

PONDÉRATION DES DONNÉES DU VOLET 2004

Nathalie Plante, Catherine Fontaine et Robert Courtemanche
Direction de la méthodologie, de la démographie et des enquêtes spéciales
Institut de la statistique du Québec
Septembre 2005

Le présent rapport a pour but de décrire la méthode de pondération utilisée pour les données de l'Étude longitudinale du développement des enfants du Québec (ÉLDEQ) au volet 2004. La section 1 présente une description des étapes ayant mené au choix de la stratégie de pondération pour chacun des instruments d'enquête. Les taux de réponse pondérés obtenus, par instrument, sont détaillés à la section 2. À la section 3 se trouve l'analyse de la non-réponse ayant donné lieu à chacune des pondérations. Finalement, la section 4 renseigne l'utilisateur sur le fichier de pondération ainsi que sur la façon d'utiliser les poids échantillonnaux dans les analyses statistiques. Cette section comporte également quelques mises en garde sur l'utilisation de ces poids.

1 Stratégie de pondération :

1.1 Questionnaire informatisé rempli par l'intervieweure (QIRI)

Tout comme au volet 2003, l'érosion de l'échantillon au volet 2004 a été importante. Parmi les 2 120 enfants répondants du volet initial (1998), on compte 1 492 répondants au volet 2004 comparativement à 1 944 répondants au volet 2002 et à 1 759 au volet 2003 (tableaux I et II). En 2004, on peut penser que ce phénomène est en partie attribuable au fait que les parents, engagés pour une première phase (de 1998 à 2002), ont considéré avoir respecté leur engagement initial; il peut aussi être attribuable à l'allongement progressif de l'entrevue en face à face (de 45 minutes en 2002 à environ deux heures en 2004).

Comme pour les volets précédents, la presque totalité des répondants au volet 2004 étaient également répondants aux autres volets (1 434 enfants sur 1 492 répondants, soit 96 %). Pour cette raison, une seule pondération a été créée pour le QIRI, tant pour les analyses transversales que longitudinales¹.

¹ Notons que la non-réponse au volet 2004 a été examinée par rapport aux volets 2003 et 2002. Ainsi, en se basant sur l'année 2003, on a l'avantage d'utiliser des données actualisées pour expliquer la non-réponse survenue en 2004. Par ailleurs, en se basant sur les données de l'année 2002, on évite une modélisation additionnelle en 2003, bien que ce choix présuppose que les mécanismes de réponse en 2003 et 2004 soient semblables.

Tableau I - Nombre de répondants au QIRI, volets 1998 à 2004

volet 1998	volet 1999	volet 2000	volet 2001	volet 2002	volet 2003	volet 2004	Nombre de répondants
Oui	1 434						
Oui	Non	Oui	Oui	Oui	Oui	Oui	5
Oui	Oui	Non	Oui	Oui	Oui	Oui	4
Oui	Oui	Oui	Non	Oui	Oui	Oui	9
Oui	Oui	Oui	Oui	Non	Oui	Oui	5
Oui	Oui	Oui	Non	Non	Oui	Oui	3
Oui	Oui	Oui	Oui	Oui	Non	Oui	28
Oui	Non	Oui	Oui	Oui	Non	Oui	1
Oui	Oui	Non	Oui	Oui	Non	Oui	1
Oui	Oui	Oui	Non	Non	Non	Oui	1
Oui	Oui	Oui	Non	Oui	Non	Oui	1
Oui	Oui	Oui	Oui	Oui	Oui	Non	278
Oui	Oui	Oui	Oui	Oui	Non	Non	154
Oui	Oui	Oui	Oui	Non	Non	Non	18
Oui	Oui	Oui	Non	Non	Non	Non	31
Oui	Oui	Non	Non	Non	Non	Non	43
Oui	Non	Non	Non	Non	Non	Non	63
Oui	Non	Oui	Non	Non	Non	Non	2
Oui	Non	Oui	Non	Non	Oui	Non	1
Oui	Non	Oui	Oui	Oui	Non	Non	1
Oui	Non	Oui	Oui	Oui	Oui	Non	2
Oui	Oui	Non	Oui	Non	Non	Non	1
Oui	Oui	Non	Oui	Non	Oui	Non	1
Oui	Oui	Non	Oui	Oui	Non	Non	5
Oui	Oui	Non	Oui	Oui	Oui	Non	5
Oui	Oui	Oui	Non	Oui	Non	Non	11
Oui	Oui	Oui	Non	Oui	Oui	Non	5
Oui	Oui	Oui	Oui	Non	Oui	Non	7

Tableau II- Nombre de répondants au QIRI, volets 1998 à 2004

	volet 1998	volet 1999	volet 2000	volet 2001	volet 2002	volet 2003	volet 2004
Nombre de répondants pour un volet donné	2 120	2 045	1 997	1 950	1 944	1 759	1 492
Nombre de répondants longitudinaux (pour un volet donné et ses précédents)	2 120	2 045	1 985	1 924	1 894	1 712	1 434

Le choix de la stratégie de pondération s'appuie sur deux modélisations sommaires de la non-réponse au QIRI au volet 2004: une par rapport aux données du volet 2002 et une autre par rapport aux données du volet 2003. Les variables considérées sont, pour la plupart, de nature socioéconomique pour l'année de référence (2002 ou 2003). Des variables dynamiques (c'est-à-dire qui varient d'une année à l'autre) sous forme longitudinale ont aussi été considérées. La méthode de pondération sera décrite plus en détail à la section 3.

L'analyse de la non-réponse selon les deux stratégies envisagées a fait ressortir des résultats intéressants et logiques dans les deux cas. Le choix de la stratégie de pondération s'est porté sur le volet 2002 pour diverses raisons. Tout d'abord, l'année 2002 comme année de référence permet de s'appuyer sur la dernière année de la première phase de l'ÉLDEQ. On sait qu'en 2003, il y a eu des changements qui peuvent expliquer, en partie, l'érosion importante de l'échantillon de répondants. De plus, un autre avantage est d'éviter la sur-modélisation, dans un contexte où l'on fait des ajustements de la non-réponse additionnels pour les instruments autres que le QIRI. La modélisation de la non-réponse comporte par conséquent quatre étapes:

- Ajustement de l'inverse des probabilités de sélection pour la non-réponse à l'enquête au volet 1998 → pondération QIRI du volet 1998
- Ajustement des poids QIRI du volet 1998 pour la non-réponse à l'enquête au volet 2000 parmi les répondants du volet 1998 toujours admissibles à l'enquête au volet 2000 → pondération QIRI transversale du volet 2000
- Ajustement des poids transversaux du volet 2000 pour la non-réponse à l'enquête au volet 2002 parmi les répondants du volet 2000 toujours admissibles à l'enquête au volet 2002 → pondération QIRI transversale du volet 2002
- Ajustement des poids transversaux du volet 2002 pour la non-réponse à l'enquête au volet 2004 parmi les répondants du volet 2002 toujours admissibles à l'enquête au volet 2004 → pondération QIRI transversale du volet 2004.

Afin d'obtenir des poids pour l'ensemble des 1 492 enfants répondants au QIRI du volet 2004, les enfants qui étaient répondants au volet 2004 mais non-répondants au volet 2002 ont été considérés répondants au volet 2002 et se sont vu attribuer un poids QIRI pour le volet 2002², ce dernier constituant le poids de base de la dernière étape d'ajustement selon la stratégie de pondération décrite précédemment.

