valeur manquante. Lorsqu'une réponse a été proposée puis rayée, un score de « 2 » est attribué (question omise) par opposition à « 9 » (valeur manquante). Cela englobe les réponses effacées ou rayées, sauf s'il est clair que la réponse effacée ou rayée est exacte, auquel cas un score de « 1 » est attribué. Si le répondant a écrit quelque chose d'illisible pour le correcteur (du gribouillage, par exemple), un score de « 2 » est attribué.

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.

Pour obtenir le score des exercices de résolution de problèmes (FMAYdS01), on a utilisé le modèle à trois paramètres issu de la théorie de la réponse d'item. Dans le cas des items à réponses libres, on a fixé à 0 la valeur du paramètre de réponse aléatoire (permettant de modéliser les réponses données au hasard aux questions à choix multiples). On a effectué une estimation du maximum de vraisemblance (EMV) des trois paramètres d'item (discrimination, difficulté et hasard) suivie d'une estimation *prévue a priori* (EPP) du score, à l'aide d'un processus itératif, jusqu'à ce que les scores convergent suffisamment. Ces calculs ont été effectués par le logiciel statistique SAS.

Enfin, on a établi une borne inférieure pour le score, ce qui a eu pour effet d'augmenter les scores des 35 derniers répondants au classement. L'élimination ou la diminution de cette borne réduit la corrélation du score et de tous les scores des tests de mathématiques des trois premiers cycles de l'ELNEJ. Les 35 scores indûment faibles semblent être le produit de réponses données au hasard ou du peu d'importance accordée au test.

Traitement des items avec réponse

Avant d'entreprendre le processus itératif d'estimation, on a calculé les scores bruts de chaque item. Les réponses erronées ont reçu un score de 0 et les réponses correctes, un score de 1, ce qui est la pratique normale dans le cadre de la TRI. Le score des réponses partiellement correctes correspond aux points obtenus divisés par le nombre maximum de points (p. ex. 1/2 pour un demi-point).

Traitement des items sans réponse

Les items sans réponse sont soit des items « omis », soit des items « non atteints ». Les items omis sont ceux que le répondant a probablement lus mais auxquels il n'a pas répondu. Dans le calcul des exercices de résolution de problèmes, un item sans réponse placé avant le dernier item pour lequel le répondant a fourni une réponse est considéré comme un item omis. Le premier et le deuxième item sans réponse suivant le dernier item avec réponse étaient également considérés comme items omis. Dans ce cas, le répondant a probablement lu les questions, a décidé qu'elles étaient trop difficiles et a interrompu le test.

Généralement, les répondants omettent des items parce qu'ils sont incapables de fournir la réponse correcte. Par conséquent, les items sans réponse ont été associés à un score de 0. Pour les items à choix multiples sans réponse, on a attribué un point divisé par le nombre de choix.

Les items non atteints sont ceux que le répondant n'a probablement pas lus. Ces items ne fournissent aucun renseignement sur les capacités du répondant. Dans le calcul des scores, tous les items avec réponse de même que l'item suivant le dernier item avec réponse sont considérés

comme des items atteints. Les items non atteints n'ont pas été pris en considération dans l'estimation des capacités du répondant.

Nota : Les exercices de résolution de problèmes ne sont pas minutés. Le répondant peut prendre aussi longtemps que nécessaire pour compléter le test.

Scores

(FMAYdS01)

Les scores apparaissant dans le fichier de données ont une moyenne de 0 et un écart-type de 1; il s'agit là de l'échelle normale dans le cadre de la TRI.

(FMAYfS02)

Ce score est souvent appelé le « score brut » et représente le nombre d'items répondu correctement.

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 (EIIACA) 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 : la compréhension de textes suivis et la compréhension de textes schématiques.

- ◆ Compréhension de textes suivis : 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 : 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 avec le 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.

Trois notes sont possibles à chaque question. La note « 1 » est attribuée lorsqu'il y a une réponse, la note « 7 » est attribuée lorsque la question est omise et la note « 0 » est attribuée lorsqu'il y a une valeur manquante. Ces notes sont inscrites sur la feuille des scores à la fin de chaque livret. Une fois l'évaluation terminée, les données des feuilles des scores sont saisies et transmises à l'équipe de l'ELNEJ à des fins d'analyse.

Le score attribué à la littératie (FLIYfS01) correspond tout simplement au nombre de réponses correctes données aux 36 questions.

Même si les questions du test d'évaluation de la littératie de l'ELNEJ ont été choisies à même celles de l'EIACA de 2003, le score de l'évaluation de la littératie de l'ELNEJ et celui de l'évaluation de la compréhension de textes schématiques 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.

Nota : 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 d'ordre mathématique. Certains chercheurs ont mis au jour des liens entre la littératie et la numératie, et ces deux concepts sont ensemble 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 quantité dans les 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.

Trois notes sont possibles à chaque question. La note « 1 » est attribuée lorsqu'il y a une réponse, la note « 7 » est attribuée lorsque la question est omise et la note « 0 » est attribuée lorsqu'il y a une valeur manquante. Ces notes sont inscrites sur la feuille des scores à la fin de chaque livret. Une fois l'évaluation terminée, les données des feuilles des scores sont saisies et transmises à l'équipe de l'ELNEJ à des fins d'analyse.

Le score attribué à la numératie (FNUYfS01) correspond tout simplement au nombre de réponses correctes données aux 32 questions.

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.

Nota : 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 Enquête sur les enfants du Nord, cycle 6

15.1 Introduction

L'Enquête sur les enfants du Nord (EEN) a été menée par Statistique Canada entre février et avril 2005, en tant qu'enquête complémentaire à l'Enquête longitudinale nationale sur les enfants et les jeunes (ELNEJ). L'enquête a été réalisée avec la collaboration et le soutien de Ressources humaines et Développement social Canada.

L'EEN est une enquête transversale (du type recensement) auprès des enfants du Yukon et du Nunavut qui sont nés en 1999 et qui étaient inscrits à la maternelle en septembre 2004. L'enquête a été menée en même temps que le cycle 6 de l'ELNEJ, avec les mêmes intervieweurs et un questionnaire similaire.

Le plan d'élaboration de l'EEN était fondé sur l'hypothèse que les ministères de l'Éducation du Yukon, des Territoires du Nord-Ouest et du Nunavut fourniraient à Statistique Canada les listes d'inscription des enfants dans les écoles ainsi que le numéro de téléphone du parent ou du tuteur. Malheureusement, les Territoires du Nord-Ouest n'ont pas pu fournir de base de sondage avant le début de la collecte.

15.2 Contexte

Cycles 1, 2 et 3 de l'ELNEJ, composante du Nord

La collecte de renseignements sur les enfants canadiens vivant dans les territoires a commencé avec le cycle 1 de l'ELNEJ en 1994 et 1995. Les données du cycle 2 ont été recueillies en 1996 et 1997, et la collecte du cycle 3 a eu lieu entre l'automne 1998 et le printemps 1999.

Comme les responsables de l'ELNEJ et de l'Enquête nationale sur la santé de la population (ENSP) désiraient recueillir des données sur les résidents du Yukon et des Territoires du Nord-Ouest, les deux enquêtes ont été réunies et raccourcies aux cycles 1, 2 et 3 afin que l'interview soit d'une longueur raisonnable. Les interviews ont été effectuées par des intervieweurs qui avaient été engagés et formés par les bureaux de la statistique du Yukon et des Territoires du Nord-Ouest. Les ménages ont été interviewés par téléphone ou sur place, selon l'endroit où ils vivaient, et les intervieweurs ont utilisé un questionnaire sur support papier.

Pour tous les cycles de ces enquêtes dans le Nord, un parent a fourni l'information pour l'ELNEJ. Un certain nombre de sections figurant sur les questionnaires principaux de l'ELNEJ, notamment sur le tempérament, le fonctionnement de la famille et les soins aux enfants, ont été exclues de l'enquête dans le Nord, et d'autres questions plus pertinentes ont été ajoutées. On voulait notamment savoir si le répondant était un membre des Premières nations et, en cas d'adoption, s'il s'agissait d'une adoption selon les coutumes autochtones.

Cycle 4 de l'ELNEJ, composante du Nord

Pour la première fois en 2001-2002, des données ont été recueillies dans les territoires par les intervieweurs de Statistique Canada, en même temps que celles de l'ELNEJ dans les provinces. L'échantillon de l'EEN était fondé sur les enfants répondants du cycle précédent.

La collecte des données s'est faite à partir des bureaux régionaux, et les intervieweurs ont pu utiliser une application d'interview téléphonique assistée par ordinateur (ITAO), ce qui signifie que les données pouvaient être traitées en même temps et plus rapidement avec celles de l'ELNEJ.

Sommaire du cycle 4

Le dépistage des répondants dans les territoires peut être très difficile en raison de la nature nomade typique de la population. Dans le cadre d'une enquête longitudinale, cette mobilité donne lieu à de sérieux problèmes de couverture, étant donné que les enfants qui ne vivent plus dans les territoires ne sont plus considérés comme faisant partie de l'échantillon. On a perdu bien des répondants au cours des quatre cycles en raison de ce type d'attrition, et il est devenu évident

après la fin du cycle 4 que la qualité des données, en particulier des données longitudinales, constituait un grave problème dans les territoires. Les données du cycle 4 ne pouvaient être diffusées que pour le Yukon et les Territoires du Nord-Ouest. Les faibles taux de réponse au Nunavut interdisent la diffusion des données concernant ce territoire.

