

Chapitre 2

La méthodologie

Carl Drouin

Direction Santé Québec
Institut de la statistique du Québec

Lyne Des Groseilliers

Mario Haché

Direction de la méthodologie et des enquêtes spéciales
Institut de la statistique du Québec

Hélène Desrosiers

Direction Santé Québec
Institut de la statistique du Québec

Louise Bourgon

Département de techniques d'éducation à l'enfance
Collège Edouard-Montpetit

Carole Lavallée

Département de techniques d'éducation à l'enfance
CÉGEP du Vieux Montréal

Les résultats présentés dans ce rapport s'appuient sur un ensemble de procédures ayant permis de recueillir et d'analyser des données sur les services de garde éducatifs à la petite enfance régis au Québec. Rappelons que l'objectif de l'enquête *Grandir en qualité 2003* est de dresser un portrait de la qualité des services offerts dans les centres de la petite enfance (CPE) et les garderies que fréquentaient les enfants québécois en 2003 et de déterminer les facteurs associés à cette qualité. Pour cela, une phase d'observation a eu lieu au printemps 2003 dans les différents types de services visés par l'enquête. Des questionnaires visant à recueillir des renseignements sur les facteurs potentiellement associés à la qualité ont également été administrés aux éducatrices ou aux responsables de services de garde en milieu familial (communément appelées RSG) des groupes observés et aux gestionnaires des établissements concernés.

Ce chapitre présente les méthodes ayant permis d'assurer la qualité et la représentativité des données recueillies. La première section du chapitre précise la méthodologie retenue pour définir la population à l'étude et pour déterminer l'échantillon de l'enquête, soit le plan de sondage, ainsi que le déroulement de

l'enquête. Comme les données recueillies l'ont été essentiellement à partir d'échelles d'observation, la section suivante en décrit le contenu de façon détaillée ainsi que les diverses étapes de construction et de validation. Les questionnaires y sont également présentés de façon plus succincte. La section intitulée « traitement des données » résume les étapes de validation, de pondération des données et présente les taux de réponse ainsi que les méthodes d'analyse utilisées. Pour faciliter la lecture du rapport, la quatrième section se présente sous la forme d'un guide dans lequel sont exposées les diverses règles de présentation des résultats. Finalement, la portée et les limites de l'enquête sont identifiées.

2.1 Méthodes d'enquête

2.1.1 Plan de sondage

2.1.1.1 Population visée

La population visée par la présente enquête est constituée des enfants¹ qui fréquentaient les services de garde éducatifs à la petite enfance régis par l'État québécois² à l'exclusion de ceux qui étaient en opération depuis moins de seize mois environ, à la date

du début de l'enquête, soit au début mars 2003³. Ces services sont composés de ceux coordonnés par les centres de la petite enfance (CPE)⁴ et de ceux offerts par les garderies⁵. La population visée exclut également les enfants qui fréquentent :

- ↓ les services établis dans les réserves indiennes,
- ↓ les services situés aux Îles-de-la-Madeleine et ceux de la région administrative Nord-du-Québec, et
- ↓ les services spécialisés dans l'accueil d'enfants handicapés.

Ces exclusions représentent 10,7 % de tous les enfants fréquentant les services de garde éducatifs régis au Québec. Ainsi, la population visée représente environ 89,3 % de la population totale, soit un peu plus de 145 000 enfants. Le critère d'inadmissibilité basé sur la date du début d'opération constitue le facteur limitant le plus cette couverture.

Il est important de noter que les CPE doivent être titulaires d'un permis et comprendre un volet en installation de CPE et un volet en milieu familial⁶. Un même CPE peut être responsable de plus d'une installation, c'est-à-dire de plus d'une entité physique offrant des services de garde éducatifs dans un lieu strictement réservé à cet effet. Ces installations peuvent offrir des services à des groupes d'enfants de moins de 18 mois (poupons) ou à des groupes composés d'enfants dits d'âge préscolaire, soit à des enfants généralement âgés entre 18 mois et 5 ans. Quant au volet milieu familial, il regroupe, sous la

coordination d'un CPE, tous les services offerts par des responsables d'un service de garde en milieu familial (RSG). Sauf exception, ces services sont offerts aux enfants de la naissance jusqu'à la fréquentation de la maternelle, soit jusqu'à l'âge de 5 ans environ⁷. Cependant, le nombre de RSG peut varier considérablement d'un CPE à l'autre.

Selon ce découpage basé sur les classes d'âge et le volet, un CPE peut offrir jusqu'à trois types de services de garde éducatifs :

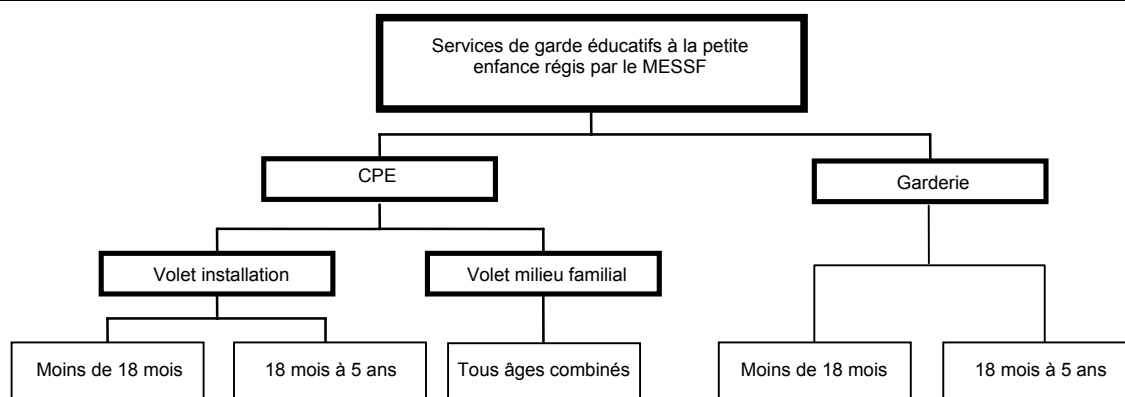
1. En installation accueillant les enfants de moins de 18 mois,
2. En installation accueillant les enfants de 18 mois à 5 ans,
3. En milieu familial.

L'offre de services est complétée par les garderies qui se distinguent par leur statut juridique. Ces dernières offrent les services suivants :

1. Garderie accueillant les enfants de moins de 18 mois,
2. Garderie accueillant les enfants de 18 mois à 5 ans.

Au total, il est donc possible de présenter le réseau de garde de la petite enfance régi au Québec selon cinq différents types de services tel qu'illustré à la figure 2.1. C'est cette division que nous avons retenue dans le cadre de l'enquête *Grandir en qualité 2003*.

Figure 2.1
Types de services de garde éducatifs régis au Québec



Source : Institut de la statistique du Québec, *Grandir en qualité 2003*.

2.1.1.2 Base de sondage

Le fichier *CAFE*⁸ du ministère de l'Emploi, de la Solidarité sociale et de la Famille (MESSF) a été retenu comme base de sondage pour effectuer l'enquête, c'est-à-dire comme base de données permettant la sélection aléatoire d'un échantillon. Ce fichier contient certains renseignements de gestion servant à la délivrance, au renouvellement des permis de services de garde éducatifs ainsi qu'à l'administration des différents programmes de subvention du Ministère. Le fichier contient, entre autres, les noms et adresses des services, le numéro de permis, le nombre de places au permis par installation et garderie (moins de 18 mois et 18 mois à 5 ans), le nombre de places au permis en milieu familial, l'entité juridique (but lucratif, sans but lucratif, coopérative), la date de fondation, etc. La version du fichier mise à jour le 20 février 2002 a été utilisée pour la sélection de l'échantillon des installations de CPE et des garderies. Toutefois, la comparaison des données concernant le nombre de places au permis pour le volet en milieu familial (fichier *CAFE* 2002) avec celles sur le nombre de places occupées (données des rapports d'activités 2000-2001 des centres de la petite enfance) a montré que le nombre de places au permis était supérieur au nombre de places occupées. En conséquence, le fichier des *Rapports d'activités 2000-2001 des centres de la petite enfance* (mis à jour le 31 mars 2001) a été utilisé pour la sélection des services en milieu familial.

2.1.1.3 Stratification et degré d'échantillonnage

Pour répondre aux exigences analytiques de l'enquête, les cinq types de services mentionnés ont été traités comme des échantillons distincts à cause de leurs différences tant sur le plan de l'organisation des services que sur celui des caractéristiques propres aux classes d'âge desservies. C'est aussi en raison de ces différences que les échelles d'observation qui permettent d'évaluer la qualité éducative des services de garde éducatifs diffèrent selon que l'on mesure la qualité en milieu familial, dans les groupes de poupons ou dans les groupes d'enfants âgés de 18 mois à 5 ans des installations ou des garderies. En conséquence, les groupes d'enfants de moins de 18 mois et ceux de

18 mois à 5 ans sont sélectionnés séparément dans les installations et les garderies.

La variable de stratification⁹ utilisée est la région administrative ou un regroupement de régions administratives lorsque le nombre de services dans la région est faible. Ce découpage du Québec permet d'assurer une meilleure représentativité de l'ensemble de la population visée. Le tableau 2.1 présente la répartition des services visés selon le type et la région administrative en date du 20 février 2002.

Le plan de sondage de l'enquête pour chacun des cinq types de services est un plan stratifié proportionnel à deux degrés de grappes. Le premier degré consiste en un échantillon de services stratifié proportionnel au nombre d'enfants pour chacun des cinq types de services. Le deuxième degré est formé d'un groupe d'enfants dans le cas des volets en installation ou en garderie, et de deux RSG pour le milieu familial pour les raisons évoquées plus loin dans le texte. Les groupes ou services en milieu familial sélectionnés constituent ainsi les unités d'observation à l'intérieur desquelles se trouvent des enfants à qui est associée la qualité observée. À titre d'exemple, pour les enfants d'âge préscolaire en garderie, au premier degré, on a sélectionné des garderies et, au deuxième degré, on a choisi un groupe d'enfants parmi les groupes d'âge préscolaire de chaque garderie sélectionnée. Cette méthode permet alors d'inférer la qualité de l'expérience vécue à tous les enfants qui appartiennent à un groupe d'âge préscolaire d'une garderie.

En d'autres termes, nous avons d'abord sélectionné des installations de CPE, des garderies ou des CPE (pour les services en milieu familial). Parmi ces services, nous avons ensuite sélectionné un groupe d'enfants (ou deux services en milieu familial) dans lequel devait se dérouler l'observation. Finalement, il nous est possible d'inférer sur la qualité des services de garde éducatifs dont bénéficient les enfants québécois en considérant *l'enfant* comme unité d'analyse.

Tableau 2.1

Répartition des services de garde éducatifs visés par l'enquête *Grandir en qualité 2003* selon le type de services et la région administrative¹

| Région administrative | CPE | | | Garderie | |
|--|---------------------------|---------------------------|---------------------------------|-----------------------|-----------------------|
| | Installation - 18 mois | Installation 18 mois + | Milieu familial ² | Garderie - 18 mois | Garderie 18 mois + |
| Bas-Saint-Laurent (01) | 13 | 23 | 19 | 0 | 0 |
| Saguenay (02) | 27 | 31 | 26 | 7 | 7 |
| Capitale-Nationale (03) | 59 | 109 | 77 | 16 | 41 |
| Mauricie (04) – Centre-du-Québec (17) ³ | 38 | 66 | 49 | 6 | 7 |
| Estrie (05) | 32 | 57 | 41 | 3 | 4 |
| Montréal (06) | 138 | 267 | 104 | 71 | 222 |
| Outaouais (07) | 30 | 47 | 30 | 2 | 6 |
| Abitibi-Témiscamingue (08) | 9 | 15 | 16 | 1 | 1 |
| Côte-Nord (09) | 9 | 11 | 10 | 0 | 0 |
| Gaspésie (11) | 5 | 8 | 8 | 0 | 1 |
| Chaudière-Appalaches (12) | 37 | 51 | 40 | 7 | 12 |
| Laval (13) | 20 | 34 | 26 | 17 | 37 |
| Lanaudière (14) | 25 | 42 | 35 | 9 | 26 |
| Laurentides (15) | 26 | 44 | 43 | 7 | 18 |
| Montérégie (16) | 121 | 149 | 106 | 35 | 89 |
| Total | 589 | 954 | 630 | 181 | 471 |

1. Un même établissement peut être compté plus d'une fois s'il offre plus d'un type de services.

2. Représente le nombre de CPE ayant un volet « milieu familial » et non pas le nombre total de RSG.

3. Les valeurs de ces deux régions ont été regroupées pour fins de confidentialité.

Sources : MESSF, Ficher *CAFE*, 20 février 2002.

MESSF, *Rapports d'activités 2000-2001 des centres de la petite enfance*.

2.1.1.4 Taille, répartition et sélection de l'échantillon

La taille de l'échantillon a été fixée en vue d'assurer un certain niveau de précision sur les estimations produites, tout en respectant les contraintes budgétaires.

Installation et garderie

Pour les installations et les garderies offrant des services aux enfants de moins de 18 mois, des échantillons de 125 installations et de 125 garderies ont été jugés nécessaires pour assurer une marge d'erreur de 9 % pour une proportion de l'ordre de 30 % associée à un niveau de confiance de 95 % pour des statistiques applicables à l'ensemble du Québec.

Pour les installations et les garderies offrant des services aux enfants de 18 mois à 5 ans, des échantillons de 225 installations et de 225 garderies ont été jugés nécessaires pour assurer une marge d'erreur de 5 % pour une proportion de l'ordre de 14 % associée

à un niveau de confiance de 95 % pour des statistiques applicables à l'ensemble du Québec. Comparativement aux échantillons pour les groupes de poupons, le nombre d'unités est plus élevé pour cette classe d'âge afin d'offrir un plus grand potentiel d'analyse, c'est-à-dire pour permettre la production d'indices de qualité selon deux sous-classes d'âge chez les enfants de 18 mois à 5 ans dans le cadre de recherches ultérieures.

Comme mentionné précédemment, une stratification proportionnelle a été privilégiée par région administrative ou regroupement de régions administratives. Au premier degré, un tirage aléatoire de services proportionnel au nombre d'enfants a été utilisé. Au deuxième degré, un seul groupe d'enfants a ensuite été sélectionné aléatoirement parmi chaque installation et chaque garderie sélectionnées; cette approche était justifiée car beaucoup d'éléments observés concernent tous les enfants d'un même groupe d'âge pour un milieu de garde donné. Ainsi, selon l'hypothèse d'une faible variation sur le plan des divers aspects évalués

entre les groupes d'enfants d'un même type de services et appartenant à une même installation ou garderie, un seul groupe a été sélectionné.

Milieu familial

Pour les services offerts en milieu familial, tous âges combinés, un échantillon de 200 RSG a été évalué nécessaire pour assurer une marge d'erreur de 6,5 % pour une proportion de l'ordre de 22 % associée à un niveau de confiance de 95 % pour des statistiques applicables à l'ensemble du Québec.

Comme pour les autres types de services, une stratification proportionnelle a été privilégiée par région administrative ou regroupement de régions administratives. Au premier degré, un tirage aléatoire proportionnel au nombre de RSG a été utilisé pour sélectionner 100 CPE offrant des services en milieu familial. Au deuxième degré, deux RSG par CPE choisis ont été sélectionnés aléatoirement car il y a lieu de croire que, comparativement aux groupes appartenant à une même installation, les éléments observés sont plus variables entre les RSG rattachées à un même CPE. Ce choix se fonde notamment sur le fait que les RSG jouissent d'un niveau d'autonomie élevé par rapport au CPE auquel elles sont rattachées et, qu'à ce titre, elles sont individuellement responsables de la qualité offerte dans leur service.

En conséquence, pour répondre aux exigences de l'enquête, la taille initiale de l'échantillon total a été fixée à 350 services en installation, 350 services en garderie et 200 responsables de service de garde en milieu familial (RSG). Aussi, selon cette méthode de sélection par type de services, plusieurs groupes d'enfants d'un même CPE pouvaient être sélectionnés selon que celui-ci offrait des services aux poupons, aux enfants d'âge préscolaire ou en milieu familial. De même, les garderies pouvaient être choisies au hasard pour les deux types de service qu'elles offrent.