La pondération ainsi créée peut être utilisée tant pour les analyses transversales des données du volet 2004 que pour les analyses longitudinales des données des volets 1998 à 2004.

1.2 Les autres instruments de collecte

Le tableau III présente le nombre de répondants obtenus pour chacun des autres instruments de collecte, le nombre de répondants obtenus parmi les répondants au QIRI et le nombre de répondants obtenus parmi les enfants ayant participé aux tests cognitifs. La hiérarchie observée quant aux répondants à un instrument par rapport aux

² La méthode utilisée pour attribuer un poids QIRI à ces enfants aux volets antérieurs sera décrite à la section 3.

autres détermine la stratégie de pondération à privilégier pour les instruments autres que le QIRI. Le schéma 1 illustre cette hiérarchie.

Tableau III - Nombre de répondants par instrument au volet 2004

	Nb répondants	Nb répondants parmi rép. QIRI	Nb répondants parmi rép. tests cognitifs
QIRI	1 492	1 492	n/a
QAAM³	1 319	1 319	n/a
QAAP⁴	1 002	1 002	n/a
Tests cognitifs	1 209	1 172	1 209
QPAE⁵	1 179	1 144	1 179
QAAENS⁶	968	942	962
Test de motricité	1 066	1 034	1 066

1.2.1 Instruments autoadministrés des parents (QAAM, QAAP)

Les répondantes au QAAM constituent un sous-ensemble des mères/conjointes des familles répondantes au QIRI. On demandait en effet aux mères biologiques ou conjointes du père de remplir le QAAM seulement pour les familles ayant répondu au QIRI. La pondération des données du QAAM consiste donc en un ajustement de la pondération QIRI pour tenir compte de la non-réponse au QAAM, parmi les répondants au QIRI dont la mère biologique ou la conjointe du père est présente au sein de la famille. On obtient ainsi 1 319 enfants avec un poids pour l'analyse des données provenant du QAAM. La situation est similaire pour le QAAP et l'on obtient 1 002 enfants avec un poids pour l'analyse des données issues de cet instrument.

1.2.2 Tests cognitifs et test de motricité

Les enfants répondants aux tests cognitifs font presque tous partie des familles répondantes au QIRI. En effet, parmi les 1 209 enfants ayant participé à au moins un des tests cognitifs, seuls 37 enfants font partie de familles n'ayant pas répondu au QIRI.⁷ La stratégie de pondération retenue consiste à faire un ajustement de la pondération QIRI pour tenir compte de la non-réponse aux tests cognitifs. Les non-répondants sont ici les enfants n'ayant participé à aucune tâche cognitive.

Les enfants dont les parents n'ont pas répondu au QIRI ont tout de même été considérés répondants à cet instrument, pour les besoins de la pondération.⁸ Cela a permis de leur attribuer un poids QIRI, ce dernier étant le poids de référence pour la pondération aux tests cognitifs. Par ailleurs, deux enfants n'ayant pas participé aux tests cognitifs et pour lesquels les parents ont déclaré un diagnostic d'autisme ou d'incapacité mentale ont également été considérés répondants. S'ils sont inclus dans

³ Questionnaire autoadministré de la mère/conjointe

⁴ Questionnaire autoadministré du père/conjoint

⁵ Questionnaire papier administré à l'enfant

⁶ Questionnaire autoadministré de l'enseignant

⁷ La participation de l'enfant aux tests cognitifs est conditionnelle au consentement de ses parents.

⁸ Seuls 35 des 37 enfants dont les familles n'ont pas répondu au QIRI du volet 2004 ont été considérés répondants aux fins de la pondération des données des tests cognitifs puisque les deux enfants restants étaient également non-répondants au QIRI du volet 2002. Ces deux enfants n'ont par conséquent pas de poids ni au QIRI, ni aux tests cognitifs au volet 2004.

les analyses, un score doit leur être attribué⁹. On obtient ainsi 1 209 enfants avec un poids pour l'analyse des données des tests cognitifs.

Parmi les enfants ayant participé aux tests cognitifs, 143 d'entre eux n'ont pas participé au test de motricité. Un ajustement additionnel a par conséquent été fait à partir du poids des tests cognitifs pour tenir compte de cette non-réponse. On obtient 1 069 enfants avec un poids pour l'analyse des données du test de motricité¹⁰.

1.2.3 Questionnaire papier administré à l'enfant (QPAE) et Questionnaire autoadministré de l'enseignant (QAAENS)

Parmi les enfants ayant participé aux tests cognitifs, seulement 30 d'entre eux n'ont pas répondu au questionnaire qui leur était adressé. Aucune pondération spécifique n'a par conséquent été créée pour cet instrument. Le poids des tests cognitifs peut être utilisé pour l'analyse des données du QPAE.

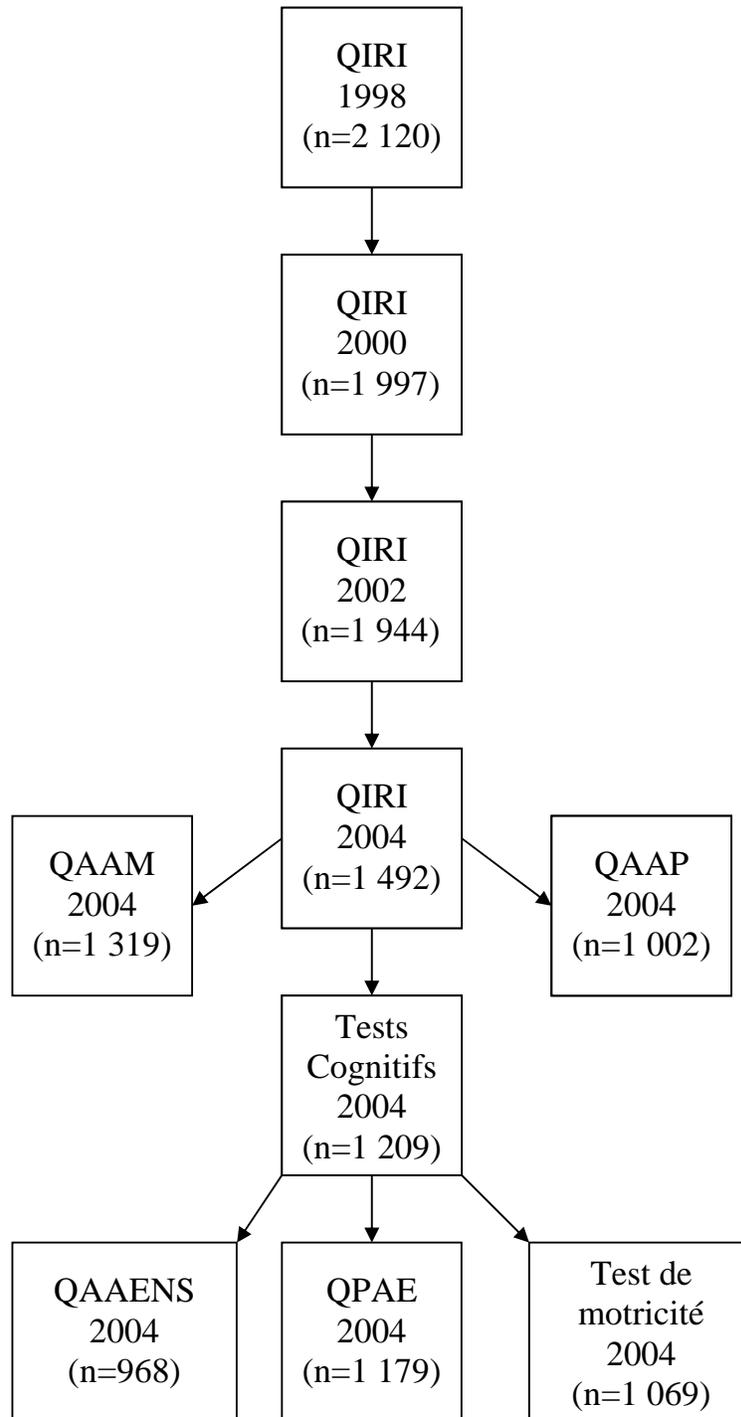
En ce qui concerne le QAAENS, la situation est différente. Parmi ceux ayant participé aux tests cognitifs, on compte en effet 240 enfants pour lesquels le professeur n'a pas rempli le questionnaire. Par ailleurs, on compte six enfants pour lesquels le professeur a rempli le QAAENS sans qu'ils aient participé aux tests cognitifs.¹¹ Le poids de référence aux tests cognitifs a donc été ajusté afin tenir compte de la non-réponse au QAAENS parmi les enfants ayant participé aux tests. Les six enfants n'ayant pas participé aux tests cognitifs ont été considérés répondants à ces tests aux fins de la pondération QAAENS. On obtient finalement 966 enfants avec un poids pour l'analyse des données du QAAENS.