Sommaire du cycle 5

Les difficultés liées à la collecte des données au Yukon, dans les Territoires du Nord-Ouest et au Nunavut, les faibles taux de réponse, l'absence de connaissance du plan d'échantillonnage et les problèmes rencontrés à l'étape du traitement des données, tout cela a contribué à compromettre l'exactitude des estimations longitudinales et transversales.

Afin de maintenir l'élan des cycles continus de l'ELNEJ ainsi que de créer une base de données de qualité sur les enfants, la collecte des données dans le Nord a fait l'objet d'un examen approfondi dans le contexte duquel un certain nombre de solutions de rechange ont été envisagées. Il a été établi qu'une nouvelle base de sondage contenant des renseignements à jour et fiables s'imposait.

À l'été 2002, des pourparlers ont commencé avec des représentants de chaque territoire afin d'obtenir les coordonnées d'une personne-ressource concernant les enfants de 5 ans inscrits dans les conseils scolaires. L'expérience menée au Yukon au cycle 5 (EEN) a montré que les dossiers du ministère de l'Éducation constituaient une excellente base de sondage. Les autorités des Territoires du Nord-Ouest et du Nunavut ont malheureusement été dans l'impossibilité de fournir ces renseignements avant le début de la période de collecte. Aucune donnée n'est donc disponible pour ces territoires.

Une base de sondage à jour et tout aussi fiable n'a pas pu être trouvée à ce moment pour les autres groupes d'âge dans les territoires. Mais nous espérons que cette question sera réglée d'ici à la prochaine Enquête sur les enfants autochtones.

Cycle 6

Les difficultés concernant la collecte des données dans le Nord se poursuivent, particulièrement au Nunavut. Même si l'enquête était au début menée par téléphone, le taux de réponse était si faible que la technique des interviews personnelles a été adoptée. Les intervieweurs se sont donc rendus dans les collectivités en avion pour dépister les répondants et mener des interviews personnelles.

Les taux de réponse par téléphone étaient faibles parce que de nombreux répondants dans ces régions n'ont pas le téléphone. Le numéro de téléphone fourni au conseil scolaire était celui du centre communautaire ou d'un autre membre de la famille. Toutefois, les interviews personnelles ont permis d'atteindre un bon taux de réponse, mais le coût de la collecte des données a également augmenté.

Prochain cycle

En 2006, l'EEN sera menée en même temps que l'Enquête auprès des enfants autochtones. Le fait de regrouper ces deux enquêtes nous aidera à mieux comprendre le développement des jeunes enfants vivant dans le Nord et fournira de précieux renseignements sur lesquels s'appuieront les responsables de l'élaboration de politiques. Cela ira également dans le sens de notre engagement actuel en vertu de l'entente sur le développement de la petite enfance qui consiste à faire état d'un ensemble commun d'indicateurs sur le bien-être des jeunes enfants.

15.3 Établissement de l'Enquête auprès des enfants du Nord

Objectifs

On s'entend de plus en plus pour dire que les cinq premières années de la vie sont décisives dans le développement de l'enfant. Ces premières années modèlent les résultats à long terme en ce qui concerne le succès scolaire, la réussite professionnelle, la santé, la qualité de vie et la capacité de s'adapter.

C'est pourquoi cette enquête sur les enfants porte sur une vaste gamme de sujets incluant notamment la santé, le développement physique, l'apprentissage, le comportement et le milieu social (famille, amis, école et collectivité). De plus, l'enquête contient une question sur l'identité autochtone qui offre la possibilité de comparer la façon dont on a répondu aux questions pour les enfants autochtones et les enfants non autochtones.

Contenu

En général, l'application de l'EEN (cycle 6) était similaire à celle de l'ELNEJ dans les provinces à quelques exceptions près.

- 1) Seules des questions s'appliquant aux enfants de 5 ans étaient posées à la PMR ou à son conjoint.
- 2) Comme la collecte des données était effectuée par téléphone à l'aide d'une application d'ITAO (à l'exception de 191 répondants au Nunavut et 47 répondants au Yukon pour lesquels des interviews personnelles ont été menées), les mesures directes utilisées dans le cadre de l'ELNEJ n'ont pas été intégrées à l'interview dans les territoires. Par conséquent, les tests suivants de l'ELNEJ ne faisaient pas partie de l'EEN :
 - l'EVIP;
 - le test sur la connaissance des nombres;
 - le test « Qui suis-je? »
- 3) Enfin, afin de réduire le fardeau de réponse et d'éliminer les questions non pertinentes, les modules suivants ont été supprimés de l'application :
 - sécurité du quartier;
 - travail après la naissance;
 - tempérament;
 - sommeil;
 - garde des enfants;
 - garde légale des enfants;
 - antécédents relatifs à la famille et à la garde des enfants.

Pour plus de renseignements sur les concepts, les définitions, les méthodes de collecte et les étapes de traitement des données de l'EEN (qui sont identiques à ceux de l'ELNEJ) veuillez vous reporter aux chapitres appropriés du présent guide.

Méthodologie de l'enquête

La population cible de l'EEN (cycle 6) était composée d'enfants qui sont nés en 1999 et qui étaient inscrits à la maternelle en septembre 2004. Comme il s'agissait d'une enquête du type recensement, il n'était pas nécessaire de faire un plan d'échantillonnage ni de procéder à la sélection ou à la répartition d'un échantillon.

Le ministère de l'Éducation du Yukon a fourni une liste de 326 enfants. De ce nombre, six étaient considérés comme hors du champ de l'enquête et n'ont pas été inclus dans le processus de pondération. Sur les 320 enfants restant dans le fichier de l'échantillon, 286 répondants (la PMR ou son conjoint) ont répondu aux questions de l'EEN (cycle 6).

Le ministère de l'Éducation du Nunavut a fourni pour sa part une liste de 683 enfants. De ce nombre, 38 étaient considérés comme hors du champ de l'enquête. Sur les 645 enfants restant dans l'échantillon, 485 répondants (la PMR ou son conjoint) ont répondu aux questions de l'EEN (cycle 6).

Pondération

L'EEN (cycle 6) est une enquête du type recensement et non une enquête probabiliste, ce qui signifie que la pondération est assez simple. De plus, il ne fallait produire que des poids transversaux, étant donné que cette enquête n'est plus longitudinale.

On a calculé des poids transversaux pour tous les enfants qui ont répondu à l'EEN au cycle 6. Encore une fois, comme il s'agissait d'une enquête du type recensement, le poids transversal initial attribué à chaque enfant était de 1.

Ajustement pour la non-réponse

Afin de calculer le poids transversal, on a fait un simple ajustement pour la non-réponse, c'est-à-dire qu'on a ajusté les poids des répondants afin de tenir compte des enfants sélectionnés qui n'ont pas répondu.

Les seuls renseignements disponibles sur les enfants non répondants sont ceux qui se trouvent sur les listes d'inscription des conseils scolaires. Il est donc difficile de faire des inférences au sujet de ces enfants. La seule variable dérivée qui s'est révélée importante est le fait que l'enfant vivait à Whitehorse ou ailleurs, ce qui permet de créer deux groupes différents. Le facteur d'ajustement pour la non-réponse de chaque groupe est la somme des répondants et des non-répondants divisée par le nombre de répondants, c'est-à-dire qu'il est plus grand que ou égal à 1.

Stratification a posteriori

La population cible de cette enquête est formée des enfants de 5 ans (enfants nés en 1999 et qui étaient inscrits à la maternelle en septembre 2004). Cette population cible est différente d'une population cible qui comprendrait tous les enfants de 5 ans du Yukon et du Nunavut (y compris ceux qui ne sont pas inscrits à l'école). C'est pourquoi les poids transversaux ne sont pas stratifiés a posteriori pour qu'ils soient conformes aux totaux connus de la population, étant donné que ces totaux incluraient tous les enfants de 5 ans du Yukon et du Nunavut. Cette enquête est du type recensement, ce qui signifie que l'enregistrement de chaque enfant est compté une fois et a un poids égal dans les résultats de l'enquête.

Pour plus de renseignements au sujet du processus de pondération, veuillez consulter le chapitre 12.0.

Qualité des données

Taux de réponse

Dans le cadre de l'EEN (cycle 6), un répondant est défini comme un enfant pour lequel au moins une des composantes suivantes est remplie : composante de l'adulte ou composante de l'enfant.

Parmi les enfants qui faisaient partie de l'échantillon de l'EEN (cycle 6), l'échantillon du Yukon a produit un taux de réponse de 89,4 %, tandis que l'échantillon du Nunavut a permis d'obtenir un taux de réponse de 75,1 %.

Tableau 1 – Taux de réponse transversaux

Territoire	Nombre d'enfants	Nombre de répondants	Taux de réponse transversal
Yukon	320	286	89,4 %
Nunavut	645	485	75,1 %
Total	965	771	79,8 %

Taux de réponse partielle

Selon la définition donnée ci-dessus, une seule composante devait être remplie pour qu'un ménage soit considéré comme répondant. Presque tous les répondants ont répondu aux deux composantes.

