

Étude sur le développement des enfants en contexte de pandémie – volet « enseignant »

Méthodologie de l'étude

Institut de la statistique du Québec

Mars 2023

Publication réalisée

à l'Institut de la statistique du Québec par :

Marilène Courteau et Amélie Lavoie

Avec la collaboration de :

Catherine Côté, Raphaëlle Élément et Amélie Ducharme

Sous la coordination de :

Marie-Eve Tremblay et Nathalie Audet

Révision linguistique et mise en page :

Direction de la diffusion et des communications

Étude financée par :

Ministère de la Santé et des Services Sociaux

Université de Montréal

Ministère de l'Éducation

Étude sous la responsabilité de :

Direction des enquêtes et des indicateurs sociaux

Institut de la statistique du Québec

Pour tout renseignement concernant

le contenu de ce document :

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

Notice bibliographique suggérée pour le rapport

COURTEAU, Marilène, et Amélie LAVOIE (2023). *Enquête sur le développement des enfants en contexte de pandémie – volet « enseignant »*. *Méthodologie de l'enquête*, [En ligne], Québec, Institut de la statistique du Québec, 20 p.

Avertissement

En raison de l'arrondissement des données, le total ne correspond pas nécessairement à la somme des parties. Afin de faciliter la lecture des résultats, les proportions de 5 % et plus sont arrondies à l'unité dans le corps du texte. Ainsi, les proportions dont la décimale est de ,5 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 % est arrondie à 21 %. Les arrondissements des proportions se terminant par ,5 dans les tableaux et figures varient donc dans le corps du texte.

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

Table des matières

Introduction	Erreur ! Signet non défini.
1.1 Plan d'échantillonnage	6
1.1.1 Population visée	6
1.1.2 Base de sondage	7
1.1.3 Méthode de sélection, taille et répartition de l'échantillon	7
1.2 Collecte de données	9
1.2.1 Questionnaire d'étude	9
1.2.2 Stratégie de collecte.....	9
1.2.3 Résultats de la collecte	11
1.3 Traitement des données.....	12
1.3.1 Validation	12
1.3.2 Pondération.....	13
1.3.3 Non-réponse partielle	15
1.4 Précision des estimations.....	16
1.5 Portée et limites de l'étude	17
Conclusion	18
Références bibliographiques	19
Annexe	20

Introduction

L'Institut de la statistique du Québec (ISQ) s'est fait confier par l'Observatoire pour l'éducation et la santé des enfants (OPES) le mandat de réaliser l'*Étude sur le développement des enfants en contexte de pandémie* en collaboration avec le ministère de l'Éducation du Québec (MEQ). Créé en 2021, l'OPES cherche à comprendre les répercussions qu'aura eue la pandémie de COVID-19 sur les enfants. Ces derniers peuvent en effet avoir subi des perturbations tant sur le plan scolaire que sur le plan familial et un portrait juste de la situation aidera les décideurs à prendre les mesures qui s'imposent en matière d'éducation, de développement psychosocial, de santé et de bien-être afin d'atténuer les conséquences négatives de la crise sanitaire sur les jeunes générations d'aujourd'hui. L'OPES s'intéresse à différents thèmes tels la santé mentale, les habitudes de vie, le retard scolaire, le décrochage scolaire, les relations sociales et familiales ainsi que l'utilisation des services sociaux et de santé.

Pour atteindre ses objectifs de recherche, l'OPES utilisera différentes données, dont des données administratives et des données d'enquêtes déjà menées ou qui seront menées dans le futur, et s'inspirera des travaux de l'ISQ et de chercheurs. L'OPES souhaite, entre autres, suivre dans le temps les enfants qui fréquentaient la maternelle 5 ans en 2016-2017 et qui avaient été visés par l'*Enquête québécoise sur le développement des enfants à la maternelle 2017* (EQDEM 2017). Grâce aux données recueillies dans le cadre de cette étude, laquelle a été réalisée quelques années avant la crise sanitaire, il sera possible de mettre en relation l'état de développement de ces enfants tel qu'il était avant la pandémie avec celui qu'il était au moment de l'étude.

L'*Étude sur le développement des enfants en contexte de pandémie* se divise en deux volets : le volet « parent » et le volet « enseignant ». La collecte de données du volet « parent » a été réalisée au cours de l'été 2021 auprès des parents d'enfants pour lesquels un questionnaire avait été rempli dans le cadre de l'EQDEM en 2017¹.

Pour le deuxième volet, des données ont été recueillies auprès d'enseignantes et enseignants de cinquième année du primaire au printemps 2022. Le présent rapport porte sur le deuxième volet, soit le volet « enseignant ».

