

INSTITUT
DE LA STATISTIQUE
DU QUÉBEC

www.stat.gouv.qc.ca

ÉDUCATION ET MILIEUX DE GARDE

Enquête québécoise sur le parcours préscolaire des enfants de maternelle 2017

Méthodologie de l'enquête



Pour tout renseignement concernant
l'Institut de la statistique du Québec (ISQ)
et les données statistiques dont il dispose,
s'adresser à :

Institut de la statistique du Québec
200, chemin Sainte-Foy
Québec (Québec)
G1R 5T4
Téléphone : 418 691-2401

ou

Téléphone : 1 800 463-4090
(sans frais d'appel au Canada et aux États-Unis)

Site Web : www.stat.gouv.qc.ca

Ce document est disponible seulement en version électronique.

Dépôt légal
Bibliothèque et Archives nationales du Québec
1^{er} trimestre 2019
ISBN : 978-2-550-83453-3 (en ligne)

© Gouvernement du Québec, Institut de la statistique du Québec, 2019

Toute reproduction autre qu'à des fins de consultation personnelle
est interdite sans l'autorisation du gouvernement du Québec.
www.stat.gouv.qc.ca/droits_auteur.htm.

Février 2019

Cette publication a été réalisée par : France Lapointe, Katlyn Thibodeau et Lucie Gingras
Institut de la statistique du Québec

Avec la collaboration de : Robert Courtemanche
Charles Alleyn
Marie-Eve Tremblay
Nathalie Audet
Institut de la statistique du Québec

Sous la coordination de : Robert Courtemanche
Institut de la statistique du Québec

Direction de la méthodologie : Éric Gagnon, directeur (intérimaire)
Institut de la statistique du Québec

Comité de relecture interne : Nathalie Audet
Bertrand Perron
Lucie Gingras
Robert Courtemanche

Comité de relecture externe : Membres du comité scientifique de l'EQPPEM
Personnes nommées au :
Ministère de l'Éducation et de l'Enseignement supérieur
Ministère de la Famille
Ministère de la Santé et des Services sociaux
Avenir d'Enfants

Ont assuré la révision linguistique et l'édition : Sarah Bélanger, à la révision linguistique
Andrée-Ann Sénéchal, à l'édition
Institut de la statistique du Québec

Enquête financée par : Ministère de la Santé et des Services sociaux
Ministère de l'Éducation et de l'Enseignement supérieur
Ministère de la Famille
Avenir d'enfants
Institut de la statistique du Québec

Enquête sous la responsabilité de : Direction des enquêtes longitudinales et sociales

Pour tout renseignement concernant le contenu de cette publication : Direction de la méthodologie
Institut de la statistique du Québec
200, chemin Sainte-Foy, 3^e étage
Québec (Québec) G1R 5T4
Téléphone : 418 691-2410
ou
1 800 463-4090
(sans frais d'appel au Canada et aux États-Unis)
Télécopieur : 418 643-4129
Site Web : www.stat.gouv.qc.ca

Notice bibliographique suggérée

LAPOINTE, France, Katlyn THIBODEAU et Lucie GINGRAS (2019). Enquête québécoise sur le parcours préscolaire des enfants de maternelle 2017. Méthodologie de l'enquête, [En ligne], Québec, Institut de la statistique du Québec, 26 p. [www.stat.gouv.qc.ca/statistiques/education/prescolaire-primaire/eqppem_methodologie.pdf].

Note : Le genre masculin a été utilisé dans le but d'alléger le texte.

Acronymes

CLSC	Centre local de services communautaires
DSP-ASSSM	Direction de santé publique de l'Agence de la santé et des services sociaux de Montréal
EHDA	Enfant handicapé ou en difficulté d'adaptation ou d'apprentissage
ÉÉPPP	Étude pour l'évaluation du programme Passe-Partout
ELDEQ	Étude longitudinale du développement des enfants du Québec
EQDEM	Enquête québécoise sur le développement des enfants à la maternelle
EQPPM	Enquête québécoise sur le parcours préscolaire des enfants de maternelle
ICIDJE	Initiative concertée d'intervention pour le développement des jeunes enfants
IMDPE	Instrument de mesure du développement de la petite enfance
ISQ	Institut de la statistique du Québec
MEES	Ministère de l'Éducation et de l'Enseignement supérieur
MELS	Ministère de l'Éducation, du Loisir et du Sport
MRC	Municipalité régionale de comté
MSSS	Ministère de la Santé et des Services sociaux
OCCS	Offord Center for Child Studies
RLS	Réseau local de services

TABLE DES MATIÈRES

9	INTRODUCTION
10	1 PLAN D'ÉCHANTILLONNAGE
10	1.1 Population visée
11	1.2 Base de sondage
11	1.3 Taille et répartition de l'échantillon
14	2 COLLECTE DES DONNÉES
14	2.1 Prétest
14	2.2 Stratégie de collecte
14	2.2.1 Méthode de collecte de données
15	2.2.2 Questionnaire
15	2.3 Résultats de collecte
17	3 TRAITEMENT DES DONNÉES
17	3.1 Validation et saisie
17	3.2 Pondération
17	3.2.1 Stratégie de pondération
18	3.2.2 Description des étapes
19	3.3 Non-réponse partielle
19	3.3.1 Taux de non-réponse partielle
20	3.3.2 Imputation
21	4 ANALYSE, PRÉCISION ET TESTS STATISTIQUES DANS LE TOME 1 DU RAPPORT DE L'ENQUÊTE
21	4.1 Analyse descriptive
21	4.2 Précision des estimations
21	4.3 Tests statistiques
22	5 PRÉSENTATION DES RÉSULTATS DANS LE TOME 1 DU RAPPORT DE L'ENQUÊTE
22	6 PORTÉE ET LIMITES DE L'ENQUÊTE
24	BIBLIOGRAPHIE

INTRODUCTION

Afin de faire une utilisation adéquate des données et des résultats issus de l'*Enquête québécoise sur le parcours préscolaire des enfants de maternelle* (EQPPEM), il importe de connaître la méthodologie utilisée. En effet, le plan d'échantillonnage de l'enquête, les procédures de collecte et le traitement des données ont tous une incidence sur les estimations produites. La connaissance de ces éléments permettra à l'utilisateur des données d'en faire une analyse judicieuse, laquelle pourra aider à l'élaboration de plans d'action appropriés et adaptés à la population visée.

La première section de ce rapport présente le plan d'échantillonnage de l'enquête, alors que la deuxième décrit brièvement les opérations de collecte de données et le bilan de celles-ci. La troisième section fournit une description des différentes étapes du traitement des données. La quatrième section porte sur le type d'analyses réalisées, la précision des données et les tests statistiques. Ensuite, la cinquième section décrit les normes de présentation des résultats. Le rapport se conclut par la portée et les limites de l'enquête.

Un des objectifs de l'EQPPEM est de fournir plus d'information sur les divers facteurs qui pourraient être associés au niveau de développement de l'enfant à la maternelle 5 ans, par exemple la fréquentation d'un service de garde éducatif avant son entrée à l'école et certaines caractéristiques de son environnement familial. C'est pourquoi cette enquête est menée à la suite de l'*Enquête québécoise sur le développement des enfants à la maternelle* (EQDEM¹), qui a recueilli, auprès de l'enseignante ou de l'enseignant, les renseignements nécessaires à la mesure du niveau de développement des enfants selon les cinq domaines évalués par l'Instrument de mesure du développement de la petite enfance (IMDPE)².

De manière plus détaillée, l'EQPPEM vise à :

- dresser un portrait du parcours préscolaire des enfants québécois (s'inscrivent dans ce parcours les services de garde, la maternelle 4 ans, le programme Passe-Partout et les haltes-garderies communautaires);
- relever d'autres facteurs qui pourraient être associés au développement de l'enfant (on retient ici des renseignements sur la santé de l'enfant, sur son environnement familial et sur certaines pratiques parentales, par exemple le fait pour les parents d'initier leur enfant à la lecture);
- fournir de l'information sur les caractéristiques socio-économiques des parents et des familles;
- mettre en lien l'expérience préscolaire des enfants et leurs conditions de vie avec les différents domaines de développement mesurés dans l'EQDEM.

C'est par la combinaison des données de l'EQPPEM et de l'EQDEM qu'il est possible d'atteindre ce dernier objectif.

L'EQPPEM tire également profit de la réalisation de l'*Étude pour l'évaluation du programme Passe-Partout* (ÉÉPPP), qui a été menée par l'Institut de la statistique du Québec (ISQ) à la demande du ministère de l'Éducation et de l'Enseignement supérieur (MEES). Cette étude vise une population qui correspond à une part importante de celle visée par l'EQPPEM (environ 33 %). De plus, les deux enquêtes présentent des contenus communs. Avec l'aval des partenaires du projet, l'ISQ a donc intégré l'échantillon de l'ÉÉPPP à celui de l'EQPPEM pour l'examen du parcours préscolaire des enfants de maternelle. La présentation du plan d'échantillonnage de l'enquête met en évidence les bénéfices obtenus par le jumelage avec l'ÉÉPPP.

