



CAPACITÉ D'APPRENTISSAGE DANS LES COMMUNAUTÉS FRANCOPHONES EN SITUATION MINORITAIRE

RAPPORT DES RÉSULTATS DE LA PHASE PRÉSCOLAIRE

Conseil d'administration de la SRSA

Richard A. Wagner
Associé principal, Norton Rose Fulbright S.E.N.C.R.L., s.r.l.

Maria David-Evans
Présidente sortante de l'IAPC et ancienne sous-ministre,
Gouvernement de l'Alberta

Robert Flynn
Professeur émérite, École de psychologie, Université d'Ottawa

John Helliwell
Codirecteur, programme Interactions sociales, identité et
mieux-être, Institut canadien de recherches avancées

Suzanne Herbert
Ancienne sous-ministre, Gouvernement de l'Ontario

Guy Lacroix, Ph. D.
Professeur d'économie, Université Laval

Renée F. Lyons, Ph. D.
Présidente de la recherche sur les maladies chroniques
complexes et directrice scientifique de TD du Bridgepoint
Collaboratory for Research and Innovation, Université de Toronto

Sharon Manson Singer
Ancienne présidente des Réseaux canadiens de recherche en
politiques publiques

Jim Mitchell
Partenaire fondateur de la firme d'experts-conseils en politiques
Sussex Circle

Président et chef de la direction de la SRSA

Jean-Pierre Voyer

Auteurs

Louise Legault, Ph. D.

Glenn Thompson, Ph. D.

Danielle Patry, Ph. D.

Robyn Carson

Lorraine Lefebvre

Paul Lalonde

La Société de recherche sociale appliquée (SRSA) est un organisme de recherche sans but lucratif, créé dans le but précis d'élaborer, de mettre à l'essai sur le terrain et d'évaluer rigoureusement de nouveaux programmes. Notre mission, qui comporte deux volets, consiste à aider les décideurs et les intervenants à déterminer les politiques et programmes qui améliorent le bien-être de tous les Canadiens, en se penchant particulièrement sur les effets qu'ils auront sur les personnes défavorisées, et à améliorer les normes relatives aux éléments probants utilisées pour évaluer ces politiques.

Depuis sa création en décembre 1991, la SRSA a réalisé plus de 200 projets et études pour différents ministères fédéraux et provinciaux, des municipalités ainsi que d'autres organismes publics et sans but lucratif. La SRSA a des bureaux à Ottawa, à Toronto et à Vancouver.

Promoteur du projet Capacité d'apprentissage

Le projet Capacité d'apprentissage a été financé par Emploi et Développement social Canada.



Le rapport d'analyse contenu dans la présente publication est la responsabilité des auteurs et ne représente pas le point de vue du Gouvernement du Canada.

Table des matières

	Page
Table des matières	i
Tableaux et figures	v
Sommaire.....	1
1. Présentation du projet Capacité d'apprentissage dans les communautés francophones en situation minoritaire	3
1.1. Genèse du projet	3
1.2. Le projet Capacité d'apprentissage	5
1.3. L'apport du projet	6
1.4. Le modelage du développement des enfants en milieu minoritaire	7
1.4.1. Le microsystème : environnement familial (variables contextuelles)	8
1.4.2. Le microsystème : environnement familial (processus familiaux)	10
1.4.3. Le microsystème : milieu de garde.....	11
1.4.4. Le mésosystème : liens entre la famille et le milieu de garde	12
1.4.5. Le macrosystème : variables communautaires	13
1.5. Une définition de la préparation à la scolarisation	14
1.6. Le présent rapport	14
2. Activités de mise en œuvre.....	17
2.1. Recrutement des communautés.....	17
2.2. Recrutement des familles participantes.....	18
2.3. Effectifs humains requis pour la mise en œuvre	19
2.3.1. Les champions.....	19
2.3.2. Les coordonnatrices communautaires	19
2.3.3. Les éducatrices	19
2.3.4. Les intervenantes en alphabétisation familiale.....	20
2.3.5. Les formatrices.....	21
2.3.6. Les évaluatrices	21
2.4. Formation.....	22
2.4.1. Formation de base des éducatrices pour le programme en garderie	22
2.4.2. Formations de suivi des éducatrices pour le programme en garderie.....	22
2.4.3. Formation de base pour le programme des Ateliers familles	23
2.4.4. Formation de suivi pour les Ateliers familles.....	23
2.5. Description du volet en garderie	24
2.5.1. Déroulement du programme en garderie.....	25
2.6. Description du volet Ateliers familles.....	26
2.6.1. L'arrimage des deux volets du programme	27
2.6.2. Déroulement du programme Ateliers familles	27
2.6.3. Centre de ressources — achat de matériel.....	29
3. Méthodologie.....	31
3.1. Population ciblée.....	31
3.2. Devis expérimental	31
3.3. Validité interne	32
3.3.1. La taille de l'échantillon.....	32
3.3.2. Stratégie d'échantillonnage ciblé	32
3.3.3. Les mesures pré-intervention	33
3.3.4. Attrition.....	34
3.3.5. Prévention de la contamination des groupes témoins.....	34