Toutefois, le déroulement de l'enquête a dû être reporté de la mi-octobre 2002 au début mars 2003. Étant donné que la sélection de l'échantillon avait déjà été faite à la fin février 2002, une vérification a été effectuée afin de s'assurer que la population visée n'avait subi que peu de modifications entre février 2002 et mars 2003. Cette analyse comparative a permis de détecter que, parmi les établissements sélectionnés, cinq garderies ont été transformées en CPE et que deux garderies ont cessé leurs opérations. Dans l'ensemble, ces changements ne sont pas majeurs; ils représentent un peu plus de 1 % des garderies sélectionnées. Par ailleurs, afin de s'assurer d'avoir un nombre suffisant de garderies dans l'échantillon, sept garderies ont été ajoutées de la façon suivante :

- ↓ 3 garderies accueillant les enfants de moins de 18 mois;
- ↓ 4 garderies accueillant les enfants de 18 mois à 5 ans.

De plus, au cours de la collecte, une autre garderie a été transformée en CPE.

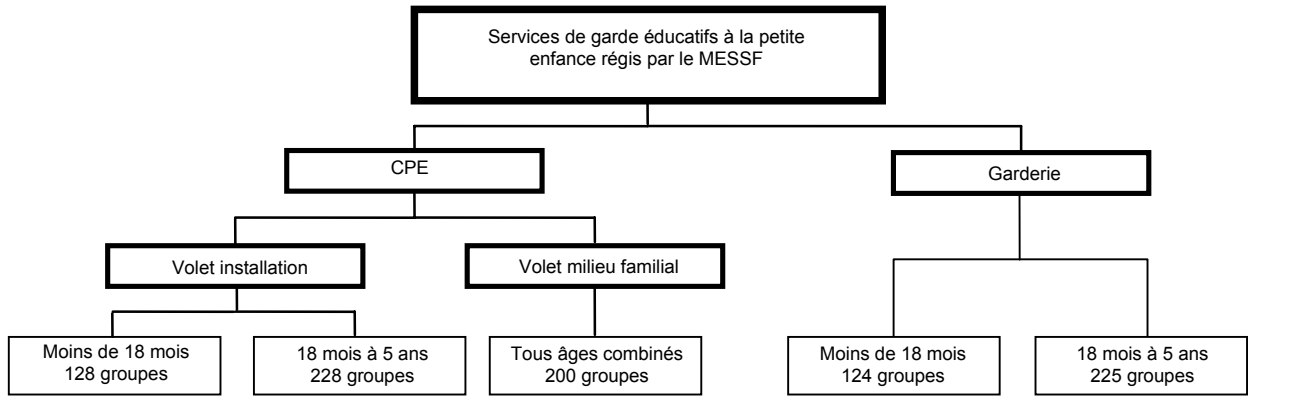
Ces divers changements expliquent l'écart entre la taille de l'échantillon initiale (900 services) et le nombre de services sélectionnés pour l'enquête (905 services) (figure 2.2).

On anticipait un effet de plan¹⁰ faible étant donné qu'une stratification proportionnelle avait été privilégiée par région administrative ou regroupement de régions, et ce, pour chaque type de service.

Pour l'ensemble des services, on supposait un taux de réponse global de 80 %. Le tableau 2.2 présente la répartition de l'échantillon par région administrative.

Figure 2.2

Nombre de services sélectionnés dans l'échantillon de l'enquête *Grandir en qualité 2003* selon le type de services



Source : Institut de la statistique du Québec, *Grandir en qualité 2003*.

Tableau 2.2

Répartition de l'échantillon de l'enquête *Grandir en qualité 2003* selon le type de services et la région administrative¹

| Région administrative | CPE | | | Garderie | | Total |
|--|---------------------------|---------------------------|---------------------------------|-----------------------|-----------------------|------------|
| | Installation - 18 mois | Installation 18 mois + | Milieu familial ² | Garderie - 18 mois | Garderie 18 mois + | |
| Bas-Saint-Laurent (01) | 2 | 4 | 8 | 0 | 0 | 14 |
| Saguenay (02) | 5 | 8 | 10 | 3 | 3 | 29 |
| Capitale-Nationale (03) | 13 | 25 | 22 | 11 | 16 | 87 |
| Mauricie (04) – Centre-du-Québec (17) ³ | 8 | 14 | 16 | 3 | 4 | 45 |
| Estrie (05) | 6 | 12 | 12 | 2 | 2 | 34 |
| Montréal (06) | 35 | 69 | 20 | 54 | 111 | 289 |
| Outaouais (07) | 8 | 11 | 14 | 1 | 3 | 37 |
| Abitibi-Témiscamingue (08) | 2 | 4 | 4 | 0 | 0 | 10 |
| Côte-Nord (09) | 1 | 2 | 6 | 0 | 0 | 9 |
| Gaspésie (11) | 1 | 2 | 4 | 0 | 0 | 7 |
| Chaudière-Appalaches (12) | 8 | 10 | 14 | 5 | 4 | 41 |
| Laval (13) | 4 | 8 | 10 | 13 | 18 | 53 |
| Lanaudière (14) | 5 | 10 | 12 | 4 | 11 | 42 |
| Laurentides (15) | 6 | 10 | 12 | 5 | 9 | 42 |
| Montérégie (16) | 24 | 39 | 36 | 23 | 44 | 166 |
| Total | 128 | 228 | 200 | 124 | 225 | 905 |

1. Rappelons que la stratification de l'échantillon par service de garde éducatif peut être composée d'une région administrative ou d'un regroupement de régions administratives.

2. Représente le nombre de CPE ayant un volet « milieu familial » et non pas le nombre total de RSG.

3. Les valeurs de ces deux régions ont été regroupées pour fins de confidentialité.

Source : Institut de la statistique du Québec, *Grandir en qualité 2003*.

2.1.2 Déroulement de l'enquête

La collecte de données de l'enquête *Grandir en qualité 2003* repose sur deux types d'instruments : d'une part, des échelles d'observation permettant d'évaluer la qualité éducative dans les différents types de services, et, d'autre part, des questionnaires visant à recueillir de l'information sur les caractéristiques des éducatrices, des RSG, des gestionnaires responsables des groupes sélectionnés ainsi que des établissements offrant les services.

L'enquête s'est déroulée dans un laps de temps le plus court possible afin d'éviter des distorsions pouvant être engendrées au fil du temps. La phase d'observation dans les groupes sélectionnés, complétée par l'administration des questionnaires aux éducatrices, le jour même de l'observation, a été réalisée du 3 mars au 5 juin 2003 inclusivement. Alors que les questionnaires destinés aux éducatrices étaient autoadministrés, ceux destinés aux RSG et aux gestionnaires ont été administrés par téléphone par des intervieweurs de l'ISQ dans les jours suivant l'observation, de sorte que leur administration s'est terminée dans la semaine suivant la fin des dernières observations. Ces questionnaires sont décrits à la section 2.2.7.

2.1.2.1 Procédure de contact et de sélection des groupes

Une campagne de sensibilisation a d'abord été menée auprès des principaux regroupements de services de garde éducatifs pour s'assurer de leur collaboration et de celle de leurs membres. Par la suite, les services admissibles de la province ont été informés par écrit de la tenue de l'enquête et de la possibilité qu'ils soient sélectionnés pour y participer.

Dans un troisième temps, des contacts ont été établis plus spécifiquement auprès des établissements sélectionnés au hasard pour participer à l'enquête, selon les modalités du plan de sondage. Ainsi, des lettres sollicitant le consentement ainsi que des documents expliquant la nature de la contribution attendue ont été expédiés aux conseils d'administration, directrices ou propriétaires des établissements sélectionnés.

Par la suite, les gestionnaires de ces mêmes établissements ont été joints par téléphone de façon à procéder à la sélection aléatoire des groupes d'enfants devant être observés. Dans le cas des services en milieu familial, les RSG étaient elles-mêmes contactées pour l'obtention du consentement. L'entretien téléphonique avec les gestionnaires ou les RSG permettait également de fixer le rendez-vous pour la journée d'observation. Des renseignements complémentaires étaient ensuite transmis par courrier aux différentes personnes concernées directement ou indirectement par la collecte de données (incluant des notes aux gestionnaires, aux éducatrices, aux RSG et aux parents). Finalement, les rendez-vous ont été confirmés par téléphone dans les jours précédant la visite. Les téléphonistes responsables de la sélection des groupes observés et de la prise de rendez-vous avaient été formées et familiarisées au préalable avec les grandes étapes du déroulement de l'enquête de façon à pouvoir informer adéquatement les répondants et à répondre aux questions qui leur étaient adressées.

2.1.2.2 Journée d'observation

L'observation à l'aide d'une échelle consiste à recueillir de l'information sur la qualité de l'expérience vécue d'un groupe d'enfants au cours d'une journée dans un service de garde éducatif. Les échelles d'observation abordent donc toutes les expériences qu'un enfant est susceptible de vivre, au cours d'une journée dans un tel service, à partir de son arrivée jusqu'à son départ.

L'observation couvre presque toutes les activités de la journée. Un protocole d'observation accompagnant les échelles a été développé par les auteures pour encadrer et systématiser le travail des observatrices¹¹. L'observatrice se présente au service suffisamment tôt pour observer les activités d'accueil et ne quitte qu'après le départ d'au moins trois enfants du groupe.

Pour que l'observation puisse être effectuée, il devait s'agir d'une journée « normale » où les enfants se retrouvent la plupart du temps dans leur environnement régulier, c'est-à-dire dans les locaux ou dans l'aire de jeu extérieure du milieu. Le programme régulier d'activités devait être respecté : il ne devait pas y avoir, à titre d'exemple, d'activités spéciales ou de sorties

éducatives. De plus, pour qu'une journée soit considérée « normale », le nombre d'enfants présents était aussi pris en considération : la présence de plus de la moitié des enfants du groupe constituait une exigence préalable à l'observation. Dans les cas où une maladie ou toute autre situation faisait en sorte que plus de la moitié des enfants du groupe étaient absents, l'observation était reportée.

Soulignons également que pour compléter l'échelle d'évaluation, la journée d'observation impliquait également une courte entrevue réalisée avec l'éducatrice ou la RSG du groupe sélectionné. L'entrevue, d'une durée approximative de vingt minutes, a servi à recueillir de l'information sur des thèmes difficiles ou impossibles à observer au cours d'une seule journée, comme les items qui concernent les interactions entre l'éducatrice ou la RSG et les parents.

2.2 Instruments d'enquête

2.2.1 Présentation et contenu des échelles d'observation de la qualité éducative

Pour atteindre les buts visés par l'enquête, l'information recueillie doit être organisée de façon à déterminer des degrés de qualité en référence aux observations réalisées. La mesure de la qualité demande aussi que l'on définisse avec précision le concept de la qualité recherchée dans le contexte des services de garde éducatifs québécois. Les principes théoriques et pratiques relatifs à cette définition de la qualité ont d'ailleurs été exposés au chapitre 1.

Une recension des recherches menées sur la qualité des services de garde éducatifs révèle également que la technique d'observation directe est la plus couramment utilisée pour évaluer la qualité de ces services (Bellemare, Gravel et Briand, 2002; Tremblay, 2003). Cette méthode permet à un ou à plusieurs observateurs neutres de recueillir des renseignements multiples sur une situation observable. Pour que plusieurs observations servent à des fins d'enquête statistique, elles doivent pouvoir être organisées dans une échelle de mesure de la qualité dont les données seront compilées et analysées. Plus spécifiquement, l'instrument de mesure élaboré dans le

cadre du présent projet devait répondre à l'ensemble des critères suivants :

- ↓ utiliser la technique d'observation directe;
- ↓ permettre de mesurer la qualité éducative de façon à ce qu'elle soit représentative de l'expérience quotidienne vécue par les enfants;
- ↓ permettre de mesurer la qualité selon les domaines d'application du programme éducatif (soit la structuration des lieux, la structuration des activités et l'intervention du personnel éducateur auprès des enfants et des parents) et selon certaines activités de base;
- ↓ assurer des mesures de qualité adaptées selon les différentes classes d'âge desservies;
- ↓ être adapté à l'organisation des services de garde éducatifs québécois, à la réglementation ainsi qu'à l'approche éducative privilégiée par le ministère de l'Emploi, de la Solidarité sociale et de la Famille et le réseau de services de garde éducatifs en général.

Trois échelles d'observation adaptées aux particularités des différents types de services observés ont ainsi été créées pour répondre aux critères susmentionnés (Bourgon, Lavallée et collaborateurs, 2003a, 2003b, 2003c)¹². Il s'agit de :

- ↓ Échelle d'observation de la qualité éducative : le service de garde préscolaire
- ↓ Échelle d'observation de la qualité éducative : le service de garde en milieu familial
- ↓ Échelle d'observation de la qualité éducative : le service de garde en pouponnière

Ces trois échelles ont une base commune. Elles s'apparentent de très près aux trois domaines d'application définis par le programme éducatif en les subdivisant sous la forme de quatre dimensions de la qualité éducative : i) la structuration des lieux; ii) la structuration et la variation de type d'activités; iii) l'interaction de l'éducatrice ou de la RSG avec les

enfants; iv) l'interaction de l'éducatrice ou de la RSG avec les parents. L'organisation même des échelles suit cette division puisque chaque item est classé dans l'une ou l'autre de ces dimensions. De façon plus détaillée, les quatre dimensions traitent des sujets suivants :

2.2.1.1 La structuration des lieux

La structuration des lieux est observée pour rendre compte de l'aménagement du service de garde éducatif et recenser le matériel mis à la disposition des enfants. Les lieux doivent répondre aux besoins et aux intérêts de tous (enfants, membres du personnel, RSG et parents) tout en favorisant le développement global des enfants.

2.2.1.2 La structuration et la variation des types d'activités

La deuxième dimension, la qualité de la structuration et de la variation des types d'activités, porte sur la planification des activités et l'observation des enfants par l'éducatrice ou la RSG, sur l'horaire de la journée et sur les activités ludiques. Encore une fois, les critères de qualité privilégiés sont tirés du programme éducatif, à savoir : la régularité de l'horaire de base, la possibilité pour les enfants de déterminer eux-mêmes leurs activités et de faire des choix ainsi que la variation des modes d'encadrement des enfants.

2.2.1.3 L'interaction de l'éducatrice ou de la RSG avec les enfants

La dimension qui traite de l'intervention de l'éducatrice ou de la RSG auprès des enfants tient compte de l'approche préconisée par le programme éducatif laquelle valorise l'intervention démocratique et souligne l'importance du jeu dans l'apprentissage chez l'enfant ainsi que l'importance d'une saine communication et de relations interpersonnelles positives.