1. Pour plus de renseignements concernant cette enquête, consulter le rapport de l'enquête, accessible sur le site Web de l'Institut de la statistique du Québec (ISQ) à l'adresse suivante : www.stat.gouv.qc.ca.
2. Copyright © McMaster University, Ontario. L'IMDPE est largement utilisé dans les autres provinces canadiennes et à l'étranger sous l'appellation Early Development Instrument (EDI). Pour plus d'information à propos des expériences canadiennes, se référer au site Web de l'Offord Centre for Child Studies (OCCS) : www.offordcentre.com/readiness/index.html.

1 PLAN D'ÉCHANTILLONNAGE

Cette section présente la population visée, la base de sondage à partir de laquelle l'échantillon d'enfants nécessaire à l'enquête a été sélectionné ainsi que la taille et la répartition de l'échantillon préparé pour l'EQPPEM. Les cibles de répondants associées aux objectifs de précision des estimations à produire sont également mentionnées dans la section 1.3.

1.1 Population visée

La population visée par l'EQPPEM est la même que celle visée par l'EQDEM. Celle-ci correspond à l'ensemble des enfants fréquentant la maternelle 5 ans³ à temps plein⁴ dans les écoles francophones et anglophones, publiques et privées (subventionnées ou non), du Québec.

Quelques enfants de maternelle sont cependant exclus de la population visée. Les enfants fréquentant une école faisant partie des commissions scolaires crie et Kativik ou étant située dans les régions sociosanitaires du Nunavik (17) et des Terres-Cries-de-la-Baie-James (18) ne font pas partie de la population à l'étude⁵. Il en est de même pour les enfants fréquentant des établissements relevant du gouvernement fédéral dans les réserves autochtones. Ces élèves ont été exclus pour les raisons suivantes : les

difficultés à les joindre, les modifications qu'il aurait fallu apporter au questionnaire et les modalités de fréquentation de la maternelle qui peuvent être différentes.

Tous les élèves fréquentant une école spéciale⁶ sont également exclus de l'enquête. À ceux-ci s'ajoutent les enfants qui, selon les fichiers administratifs du MEES ou selon les renseignements obtenus de l'école, ont un code d'élèves handicapés ou en difficulté d'adaptation ou d'apprentissage (EHDAA)⁷ et font partie d'une classe composée de plus de 50 % d'EHDAA. Par contre, les EHDAA fréquentant une classe ordinaire où 50 % et moins des élèves sont considérés comme des EHDAA ont été conservés dans la population visée et ont été retenus pour la collecte des données. Ils sont donc pris en compte dans le calcul des taux de réponse présentés. Cependant, ils sont exclus du domaine retenu pour l'analyse présentée dans les tomes 1 et 2 de l'EQPPEM en raison des recommandations émises par l'Offord Centre for Child Studies (OCCS).

Somme toute, la population visée par l'EQDEM et l'EQPPEM représente 98 % de l'ensemble des enfants de maternelle en 2016-2017⁸. Le tableau 1.1 présente la distribution de ceux-ci selon le secteur d'éducation, en termes d'écoles visées et d'élèves de maternelle 5 ans. La colonne « Domaine d'analyse » correspond au nombre d'enfants de la population visée, à l'exception des EHDAA fréquentant une classe admissible à l'enquête.

Tableau 1.1
Population visée par l'EQPPEM 2017 et domaine d'analyse retenu pour les tomes 1 et 2

Unités visées ou analysées		Secteur public	Secteur privé	Population visée	Domaine d'analyse (EHDAA des classes admissibles exclus)
EQPPEM	Écoles	1 748	147	1 895	1 895
	Élèves	82 567	5 372	87 939	85 851

- Bien que toutes les commissions scolaires du Québec soient tenues d'offrir la maternelle 5 ans, sa fréquentation demeure facultative. Soulignons toutefois que la presque totalité des enfants de 5 ans y sont inscrits.
- Il n'est pas possible d'inscrire un enfant à la maternelle 5 ans à demi-temps, sauf dans de rares cas (MEESR, 2015).
- Un peu plus de 600 enfants étaient inscrits à la maternelle dans ces commissions scolaires en 2016-2017 (renseignements transmis par la Direction de la méthodologie et des études du MEES, fichier d'avril 2017).
- Sont considérées comme « écoles spéciales » les écoles des hôpitaux et des centres de réadaptation.
- Pour obtenir la liste des codes de difficulté accompagnés des définitions, voir MELS, 2007. Consultez Tremblay et Simard (2018) pour les détails sur la définition retenue en 2017.
- Compilation de l'ISQ à partir de renseignements transmis par la Direction de la méthodologie et des études du MEES (fichier d'avril 2017).

1.2 Base de sondage

La base de sondage utilisée par l'EQPPEM a été construite à partir de l'information tirée des systèmes administratifs du MEES. L'ISQ a entre autres reçu divers renseignements nominaux sur les commissions scolaires, les écoles⁹ et les élèves¹⁰. Il faut savoir que les fichiers des systèmes administratifs du MEES sont mis à jour de façon continue pendant l'année scolaire. C'est pourquoi deux extractions ont été faites des banques administratives, soit une en décembre 2016 et une autre en avril 2017. Ce sont les renseignements de l'extraction de décembre 2016 qui ont servi à élaborer la base de sondage de l'enquête. L'information du fichier du MEES extraite en avril 2017, mise à jour avec les renseignements recueillis lors de la collecte de l'EQDEM, a servi à établir de manière définitive la taille de la population visée et a permis d'évaluer la couverture de la base de sondage.

Il faut dire que cette base de sondage comporte un peu de sous-couverture de même qu'une très faible surcouverture. La sous-couverture de la base est associée au fait que des élèves à la maternelle 5 ans ne se trouvaient pas dans les systèmes administratifs du MEES en décembre 2016, mais y figuraient en avril 2017, ou ont été ajoutés par les enseignants de classes de maternelle 5 ans en 2016-2017 sur la liste des élèves de leurs classes respectives. Quant à la surcouverture de la base, elle est due au fait que certains élèves, présents dans le fichier du MEES de décembre 2016, ne figuraient plus dans le fichier du MEES après la mise à jour d'avril 2017 et n'avaient pas non plus été retrouvés lors de la collecte de l'EQDEM. De plus, certains élèves de la base de sondage peuvent avoir quitté le Québec, s'être retrouvés dans une école spéciale, s'être déplacés vers des régions ou des commissions scolaires exclues de l'enquête ou encore s'être trouvés, en 2016-2017, dans une classe composée de plus de 50 % d'EHDAA (classe non admissibles à l'enquête).

Malgré cela, l'utilisation de l'information provenant du MEES dans la construction de la base de sondage garantit une bonne couverture de la population visée et des renseignements de qualité sur les écoles et les enfants. En effet, le fichier du MEES de décembre 2016 a permis d'obtenir une base de sondage comportant une surcouverture¹¹ estimée à moins de 2 % et une sous-couverture¹² à peu près équivalente des enfants de la population visée. De plus, la comparaison de l'information du fichier de décembre 2016 du MEES concernant chaque enfant avec celle du fichier mis à jour en avril 2017 a montré des différences marginales. Par ailleurs, les renseignements nécessaires pour joindre les parents des enfants de l'échantillon étaient rarement manquants dans la base de sondage.

1.3 Taille et répartition de l'échantillon

Avant de détailler l'information concernant la taille et la répartition de l'échantillon de l'EQPPEM, il faut rappeler que cette enquête fait suite à l'EQDEM et qu'elle tire profit de l'échantillon de l'ÉÉPPP. La combinaison des trois enquêtes a une influence sur les paramètres du plan de sondage de l'EQPPEM, qui est le dernier à être élaboré. Celui-ci doit d'abord tenir compte de l'effet de la combinaison de l'EQPPEM avec l'EQDEM. Puis, la contribution de l'échantillon de l'ÉÉPPP à l'EQPPEM est examinée afin que soient déterminées la taille et la répartition de l'échantillon d'enfants qui s'ajoute à celui de l'ÉÉPPP.

Bien que l'EQDEM soit une enquête par recensement à caractère obligatoire menée auprès des commissions scolaires et des écoles privées, elle comporte tout de même un certain niveau de non-réponse globale. Donc, pour les enfants de l'échantillon de l'EQPPEM, on a établi la taille de l'échantillon de départ pour l'enquête en prévoyant des pertes associées à la non-réponse à l'EQDEM¹³.