Tableau 2 – Taux de réponse aux composantes

Territoire	Nombre de répondants	Les deux composantes remplies	Composante de l'enfant seulement	Composante de l'adulte seulement
Yukon	286	276	10	0
Nunavut	485	443	42	0
Total	771	719	52	0

Erreurs relatives à l'enquête

Les estimations tirées de cette enquête sont fondées sur un recensement des enfants et, par conséquent, il n'y a pas d'erreur d'échantillonnage.

En revanche, des erreurs non liées à l'échantillonnage peuvent se produire à chaque étape d'une enquête ou presque. Il arrive que des intervieweurs comprennent mal les instructions, que des répondants répondent aux questions de façon erronée, que des réponses soient mal inscrites sur le questionnaire et que des erreurs se produisent au moment du traitement et de la totalisation des données. Ce sont là des exemples d'erreurs non dues à l'échantillonnage.

Il y a bien des raisons pour lesquelles certains ménages n'ont pas répondu à l'enquête. Dans certains cas, les intervieweurs ont été incapables de retracer le ménage ou d'entrer en contact avec un ménage sélectionné pendant la période de collecte. Dans d'autres cas, le ménage a refusé de participer à l'enquête.

Lignes directrices pour la totalisation, l'analyse et la diffusion de données

Les estimations établies d'après les données de la présente enquête doivent respecter les lignes directrices relatives à l'analyse et à la diffusion telles qu'énoncées au chapitre 17.0.

16.0 Enjeux analytiques

Le but de ce chapitre est de donner 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). Quelques points soulevés dans ce chapitre l'ont déjà été avec plus de détails dans d'autres chapitres de ce guide. L'objectif de ce chapitre 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;
- 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.

Pour une description détaillée de l'échantillon de l'ELNEJ, voir le chapitre 5.0 (Méthodologie de l'enquête - Échantillon). Pour une description de la méthode de calcul des poids de l'enquête, voir le chapitre 11.0 (Pondération et traitement de la non-réponse). Pour une description de la méthode d'estimation de l'erreur d'échantillonnage, voir le chapitre 13.0 (Calcul de la variance). 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, etc.), voir le chapitre 12.0 (Qualité des données, taux de réponse et couverture).

16.1 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 de certaines caractéristiques 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 structure 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, etc. (Un nombre disproportionné d'enfants proviennent des petites provinces dans l'échantillon.) 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 structure en grappes de l'échantillon entraîne une non-indépendance des unités : les enfants d'un même ménage ne sont pas indépendants.

Supposons que l'analyste désire connaître la répartition des enfants de la cohorte initiale au Canada (par 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 dix 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 autres que les poids entonnoir (pour les enfants qui ont répondu aux cycles 1 et 6, 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.

Tableau – Répartition des enfants dans la population, pour la cohorte initiale, estimations pondérées et non pondérées en utilisant les poids entonnoir du cycle 6 (fwtcwd1l)

Province	Nombre d'enfants (non pondéré)	Pourcentage d'enfants (non pondéré)	Nombre d'enfants (pondéré)	Pourcentage d'enfants (pondéré)
Terre-Neuve-et-Labrador	689	6,00	88 986	1,91
Île-du-Prince-Édouard	339	2,95	23 148	0,50
Nouvelle-Écosse	839	7,31	144 088	3,09
Nouveau-Brunswick	710	6,18	115 131	2,47
Québec	2 108	18,36	1 090 582	23,41
Ontario	2 834	24,68	1 773 616	38,08
Manitoba	905	7,88	182 869	3,93
Saskatchewan	958	8,34	173 611	3,73
Alberta	1 109	9,66	489 913	10,52
Colombie-Britannique	992	8,64	576 125	12,37
Total	11 483	100,00	4 658 069	100,00

S'il n'utilise pas les poids, l'analyste en viendrait à la conclusion inexacte que 22,44 % 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). Sans les poids, l'analyste conclurait également de façon erronée qu'il n'y a que 11 483 enfants dans la population. En réalité, il y en a plus de 4,6 millions.

16.2 L'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 il n'est pas possible de faire d'estimations au niveau du ménage. Toutes les estimations doivent être faites au niveau de l'enfant. Par exemple, nous pouvons estimer le nombre d'enfants vivant dans un ménage avec un seul parent mais nous **ne pouvons pas** estimer le nombre de ménages avec un seul parent.

16.3 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 toujours la population au moment de la sélection initiale de l'enfant. En ce qui concerne les enfants de la cohorte initiale (âgés de 10 à 21 ans au cycle 6), deux ensembles de poids longitudinaux sont disponibles, c'est-à-dire les variables FWTCW01L et FWTCWd1L. Le premier ensemble de poids comprend les enfants longitudinaux qui ont répondu (ou leur PMR a répondu) au cycles 1 et 6, sans égard à leur réponse aux cycles intermédiaires. Le deuxième ensemble de poids longitudinaux, appelé poids « entonnoir », calculé pour les enfants qui ont répondu (ou leur PMR a répondu) à tous les cycles (de 1 à 6).

Par exemple, si vous faites une analyse sur les enfants introduits au cycle 1 et voulez regarder leurs caractéristiques au cycle 6, sans avoir besoin de leur information aux cycles 2 à 5, 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 après la population initiale longitudinale. Pour le cycle 6, des poids transversaux ne sont calculés que pour les enfants âgés de 0 à 5 ans (pour une description de l'échantillon formé de ces enfants, voir le chapitre 5.0, Méthodologie de l'enquête - Échantillon). Il n'est pas recommandé d'utiliser la cohorte initiale pour faire des inférences concernant la population transversale de tous les jeunes de 10 à 21 ans vivant au Canada lors de la collecte du cycle 6 puisque cet échantillon n'a jamais été complété pour tenir compte des immigrants arrivés après 1994. (Pour plus de précisions sur l'incidence de l'immigration, voir le chapitre 12.0, Qualité des données, taux de réponse et couverture.)

16.4 Estimations pondérées simples (totaux, moyennes et proportions)

La présente section 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 toutes 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.

16.4.1 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 ajustés 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 le poids final ajusté 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 :

$$\hat{Y} = \sum_{i \in S_r} w_i y_i$$

16.4.2 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 et 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.

$$\begin{aligned}\hat{Y} &= \frac{\sum_{i \in S_r} w_i y_i}{\sum_{i \in S_r} w_i} \\ &= \frac{\hat{Y}}{\hat{N}}\end{aligned}$$

16.4.3 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 et 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 possédant la caractéristique donnée.

16.4.4 Estimation de domaines de la population

Des estimations peuvent être requises pour des domaines (groupe d'âge, sexe, etc.).

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 de domaines 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 de domaines 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 de domaines pour des données qualitatives ou quantitatives est :

$$\begin{aligned} \hat{P}_{\text{domaine}} &= \frac{\sum_{i \in S_r \cap \text{domaine} \cap C} w_i}{\sum_{i \in S_r \cap \text{domaine}} w_i} \\ &= \frac{\hat{N}_{\text{domaine} \cap C}}{\hat{N}_{\text{domaine}}} \end{aligned}$$

16.5 Poids normalisés

Pour garantir que les estimations des caractéristiques de la population observée finie sont plus ou moins non biaisées au regard du plan d'enquête, 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 à base de modèles, 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.) Le nombre d'unités observées est alors à tout le moins correct. Toutefois, comme le logiciel à base de modèles traite encore les unités comme si elles étaient réparties de façon indépendante et identique, les estimations de la variance d'échantillon 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. Les poids normalisés ne doivent être utilisés que s'il est impossible de réaliser correctement une analyse statistique à l'aide d'un logiciel à base de modèles, 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 s'établissent à 1, 3, 4 et 6. Les poids normalisés sont présentés dans le tableau ci-dessous.

Tableau – 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).

Tableau – 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 la non-réponse)	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		
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 par 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).

Tableau – Poids normalisés par 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
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
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 structure 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 (la taille réelle du test est plus importante que prévu).

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 : elle est parfois trop rigoureuse et parfois pas suffisamment.

L'exemple ci-dessous illustre comment la valeur de p estimée, produite au moyen de poids normalisés, peut être incorrecte. Nous avons 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 sont indépendants.

Conclusion

Avec un logiciel à base de modèles, 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 bien 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 suffisamment prudent.

Le domaine d'analyse et le nombre de répondants doivent être pris en considération dans le calcul des poids normalisés.

Si on utilise un logiciel fondé sur le plan de sondage (comme SUDAAN ou Bootvar), la normalisation n'est pas nécessaire.

16.6 Mesures répétées

Les mesures répétées sont des observations multiples des mêmes variables de la même unité échantillonnée. Les mesures répétées surviennent 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, etc.). 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 façon 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 autre que le poids entonnoir, par exemple) et disons que $bs1_1$, $bs2_1$, etc., correspondent aux poids bootstrap de l'enfant.

Supposons que, pour le premier enfant, nous disposions des données des cycles 1, 2 et 3 et que, pour le deuxième enfant, nous disposions uniquement des données du cycle 1. Dans le cas du troisième enfant, nous avons les données des cycles 1 et 2 et, pour ce qui est du quatrième enfant, nous avons uniquement les données des cycles 2 et 3. Le fichier serait alors construit comme ci-dessous, et la variance d'échantillonnage estimée à l'aide des poids bootstrap serait calculée comme d'habitude.

Enregistrement	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}	$bs1_4$	$bs2_4$

À 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.

16.7 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 des inférences sont faites.
- 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).