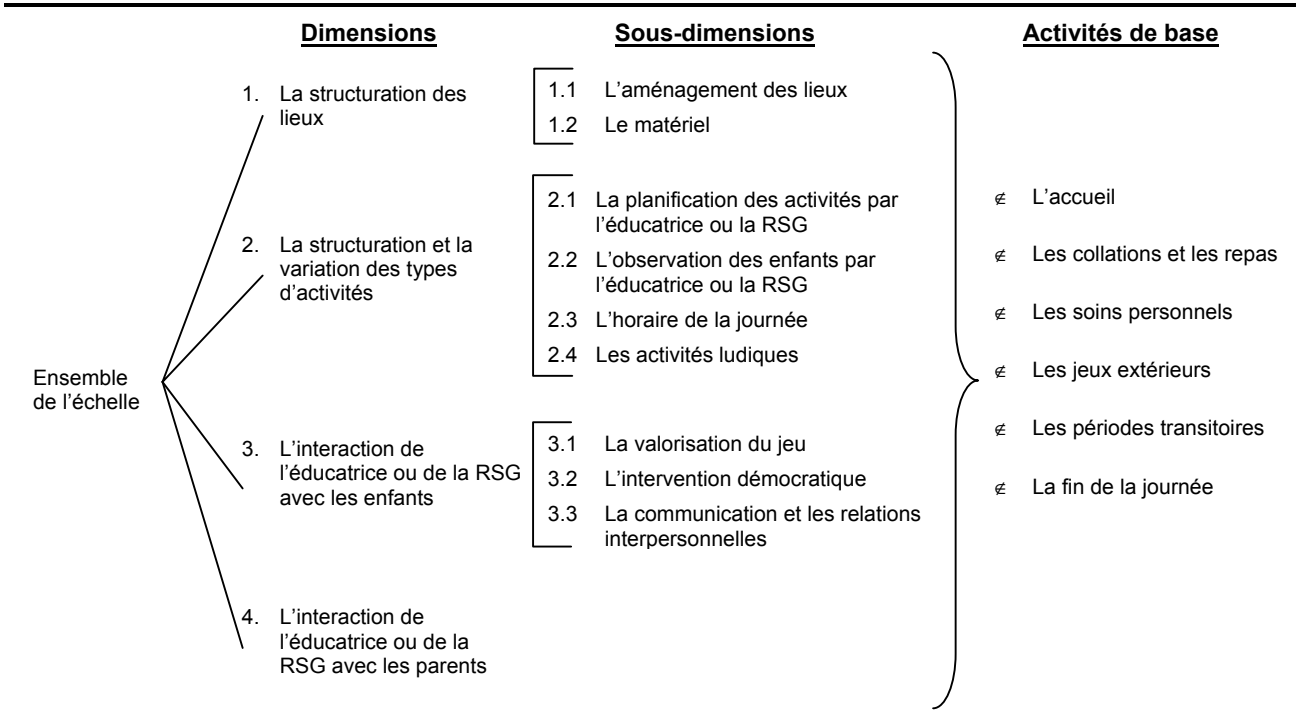
2.2.1.4 L'interaction de l'éducatrice ou de la RSG avec les parents

La quatrième dimension concerne la qualité des relations de l'éducatrice ou de la RSG avec les parents. Cette dernière dimension, qui porte sur la collaboration entre le personnel du service de garde éducatif et les parents, est mesurée notamment à partir de l'observation de la période d'arrivée et de celle du départ des enfants.

Ainsi, comme illustré à la figure 2.3, chacune des trois échelles d'observation de la qualité éducative inclut un ensemble d'items regroupés non seulement dans l'une ou l'autre des quatre dimensions de la qualité éducative, mais également selon neuf sous-dimensions. De plus, une section complémentaire des échelles permet d'évaluer la qualité selon six activités de base, soit des périodes habituelles de la journée ou des activités routinières. Plus précisément, l'accueil, les repas, les soins personnels, les périodes transitoires, les jeux extérieurs et la fin de la journée sont les activités de base observées. En fait, chaque item de cette section fait également partie intégrante de l'une ou l'autre des dimensions précédentes de sorte que ces items servent à la fois à mesurer la qualité des activités de base et la qualité par dimension et sous-dimension.

Figure 2.3

Organisation et contenu des trois échelles d'observation de la qualité éducative de l'enquête *Grandir en qualité 2003*



Source : Institut de la statistique du Québec, *Grandir en qualité 2003*.

2.2.2 Entrevue en complément à l'observation

Certains éléments jugés essentiels à la mesure de la qualité étant difficilement observables, une entrevue dirigée d'environ vingt minutes avec l'éducatrice responsable du groupe ou la RSG a été réalisée pour compléter les échelles d'observation. Il s'agit, par exemple, d'aspects liés à la planification quotidienne des activités par les éducatrices ou les RSG et aux interactions entre ces dernières et les parents. L'entrevue est effectuée par l'observatrice lors de la journée d'observation. Quatre thèmes y sont abordés : i) la communication avec les parents, ii) l'observation des enfants, iii) la planification des activités et iv) les sorties à l'extérieur du service. Après l'entrevue, l'observatrice doit cocher dans l'échelle les éléments qui correspondent aux réponses fournies par l'éducatrice ou la RSG.

2.2.3 Fonctionnement des échelles

Chaque section des échelles est composée d'items comprenant un énoncé et des éléments à cocher. L'énoncé est formulé comme un objectif à atteindre. Suivant l'exemple de la figure 2.4, l'énoncé de l'item 3.1.2 stipule que « l'éducatrice soutient les initiatives personnelles des enfants dans leurs jeux ». Chaque énoncé est suivi d'une liste à cocher composée d'éléments observables. Dans le présent exemple, il s'agit de stratégies pouvant être employées par l'éducatrice pour soutenir les initiatives des enfants. Il peut s'agir également des caractéristiques des lieux, d'éléments du matériel, de types d'activités, etc. Ce sont ces éléments qui décrivent des conditions, des objets ou des comportements observés tout au long de la journée.

Figure 2.4

Exemple d'item mesuré par l'enquête *Grandir en qualité 2003* : l'item 3.1.2 de l'échelle préscolaire

| 3.1.2 L'éducatrice soutient les initiatives personnelles des enfants dans leurs jeux | |
|---|--|
| Stratégies pour soutenir les initiatives personnelles des enfants : | |
| L'éducatrice : | |
| <input type="checkbox"/> | 1. Met en œuvre les suggestions des enfants pour faire évoluer le jeu |
| <input type="checkbox"/> | 2. Invite les enfants à parler de ce qu'ils font et des projets qu'ils imaginent |
| <input type="checkbox"/> | 3. Permet aux enfants de prendre des risques calculés, c'est-à-dire qui ne mettent pas en cause leur sécurité |
| <input type="checkbox"/> | 4. Incite les enfants à innover et à tester leurs idées |
| <input type="checkbox"/> | 5. Commente de façon précise et concrète les réussites des enfants |
| <input type="checkbox"/> | 6. Décrit de façon précise et concrète les difficultés rencontrées par les enfants |
| <input type="checkbox"/> | 7. Si nécessaire, fait des suggestions pour poursuivre le jeu, sans insister pour que les enfants adoptent celles-ci |
| <input type="checkbox"/> | 8. Imite ce que font les enfants pour s'intégrer à leur jeu |
| <input type="checkbox"/> | 9. Joue le rôle que les enfants lui proposent dans leurs jeux |
| <input type="checkbox"/> | 10. Demande aux enfants de décrire les « règles » des jeux qu'ils inventent |
| <input type="checkbox"/> | 11. Suit les indices fournis par les enfants sur le contenu du jeu ou la direction à donner au jeu |
| <input type="checkbox"/> | 12. Fournit le matériel nécessaire pour enrichir et soutenir le jeu, selon la direction que les enfants lui donnent |
| .../12 | |

Source : Bourgon, Lavallée et collaborateurs, 2003a.

Essentiellement, le travail de l'observatrice consistait donc à relever la présence des éléments des listes à cocher pour chaque item. Comparativement à d'autres échelles de mesure de la qualité (par exemple, l'*Échelle d'évaluation de l'environnement préscolaire - EEEP-R*, *CaMIE* et *OLiVE*, *High/Scope Program Quality Assessment* et *Kaléidoscope*), ce travail était réalisé sans égard à l'évaluation de la qualité puisque l'observatrice ignorait, sauf exception, les scores de qualité induits par la présence d'éléments cochés. En fait, les résultats rapportés par l'observatrice étaient par la suite compilés par un programme informatique qui

calculait le niveau de qualité selon des critères prédéterminés par les auteures de l'échelle. Cette façon de faire, basée sur l'observation factuelle de l'environnement de garde, allégeait la tâche de l'observatrice et permettait une plus grande objectivité de sa part.

Le score de qualité ainsi déterminé est fonction d'une échelle de qualité à quatre niveaux où la qualité d'un item fait référence spécifiquement à une cotation comme illustré à la figure 2.5. Ces quatre niveaux font référence consécutivement à un niveau de qualité :

Figure 2.5

Exemple de cotation d'un item des échelles d'observation, enquête *Grandir en qualité 2003*

| 3.1.2 L'éducatrice soutient les initiatives personnelles des enfants dans leurs jeux | | | |
|---|---|--|---|
| 1 - Niveau inadéquat | 2 - Niveau minimal | 3 - Niveau bien | 4 - Niveau très bien |
| L'éducatrice ne soutient pas les initiatives personnelles des enfants. | L'éducatrice soutient peu les initiatives personnelles des enfants. | L'éducatrice soutient bien les initiatives personnelles des enfants. | L'éducatrice soutient très bien les initiatives personnelles des enfants. |
| 2 éléments ou moins sont cochés. | De 3 à 5 éléments sont cochés. | De 6 à 8 éléments sont cochés. | 9 éléments ou plus sont cochés. |

Source : Bourgon, Lavallée et collaborateurs, 2003a.

- ↓ **Une cote « 1 » - inadéquat** - est attribuée lorsque l'observation démontre la présence d'aucun ou de très peu d'éléments favorisant la qualité éducative. Elle est attribuée aussi quand on constate la présence d'au moins un élément mettant en cause la sécurité, la santé ou l'intégrité des enfants ou encore l'absence d'un ou de plusieurs éléments les garantissant.
- ↓ **Une cote « 2 » – minimal** – est attribuée lorsque les conditions de base pour assurer la santé, la sécurité et l'intégrité des enfants sont présentes mais que peu d'éléments favorisant la qualité éducative ont été observés.
- ↓ **Une cote « 3 » – bien** – est attribuée lorsqu'un nombre assez satisfaisant d'éléments favorisant la qualité éducative ont été observés.
- ↓ **Finalement, une cote « 4 » – très bien** – est attribuée lorsqu'un nombre très satisfaisant d'éléments favorisant la qualité éducative ont été observés.

Bien que le nombre d'éléments observables varie d'un item à l'autre, le nombre d'éléments à cocher pour atteindre un niveau de qualité n'est pas déterminé en fonction d'une proportion constante d'un item à l'autre¹³. En effet, des critères qualitatifs reliés à différents niveaux de l'échelle ont été pris en considération pour chacun des items. Par ailleurs, certaines listes regroupent des éléments à proscrire, à éviter ou obligatoires¹⁴. Dans les deux premiers cas, la présence de conditions physiques inadéquates, de comportements répressifs ou qui démontrent une négligence entraîne un classement au niveau inadéquat, peu importe le nombre d'éléments positifs observés. Par exemple, une ouverture dans la clôture permettant aux enfants de sortir de la cour extérieure est un élément à proscrire dans l'évaluation de l'aménagement de la cour. Pour les éléments obligatoires, leur poids relatif est plus important comparativement aux autres éléments de l'item. Par exemple, ils peuvent être requis pour atteindre les niveaux bien ou très bien ou encore pour obtenir une cote supérieure à 1.

2.2.4 Processus d'élaboration et de révision des échelles

2.2.4.1 Production de la version préliminaire

La production des échelles d'observation relève du MESSF bien que l'ISQ ait agi à titre de conseiller dans les dernières étapes du processus d'élaboration des instruments. Au départ, le Ministère a mis sur pied un comité d'élaboration des instruments de l'enquête auquel siégeaient des professionnels du MESSF et des spécialistes externes de la petite enfance. Deux de ces spécialistes ont par la suite reçu le mandat de rédiger le contenu des échelles d'observation. Pour appuyer les auteurs, le Ministère a procédé à un exercice de validation en faisant appel à des experts en mesures et évaluation et à d'autres spécialistes du domaine des services de garde éducatifs à l'enfance. Ces derniers ont été appelés à commenter et à juger de la pertinence, de l'exhaustivité, de la compréhension et de la cohérence des énoncés de chacune des trois échelles.

Comme l'enquête était encadrée par un comité d'orientation sous la responsabilité de l'ISQ, les échelles d'observation ont également été revues par les membres de ce comité. Ces derniers, en provenance de l'ISQ, du Ministère, du milieu universitaire et de la recherche, du domaine de la formation du personnel éducateur ainsi que des principaux regroupements de services de garde éducatifs au Québec ont proposé plusieurs modifications. C'est à la suite de ces étapes initiales qu'une version préliminaire a pu être expérimentée dans le cadre d'un prétest.

2.2.4.2 Réalisation du prétest

Comme il est d'usage dans le cadre d'enquêtes de cette envergure, l'enquête principale a été précédée d'un prétest. Celui-ci a été réalisé en mai 2002 auprès de 87 groupes d'enfants appartenant aux différents types de services visés dans les régions de Montréal et Québec. Les procédures de contacts, l'observation et l'administration des questionnaires ont été faites dans des conditions similaires à celles alors prévues pour l'enquête. Cet exercice de préparation à l'enquête avait un double objectif : d'une part, il visait un ensemble

d'objectifs liés à la méthodologie de l'enquête et aux pratiques de collecte, soit,

- ↓ obtenir des estimations plus précises des taux de réponse attendus pour chaque type de services visé,
- ↓ mettre à l'épreuve le contenu et la forme de la formation offerte aux observatrices responsables de remplir les échelles d'observation,
- ↓ valider les méthodes de contact et d'obtention de consentement avec les différents types de services et intervenants,
- ↓ tester le protocole d'observation,
- ↓ optimiser l'ensemble de la procédure de terrain.

D'autre part, compte tenu de l'utilisation de nouveaux instruments, le prétest devait également servir à valider le contenu des échelles, à uniformiser leur interprétation et, donc, à orienter les ajustements à apporter avant le début de l'enquête. Les échelles ont donc été soumises à un ensemble de mesures et d'inspections, sous la supervision d'un conseiller en mesures et évaluation. Plus précisément, les données recueillies lors du prétest ont permis i) un examen de la fidélité inter-juges entre les juges expertes et les observatrices du prétest; ii) une analyse de la consistance interne des dimensions et des items des échelles; et iii) une validation concurrente avec une autre échelle de mesure de la qualité, l'*Échelle d'évaluation de l'environnement préscolaire (EEEP-R)*¹⁵ (Harms, Clifford et Cryer, 1998).

De façon générale, les résultats de ces tests se sont avérés satisfaisants au regard des objectifs d'un prétest¹⁶. En résumé, l'analyse des accords inter-juges aura permis notamment d'établir des critères plus sévères quant aux seuils d'accord minimums à atteindre dans le cadre de l'enquête. L'analyse de la consistance interne aura révélé des niveaux de consistance satisfaisants dans l'ensemble et selon les dimensions en plus d'aider à cerner des besoins d'ajustements à certaines dimensions ou certains items des échelles.

Quant à l'exercice de validation concurrente, il aura révélé une corrélation de 0,36 entre l'*EEEP-R* et l'*Échelle d'observation de la qualité éducative : le service de garde préscolaire* (version utilisée au prétest). Ce degré de corrélation s'est avéré conforme aux attentes initiales et s'explique notamment par le fait que les échelles employées dans la présente enquête visent à mesurer une qualité éducative qui correspond de près aux principes du programme éducatif en vigueur au Québec. En ce sens, l'échelle de l'enquête met l'accent sur certains aspects qui sont peu ou pas couverts par l'*EEEP-R* (ex. : la possibilité pour les enfants de faire des choix d'ateliers ou d'activités) ou, à l'inverse, qui ne considèrent pas des critères pris en compte dans l'*EEEP-R* (ex. : la fréquence d'activités centrées sur des didactiques précises). En deuxième lieu, les différences entre les deux instruments comparés peuvent s'expliquer par la méthode de mesure propre à chacune des échelles. Comme mentionné précédemment, la méthode de mesure employée dans l'échelle préscolaire vise à faire l'inventaire des conditions, des objets ou des comportements observés tout au long de la journée. Sauf, pour certains éléments obligatoires ou à proscrire, cet inventaire est fait sans accorder d'ordre de priorité à chacun des éléments observés. Contrairement à cette méthode, celle de l'*EEEP-R* est une échelle hiérarchique et cumulative qui accorde un ordre d'importance à tous les éléments d'un même item en les classant de 1 à 7. En conséquence, lorsque l'ensemble des conditions des niveaux inférieurs n'ont pas été rencontrées, l'évaluation faite à l'aide de cet outil ne considère pas les éléments plus positifs des niveaux supérieurs qui auraient pu être observés.