-
9. Comme les coordonnées des commissions scolaires et des écoles, la langue d'enseignement de l'école, le réseau d'enseignement, etc.
 10. Comme le nom, le prénom, l'âge, le sexe, la langue maternelle, le code postal, le code d'admissibilité exceptionnelle, etc.
 11. La surcouverture correspond au rapport entre le nombre d'enfants de la base de sondage inadmissibles à l'EQPPEM et le nombre d'enfants de la population visée tel qu'établi à la fin de la collecte de l'EQDEM et après la mise à jour du fichier du MEES en avril 2017.
 12. On évalue la sous-couverture en comparant le nombre d'enfants présents dans la base de sondage avec le nombre d'enfants de la population visée tel qu'établi à la fin de la collecte de l'EQDEM et après la mise à jour du fichier du MEES en avril 2017.
 13. Parmi les répondants à l'EQDEM, seuls les enfants pour qui le jumelage d'une réponse de l'enseignant à celle des parents est possible sont considérés ici comme des répondants.

Quant à la combinaison de l'EQPPEM avec l'ÉÉPPP, on doit considérer que cette dernière est une étude de type cas-témoin ayant pour objectif principal de mettre en lien le niveau de développement des enfants à la maternelle avec la participation au programme Passe-Partout. Il faut dire que la population visée par l'ÉÉPPP correspond à un peu moins du tiers de celle visée par l'EQPPEM, puisque le programme n'est pas offert dans toutes les écoles. Par ailleurs, le plan de sondage de l'étude comporte une stratification très fine de la base de sondage de manière à assurer le meilleur « pairage » des enfants ayant participé au programme Passe-Partout avec les témoins. De fait, on a construit quelque 342 strates secondaires en considérant l'école ou le regroupement d'écoles fréquentées par les enfants de maternelle 5 ans en 2016-2017 (stratification secondaire). Le tirage des enfants Passe-Partout est fait indépendamment du tirage des enfants témoins; une stratification (principale) des enfants est ainsi d'abord faite selon la participation (ou non) au programme. Le plan d'échantillonnage de l'ÉÉPPP comporte donc 684 strates. La répartition de l'échantillon des témoins entre les strates est choisie de manière à ce que l'atteinte de l'objectif de la comparaison avec les enfants Passe-Partout soit optimisée et qu'une description fiable de la sous-population des enfants qui n'ont pas participé au programme en 2015-2016 soit obtenue. Ce choix optimal pour l'ÉÉPPP rend l'échantillon des enfants témoins moins efficace pour la description de cette sous-population des enfants de maternelle. Pour la combinaison de l'ÉÉPPP à l'EQPPEM, il faut aussi se rappeler que le plan de sondage de l'ÉÉPPP n'est pas conçu pour répondre à des objectifs régionaux. Mentionnons enfin que le nombre de répondants visés pour l'ÉÉPPP est de 4 300 répondants (2 200 enfants ayant participé au programme et 2 100 enfants témoins).

Pour sa part, l'EQPPEM doit fournir des estimations à l'échelle des régions administratives (RA)¹⁴. En plus, pour les régions sociosanitaires (RSS) de la Capitale-Nationale et de la Chaudière-Appalaches¹⁵, des estimations sont désirées pour des regroupements de réseaux locaux de services (RLS)¹⁶. Le plan de sondage a donc été construit de manière à ce qu'il soit possible :

- d'estimer un phénomène présent chez environ 10 % des enfants avec une bonne précision relative, c'est-à-dire qu'un coefficient de variation¹⁷ (CV) d'au plus 15 % dans chacune des dix-sept RA visées par l'enquête est obtenu;
- d'estimer un phénomène présent chez environ 10 % des enfants avec une bonne précision relative, c'est-à-dire qu'un coefficient de variation (CV) d'au plus 15 % dans chacun des regroupements de RLS définis pour les RSS de la Capitale-Nationale et de la Chaudière-Appalaches est obtenu.

Le nombre de répondants nécessaires à l'atteinte de ces objectifs de précision a été fixé à 440 pour chaque RA et pour chacun des cinq regroupements de RLS des RSS de la Capitale-Nationale et de la Chaudière-Appalaches pour lesquels des estimations sont requises. Pour qu'il soit possible d'offrir des estimations provinciales ayant une meilleure précision statistique, la cible minimale de répondants est fixée à 990 pour la RA de Montréal et à 770 pour la Montérégie. En tout, le nombre total de répondants nécessaires à l'atteinte des objectifs de l'EQPPEM s'élève donc à près de 9 700.

Comme indiqué précédemment, une partie des besoins statistiques de l'EQPPEM sont comblés par l'utilisation des répondants de l'ÉÉPPP. En tenant compte de l'efficacité de l'échantillon de cette étude pour les objectifs de l'EQPPEM¹⁸, l'ISQ a établi qu'environ 6 400 répondants

-
14. La couverture géographique de l'EQPPEM se situe au niveau de la RA ou du RLS (ou regroupement de RLS) *de résidence* des enfants. Puisque nous ne disposons pas de renseignements détaillés à cet égard pour tous les enfants de la base de sondage, la stratification a été faite à partir de l'adresse *de l'école*. Une vérification de la correspondance géographique des deux adresses lorsque disponibles a montré une très forte corrélation entre elles : dans moins de 2 % de cas, on a observé une différence entre les deux.
 15. Il y a une correspondance exacte entre la RA et la RSS pour les régions de la Capitale-Nationale et de la Chaudière-Appalaches. Le découpage en RLS n'impose donc pas de réaménagement particulier.
 16. Pour la RSS de la Capitale-Nationale, l'EQPPEM doit combler des besoins statistiques précis pour le RLS Québec-Sud et le RLS Québec-Nord. La RSS de la Chaudière-Appalaches est découpée en trois groupes : le RLS Alphonse-Desjardins, le RLS de Beauce et les autres RLS de la RSS.
 17. Pour plus de renseignements, voir la section 4.2.
 18. Rappelons que l'ÉÉPPP ne couvre pas l'ensemble du territoire parce que le programme Passe-Partout n'est pas offert partout.

seraient nécessaires en plus de ceux de l'ÉÉPPP (4 300 répondants). Finalement, en tout, le nombre de répondants visés pour l'EQPPEM avoisine donc 10 700. Grâce à la combinaison de l'EQPPEM avec l'ÉÉPPP, une économie d'environ 3 300 répondants a pu être réalisée.

Afin que la cible de répondants déterminée soit atteinte, la taille d'échantillon préparé pour l'EQPPEM est ajustée de manière à ce qu'elle tienne compte du taux d'admissibilité prévu et des taux de réponse attendus pour les questionnaires destinés à l'enseignant (EQDEM) et aux parents (EQPPEM). Un échantillon total d'un peu plus de 18 000 enfants a été sélectionné pour l'EQPPEM. Le tableau 1.2 présente les hypothèses retenues pour les différents facteurs d'ajustement de la taille d'échantillon et le tableau 1.3 présente la composition de l'échantillon préparé ainsi que le nombre de répondants attendu par RA et regroupements de RLS pour les RSS de la Capitale-Nationale et de la Chaudière-Appalaches.

Tableau 1.2
Hypothèses sur le taux d'admissibilité et les taux de réponse pour le calcul de la taille d'échantillon initiale de l'EQPPEM

Taux d'admissibilité attendu	Taux de réponse attendu EQDEM (enseignants)	Taux de réponse attendu EQPPEM (parents)	Taille de l'échantillon EQPPEM
99 %	90 %	66 %	18 062

Tableau 1.3
Taille d'échantillon et nombre de répondants attendu selon le découpage géographique, Québec, 2017

Région administrative/ réseau local de services de l'école	Taille d'échantillon	Nombre de répondants attendu ¹
01 Bas-Saint-Laurent	749	450
02 Saguenay-Lac-Saint-Jean	936	555
03 Capitale-Nationale		
RLS Québec-Sud	789	450
RLS Québec-Nord	860	515
Autres RLS ²	165	100
04 Mauricie	791	460
05 Estrie	893	530
06 Montréal	1 764	995
07 Outaouais	803	460
08 Abitibi-Témiscamingue	723	450
09 Côte-Nord	775	450
10 Nord-du-Québec	171	100
11 Gaspésie-Îles-de-la-Madeleine	686	405
12 Chaudière-Appalaches		
RLS Alphonse-Desjardins	971	580
RLS de Beauce	794	460
Autres RLS	756	450
13 Laval	787	445
14 Lanaudière	873	510
15 Laurentides	848	490
16 Montérégie	2 147	1 360
17 Centre-du-Québec	781	485
Ensemble du Québec	18 062	10 700

1. Le nombre de répondants attendu est parfois plus élevé que les cibles de départ en raison de la combinaison avec l'ÉÉPPP qui n'est pas efficace. C'est le cas par exemple pour l'échantillon de la Montérégie. Pour la région 10, un recensement des enfants de la population visée mène à 100 répondants attendus, si l'on considère les hypothèses retenues pour les taux d'admissibilité et de réponse.