À 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.

Il y a souvent 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 ainsi qu'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 BHLCQ01 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 façon 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 de 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.

16.8 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 dépister l'enfant. La non réponse totale est prise en considération et corrigée lors de la pondération. Voir le chapitre 11.0 pour de plus amples détails.

Il y a non-réponse partielle lorsque des questions précises sont laissées sans réponse. À noter qu'un enfant ou un jeune ne doit pas avoir répondu à toutes les questions au cycle 6 afin d'être considéré comme un répondant. Par exemple, 90 % des enfants répondants ont terminé le test de l'EVIP (voir le chapitre 11.0 pour plus d'exemples de non-réponse partielle). Certaines réponses sont imputées, mais la plupart de 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. 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 que les enregistrements pour lesquels des renseignements complets sont disponibles pour l'analyse principale 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 de vivre 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 avertit les gens qui prennent connaissance des résultats que la non-réponse a posé un problème et les aide à interpréter correctement les conclusions analytiques.
2. Déclarer la non-réponse partielle comme une catégorie.
3. 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).
4. Répondérer les enregistrements comportant une réponse afin de tenir compte des répondants affichant une non-réponse partielle.
5. 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.
6. 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.
7. À noter que, pour l'ELNEJ, les poids sont stratifiés a posteriori en des 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).
8. Imputer la non-réponse partielle (remplacer les valeurs manquantes par d'autres valeurs).
 - a. 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.

- b. 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, Analyse de la non-réponse partielle, 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.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 coll., New York : Academic Press, 1983)
- *Nonresponse in Household Surveys* (Groves, R. et Cooper, M., New York : Wiley, 1998)
- *Statistical Analysis with Missing Data* (Little, R.J.A. et Rubin, D.B., New York : Wiley, 1987)

Il est aussi possible d'obtenir plus d'information sur la façon de traiter la non-réponse partielle dans l'ELNEJ à l'occasion d'un atelier offert par Statistique Canada. Les transparents de cet atelier sont disponibles dans les centres de données de recherche. De plus, l'atelier peut être présenté par un méthodologiste de Statistique Canada sur demande.

16.9 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, aux erreurs de réponse, etc. Les erreurs non dues à l'échantillonnage sont décrites au chapitre 12.0 – Qualité des données, taux de réponse et couverture.

16.10 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 des effets de plan de sondage de l'ELNEJ lors du calcul de la variance d'échantillonnage. En effet, ces logiciels ne peuvent pas tenir compte de la complexité de l'ELNEJ (plan de sondage complexe, ajustement des poids pour la non-réponse et post-stratification) pour calculer la variance. Il en résulte habituellement une sous-estimation de celle-ci, sous-estimation qui peut être quelquefois importante. L'utilisateur est donc encouragé à utiliser les poids Bootstrap pour l'estimation de la variance.

Certains logiciels sont capables d'utiliser les poids Bootstrap. SUDAAN, WesVar et STATA9 calculent des estimations de variance basées 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 avec SAS. La dernière accorde la flexibilité 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 et d'homogénéité, et d'ajuster des modèles de régression linéaire, log-linéaire, logistique et de survie. SUDAAN peut faire la lecture de fichiers SAS, SPSS et d'autres formats communs.

Le logiciel WesVar présente une approche « pointer-cliquer » à l'analyse, ce qui le rend très facile à apprendre. Les types d'analyse sont plus limités que dans SUDAAN, mais WesVar offre quand même le calcul des estimations de moyennes, ratios et totaux, des tests d'indépendance et d'homogénéité, et la modélisation linéaire et logistique. La lecture de fichiers SAS, SPSS et d'autres formats communs est possible avec WesVar, mais le logiciel requiert l'étape supplémentaire de 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 par 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.ca/francais/freepub/12-002-XIF/12-002-XIF2004002.pdf>. Voir le chapitre 13.0 pour plus de renseignements sur l'estimation de la variance d'échantillonnage, y compris des précisions sur un autre outil appelé Bootvar.

16.11 Coefficients de variation pour des proportions

Il est d'usage d'utiliser le coefficient de variation (CV) pour déterminer la qualité d'une estimation, 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 devrait faire preuve de grande prudence. Le CV n'est pas dans tous les cas une bonne mesure de la qualité pour des proportions estimées. Plus de détails à ce sujet se trouvent dans le chapitre 13.0 (Calcul de la variance) et le chapitre 17.0 (Lignes directrices concernant la totalisation, l'analyse et la diffusion de données).

16.12 Écart-type et erreur-type

Les expressions « écart-type » et « erreur-type » sèment parfois la confusion. Voir le chapitre 13.0 sur le calcul de la variance pour des définitions claires de ces expressions et des explications sur la façon de les appliquer à l'ELNEJ.

16.13 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. Lorsque vous analysez une population en particulier, assurez-vous que les questions qui vous intéressent s'appliquent à votre population. De plus, lors de 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 que la question n'a pas été répondue, c'est qu'elle ne s'applique pas. À l'occasion, « Enchaînement valide » peut signifier « 0 » ou « pas à l'école ». Consultez le questionnaire pour connaître les détails.

17.0 Lignes directrices concernant la totalisation, l'analyse et la diffusion des données

Ce chapitre du document décrit les lignes directrices que doivent suivre les utilisateurs qui totalisent, analysent, publient ou diffusent des données tirées du fichier 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 chiffres qui n'ont pas été publiés.

17.1 Lignes directrices concernant l'arrondissement

Afin que les estimations destinées à la publication ou à toute autre forme de diffusion qui sont tirées du fichier de microdonnées de l'Enquête longitudinale nationale sur les enfants et les jeunes (ELNEJ) correspondent à celles de Statistique Canada, nous conseillons vivement aux utilisateurs de se conformer aux lignes directrices suivantes en ce qui concerne leur arrondissement.

- a) Les estimations dans le corps d'un tableau statistique doivent être arrondies à la centaine près au moyen de 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 normal 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 à l'aide de 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 à l'aide de 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) à l'aide de la technique d'arrondissement normale.
- e) Si, en raison de contraintes d'ordre technique ou de toute autre nature, une technique d'arrondissement autre que la technique normale est utilisée et que les estimations à publier ou à diffuser sous une forme quelconque diffèrent des estimations correspondantes publiées par Statistique Canada, nous conseillons vivement aux utilisateurs d'expliquer ces divergences dans le ou les documents à publier ou à diffuser.
- f) Les utilisateurs ne doivent en aucun cas publier ou diffuser des estimations non arrondies. Des estimations non arrondies laissent supposer qu'elles sont plus précises qu'elles ne le sont en réalité.

17.2 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 pas seulement elle-même, mais aussi plusieurs autres unités au sein de la population observée.

Dans le cas de l'ELNEJ, le poids attribué à chaque enfant tient compte du nombre d'enfants que représente un répondant donné.

Lorsqu'il produit des estimations simples, y compris de tableaux statistiques ordinaires, l'utilisateur doit appliquer le poids d'échantillonnage approprié. Si 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 visée par l'enquête et ces estimations ne correspondront pas à celles produites par Statistique Canada.

17.3 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. 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 dans la mesure du possible. Pour de plus amples détails au sujet des poids de l'enquête, veuillez vous référer au chapitre 11.0.

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.

17.4 Lignes directrices concernant la diffusion

L'utilisateur des données ne doit diffuser ni publier aucune estimation permettant de déterminer 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, ce qui implique de calculer l'erreur type qui y est associée (voir le chapitre 13.0). L'utilisateur devrait également déterminer l'incidence que des erreurs non dues à l'échantillonnage (voir le chapitre 11.0) peuvent avoir sur l'estimation.

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 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 des 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.

Annexe I

Outils pour chercheurs et analystes de l'Enquête longitudinale nationale sur les enfants et les jeunes

Des analystes et chercheurs qui utilisent les données de l'Enquête longitudinale nationale sur les enfants et les jeunes (ELNEJ) ont exprimé le besoin d'avoir des outils qui leur permettraient de mieux comprendre et utiliser les données de l'enquête. Le but de ce document est de répondre à ces besoins.

Ce document comprend deux sections principales. La première section traite des différents éléments méthodologiques à considérer lors de la soumission de projets de recherche qui utilisent les données de l'ELNEJ. La deuxième section donne des recommandations sur ce qui doit être considéré lorsque les données de l'ELNEJ sont utilisées pour écrire un article. Tel qu'on devrait s'y attendre, beaucoup d'éléments qui se trouvent dans la section sur les propositions 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, si nécessaire.

Avant de soumettre un projet de recherche :

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 de ce qui doit être considéré lorsqu'ils veulent utiliser les données de l'ELNEJ.

Ce document identifie les éléments essentiels qui doivent être considérés lors de la rédaction d'une proposition de recherche. Les auteurs sont fortement encouragés à utiliser cette liste afin de s'assurer qu'ils ont couvert tous les éléments avant de soumettre leur proposition de recherche.

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 ce cas, tous les éléments présentés dans ce document s'appliquent. Dans les cas où les données de l'ELNEJ sont utilisées à d'autres fins que pour faire des inférences (par exemple, des études de cas sans inférence à la population), 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.