⁹ La question des enfants souffrant de problèmes de développement est soulevée à la section 3.5.3.1.

¹⁰ Comme pour la pondération des données des tests cognitifs, les deux enfants répondants au test de motricité mais dont les parents n'ont pas répondu au QIRI du volet 2004 se sont vu attribuer un poids nul pour le test de motricité. Les cinq enfants souffrant d'autisme ou d'incapacité mentale diagnostiquée, tel que déclaré par les parents, et qui n'ont pas participé au test de motricité, ont été considérés répondants à ce test. Un score doit leur être attribué pour l'analyse, le cas échéant.

¹¹ Le consentement des parents est requis pour administrer aux enfants les tests cognitifs à l'école et demander à l'enseignant de remplir le questionnaire autoadministré. Ainsi, lorsque le parent ne donne pas son accord, on ne peut pas demander à l'enseignant de remplir le questionnaire autoadministré. Dans quelques cas, le consentement des parents a été obtenu mais les tests n'ont pu être administrés.

Schéma 1. Hiérarchie des répondants aux différents instruments de collecte



2 Taux de réponse :

Le tableau IV présente les taux de réponse pondérés obtenus au volet 2004 pour chacun des instruments. Ces taux sont obtenus en multipliant les taux obtenus aux différentes étapes de pondération, selon le cas. La démarche de modélisation ainsi que les résultats obtenus pour chacun des instruments sont présentés à la section 3.

Tableau IV - Taux de réponse pondérés au volet 2004

Instrument	Proportion de répondants parmi les répondants au QIRI	Proportion de répondants parmi les répondants aux tests cognitifs	Taux de réponse pondéré relatif à l'échantillon initial du volet 1998
QIRI	100 %	n/a	54 %
QAAM	87 %	n/a	47 %
QAAP	77 %	n/a	42 %
Tests cognitifs	n/a	100 %	42 %
Test de motricité	n/a	87 %	36 %
QPAE	n/a	97 %	40 %
QAAENS	n/a	80 %	33 %

On note que pour le QAAENS et le test de motricité, le taux de réponse est inférieur à 40 %. Le fait que les enfants répondants à ces instruments soient caractérisés fait en sorte que certains sous-groupes sont sous-représentés de façon très importante au sein de l'échantillon des répondants. Cela signifie que le taux de réponse est encore plus faible dans certains sous-groupes dont les caractéristiques sont liées aux mesures étudiées.¹²

Par ailleurs, au cours des ans, l'érosion semble toujours se produire au sein des sous-groupes des familles plus défavorisées sur le plan économique ou dont la mère est jeune, peu scolarisée ou immigrante. Cela fait en sorte que la proportion de répondants peut devenir très faible dans certains sous-groupes. Il est alors risqué de penser qu'un petit nombre de répondants représente adéquatement les non-répondants. La pondération est dans ce cas un outil qui minimise les risques de biais mais il est peu probable qu'elle permette à elle seule de corriger les biais dus à la non-réponse. Bien que l'ÉLDEQ fournisse bon nombre de variables pour tenter de limiter les biais dus à la non-réponse, les résultats basés sur ces instruments devraient être interprétés avec prudence.

¹² Il est à noter que le QAAENS compte par ailleurs un nombre important de variables pour lesquelles la non-réponse partielle est importante. Les analyses basées sur ces variables doivent donc faire l'objet d'une attention particulière. Pour de plus amples renseignements on peut consulter « Fontaine, C. et Plante, N., 2005. Étude de la non-réponse partielle au volet 2004 » à l'adresse suivante www.jesuiseserai.stat.gouv.qc.ca sous l'onglet « Documentation technique ».

3 Analyse de la non-réponse :

3.1 Démarche générale d'analyse

La création de pondérations ajustées pour la non-réponse aux différents instruments est basée sur la formation de classes de pondération. Comme aux volets précédents, la technique employée est la modélisation par segmentation fondée sur l'algorithme CHAID (« Chi-square automatic interaction detection ») mis au point par Kass (1980). Les classes de pondération sont créées sous forme d'arborescence; elles ne résultent donc pas nécessairement du croisement de toutes les variables considérées pour la modélisation. Pour une famille donnée, l'ajustement de la pondération consiste à diviser le poids de référence par la proportion pondérée de familles répondantes observée au sein de la classe à laquelle elle appartient.

Des tests du khi-deux approximatifs sont effectués au préalable à l'aide du logiciel SAS pour choisir un sous-ensemble de variables les plus liées à la non-réponse. Les variables considérées pour la modélisation avec CHAID sont celles dont le seuil est égal ou inférieur à 0,20.

Certaines des variables retenues par la modélisation ont par ailleurs fait l'objet d'une analyse supplémentaire. Cette analyse consistait à vérifier si celles-ci étaient reliées aux mesures provenant de l'instrument en question¹³. En effet, lors de la création des classes de pondération, il faut s'assurer que les variables choisies soient liées non seulement à la probabilité de répondre, mais aussi aux variables mesurées dans l'enquête. Autrement, la réduction du biais potentiel dû à la non-réponse pourrait être négligeable (Beaumont, 2002); elle ne compenserait pas dans ce cas la perte de précision due à la perturbation des poids¹⁴.

3.2 Pondération transversale des données du QIRI

3.2.1 Admissibilité à l'enquête au volet 2004

Parmi les 2 120 répondants au volet initial, on compte 22 familles ayant quitté définitivement le Québec entre 1998 et 2004 et deux familles dont l'enfant cible est décédé entre les volets 1998 et 2004. Ces familles, considérées comme inadmissibles à l'enquête, ne sont plus visées par l'enquête en ce sens qu'elles ne font plus partie de la population sur laquelle porte l'inférence. La population visée est par conséquent composée des enfants survivants qui sont demeurés au Québec entre les volets 1998 et 2004 ou qui ont quitté la province temporairement.