2. Pour ce regroupement de RLS, il n'y a pas d'objectif particulier de précision.

Source : Institut de la statistique du Québec, *Enquête québécoise sur le parcours préscolaire des enfants de maternelle 2017*.

2 COLLECTE DES DONNÉES

2.1 Prétest

Aucun prétest n'était prévu pour l'enquête. Par contre, le questionnaire entier a été vérifié auprès d'un échantillon de parents de 100 enfants entre le 3 et le 5 avril 2017. Cette vérification a permis de valider la compréhension des questions après la réalisation de 58 entrevues par téléphone. À la suite de cette opération, quelques modifications très mineures ont été apportées au questionnaire de l'enquête. Il faut dire que les changements apportés au questionnaire sont marginaux, de sorte que les 58 entrevues réalisées sont conservées dans le fichier de données de l'enquête et servent aux analyses produites dans les tomes 1 et 2 du rapport de l'enquête.

2.2 Stratégie de collecte

La collecte de données de l'EQPPEM a été menée auprès des parents une fois la collecte de l'EQDEM pour leur enfant réalisée auprès des enseignants. En effet, seuls les enfants pour lesquels un questionnaire avait été rempli pour l'EQDEM étaient utilisés pour la collecte auprès des parents.

Il s'agissait d'une collecte multimodale utilisant les modes Web et téléphonique. Les parents avaient ainsi la possibilité de répondre au questionnaire informatisé autoadministré via une page Web, hébergée sur le site sécurisé de l'ISQ (mode IWAO – interview Web assistée par ordinateur), ou encore de prendre part à une entrevue téléphonique assistée par ordinateur menée par un intervieweur (mode ITAO – interview téléphonique assistée par ordinateur). Comme pour les autres enquêtes réalisées à l'ISQ utilisant une collecte multimodale, le mode Web était d'abord offert aux parents des enfants de l'échantillon de l'EQPPEM; le mode téléphonique était utilisé lorsque les parents signalaient à l'ISQ ne pas vouloir remplir le questionnaire en ligne ou être dans l'impossibilité de le faire, ou après des relances téléphoniques ultimes infructueuses effectuées auprès des non-répondants. Avec cette approche, mentionnons qu'environ 7 répondants sur 10 (73 %) ont rempli le questionnaire en mode Web.

Les étapes de la stratégie de collecte adoptée pour l'EQPPEM sont les suivantes : envoi d'une lettre d'information aux parents des enfants pour lesquels un questionnaire a été rempli pour l'EQDEM pour les inviter à participer à l'EQPPEM et les inciter à répondre par Internet en utilisant l'identifiant attribué par l'ISQ¹⁹. Par la suite, une première lettre de relance visant à réitérer aux parents non répondants l'invitation à remplir le questionnaire via Internet a été envoyée. Après une période de temps jugée suffisante, des rappels téléphoniques ont débuté; si la personne mentionnait vouloir toujours remplir le questionnaire sur le Web, on demandait le courriel pour effectuer les relances suivantes. Sinon, les parents étaient invités à réaliser l'entrevue au téléphone s'ils étaient disponibles.

► 2.2.1 Méthode de collecte de données

La collecte de donnée auprès des parents s'est échelonnée sur près de 12 semaines, soit du 19 avril au 3 juillet 2017.

Les intervieweurs ont reçu une formation spécifique à l'enquête pour se familiariser avec le contexte, les objectifs, le contenu du questionnaire, les codes de résultat et le questionnaire informatisé associés à l'enquête. Les intervieweurs étaient chargés de fournir une assistance aux participants désirant remplir le questionnaire en ligne, procédaient à des relances téléphoniques ou réalisaient des entrevues téléphoniques.

Le nombre de questionnaires remplis a été suivi de près, globalement et pour chacun des découpages géographiques du plan de sondage. En outre, plusieurs stratégies ont été mises en place pour que la collaboration des parents des enfants sélectionnés soit favorisée et que la gestion de la collecte se trouve facilitée. Par exemple, en tout temps, les parents disposaient d'un numéro de téléphone et d'une adresse courriel réservés à l'enquête.

Quant aux parents pour lesquels le MEES n'avait pas de numéro de téléphone (environ 5 % des personnes échantillonnées), des recherches ont été effectuées afin que des numéros valides soient trouvés. Les parents pour qui le numéro de téléphone était resté inconnu après la recherche ont reçu une lettre leur demandant de contacter l'ISQ.

19. Ceux n'ayant pas accès à Internet ou ne désirant pas faire l'entrevue en mode Web avaient dès lors la possibilité de contacter l'ISQ par téléphone pour répondre au questionnaire administré par un intervieweur.

Notons que l'on incluait les parents biologiques, adoptifs, les tuteurs et les tuteurs parmi les répondants-clés potentiels pour recueillir l'information concernant l'enfant échantillonné.

Soulignons enfin que l'on a tout mis en place pour alléger le fardeau de réponse des parents. Par exemple, les questionnaires informatisés Web et téléphonique comportaient des passages complexes préprogrammés pour que le parent ne voie que les questions qui correspondent à sa situation. Plusieurs tests ont été réalisés avant le début de la collecte pour que l'on s'assure du bon fonctionnement des applications informatiques IWAO et ITAO.

► 2.2.2 Questionnaire

Le tableau 1.4 présente les thèmes de chacune des sections du questionnaire de l'EQPPEM²⁰.

Tableau 1.4
Contenu du questionnaire adressé aux parents des enfants sélectionnés pour l'EQPPEM

Section A	Description de la famille
Section E	Parcours préscolaire en milieu de garde
Section F	Autres activités ou programmes éducatifs (y compris les activités de lecture et de numératie en contexte familial)
Section G	Santé de l'enfant
Section H	Contexte familial (santé des parents, stress familial et soutien social)
Section I	Résidence et perception de la sécurité du quartier
Section J	Caractéristiques socioéconomiques

Note : Les sections B à D du questionnaire étaient exclusivement destinées aux parents des enfants de l'échantillon de l'ÉÉPPP.

Toutes les questions ont été révisées par une linguiste et traduites en anglais pour qu'il soit possible de réaliser des entrevues dans les deux langues.

Pour le contenu particulier de l'EQPPEM, la durée moyenne du questionnaire lors des appels téléphoniques a été de 16 minutes. La durée du questionnaire Web n'a pas pu être déterminée, puisqu'il est impossible de considérer les pauses ; seul le temps de connexion peut être comptabilisé. Étant donné cette limite importante, la durée moyenne des questionnaires Web n'a pas été calculée.

2.3 Résultats de collecte

À l'instar des taux de réponse généralement présentés par l'ISQ, les taux présentés pour l'EQPPEM 2017 sont pondérés. Ils tiennent compte, entre autres facteurs, de la non-proportionnalité régionale et locale de l'échantillon de l'enquête.

Concernant les hypothèses de départ de l'EQPPEM, le taux de réponse obtenu à l'EQDEM (enseignants) est supérieur à celui anticipé. Le caractère obligatoire de l'EQDEM a certainement favorisé ce bon résultat. La non-réponse est en général associée au fait que les enseignants ne peuvent fournir de réponse pour les élèves qui n'étaient pas dans leur classe depuis au moins un mois ou encore au fait que certains parents ont refusé de faire participer leur enfant à l'EQDEM. Alors que le scénario initial prévoyait l'envoi d'un peu moins de 16 100 unités de l'échantillon de départ à la collecte auprès des parents, ce sont plutôt les parents de près de 17 000 enfants qui ont reçu l'invitation à répondre au questionnaire sur le parcours préscolaire. En fait, le taux d'admissibilité combiné au taux de réponse à l'EQDEM a atteint 93,7 % au lieu du 89,1 % anticipé.

Au total, ce sont 11 827 parents d'enfants de maternelle²¹ qui ont répondu à l'EQPPEM, ce qui correspond à un taux de réponse pondéré global de 66,3 %. Ce taux prend en considération la non-réponse à l'EQDEM et la non-réponse à l'EQPPEM.

Le tableau 1.6 présente le nombre de répondants et le taux de réponse pondéré obtenu par RA ainsi que par RLS (regroupement de RLS) pour les deux régions où un découpage géographique plus fin était requis. La première colonne rend compte du respect des cibles présentées au tableau 1.3, qui ont toutes été atteintes ou dépassées.

20. Le questionnaire de l'EQPPEM est disponible sur le site Web de l'ISQ à l'adresse www.stat.gouv.qc.ca/.

21. De ce nombre, 11 545 enfants appartiennent au domaine d'analyse couvert dans les tomes 1 et 2 du rapport de l'enquête. En effet, 282 enfants, correspondant à des EHDAA fréquentant une classe ordinaire où 50 % et moins des élèves sont considérés comme des EHDAA, sont exclus de l'ensemble de données analysé.