Les éléments du cadre analytique

Il y a six éléments principaux à considérer lors de la préparation d'une proposition de recherche. 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 à être utilisée pour l'analyse.
 - ELNEJ
 - Autre (spécifiez)
- b Indiquez quelles autres sources de données, si applicable, seront utilisées.
- 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 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'impact potentiel sur votre analyse de cette différence.
 - Si d'autres sources de données sont utilisées pour fins de comparaison, vous devez inclure un énoncé décrivant les populations cibles des différentes sources de données utilisées.
- b Traitement de la non-réponse :
 - Si des variables que vous prévoyez utiliser présentent de la non-réponse, inclure un énoncé indiquant le niveau de non-réponse de ces variables et l'impact potentiel de cette non-réponse sur votre analyse.
 - Identifiez de quelle façon la non-réponse sera traitée :
 - Imputation
 - Repondération
 - Rapportée comme une valeur
 - Ignorée, analyse faite uniquement avec les répondants.
 - Une analyse des caractéristiques des répondants vs les non-répondants devrait être faite afin d'identifier des 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 voyez des restrictions à ce que vous pouvez faire avec les données de l'ELNEJ.

3) Variables

- Fournissez une liste préliminaire des variables de l'ELNEJ qui seront utilisées dans votre analyse.
- Lors de la construction de modèle mathématique, identifiez les variables, si elles sont connues, qui feront partie du modèle.

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 ou longitudinaux, que vous utiliserez dans votre analyse. Notez que si des estimations transversales et longitudinales sont requises, 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, étant donné les probabilités de sélection très inégales des unités.

5) Estimation de la variance

Différentes méthodes sont disponibles pour calculer la précision des estimations, telles que le calcul de la variance et du coefficient 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 :

- Approximations utilisant les tables d'approximation de variabilité échantillonnale (disponibles pour les 4 premiers cycles).
- Utilisation de l'interface Excel qui contient la variance et le CV pour une série de domaines d'intérêt.
- Utilisation des poids Bootstrap avec le programme Bootvar, SUDAAN ou autres programmes qui peuvent utiliser les poids Bootstrap.
- Variance et coefficient de variation non calculés. Nota : ceci signifie qu'aucune inférence statistique ne peut être faite.
- Utilisation d'un autre logiciel (précisez le logiciel : _____). Nota : très peu de logiciels sont capables de tenir compte du plan de sondage complexe de l'ELNEJ pour calculer la variance.

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 que vous allez comparer sont statistiquement différentes.
- c Prévoyez inclure les intervalles de confiance de vos estimations, basés sur un calcul approprié de la variance.

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 de ce qui doit être considéré lorsqu'ils veulent utiliser les données de l'ELNEJ.

Ce document identifie les éléments essentiels qui doivent être considérés lors de la soumission d'un article analytique pour révision. Les auteurs sont fortement encouragés à utiliser cette liste afin de s'assurer qu'ils ont couvert tous les éléments avant de soumettre leur article.

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 ce cas, tous les éléments présentés dans ce document s'appliquent. Dans les cas où les données de l'ELNEJ sont utilisées à d'autres fins que pour faire des inférences (par exemple, des études de cas sans inférence à la population), 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.

Les éléments du cadre analytique

Il y a six éléments principaux à considérer lors de la préparation d'une proposition de recherche. 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 source de données, si applicable, ont été utilisées.
- 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 les populations cibles des différentes sources de données utilisées.
- b Traitement de la non-réponse:
 - 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 a été traitée:
 - Imputation
 - Repondération
 - Rapportée comme une valeur
 - Ignorée, analyse faite uniquement avec les répondants.
 - Analyse des caractéristiques des non-répondants vs les répondants devrait être faite afin de dissiper tout doute causé par un biais potentiel.
- c Restrictions avec les données de l'ELNEJ:
 - 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 des restrictions à ce que vous avez fait dans votre projet avec les données de l'ELNEJ (p. ex. : les variables utilisées).

3) Variables

- Décrivez les variables 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 vous avez produit des estimations transversales et longitudinales, vous devez utiliser les poids appropriés dans chaque cas.
- 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, étant donné les probabilités de sélection très inégales des unités.

5) Estimation de la variance

Décrivez la méthode que vous avez utilisée pour faire l'estimation de la variance, ainsi que la variance et/ou le coefficient de variation pour chaque estimation. Les options possibles pour calculer la variance et/ou le CV sont :

- Approximations utilisant les tables d'approximation de variabilité échantillonnale (disponibles pour les 4 premiers cycles)

- Utilisation de l'interface Excel qui contient la variance et le 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 : _____). Nota : très peu de logiciels sont capables de tenir compte du plan de sondage complexe de l'ELNEJ pour calculer la variance.

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 que vous avez comparées 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, si la méthodologie que vous avez utilisée est correcte et bien expliquée dans votre article.

Annexe II

Analyse de la non-réponse partielle

L'Analyse de la non-réponse partielle qui suit, laquelle s'est fondée sur les résultats du cycle 5 de l'**Enquête longitudinale nationale sur les enfants et les jeunes**, 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 prendre en considération 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 répondant. 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 dans le cycle 1 en 1994-1995 et de 10 à 13 ans dans le cycle 5 en 2002-2003, pour qui 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. 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 dans l'ensemble, 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 autres réponses ont été données par les enfants eux-mêmes, à l'aide de questionnaires autoadministrés. Les taux de réponse étaient plus bas que pour les informations données par les parents, se situant entre 74 % et 78 % pour les variables prédictives (pratiques parentales répressives, éducation parentale, 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 cette analyse 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 réponses 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 en ce qui a trait à la variable des compétences parentales (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. Les répondants de la région des Prairies ont eu un taux de réponse 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

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 compétences parentales - 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 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 la région 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

Annexe III

Concordance entre les noms de variables du traitement et les noms de variables de la diffusion, pour les questionnaires auto administrés du cycle 6 de l'ELNEJ

Variable	Longueur	Type	Livret20	Livret21	Livret22	Livret23
PERSRUK	14	C	PERSRUK	PERSRUK	PERSRUK	PERSRUK
FMMCQ01	3	N	EFFAGEYR	EFFAGEYR	EFFAGEYR	EFFAGEYR
FMMCbQ1A	3	N	AGE	AGE	AGE	AGE
FMMCQ02	1	C	GENDER	GENDER	GENDER	GENDER
FFFCQ01	1	N	A1	A1	A1	A1
FFFCQ02	1	N	A2	A2	A2	A2
FFFCQ03	1	N	A3	A3	A3	A3
FFFCQ04	1	N	A4	A4	A4	A4
FFFCcQ4A	1	N			A5	A5
FFFCQ05	2	N	A5	A5	A6	A6
FFFCQ06	2	N	A6_2			
FFFCbQ13	2	N		A6_2	A7_2	A7_2
FFFCbQ14	2	N		A7_2	A8_2	A8_2
FFFCbQ15	1	N		A8	A9	A9
FFFCQ07	1	N	A8	A10	A13	A13
FFFCQ08A	1	N	A9A	A11A	A14A	A14A
FFFCQ08B	1	N	A9B	A11B	A14B	A14B
FFFCQ08C	1	N	A9C	A11C	A14C	A14C
FFFCQ08D	1	N	A9D	A11D	A14D	A14D
FFFCQ08E	1	N	A9E	A11E	A14E	A14E
FFFCQ08F	1	N	A9F	A11F	A14F	A14F
FFFCQ08G	1	N	A9G	A11G	A14G	A14G
FFFCQ08H	1	N	A9H	A11H	A14H	A14H
FFFCc08I	1	N	A9I	A11I	A14I	A14I
FFFCQ08J	1	N	A9J	A11J		
FFFCQ08K	1	N	A9K	A11K	A14J	A14J
FFFCQ08L	1	N	A9L	A11L	A14K	A14K
FFFCQ08M	1	N	A9M	A11M	A14L	A14L
FFFCQ08N	1	N	A9N	A11N	A14M	A14M
FFFCQ09	1	N	A10	A12		
FFFCQ12	2	N	A11	A13		
FFFCd12A	1	N			A15	A15
FDRCdQ05	1	N	A7A	A9A	A10A	A10A
FDRCdQ09	1	N	A7B	A9B	A10B	A10B
FFFCd16C	1	N		A9C	A10C	A10C
FFFCc16D	1	N	A7C	A9D	A10D	A10D
FFFCc16E	1	N	A7D	A9E	A10E	A10E
FFFCc19A	1	N			A11A	A11A_1
FFFCc19B	1	N			A11B	A11B_2
FFFCc19C	1	N			A11C	A11C_3
FFFCc19D	1	N			A11D	A11D_4
FFFCc20A	1	N			A12A	A12A
FFFCc20B	1	N			A12B	A12B
FFFCc20C	1	N			A12C	A12C