Par ailleurs, ces étapes de préparation à l'enquête ont été très utiles au raffinement du protocole d'observation, c'est-à-dire à la description avec grande précision de l'ensemble des procédures à suivre lors de l'observation dans les services de garde éducatifs. Les échelles d'observation ont fait l'objet de certains changements au terme du prétest et des étapes de validation subséquentes, de sorte que des exercices d'accords inter-juges entre les expertes ont dû être réalisés, avant la formation des observatrices qui a précédé l'enquête, à partir des versions modifiées des échelles préscolaire et pouponnière. Les résultats de

ces accords ont clairement démontré la cohérence dans l'utilisation de ces deux échelles puisque les coefficients de fidélité inter-juges *kappas pondérés* ont été respectivement de 0,92 pour l'échelle préscolaire et de 0,87 pour l'échelle pouponnière¹⁷. Sur la base du nombre total d'items, ces exercices ont indiqué que les expertes ont désigné des cotes identiques dans 88 % des cas (pourcentage d'accord parfait) et à un point près dans 100 % des cas (pourcentage d'accord proche) pour l'échelle préscolaire. Pour l'échelle pouponnière, c'est 80 % des items que les juges ont coté de façon identique et 99 % à un point près. À cette étape, il n'a pas été jugé nécessaire d'effectuer un test d'accord en milieu familial, puisque l'échelle d'observation pour ce type de services renferme généralement des items dont les concepts sont couverts dans l'une ou l'autre des deux autres échelles.

2.2.5 Formation des observatrices et fidélité inter-juges

La formation des observatrices de l'enquête *Grandir en qualité 2003* s'est déroulée sur une période de dix jours, du 17 au 28 février 2003, soit seulement quelques jours avant le début des observations. Cette formation, sous la responsabilité des auteures des échelles, visait essentiellement à assurer la compréhension et la maîtrise de chacune des trois échelles d'observation par les observatrices. Cet objectif central devait toutefois être atteint parallèlement avec l'obtention d'un degré élevé d'uniformité dans l'utilisation des échelles et dans l'interprétation des éléments observables. Des lectures préparatoires, des exposés théoriques, des exercices pratiques et des exercices d'accords inter-juges ont été réalisés au cours de la formation pour s'assurer de l'atteinte de cette uniformité.

Soulignons également que les observatrices sélectionnées pour prendre part à l'enquête l'ont été sur la base de leur formation et de leur expérience de travail. Celles-ci possédaient généralement au minimum une formation universitaire de premier cycle en éducation préscolaire, en psychologie, en psycho-éducation, en enfance inadaptée, en éducation spécialisée, en orthopédagogie ou en adaptation préscolaire. Les entrevues de sélection ont aussi

permis de cibler les candidates selon des critères liés à la capacité d'observation et à la connaissance pratique de la petite enfance.

2.2.5.1 Résultats d'accords inter-juges

Lors de la formation, les résultats obtenus à partir des exercices d'accords inter-juges ont établi à la fois la fidélité des instruments d'observation et le degré élevé d'uniformité atteint par les 28 observatrices de l'enquête dans l'utilisation de ceux-ci. Dans un premier temps, les degrés d'accord ont été mesurés à partir de deux exercices d'observation survenus les septième et dixième jours de la formation dans des laboratoires d'observation recevant des enfants de 18 mois à 5 ans (échelle préscolaire); ces services disposaient de vitres sans tain permettant l'observation et l'écoute simultanées de la part des 28 observatrices. Pour chaque observation, les cotes de chacune des observatrices ont été comparées avec celles d'une échelle maîtresse (remplie par les formatrices) afin de calculer des mesures de fidélité. L'atteinte d'un coefficient *kappa pondéré* de 0,50 au dernier exercice d'accord était jugée nécessaire pour qu'une observatrice soit retenue pour faire partie de l'enquête. De façon générale, il était toutefois souhaité que la majorité des coefficients se situent au-delà de 0,60, soit le seuil minimal pour que l'accord soit considéré élevé selon la grille d'interprétation de Landis et Koch (1977). Le sommaire de ces résultats, présenté au tableau 2.3, démontre l'atteinte de ces objectifs de même que la progression obtenue entre le premier et le second test d'accord. Soulignons, de plus, que 27 des 28 observatrices ont obtenu un *kappa pondéré* supérieur à 0,70 lors du deuxième exercice.

Malgré que la formation ait abordé en parts égales le contenu des échelles en pouponnière et en milieu familial lors des exposés théoriques et à l'aide d'exercices pratiques et de validation, les conditions de l'enquête rendaient difficile l'évaluation des degrés d'accord pour ces deux échelles. Les problèmes de faisabilité relèvent notamment de l'inexistence de laboratoires d'observation pour ces deux types de milieux, du fait que seules les deux formations pouvaient être désignées « juges expertes » et que ces services se prêtent généralement moins bien à la présence d'au moins deux personnes supplémentaires dans le local d'appartenance.

Tableau 2.3

Sommaire des accords inter-juges pour l'échelle préscolaire lors de la formation des observatrices pour l'enquête *Grandir en qualité 2003*

| | Nombre d'observatrices évaluées | Kappa pondéré | | | Accord parfait ¹ | Accord proche ² |
|----------|---------------------------------------|---------------|---------|---------|--------------------------------|-------------------------------|
| | | Médiane | Minimum | Maximum | | |
| Accord 1 | 28 | 0,65 | 0,42 | 0,79 | 0,66 | 0,92 |
| Accord 2 | 28 | 0,78 | 0,66 | 0,87 | 0,75 | 0,97 |

1. L'accord parfait indique le pourcentage d'atteinte de cotes identiques entre deux juges.

2. L'accord proche indique le pourcentage d'atteinte de cotes à un point près.

Source : Institut de la statistique du Québec, *Grandir en qualité 2003*.

Pour ces deux échelles, nous avons donc opté pour une procédure de validation différente et moins lourde. En fait, un total de dix tests d'accord inter-juges (cinq par échelle) ont été réalisés à la suite de la sélection aléatoire d'observatrices. Cette procédure visait particulièrement à tester l'hypothèse selon laquelle la formation rend les observatrices aptes à observer et à interpréter les trois types de services visés.

Ces exercices d'accord ont été réalisés en début d'enquête dans des services en pouponnière et en milieu familial sélectionnés pour l'enquête et toute autre

observation dans ces services devait être retardée jusqu'à l'obtention des scores d'accord attendus. Les résultats de ces tests, présentés au tableau 2.4, ont confirmé l'hypothèse soulevée. En milieu familial, les coefficients *kappas pondérés* se sont tous situés au-delà de 0,60, soit un degré d'accord élevé. Bien que généralement un peu plus faibles, les scores obtenus lors des tests d'accord en pouponnière ont également satisfait aux exigences fixées puisque ceux-ci ont varié de 0,58 à 0,78.

Tableau 2.4

Accords inter-juges de l'enquête *Grandir en qualité 2003* pour les échelles en milieu familial et en pouponnière

| | Kappa pondéré | % | |
|---------------------------|---------------|-----------------------------|----------------------------|
| | | Accord parfait ¹ | Accord proche ² |
| En milieu familial | | | |
| Observatrice 1 | 0,84 | 0,77 | 0,98 |
| Observatrice 2 | 0,62 | 0,61 | 0,94 |
| Observatrice 3 | 0,83 | 0,79 | 1,00 |
| Observatrice 4 | 0,77 | 0,74 | 0,94 |
| Observatrice 5 | 0,75 | 0,66 | 0,97 |
| En pouponnière | | | |
| Observatrice 1 | 0,61 | 0,61 | 0,91 |
| Observatrice 2 | 0,78 | 0,75 | 0,98 |
| Observatrice 3 | 0,74 | 0,77 | 0,94 |
| Observatrice 4 | 0,60 | 0,63 | 0,86 |
| Observatrice 5 | 0,58 | 0,59 | 0,90 |

1. L'accord parfait indique le pourcentage d'atteinte de cotes identiques entre deux juges.

2. L'accord proche indique le pourcentage d'atteinte de cotes à un point près.

Source : Institut de la statistique du Québec, *Grandir en qualité 2003*.

2.2.5.2 Suivi et soutien des observatrices

De nombreuses mesures ont été mises en place afin d'assurer le maintien d'un niveau d'accord acceptable tout au long de l'enquête. En premier lieu, deux séances de formation complémentaire ont été tenues. Les observatrices ont été réunies au terme des premier et deuxième mois d'observation pour participer à chaque fois à une formation de deux jours permettant notamment de faire le point sur des questions soulevées à la suite de leurs observations, de renforcer certaines notions et de partager leurs expériences.

En plus de cette démarche, un suivi intensif et régulier auprès des observatrices a été assuré lors du déroulement de l'enquête. En particulier, une procédure systématique de révision des échelles par des observatrices-vérificatrices permettait de valider le contenu des échelles et de signaler certains manques ou incohérences et était l'occasion de retours réguliers auprès de toutes les observatrices. Au besoin, ces dernières pouvaient également consulter les formatrices ou une autre personne-ressource. Des notes pouvaient également être envoyées à l'ensemble des observatrices si la situation l'exigeait.

Les formatrices ont également visité chacune des observatrices lors d'une observation pour assurer un soutien plus personnalisé encore. Ces visites d'une demi-journée ont permis aux observatrices de discuter de faits concrets survenus en présence de la formatrice. Un retour sur les observations a été fait de façon systématique lors de cette même journée.

Enfin, la deuxième formation complémentaire a aussi été l'occasion de mesurer, une dernière fois, le degré d'accord inter-juges par un exercice tenu dans un laboratoire d'enfants. Les résultats ont encore une fois confirmé le maintien d'un degré satisfaisant d'accord entre les expertes et toutes les observatrices, démontrant, par le fait même, l'utilité des mesures de soutien déployées en cours d'enquête. Plus précisément, toutes les observatrices ont obtenu un accord mesuré par le *kappa pondéré* supérieur à 0,55, dont 23 sur 26¹⁸ avec un degré d'accord jugé élevé (supérieur à 0,60).

2.2.6 Évaluation de la consistance interne des échelles

Les données recueillies en cours d'enquête ont également permis de revenir sur l'évaluation de la consistance interne des construits des échelles ou, en d'autres termes, de valider dans quelle mesure un groupe d'items est cohérent par rapport aux concepts mesurés. Pour chacune des échelles, la cohérence des principaux construits que sont les quatre dimensions de la qualité a donc été évaluée à l'aide du coefficient alpha (ζ) de Cronbach (Cronbach, 1951).

Les résultats de cette analyse de fidélité démontrent que les dimensions 1, 2 et 3 possèdent un degré de consistance interne de « très bon » à « élevé »¹⁹ selon une grille d'interprétation proposée par Murphy et Davidshofer (1994) (tableau 2.5)²⁰. Toutefois, comme le coefficient alpha peut varier de façon importante en fonction du nombre d'items, cette mesure rend difficile l'interprétation (et la comparaison avec les autres résultats) de la dimension 4 qui ne contient que sept items. La mesure de la consistance de ce construit est donc fournie à titre indicatif seulement.

2.2.7 Questionnaires et autres sources de données

2.2.7.1 Questionnaires de l'enquête

L'enquête *Grandir en qualité 2003* vise également à déterminer les facteurs associés à la qualité des services de garde éducatifs offerts aux enfants québécois. C'est dans cette perspective que des renseignements divers concernant non seulement les caractéristiques et le fonctionnement des services eux-mêmes, mais aussi au sujet des éducatrices, des RSG, des gestionnaires des groupes observés, des établissements auxquels ils sont rattachés et des milieux géographiques d'appartenance ont été recueillis dans le cadre de l'enquête ou à partir d'autres sources de données.

Tout d'abord, quatre questionnaires²¹ ont été conçus, pour les besoins de l'enquête, par le MESSF en collaboration avec l'ISQ et les membres du comité aviseur pour recueillir des renseignements relatifs aux caractéristiques des éducatrices, des RSG et des

Tableau 2.5

Analyse de la consistance interne des échelles d'observation de l'enquête *Grandir en qualité 2003* selon les dimensions

| | Nombre d'items ¹ | Alpha de Cronbach (standardisé) |
|--|-----------------------------|---------------------------------|
| Échelle préscolaire (n = 404) | 121 | 0,954 |
| Dimension 1 : Structuration des lieux | 39 | 0,849 |
| Dimension 2 : Structuration et variation des types d'activités | 30 | 0,842 |
| Dimension 3 : Interaction éducatrice-enfants | 45 | 0,930 |
| Dimension 4 : Interaction éducatrice-parents | 7 | 0,648 |
| Échelle pouponnière (n = 236) | 115 | 0,956 |
| Dimension 1 : Structuration des lieux | 37 | 0,871 |
| Dimension 2 : Structuration et variation des types d'activités | 30 | 0,831 |
| Dimension 3 : Interaction éducatrice-enfants | 41 | 0,938 |
| Dimension 4 : Interaction éducatrice-parents | 7 | 0,641 |
| Échelle milieu familial (n = 153) | 130 | 0,943 |
| Dimension 1 : Structuration des lieux | 44 | 0,814 |
| Dimension 2 : Structuration et variation des types d'activités | 30 | 0,819 |
| Dimension 3 : Interaction RSG-enfants | 49 | 0,924 |
| Dimension 4 : Interaction RSG-parents | 7 | 0,561 |

1. Le nombre d'items correspond au nombre d'items utilisés dans les calculs de consistance et considère ainsi les items pouvant être décomposés (ex. : 5.4.1S et 5.4.1N).

Source : Institut de la statistique du Québec, *Grandir en qualité 2003*.

gestionnaires responsables des groupes d'enfants observés, de même que, plus largement, de l'information concernant les établissements fréquentés par les enfants :

↓ QP-1 : *Questionnaire autoadministré destiné à l'éducatrice*; ce questionnaire s'adresse à l'éducatrice attitrée du groupe d'enfants observé²². Les questions traitent notamment des conditions de travail, de l'expérience professionnelle, de la satisfaction au travail et de certaines caractéristiques personnelles des éducatrices tels l'âge et le niveau de scolarité.

↓ QP-2 : *Questionnaire téléphonique destiné au gestionnaire de la garderie ou de l'installation*; ce questionnaire est administré à la personne la plus directement responsable de la gestion quotidienne du groupe d'enfants sélectionné. Les services sélectionnés détenaient la responsabilité de désigner cette personne. En premier lieu, ce questionnaire permet d'obtenir de l'information spécifique aux gestionnaires du même ordre que celle recueillie pour les éducatrices. De plus, c'est par le biais du QP-2 que sont récoltées certaines données sur le fonctionnement de l'installation ou

de la garderie où s'est déroulée l'observation. Il s'agit, par exemple, de questions relatives à la présence de clientèles prioritaires, d'une conseillère pédagogique ou de stagiaires, ou encore, portant sur les moyens de communication utilisés par les services pour informer les parents des politiques en cours.

↓ QP-3 : *Questionnaire téléphonique destiné à la responsable ou coordonnatrice au CPE de la garde en milieu familial*; suivant la logique appliquée pour le QP-2, ce questionnaire s'adresse à la personne la plus directement responsable du soutien et de la communication avec la RSG du groupe d'enfants sélectionné. Des renseignements semblables à ceux recueillis au QP-2 permettent de dresser sommairement le profil de ces personnes (niveau de formation, expérience, satisfaction au travail, etc.) que nous définissons comme les coordonnatrices du volet en milieu familial. De plus, ce questionnaire sert à décrire certaines caractéristiques du volet en milieu familial dans les CPE (nombre de nouvelles RSG, services offerts par le CPE aux RSG, etc.).

↓ QP-4 : *Questionnaire téléphonique destiné aux responsables de services en milieu familial*; ce questionnaire s'adresse aux RSG des groupes sélectionnés et vise à préciser certaines de leurs caractéristiques professionnelles et personnelles, notamment les conditions de travail, l'expérience professionnelle, le soutien offert par le CPE, leur degré de satisfaction face au travail, l'âge, le revenu et le niveau de scolarité.

Rappelons que les données recueillies à l'aide de ces questionnaires ne constituent pas des critères de qualité comme tels mais sont traitées comme des facteurs pouvant être associés à la qualité telle que mesurée dans l'enquête.

Les versions préliminaires de ces questionnaires ont été administrées auprès des services de garde éducatifs ayant pris part au prétest de l'enquête réalisé au printemps 2002, soit 142 questionnaires administrés dans le cadre de 69 observations. Dans leur format initial respectif, tous les questionnaires étaient de type autoadministré. Par la suite, des ajustements ont été apportés dans le but de réduire le fardeau des répondants et de diminuer la non-réponse en clarifiant les énoncés de certaines questions. Un deuxième prétest a été effectué, comptant une vingtaine de questionnaires avec les versions modifiées de façon à s'assurer que les changements apportés amélioreraient la qualité des données recueillies.