Objectifs de l'étude

Le volet « enseignant » de l'étude vise principalement à mesurer certains aspects de l'état du développement des enfants alors qu'ils sont en cinquième année du primaire. Un jumelage des données recueillies dans le cadre de ce deuxième volet de l'étude avec celles obtenues dans le cadre du volet « parent », avec les données tirées de

¹ Le fichier de données ainsi que le rapport méthodologique décrivant les principaux paramètres du volet « parent » de l'étude sont disponibles au Centre d'accès aux données de recherche de l'ISQ (CADRISQ).

l'EQDEM 2017, ainsi qu'avec certains renseignements de banques de données administratives liés notamment à l'éducation, à la santé et aux services sociaux, permettra de vérifier si la pandémie de COVID-19 a pu affecter différents aspects du développement des enfants qui étaient en cinquième année du primaire durant l'année scolaire 2021-2022.

L'étude vise plus spécifiquement à recueillir de l'information sur les aspects suivants :

- le cheminement scolaire de l'enfant ;
- son état de développement (bien-être physique, développement socioaffectif, comportements, etc.) ;
- la relation de l'enseignant avec l'enfant ;
- certains problèmes identifiés chez l'enfant et les services reçus à l'école.

La première section de ce rapport présente le plan d'échantillonnage de l'étude (population visée, base de sondage, échantillonnage), et 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 la précision des données et la présentation des résultats. Enfin, le rapport se conclut par la portée et les limites de l'étude.

1.1 Plan d'échantillonnage

1.1.1 Population visée

La population visée par le volet « enseignant » de l'*Étude sur le développement des enfants en contexte de pandémie* correspond à l'ensemble des enfants qui avaient été visés par l'EQDEM 2017 et qui étaient en cinquième année du primaire dans une école francophone ou anglophone, publique ou privée (subventionnée ou non), du Québec durant l'année scolaire 2021-2022. Précisons que les enfants qui avaient été visés par l'EQDEM 2017 étaient ceux qui fréquentaient la maternelle 5 ans à temps plein en 2016-2017, dans une école francophone ou anglophone, publique ou privée (subventionnée ou non), du Québec. Toutefois, les enfants qui résidaient dans les régions sociosanitaires 17 (Nunavik) ou 18 (Terres-Cries-de-la-Baie-James) étaient exclus².

Sont donc exclus de la population visée par l'*Étude sur le développement des enfants en contexte de pandémie* les enfants ayant redoublé une année scolaire, ceux faisant l'école à la maison, ceux ayant déménagé à l'extérieur du Québec après la maternelle et les enfants de cinquième année du primaire qui ne faisaient pas partie de la population

² Pour obtenir plus de renseignements sur la méthodologie utilisée pour l'EQDEM 2017, voir Tremblay et Simard (2018).

visée par l'EQDEM 2017, comme les nouveaux arrivants et les enfants résidant dans les régions sociosanitaires 17 (Nunavik) ou 18 (Terres-Cries-de-la-Baie-James).

Il est à noter que bien que la population visée soit composée d'enfants, les réponses au questionnaire de l'étude étaient fournies par leur enseignante ou enseignant. Les questions portaient sur des comportements simples et observables par le personnel enseignant.

1.1.2 Base de sondage

La base de sondage utilisée pour sélectionner l'échantillon de l'étude a été élaborée à partir des données recueillies dans le cadre de l'EQDEM 2017. Des informations provenant des fichiers du ministère de l'Éducation du Québec ont été ajoutées à la base de sondage après la sélection de l'échantillon. Ces informations ont entre autres été utilisées pour contacter les écoles et les enseignantes et enseignants des enfants sélectionnés.

1.1.3 Méthode de sélection, taille et répartition de l'échantillon

L'un des principaux objectifs de l'étude est de fournir des estimations fiables à la fois pour l'ensemble de la province, pour le territoire de la région métropolitaine de recensement (RMR) de Montréal, et pour le territoire hors de la RMR de Montréal et ce, tant pour l'ensemble de la population visée que pour la sous-population comptant les 10 % d'enfants les plus vulnérables. Le plan de sondage de cette étude a donc été construit de manière à ce qu'il soit possible³ :

- d'estimer une proportion de 10 % ou plus pour la sous-population comptant les 10 % d'enfants les plus vulnérables avec une bonne précision relative, c'est-à-dire un coefficient de variation (CV)⁴ d'au plus 15 % à l'échelle provinciale;
- d'estimer une proportion de 10 % ou plus pour la sous-population comptant les 10 % d'enfants les plus vulnérables avec une bonne précision relative, c'est-à-dire un coefficient de variation (CV)⁵ d'au plus 15 % à l'échelle régionale (c'est-à-dire pour la RMR de Montréal et pour le territoire hors de la RMR de Montréal).