FSCCQ01	1	N	B1	B1	B1	
FSCCbQ18	1	N		B2	B2	
FSCCb19A	1	N		B3A	B3A	
FSCCb19B	1	N		B3B	B3B	
FSCCb19C	1	N		B3C	B3C	
FSCCb19D	1	N		B3D	B3D	
FSCCc19F	1	N		B3E	B3E	
FSCCb19E	1	N		B3F	B3F	
FSCCb20A	1	N		B4A	B4A	
FSCCb20B	1	N		B4B	B4B	
FSCCb20C	1	N		B4C	B4C	
FSCCb20E	1	N		B4D	B4D	
FSCCc20H	1	N		B4E	B4E	
FSCCc20I	1	N		B4F	B4F	
FSCCc20J	1	N		B4G	B4G	
FSCCb20G	1	N		B4H	B4H	
FSCCQ02	1	N	B2	B5	B5	
FSCCcQ3a	1	N	B3A	B6A	B6A	
FSCCcQ03	1	N	B3B	B6B	B6B	
FSCCcQ3b	1	N	B3C	B6C	B6C	
FSCCcQ3D	1	N	B3D	B6D	B6D	
FSCCcQ3C	1	N		B6E	B6E	
FSCCcQ3E	1	N		B6F	B6F	
FSCCcQ3F	1	N		B6G	B6G	
FSCCd3G	1	N			B6H	
FSCCc21A	1	N	B4A	B7A	B7A	
FSCCc21B	1	N		B7D	B7D	
FSCCc21C	1	N	B4B	B7B	B7B	
FSCCc21D	1	N	B4C	B7C	B7C	
FSCCc21E	1	N		B7E	B7E	
FSCCc21F	1	N		B7F	B7F	
FSCCcQ26	1	N			B8	
FSCCcQ27	1	N			B9	
FSCCcQ10	1	N	B5F	B8	B10	B7
FSCCbQ22	1	N		B9A	B12A	
FSCCc22A	1	N		B9B	B12B	
FSCCcQ28	1	N			B13	
FSCCcQ29	1	N			B14	
FSCCQ12	1	N	B6A	B10A	B15A	
FSCCQ11	2	N	B6B	B10B	B15B	
FSCCcQ16	2	N	B6C	B10C	B15C	
FSCCQ17	2	N	B6D	B10D	B15D	
FSCCcQ30	2	N		B11	B16	
FSCCQ13	2	N	B7A	B12A	B17A	
FSCCQ14	1	N	B7B	B12B	B17B	
FSCCQ15	1	N	B7C	B12C	B17C	
FSCCcQ24	2	N		B13	B18	
FSCCc31A	1	N			B11A	
FSCCc31B	1	N			B11B	
FSCCc31C	1	N			B11C	
FSCCc31D	1	N			B11D	
FSCCc31E	1	N			B11E	

FSCCQ05	1	N	B5A			
FSCCQ06	1	N	B5B			
FSCCQ07	1	N	B5C			
FSCCQ08	1	N	B5D			
FSCCQ09	1	N	B5E			
FAMCQ01A	1	N	C1A	C1A	C1A	B1A
FAMCQ01B	1	N	C1B	C1B	C1B	B1B
FAMCQ01C	1	N	C1C	C1C	C1C	B1C
FAMCQ01D	1	N	C1D	C1D	C1D	B1D
FAMCQ01E	1	N	C1E	C1E	C1E	B1E
FAMCe25A	1	N	C2A	C2A	C2A	B2A
FAMCe25B	1	N	C2B	C2B	C2B	B2B
FAMCe25C	1	N	C2C	C2C	C2C	B2C
FAMCe25D	1	N	C2D	C2D	C2D	B2D
FAMCe25E	1	N	C2E	C2E	C2E	B2E
FAMCe25F	1	N	C2F	C2F	C2F	B2F
FAMCe25G	1	N	C2G	C2G	C2G	B2G
FAMCe25H	1	N	C2H	C2H	C2H	B2H
FAMCe25I	1	N	C2I	C2I	C2I	B2I
FAMCe25J	1	N	C2J	C2J	C2J	B2J
FAMCe25K	1	N	C2K	C2K	C2K	B2K
FAMCe25L	1	N	C2L	C2L	C2L	B2L
FAMCe25M	1	N	C2M	C2M	C2M	B2M
FAMCe25N	1	N	C2N	C2N	C2N	B2N
FAMCe25O	1	N	C2O	C2O	C2O	B2O
FAMCbQ02	1	N		C3	C3	B3
FAMCcQ03	1	N		C4	C4	B4
FAMCcQ6A	1	N		C5A	C7A	
FAMCcQ6B	1	N		C5B	C7B	
FAMCdQ6C	1	N				B8A
FAMCcQ7A	1	N		C6A	C8A	
FAMCcQ7B	1	N		C6B	C8B	
FAMCdQ7C	1	N				B8B
FAMCcQ8A	1	N		C7A	C9A	
FAMCcQ8B	1	N		C7B	C9B	
FAMCdQ8C	1	N				B8C
FAMCdQ4A	1	N			C5A	
FAMCcQ4B	1	N			C5B	
FAMCcQ4C	1	N			C5C	
FAMCcQ4D	1	N			C5D	
FAMCeQ4E	1	N			C5E	
FAMCe26A	1	N				B5A
FAMCe26B	1	N				B5B
FAMCe26C	1	N				B5C
FAMCe26D	1	N				B5D
FAMCe26E	1	N				B5E
FAMCe26F	1	N				B5F
FAMCcQ5A	1	N			C6A	B6A
FAMCcQ5B	1	N			C6B	B6B
FAMCcQ5C	1	N			C6C	B6C
FAMCcQ5D	1	N			C6D	B6D
FAMCdQ09	1	N				B9

FAMCdQ10	1	N				B10
FFBCQ01A	1	N	D1A	D1A	D1A	
FFBCc01B	1	N	D1B	D1B	D1B	
FFBCQ01C	1	N	D1C	D1C	D1C	
FFBCQ01D	1	N	D1D	D1D	D1D	
FFBCQ01E	1	N	D1E	D1E	D1E	
FFBCQ01F	1	N	D1F	D1F	D1F	
FFBCQ01G	1	N	D1G	D1G	D1G	
FFBCQ01H	1	N	D1H	D1H	D1H	
FFBCQ01I	1	N	D1I	D1I	D1I	
FFBCQ01J	1	N	D1J	D1J	D1J	
FFBCQ01K	1	N	D1K	D1K	D1K	
FFBCQ01L	1	N	D1L	D1L	D1L	
FFBCQ01M	1	N	D1M	D1M	D1M	
FFBCQ01P	1	N	D1N	D1N	D1N	
FFBCQ01Q	1	N	D1O	D1O	D1O	
FFBCQ01R	1	N	D1P	D1P	D1P	
FFBCQ01S	1	N	D1Q	D1Q	D1Q	
FFBCQ01T	1	N	D1R	D1R	D1R	
FFBCQ01U	1	N	D1S	D1S	D1S	
FFBCc01V	1	N	D1T	D1T	D1T	
FFBCQ01W	1	N	D1U	D1U	D1U	
FFBCd01X	1	N	D1V	D1V	D1V	
FFBCQ01Z	1	N	D1W	D1W	D1W	
FFBCQ1AA	1	N	D1X	D1X	D1X	
FFBCQ1BB	1	N	D1Y	D1Y	D1Y	
FFBCQ1CC	1	N	D1Z	D1Z	D1Z	
FFBCQ1DD	1	N	D1AA	D1AA	D1AA	
FFBCQ1FF	1	N	D1BB	D1BB	D1BB	
FFBCQ1GG	1	N	D1CC	D1CC	D1CC	
FFBCQ1HH	1	N	D1EE	D1EE	D1EE	
FFBCQ1JJ	1	N	D1DD	D1DD	D1DD	
FFBCQ1LL	1	N	D1FF	D1FF	D1FF	
FFBCQ1MM	1	N	D1GG	D1GG	D1GG	
FFBCQ1NN	1	N	D1HH	D1HH	D1HH	
FFBCQ1OO	1	N	D1II	D1II	D1II	
FFBCQ1PP	1	N	D1JJ	D1JJ	D1JJ	
FFBCQ1QQ	1	N	D1KK	D1KK	D1KK	
FFBCQ1RR	1	N	D1LL	D1LL	D1LL	
FFBCQ1SS	1	N	D1MM	D1MM	D1MM	
FFBCQ1TT	1	N	D1NN	D1NN	D1NN	
FFBCc1UU	1	N	D1OO	D1OO	D1OO	
FFBCd10A	1	N				C1A
FFBCd10B	1	N				C1B
FFBCd10C	1	N				C1C
FFBCd10D	1	N				C1D
FFBCd10E	1	N				C1E
FFBCd10F	1	N				C1F
FFBCd10G	1	N				C1G
FFBCd10H	1	N				C1H
FFBCd10I	1	N				C1I
FFBCd10J	1	N				C1J