Finalement, dans leur version finale, trois des quatre questionnaires ont été administrés par téléphone au cours des journées suivant la visite de l'observatrice dans les groupes d'enfants sélectionnés. Les intervieweurs téléphoniques chargés d'administrer ces trois questionnaires ont reçu au préalable une formation, visant à parfaire leur connaissance du contenu des questionnaires ainsi que du contexte général de l'enquête, de façon à pouvoir répondre aux questions soulevées par les répondants. L'unique questionnaire autoadministré est celui s'adressant à l'éducatrice attirée au groupe sélectionné; il a été distribué par l'observatrice et recueilli lors de la journée d'observation.

2.2.7.2 Sources de données complémentaires

Une sélection de variables issues des *Rapports d'activités 2002-2003 des centres de la petite enfance et des garderies*²³ ainsi que des *Rapports financiers annuels des centres de la petite enfance et des garderies 2002-2003*²⁴, représentant la situation au 31 mars 2003, s'ajoutent à l'information recueillie directement dans le cadre de l'enquête. Ces rapports remplis annuellement par l'ensemble des services régis du Québec fournissent des renseignements complémentaires quant à l'organisation des services enquêtés. En particulier, les données des rapports d'activités fournissent à l'enquête des renseignements supplémentaires sur les conditions de travail des éducatrices et des gestionnaires (ex. : nombre de congés de divers types), sur le fonctionnement de l'installation ou de la garderie (ex. : nombre de réunions du conseil d'administration dans l'année, nombre d'enfants dans le volet) ou sur le volet milieu familial du CPE (ex. : nombre de RSG dans le volet).

Quant aux variables issues des rapports financiers, celles-ci ont pour but de décrire plus spécifiquement la structure financière des installations et des garderies à partir d'indices portant sur les charges et les produits de chacun des services enquêtés. L'ISQ et le MESSF ont donc convenu d'extraire certaines données de ces rapports dans le but de ne pas surcharger les participants à l'enquête et de réduire les coûts de collecte. Notons que des mesures ont été prises par l'ISQ pour éviter toute possibilité de croisement de manière à préserver l'anonymat des services participants.

Finalement, des données issues du Recensement du Canada 2001 ont également été utilisées afin de décrire certaines caractéristiques démographiques et socioéconomiques propres à la population résidant dans les secteurs où sont localisés les services de garde éducatifs. Ces données comprennent, par exemple, le taux d'activité de la population de 15 ans et plus, la part d'unités familiales à faible revenu et celle des résidents nés hors Québec. La localisation des services enquêtés, à partir du code postal, a aussi permis de regrouper les services selon leur appartenance à différents territoires administratifs (ex. : régions administratives) ou statistiques (ex. : régions métropolitaines de recensement).

2.3 Traitement des données

2.3.1 Validation des données

La validation des données recueillies à l'aide des échelles d'observation s'est déroulée en trois étapes. D'abord, une procédure systématique de validation manuelle a été instaurée sur réception des échelles d'observation à l'ISQ. Des observatrices mandatées à la vérification des échelles devaient s'assurer de la présence et de la cohérence de l'information relative à l'identification des services et aux contextes d'observation. La présence des items remplis par entrevue ainsi que de certains items spécifiques était également validée afin de réduire la non-réponse partielle. Au besoin, les vérificatrices devaient consulter les observatrices afin de compléter l'information manquante. Dans un deuxième temps, le programme de saisie informatique fixait des seuils minimum et maximum pour l'entrée des variables administratives et de contextes d'observation. Le cas échéant, les erreurs étaient corrigées par les responsables de terrain ou après avoir consulté les observatrices. Finalement, la banque de données brutes, c'est-à-dire celle contenant l'ensemble des éléments à cocher, ainsi que la banque de données contenant les cotes de qualité ont fait l'objet d'une validation informatique de manière à tester les liens logiques entre certains items et à s'assurer de la cohérence des cotes produites par le système d'attribution des cotes. Ces validations ont permis de corriger manuellement quelques erreurs de même que de revoir des formules d'attribution de cotes pour les rendre mutuellement exclusives.

Le *Questionnaire autoadministré destiné à l'éducatrice* (QP-1) a été soumis sensiblement au même processus de vérification que les échelles d'observation. En bref, les vérifications des codeurs ont porté sur la présence des réponses, la correction des sauts et l'entrée de codes de non-réponse partielle pour faciliter la saisie informatique du questionnaire. Les étapes subséquentes de validation informatique à la saisie et sur les banques de données résultantes ont plutôt permis de vérifier les valeurs extrêmes. La validation des questionnaires QP-2, QP-3 et QP-4 a été largement facilitée par leur administration téléphonique puisque ce mode de collecte implique l'informatisation des questionnaires, ce qui élimine d'emblée les omissions et les erreurs de chemin. Pour l'ensemble des questionnaires, la codification des questions ouvertes (par exemple « autres ») a été faite en cours d'analyse à l'aide des experts dans le domaine.

2.3.2 Non-réponse et items non observés

2.3.2.1 Non-réponse globale

Comme dans toute enquête, les unités échantillonnées n'ont pas toutes accepté de prendre part à l'observation ou de répondre aux différents questionnaires. Le taux de réponse se définit comme le rapport entre le nombre d'unités répondantes et le nombre d'unités admissibles à l'enquête. Dans le cadre de l'enquête, même si l'unité d'analyse est l'enfant, les taux de réponse sont présentés en lien avec les groupes d'enfants sélectionnés, et ce, pour chaque type de services de l'étude (tableau 2.6). Ces taux de réponse ne sont pas pondérés et font référence à la réalisation complète de l'observation.

Tableau 2.6

Taux de réponse de l'enquête *Grandir en qualité 2003* selon le type de services et l'instrument de mesure

| | Échelle d'observation | Questionnaire destiné à l'éducatrice (QP-1) ou à la RSG (QP-4) | Questionnaire destiné à la gestionnaire (QP-2 ou QP-3) |
|--------------------------------------|--------------------------|--|--|
| Installation moins de 18 mois | | | |
| Répondants | 124 | 124 | 124 |
| Non-répondants | 3 | 3 | 3 |
| Inadmissibles | 1 | 1 | 1 |
| Taux de réponse (%) | 97,6 | 97,6 | 97,6 |
| Installation 18 mois à 5 ans | | | |
| Répondants | 220 | 220 | 220 |
| Non-répondants | 8 | 8 | 8 |
| Inadmissibles | 0 | 0 | 0 |
| Taux de réponse (%) | 96,5 | 96,5 | 96,5 |
| Garderie moins de 18 mois | | | |
| Répondants | 112 | 112 | 111 |
| Non-répondants | 12 | 12 | 13 |
| Inadmissibles | 0 | 0 | 0 |
| Taux de réponse (%) | 90,3 | 90,3 | 89,5 |
| Garderie 18 mois à 5 ans | | | |
| Répondants | 184 | 184 | 184 |
| Non-répondants | 40 | 40 | 40 |
| Inadmissibles | 1 | 1 | 1 |
| Taux de réponse (%) | 82,1 | 82,1 | 82,1 |
| Milieu familial | | | |
| Répondants | 153 | 153 | 153 |
| Non-répondants | 47 | 47 | 47 |
| Inadmissibles | 0 | 0 | 0 |
| Taux de réponse (%) | 76,5 | 76,5 | 76,5 |
| Ensemble des services | | | |
| Répondants | 793 | 793 | 792 |
| Non-répondants | 110 | 110 | 111 |
| Inadmissibles | 2 | 2 | 2 |
| Taux de réponse (%) | 87,7 | 87,7 | 87,7 |

Source : Institut de la statistique du Québec, *Grandir en qualité 2003*.

Au total, 793 des 903 services admissibles ont accepté de participer à l'observation pour un taux de réponse global de 87,8 %. On remarque que le taux de réponse aux échelles d'observation varie selon les types de services de l'enquête. Le taux de réponse des services en installation de CPE est très élevé, atteignant respectivement 97,6 % et 96,5 % dans la classe d'âge des poupons et chez les groupes d'âge préscolaire. Le taux de réponse des garderies est de 90,3 % et de 82,1 % respectivement pour les groupes de poupons et ceux d'enfants de 18 mois à 5 ans. Ces taux de réponse sont tous supérieurs au taux de réponse global attendu de 80 %. Quant aux services en milieu familial,

le taux de réponse obtenu est de 76,5 %. Ce taux est légèrement inférieur au taux visé; il est toutefois bien supérieur à celui obtenu lors du prétest de l'enquête, soit 60 %.

Parmi les services ayant participé à l'observation, tous ont accepté de remplir les questionnaires de l'enquête à l'exception d'un seul cas²⁵. En conséquence, sauf pour cette exception, les taux de réponse aux questionnaires sont les mêmes que pour les échelles d'observation.

Par ailleurs, on dispose des données sur les rapports d'activités et les rapports financiers 2002-2003 en provenance du MESSF pour la grande majorité des services de garde éducatifs ayant participé à l'enquête de sorte que les taux de réponse combinés obtenus ne sont que légèrement inférieurs à ceux de l'enquête (tableau 2.7). Ainsi, selon le type de services, les taux de réponse suivent les tendances observées pour les questionnaires, excédant 95 % pour les groupes en installation de CPE, alors qu'ils se situent entre 76 % et

88 % en garderie et qu'ils sont légèrement inférieurs à 75 % pour les groupes en milieu familial. Soulignons qu'au moment de l'enquête, quelques garderies étaient non conventionnées (n = 6), c'est-à-dire qu'elles ne détenaient pas d'accord avec le Ministère pour l'offre de places à contribution réduite. C'est ce qui explique le nombre élevé de garderies considérées non admissibles à remplir le rapport financier annuel 2002-2003 car les garderies non conventionnées n'ont pas l'obligation de remplir ce rapport.

Tableau 2.7

Taux de réponses combinés à l'enquête *Grandir en qualité* 2003 et aux rapports d'activités et financiers 2002-2003 des CPE et des garderies selon le type de services

| | Rapports d'activités | Rapports financiers |
|--------------------------------------|----------------------|---------------------|
| Installation moins de 18 mois | | |
| Répondants | 123 | 124 |
| Non-répondants | 4 | 3 |
| Inadmissibles | 1 | 1 |
| Taux de réponse (%) | 96,9 | 97,6 |
| Installation 18 mois à 5 ans | | |
| Répondants | 217 | 218 |
| Non-répondants | 11 | 10 |
| Inadmissibles | 0 | 0 |
| Taux de réponse (%) | 95,2 | 95,6 |
| Garderie moins de 18 mois | | |
| Répondants | 105 | 104 |
| Non-répondants | 19 | 14 |
| Inadmissibles | 0 | 6 |
| Taux de réponse (%) | 84,7 | 88,1 |
| Garderie 18 mois à 5 ans | | |
| Répondants | 171 | 177 |
| Non-répondants | 53 | 41 |
| Inadmissibles | 1 | 7 |
| Taux de réponse (%) | 76,3 | 81,6 |
| Milieu familial | | |
| Répondants | 149 | 145 |
| Non-répondants | 51 | 55 |
| Inadmissibles | 0 | 0 |
| Taux de réponse (%) | 74,5 | 72,5 |

Source : Institut de la statistique du Québec, *Grandir en qualité* 2003.

2.3.2.2 Non-réponse partielle

En plus de la non-réponse pour l'ensemble d'un instrument, on retrouve également de la non-réponse partielle. La non-réponse est partielle lorsque seulement une partie du questionnaire a été remplie. Celle-ci peut également entraîner des biais dans les estimations puisque les non-répondants peuvent présenter des caractéristiques différentes de celles des répondants. Plus la non-réponse partielle est importante, plus le risque de biais est grand. Dans le cas présent, l'analyse de la non-réponse partielle ne concerne pas les échelles d'observation puisque les processus de collecte, de vérification et de saisie ont permis de s'assurer que tous les items observés étaient complétés. Cette analyse a donc été effectuée seulement pour les questionnaires et ce, dans chacun des types de services. Elle consiste à évaluer la non-réponse pour chaque question, à cibler les cas problématiques et à déterminer l'importance de la non-réponse partielle. À noter que le taux de non-réponse à chaque question est obtenu par le rapport entre le nombre pondéré d'individus n'ayant pas répondu à la question et le nombre pondéré d'individus devant répondre à cette question (répondants potentiels). Prenant en compte les différentes tailles d'échantillon selon les types de services, nous avons considéré qu'une question était susceptible d'être problématique si le taux de non-réponse était supérieur à 10 % pour le volet milieu familial et pour les services accueillant les enfants de moins de 18 mois et à 5 % pour ceux accueillant les enfants de plus de 18 mois à 5 ans. Cette analyse a permis de détecter des problèmes de traitement de certaines questions, dont les principaux sont discutés à l'annexe A. À noter qu'aucun problème n'a été observé en milieu familial.

2.3.2.3 Impact des items non observés sur le score moyen d'une sous-dimension dans les échelles d'observation²⁶

Dans les échelles, les observatrices peuvent ne pas être en mesure d'observer les éléments qui composent certains items. Ce sont des items qui décrivent des situations particulières qui peuvent ne pas survenir le jour de l'observation. Par exemple, les observatrices ne pouvaient juger de la qualité des périodes en groupes jumelés si au cours de la journée d'observation il n'y

avait aucune période de ce genre. Le cas échéant, une valeur de 99 (« ne s'applique pas ») est attribuée au score de l'item dans le but d'indiquer qu'il ne contribuera pas au score moyen de qualité de la sous-dimension.

Une étude de sensibilité a donc été effectuée afin de déterminer si les items comportant des 99 peuvent influencer la moyenne du score de qualité d'une sous-dimension. En effet, le score de qualité moyen pourrait être meilleur si on ne peut évaluer un item qui, en raison de la nature du thème, aurait tendance à recevoir des cotes généralement faibles. C'est le cas notamment de l'item portant sur les changements de couche, un item où les cotes obtenues sont souvent faibles. L'absence d'enfants aux couches occasionnerait donc une hausse du score moyen de la sous-dimension concernée. À l'inverse, le score de qualité pourrait être plus faible si on ne peut évaluer un item qui aurait tendance à obtenir de bons résultats, à cause de la nature du thème.

Les résultats de cette étude de sensibilité sont présentés au tableau 2.8. L'analyse démontre que, parfois, quelques items comportant des 99 ont tendance à faire diminuer le score moyen de qualité de certaines sous-dimensions, alors que d'autres ont tendance à le faire augmenter. Notamment, on observe que pour tous les types de services la manifestation de comportements dérangeants a un impact négatif sur les scores moyens de certaines sous-dimensions et que, en milieu familial, la présence de poupons contribue à faire baisser les scores moyens à plusieurs sous-dimensions. À l'inverse, dans la classe d'âge de moins de 18 mois, la présence d'enfants capables de ramper ou de s'asseoir a une incidence favorable sur l'évaluation de la sous-dimension 1.2 portant sur l'aménagement des lieux. Enfin, soulignons que l'analyse des résultats ne comporte pas d'ajustements visant à compenser pour ces facteurs puisque les spécialistes du domaine des services de garde éducatifs jugent les tendances révélées par l'étude de sensibilité explicables et plausibles. Cependant, les nuances nécessaires sont apportées dans le texte là où il est pertinent de le faire.