Les familles n'ayant pu être retracées, ayant refusé de répondre ou ayant été dans l'impossibilité de le faire sont toutes considérées admissibles à l'enquête. Bien que l'on sache que parmi celles n'ayant pu être retracées il pourrait y avoir des familles démenagées définitivement hors du Québec, leur nombre est trop petit pour que l'on en

¹³ Cet examen n'a pas été fait pour la pondération des données du QIRI étant donné le grand nombre de mesures recueillies par le biais de cet instrument.

¹⁴ Puisque les mesures sont manquantes pour les non-répondants, on doit étudier ce lien à partir du sous-ensemble de répondants seulement. Dans ce cas, on fait l'hypothèse que si un lien est détecté pour le sous-ensemble des répondants, il le serait aussi si le même test était effectué à partir de l'échantillon total.

tienne compte dans la pondération. Sur cette base, l'échantillon admissible à l'enquête au volet 2004 est composé de 2 096 familles.

3.2.2 Enfants répondants au volet 2004 mais non-répondants au volet 2002

Neuf enfants dont la famille était répondante au QIRI du volet 2004 ne l'étaient pas au volet 2002. Un nouveau poids QIRI a par conséquent été recalculé pour l'ensemble des répondants du volet 2002, en incluant ces neuf enfants. Pour ce faire, les classes de pondération définies au volet 2002 ont été conservées ; seules les proportions pondérées de répondants ont été recalculées dans chaque classe. Pour les variables servant à créer les classes de pondération, des valeurs ont été imputées pour les non-répondants du volet 2002, aux seules fins de la pondération.

3.2.3 Variables considérées et résultats

Pour tenir compte de la non-réponse au volet 2004, un ajustement a été fait à partir de la pondération transversale modifiée du volet 2002 (section 3.2.2). Cet ajustement est requis puisque les répondants au volet 2004 présentent des caractéristiques différentes des non-répondants. On minimise ainsi les risques de biais dus à la non-réponse dans les estimations qui seront produites. La nouvelle variable de pondération pour le QIRI est appelée PEPCM7; elle est appropriée tant pour les analyses transversales des données du QIRI du volet 2004 que pour les analyses longitudinales impliquant des variables provenant du QIRI des volets 1998 à 2004. Qui plus est, la pondération PEPCM7 peut être utilisée pour analyser les variables du QIRI du volet 2004 en combinaison avec des variables du QIRI d'un ou plusieurs des volets précédents.

Les variables considérées pour la modélisation sont principalement de nature socioéconomique et sont tirées du QIRI du volet 2002 ; des variables dites longitudinales ont également été étudiées en créant un indice à partir de la même mesure prise de 1998 à 2002. Ces variables sont: le revenu du ménage (revenu faible à au moins un des 5 volets, soit moins de 10 000\$, versus autres); le type de famille (monoparentalité ou nouveau conjoint à au moins un volet versus autres); la présence du père biologique (le père biologique est absent à au moins un volet versus autres); le niveau de suffisance du revenu (insuffisance du revenu à au moins un volet versus autres); le travail du père au cours des douze derniers mois (n'a pas travaillé au cours des douze mois précédant l'enquête pour plus d'un volet versus autres); la principale source de revenu (aide sociale comme principale source de revenu à aucun volet; à 1 ou 2 volets, à 3 volets ou plus); la situation en emploi des parents (aucun parent en emploi à aucun volet, à 1 ou 2 volets, à 3 volets ou plus); la santé perçue par la mère (santé passable ou mauvaise à aucun volet, à 1 volet, à 2 volets et plus); fréquentation d'une école par le père au moment de l'enquête (n'a jamais fréquenté d'école; a fréquenté une école à 1 volet; a fréquenté une école à 2 volets ou plus).

Parmi l'ensemble des variables considérées pour la modélisation, voici celles qui ont été retenues:

- le statut d'immigrante de la mère (ESDMD1A)
- le plus haut niveau de scolarité de la mère/conjointe (EEDMD01)

- le niveau de suffisance du revenu¹⁵ (variable longitudinale)

Au sein des différentes classes d'ajustement de la pondération, la proportion de répondants varie de 45 % à 81 %. La proportion la plus faible est observée dans la classe des mères immigrantes ayant un faible niveau de scolarité (diplôme d'études secondaires ou moins).

3.3 Pondération transversale des données du QAAM

3.3.1 Non-réponse au QAAM : état de la situation

Aux cinq premiers volets (de 1998 à 2002), l'intervieweuse demandait à la mère de remplir le QAAM lors de sa visite au domicile de l'enfant cible. La mère effectuait le plus souvent cette tâche pendant que l'intervieweuse administrait les tâches d'évaluation du développement cognitif à l'enfant. Lorsque la mère n'avait pas rempli le questionnaire durant la visite de l'intervieweuse, elle était invitée à acheminer son questionnaire complété par la poste. Au volet 2003, l'entrevue auprès de la personne qui connaît le mieux l'enfant (PCM) a été réalisé par téléphone et de ce fait, on a demandé à la mère de réacheminer le questionnaire autoadministré par la poste. Au volet 2004 les entrevues ont de nouveau été réalisées en face à face au domicile de l'enfant mais les tests cognitifs ont, comme au volet 2003, été passés à l'école pour la vaste majorité des enfants. Le QAAM était donc remis à la mère au moment de l'entrevue et celle-ci était invitée à l'expédier par la poste une fois le questionnaire complété.

On remarque que la proportion non pondérée de questionnaires complétés, parmi les familles répondantes où la mère ou conjointe du père est présente, a augmenté au volet 2004 comparativement au volet précédent (tableau V). En effet, cette proportion non pondérée était de 83 % au volet 2003 (1 444 / 1 749), comparativement à 89 % au volet 2004 (1 319/1 480). Malgré cette hausse, cette proportion non pondérée demeure en deçà des proportions observées aux cours des cinq premiers volets (96 % à 98 %).

Puisque seules les mères des familles répondantes au QIRI ont eu à remplir le QAAM, la pondération transversale du QAAM du volet 2004 est basée sur la pondération QIRI de ce volet. Cette dernière doit être ajustée pour tenir compte de la non-réponse au QAAM au volet 2004, parmi les mères ou conjointes présentes à ce volet. Cet ajustement est nécessaire puisque les caractéristiques des mères répondantes au QAAM diffèrent de celles des mères non-répondantes. Cet ajustement a donné lieu à une pondération transversale QAAM pour le volet 2004.

Cette pondération transversale peut être utilisée pour les analyses impliquant des variables du QAAM au volet 2004 seulement (en plus de variables du QIRI des volets de 1998 à 2004)¹⁶. Les résultats de ces analyses peuvent être inférés à la sous-population des enfants dont la mère ou la conjointe du père est présente au volet 2004. Cette sous-population représente plus de 99 % de la population totale visée par l'ÉLDEQ.

¹⁵ En acceptant un volet avec données manquantes.

¹⁶ À noter qu'aucune pondération longitudinale n'a été créée pour les volets 1998 à 2004 (contrairement à ce qui a été fait en 2003). Si des analyses particulières faisaient en sorte qu'une telle pondération était nécessaire, une demande pourrait être adressée à l'ISQ.