Les colonnes suivantes fournissent l'information pertinente pour la qualité des estimations produites en utilisant le lieu de résidence des enfants.

Tableau 1.6

Nombre de répondants selon le découpage géographique associé à l'adresse de l'école, nombre de répondants (EHDDA inclus ou exclus) et taux de réponse pondéré selon le découpage géographique de l'adresse de résidence, Québec, 2017

Région administrative/ RLS ou regroupement de RLS	EHDDA inclus			EHDDA exclus n
	Nombre de répondants (adresse de l'école) n	Nombre de répondants (adresse de résidence) n	Taux de réponse pondéré ¹ (adresse de résidence) (%)	
01 Bas-Saint-Laurent	547	551	73,4	538
02 Saguenay–Lac-Saint-Jean	658	658	70,6	644
03 Capitale-Nationale				
RLS Québec-Sud	540	485	68,9	465
RLS Québec-Nord	593	634	69,1	623
Autres RLS	106	106	62,2	104
04 Mauricie	519	515	66,7	499
05 Estrie	550	560	62,5	546
06 Montréal	1 036	1 023	61,2	1 014
07 Outaouais	502	503	63,4	483
08 Abitibi-Témiscamingue	464	463	64,9	448
09 Côte-Nord	470	470	61,6	457
10 Nord-du-Québec	110	111	63,6	108
11 Gaspésie–Îles-de-la-Madeleine	428	427	63,2	408
12 Chaudière-Appalaches				
RLS Alphonse-Desjardins	718	734	73,0	712
RLS de Beauce	537	525	72,0	517
Autres RLS	570	565	72,8	551
13 Laval	478	491	60,3	478
14 Lanaudière	553	562	64,0	546
15 Laurentides	558	543	65,9	539
16 Montérégie	1 401	1 406	66,8	1 379
17 Centre-du-Québec	489	495	64,4	486
Ensemble du Québec	11 827	11 827	66,3	11 545

1. Ce taux de réponse tient compte de la réponse à l'EQDEM et à l'EQPPM; il s'agit du rapport entre le nombre pondéré de répondants et le nombre pondéré d'enfants admissibles.

Source: Institut de la statistique du Québec, *Enquête québécoise sur le parcours préscolaire des enfants de maternelle 2017*.

3 TRAITEMENT DES DONNÉES

Cette section fournit de l'information sur le traitement des données de l'EQPPEM. On y présente d'abord les éléments de validation des données de l'enquête, la méthode de pondération retenue et l'analyse de la non-réponse partielle.

3.1 Validation et saisie

Le recours à la collecte assistée par ordinateur, que ce soit Web ou téléphonique, permet une saisie automatique des réponses fournies par les parents. La très grande majorité des questions sont en outre autocodées, c'est-à-dire que l'intervieweur ou le parent n'a qu'à cocher le choix de réponse approprié. Pour certaines questions comportant plusieurs possibilités de réponse, des menus déroulants offrant des choix de réponse prédéterminés ont été conçus.

De plus, les collectes téléphoniques et Web étant entièrement informatisées, plusieurs validations de base sont donc réalisées au fur et à mesure que le questionnaire est rempli, notamment la vérification du respect des choix de réponse pour les questions fermées ou celle de l'adéquation des sauts de section pour les blocs de questions ne s'adressant qu'à un sous-groupe de répondants. Bien que les questionnaires informatisés aient été amplement testés avant le début de la collecte, tous les codes de réponses et les sauts de section, de même que la cohérence des réponses d'un même questionnaire, ont quand même été validés a posteriori.

Par ailleurs, plusieurs questions comportent un choix de réponse ouvert représenté par la catégorie « Autre ». C'est le cas par exemple de la question QJ4, soit « Quel est le plus haut diplôme que vous avez obtenu? » ou encore de la question Q3, soit « Quel est votre lien avec "prénom"? ». Un certain nombre de cas « Autre » a été reclassé à la suite de l'examen de la réponse fournie par le parent. En effet, malgré les consignes données aux parents, certains ont mal codé leur réponse. Par exemple, il était indiqué que si le parent avait une garderie de type familial à la maison, il devait cocher « Dans la maison de quelqu'un » et non « Autre » pour les questions QE3A, QE4A, QE5A, QE6A et QE7A. Ces cas « Autre » ont donc été reclassés dans cette catégorie.

Enfin, les réponses fournies par les parents aux questions sur la fréquentation d'une maternelle 4 ans (QF1) ou sur la participation au programme Passe-Partout (QF2) en

2015-2016 dans une école publique ont été comparées aux informations provenant du fichier administratif du MEES. Pour près de 10 % des enfants, il y avait incohérence entre l'information provenant du système scolaire et celle donnée par le parent. Après examen, les informations de la base de sondage confirmées par les enseignants seraient plus exactes que les réponses fournies par les parents. L'ordre des questions présentées aux parents, le manque de description précise des programmes ou la méconnaissance des parents auraient eu comme effet de créer une surdéclaration de la fréquentation de la maternelle 4 ans temps partiel et temps plein. Ainsi, les résultats des questions QF1 et QF2 ont été retirés du fichier de données pour tous les répondants; une variable dérivée a été construite à partir de la base de sondage et des renseignements fournis par les enseignants, ce qui a permis de pallier ce retrait.

3.2 Pondération

► 3.2.1 Stratégie de pondération

La pondération consiste à attribuer à chaque enfant pour qui un questionnaire a été rempli par le parent, une valeur, c'est-à-dire un poids statistique, qui correspond au nombre d'enfants qu'il « représente » dans la population visée. La pondération est essentielle pour qu'une inférence adéquate soit faite à cette population, puisque cette dernière n'a pas été observée dans sa totalité. En effet, l'EQPPEM est menée à l'aide d'un échantillon et, comme toutes les enquêtes, elle n'est pas exempte de non-réponse. Le poids statistique doit donc tenir compte tant de la probabilité de sélection des enfants, prédéterminée par le plan d'échantillonnage de l'EQPPEM, que de la non-réponse à l'EQDEM et à l'EQPPEM des enfants échantillonnés.

Il faut dire que dans l'EQPPEM, les objectifs régionaux et infrarégionaux de même que la combinaison avec l'échantillon de l'ÉÉPPP entraînent une variation considérable des probabilités de sélection et font que la répartition de l'échantillon selon les RA et les RLS ne représente pas bien celle de la population visée. D'où la nécessité de tenir compte de la pondération pour la production des estimations.

De plus, la non-réponse a touché davantage certains groupes d'enfants de l'échantillon, ce qui a nécessité un ajustement de la pondération visant à réduire un biais potentiel de non-réponse dans les estimations. Cette non-réponse peut survenir à différents moments

du processus de collecte puisque les commissions scolaires et les écoles privées peuvent refuser le jumelage des données recueillies auprès des enseignants avec celles fournies par les parents ; de leur côté, les parents peuvent également refuser ce jumelage ou simplement décider de ne pas participer à l'EQDEM ou à l'EQPPEM. Il s'agit dans tous ces cas d'une non-réponse à l'EQPPEM.

Enfin, la pondération des répondants de l'EQPPEM comporte une dernière étape qui permet de corriger les défauts de couverture de la base de sondage dont il a été question à la section 1.2 et d'améliorer la qualité des estimations produites à l'échelle des régions de résidence des enfants et de celles concernant le niveau de développement des enfants à la maternelle en 2016-2017.

La stratégie de pondération qui a été élaborée par l'ISQ pour l'EQPPEM tient ainsi compte de la probabilité de sélection des enfants des échantillons constitués pour l'enquête, de l'ajustement de la non-réponse totale à l'EQDEM et à l'EQPPEM et d'un ajustement à certains totaux de population connus ou estimés concernant les enfants de maternelle 5 ans de la population visée.

► 3.2.2 Description des étapes

La première étape consiste à calculer le poids de départ de chaque enfant, qui correspond à l'inverse de sa probabilité de sélection. Chacun des 18 062 enfants échantillonnés pour l'EQPPEM reçoit donc un poids équivalent à l'inverse de sa probabilité de sélection dans la strate à laquelle il appartient. Les pondérations de départ des enfants de l'échantillon varient donc d'une strate à l'autre²². Pour chacune, le *poids de base* correspond au résultat de la division du nombre d'enfants dans la base de sondage par le nombre d'enfants sélectionnés dans l'échantillon.