3.4. Échantillon par communautés et par groupes expérimentaux.....	35
3.4.1. Taux de rétention.....	37
3.5. Mesures.....	37
3.5.1. Plan des collectes de données.....	37
3.5.2. Mesures auprès des enfants.....	39
3.5.3. Mesures auprès des parents.....	45
3.5.4. Connaissances, attitudes et croyances des parents.....	49
3.5.5. Dosage.....	51
3.6. Hypothèses à l'étude.....	52
3.6.1. L'impact du volet en garderie sur les dimensions linguistiques des enfants.....	52
3.6.2. L'impact du volet en garderie sur les dimensions identitaire et culturelle des enfants.....	53
3.6.3. L'impact du volet Ateliers familles sur les attitudes et comportements des parents.....	53
4. Méthodologie employée pour l'étude de la mise en œuvre.....	55
4.1. Modèle d'analyse des données qualitatives pour l'évaluation de la mise en œuvre.....	55
4.2. Outils de collecte de données.....	56
4.2.1. Étude de la mise en œuvre du volet en garderie.....	56
4.2.2. Étude de la mise en œuvre des Ateliers familles.....	60
4.3. Méthode d'analyse des données.....	64
4.3.1. Stratégies d'analyse.....	64
5. Résultats de la mise en œuvre du programme en garderie.....	67
5.1. Stratégie d'analyse des observations en garderie.....	68
5.2. Qualité Structurale et différenciation des garderies observées.....	70
5.3. Éléments structurels en place dans les milieux observés : fidélité et différenciation.....	72
5.4. Contenu éducatif : fidélité, qualité et différenciation du contenu des activités.....	75
5.5. La qualité des interactions entre les éducatrices et les enfants.....	82
5.6. Dosage.....	83
5.7. Réactions des éducatrices du groupe programme.....	84
5.8. Facilitateurs et défis.....	85
5.9. Retombées perçues.....	89
5.10. Sommaire.....	91
6. Étude de la mise en œuvre du programme d'Ateliers familles.....	95
6.1. La fidélité du contenu abordé dans les Ateliers familles.....	96
6.2. La qualité de l'animation des Ateliers familles.....	98
6.3. Dosage.....	100
6.4. Réactions des parents lors des ateliers.....	102
6.5. Facilitateurs et obstacles.....	104
6.6. Retombées perçues.....	106
6.7. Sommaire.....	108
7. Analyses préliminaires.....	111
7.1. Processus du contrôle de la qualité.....	111
7.2. Analyse des valeurs manquantes.....	112
7.2.1. Répartition des valeurs manquantes.....	112
7.2.2. Patron des données manquantes.....	114
7.2.3. Stratégie d'imputation des données.....	118
7.3. Variables confusionnelles.....	121
7.3.1. Associations avec les groupes expérimentaux et les résultantes des enfants.....	121
7.3.2. Effet instable des covariables dans le temps.....	122
7.4. Représentativité de l'échantillon (projet Capacité d'apprentissage versus EVMLO).....	123
7.4.1. Problématique.....	123
7.4.2. Statut d'immigrant et profil linguistique.....	124
7.4.3. Caractéristiques sociodémographiques.....	127
7.5. Sommaire des implications pour les analyses d'impact.....	131