Tableau V – Nombre de répondantes au QAAM, volets de 1998 à 2004

	volet 1998	volet 1999	volet 2000	volet 2001	volet 2002	volet 2003	volet 2004
nombre de mères/conjointes présentes parmi les familles répondantes au QIRI	2 119	2 042	1 990	1 936	1 930	1 749	1 480
nombre de répondantes au QAAM pour un volet donné	2 051	1 966	1 934	1 897	1 877	1 444	1 319
nombre de répondantes longitudinales au QAAM (pour un volet donné et les précédents)	2 051	1 925	1 839	1 766	1 719	1 338	1 075

3.3.2 Variables considérées et résultats

Les variables considérées pour la modélisation proviennent essentiellement du QIRI du volet 2004. Ces variables concernent la mère ou la conjointe du père, par exemple le statut d'immigrante, le plus haut niveau de scolarité, le groupe d'âge, etc. Les variables considérées peuvent porter également sur la famille, telles le revenu du ménage, la présence des deux parents biologiques dans le ménage, le type de famille, etc.

Des variables dites longitudinales combinant des données des volets 1998 à 2004 ont aussi été considérées pour la modélisation. Ces variables sont: le revenu du ménage (revenu faible à aucun volet (moins de 10 000\$) versus autres ou revenu faible (moins de 15 000\$) à aucun volet, à 1 ou 2 volets, à 3 volets ou plus); le type de famille (monoparentalité à aucun volet, à 1 ou 2 volets, à 3 ou 4 volets, à 5 volets ou plus); la présence du père biologique (père biologique absent à aucun volet, à 1 volet, à 2 ou 3 volets, à 4 volets, à 5 ou 6 volets, à 7 volets); le niveau de suffisance du revenu (insuffisance du revenu à aucun volet, à 1 volet, à 2 ou 3 volets, à 4 volets ou plus); le travail du père au cours des douze derniers mois (n'a pas travaillé au cours des douze mois précédant l'enquête pour plus d'un volet versus autres); la principale source de revenu du ménage (aide sociale à aucun volet, à 1 ou 2 volets, à 3 volets ou plus); la situation en emploi des parents (aucun parent en emploi à aucun volet, à 1 ou 2 volets, à 3 volets ou plus); fréquentation d'une école par la mère au moment de l'enquête (n'a jamais fréquenté d'école au moment de l'enquête; a fréquenté une école à 1 ou 2 volets; a fréquenté une école à 3 volets ou plus).

Les variables retenues pour la modélisation de la non-réponse au QAAM du volet 2004 sont:

- les langues parlées à la maison par la mère/conjointe (ni français, ni anglais versus les autres);

- le revenu du ménage faible (moins de 10 000 \$) pour aucun volet versus les autres;
- le nombre de frères et sœurs de l'enfant-cible (moins de trois ou trois et plus);
- la fréquentation d'une école par la mère au moment de l'enquête (oui ou non).

Au sein des différentes classes d'ajustement de la pondération, la proportion de répondants varie de 64 % à 93 % (relativement à une proportion globale de 87 %). La proportion la plus faible est observée dans la classe où la mère/conjointe ne parle ni français, ni anglais à la maison.

3.4 Pondération transversale des données du QAAP

3.4.1 Non-réponse au QAAP : état de la situation

On remarque que la proportion non pondérée de questionnaires complétés, parmi les familles répondantes où le père ou conjoint de la mère est présent, est pratiquement la même aux volets 2003 et 2004 (tableau VI). En effet, cette proportion non pondérée était de 78 % au volet 2003 (1 184 / 1 510) et de 79 % au volet 2004 (1 002 / 1 271). Malgré ce statut quo, cette proportion non pondérée demeure en deçà des proportions observées aux cours des cinq premiers volets (86 % à 91 %).

Étant donné l'importance de la non-réponse au QAAP au volet 2004, un ajustement de la pondération du QIRI (PEPCM7) a été fait. Cette pondération transversale peut être utilisée pour les analyses impliquant des variables du QAAP au volet 2004 seulement (en plus de variables du QIRI des volets de 1998 à 2004). Les résultats de ces analyses peuvent être inférés à la sous-population des enfants dont le père ou le conjoint de la mère est présent au volet 2004. Cette sous-population représente environ 84 % de la population totale visée par l'ÉLDEQ.

Tableau VI – Nombre de répondants au QAAP, volets de 1998 à 2004

	volet 1998	volet 1999	volet 2000	volet 2001	volet 2002	volet 2003	volet 2004
nombre de pères/conjoints présents parmi les familles répondantes au QIRI	1 950	1 855	1 773	1 723	1 700	1 510	1 271
nombre de répondants au QAAP pour un volet donné	1 775	1 673	1 616	1 519	1 475	1 184	1 002
nombre de répondants longitudinaux au QAAP (pour un volet donné et les précédents)¹⁷	1 775	1 581	1 431	1 311	1 204	966	752

Comme dans le cas du QAAM, seuls les pères des familles répondantes au QIRI ont eu à remplir le QAAP. De ce fait, la pondération transversale du QAAP du volet 2004 est basée sur la pondération QIRI de ce volet. Cette dernière doit être ajustée pour tenir compte de la non-réponse au QAAP au volet 2004, parmi les pères ou conjoints présents à ce volet.

¹⁷ Ce nombre est calculé parmi les pères biologiques ou conjoints présents à ces mêmes volets.

3.4.2 Variables considérées et résultats

Étant donné que les caractéristiques des répondants au QAAP diffèrent de celles des pères/conjoints n'ayant pas répondu, un ajustement de la pondération QIRI du volet 2004 (variable PEPCM7) est requis pour tenir compte de la non-réponse au QAAP de ce même volet. La nouvelle variable de pondération transversale pour le QAAP est PEQAAP7. Rappelons que cette variable est appropriée pour les analyses impliquant des variables tirées du QAAP au volet 2004, en combinaison ou non avec des variables provenant du QIRI aux différents volets.

Encore une fois les variables considérées pour la modélisation proviennent essentiellement du QIRI du volet 2004. Ces variables concernant le père ou le conjoint de la mère sont par exemple la principale activité du père au moment de l'enquête, le plus haut niveau de scolarité, le groupe d'âge, etc. D'autres variables portant sur la famille, telles le niveau de revenu du ménage, la présence des deux parents biologiques dans le ménage, le type de famille, la principale source de revenu du ménage, etc. ont également été prises en compte.

Des variables dites longitudinales ont aussi été considérées pour la modélisation ; celles-ci sont, à peu de chose près, les mêmes que celles considérées pour le QAAM . À cela s'ajoute la santé perçue par le père (santé passable ou mauvaise à aucun volet, à 1 volet, à 2 volets et plus) et le type de famille (monoparentale depuis 2003 ou 2004 ou recomposée avec un nouveau conjoint depuis 2003 ou 2004).

Les variables retenues pour la modélisation de la non-réponse au QAAP du volet 2004 sont:

- La présence du père biologique (tous les volets versus moins de 7 volets);
- Langue parlée à la maison par le père (ni français, ni anglais versus autres);
- Principale activité du père au moment de l'enquête (soin de la famille, soin de la famille et travail, retraite versus travail contre rémunération, études, convalescence, recherche d'emploi, autre);
- Niveau de suffisance de revenu (suffisant versus insuffisant ou très insuffisant).

Au sein des différentes classes d'ajustement de la pondération, la proportion de répondants varie de 49 % à 87 %, relativement à 77 % globalement. La proportion la plus faible est observée dans la classe où le père biologique est présent à moins de 7 volets.