Le nombre total de répondants nécessaire à l'atteinte des objectifs mentionnés précédemment a été fixé à 8 800 enfants. En faisant des hypothèses quant aux taux de réponse et d'admissibilité⁶, on a pu déterminer la taille d'échantillon nécessaire pour chacune des strates. Ces strates, visant à assurer un certain contrôle de la composition

³ Dans la mesure où les hypothèses de collecte (taux de réponse et taux d'admissibilité) sont valides.

⁴ Des renseignements sur le calcul du coefficient de variation sont fournis à la section 1.4.

⁵ Des renseignements sur le calcul du coefficient de variation sont fournis à la section 1.4.

⁶ Les hypothèses pour le taux d'admissibilité et pour le taux de réponse à l'échelle provinciale sont respectivement de 87 % et 90 %.

de l'échantillon, ont été formées par le croisement des variables suivantes : le découpage géographique principal (qui se divise en deux : la RMR de Montréal et le territoire hors de la RMR de Montréal) et les régions administratives se trouvant à l'intérieur du découpage géographique principal. Au total, 20 strates ont été formées et un échantillon de 11 201 enfants⁷ a été sélectionné pour l'étude. Ces enfants sont issus de 4 152 classes réparties dans 1 715 écoles publiques et 125 écoles privées. Soulignons que des enfants ont été sélectionnés dans l'ensemble des 70 centres de services scolaires et commissions scolaires. Le 0 présente la taille de l'échantillon et le nombre de répondants visés par strate. Il est à noter qu'un enfant est considéré comme un répondant si son enseignante ou enseignant a rempli un questionnaire à son sujet.

Tableau 1.1. Taille de l'échantillon et nombre de répondants visés pour le volet « enseignant » de l'Étude sur le développement des enfants en contexte de pandémie

Découpage géographique principal	Région administrative	Taille d'échantillon	Répondants visés (n)
RMR de Montréal	Montréal (06)	2 381	1 908
	Laval (13)	559	463
	Lanaudière (14)	481	396
	Laurentides (15)	559	440
	Montérégie (16)	1 496	1 193
Hors RMR de Montréal	Bas-Saint-Laurent (01)	258	205
	Saguenay–Lac-Saint-Jean (02)	358	296
	Capitale-Nationale (03)	1 048	807
	Mauricie (04)	320	248
	Estrie (05)	451	342
	Outaouais (07)	657	464
	Abitibi-Témiscamingue (08)	218	172
	Côte-Nord (09)	116	90
	Nord-du-Québec (10)	24	19
	Gaspésie–Îles-de-la-Madeleine (11)	93	73
	Chaudière-Appalaches (12)	660	513
	Lanaudière (14)	233	176
	Laurentides (15)	220	175
	Montérégie (16)	741	552
	Centre-du-Québec (17)	348	269
Ensemble de la population visée		11 201	8 800

Source : Institut de la statistique du Québec, volet « enseignant » de l'Étude sur le développement des enfants en contexte de pandémie, 2022

⁷ Il est à noter que l'échantillon du volet « parent » de l'étude est un sous-échantillon du volet « enseignant ».

1.2 Collecte de données

1.2.1 Questionnaire d'étude

Le questionnaire de l'étude, développé par l'Observatoire et l'ISQ, visait à recueillir des renseignements sur différents aspects du développement des enfants de cinquième année du primaire.

En moyenne, les répondants ont pris environ 14 minutes pour remplir le questionnaire sur le Web. Outre l'introduction et les remerciements, le questionnaire comportait les grandes sections suivantes :

- **Section A** : À propos de l'enfant
- **Section B** : À propos du cheminement scolaire de l'enfant
- **Section C** : À propos du bien-être physique de l'enfant
- **Section D** : À propos du développement socioaffectif de l'enfant
- **Section E** : À propos des habitudes de travail de l'enfant
- **Section F** : À propos des comportements de l'enfant
- **Section G** : À propos de votre relation avec l'enfant
- **Section H** : Préoccupations particulières à propos de l'enfant

1.2.2 Stratégie de collecte

La collecte de données s'est déroulée du 24 mars au 16 juin 2022. Les enseignantes et les enseignants devaient alors remplir un questionnaire sur le Web pour chacun des enfants de leur classe sélectionné pour l'étude. Toutefois, avant le début de la collecte de données, plusieurs étapes préalables ont dû être réalisées, et ce, dès l'automne 2021.

En premier lieu, avant même de contacter les enseignants des enfants sélectionnés, les centres de services scolaires (CSS), les commissions scolaires (CS) et les écoles privées devaient d'abord signer le formulaire de consentement⁸ pour le partage des renseignements de l'étude avec le MEQ et pour leur jumelage avec :

- les données recueillies dans le cadre de l'EQDEM 2017 ;
- les informations recueillies dans le cadre du volet « parent » de l'étude ;
- certains renseignements de banques de données administratives liées, notamment, à l'éducation, à la santé et aux services sociaux.