8. Approche adoptée pour les analyses	135
8.1. Modélisation linéaire hiérarchique (MLH)	135
8.2. La méthode différence des différences	136
8.3. Conceptualisation du traitement et du dosage	137
8.3.1. Exposition hétérogène au traitement	137
8.3.2. Évaluation de l'effet du dosage ou de l'exposition	138
8.3.3. Évaluation de l'effet de la fidélité/qualité du programme en garderie	140
8.4. Contrôle de biais dans le cadre d'une étude longitudinale	142
8.4.1. Changements dans la composition des groupes	142
8.4.2. Pentes développementales non-parallèles	143
8.4.3. Validité des impacts basés sur une variable autre que le groupe	143
8.4.4. Ateliers familles : Un cas spécial	144
8.4.5. Tests de modération et de médiation	146
8.5. Plan des chapitres à suivre	147
9. Impact du programme testé sur les enfants	149
9.1. Traitement des variables	150
9.1.1. Mesures de résultantes	150
9.1.2. Variables cibles et covariables	153
9.1.3. Contre-validation non-paramétrique	156
9.2. Lors de la première année	157
9.2.1. Détails techniques	157
9.2.2. Analyses par groupes	159
9.2.3. Analyses par dosage	168
9.2.4. Analyses par qualité et fidélité du programme en garderie	176
9.2.5. Analyses par profil linguistique	185
9.3. Lors de la deuxième année	189
9.3.1. Détails techniques	189
9.3.2. Analyses par groupe	191
9.3.3. Analyses par dosage	200
9.3.4. Analyses par qualité et fidélité du programme en garderie lors de la première année	204
9.3.5. Analyses par profil linguistique	213
9.4. Sommaire	220
10. Impact du volet Ateliers familles	223
10.1. Impact direct sur les parents	223
10.1.1. Variables retenues pour l'analyse	223
10.1.2. Analyses par groupe	224
10.1.3. Analyses portant sur les parents du groupe programme	232
10.1.4. Sommaire des analyses d'impact sur les parents	241
10.2. Impact indirect sur les enfants	241
10.2.1. Lors de la première année	241
10.2.2. Lors de la deuxième année	242
10.3. Sommaire	243
11. Discussion	245
11.1. La mise en oeuvre du projet	245
11.1.1. Le programme en garderie	246
11.1.2. Les Ateliers familles	247
11.1.3. Recommandations pour la mise en oeuvre du programme à l'avenir	249
11.2. l'impact du programme testé	249
11.2.1. Analyse par groupe de l'impact du programme	250
11.2.2. Programme en garderie versus Ateliers familles	251
11.2.3. Le profil linguistique des enfants	253
11.2.4. Limites des résultats et prochains travaux	255
11.3. Conclusion	257
Références	259

Annexe A: Calendrier et outils d'évaluation	271
Annexe B : Contenu des trois versions de l'ÉPE-AD.....	273
Annexe C : Sommaire des thèmes abordés dans les Ateliers familles.....	279
Annexe D : Procédure pour l'administration de l'ÉPE-AD (Mesure pré-intervention).....	281
Annexe E : Validation de l'imputation de l'ÉPE-AD	283
Annexe F : Résultat de l'analyse des variables confusionnelles.....	285
Annexe G : Résultats de l'analyse des échelles ÉVIP-R et EOWPVT-F.....	295
Annexe H : Résultats de l'analyse des échelles de l'ÉPE-AD (deuxième année).....	297
Annexe I : Résultats de l'analyse des impacts en fonction du profil linguistique.....	299
Annexe J : Résultats de l'analyse de la fréquence et de la langue des activités de littératie.....	303