3.5 Pondération transversale des données du QPAE et des tests cognitifs

3.5.1 Enfants répondants aux tests cognitifs mais non-répondants au QIRI

On compte 37 enfants ayant participé à au moins un des tests cognitifs au volet 2004 mais dont la famille est non-répondante au QIRI. Parmi ceux-ci, un certain nombre ont également participé soit au test de motricité, soit au QPAE, ou encore ont eu un enseignant ayant rempli le QAAENS. Puisque la pondération QIRI du volet 2004 constitue la base pour la pondération des autres instruments de collecte, ces enfants ont

été considérés répondants aux fins de la pondération QIRI, de sorte qu'un poids QIRI a pu leur être attribué, au même titre que les vrais répondants¹⁸.

3.5.2 Questionnaire papier administré à l'enfant (QPAE)

Tel que discuté à la section 1.2.3, la non-réponse au QPAE, relativement aux enfants ayant participé à au moins un test cognitif, peut être négligée aux fins de la pondération (2,5 %). La pondération des données des tests cognitifs peut par conséquent être utilisée pour l'analyse des données du QPAE.

3.5.3 Tests cognitifs

3.5.3.1 Problèmes chroniques liés au développement

La question des enfants souffrant de problèmes chroniques liés au développement a été étudiée sur la base des troubles diagnostiqués déclarés par le parent à l'un ou l'autre des volets de 2001 à 2004. On constate que parmi les 1 492 répondants au QIRI au volet 2004, sept enfants souffrent d'autisme ou d'incapacité mentale. Parmi ces derniers, cinq enfants ont participé aux tâches cognitives, en tout ou en partie, au volet 2004; parmi ceux-ci, seulement trois enfants ont complété au moins une tâche, et un seul de ces enfants a complété le jeu des nombres et l'ÉVIP. De plus, seulement deux enfants ont participé au test de motricité. Par ailleurs, les enfants souffrant de problèmes chroniques tels que le déficit d'attention, la dysphasie, le retard intellectuel et les troubles de développement, d'apprentissage ou émotif présentent un taux de réponse légèrement inférieur à l'ensemble des enfants.

Les enfants souffrant de problèmes de santé chroniques diagnostiqués tels que l'autisme ou l'incapacité mentale ont été considérés répondants « inaptes » pour l'ensemble des jeux. Dans la suite du texte, le terme « inapte » sera utilisé pour désigner ces enfants. **Les enfants « inaptes » peuvent être considérés répondants aux fins de l'analyse; il est alors suggéré d'imputer les scores les plus faibles (théoriques ou rencontrés) aux essais pour lesquels les données sont manquantes. Autrement, ces enfants pourraient également être exclus de certaines analyses pour lesquelles l'inférence est faite à la sous-population des enfants qui ne présentent pas de telle incapacité chronique.**

Puisque les enfants souffrant d'un autre type de trouble chronique présentent un taux de réponse du même ordre de grandeur que celui de l'ensemble des enfants, leur non-réponse a été modélisée de la même façon que l'ensemble; la variable indiquant la présence ou non de trouble chronique a toutefois été considérée comme variable potentielle pour la formation des classes de pondération. Toutefois, elle n'est pas ressortie dans les résultats de modélisation de la non-réponse.

Pour les 35 enfants n'ayant participé qu'aux tâches cognitives au volet 2004, des valeurs ont été imputées pour les variables socioéconomiques définissant les classes de pondération pour le QIRI. Cette imputation a été faite sur la base des informations disponibles aux autres volets et ne sert qu'aux fins de la pondération.

¹⁸ Plus précisément, deux enfants ont été exclus de ce processus puisqu'ils étaient non-répondants au volet 2002, volet de base pour la pondération des données du QIRI du volet 2004. Ces deux enfants sont par conséquent considérés non-répondants tant au QIRI qu'aux autres instruments de collecte pour le volet 2004.

3.5.3.2 Variables considérées et résultats

Pour l'ensemble des tâches cognitives, les enfants classifiés « inaptes » sont exclus du groupe de ceux sur lesquels porte l'analyse de la non-réponse. Nous supposons que nous avons une liste exhaustive des enfants déclarés souffrant de problèmes tels que l'autisme ou l'incapacité mentale (parmi les répondants à l'enquête des volets de 2001 à 2004). Le poids QIRI de base du volet 2004 est attribué à ces enfants; ils ne représentent donc aucun enfant considéré non-répondant aux tâches cognitives.

Pour les autres enfants, les variables retenues pour la modélisation proviennent généralement du QIRI du volet étudié et sont principalement de nature socioéconomique. Quelques variables portent également sur les habitudes ou le comportement de l'enfant. Finalement, quelques variables ont été étudiées sous forme longitudinale, c'est-à-dire qu'elles ont été créées à partir d'une même mesure prise à différents volets (volet étudié et volets antérieurs). Ces variables portent sur le niveau de suffisance du revenu, le revenu du ménage, le type de famille, le travail de la mère au cours des douze derniers mois et la présence du père biologique.

Pour une variable donnée, les proportions de répondants ont été comparées entre les différentes catégories de cette variable à l'aide d'un test du khi-deux. Par la suite, les variables pour lesquelles le seuil observé est inférieur à 0,15 ont été considérées pour la modélisation de la non-réponse.

Parmi l'ensemble des variables considérées pour l'ajustement pour la non-réponse aux tâches cognitives au volet 2004, les variables suivantes ont été retenues:

- la langue parlée à la maison par la mère (GSDMD6A)
- le statut d'immigrante de la mère (GSDMD1A)
- la présence du père biologique (père absent à au moins un volet versus autres)
- le sexe de l'enfant (ASEXE_2)
- le fait que l'enfant préfère jouer seul (GBEEQ6J2)
- la fréquence d'utilisation de l'ordinateur par l'enfant (GACEQ04A)
- le statut d'emploi de la mère/conjointe au cours des 12 derniers mois (GLFMD1C)

Au sein des différentes classes d'ajustement de la pondération, la proportion de répondants varie de 47 % à 98 % (relativement à une proportion globale de 77 %). Les proportions les plus faibles sont observées lorsque la mère ne parle ni le français, ni l'anglais à la maison, ou encore lorsque la mère est immigrante et ne parle que l'anglais à la maison.

3.6 Pondération transversale des données du test de motricité

La pondération des données du test de motricité consiste à faire un ajustement additionnel à la pondération des tests cognitifs pour tenir compte de la non-réponse partielle au test de motricité (voir schéma 1). La méthode de création des classes de pondération utilisée est semblable à celle utilisée pour la pondération des tests cognitifs. Le même ensemble de variables a également été considéré pour la formation des classes.

Parmi l'ensemble des variables considérées pour l'ajustement pour la non-réponse au test de motricité, seul le statut d'immigrant du père (incluant une catégorie pour les enfants dont le père/conjoint est absent) a été retenu. Au sein des classes de pondération, la proportion de répondants varie de 72 % à 90 % (relativement à une proportion globale de 87 %), la proportion la plus faible étant observée chez les enfants dont le père est immigrant non européen.

3.7 Pondération transversale des données du QAAENS

Tout comme dans le cas du test de motricité, la pondération des données du QAAENS consiste à ajuster les poids des tests cognitifs pour tenir compte de la non-réponse au QAAENS. Les variables considérées pour la modélisation sont les mêmes que dans le cas des tests cognitifs.