⁸ Au total, six écoles n'ont pas signé le formulaire de consentement et trois ont refusé le partage et le jumelage des données.

L'envoi du formulaire de consentement ainsi que de la lettre d'information sur l'étude aux CSS et aux CS a été fait le 26 novembre 2021⁹. L'envoi aux écoles privées et aux écoles publiques a pour sa part eu lieu dans la semaine du 14 février 2022, et les appels dans les écoles pour obtenir les informations nécessaires à la collecte de donnée ont commencé le 14 mars 2022.

Après l'obtention de l'information requise auprès des écoles, des trousse d'information ont été envoyées aux enseignantes et enseignants de cinquième année concernés. Chaque trousse contenait :

- une lettre d'invitation (pour chaque enseignante ou enseignant concerné) ;
- une note portant sur la compensation financière (pour chaque enseignante ou enseignant concerné);
- la procédure de connexion (pour chaque enseignante ou enseignant concerné);
- une note d'informations supplémentaires sur l'étude (pour chaque enseignante ou enseignant concerné);
- une lettre d'information comprenant un coupon de refus à remettre aux parents des enfants visés par l'étude;
- un mémo à la direction pour expliquer quoi faire avec toute cette documentation.

Chaque personne a donc reçu une enveloppe contenant l'information nécessaire pour participer à l'étude, y compris les identifiants Web pour se connecter au questionnaire des enfants sélectionnés dans sa classe.

Pour faciliter le contact et encourager la participation, des relances courriel ont directement été envoyées au personnel enseignant afin de l'inciter à remplir les questionnaires manquants. Des relances courriel ont également été envoyées pour annoncer la date de fin de la collecte (initialement prévue le 31 mai) et la prolongation de la collecte (d'abord jusqu'au 10 juin, et ensuite jusqu'au 15 juin). Un courriel a enfin été envoyé aux enseignantes et enseignants participants ainsi qu'aux directions d'écoles, aux CSS et aux CS pour les remercier d'avoir participé à l'étude.

À noter également qu'une compensation financière a été versée aux CS, aux CSS et aux écoles privées pour la participation de leur personnel enseignant à l'étude, soit le remboursement des frais de suppléance occasionnés par la libération des enseignants. Ces paiements ont été attribués selon le nombre de questionnaires remplis par chaque enseignant (p. ex. : paiement d'une demi-journée pour les enseignants ayant rempli de trois à six questionnaires).

Soulignons enfin que pour s'assurer de répondre aux objectifs de l'étude, l'ISQ a attribué à celle-ci un caractère obligatoire¹⁰, et ce, conformément à la Loi sur l'Institut de la statistique du Québec (RLRQ, chapitre I-13.011).

⁹ Il y a eu un deuxième envoi au CSS/CS le 14 février, concernant des changements dans les paramètres de l'étude (compensation).

¹⁰ Le caractère obligatoire de l'étude ne s'appliquait pas aux parents des enfants visés par le volet « enseignant » de l'étude. Les parents avaient le droit de refuser que l'enseignante ou l'enseignant remplisse un questionnaire au sujet de leur enfant.

1.2.3 Résultats de la collecte

Comme les taux de réponse généralement présentés par l'ISQ, ceux présentés pour l'étude sont pondérés. Ils tiennent compte, entre autres facteurs, de la non-proportionnalité régionale¹¹ de l'échantillon de l'étude.

Taux d'admissibilité

Le taux d'admissibilité pondéré¹² à l'étude est de 90 %. Parmi les enfants non admissibles à l'étude, 63 % n'étaient pas en cinquième année du primaire (il est possible qu'ils aient redoublé ou sauté une année scolaire) et 18 % étaient scolarisés en dehors du Québec. Les autres causes d'inadmissibilité étaient le fait de résider dans les régions sociosanitaires 17 ou 18 et celui d'être scolarisé à la maison.

Taux de réponse

Au total, 8 217 enfants sont répondants à l'étude, ce qui correspond à un taux de réponse pondéré¹³ de 81,7 %. Ces enfants répondants sont issus de 3 661 classes réparties dans 1607 écoles publiques et 104 écoles privées. Parmi les non-répondants à l'étude, 54 % ont un parent qui a refusé que l'enseignante ou l'enseignant remplisse un questionnaire au sujet de son enfant, 6 % ont changé d'école, 5 % sont des enfants pour lesquels le MEQ n'a pas pu trouver d'information dans ses banques de données et 0,4 % fréquentent une école qui a refusé de consentir au partage et au jumelage des données. Les autres cas sont des élèves pour lesquels l'enseignante ou l'enseignant n'a pas rempli de questionnaire sans mention de la raison.