Tableau 2.8

Résultats de l'étude de sensibilité pour les cinq types de services de l'enquête *Grandir en qualité 2003*

| Sous-dimension | Description de l'événement | Tendance générale (impact de l'événement sur le score moyen de la sous-dimension) |
|--------------------------------------|--|--|
| Installation moins de 18 mois | | |
| 1.2 | Présence d'un enfant capable de s'asseoir | Augmentation |
| 1.2 | Présence d'un enfant capable de ramper | Augmentation |
| 3.2 | Manifestation de comportements dérangeants | Diminution |
| 3.2 | Déplacement du groupe | Diminution |
| Installation 18 mois et plus | | |
| 1.1 | Présence d'un enfant aux couches | Diminution |
| 3.2 | Manifestation de comportements dérangeants | Diminution |
| Garderie moins de 18 mois | | |
| 1.2 | Présence d'un enfant capable de s'asseoir | Augmentation |
| 1.2 | Présence d'un enfant capable de ramper | Augmentation |
| 3.2 | Manifestation de comportements dérangeants | Diminution |
| Garderie 18 mois et plus | | |
| 1.1 | Présence d'un enfant aux couches | Diminution |
| 3.2 | Manifestation de comportements dérangeants | Diminution |
| Milieu familial | | |
| 1.1 | Présence d'un poupon au biberon | Diminution |
| 1.1 | Présence d'un enfant aux couches | Diminution |
| 1.1 | Sortie à l'extérieur de la cour | Diminution |
| 2.3 | Présence d'un poupon | Diminution |
| 2.4 | Présence d'une assistante | Augmentation |
| 3.2 | Manifestation de comportements dérangeants | Diminution |
| 3.2 | Présence d'un poupon | Diminution |
| 3.3 | Présence d'un poupon | Diminution |

Source : Institut de la statistique du Québec, *Grandir en qualité 2003*.

2.3.3 Pondération

Pour les échelles d'observation, afin d'inférer les données de l'échantillon à la population visée, chaque unité répondante doit se voir associer un poids. Ce poids, appelé poids populationnel, est le nombre d'unités visées par l'enquête et « représentées » par le répondant.

Les différentes étapes de la pondération dans l'enquête sont au nombre de quatre. Il s'agit du calcul des poids initiaux correspondant à l'inverse de la probabilité de sélection d'un groupe d'enfants dans sa strate, de l'ajustement des poids pour tenir compte de la non-réponse, du calcul des poids attribués à chaque enfant observé et de la poststratification.

2.3.3.1 Probabilité de sélection

Comme l'étude repose sur un échantillon probabiliste, elle permet d'évaluer pour chaque unité de la population, la probabilité de faire partie de l'échantillon. L'inverse de la probabilité de sélection est utilisé comme pondération initiale. Cette première étape de pondération permet de tenir compte adéquatement de la non-proportionnalité de l'échantillon par rapport à la distribution de la population. On trouvera, à l'annexe A, une présentation détaillée du calcul des poids initiaux.

2.3.3.2 Ajustement pour la non-réponse²⁷

Le taux de réponse est un élément important quant à la qualité des résultats d'une enquête. Ce taux est un indicateur de la présence potentielle de biais dans les résultats. En effet, dans toute enquête, il est possible

que les non-répondants possèdent des caractéristiques différentes de celles des répondants. Ainsi, plus la non-réponse est élevée, plus il y a des risques que des biais soient introduits dans les estimations inférées à l'ensemble de la population à partir des réponses reçues.

L'ajustement pour la non-réponse consiste à redresser l'échantillon des répondants par une modification de la pondération afin de le rendre, dans la mesure du possible, semblable à l'échantillon tiré initialement. Cette technique nécessite de l'information complémentaire sur les répondants et les non-répondants. Pour que l'ajustement effectué soit efficace, il est important que l'information auxiliaire dont on dispose soit liée aux variables mesurées dans l'enquête, sinon l'impact de l'ajustement sera négligeable sur la réduction d'un biais potentiel.

Ainsi l'ajustement s'appuie sur la création de groupes homogènes de pondération à l'aide de variables provenant de la base de sondage. Ces variables sont la vague de la collecte²⁸, le nombre d'enfants dans le service, l'âge du service, le nombre de services sélectionnés dans un même établissement et la strate. Afin de déterminer la ou les variables expliquant le mieux la non-réponse, l'algorithme CHAID (**Chi-Square Automatic Interaction Detection**/Détection automatique d'interactions du khi-carré), mis au point par Kass (1980), a été utilisé pour aider à constituer des classes de pondération ayant des taux de réponse différents. Après l'application de la procédure CHAID à l'échantillon et en tenant compte de certaines contraintes, il s'est avéré que la seule variable pertinente pour expliquer la non-réponse des cinq types de services était un regroupement de strates en deux ou trois modalités. Rappelons que les strates sont des régions administratives ou des regroupements de régions administratives. Ainsi, les classes de pondération correspondent à des regroupements de régions administratives.

Il est important de souligner que la pondération calculée pour les échelles d'observation est aussi valable pour les instruments complémentaires (QP-1 à QP-4, les données des rapports d'activités 2002-2003 et des

rapports financiers 2002-2003 des CPE et des garderies) étant donné leur taux de réponse élevé.

2.3.3.3 Poids attribué à chaque enfant observé

Puisque tous les enfants des groupes sélectionnés sont observés, il est possible d'attribuer un poids à chaque enfant de façon à inférer les résultats à l'ensemble des enfants de la population visée. On trouvera à l'annexe A une présentation du calcul des poids attribués à chaque enfant observé.

2.3.3.4 Poststratification

La poststratification, pour sa part, consiste à redresser l'échantillon des répondants par une modification de la pondération afin de le rendre similaire à la population visée. Ainsi, on découpe la population en groupes (aussi appelés poststrates) dans lesquels les réponses au questionnaire sont homogènes. Tout comme la stratification, la poststratification augmente la précision des estimations; elle contribue simultanément à diminuer le biais de la non-réponse et à corriger pour la sous-couverture.

La variable qui a servi à l'ajustement est la strate; les données utilisées pour les installations et les garderies proviennent du fichier CAFE mis à jour le 20 février 2003, soit la plus récente mise à jour du fichier disponible au moment de la pondération. Afin de respecter la définition de la population visée, les modifications appropriées ont été apportées au fichier.

En ce qui concerne le volet en milieu familial, la poststratification diffère quelque peu de celle des installations de CPE et des garderies. Nous avons constaté que le nombre de places au permis du fichier CAFE est supérieur au nombre de places occupées dans le cas des milieux familiaux, ce qui rend son utilisation pour la poststratification hasardeuse. Quant au plus récent fichier des rapports d'activités du volet en milieu familial disponible au moment de la pondération, il représentait la situation au 31 mars 2002. Étant donné le temps écoulé entre cette date et la collecte, il est fort probable que l'information contenue dans ce fichier sous-estime la population visée, ce qui empêche aussi son utilisation dans une poststratification habituelle.

Malgré le fait que les tailles de la population visée par poststrate obtenues du fichier des rapports d'activités ne soient pas représentatives de la situation au moment de la collecte, il est raisonnable de penser que la répartition relative n'a pas beaucoup changé en une année. Pour cette raison, la poststratification pour le volet en milieu familial a été effectuée de manière à reproduire seulement la répartition relative de la population visée. Le détail des calculs de la poststratification est présenté à l'annexe A.

2.3.4 Méthodes d'analyse

Les résultats de la présente enquête ont été produits par l'ISQ; deux approches ont été adoptées. La première, qui est approximative sur le plan du calcul de précision et des tests, permet la production à grande échelle d'estimations et de tests avec le logiciel d'analyse statistique SAS. La seconde approche utilise plutôt un logiciel statistique spécialisé dans le traitement de données provenant d'un plan de sondage complexe, SUDAAN (Shah, Barnwel et Bieler, 1997).

La majorité des estimations produites dans le rapport sont des proportions ou des moyennes. On aborde donc la production d'estimations simples, puis le calcul de la précision associée à ces estimations. Finalement, les méthodes ayant servi à la construction des tests du khi-carré et des tests de comparaison de proportions et de moyennes sont présentées.

2.3.4.1 Estimations de proportions et de moyennes

Comme nous l'avons déjà précisé, des cotes selon une échelle à quatre niveaux (1 – inadéquat; 2 – minimal; 3 – bien; 4 – très bien) sont assignées à chaque item selon les observations effectuées. C'est à partir de ces mêmes cotes qu'il est possible de calculer un score de qualité pour une observation et que, de là, il est possible de calculer des moyennes par type de services

(moyennes d'un item, d'une sous-dimension, d'une dimension ou pour l'ensemble de l'échelle), ou encore, produire la fréquence simple des scores attribués aux services et qui sont ensuite inférés aux enfants.

Toutes les estimations simples de scores de qualité, sous la forme de proportion ou de moyenne, produites à partir des données de l'enquête utilisent les pondérations. Cette façon de faire permet l'inférence à la population visée.

2.3.4.2 Précision des estimations de proportions

Dans la majorité des enquêtes statistiques, les estimations sont entachées d'erreurs dites d'échantillonnage, c'est-à-dire celles dues au fait que l'on n'enquête qu'une partie de la population visée. La complexité du plan de sondage utilisé pour produire l'échantillon influence la précision des résultats; une mesure de la précision de chaque estimation est donc nécessaire à l'inférence.

Le recours à un plan stratifié proportionnel à deux degrés de grappes a un impact sur la précision des estimations. Si on fait l'hypothèse que l'échantillonnage est aléatoire simple, on prend le risque d'engendrer un biais dans les estimations et de sous-estimer leur variance. En conséquence, il peut en résulter des conclusions erronées sur le plan des tests statistiques.

L'effet de plan nous renseigne donc sur la précision relative par rapport à un plan aléatoire simple. Il peut être calculé pour chacune des estimations produites mais, étant donné leur grand nombre, des effets de plan moyens ont été calculés pour une série de variables homogènes par rapport à la variance, pour chacun des cinq types de services de l'enquête²⁹. Les effets de plan produits pour l'analyse de données sont présentés au tableau 2.9.

Tableau 2.9

Effet de plan selon les cinq types de services de l'enquête *Grandir en qualité 2003*

| | Effet de plan |
|---|---------------|
| Installation (moins de 18 mois/18 mois à 5 ans) et garderie (18 mois à 5 ans) | 1,05 |
| Garderie (moins de 18 mois) | 0,7 |
| Milieu familial | 1,15 |

Source : Institut de la statistique du Québec, *Grandir en qualité 2003*.

En conséquence, l'estimation de la variance des proportions, pour le volet milieu familial, est environ 15 % plus grande que celle d'un plan aléatoire simple de même taille.

Notons que dans tous les cas, les calculs de précision nécessaires pour évaluer les effets de plan ont été faits avec le logiciel SUDAAN. La méthode de linéarisation de Taylor sert à faire l'estimation de la variance et l'estimateur de Yates-Grundy-Sen (Yates et Grundy-Sen, 1953) a été utilisé.

Ces effets de plan ont été pris en considération lors de l'analyse des données de l'enquête, notamment pour la construction d'intervalles de confiance approximatifs des proportions et pour un ajustement simple aux tests du khi-carré produits par le logiciel statistique SAS. Cet ajustement est rendu possible en divisant les poids échantillonnaux³⁰ par l'effet de plan.

2.3.4.3 Précision des estimations de moyennes

Dans toutes les situations, la précision des estimations de moyennes a été obtenue à l'aide du logiciel SUDAAN.

2.3.4.4 Tests statistiques

Le test d'indépendance du khi-carré, qui sert à établir s'il existe ou non un lien entre deux variables catégorielles, ainsi que les tests de comparaison de proportions ou de moyennes sont les principaux tests statistiques utilisés dans l'enquête.

En ce qui concerne la relation entre deux variables continues, des tests de comparaison de moyennes ont été effectués sur la base des intervalles de confiance obtenus à l'aide du logiciel SAS. L'absence de chevauchement entre ces intervalles était le seul critère retenu pour signaler une différence entre les estimations.

La relation entre deux variables catégoriques a été mesurée par le test du khi-carré. Le seuil de signification est fixé à 5 % pour tous les tests. Dans tous les cas, un ajustement a été effectué pour tenir compte de la complexité du plan de sondage³¹. Pour les tableaux comptant deux catégories par variable, ce

test s'avère à la fois valide pour indiquer qu'il existe des différences significatives entre les modalités et, en conséquence, ne nécessite pas d'effectuer des tests de comparaison de proportions.

Dans le cas où des variables comptant plus de deux catégories ont été utilisées pour les croisements, des tests de comparaison ont été effectués lorsque le test du khi-carré révélait une association significative entre les deux variables. Les tests d'égalité de deux proportions étaient effectués à partir de deux intervalles de confiance, un pour chaque paramètre d'intérêt. La règle à appliquer était la suivante : l'absence de chevauchement entre ces intervalles obtenus à l'aide du logiciel SAS indiquait une différence significative. Dans les cas où les différences semblaient importantes, sans avoir été déclarées significatives, on établissait la comparaison en effectuant directement avec le logiciel SUDAAN un test sur la différence des proportions. Les intervalles de confiance associés aux estimations de proportions sont fournis aux annexes B et C.

Analyse de régression logistique des facteurs associés à la qualité d'ensemble

Pour approfondir l'analyse des facteurs associés à la qualité d'ensemble des services de garde éducatifs offerts aux enfants dans chacun des types de services, une analyse de régression logistique a été entreprise de façon exploratoire. Ces régressions ont été réalisées en fonction des résultats issus des analyses bivariées.

Plus précisément, sauf exception³², toutes les variables associées à la variable d'intérêt au seuil de 10 %, sur la base des analyses bivariées, ont été conservées. Étant donné la nature exploratoire de l'analyse et le grand nombre de variables déjà prises en compte dans l'analyse bivariée, aucune variable additionnelle n'a été ajoutée à celles déjà retenues afin de vérifier notamment si elle pouvait jouer un rôle confondant. Compte tenu également des tailles d'échantillon restreintes aucun test d'interaction n'a été effectué entre les variables retenues.

On a procédé dans un premier temps à l'aide du logiciel SAS, sans tenir compte de l'effet de plan au moment des estimations de variances, mais tout en pondérant avec un poids à moyenne 1 (poids échantillonnaux). De

façon générale, les variables indépendantes ont été testées simultanément et non par blocs ou groupes de variables afin de déterminer lesquelles s'avéraient liées à la mesure de qualité au seuil de 0,05, une fois l'ensemble des variables prises en compte. Dans certains cas, l'analyse a toutefois été menée par étapes afin de dégager certaines pistes d'interprétation quant aux processus à l'œuvre (variables potentiellement confondantes) car les estimations des paramètres semblaient plus sensibles à l'ajout ou au retrait de certaines variables. Dans un deuxième temps, les modèles finaux incluant seulement les variables significatives au seuil de 0,05 ont été repris avec le logiciel SUDAAN afin de tenir compte de l'effet de plan de l'enquête. Aucune variable n'a alors été ajoutée; certains paramètres ont pu toutefois devenir alors non significatifs.

Signalons que comme les événements étudiés ne sont pas rares (le fait de fréquenter des services de garde éducatifs de meilleure qualité), il n'est pas recommandé d'interpréter les rapports de cote dérivés des analyses de régression logistique comme des risques relatifs; on considérera seulement que la probabilité d'occurrence de la variable dépendante est augmentée ou diminuée par un facteur donné (Hosmer et Lemeshow, 1989).

2.3.4.5 Choix des catégories des variables de croisement pour l'étude des facteurs associés

Pour l'analyse des facteurs associés, une mesure de qualité comportant deux catégories (voir plus loin dans le texte) a été mise en relation avec chacune des variables retenues. Les catégories des variables de croisement ont été établies selon une base théorique ou en fonction de considérations pratiques lorsque la nature de la variable le permettait. Les données sur le rapport enfants/éducatrice où il était possible de former des catégories en se basant sur la réglementation en vigueur en est un exemple. Toutefois, vu l'absence de telles références dans la majorité des cas, les seuils des variables de croisement ont été choisis en fonction de critères statistiques soit, plus précisément, à partir de la moyenne, de la médiane ou encore de l'étude des distributions. Par exemple, les catégories de la variable décrivant la taille des établissements ont été déterminées en fonction de seuils où se cumulent

suffisamment d'observations dans la distribution pour les fins de l'analyse. Cette méthode offre l'avantage de proposer des catégories neutres afin de noter si une association existe ou non entre la variable de croisement considérée et le fait de se classer dans l'une ou l'autre des deux catégories de qualité. Par contre, selon cette méthode, les catégories d'une même variable peuvent être différentes si les distributions par type de services le justifient.

2.4 Présentation des résultats

Cette section énonce brièvement les règles générales appliquées à la présentation des résultats, notamment la formulation générale des résultats et les modes de présentation. Rappelons que toutes les estimations de la qualité présentées dans ce rapport font référence à la qualité de l'expérience vécue par les enfants fréquentant l'un ou l'autre des cinq types de services de garde éducatifs retenus aux fins de l'enquête.