Parmi l'ensemble des variables considérées pour l'ajustement pour la non-réponse totale au QAAENS, les variables suivantes ont été retenues:

- le fait que l'enfant préfère jouer seul (GBEEQ6J2)
- le statut d'immigrante de la mère (GSDMD1A)
- la présence de problèmes émotifs ou comportementaux de l'enfant à l'école (GEEEEQ36D)
- la situation en emploi des parents au cours des 12 derniers mois (GLFFD01)
- le fait que l'enfant ne peut rester en place, est agité (GBEEQ6B)

Au sein des différentes classes d'ajustement de la pondération, la proportion de répondants varie de 62 % à 99 % (relativement à une proportion globale de 80 %). Les proportions les plus faibles sont observées lorsque l'enfant préfère souvent jouer seul, ou encore lorsque l'enfant éprouve des problèmes émotifs ou comportementaux à l'école et qu'il ne fait pas partie d'une famille biparentale dont les deux parents sont en emploi (en plus d'avoir une mère non immigrante ou immigrante européenne et de ne pas préférer jouer souvent seul).

4 Fichier de pondération, mises en garde et recommandations aux fins de l'analyse :

4.1 Fichier de pondération

Le fichier SAS E7_DIST.POIDS701 contient les variables de pondération suivantes: PEPCM7 (poids QIRI transversal 2004 et longitudinal 1998-2004), PEQAAM7 (poids QAAM transversal 2004), PEQAAP7 (poids QAAP transversal 2004), PEQAAENS7 (poids QAAENS transversal 2004) et PEQPAE7 (poids QPAE transversal 2004). Le fichier SAS E7_DIST.POIDS_JEUX701 contient, pour sa part, les variables de pondération suivantes: PEJEUX7 (poids transversal aux JEUX de 2004) et PEMOTRI7 (poids transversal au test de motricité 2004). Ces poids doivent être utilisés dans les analyses afin d'inférer les résultats à la population visée tout en minimisant les biais dans les estimations.

Le fichier de pondération (E7_DIST.POIDS_JEUX701) contient également une variable (INAPTE_E7) permettant d'identifier les enfants définis comme étant « inaptes », soit ceux souffrant de problèmes de santé chroniques diagnostiqués tels

que l'autisme ou l'incapacité mentale. Les enfants « inaptes » peuvent être considérés répondants aux fins de l'analyse des données des tests cognitifs et de motricité, qu'ils aient complété une tâche ou non; il est alors suggéré d'imputer les scores les plus faibles (théoriques ou rencontrés) aux essais pour lesquels les données sont manquantes. Autrement, ces enfants pourraient également être exclus de certaines analyses portant sur les tests cognitifs ou de motricité et pour lesquelles l'inférence est faite à la sous-population des enfants qui ne présentent pas de telle incapacité chronique.

4.2 Tests statistiques

Les poids contenus dans les fichiers POIDS701 et POIDS_JEUX701 sont des poids échantillonnaires, c'est-à-dire des poids qui ont été multipliés par une constante de sorte que la somme des poids soit égale à la taille de l'échantillon. Ces poids peuvent par conséquent être utilisés pour faire des tests approximatifs avec des logiciels qui ne tiennent pas compte du plan de sondage complexe dans l'estimation de la variance et les tests statistiques.

Afin de pallier au caractère approximatif des tests statistiques réalisés à l'aide de poids échantillonnaires, il est recommandé d'adopter une approche conservatrice en abaissant le seuil théorique des tests. Par exemple, si l'on souhaite faire des tests au seuil théorique de 0,05, on peut choisir de n'interpréter que les résultats significatifs au seuil 0,01.

Dans le cas particulier de tests du khi-deux sur un tableau de fréquences, l'utilisation des poids échantillonnaires divisés par un effet de plan moyen égal à 1,3 demeure appropriée pour obtenir un test approximatif. Il n'est alors pas nécessaire d'abaisser le seuil des tests.

L'utilisation de poids échantillonnaires comporte toutefois certaines limites. En fait, les poids ramenés à la taille de l'échantillon permettent d'obtenir des proportions estimées non biaisées par rapport au plan de sondage ainsi qu'une taille d'échantillon globale égale à la taille réelle. Ces poids ne préservent toutefois pas la taille d'échantillon de chacune des catégories d'une variable, c'est-à-dire des sous-groupes au sein de la population. En présence de poids peu variables, la somme des poids échantillonnaires pour un sous-groupe est approximativement égale à la taille de celui-ci; l'utilisation de ces poids permet de faire des tests approximatifs valides. Dans le cas contraire, la somme des poids échantillonnaires peut différer de façon importante de la taille d'échantillon pour un sous-groupe. Cela a pour conséquence d'invalider les tests statistiques, à moins qu'ils ne soient réalisés à l'aide d'un logiciel qui permet de tenir compte de l'effet du plan de sondage dans l'estimation des paramètres ainsi que de leur variance.

Par exemple, supposons que l'on analyse les données du QAAP du volet 2004 et que l'on désire étudier les caractéristiques du père selon qu'il est ou non le père biologique de l'enfant cible. L'échantillon des répondants au QAAP comprend 1 002 pères ou conjoints. Parmi ceux-ci, on compte 943 pères biologiques et 56 conjoints ou autres¹⁹, la somme des poids échantillonnaires étant de 911 et 86 respectivement pour ces sous-groupes. C'est donc dire que si l'on ne tient pas compte du plan de sondage, on sous-

¹⁹ On compte également trois cas où ce statut est un grand-parent ou une famille d'accueil.

estimera la variance des estimations pour le sous-groupe des conjoints ou autres puisque la taille d'échantillon sera faussement surévaluée (86 au lieu de 56).

Le seuil observé des tests statistiques sera également sous-estimé dans ce cas. Ainsi, il se pourrait que l'on déclare significatifs des écarts entre les sous-groupes qui ne sont pas réels. Dans le contexte d'une telle analyse, il faudrait plutôt faire une analyse pour chacun des sous-groupes séparément en réajustant les poids de telle sorte que la somme des poids pour chaque sous-groupe soit égale à la taille d'échantillon. Il suffit pour ce faire de diviser les poids par la moyenne des poids pour un sous-groupe. Cette recommandation vaut pour toute analyse portant sur un sous-groupe. Il est important dans ces cas de s'assurer que la somme des poids est approximativement égale à la taille d'échantillon de ce sous-groupe; autrement, un ajustement des poids est requis.

4.3 Choix de la pondération

Pour les analyses n'impliquant que des variables provenant du QIRI, la pondération PEPCM7 est indiquée, peu importe les volets considérés simultanément au volet 2004. Pour les analyses portant sur des variables tirées des autres instruments de collecte (et possiblement du QIRI), la question du choix d'une pondération appropriée se pose. Les poids transversaux des données provenant de ces instruments ont été ajustés pour tenir compte de la non-réponse totale à ces instruments au volet 2004 uniquement. Puisqu'on observe une certaine hiérarchie dans les sous-groupes de répondants d'un instrument à l'autre, il est possible d'identifier une pondération appropriée pour un certain nombre de situations d'analyse. Le tableau VII présente les situations où une pondération est disponible.

Pour les autres situations, le choix d'une pondération appropriée doit tenir compte tant de la perte d'unités d'analyse due à l'absence de poids pour ces unités que de la qualité de l'ajustement pour la non-réponse. En ce sens, une évaluation des pondérations existantes doit être faite afin de déterminer si l'une d'elle peut être utilisée pour l'analyse. Deux exemples illustrant la démarche à suivre pour choisir une pondération adéquate sont présentés en annexe.