FFBCd10K	1	N				C1K
FFBCd10L	1	N				C1L
FFBCcQ04	1	N		D2	D2	C2
FFBCcQ4A	1	N		D3	D3	C3
FFBCcQ05	1	N		D4	D4	C4
FFBCcQ07	1	N		D5	D5	C5
FFBCcQ08	1	N		D6	D6	C6
FFBCbQ2A	1	N	D2A			
FFBCbQ2B	1	N	D2B	D7A	D7A	C7A
FFBCbQ2C	1	N	D2C			
FFBCbQ2D	1	N	D2D			
FFBCbQ2E	1	N	D2E	D7B	D7B	C7B
FFBCbQ2F	1	N	D2F	D7C	D7C	C7C
FFBCcQ2H	1	N				C7D
FFBCeQ2O	1	N		D7D	D7D	C7E
FFBCbQ2P	1	N		D7E	D7E	C7F
FFBCdQ2Z	1	N				C7G
FFBCe2ZZ	1	N		D7F	D7F	C7H
FFBCbQ2Y	1	N		D7G	D7G	C7I
FFBCe2BB	1	N		D7H	D7H	C7J
FFBCcQ3A	1	N		D8	D8	C8
FFBCQ03	1	N	D3			
FATCbQ1A	1	N	H1A	E1A		
FATCc1AA	1	N			E1A	
FATCbQ1B	1	N	H1B	E1B		
FATCc1BB	1	N			E1B	
FATCbQ1C	1	N	H1C	E1C		
FATCc1CC	1	N			E1C	
FATCbQ1D	1	N	H1D	E1D		
FATCc1DD	1	N			E1D	
FATCbQ1E	1	N	H1E	E1E		
FATCc1EE	1	N			E1E	
FATCbQ1I	1	N	H1F	E1F		
FATCc1II	1	N			E1F	
FATCbQ1F	1	N	H1G			
FATCeQ20	2	N	H2	E2	E2	
FATCbQ04	1	N		E3	E3	
FATCdQ07	2	N	H3	E7	E7	
FATCQ03	2	N	H4	E4		
FATCeQ12	1	N	H5A	E8A	E8A	
FATCe12B	1	N	H5B	E8B	E8B	
FATCe12C	1	N	H5C	E8C	E8C	
FATCeQ13	1	N	H6A	E9A	E9A	
FATCe13B	1	N	H6B	E9B	E9B	
FATCe13C	1	N	H6C	E9C	E9C	
FATCeQ21	2	N	H7	E10	E10	
FATCeQ22	1	N	H8	E11	E11	
FATCdQ10	2	N		E12	E12	
FATCbQ11	2	N		E13		
FATCdQ5A	1	N		E5A	E5A	
FATCbQ5B	1	N		E5B	E5B	
FATCbQ5C	1	N		E5C	E5C	

FATCbQ5D	1	N		E5D	E5D	
FATCbQ5E	1	N		E5E	E5E	
FATCbQ5F	1	N		E5F	E5F	
FATCbQ5G	1	N		E5G	E5G	
FATCdQ05	1	N		E6	E6	
FATCc14A	1	N			E4A	
FATCc14B	1	N			E4B	
FATCc14C	1	N			E4C	
FATCc14D	1	N			E4D	
FDRCdQ01	2	N	G1	F1	F1	D1
FDRCQ04	2	N		F2	F2	D2
FDRCQ03	2	N	G2_2	F3_2	F3_2	
FDRCcQ06	1	N	G3			
FDRCdQ6A	2	N		F4	F4	D3
FDRCdQ07	2	N	G4	F5	F5	
FDRCdQ9A	1	N		F6	F6	
FDRCbQ9B	2	N		F7	F7	
FDRCdQ9C	2	N		F8	F8	D4
FDRCdQ15	2	N		F9	F9	D5
FDRCd14A	2	N		F11A_3	F11A_3	
FDRCc14C	2	N		F11B_3	F11B_3	
FDRCc14B	2	N		F11C_3	F11C_3	
FDRCd14F	2	N		F11D_3	F11D_3	
FDRCc14D	2	N		F11E_3	F11E_3	
FDRCQ10	1	N	G5			
FDRCQ12	2	N	G6			
FDRCc18A	2	N		F10A	F10A	D6A
FDRCd18B	2	N		F10B	F10B	D6B
FDRCd18C	2	N		F10C	F10C	D6C
FDRCc18D	2	N		F10D	F10D	D6D
FDRCdQ19	1	N				D7
FDRCdQ20	1	N				D8
FPMCdQ04	2	N	E1	G1	G1	G1
FPMCCQ5A	1	N	E2A	G2A	G2A	G2A
FPMCCQ5B	1	N	E2B	G2B	G2B	G2B
FPMCCQ5C	1	N	E2C	G2C	G2C	G2C
FPMCCQ06	1	N	E3	G3	G3	G3
FPMCdQ6A	2	N				G4A
FPMCdQ6B	2	N				G4B
FPMCdQ6C	1	N				G5A
FPMCdQ6D	1	N				G5B
FPMCdQ6E	1	N				G5C
FPMCdQ6F	1	N				G5D
FPMCdQ6G	1	N				G5E
FPMCdQ6H	1	N				G5F
FPMCdQ6I	1	N				G5G
FPMCdQ6J	1	N				G5H
FPMCdQ6K	1	N				G5I
FPMCdQ6L	1	N				G5J
FPMCdQ07	2	N	E4	G4	G4	G6
FPMCCQ8A	1	N	E5A	G5A	G5A	G7A
FPMCCQ8B	1	N	E5B	G5B	G5B	G7B

FPMCcQ8C	1	N	E5C	G5C	G5C	G7C
FPMCcQ09	1	N	E6	G6	G6	G8
FPMCdQ9A	2	N				G9A
FPMCdQ9B	2	N				G9B
FPMCdQ9C	1	N				G10A
FPMCdQ9D	1	N				G10B
FPMCdQ9E	1	N				G10C
FPMCdQ9F	1	N				G10D
FPMCdQ9G	1	N				G10E
FPMCdQ9H	1	N				G10F
FPMCdQ9I	1	N				G10G
FPMCdQ9J	1	N				G10H
FPMCdQ9K	1	N				G10I
FPMCdQ9L	1	N				G10J
FPMCcQ10	1	N	E7	G7	G7	G12
FPMCd11A	2	N	E8	G8	G8	G13
FPMCc11B	2	N	E9	G9	G9	G14
FPMCcQ1A	1	N	E10A	G10A	G10A	
FPMCcQ1B	1	N	E10B	G10B	G10B	
FPMCcQ1C	1	N	E10C	G10C	G10C	
FPMCcQ1D	1	N	E10D	G10D	G10D	
FPMCcQ1E	1	N	E10E	G10E	G10E	
FPMCcQ1F	1	N	E10F	G10F	G10F	G11A
FPMCcQ1G	1	N	E10G	G10G	G10G	
FPMCcQ1H	1	N	E10H	G10H	G10H	
FPMCcQ1I	1	N	E10I	G10I	G10I	
FPMCcQ1J	1	N	E10J	G10J	G10J	
FPMCcQ1R	1	N	E10K	G10K	G10K	
FPMCcQ1K	1	N	E10L	G10L	G10L	
FPMCcQ1L	1	N	E10M	G10M	G10M	
FPMCcQ1M	1	N	E10N	G10N	G10N	
FPMCcQ1N	1	N	E10O	G10O	G10O	
FPMCcQ1O	1	N	E10P	G10P	G10P	
FPMCcQ1P	1	N	E10Q	G10Q	G10Q	
FPMCcQ1Q	1	N	E10R	G10R	G10R	
FPMCdQ1S	1	N	E10S	G10S	G10S	
FPMCdQ1T	1	N	E10T	G10T	G10T	G11B
FPMCdQ9U	1	N				G11C
FPMCdQ9V	1	N				G11D
FPMCbQ3A	1	N	E11A	G11A		
FPMCbQ3C	1	N	E11B	G11B		
FPMCbQ3E	1	N	E11C	G11C		
FPMCd12A	1	N				G15
FPMCd12B	1	N				G16
FPMCd12C	1	N				G17A
FPMCd12D	1	N				G17B
FPMCd12E	1	N				G17C
FPMCd12F	1	N				G17D
FPMCd12G	1	N				G17E
FPMCd12H	1	N				G17F
FPMCd12I	1	N				G17G
FPMCd12J	1	N				G17H

FPMCd12K	1	N					G17I
FPMCd12L	1	N					G17J
FHTCcQ03	1	N		H1	H1		
FHTCbQ3A	1	N		H4A	H4A		
FHTCbQ3B	1	N		H4B	H4B		
FHTCbQ3C	1	N		H4C	H4C		
FHTCbQ3G	1	N		H4D	H4D		
FHTCbQ04	1	N		H5			
FHTCbQ05	1	N		H6			
FHTCdQ06	1	N		H7	H5		
FHTCeQ5A	1	N		H8	H6		E8
FHTCeQ5B	1	N					E8_1A
FHTCeQ5C	1	N					E8_1B
FHTCeQ5D	1	N					E8_1C
FHTCeQ5E	1	N					E8_1D
FHTCdQ5F	1	N					E8_1E
FHTCdQ5G	1	N					E8_2A
FHTCdQ5H	1	N					E8_2B
FHTCdQ5I	1	N					E8_2C
FHTCdQ5J	1	N					E8_2D
FHTCeQ5K	1	N					E8_3A
FHTCeQ5L	1	N					E8_3B
FHTCeQ5M	1	N					E8_3C
FHTCeQ5N	1	N					E8_3D
FHTCeQ5O	1	N					E8_3E
FPUCQ01	1	N	F1	H9	H7		E3
FPUCQ02	1	N	F2	H10	H8		E4
FPUCQ03	1	N	F3	H11			
FPUCdQ04	1	N	F4	H12	H10		E6
FPUCQ05	1	N	F5	H13	H11		E7
FDACcQ1A	2	N		H14_2	H12_2		F1_2
FDACcQ1B	1	N		H15	H13		F2
FDACcQ1C	1	N			H14		
FDACcQ02	2	N		H16	H15		F3
FDACcQ2A	1	N			H16		F4
FDACcQ05	1	N			H17		F5
FDACdQ23	1	N					F8_1
FDACeQ26	1	N					F9
FDACeQ25	2	N					F11
FDACcQ06	2	N			H18		F6
FDACcQ07	2	N			H19_1		F7
FDACcQ08	1	N			H20		
FDACcQ09	1	N			H21		
FDACd22A	1	N					F10A
FDACd22B	1	N					F10B
FDACd22C	1	N					F10C
FDACd22D	1	N					F10D
FDACd22E	1	N					F10E
FDACd22F	1	N					F10F
FDACd22G	1	N					F10G
FDACd22H	1	N					F10H
FDACbQ4C	1	N		H17A			