Le taux de réponse pondéré pour la RMR de Montréal est de 80 % contre 84 % pour le territoire hors de la RMR de Montréal. Le tableau 1.2 présente le nombre de répondants et le taux de réponse pondéré pour les découpages géographiques et pour l'ensemble de la population visée.

¹¹ Étant donné les objectifs de précision pour le découpage géographique principal (qui se divise en deux : la RMR de Montréal et le territoire hors de la RMR de Montréal), la proportion de la taille de l'échantillon par rapport à la taille de la population diffère d'une division à l'autre. Les régions situées en dehors de la RMR de Montréal sont légèrement surreprésentées dans l'échantillon par rapport à leur poids dans la population.

¹² Taux d'admissibilité pondéré (avec le poids de sélection). Plus de renseignements sur la pondération sont fournis à la section 1.3.2.

¹³ Taux de réponse pondéré avec le poids initial ajusté pour l'admissibilité (ajustement décrit à la section 1.3.2).

Tableau 1.2. Nombre de répondants et taux de réponse pondéré pour chacune des deux divisions du découpage géographique principal et pour l'ensemble de la population visée

Découpage géographique principal	Répondants (n)	Taux de réponse pondéré (%)
RMR de Montréal	4 025	80,0
Hors RMR de Montréal	4 192	83,7
Ensemble de la population visée	8 217	81,7

Source : Institut de la statistique du Québec, volet « enseignant » de l'*Étude sur le développement des enfants en contexte de pandémie, 2022*.

1.3 Traitement des données

1.3.1 Validation

Le recours aux questionnaires informatisés (AIAO) permet une saisie automatique des réponses. La très grande majorité des questions sont autocodées, c'est-à-dire que l'enseignante ou l'enseignant n'a qu'à cocher le choix de réponse : la saisie s'en trouve donc réduite au minimum. De plus, comme les questionnaires sont entièrement informatisés, plusieurs validations de base sont réalisées pendant que l'enseignant répond aux questions sur le Web¹⁴, 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.

Des vérifications concernant le statut de réponse des enfants (répondants, non répondants ou inadmissibles) ont aussi été effectuées *a posteriori*. Celles-ci nous ont principalement permis de conserver comme répondants certains enfants dont l'enseignant n'avait pas rempli le questionnaire en entier et d'en exclure d'autres pour lesquels l'enseignant n'avait pas répondu à une proportion suffisante de questions.

De plus, le fichier des données de l'étude a fait l'objet de vérifications. La première avait pour but de s'assurer que, pour chaque question, tous les répondants qui devaient y répondre l'avaient fait et que tous ceux qui ne devaient pas y répondre ne l'avaient pas fait.

Enfin, les codes postaux fournis par les enseignants ont également fait l'objet de vérifications, afin de s'assurer que chaque répondant était classé dans la bonne région administrative et que les indicateurs qui seraient créés, tel que l'indice de défavorisation matérielle et sociale, allaient être adéquats.

¹⁴ La collecte pour cette étude était effectuée uniquement sur le Web.

1.3.2 Pondération

La pondération consiste à attribuer à chaque répondant de l'étude une valeur, c'est-à-dire un poids qui correspond au nombre d'enfants qu'il « représente » dans la population. La pondération est essentielle pour la production de diverses estimations à partir des données recueillies dans le cadre de l'étude. Elle nous permet de rapporter les données des répondants à la population visée et, ainsi, de faire des inférences adéquates à cette population, bien que celle-ci n'ait pas été observée dans sa totalité.

En raison des objectifs de diffusion de l'étude selon le découpage géographique principal (la RMR de Montréal et le territoire hors de la RMR de Montréal), l'échantillon a non seulement été réparti selon ce découpage, mais aussi selon les régions administratives se trouvant à l'intérieur de celui-ci. Toutefois, cette répartition est légèrement différente de celle de la population du Québec. La pondération permet la correction de la distorsion créée par le plan de sondage utilisé pour la sélection de l'échantillon.

De plus, on sait que dans les enquêtes, la probabilité de répondre varie selon plusieurs caractéristiques sociodémographiques. Il est donc nécessaire de tenir compte de ces éléments en les intégrant à la pondération qui sera utilisée lors de la production de statistiques à partir des données de l'étude. La pondération est également l'un des éléments à prendre en compte pour estimer correctement la précision des données.

La stratégie de pondération établie pour cette étude tient compte, entre autres, de la probabilité qu'un enfant soit sélectionné dans l'échantillon et de la portion de l'échantillon qui s'est avérée inadmissible lors de la collecte des données. Elle comprend également un ajustement pour la non-réponse totale à l'étude.