2.4.1 Niveaux de la qualité éducative selon les scores et les scores moyens

Il convient d'abord de préciser les termes utilisés dans les chapitres d'analyse pour décrire la qualité car ceux-ci varient selon les niveaux d'agrégation des données :

- ↓ SCORE : mesure de qualité calculée caractérisant une dimension ou une sous-dimension pour un groupe observé. Ce score est obtenu par la moyenne des cotes des items composant la dimension ou la sous-dimension. Exemple : « le score obtenu par un groupe observé à la sous-dimension 1.1. »
- ↓ SCORE TOTAL : mesure de qualité estimée caractérisant tous les items de l'échelle pour un groupe observé. Ce score est obtenu par la moyenne des cotes de tous les items. Exemple : « le score total obtenu par un groupe pour l'ensemble de l'échelle. »
- ↓ SCORE MOYEN D'ITEM : mesure de qualité estimée caractérisant un item pour l'ensemble des enfants fréquentant un même type de services. Ce score est obtenu par l'estimation de la moyenne des cotes de l'item. Exemple : « le score moyen à

l'item 112 estimé pour l'ensemble des poupons en installation de CPE. »

- ↓ SCORE MOYEN : mesure de qualité estimée caractérisant une dimension ou une sous-dimension pour l'ensemble des enfants fréquentant un même type de services. Ce score est obtenu par l'estimation de la moyenne des scores d'une dimension ou d'une sous-dimension. Exemple : « le score moyen de la dimension 3 estimé pour les enfants d'âge préscolaire en garderie. »
- ↓ SCORE MOYEN TOTAL : mesure de qualité estimée caractérisant tous les items de l'échelle pour l'ensemble des enfants fréquentant un même type de services. Ce score est obtenu par l'estimation de la moyenne des scores totaux. Exemple : « le score moyen total estimé pour l'ensemble des enfants en milieu familial. »

À la lumière des définitions précédentes, on remarque que tous les items ont un poids équivalent peu importe le type de score calculé. Ainsi, le score d'une dimension s'obtient simplement en calculant la moyenne des cotes de tous les items qui font partie de cette dimension.

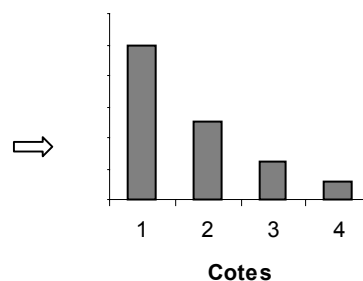
Les résultats décrivant les niveaux de qualité mesurés sont présentés sous forme de tableaux et de figures. Dans les chapitres subséquents, les tableaux présentent les scores moyens de qualité par dimension et sous-dimension. Les scores moyens par item sont fournis dans des tableaux complémentaires à la fin de chaque chapitre.

De plus, un indice de qualité en six classes accompagne les scores et les scores moyens afin de les décrire qualitativement. À partir de ces scores pour un univers mesuré, c'est-à-dire un item, une sous-dimension, une dimension ou pour l'ensemble d'une échelle, les classes de qualité indiquent donc dans quelle mesure ce qui est observé correspond aux principes qui sous-tendent le programme éducatif. Dans les cas de scores et de scores moyens de niveaux très bas, bas et moyen-bas la qualité n'est pas satisfaisante au regard des principes du programme éducatif; par contre, la qualité est satisfaisante par rapport à ces mêmes principes lorsque ces scores se situent aux niveaux moyen-élevé, élevé ou très élevé. Voici, plus spécifiquement, comment chacune de ces classes est définie.

QUALITÉ DONT LE NIVEAU NE CORRESPOND PAS AUX PRINCIPES DU PROGRAMME ÉDUCATIF

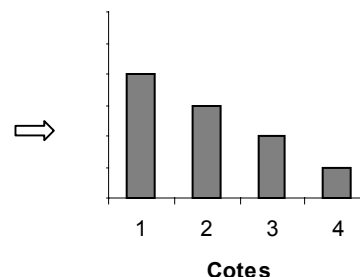
Niveau de qualité très bas

Un score ou un score moyen variant de 1,00 à 1,49 indique qu'en général pour un item, une sous-dimension, une dimension ou pour la totalité d'une échelle, la qualité éducative est *médiocre* au regard des principes du programme éducatif. Comme l'illustre la figure de la distribution typique des cotes pour l'ensemble des items de l'univers mesuré, le qualificatif très bas implique un très grand nombre de cotes de niveau 1 (inadéquat) et une diminution marquée du nombre de cotes à mesure que le niveau augmente, le nombre de cotes de niveau 4 (très bien) étant très peu élevé, voire nul. En conséquence, des changements majeurs et nombreux sont à faire pour améliorer la qualité éducative d'un univers mesuré qui se situe au niveau très bas de l'indice.



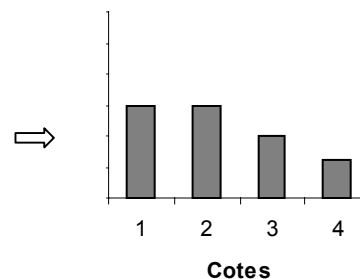
Niveau de qualité bas

Un score ou un score moyen variant de 1,50 à 1,99 indique qu'en général pour un item, une sous-dimension, une dimension ou pour la totalité d'une échelle, la qualité éducative est *très faible* au regard des principes du programme éducatif. L'illustration graphique révèle une décroissance assez constante des cotes pour l'ensemble des items de l'univers mesuré. Les items cotés 1 et 2 demeurent les plus nombreux alors qu'encore très peu d'items sont cotés 4, bien qu'ils soient en général plus nombreux qu'au niveau précédent. En conséquence, bien que l'on perçoive certaines tendances positives par rapport au niveau de qualité très bas, des changements majeurs et nombreux sont à faire pour améliorer la qualité éducative d'un univers mesuré qui se situe au niveau bas de l'indice.



Niveau de qualité moyen-bas

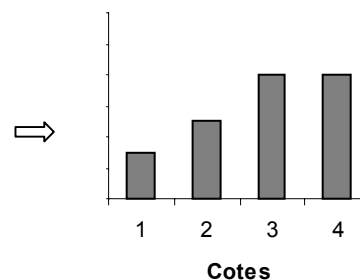
Un score ou un score moyen variant de 2,00 à 2,49 indique qu'en général pour un item, une sous-dimension, une dimension ou pour la totalité d'une échelle, la qualité éducative est *faible* au regard des principes du programme éducatif. L'illustration graphique de la distribution des cotes pour l'ensemble des items de l'univers mesuré se caractérise généralement par une majorité d'items cotés 1 ou 2, et, en comparaison, moins d'items cotés 3 ou 4. En conséquence, pour améliorer la qualité éducative d'un univers mesuré qui se situe au niveau moyen-bas de l'indice et la rendre satisfaisante au regard du programme éducatif, plusieurs changements majeurs restent toujours à faire quoiqu'ils soient moins nombreux qu'aux autres niveaux inférieurs.



QUALITÉ DONT LE NIVEAU CORRESPOND AUX PRINCIPES DU PROGRAMME ÉDUCATIF

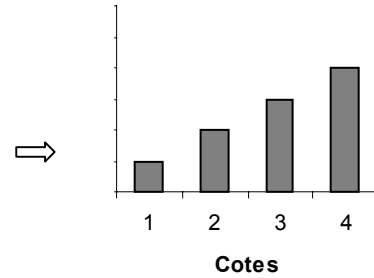
Niveau de qualité moyen-élevé

Un score ou un score moyen variant de 2,50 à 2,99 indique qu'en général pour un item, une sous-dimension, une dimension ou pour la totalité d'une échelle, la qualité éducative est *passable* au regard des principes du programme éducatif. L'illustration graphique de la distribution des cotes pour l'ensemble des items de l'univers mesuré montre dans ce cas qu'il y a plus d'items recevant la cote de 3 ou 4 que d'items cotés 1 ou 2. En conséquence, pour améliorer la qualité éducative d'un univers mesuré qui se situe au niveau moyen-élevé de l'indice, quelques changements sont à faire sans exclure la possibilité que certains de ceux-ci soient majeurs.



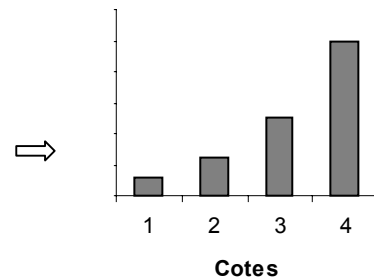
Niveau de qualité élevé

Un score ou un score moyen variant de 3,00 à 3,49 indique qu'en général pour un item, une sous-dimension, une dimension ou pour la totalité d'une échelle, la qualité éducative est *bonne* au regard des principes du programme éducatif. L'illustration type de la distribution des cotes pour l'ensemble des items de l'univers mesuré révèle ainsi une croissance constante du nombre d'items à mesure que le niveau de la cote augmente (de 1 à 4). En conséquence, assez peu de changements sont à faire pour améliorer la qualité éducative d'un univers mesuré qui se trouve au niveau élevé de l'indice, toujours sans exclure la possibilité qu'un petit nombre d'entre eux soient majeurs.



Niveau de qualité très élevé

Un score ou un score moyen variant de 3,50 à 4,00 indique qu'en général pour un item, une sous-dimension, une dimension ou pour la totalité d'une échelle, la qualité éducative est *très bonne* au regard des principes du programme éducatif. La distribution des cotes pour l'ensemble des items de l'univers mesuré se caractérise par une augmentation marquée du nombre d'items selon le niveau de la cote, allant de l'absence ou d'un très faible nombre d'items cotés 1 à un très grand nombre d'items cotés 4. En conséquence, très peu de changements sont à faire pour améliorer ou maintenir la qualité éducative d'un univers mesuré qui se situe au niveau très élevé de l'indice.



Cet indice à six classes est aussi utilisé pour la construction des histogrammes qui illustrent la répartition des enfants selon le niveau de qualité estimé pour l'ensemble de l'échelle, une dimension ou une sous-dimension, pour chacun des types de services couverts par l'enquête.

Ajoutons que les scores moyens sur lesquels s'appuie l'indice de qualité sont imprécis puisqu'ils représentent une valeur estimée. En conséquence, ils ne peuvent pas toujours être associés strictement à un seul niveau de l'indice de qualité. Pour tenir compte de cette imprécision, les tableaux présentent la marge d'erreur de chacun des scores moyens. De plus, pour les scores moyens dont la marge d'erreur signifie qu'ils chevauchent plus d'un niveau de qualité, une note précise que le score moyen tend vers la catégorie supérieure (\Rightarrow) ou inférieure (\Leftarrow) au niveau indiqué dans le tableau. Pour rendre compte de cette imprécision dans le texte, on fait, au besoin, référence directement aux seuils délimitant les classes de l'indice tout en

regroupant les scores moyens dans une seule classe selon la valeur estimée afin d'alléger la syntaxe et de faciliter la compréhension. Enfin, par souci de précision, nous fournissons à l'annexe B les intervalles de confiance associés à chacune des estimations présentées dans les histogrammes de fréquence qui illustrent les proportions d'enfants concernés par chaque niveau de qualité.

2.4.2 Étude des facteurs associés

Suivant l'objectif de déterminer les facteurs associés à la qualité, les scores pour l'ensemble de l'échelle ainsi que pour chacune des quatre dimensions sont croisés par analyse bivariée avec un ensemble de caractéristiques qu'il s'avère pertinent d'étudier (voir chapitre 1). Toutefois, étant donné les effectifs restreints de certains types de services, nous avons opté pour la division des scores de qualité en deux catégories en tenant compte de leur distribution respective :

- ↓ **33 % inférieur** : une première catégorie regroupe le tiers inférieur des scores de qualité au sein d'une distribution de scores pour l'ensemble de l'échelle ou pour une dimension. Le terme « tercile inférieur »³³ est parfois employé pour désigner cette catégorie.
- ↓ **67 % supérieur** : en conséquence, la deuxième catégorie regroupe le reste des scores de qualité au sein d'une distribution de scores pour l'ensemble de l'échelle ou pour une dimension, soit les deux tiers supérieurs.

Contrairement à une méthode qui aurait fait appel aux classes de l'indice de qualité, cet indice dichotomique permet d'améliorer la puissance des tests du khi-carré puisque l'on s'assure d'obtenir un nombre relativement élevé et constant d'observations dans chacune des deux catégories de la qualité. Ainsi, on augmente les chances de noter des liens significatifs. Il faut cependant mentionner que la variable catégorique ainsi créée ne peut être utilisée que pour l'étude des facteurs associés; elle ne peut être employée pour comparer des dimensions ou des sous-dimensions à l'intérieur d'un même type de services car les seuils varient en fonction des scores obtenus³⁴.

De plus, une attention particulière est portée à l'inférence faite à la suite des tests du khi-carré. Dans l'interprétation des résultats, on dira, par exemple, qu'il y a un lien entre la variable de croisement « x » et la fréquentation par les enfants de services de garde éducatifs de *moindre* qualité. On ne peut toutefois pas dire qu'il y a un lien entre la variable « x » et la fréquentation par les enfants de services de garde éducatifs de *faible* qualité, car le tercile inférieur rassemble les résultats correspondant à une moins bonne qualité par rapport à la moyenne et non pas nécessairement des scores équivalant à un niveau de qualité faible. La nuance est importante car le seuil délimitant les deux catégories de l'indice de qualité varie selon le type de services et la dimension de la qualité éducative prise en compte.

Dans l'analyse bivariée, tous les facteurs associés à la qualité au seuil de signification de 5 % ($p < 0,05$) sur la base du test du khi-carré sont présentés sous forme de

tableaux. De plus, les associations observées au seuil de 10 % ($0,05 \leq p < 0,10$) sont évoquées dans le texte sous forme de tendances lorsqu'elles sont jugées intéressantes ou lorsqu'elles apparaissent complémentaires à des résultats significatifs au seuil de 5 %. Cette approche est justifiée par le fait que l'étude porte sur un domaine exploratoire et complexe pour lequel il pourrait s'avérer pertinent de définir des pistes de recherche même si, ce faisant, la probabilité que l'association révélée soit fautive est plus élevée.

2.4.3 Autres éléments de forme

- ↓ Les scores moyens de qualité présentés dans le texte, les tableaux et les figures ont été arrondis à deux décimales. Quant aux proportions, elles ont été arrondies à l'unité lorsqu'elles sont mentionnées dans le texte, et à une décimale dans les tableaux et figures, à l'exception des proportions inférieures à 5 % pour lesquelles une décimale a aussi été conservée dans le texte.
- ↓ Dans ce rapport, lorsque l'on fait référence aux items des échelles, la numérotation omet les points afin d'alléger le texte et d'en faciliter la lecture; par exemple, un item « 2.2.2 » dans une échelle est plutôt numéroté « 222 » dans le texte et les tableaux.
- ↓ Dans certains cas, il n'est pas pertinent de présenter les données sous forme de tableaux ou de figures. C'est le cas notamment des associations dont le seuil de signification se situe entre 5 % et 10 %. Pour bien préciser que ces données n'apparaissent pas dans les tableaux et figures publiés, la mention « données non présentées » accompagne de tels résultats lorsqu'ils sont évoqués dans le texte.
- ↓ Dans les tableaux et les figures, une estimation de 0 % signifie que dans l'échantillon aucun enfant ne présentait les caractéristiques décrites. Toutefois, il n'est pas exact d'inférer ce résultat à la population en disant qu'aucun enfant ne présente cette caractéristique car il y a un intervalle de confiance associé à cette estimation. En conséquence, dans la population visée, il pourrait y avoir une très faible

proportion d'enfants ayant ces mêmes caractéristiques. Afin de refléter cette possibilité, il convient donc de mentionner *une proportion infime* d'enfants plutôt que de dire *aucun* enfant.

2.5 Portée et limites de l'enquête

En raison des taux de réponse obtenus et de la taille considérable de son échantillon par rapport à la taille du réseau québécois de services de garde éducatifs régis, l'enquête *Grandir en qualité 2003* assure que les données produites sont représentatives de la qualité de l'expérience vécue par les enfants dans chacun des cinq types de services visés par l'étude.