En ce qui a trait à l'EQDEM, bien qu'il y ait une légère non-réponse à l'échelle des écoles et des enseignants, elle est principalement observée à l'échelle des enfants

et correspond au refus de participation de la part de parents. Dans quelques rares cas, l'école a refusé le jumelage des renseignements fournis par l'enseignant avec ceux de l'EQPPEM. Pour la pondération, il a été convenu de traiter cette non-réponse en une seule étape d'ajustement étant donné les taux de réponse très élevés obtenus (environ 95 %). Ainsi, les enfants « répondants », c'est-à-dire ceux pour lesquels un questionnaire ayant été rempli dans le cadre de l'EQDEM est disponible pour un jumelage avec celui de l'EQPPEM, sont pondérés et « représentent » ainsi ceux pour qui ce n'est pas le cas. Pour ce faire, on effectue une analyse pour définir les variables ou caractéristiques selon lesquelles la probabilité d'avoir un questionnaire « répondant » pourrait varier. On utilise la méthode du score²³ pour déterminer quels sont les groupes homogènes de pondération, c'est-à-dire des sous-groupes de l'échantillon à l'intérieur desquels la propension à répondre est assez uniforme²⁴. L'ajustement s'effectue ensuite à l'intérieur de chacun de ces groupes : le poids de base des enfants admissibles à l'EQPPEM pour lesquels un questionnaire EQDEM est disponible est multiplié par l'inverse du taux de réponse (pondéré) dans le groupe homogène de pondération. Le premier ajustement pour la non-réponse est ainsi terminé.

Ensuite, un ajustement pour la non-réponse à l'EQPPEM est effectué : on emploie la méthode du score de nouveau pour constituer des groupes d'enfants homogènes en ce qui a trait à la propension à répondre au questionnaire adressé aux parents. Les probabilités de répondre à ce questionnaire sont encore ici estimées au moyen d'un modèle de régression logistique²⁵.

22. En tout, 725 strates ont été construites pour l'EQPPEM. Elles correspondent au croisement des 684 strates utilisées pour l'ÉÉPPP avec les découpages géographiques particuliers de l'EQPPEM, c.-à-d. 704 strates ; s'ajoutent à ce nombre 21 strates pour les échantillons régionaux et infrarégionaux sélectionnés en plus de l'échantillon de l'ÉÉPPP.

23. Pour plus de renseignements sur la méthode du score, veuillez consulter Haziza et Beaumont, 2007 et Eltinge et Yansaneh, 1997.

24. Les variables retenues pour la construction des groupes homogènes sont : le nombre de parents inscrits dans le fichier du MEES, le lieu de naissance des parents (Québec ou hors Québec), le réseau de l'école des enfants (privé, public) et la combinaison de la région administrative et du RLS/regroupement de RLS de la résidence des enfants.

25. Le modèle retenu comporte plusieurs indicateurs : le lieu de résidence de l'enfant tel que défini par la région administrative ou encore, pour certains, par le réseau local de services ; le nombre de parents inscrits dans le fichier du MEES ; la langue maternelle de l'enfant ; un indicateur calculé à partir du rang de l'indice de défavorisation associé à l'école (IMSE) ; et un indicateur de développement de l'enfant divisé en trois catégories (vulnérable dans aucun domaine, vulnérable dans un domaine et vulnérable dans deux domaines ou plus) produit à partir de l'information fournie par l'enseignant.

Enfin, la dernière étape de la pondération est la post-stratification. Cet ajustement permet de s'assurer que la somme des poids de tous les répondants de l'enquête correspond bien à la population des enfants de maternelle visée par l'EQPPEM, notamment pour certaines caractéristiques associées aux mesures d'intérêt. Comme mentionné en introduction du rapport, l'analyse de la relation entre le parcours préscolaire des enfants, leurs conditions de vie et les différents aspects de leur développement à la maternelle 5 ans fait partie des objectifs de l'EQPPEM. L'EQDEM, qui est basée sur un recensement des enfants inscrits en maternelle en 2016-2017, offre des estimations très précises du niveau de développement de ces enfants à leur entrée à l'école. En effet, des totaux estimés peuvent être produits à partir de l'information obtenue des enseignants au sujet des 83 335 enfants répondants de l'EQDEM, qui représentent près de 95 % de la population visée. Ainsi, les estimations de l'EQDEM résultant du croisement d'un indicateur de développement divisé en trois catégories (vulnérable dans aucun domaine, vulnérable dans un domaine et vulnérable dans deux domaines ou plus²⁶) avec une variable de regroupements de régions/RLS de la résidence des enfants ont servi à la poststratification des répondants de l'EQPPEM.

3.3 Non-réponse partielle

► 3.3.1 Taux de non-réponse partielle

La pondération permet de réduire les biais associés à la non-réponse totale, mais ne comporte pas d'ajustement pour la non-réponse partielle. Celle-ci se produit lorsque pour une question, les réponses sont manquantes ou invalides pour certains des répondants à l'enquête. Il est connu qu'une non-réponse partielle importante, au même titre que la non-réponse totale, peut entraîner certains biais dans les estimations s'il s'avère que les non-répondants (partiels) présentent des caractéristiques différentes de celles des répondants et, de surcroît, si ces caractéristiques sont liées à ce qui est mesuré.

Pour chaque question ou indicateur (construit à partir de plusieurs questions), le taux de non-réponse partielle pondéré se définit comme le rapport entre le nombre pondéré de personnes qui n'ont pas répondu ou qui ont fourni une réponse invalide à la question (ou pour lesquelles l'indicateur ne peut être construit) et le nombre pondéré de personnes admissibles à la question (ou à l'indicateur). Plus ce taux est élevé, plus le risque de biais induits par la non-réponse partielle est grand.

Pour l'EQPPEM, compte tenu de la précision attendue pour les estimations, on fait l'hypothèse qu'une non-réponse partielle inférieure à 5 % a une incidence négligeable sur les estimations sauf dans la situation où cette non-réponse est très caractérisée.

L'examen des taux de non-réponse partielle a montré qu'une seule question dépasse à cet égard le seuil critique, soit celle portant sur les raisons principales pour ne pas avoir fait garder son enfant de façon régulière avant la maternelle (QE11), pour laquelle le taux de non-réponse partielle est de 6,9 %. Un examen des caractéristiques des non-répondants comparativement à celles des répondants a été réalisé afin que les risques potentiels de biais dû à cette non-réponse soient évalués. Cet examen révèle que, comparativement aux répondants, les parents qui n'ont pas répondu à cette question sont proportionnellement surreprésentés chez les parents d'enfants nés à l'extérieur du Canada, parmi les parents qui ont l'anglais (avec ou sans autres langues, sauf le français) comme langue maternelle ou qui résident dans la RMR de Montréal (données non présentées). Les non-répondants sont par ailleurs moins nombreux en proportion que les répondants chez les bénéficiaires de l'assurance-emploi et chez les résidents de petites villes ou de milieux ruraux. Ainsi, l'EQPPEM sous-estimerait très légèrement la proportion d'enfants qui n'ont pas été gardés parce que l'un des parents a fait le choix de demeurer à la maison et surestime un peu la proportion d'enfants n'ayant pas été gardés parce qu'il n'y avait pas de place disponible en service de garde. Dans les deux cas, compte tenu de la précision des estimations, le biais dû à la non-réponse partielle n'a aucune incidence sur les conclusions des analyses effectuées.

26. Pour plus d'information sur le concept de vulnérabilité, sur les domaines de développement et sur les indicateurs qui y sont associés, le lecteur peut consulter le rapport de l'EQDEM 2017, disponible sur le site Web de l'ISQ à l'adresse suivante : www.stat.gouv.qc.ca/.

Quelques indicateurs rendant compte de l'intensité de l'utilisation des services éducatifs présentent également des taux de non-réponse partielle avoisinant le seuil critique de 5 %. Il s'agit de la durée cumulative en services de garde éducatifs (CUMUL_SGER), de l'âge au début de l'utilisation d'un service éducatif (AG_DEB_SERV_EDU_R) et de la dose d'exposition moyenne en services éducatifs (DOSE_EDU_R2). L'examen des biais potentiels liés à la non-réponse partielle a montré qu'il y a proportionnellement plus de non-répondants parmi les familles à faible revenu et parmi les enfants vivant au sein de familles monoparentales ou recomposées. Ce faisant, l'EQPPEM sous-estimerait très légèrement la proportion d'enfants les moins intensément exposés aux services éducatifs, c'est-à-dire ceux qui ont bénéficié de tels services 12 mois ou moins, ceux qui ont commencé à en bénéficier à partir de 4 ans seulement ou ceux pour qui la dose moyenne d'exposition à ces services est jugée faible (24 mois ou moins). Dans chaque cas, le biais lié à la non-réponse partielle serait inférieur à un point de pourcentage et n'aurait aucune incidence sur les conclusions des analyses, une fois la précision des estimations prise en compte.