Description des étapes

La première étape consiste à calculer le poids initial de chaque individu faisant partie de l'échantillon. Étant donné que la base de sondage est constituée des enfants pour lesquels un questionnaire a été rempli dans le cadre de l'EQDEM 2017, le poids initial d'un individu est égal au produit obtenu lorsqu'on multiplie le poids qu'il avait dans le cadre de l'EQDEM 2017 (« poids EQDEM 2017 ») par un facteur d'ajustement. Ce facteur d'ajustement est le même pour tous les individus d'une même strate. En effet, pour chacune des strates, la somme des poids initiaux des individus de l'échantillon doit être égale à la somme des poids EQDEM 2017 de tous les individus de la strate. Ainsi, le facteur d'ajustement est égal au quotient que l'on obtient lorsqu'on divise la somme des poids EQDEM 2017 de tous les individus de la strate par la somme des poids EQDEM 2017 des individus de la strate faisant partie de l'échantillon.

Ensuite, on doit ajuster le poids initial de chaque individu pour tenir compte de l'inadmissibilité de certains enfants à l'étude. Cet ajustement se fait en deux étapes : on doit d'abord le faire en fonction de l'inadmissibilité établie

préalablement à la collecte et ensuite selon l'inadmissibilité établie en cours de collecte. En effet, pour cette étude, les données obtenues du MEQ après la sélection de l'échantillon permettent d'identifier a priori certains enfants non admissibles avant même la collecte de données auprès des enseignantes et enseignants (enfant n'étant pas en cinquième année du primaire, enfant scolarisé en dehors du Québec, enfant résidant dans les régions sociosanitaires 17 ou 18). Pour ces enfants, considérés comme inadmissibles avant même la collecte, le poids est égal à zéro. Par la suite, pour les enfants restants, le statut d'admissibilité final étant inconnu pour plusieurs non-répondants, il faut réduire leur poids pour tenir compte du fait qu'une partie d'entre eux est probablement inadmissible. Pour ce faire, on multiplie le poids de ces enfants par le taux d'admissibilité établi en cours de collecte pondéré, qui correspond à la proportion pondérée (avec le poids initial) d'enfants admissibles parmi ceux pour lesquels l'admissibilité a pu être déterminée lors de l'étude.

L'étape suivante est l'ajustement pour la non-réponse totale. À cet effet, on effectue une analyse pour définir les variables ou les caractéristiques selon lesquelles la probabilité d'avoir un questionnaire considéré comme rempli pourrait varier. On utilise la méthode du score¹⁵ pour déterminer quels sont les groupes homogènes de pondération, c'est-à-dire les sous-groupes de l'échantillon à l'intérieur desquels la propension à répondre est assez uniforme¹⁶. L'ajustement s'effectue ensuite à l'intérieur de chacun de ces groupes : le poids des enfants pour lesquels un questionnaire considéré comme rempli est disponible est ajusté par l'inverse du taux de réponse (pondéré) dans le groupe homogène de pondération.

Par la suite, on vérifie qu'aucun enfant n'a de poids très élevé comparativement au poids des enfants d'un même découpage géographique ou d'une même région administrative. Seuls deux cas de figure ont été relevés dans le cadre de cette étude. Dans ces cas, le poids sera égal au poids maximal de la strate (formée par le croisement du découpage géographique et de la région) dont ils font partie en excluant les poids extrêmes, puis les poids de la strate sont ensuite ajustés selon un même facteur (facteur = poids total avant correction/poids total après correction) pour que le poids total de la strate après correction des poids extrêmes soit le même qu'avant la correction.

¹⁵ Pour obtenir plus de renseignements sur la méthode du score, voir Haziza et Beaumont (2007) et Eltinge et Yansaneh (1997).

¹⁶ Les variables (qui sont créées à partir des données obtenues du MEQ ou tirées de l'EQDEM 2017) retenues dans le modèle sont : la région, la langue maternelle, le lieu de naissance des parents, l'indice de défavorisation matérielle et sociale, le fait de fréquenter une école publique ou privée ainsi que la présence d'une vulnérabilité dans au moins un domaine de développement selon l'EQDEM 2017.

1.3.3 Non-réponse partielle

Définition

La pondération tient compte de la non-réponse totale, mais n'a pas été ajustée en fonction de la non-réponse partielle. Cette dernière se produit lorsqu'une variable (ou lorsqu'un indicateur créé à partir de plusieurs questions) comporte des données manquantes pour les répondants à l'étude. On sait qu'une non-réponse partielle importante peut entraîner certains biais dans les estimations, au même titre que la non-réponse totale, s'il s'avère que les non-répondants présentent des caractéristiques différentes de celles des répondants et, de surcroît, si ces caractéristiques sont liées au thème abordé. Le taux de non-réponse partielle pondéré pour une variable donnée se définit comme le rapport entre le nombre pondéré d'enfants pour lesquels aucune réponse n'a été fournie à la question et le nombre pondéré d'enfants admissibles à la question. Plus ce taux est élevé, plus le risque de biais induits par la non-réponse partielle est grand.