Tableaux et figures

	Page
Figure 1.1 : Modèle écologique complet de Bronfenbrenner (1979)	8
Tableau 2.1 : Calendrier de prestation du programme préscolaire	26
Tableau 3.1 : Répartition des participants par communauté depuis l'inscription	36
Tableau 3.2 : Répartition des participants par groupe expérimental.....	37
Tableau 3.3 : Raisons justifiant le retrait de l'enfant du projet Capacité d'apprentissage	37
Tableau 3.4 : Taux de réponses des sondages auprès des parents	38
Tableau 3.5 : Taux de réponses des sondages liés aux Ateliers familles	39
Tableau 3.6 : Taux de réponses des évaluations auprès des enfants	39
Tableau 3.7 : Alpha de Cronbach des domaines de l'ÉPE-AD par vague d'évaluation.....	41
Tableau 3.8 : Alpha de Cronbach des sous-échelles de vocabulaire de l'ÉPE-AD par vague d'évaluation ...	43
Tableau 3.9 : Coefficients de corrélation Pearson des échelles de l'ÉPE-AD et des échelles de vocabulaire expressif et réceptif	45
Tableau 3.10 : Alpha de Cronbach de l'échelle des activités de littératie et de l'échelle des langues des activités d'alphabétisation à chaque période de sondage	49
Tableau 4.1 : Résultats fictifs pour la matrice sur la fidélité structurelle.....	65
Tableau 5.1 : Périodes d'observation par année	69
Tableau 5.2 : Nombre de classes faisant partie du projet selon le groupe expérimental, l'année d'implémentation et la cohorte	69
Tableau 5.3 : Score moyen de la qualité structurelle des milieux observés selon les dimensions évaluées et le groupe expérimental.....	71
Tableau 5.4 : Pourcentage moyen de fidélité des éléments structurels selon l'année de la prestation du programme et le groupe expérimental	75
Tableau 5.5 : Scores moyens pour la qualité du contenu éducatif selon les dimensions évaluées et le groupe expérimental	80
Tableau 5.6 : Pourcentages moyens de la fidélité du contenu éducatif selon l'année de la prestation du programme et selon le groupe expérimental	82
Tableau 5.7 : Dosage en heures par semaine que les enfants ont fréquenté la garderie selon le groupe expérimental et cohorte.....	84
Tableau 6.1 : Contenu abordé selon le groupe de parents	97
Tableau 6.2 : Contenu abordé selon l'atelier observé.....	98
Figure 6.1 : Présence des parents selon l'atelier et le site	101
Figure 6.2 : Taux de participation des parents selon l'atelier.....	102
Tableau 7.1 : Pourcentage de cas manquant au moins un sondage ou une évaluation lors de la première année de l'étude (quatre périodes d'évaluation)	117
Tableau 7.2 : Enfants évalués en anglais lors des cinq premières évaluations en fonction du groupe	117
Tableau 7.3 : Distribution des valeurs manquantes et des imputations de l'ÉPE-AD selon les groupes expérimentaux (2 ^e Cohorte).....	120
Tableau 7.4 : Comparaison entre le projet Capacité d'apprentissage et l'EVMLO.....	125
Tableau 7.5 : Comparaison entre le projet Capacité d'apprentissage et l'EVMLO —Enfants groupés par langue maternelle.....	126
Tableau 7.6 : Comparaison entre les mères du projet Capacité d'apprentissage et de l'EVMLO — Mères groupées par langue maternelle.....	126
Tableau 7.7 : Comparaison entre les pères du projet Capacité d'apprentissage et de l'EVMLO — Pères groupés par langue maternelle.....	127
Tableau 7.8 : Comparaison entre le projet Capacité d'apprentissage et l'EVMLO — Familles par classe de revenu.....	128
Tableau 7.9 : Comparaison des degrés de scolarité des mères du projet Capacité d'apprentissage et de l'EVMLO.....	129
Tableau 7.10 : Comparaison entre les degrés de scolarité des pères du projet Capacité d'apprentissage et de l'EVMLO	129
Tableau 7.11 : Comparaison de la taille des familles ¹ dans le projet Capacité d'apprentissage et l'EVMLO.....	130