Par ailleurs, l'examen des taux de non-réponse partielle pour certains indicateurs montre que certains d'entre eux pourraient présenter des biais dus à la non-réponse partielle, bien que les taux de non-réponse partielle soient un peu inférieurs au seuil de 5 %. Les indicateurs dichotomiques de garde régulière en services de garde éducatifs (GARD_SGE) et d'utilisation d'un service éducatif (SERVICE_EDU) pendant au moins 3 mois, de même que ceux présentant le profil longitudinal de modes de garde ou de services éducatifs utilisés entre 0 et 5 ans (PROFIL_LONG et PROFIL_EDU2_8CAT), comportent tous un peu moins de 5 % de non-réponse partielle. Cependant, puisque tous les non-répondants partiels sont des enfants gardés, ces indicateurs présentent une surestimation de la proportion d'enfants non gardés. On recommande donc aux analystes d'utiliser l'indicateur dichotomique simple de garde régulière pendant au moins 3 mois (QE1R) pour produire cette estimation (7,7 % selon cet indicateur c. 8,1 % selon PROFIL_LONG). Il importe aussi de mentionner que ces indicateurs sont tout de même appropriés pour la comparaison des proportions d'enfants selon le mode de garde ou le service éducatif utilisé ou encore pour l'examen comparatif des caractéristiques des groupes

d'enfants définis par les différentes valeurs des indicateurs. Par exemple, l'indicateur de profil longitudinal des modes de garde permet de comparer les caractéristiques des enfants ayant fréquenté exclusivement des CPE entre 0 et 5 ans avec celles des enfants qui n'ont fréquenté aucun service de garde.

Enfin, les indicateurs utilisant le revenu familial avant impôt présentent une non-réponse partielle plus importante qui a été traitée (voir section 3.3.2).

► 3.3.2 Imputation

Lorsque le risque de biais associé à la non-réponse partielle est jugé préoccupant, l'imputation des données manquantes est une solution souvent mise de l'avant. Pour une variable présentant de la non-réponse partielle, l'imputation consiste à attribuer une valeur à un individu en remplacement d'une donnée manquante ou invalide au moyen des renseignements disponibles à son sujet²⁷.

Pour l'EQPPEM, l'ISQ a procédé à l'imputation des données des non-répondants partiels pour le revenu familial avant impôt qui sert à la construction de l'indicateur de revenu en quatre catégories basé sur la mesure de faible revenu (MFR) utilisé comme variable de croisement. L'indicateur est calculé à partir des questions sur le revenu total provenant de toutes les sources, avant impôt et autres retenues, de tous les membres du ménage pour les 12 derniers mois (QJ1 et QJ2 et sous-questions QJ1 [A à E] et QJ2 [A à E]). Le répondant avait la possibilité de fournir cette information soit sous forme de valeur exacte, bien qu'approximative, soit sous forme de catégorie en indiquant la tranche de revenu. Parfois, la déclaration sous forme de catégorie de revenus n'est pas assez précise pour permettre l'attribution d'une valeur à l'indicateur de revenu. Lorsque ces cas sont ajoutés à ceux pour lesquels aucune réponse n'a été fournie au sujet du revenu du ménage, le taux pondéré de non-réponse partielle atteint 6,84 %. L'imputation déductive et par donneur (« Hot Deck ») a permis de réduire cette non-réponse partielle à 1,82 %, les cas pour lesquels aucune information sur le revenu n'est disponible étant demeurés « non-répondants » pour l'indicateur de revenu.

27. Ces informations sur les non-répondants partiels peuvent provenir du questionnaire d'enquête lui-même, ou encore de la base de sondage.

4 ANALYSE, PRÉCISION ET TESTS STATISTIQUES DANS LE TOME 1 DU RAPPORT DE L'ENQUÊTE

4.1 Analyse descriptive

Dans le premier tome du rapport de l'enquête, ce sont presque uniquement des proportions qui sont estimées, bien qu'il arrive qu'au passage, on présente des estimations de moyennes ou de médianes. Dans tous les cas, le recours à des données pondérées rend possible l'inférence adéquate à la population du domaine d'analyse retenu.

Par ailleurs, les analyses effectuées sont descriptives et généralement bivariées. Elles ne tiennent pas compte des corrélations complexes entre les variables, qui pourraient modifier ou expliquer en tout ou en partie les associations initialement détectées entre un indicateur et une variable de croisement (effet interactif ou confusionnel), comme pourrait le faire une analyse multivariée. Le tome 2 de l'EQPPEM présente des résultats d'analyses multivariées réalisées à partir des données de l'EQPPEM.

4.2 Précision des estimations

Les données tirées d'enquêtes statistiques comme l'EQPPEM comportent des erreurs dites d'échantillonnage, dues notamment au fait que seule une partie des unités de la population est sélectionnée pour participer à l'enquête et qu'il y a de la non-réponse globale. Ces erreurs se répercutent sur la précision des estimations produites. La complexité du plan d'échantillonnage, qui n'est pas un plan aléatoire simple, et la non-réponse globale à l'EQDEM et à l'EQPPEM influencent la précision des estimations. Il est nécessaire de mesurer la précision de chaque estimation (proportion ou moyenne) et d'en tenir compte dans l'interprétation des résultats inférés à la population visée.

Pour l'EQPPEM, l'estimation de la variance et les tests statistiques sont effectués à l'aide de poids d'auto-amorçage, aussi appelés « poids *bootstrap* ». Une série

de 500 poids d'autoamorçage a été créée, ce qui permet de tenir compte adéquatement non seulement du plan de sondage complexe, mais aussi de tous les ajustements de non-réponse et de poststratification apportés à la pondération dans l'estimation de la variance et les tests statistiques. Ces poids ont par la suite été utilisés avec le logiciel SUDAAN.

Le coefficient de variation (CV) est l'indicateur de précision relative recommandé par l'ISQ pour l'analyse des données de l'EQPPEM. Les estimations dont le CV est supérieur à 15 % sont accompagnées d'un astérisque dans les tableaux et les figures²⁸.

Le CV se calcule comme suit :

$$CV = \frac{\sqrt{\text{variance de l'estimation}}}{\text{estimation}}$$

4.3 Tests statistiques

Pour l'analyse des données de l'EQPPEM présentée dans le tome 1, l'ISQ a d'abord effectué un test statistique d'indépendance du khi-deux²⁹ pour déceler un lien possible entre l'indicateur analysé et une variable de croisement correspondant généralement à une caractéristique socioéconomique ou autre. Les tests dont le seuil de signification observé est inférieur au seuil fixé à 5 % montrent une association significative.

Des tests d'égalité de deux proportions peuvent ensuite être réalisés pour que les associations révélées par les tests du khi-deux significatifs soient mieux comprises. La comparaison est basée sur la construction d'une statistique de Wald à partir de la différence des transformations *logit*³⁰ des proportions estimées. La variance et la covariance de la différence des *logits* sont calculées à l'aide des poids d'autoamorçage. Notons qu'un écart important entre deux proportions n'est pas nécessairement statistiquement significatif. Tout dépend de la variabilité associée aux estimations. Le seuil de signification pour qu'un écart soit considéré comme significatif est de 5 %.

28. Pour plus de renseignements sur la présentation des CV, consulter la section 5.1.

29. On utilise une version modifiée du test du khi-deux habituel qui tient compte du plan de sondage de l'enquête : il s'agit de l'ajustement de Satterthwaite du test du khi-deux (Rao et Scott, 1981). Plus précisément, c'est la statistique F correspondant à cette correction du test du khi-deux qui est utilisée dans les analyses.

30. La transformation « logit » = $\log(p/(1-p))$ est utilisée comme transformation normalisatrice dans la construction d'intervalles de confiance afin que leurs taux de couverture soient améliorés, en particulier pour de petites proportions (Korn et Graubard, 1999).

5 PRÉSENTATION DES RÉSULTATS DANS LE TOME 1 DU RAPPORT DE L'ENQUÊTE

Des indications générales pour l'ensemble des résultats présentés dans le premier tome du rapport d'analyse de l'enquête sont fournies dans cette section. Elles concernent notamment la description et la présentation des résultats.

En général, les estimations de proportions sont arrondies à l'unité³¹ près quand elles sont mentionnées dans le texte et à une décimale près dans les tableaux et figures, à l'exception des estimations de proportions de moins de 5 %, pour lesquelles une décimale est conservée dans le texte. En raison de l'arrondissement, la somme des proportions de certains tableaux ou figures peut être légèrement supérieure ou inférieure à 100 %.

Les estimations dont le CV est inférieur ou égal à 15 % sont suffisamment précises et sont présentées sans indication à cet égard. La présentation de celles dont le CV est supérieur à 15 % est basée sur le système de classification du tableau 1.6. Par ailleurs, la présentation des résultats rend compte du fait que les statistiques fournies sont basées sur un échantillon. Des expressions appropriées qui confirment qu'il ne s'agit pas de valeurs exactes sont donc utilisées dans le rapport.