En général, lorsque l'on analyse des données d'enquête, il est raisonnable de faire l'hypothèse qu'un taux de non-réponse partielle inférieur à 5 % a un effet négligeable sur les estimations à l'échelle provinciale, sauf dans la situation où cette non-réponse est très caractérisée. Les variables pour lesquelles le taux de non-réponse partielle dépasse le seuil de 5 % doivent en général être analysées plus en profondeur au regard des biais. En raison des faibles taux de non-réponse partielle observés pour la grande majorité des variables de l'étude, les risques de biais, relativement à la précision des estimations, sont assez faibles.

Les huit variables ayant un taux de non-réponse partielle supérieur à 5 % sont présentées dans le *Tableau A1 – Variables ayant un taux de non-réponse partielle pondéré supérieur à 5 %, Québec, 2022* se trouvant en annexe. Ces variables portent toutes sur le fait que l'enfant ait reçu ou non des services de différentes ressources professionnelles non enseignantes à l'école¹⁷.

Analyse de la non-réponse partielle

Aucune des variables ayant un taux de non-réponse partielle supérieur à 5 % n'a fait l'objet d'une imputation. En effet, à la suite d'une analyse des biais potentiels, l'effet de la non-réponse partielle sur les résultats produits est jugé négligeable. Cependant, étant donné un taux de non-réponse partielle de près de 30%, nous recommandons de ne pas utiliser la sous-question H-Q4J, laquelle demandait si, oui ou non, l'enfant a reçu les services d'un autre professionnel que ceux mentionnés aux sous-questions H-Q4A à H-Q4I (« *Autre, veuillez préciser* »).

¹⁷ Infirmière, orthophoniste, psychoéducateur, etc.

1.4 Précision des estimations

Dans toute enquête statistique, les estimations produites sont entachées d'erreurs. Il y a d'une part l'erreur liée à l'échantillonnage (voir la section suivante), qui est attribuable au fait que seule une partie de la population visée est sélectionnée pour participer à l'enquête, et, d'autre part, les erreurs qu'entraînent, par exemple, la non-réponse et les erreurs de mesure associées à des questions difficiles à interpréter, à la saisie ou au traitement de données. Il a déjà été question, dans les sections précédentes, de la principale source d'erreur non liée à l'échantillonnage, soit la non-réponse.

L'erreur due à l'échantillonnage

Ce type d'erreur se répercute sur les estimations produites, dont la précision est influencée par la complexité du plan d'échantillonnage. Il est donc nécessaire de mesurer la précision de chaque estimation avant de procéder à l'analyse et à l'interprétation des résultats inférés à la population visée. Pour cette étude, une série de 500 poids d'autoamorçage, aussi appelés « poids *bootstrap* », a été créée afin de tenir compte adéquatement non seulement du plan de sondage complexe, mais aussi de tous les ajustements pour l'inadmissibilité et la non-réponse apportés à la pondération dans l'estimation de la variance et les tests statistiques.

Le coefficient de variation (CV), qui permet la mesure de la précision relative d'une estimation, peut être utilisé pour quantifier l'erreur d'échantillonnage. Plus le CV est petit, plus l'estimation est précise. Le CV s'exprime comme suit :

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

Une autre mesure de la précision d'une estimation est l'étendue de l'intervalle de confiance (IC) : plus l'intervalle est court, plus la valeur du paramètre est circonscrite. L'intervalle de confiance associé à une proportion représente la zone d'incertitude liée à l'estimation. Cette zone d'incertitude est étroitement liée au niveau de confiance choisi. Ainsi, un intervalle de confiance pour une proportion à un niveau de confiance de 95 % signifie que si l'on répétait l'étude 100 fois et qu'à chaque fois on estimait la proportion et calculait son intervalle de confiance, 95 des 100 intervalles ainsi créés contiendraient la vraie valeur de la proportion dans la population.

Présentation des résultats

Lors de la présentation de résultats tirés de l'étude, il est important de rendre compte du fait que les statistiques fournies sont des estimations et non des valeurs exactes.

1.5 Portée et limites de l'étude

La participation à l'étude a été bonne, comme le démontre le taux de réponse obtenu lors de la collecte des données (82 %). Cependant, ce taux est inférieur à celui attendu (90 %). Malgré ce fait, le potentiel analytique des données de l'étude se rapportant à l'ensemble des répondants est bon et la précision obtenue est près de celle visée au moment de la conception du plan de sondage.