Tableau 7.12 :	Comparaison entre le projet Capacité d'apprentissage et l'EVMLO — Nombre d'enfants par répondant	130
Tableau 7.13 :	Comparaison entre le projet Capacité d'apprentissage et l'EVMLO — Nombre de familles monoparentales et biparentales	131
Tableau 9.1 :	Impact du programme non-ajustés sur la préparation à la scolarisation lors de la première année — Scores standardisés	163
Tableau 9.2 :	Impact du programme (après ajustement pour covariables) sur la préparation à la scolarisation lors de la première année — Scores standardisés	164
Tableau 9.3 :	Différentiation non-ajustée de l'impact des heures en garderie (Dosage) selon le programme reçu (1 ^{ière} année) — Scores standardisés	170
Tableau 9.4 :	Différentiation ajustée de l'impact des heures en garderie (Dosage) selon le programme reçu (1 ^{ière} année) — Scores standardisés	171
Tableau 9.5 :	La différenciation de la qualité du programme testé	177
Tableau 9.6 :	La différenciation de la fidélité du programme testé	177
Tableau 9.7 :	Inter-corrélation des indices de la qualité et de la fidélité de programme en garderie	178
Tableau 9.8 :	L'effet de la qualité du programme en garderie avec ajustement (1 ^{ière} année) — Scores standardisés	181
Tableau 9.9 :	L'effet de la fidélité du programme en garderie avec ajustement (1 ^{ière} année) — Scores standardisés	182
Figure 9.1 :	Diagramme illustrant comment l'effet du programme est départagé en deux composante : la partie attribuable à la qualité et la fidélité du programme en garderie (a, b) et celle qui ne l'est pas (c')	183
Tableau 9.10 :	Test de médiation prenant la qualité comme médiateur de l'effet programme à 12 mois	184
Tableau 9.11 :	Test de médiation prenant la fidélité comme médiateur de l'effet programme à 12 mois	185
Tableau 9.12 :	L'effet modérateur du Type de foyer sur les indicateurs de la préparation à la scolarisation ..	186
Tableau 9.13 :	Impact du programme testé selon le Type de foyer : à forte ou à faible exposition au français — (1 ^{ère} année) — Scores ajustés	187
Figure 9.2 :	Effet du programme (DD) sur la Communication à chaque évaluation	193
Figure 9.3 :	Effet du programme (DD) sur la Conscience de soi à chaque évaluation	194
Figure 9.4 :	Effet du programme (DD) sur la Cognition à chaque évaluation	195
Figure 9.5 :	Effet du programme (DD) sur le Vocabulaire expressif à chaque évaluation	197
Figure 9.6 :	Intervalles de confiance de l'impact du programme standardisé (ANCOVA) sur le vocabulaire français : Échelles normalisées de vocabulaire réceptif (ÉVIP-R) et expressif (EOWPVT)	199
Tableau 9.14 :	Nombre d'heures en garderie lors de la première année (Dosage) et son effet non-ajusté sur le développement de la préparation à la scolarisation lors de la deuxième année selon les deux groupes en garderie — Scores standardisés	201
Tableau 9.15 :	Nombre d'heures en garderie lors de la première année (Dosage) et son effet ajusté sur le développement de la préparation à la scolarisation lors de la deuxième année selon les deux groupes en garderie — Scores standardisés	202
Tableau 9.16 :	Impact du dosage en garderie sur le vocabulaire français : Échelles normalisées de vocabulaire réceptif (ÉVIP-R) et expressif (EOWPVT)	203
Tableau 9.17 :	La qualité du programme en garderie lors de la première année et son effet ajusté sur le développement de la préparation à la scolarisation lors de la deuxième année — Scores standardisés	208
Tableau 9.18 :	La fidélité du programme en garderie lors de la première année et son effet ajusté sur le développement de la préparation à la scolarisation lors de la deuxième année — Scores standardisés	210
Tableau 9.19 :	Impact de la qualité et la fidélité de programme sur le vocabulaire français : Échelles normalisées de vocabulaire réceptif (ÉVIP-R) et expressif (EOWPVT)	211
Tableau 9.20 :	Test de médiation prenant la fidélité comme médiateur de l'effet programme à 24 mois	212
Tableau 9.21 :	Test de médiation prenant la qualité comme médiateur de l'effet du programme en garderie à 24 mois	213
Figure 9.7 :	Intervalles de confiance pour l'effet standardisé (DD) du programme sur la Communication par type de foyer	215