Tableau 1.6
Système de classification de la précision des estimations en vigueur à l'Institut de la statistique du Québec

Précision de l'estimation (CV)	Signe distinctif	Réserve à mentionner lors de la diffusion
$CV \leq 15\%$	(aucun signe)	Estimation diffusée sans mention
$15\% < CV \leq 25\%$	*	Estimation à interpréter avec prudence
$CV > 25\%$	**	Estimation à utiliser avec circonspection ; fournie à titre indicatif seulement

Source : Institut de la statistique du Québec.

En général, seuls les résultats significatifs au seuil de 5 % sont mentionnés dans le rapport d'analyse. En présence d'un résultat global significatif (selon le test du khi-deux), des lettres ajoutées en exposant aux statistiques présentées dans les tableaux et figures indiquent quelles sont les paires de catégories d'une variable de croisement pour lesquelles on détecte des écarts significatifs, au seuil de 5 %, pour les proportions correspondant aux phénomènes d'intérêt. Une même lettre indique un écart significatif entre deux catégories. Il peut arriver que deux proportions qui semblent différentes ne le soient pas d'un point de vue statistique, en raison notamment du petit nombre d'individus sur lequel est basée l'estimation. On dit, dans ce cas, qu'il n'y a pas de différence statistiquement significative ou que l'enquête ne permet pas de détecter de différence entre ces proportions.

6 PORTÉE ET LIMITES DE L'ENQUÊTE

Tout a été mis en place pour assurer la qualité et la représentativité de l'EQPPEM. La disponibilité d'une base de sondage construite avec des fichiers récents détenus par le MEES a, par exemple, permis de préciser la définition des enfants de la population visée. En plus d'offrir une excellente couverture, cette base comporte plusieurs renseignements facilitant le contact avec les parents des enfants des échantillons. D'ailleurs, la possibilité de réaliser une stratification adéquate des enfants de la population visée selon la région administrative ou le réseau local de services dans lequel se trouve l'école fréquentée par l'enfant assure l'élaboration d'un plan d'échantillonnage efficace pour les objectifs de l'EQPPEM.

L'échantillon disponible pour l'EQPPEM est de taille considérable à l'échelle provinciale. Le nombre de répondants disponible à l'échelle régionale offre également un potentiel intéressant pour la production d'estimations de qualité pour des phénomènes peu fréquents touchant aussi peu que 10 % de la population.

31. Dans le texte, les proportions dont la décimale est ,5 dans les tableaux et les figures ont été arrondies à l'unité inférieure ou supérieure selon la seconde décimale. Par exemple, la proportion 20,47 % est arrondie à 20 %, alors que la proportion 20,53 % l'est à 21 %. Les arrondissements des proportions se terminant par ,5 dans les tableaux et figures varient donc dans le texte.

La collecte informatisée permet aussi d'optimiser la qualité des données recueillies auprès des parents. Elle assure également un fardeau de réponse limité aux répondants. Il peut cependant y avoir certaines erreurs pour les mesures portant sur des périodes de référence plus longues, comme celles associées au congé pris par les parents au moment de la naissance ou de l'adoption de l'enfant. Dans ce cas précis, puisque la venue ou l'adoption d'un enfant est un événement marquant dans la vie, on peut penser que les risques de ces erreurs sont moins importants pour l'EQPPEM.

Les procédures inférentielles utilisées ont aussi fait l'objet d'une attention particulière. Ainsi, les pondérations utilisées dans l'analyse des données ont été calculées de façon à ce que la fiabilité de l'inférence à la population visée soit assurée. Les mesures de redressement effectuées permettent également de réduire au minimum les biais potentiels dus à la non-réponse. Il faut noter que tous les renseignements de la base de sondage et ceux fournis par les enseignants sont utilisés pour l'ajustement de la non-réponse associée aux questionnaires adressés aux parents.

Comme c'est le cas pour toutes les enquêtes populationnelles, bien que toutes les précautions possibles aient été prises pour que la qualité des données soit assurée et que les biais soient minimisés, certaines limites doivent être considérées. Ainsi, il est impossible de garantir l'exactitude des réponses fournies par les parents qui peuvent, par exemple, avoir de la difficulté à se souvenir de choses passées depuis plusieurs mois, voire quelques années, comme c'est le cas pour certains aspects particuliers (nombre d'heures de garde, par exemple) du parcours préscolaire de leur enfant. Il est aussi possible que les parents ignorent certaines informations, lesquelles peuvent, par exemple, dans le cas des parents adoptifs, concerner les premières années de vie de leur enfant. Les réponses des parents à certaines questions, par exemple celles portant sur les activités parentales d'initiation à la littératie et la numératie, pourraient également être entachées d'un biais de désirabilité sociale. À cela s'ajoute le facteur « enseignant » en ce

qui a trait à la mesure du développement des enfants étudiés, c'est-à-dire le fait que deux enseignants ne jugent pas nécessairement de la même façon un élève; leurs réponses peuvent alors être teintées de subjectivité, d'autant plus que certains peuvent avoir été plus sévères que d'autres lorsqu'ils ont répondu au questionnaire de l'IMDPE utilisé dans l'EQDEM.

Il faut aussi rappeler que les analyses présentées dans les tomes 1 et 2 du rapport de l'enquête n'incluent pas les enfants considérés EHDAA. Cela doit être gardé en mémoire, en particulier lors de la lecture des chapitres présentant la vulnérabilité des enfants en lien avec leurs caractéristiques, leur contexte familial ou leur parcours préscolaire.

Enfin, une enquête transversale³¹ comme l'EQPPEM permet de déceler des liens entre deux variables de même que des différences entre des groupes de population. Ce genre d'enquête ne permet toutefois pas d'établir de lien de causalité entre les caractéristiques étudiées. En lien avec certains objectifs de l'enquête, les analyses présentées dans le tome 1 de l'EQPPEM s'appuient essentiellement sur des méthodes bivariées d'analyse statistique. L'interprétation de certains résultats doit donc être faite avec prudence. Seules des analyses multivariées comme celles présentées dans le tome 2 du rapport permettent de contrôler de manière plus appropriée l'effet de confusion possible de certains facteurs exogènes. L'approche retenue pour le tome 1 du rapport d'analyse permet néanmoins d'explorer les données recueillies de façon utile et fournit ainsi une bonne description du parcours préscolaire des enfants visés par l'EQPPEM.

32. Ce type d'enquête mesure un phénomène au cours d'une période donnée.

BIBLIOGRAPHIE

- ELTINGE, J. L., et I. S. YANSANEH (1997). « Méthodes diagnostiques pour la construction de cellules de correction pour la non-réponse, avec application à la non-réponse aux questions sur le revenu de la U.S. Consumer Expenditure Survey », *Techniques d'enquête*, vol. 23, n° 1, p. 3745.
- HAZIZA, D., et J.-F. BEAUMONT (2007). "On the Construction of Imputation Classes in Surveys", *International Statistical Review*, vol. 75, n° 1, avril, p. 25-43.
- KORN, E. L., et B. I. GRAUBARD (1999). *Analysis of Health Surveys*, New York, John Wiley & Sons, 382 p.
- MINISTÈRE DE L'ÉDUCATION, DU LOISIR ET DU SPORT (MELS) [Québec] (2007). *L'organisation des services éducatifs aux élèves à risque et aux élèves handicapés ou en difficulté d'adaptation ou d'apprentissage (EHDAA)*, [En ligne], Québec, Gouvernement du Québec, 25 p. [www.education.gouv.qc.ca/fileadmin/site_web/documents/dpse/adaptation_serv_comp/19-7065.pdf].
- MINISTÈRE DE L'ÉDUCATION, DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE (MEESR) [Québec] (2015). *Indicateurs de l'éducation. Éducation préscolaire, enseignement primaire et secondaire. Édition 2014*, [En ligne], Québec, Gouvernement du Québec, 84 p. [www.education.gouv.qc.ca/fileadmin/site_web/documents/PSG/statistiques_info_decisionnelle/indicateurs_2014_fr.pdf].
- RAO, J. N. K., et A. J. SCOTT (1981). "The analysis of categorical data from complex surveys: Chi-squared tests for goodness of fit and independence in two-way tables", *Journal of the American Statistical Association*, vol. 76, n° 374, p. 221-230.
- SIMARD, M., A. LAVOIE et N. Audet (2018). *Enquête québécoise sur le développement des enfants à la maternelle 2017*, [En ligne], Québec, Institut de la statistique du Québec, 126 p. [www.stat.gouv.qc.ca/statistiques/sante/enfants-ados/developpement-enfants-maternelle-2017.pdf].
- TREMBLAY, M.-E., et M. SIMARD (2018). *Méthodologie de l'Enquête québécoise sur le développement des enfants à la maternelle 2017*, [En ligne], Québec, Institut de la statistique du Québec, 40 p. [www.stat.gouv.qc.ca/statistiques/sante/enfants-ados/eqdem-rapport-methodologique-2017.pdf].

Des statistiques sur le Québec d'hier et d'aujourd'hui
pour le Québec de demain