Sur le plan méthodologique, plusieurs moyens ont été utilisés pour assurer cette représentativité. Une pondération a été attribuée à l'ensemble des instruments de manière à réduire l'impact de la non-réponse totale et donc à assurer la fiabilité de l'inférence à la population visée. De plus, toutes les mesures de précision et les tests ont été produits en prenant en considération le plan de sondage de l'enquête.

En soi, cette capacité à produire des données sur la qualité des services de garde éducatifs au Québec selon différentes catégories de services constitue un pas important pour l'avancée des connaissances dans le domaine. De plus, nous croyons que d'autres particularités du projet en font une enquête dont les retombées sont larges et particulièrement utiles pour les intervenants du réseau de services de garde éducatifs à la petite enfance. En premier lieu, on note l'importance du lien étroit entre les mesures de qualité et les principes qui sous-tendent le programme éducatif mis de l'avant par le MESSF. Ce lien aura permis de produire des données qui correspondent aux principes sous-jacents au programme éducatif lesquels sont cohérents avec les éléments de la réglementation et les démarches de sensibilisation et de soutien du MESSF. Plus particulièrement, le recours au programme éducatif comme guide pour la définition de la qualité aura permis de créer des indices de la qualité selon chacun des domaines d'application que sont *la structuration des lieux, la structuration et la variation des types d'activités et les interactions des éducatrices*

et des RSG avec les enfants et avec les parents. De plus, les indices de qualité des diverses activités de base (l'accueil, les repas, les soins personnels, les jeux extérieurs et la fin de la journée) permettent une étude plus spécifique de la qualité éducative selon des moments ou des activités quotidiennes récurrentes. Dans son ensemble, l'enquête *Grandir en qualité 2003* propose donc une diversité de mesures et de nuances qui sont autant de pistes d'ajustement pour l'amélioration continue de la qualité des services de garde éducatifs au Québec.

L'enquête a aussi permis de cerner de nombreux facteurs liés à la qualité d'ensemble et des différentes dimensions, ce qui permettra sans doute de mieux peser les actions à prendre pour soutenir le développement de la qualité sous plusieurs formes. Il convient cependant de souligner que les tests d'association servent uniquement à établir s'il existe une association entre deux variables. En raison notamment de la nature transversale de l'enquête, on ne peut présumer de liens de causalité entre les variables de croisement retenues et la qualité des services de garde éducatifs offerts aux enfants québécois.

Sur le plan de la procédure de collecte, de nombreux choix se sont avérés avantageux et augmentent indirectement la portée des résultats de l'enquête. Il s'agit particulièrement du choix de couvrir l'ensemble d'une journée, de l'arrivée au départ des enfants, de la décision de recourir presque exclusivement à l'observation directe et du choix de recourir à une cotation à l'aveugle limitant ainsi la subjectivité des observatrices.

Comme dans toute enquête, en particulier les enquêtes réalisées en face-à-face, l'enquête *Grandir en qualité 2003* n'est pas exempte de biais de désirabilité sociale. Par exemple, il est possible que certains membres des établissements visités aient légèrement modifié leur environnement ou leurs comportements afin de révéler des conditions qu'ils considèrent bénéfiques pour l'évaluation faite. Par contre, on doit souligner l'exhaustivité des instruments de mesure mis au point, la formation solide donnée aux observatrices et toutes les autres mesures mises en place afin d'assurer la

production d'estimations les plus valides et les plus fiables possible.

Notons également que la période de collecte implique possiblement un biais puisqu'elle ne reflète pas les conditions en cours sur une année complète. Par exemple, il est possible que les périodes hivernale et printanière durant lesquelles s'est déroulée l'enquête engendre plus de complexité dans la gestion des arrivées et des départs en raison de l'habillement des enfants; que le matériel disponible à l'extérieur soit moins nombreux qu'en période estivale. Inversement, il est probable qu'à ce stade dans l'année l'éducatrice ou la RSG connaisse fort bien les enfants sous sa responsabilité et ait déjà établi une bonne complicité avec eux contrairement à la situation qui aurait cours à l'automne, au moment où plusieurs groupes se forment. Parallèlement, comme dans la grande majorité des études du même genre, le recours à une seule journée d'observation et dans un seul groupe d'enfants par classe d'âge limite sans doute le potentiel de généralisation.

Si le développement de nouveaux instruments d'observation s'est avéré avantageux pour mesurer la qualité de l'implantation du programme éducatif, ce choix constitue toutefois une limite sur le plan de la comparabilité des résultats avec d'autres études portant sur la qualité réalisées ici ou ailleurs. De même, au sein de l'enquête, il est difficile de comparer directement les scores de qualité à partir des trois échelles (préscolaire, pouponnière et milieu familial). Les critères ayant servi à la construction des indices de qualité diffèrent d'une à l'autre puisque celles-ci ont été créées justement pour tenir compte des particularités de chacun de ces types de services. Cependant, cette limite est aussi observée dans les autres études qui ont fait l'examen de la qualité de plusieurs types de services. Notons, de plus, que la conception des échelles d'observation est basée sur les mêmes construits théoriques relativement à la qualité éducative, ce qui permet des comparaisons d'ordre général.

Notes

1. Il s'agit en réalité du nombre de places au permis. Dans le but d'alléger le texte, nous utiliserons plutôt le terme « enfants » pour les désigner.
2. Rappelons qu'au moment de l'enquête, le ministère de la Famille et de l'Enfance (MFE) agissait comme responsable des services de garde éducatifs; aujourd'hui ceux-ci sont sous la responsabilité du ministère de l'Emploi, de la Solidarité sociale et de la Famille (MESSF).
3. Ainsi, les services devaient être en opération au moins depuis le 15 octobre 2001 pour être admissibles à l'enquête, afin d'éviter de mesurer la qualité dans des services en implantation.
4. Le CPE est un établissement qui fournit des services de garde éducatifs dans une installation et qui coordonne de tels services en milieu familial pour des enfants de la naissance à la fréquentation de la maternelle. Le CPE est administré par un conseil d'administration composé dans une proportion d'au moins les deux tiers de parents.
5. Une garderie est un établissement qui fournit des services de garde éducatifs dans un local dont le permis est généralement détenu par une personne physique ou une corporation à but lucratif. Dans de plus rares cas, ce permis est détenu par une municipalité, une commission scolaire ou un organisme sans but lucratif. Dans les deux situations, le conseil d'administration n'est pas nécessairement composé d'une majorité de parents.
6. Depuis 2002, un titulaire de permis peut obtenir une dispense afin d'offrir un seul des deux volets.
7. L'article 4 du *Règlement sur les centres de la petite enfance* [C-8.2, r.2] et l'article 5 du *Règlement sur les garderies* [C-8.2, r.5.1] précisent les détails quant aux classes d'âge couvertes par les services à la petite enfance du MESSF.
8. CAFE : *Conception administrative famille-enfance*.
9. La stratification assure la représentation dans l'échantillon de tous les éléments de la population. Plus les groupes formés sont homogènes, plus les estimations sont précises. La sélection de l'échantillon se fait de façon indépendante dans chaque groupe, aussi appelé strate.
10. L'effet de plan est le quotient de la variance estimée en tenant compte du plan de sondage par la variance obtenue à partir d'un plan de sondage aléatoire simple de même taille.
11. Ce protocole est consultable auprès des auteurs du rapport.
12. Les échelles utilisées dans le cadre de l'enquête sont consultables auprès des auteurs du rapport. Des versions révisées font l'objet d'une publication par le MESSF (Bourgon, Lavallée et collabo-rateurs, 2004a, 2004b, 2004c).
13. Les critères déterminant la cote à chacun des items peuvent être obtenus auprès des auteurs.
14. Les observatrices ont été informées que le fait de cocher au moins un élément à proscrire ou à éviter engendrait une cote plus sévère pour l'item mesuré. Bien que cette situation implique moins d'objectivité de la part des observatrices, en retour cela leur permettait de mieux évaluer l'impact de leurs décisions.

15. L'*EEEP-R* est plus souvent désignée par son abréviation anglaise *ECERS (Early Childhood Environment Rating Scale)*.
16. Le cahier technique de l'enquête résume les résultats obtenus à chacune de ces étapes de la construction des échelles d'observation. Celui-ci est consultable auprès des auteurs du rapport.
17. Notons que les valeurs des *kappas pondérés* peuvent varier de -1 à 1 . Il est recommandé d'interpréter le coefficient selon la grille d'évaluation suivante : < 0 = accord nul; $0,01$ à $0,19$ = accord faible; $0,20$ à $0,39$ = accord suffisant; $0,40$ à $0,59$ = accord modéré; $0,60$ à $0,79$ = accord élevé; $0,80$ à $1,00$ = accord presque parfait (Landis et Koch, 1977).
18. Le nombre d'observatrices ayant participé à cet exercice est de 26 puisque 2 des observatrices initiales ne travaillaient plus au projet.
19. L'échelle proposée par Murphy et Davidshofer (1994) est la suivante : Ω $0,60$ = fidélité inadéquate; $0,60$ à $0,69$ = fidélité faible; $0,70$ à $0,79$ = fidélité moyenne; $0,80$ à $0,89$ = fidélité très bonne; $0,90$ à $1,00$ = fidélité élevée.
20. Parallèlement à l'analyse de consistance des dimensions, le degré de discrimination des items, et donc leur niveau d'homogénéité par rapport à l'ensemble des items d'une même dimension, a aussi été évalué. Ces résultats sont présentés à l'annexe A.
21. Les questionnaires de l'enquête sont présentés à l'annexe D.
22. Dans le cas des groupes comptant plus d'une éducatrice attitrée (groupes fusionnés et rassemblements), le questionnaire n'était administré qu'à l'une d'entre elles, soit à celle disponible au moment désigné (durant la période de la sieste). L'entrevue menée en complément à l'observation était également réalisée avec celle-ci.
23. Les questionnaires de ces rapports sont présentés à l'annexe E avec indication des questions utilisées dans le cadre de la présente enquête.
24. L'annexe F fournit la liste et les définitions des variables utilisées dans l'enquête.
25. La gestionnaire d'une seule garderie d'enfants d'âge préscolaire n'a pas répondu au questionnaire qui lui était adressé.
26. Le cahier technique de l'enquête, consultable auprès des auteurs, fournit de plus amples renseignements sur l'impact des items non observés.
27. Pour une présentation plus détaillée des calculs d'ajustement pour la non-réponse, voir l'annexe A.
28. Les journées d'observation se sont déroulées selon trois vagues successives (vague 1 : 3 mars au 4 avril 2003, vague 2 : 7 avril au 2 mai 2003 et vague 3 : 5 mai au 5 juin 2003).
29. L'effet de plan est principalement causé par la variation des poids échantillonnaires en raison du fait que la base de sondage n'était pas récente et que le nombre d'enfants par groupe était variable. Toutefois, pour le milieu familial, l'effet de plan plus élevé est aussi lié au fait que le nombre moyen d'enfants par RSG, estimé à six, était parfois sous-estimé à cause de la présence d'une assistante permettant ainsi d'augmenter le nombre d'enfants dans le milieu familial. L'effet de plan pour les groupes de poupons en garderie est plus faible car le nombre de garderies échantillonnées est élevé par rapport à la population visée; ainsi, il est possible de tenir compte de l'effet de population finie dans le calcul de la variance.
30. Les poids échantillonnaires, aussi appelés poids à moyenne 1, sont proportionnels aux poids populationnels et leur somme correspond au nombre de répondants à l'instrument.

31. Deux approches ont été utilisées pour faire cet ajustement. La première, réalisée avec le logiciel SAS à partir des données pondérées par les poids échantillonnaires, consiste à diviser la statistique habituelle du test par l'effet de plan moyen dans un tableau croisé de fréquences. Puis, si un test plus exact s'avérait nécessaire en raison de la proximité du seuil de 5 %, la correction de Satterthwaite pour le test du khi-carré était faite à l'aide du logiciel SUDAAN.
32. Ainsi, dans certains types de services, les variables trop fortement corrélées à d'autres variables ou celles présentant un taux de non-réponse partielle plus élevé, elles-mêmes souvent liées à d'autres variables du modèle, ont dû être exclues. Ces variables sont mentionnées dans chacun des chapitres d'analyse.
33. Les quantiles sont des mesures de position qui ne tentent pas nécessairement de déterminer le centre d'une distribution d'observations, mais de décrire une position particulière.
34. Les seuils délimitant les deux catégories de qualité sont présentés dans chacun des chapitres d'analyse.

Bibliographie

- BELLEMARE, G., A.-R. GRAVEL et L. BRIAND (2002). *La qualité des services de garde à l'enfance. Un bilan de la littérature. Cahier de recherche de l'Alliance de recherche universités-communautés en économie sociale*, Hull, Université du Québec en Outaouais, R-11-2002, 56 p.
- BOURGON, L., C. LAVALLÉE et collaborateurs (2003a). *Échelle d'observation de la qualité éducative. Le service de garde préscolaire*, version utilisée dans l'enquête *Grandir en qualité 2003*, Québec, Ministère de la Famille et de l'Enfance, Gouvernement du Québec, 82 p.
- BOURGON, L., C. LAVALLÉE et collaborateurs (2003b). *Échelle d'observation de la qualité éducative. Le service de garde en milieu familial*, version utilisée dans l'enquête *Grandir en qualité 2003*, Québec, Ministère de la Famille et de l'Enfance, Gouvernement du Québec, 99 p.
- BOURGON, L., C. LAVALLÉE et collaborateurs (2003c). *Échelle d'observation de la qualité éducative. Le service de garde en pouponnière*, version utilisée dans l'enquête *Grandir en qualité 2003*, Québec, Ministère de la Famille et de l'Enfance, Gouvernement du Québec, 87 p.
- BOURGON, L., C. LAVALLÉE et collaborateurs (2004a). *Échelle d'observation de la qualité éducative. Les services de garde en installation pour les enfants de 18 mois ou plus*, Québec, Ministère de l'Emploi, de la Solidarité sociale et de la Famille, à paraître.
- BOURGON, L., C. LAVALLÉE et collaborateurs (2004b). *Échelle d'observation de la qualité éducative. Les services de garde en milieu familial*, Québec, Ministère de l'Emploi, de la Solidarité sociale et de la Famille, à paraître.
- BOURGON, L., C. LAVALLÉE et collaborateurs (2004c). *Échelle d'observation de la qualité éducative. Les services de garde en pouponnière*, Québec, Ministère de l'Emploi, de la Solidarité sociale et de la Famille, à paraître.
- CRONBACH, L. J. (1951). « Coefficient alpha and the internal structure of tests », *Psychometrika*, vol. 16, n° 3, p. 297-334.
- HARMS, T., R. M. CLIFFORD et D. CRYER (1998). *Early Childhood Environment Rating Scale*, New York, Teachers College Press, 60 p. (traduction : BAILLARGEON, M., et H. LAROCHE (1998). *Échelle d'évaluation de l'environnement préscolaire*, Sainte-Foy, Presses de l'Université du Québec, 74 p.).
- HOSMER, D. W., et S. LEMESHOW (1989). *Applied Logistic Regression*, New York, Wiley, 307 p.
- KASS, G. V. (1980). « An explanatory technique for investigating large quantities of categorical data », *Applied statistics*, vol. 29, p. 119-127.
- LANDIS, J., et G. G. KOCH (1977). « The measurement of observer agreement for categorical data », *Biometrics*, vol. 33, p. 159-174.
- MURPHY, K. R., et C. O. DAVIDSHOFER (1994). *Psychological Testing. Principles and Applications*, 3^e édition, Englewood Cliffs (N.J.), Prentice-Hall, 548 p.
- SHAH, B. V., B. G. BARNWEL et G. S. BIELER (1997). *Sudaan User's Manual, version 7.5*, Research Triangle Park (N.C.), Research Triangle Institute.
- TREMBLAY, S. (2003). *Enquête Grandir en qualité. Recension des écrits sur la qualité des services de garde*, Québec, Ministère de la Famille et de l'Enfance, Direction de la recherche, de l'évaluation et de la statistique, 37 p.

YATES, F., et P. M. GRUNDY-SEN (1953). « Selection without Replacement from within Strata with Probability Proportional to Size », *Journal of the Royal Statistical Society, Série B*, vol. 15, p. 253-261.