Figure 9.8 :	Intervalles de confiance pour l'effet standardisé (DD) du programme sur la Conscience de soi par type de foyer.....	216
Figure 9.9 :	Intervalles de confiance pour l'effet standardisé (DD) du programme sur la Cognition par type de foyer	217
Figure 9.10 :	Intervalles de confiance pour l'effet standardisé (DD) du programme sur le Vocabulaire expressif par type de foyer.....	218
Figure 9.11 :	Intervalles de confiance pour l'effet standardisé (ANCOVA) du programme sur l'ÉVIP-R et l'EOWPVT par type de foyer	219
Figure 10.1 :	Distribution de la Fréquence des activités de littératie au niveau de base	225
Figure 10.2 :	Distribution de la Fréquence des activités de littératie à 12 mois	225
Figure 10.3 :	Distribution de la Fréquence des activités de littératie à 24 mois	226
Figure 10.4 :	L'effet du programme Ateliers familles sur la Fréquence des activités de littératie.....	227
Tableau 10.1 :	Rapport des cotes des effets DD standardisés du modèle ajusté de la Fréquence des activités de littératie	228
Figure 10.5 :	Distribution de la Langue des activités de littératie au niveau de base	229
Figure 10.6 :	Distribution de la Langue des activités de littératie à 12 mois.....	230
Figure 10.7 :	Distribution de la Langue des activités de littératie à 24 mois.....	230
Figure 10.8 :	L'effet du programme testé sur la Langue des activités de littératie.....	231
Tableau 10.2 :	Rapport des cotes des effets DD standardisés du modèle ajusté de la Langue des activités de littératie	232
Tableau 10.3 :	Sommaire des impacts du programme Ateliers familles sur l'opinion des parents : le parent comme premier éducateur de l'enfant.....	233
Tableau 10.4 :	Sommaire des impacts du programme Ateliers familles sur l'opinion des parents : l'importance de créer un environnement francophone à la maison.....	234
Tableau 10.5 :	Sommaire des impacts du programme Ateliers familles sur l'opinion des parents : le parent et l'éducatrice travaillent ensemble pour préparer l'enfant à l'école	234
Tableau 10.6 :	Sommaire des impacts du programme des Ateliers familles sur la modélisation (deux semaines post programme)	235
Tableau 10.7 :	Sommaire des impacts du programme des Ateliers familles sur la modélisation (huit mois post-programme).....	235
Tableau 10.8 :	Sommaire des impacts du programme Ateliers familles sur l'auto-efficacité des parents.....	236
Tableau 10.9 :	Sommaire des impacts du programme Ateliers familles sur les connaissances des parents	236
Tableau 10.10 :	Sommaire des impacts du programme Ateliers familles sur leur appartenance à la communauté francophone.....	237
Tableau 10.11 :	Corrélations entre le dosage et trois questions ciblant l'opinion des parents.....	238
Tableau 10.12 :	Corrélations entre dosage et Modélisation, Efficacité, Connaissances et Appartenance	238
Tableau 10.13 :	Corrélations entre la qualité de la prestation du programme et trois questions ciblant l'opinion des parents	240
Tableau 10.14 :	Corrélations entre la qualité de la prestation du programme et la Modélisation, l'Efficacité, Connaissances et l'Appartenance.....	240
Tableau 11.1 :	Sommaire des résultats de l'analyse de la mise en œuvre.....	248
Tableau 11.2 :	Sommaire des résultats des analyses d'impact du programme testé.....	250
Tableau A.1 :	Calendrier des évaluations des enfants	271
Tableau B.1 :	Sous-échelle de Vocabulaire expressif	274
Tableau B.2 :	Sous-échelle de Conscience de soi	275
Tableau B.3 :	Sous-échelle de Conscience phonologique	275
Tableau B.4 :	Sous-échelle de Numératie	276
Tableau B.5 :	Mémoire pour les informations personnelles.....	276
Tableau B.6 :	Sous-échelle de connaissances alphabétiques	277
Tableau B.7 :	Lecture orale	277
Tableau D.1 :	Arbre décisionnel pour déterminer la langue d'administration du test	282
Tableau E.1 :	Moyennes et écarts types des scores originaux et imputés ainsi que leur inter-corrélation	284
Tableau F.1 :	Pertinence des variables de démographie et de composition familiale	286
Tableau F.2 :	Pertinence des variables socio-économiques	287
Tableau F.3 :	Pertinence des variables linguistiques	288
Tableau F.4 :	Pertinence des variables sociolinguistiques	289