FDACbQ4D	1	N		H17B		
FDACbQ4E	1	N		H17C		
FDACbQ4F	1	N		H17D		
FWKCbQ02	1	N		I1		
FWKcCQ4A	1	N			I1A	
FWKcCQ4B	1	N			I1B	
FWKcCQ4C	1	N			I1C	
FWKcCQ4D	1	N			I1D	
FWKCe06A	1	N			I2A	
FWKCe06B	1	N			I2B	
FWKCe06C	1	N			I2C	
FWKCe06D	1	N			I2D	
FWKCe06E	1	N			I2E	
FWKCe06F	1	N			I2F	
FWKcCQ9A	2	N			I3A_3	
FWKcCQ9B	2	N			I3B_3	
FWKCeQ10	1	N		I2		
FWKCeQ11	1	N			I4	
FWKcd12A	1	N			I5A	
FWKcd12B	1	N			I5B	
FWKcd12C	1	N			I5C	
FWKcd12D	1	N			I5D	
FWKcd13A	1	N			I6A	
FWKcd13B	1	N			I6B	
FWKcd13C	1	N			I6C	
FWKcd13D	1	N			I6D	
FWKcd13E	1	N			I6E	
FWKcd13F	1	N			I6F	
FWKcd13G	1	N			I6G	
FWKcd13H	1	N			I6H	
FWKcd13I	1	N			I6I	
FPUCcQ3A	3	N			H9	E5
FHTCbQ01	4.2	N		H2	H2	E1
FHTCbQ02	7.3	N		H3	H3	E2
FEQYeS06	2	N	RA	RA	RA	RA
FEQYeS07	2	N	ER	ER	ER	ER
FEQYeS08	2	N	SM	SM	SM	SM
FEQYeS09	2	N	AD	AD	AD	AD
FEQYeS10	2	N	GM	GM	GM	GM
FEQYeS04	2	N	EQ4	EQ4	EQ4	EQ4
FEQYeS05	2	N		EQ5	EQ5	EQ5
FbMleS01	7.3	N		selfbmi	selfbmi	selfbmi
FbMleS02	2	N		selfbmi2	selfbmi2	selfbmi2
FbMleS03	1	N		selfbmi3	selfbmi3	selfbmi3
FFFcS01	2	N	frfs01	frfs01	frfs01	frfs01
FAMcS02	2	N	abms01	abms01	abms01	abms01
FFBcS01	2	N	behs01	behs01	behs01	
FFBCdS02	2	N	behs02	behs02	behs02	
FFBcS03	2	N	behs03	behs03	behs03	
FFBCdS04	2	N	behs04	behs04	behs04	
FFBcS05	2	N	behs05	behs05	behs05	
FFBcS07	2	N	behs07	behs07	behs07	

FPMcS1	2	N	pars01	pars01	pars01	
FPMcbS2B	2	N	pars02	pars02	pars02	
FPMcS3	2	N	pars03	pars03	pars03	
FPMcdS4	2	N				pars04
FPMcdS5	2	N				pars05
FHTCbS1b	2	N				deps01
FWTCw01L	12.4	N				
Fwtc wd1l	12.4	N				

Annexe IV

Concordance entre les noms de variables du cycle 6 et les noms de variables du Livret 24 (18 à 19 ans, les questionnaires auto administrés) du cycle 5

Module	C6 Nom de la variable	C6 Nom à la collecte	C5 Nom de la variable	C5 Livret 24 auto administré
ABM	FAMYfQ01	ABMQ1	EAMCQ01A	A1-A
ABM	FAMYfQ02	ABMQ2	EAMCQ01B	A1-B
ABM	FAMYfQ03	ABMQ3	EAMCQ01C	A1-C
ABM	FAMYfQ04	ABMQ4	EAMCQ01D	A1-D
ABM	FAMYfQ05	ABMQ5	EAMCQ01E	A1-E
ABM	FAMYfQ6A	ABMQ6A	EAMCe26A	A10-A
ABM	FAMYfQ6B	ABMQ6B	EAMCe26B	A10-B
ABM	FAMYfQ6C	ABMQ6C	EAMCe26D	A10-C
ABM	FAMYfQ6E	ABMQ6E	EAMCe26E	A10-D
ABM	FAMYfQ6G	ABMQ6G	EAMCe26F	A10-E
RLY	FRLYfQ01	rlyq1	EDACcQ2A	C4
RLY	FRLYfQ02	rlyq2	EDACcQ1B	C2
RLY	FRLYfQ03	rlyq3	EDACcQ1C	C2
RLY	FRLYfQ07	rlyq7	EDACcQ05	C5
RLY	FRLYfQ08	rlyq8	EDACcQ06	C6
RLY	FRLYfQ09	rlyq9	EDACcQ07	C7
RLY	FRLYfQ10	rlyq10	EDACdQ23	C8
RLY	FRLYfQ11	rlyq11	EAMCeQ21	A6
RLY	FRLYfQ12	rlyq12	EAMCeQ22	A7
RLY	FRLYfQ13	rlyq13	EAMCeQ23	A8
RLY	FRLYfQ14	rlyq14	EAMCeQ24	A9
RLY	FRLYf15A	rlyq15a	EDACd22A	C9
RLY	FRLYf15B	rlyq15b	EDACd22B	C9
RLY	FRLYf15C	rlyq15c	EDACd22C	C9
RLY	FRLYf15F	rlyq15f	EDACd22D	C9
RLY	FRLYf15G	rlyq15g	EDACd22E	C9
RLY	FRLYf15H	rlyq15h	EDACd22F	C9
RLY	FRLYf15I	rlyq15i	EDACd22H	C9
RLY	FRLYfQ16	rlyq16	EDACeQ26	C10
RLY	FRLYf17A	rlyq17a	EDACeQ25	C11
RLY	FRLYf17B	rlyq17b	EDACeQ25	C11
RLY	FRLYf17C	rlyq17c	EDACeQ25	C11
RLY	FRLYf17D	rlyq17d	EDACeQ25	C11
RLY	FRLYf17E	rlyq17e	EDACeQ25	C11
RLY	FRLYf17F	rlyq17f	EDACeQ25	C11
RLY	FRLYf17G	rlyq17g	EDACeQ25	C11
RLY	FRLYf17H	rlyq17h	EDACeQ25	C11
RLY	FRLYf17I	rlyq17i	EDACeQ25	C11
RLY	FRLYf17J	rlyq17j	EDACeQ25	C11
HLTY	FHTYfQ06	hltyq6	EDRCdQ01	E1
HLTY	FHTYfQ07	hltyq7	EDRCQ04	E2
HLTY	FHTYfQ08	hltyq8	EDRCdQ6A	E3
HLTY	FHTYfQ10	hltyq10	EDRCdQ15	E5
HLTY	FHTYfQ11	hltyq11	EDRCdQ15	E5
HLTY	FHTYfQ12	hltyq12	EDRCc18A	E6
HLTY	FHTYfQ13	hltyq13	EDRCc18A	E6

HLTY	FHTYf14A	hltyq14a	EFBCd10A	B1-A
HLTY	FHTYf14B	hltyq14b	EFBCd10B	B1-B
HLTY	FHTYf14C	hltyq14c	EFBCd10C	B1-C
HLTY	FHTYf14D	hltyq14d	EFBCd10D	B1-D
HLTY	FHTYf14E	hltyq14e	EFBCd10E	B1-E
HLTY	FHTYf14F	hltyq14f	EFBCd10F	B1-F
HLTY	FHTYf14G	hltyq14g	EFBCd10G	B1-G
HLTY	FHTYf14H	hltyq14h	EFBCd10H	B1-H
HLTY	FHTYf14J	hltyq14j	EFBCd10I	B1-I
HLTY	FHTYf14K	hltyq14k	EFBCd10J	B1-J
HLTY	FHTYf14M	hltyq14m	EFBCd10K	B1-K
HLTY	FHTYf14N	hltyq14n	EFBCd10L	B1-L
FBH	FFBYfQ01	fbhq1	EFBCcQ4A	B2-5
FBH	FFBYfQ02	fbhq2	EFBCcQ4A	B2
FBH	FFBYfQ03	fbhq3	EFBCcQ05	B3
FBH	FFBYfQ04	fbhq4	EFBCcQ07	B4
FBH	FFBYfQ05	fbhq5	EFBCcQ08	B5
FBH	FFBYfQ06	fbhq6	EFBCcQ2H	B6-B
FBH	FFBYfQ07	fbhq7	EFBCdQ2Z	B6-E
FBH	FFBYfQ08	fbhq8	EDRCdQ19	E7
FBH	FFBYfQ09	fbhq9	EDRCdQ20	E8
FBH	FFBYfQ10	fbhq10	EFBCcQ3A	B7
SUPY	FSPYfQ14	supyq4	EPMCcQ06	D3
SUPY	FSPYfQ15	supyq5	EPMCcQ09	D8