La pondération a permis de réduire les biais engendrés par la non-réponse dans les estimations, notamment grâce à des variables explicatives présentes dans la base de sondage.

Par ailleurs, comme c'est le cas dans la plupart des enquêtes populationnelles comportant des données autodéclarées, il est difficile de garantir l'exactitude des réponses fournies par les répondants. Ces derniers peuvent être influencés par le phénomène de désirabilité sociale ou par la difficulté de se rappeler des choses du passé.

De plus, dans le cas où les résultats de l'étude permettent de déceler des liens entre deux variables, de même que des différences entre des groupes de population, il est important de noter qu'ils ne permettent pas d'établir de lien de causalité entre les caractéristiques étudiées. Mentionnons qu'il convient également de tenir compte de facteurs potentiellement confondants dans l'analyse des liens entre variables.

Finalement, il est important de tenir compte des mises en garde et des recommandations mentionnées dans les sections précédentes afin d'utiliser adéquatement les données de cette étude. La section 1.4 explique, entre autres, comment tenir compte du plan de sondage complexe lors de la mesure de la précision des résultats produits. Il est à noter qu'une analyse de la non-réponse partielle devrait être effectuée pour tous les résultats qui seront produits afin de vérifier si cette non-réponse est importante et problématique.

Conclusion

Le présent document contient les informations méthodologiques relatives au volet « enseignant » de *l'Étude sur le développement des enfants en contexte de pandémie*. La population visée y est présentée. Des informations sur le plan d'échantillonnage et la collecte de données y sont également incluses afin que les conditions et le contexte dans lequel l'étude a été réalisée soient bien situés. Comme l'étude a été faite de manière à ce que ses résultats soient inférés à la population visée, une section portant sur le traitement des données, notamment les méthodes de pondération appliquées, se trouve également dans ce rapport. Enfin, des informations relatives aux méthodes d'analyse des données issues de l'étude sont fournies.

On retiendra que l'étude a permis de joindre, entre le 24 mars et le 16 juin 2022, les enseignants de près de 10 000 enfants admissibles sélectionnés pour l'étude. Rappelons que pour qu'un enfant soit sélectionné, un questionnaire portant sur leur état de développement devait avoir été rempli en 2017 par leur enseignant alors qu'ils étaient à la maternelle 5 ans. De plus, pour être admissibles à l'étude, les enfants sélectionnés devaient fréquenter la cinquième année du primaire dans un établissement du Québec durant l'année scolaire 2021-2022. La participation à l'étude a été bonne, comme en témoigne le taux de réponse de 82 %. Le potentiel analytique de l'étude est bon, surtout pour les estimations visant l'ensemble de la population. Mentionnons aussi que très peu de questions sont associées à un taux de non-réponse partielle supérieur à 5 %.

L'ensemble des informations présentées dans ce rapport vise à guider l'utilisateur des données dans son analyse et son interprétation des résultats du volet « enseignant » de l'étude.

Références bibliographiques

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*, [En ligne], produit n° 12-001-X19970013103 au catalogue de Statistique Canada, vol. 23, n° 1, juin, p. 37-45. [www150.statcan.gc.ca/n1/fr/pub/12-001-x/1997001/article/3103-fra.pdf?st=YRiuivNK].

HAZIZA, D., et J.-F. BEAUMONT (2007). "On the Construction of Imputation Classes in Surveys", *International Statistical Review*, [En ligne], vol. 75, n° 1, avril, p. 25-43. [www.jstor.org/stable/41508447].

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. [statistique.quebec.ca/fr/fichier/methodologie-de-lenquete-quebecoise-sur-le-developpement-des-enfants-a-la-maternelle-2017.pdf]

Annexe

Tableau A.1 Variables ayant un taux de non-réponse partielle pondéré supérieur à 5 %, Québec, 2022

Nom de la variable	Libellé : Depuis le début de l'année scolaire, (nom enfant) a-t-il/elle reçu les services d'un professionnel non enseignant à l'école?	Taux de non-réponse partielle pondéré	Effet possible sur l'analyse
H_Q4A	Infirmière	9,2	Négligeable
H_Q4B	Orthophoniste	6,0	Négligeable
H_Q4C	Psychoéducateur	5,8	Négligeable
H_Q4D	Travailleur social	7,4	Négligeable
H_Q4E	Psychologue	8,3	Négligeable
H_Q4G	Hygiéniste dentaire	13,5	Négligeable
H_Q4H	Ergothérapeute	8,6	Négligeable
H_Q4J	Autres <i>Veillez préciser</i>	29,0	Risque de biais important – ne devrait pas être analysée

Source : Institut de la statistique du Québec, volet « enseignant » de l'*Étude sur le développement des enfants en contexte de pandémie*, 2022.