Les possibilités d'analyse demeurant nombreuses pour le volet 2004 en raison de la non-réponse aux nombreux instruments de collecte autres que le QIRI, le choix d'une pondération adéquate peut nécessiter un examen cas par cas. En précisant la population visée, de même que les instruments et volets considérés pour l'analyse, une pondération appropriée pourrait être suggérée par l'ISQ lorsque nécessaire.

Tableau VII – Choix d’une pondération

Provenance des variables d’analyse	Pondération appropriée
QIRI 2004 ou de 1998 à 2004 seulement	PEPCM7
QAAM 2004, seul ou avec QIRI (un ou pl. volets)	PEQAAM7
QAAP 2004, seul ou avec QIRI (un ou pl. volets)	PEQAAP7
Tests cognitifs 2004, seul ou avec QIRI (un ou pl. volets)	PEJEUX7
QPAE 2004, seul ou avec QIRI (un ou pl. volets)	PEQPAE7
QAAENS 2004, seul ou avec QIRI (un ou pl. volets)	PEQAAENS7
Test de motricité, seul ou avec QIRI (un ou pl. volets)	PEMOTRI7
Tests cognitifs et QPAE 2004, seul ou avec QIRI (un ou pl. volets)	PEQPAE7
Tests cognitifs et QAAENS 2004, seul ou avec QIRI (un ou pl. volets)	PEQAAENS7
Tests cognitifs et de motricité 2004, seul ou avec QIRI (un ou pl. volets)	PEMOTRI7
QPAE et QAAENS, seul ou avec tests cognitifs ou avec QIRI	PEQAAENS7

5 Références bibliographiques :

BEAUMONT, J.-F. (2002). Quand est-on en présence de non-réponse non-ignorable? *Le Bulletin d'imputation*, 2-1, 2-4.

KAAS, G. V. (1980). « An exploratory technique for investigating large quantities of categorical data », *Applied Statistics*, 29, p.119-127.

ANNEXE

Le choix d'une pondération – deux exemples

On peut choisir d'utiliser la pondération qui minimise la perte d'unités d'analyse afin d'optimiser la précision des estimations et la puissance des tests statistiques. En plus de diminuer la puissance statistique, une perte d'unités d'analyse pourrait entraîner certains biais dans les estimations. Mise à part la perte d'unités d'analyse, le choix d'une pondération doit également considérer la qualité de l'ajustement pour la non-réponse. En effet, un ajustement incomplet pour la non-réponse étant source de biais, il faut s'assurer que la part de non-réponse pour laquelle aucun ajustement n'a été fait est faible. Une proportion de 5 % ou moins peut être considérée faible.

Par exemple, supposons qu'un modèle inclut des variables tirées du QIRI à tous les volets ainsi que des variables tirées du QAAM et des tests cognitifs au volet 2004. Si l'on suppose qu'il n'y a aucune non-réponse partielle²⁰ pour les variables considérées, on compterait 1 082 unités pour cette analyse. Toutes ces unités ont un poids transversal QAAM (PEQAAM7) ainsi qu'un poids aux tests cognitifs (PEJEUX7) pour le volet 2004. Il n'y a donc aucune perte d'unités d'analyse, peu importe la pondération retenue. Toutefois, comparativement à la pondération PEQAAM7, la pondération PEJEUX7 a été ajustée pour une plus grande part de la non-réponse à l'un de ces deux instruments (parmi les répondants au QIRI ou aux jeux du volet 2004). La pondération des tests cognitifs serait donc la meilleure pondération à utiliser pour cette analyse, bien que très imparfaite puisqu'elle ne tiendrait pas compte d'une part importante de la non-réponse, soit la non-réponse des enfants dont la mère n'a pas complété le QAAM mais qui ont participé aux tests cognitifs (environ 30 % de la non-réponse totale, c'est-à-dire la non-réponse combinée au QAAM et aux tests cognitifs). Il faudrait par conséquent tenir compte des biais potentiels induits par cette non-réponse dans l'interprétation des résultats.

Supposons dans un deuxième temps qu'un modèle inclut des variables tirées du QIRI des volets de 1998 à 2004 ainsi que des variables du QAAM des volets de 2000 à 2004. On compterait dans ce cas, s'il n'y a aucune non-réponse partielle, 1 102 unités d'analyse. Les pondérations envisageables pour cette analyse sont la pondération QAAM transversale du volet 2004 et la pondération longitudinale QAAM des volets de 1998 à 2003.²¹ Aucune unité d'analyse ne serait perdue si l'on utilisait la pondération transversale. À l'opposé, l'utilisation de la pondération longitudinale entraînerait la perte de 27 unités d'analyse, soit environ 2 % (27/1102) de celles-ci. Cette perte est toutefois petite dans ce cas-ci et le risque de biais est donc faible. De plus, une part de la non-réponse n'est pas prise en compte avec la pondération

²⁰ En sus des problèmes dus à la non-réponse totale, la perte d'unités d'analyse due à la non-réponse partielle provenant de chacune des variables considérées pour la modélisation doit être étudiée. Si cette non-réponse est importante, les estimations pourraient être entachées d'un biais; l'interprétation des résultats devrait par conséquent en tenir compte. On peut tenter d'expliquer, par exemple, dans quel sens irait le biais s'il y a lieu.

²¹ Rappelons qu'aucune pondération longitudinale spécifique n'a été créée pour les instruments autoadministrés pour les volets de 1998 à 2004.

longitudinale puisqu'elle ne tient pas compte de la non-réponse au QAAM au volet 2004. Cette proportion est de l'ordre de 20 % (263/1 338) puisque l'on compte 263 unités non répondantes au volet 2004 parmi les 1 338 unités répondantes aux volets de 1998 à 2003.

Avec la pondération transversale, aucune unité d'analyse n'est perdue mais une part de la non-réponse n'est pas non plus prise en compte dans l'ajustement de la pondération. En effet, la pondération transversale QAAM pour le volet 2004 a été ajustée pour tenir compte de la non-réponse au QAAM au volet 2004, mais elle ne tient pas compte de la non-réponse au QAAM aux volets de 1998 à 2003²². Dans le cas qui nous occupe, cette part est de l'ordre de 16 % (217/1 319) puisque l'on compte 217 unités non répondantes au QAAM aux volets de 2000 à 2003 parmi les 1 319 unités répondantes au volet 2004. Notons qu'avec la pondération longitudinale, bien que certaines unités d'analyse seraient perdues, la non-réponse au QAAM à chacun des volets de 2000 à 2004 a été considérée dans l'ajustement de la pondération.

Finalement, puisque la pondération transversale QAAM du volet 2004 n'entraîne aucune perte d'unités d'analyse et que la portion de la non-réponse dont la pondération ne tient pas compte est légèrement inférieure à celle observée pour la pondération longitudinale, l'utilisation de la pondération transversale serait recommandée plutôt que de la pondération longitudinale pour faire cette analyse. Toutefois, étant donné la part non négligeable de la non-réponse dont la pondération transversale ne tient pas compte, il faut tenir compte des biais potentiels induits par cette non-réponse dans l'interprétation des résultats.

²² La pondération transversale QAAM pour le volet 2004 n'a pas été ajustée pour tenir compte de la non-réponse au QAAM aux volets 1998 à 2003, ni de la non-réponse au QIRI pour les volets de 1999 à 2003. La non-réponse au QIRI pour ces volets étant négligeable parmi les répondants au volet 2004, les calculs présentés ici n'en tiennent pas compte.