Tableau F.5 :	Pertinence de variables de style parental	290
Tableau F.6 :	Pertinence des caractéristiques de l'environnement en garderie.....	291
Tableau F.7 :	Pertinence des facteurs méthodologiques	292
Tableau F.8 :	Tests robustes de l'instabilité des effets de variables confusionnelles dans le temps	293
Tableau G.1 :	Impact du programme sur le vocabulaire français : Échelles normalisées de Vocabulaire réceptif (ÉVIP-R) et Vocabulaire expressif (EOWPVT).....	295
Tableau H.1 :	Impact du programme non-ajustés sur la préparation à la scolarisation lors de la deuxième année – Scores standardisés.....	297
Tableau H.2 :	Impact du programme ajustés sur la préparation à la scolarisation lors de la deuxième année – Scores standardisés	298
Tableau I.1 :	Tests de l'hypothèse que le profil linguistique du foyer agit comme modérateur de l'effet programme lors de la deuxième année.....	299
Tableau I.2 :	Impact du programme testé selon le profil linguistique du foyer – (2e année) – Scores ajustés	300
Tableau I.3 :	Impact du programme testé sur le vocabulaire selon le profil linguistique : Échelles normalisées de vocabulaire réceptif (ÉVIP-R) et expressif (EOWPVT)	301
Tableau J.1:	Impact du programme testé sur la fréquence des activités de littératie ainsi que la langue utilisée lors de ces activités (non-ajusté).....	303
Tableau J.2:	Impact ajusté du programme testé sur la Fréquence des activités de littératie ainsi que la Langue des activités de littératie.....	304

Sommaire

Ce rapport vise à documenter la mise en œuvre et l'impact du projet Capacité d'apprentissage dans les communautés francophones en situation minoritaire (titre abrégé : projet Capacité d'apprentissage; auparavant le Projet pilote de garde d'enfants, PPGE), un projet de démonstration financé par Ressources humaines et développement des compétences Canada (RHDC). Les services de la Société de recherche sociale appliquée (SRSA) ont été retenus pour la mise en œuvre, la gestion, la collecte et les analyses des données du projet. Le projet met à l'essai un programme préscolaire¹ qui jumelle un volet de service de garde spécifiquement développé pour satisfaire aux besoins des enfants francophones en milieu minoritaire à un volet d'alphabétisation familiale ciblant les parents de ces enfants. Ce programme a pour objectif de développer les compétences langagières de l'enfant, sa connaissance et son utilisation du français, sa connaissance de la culture francophone et son appartenance à celle-ci, ainsi que de favoriser sa préparation à l'école et son développement en général.

L'évaluation de ce programme se fait au moyen d'une méthodologie quasi expérimentale avec groupes témoins non équivalents. La méthodologie prévoit trois groupes expérimentaux : le groupe programme composé d'enfants inscrits dans la garderie francophone offrant le nouveau programme préscolaire; le groupe témoin en garderie comprenant les enfants inscrits dans une garderie francophone n'offrant pas le nouveau programme; et le groupe témoin hors garderie regroupant les enfants dont la garde de jour a lieu à la maison ou en garderie familiale non réglementée. Le groupe témoin en garderie vise à tenir compte de l'influence d'une garderie en installation sur le développement des enfants, ce qui constitue en soi un traitement. Le groupe témoin hors garderie vise à tenir compte de l'influence d'un milieu de garde informel sur le développement de l'enfant. Le projet compte deux cohortes de participants, la première recrutée en 2007 et la deuxième recrutée en 2008.

Le présent rapport porte sur les données collectées avant l'intervention et durant les deux années de la prestation du programme, soit de mai 2007 à octobre 2009 auprès de la première cohorte et d'octobre 2008 à octobre 2010 auprès de la deuxième cohorte. Cette période correspond au moment où les enfants sont âgés, en moyenne, de trois à cinq ans. Une méthodologie mixte de recherche a été utilisée dans le cadre du projet Capacité d'apprentissage. Cette approche préconise l'utilisation d'une diversité d'outils, de nature autant quantitative que qualitative, provenant de plusieurs sources toutes choisies en fonction des objectifs de la recherche, c'est-à-dire de vérifier si le programme a l'effet voulu et de comprendre comment il exerce son effet. Ainsi, les analyses ont été effectuées à partir des données provenant, entre autres, des évaluations des enfants, des sondages auprès des parents, des observations en garderie et des Ateliers familles ainsi que des données administratives (par ex. le registre des présences et absences des enfants à la garderie).

Dans un premier temps, nous avons étudié la mise en œuvre du programme par le biais d'observations directes des sites où le projet a eu lieu. L'analyse de ces données montre que de façon générale le programme reçu par les enfants et les parents ressemblait à ce qui était établi

¹ Officiellement connu sous le nom de services de garde enrichis dans les documents de RHDC, la Société de recherche sociale appliquée, ou SRSA, en accord avec RHDC, nomme dorénavant le programme « programme préscolaire de service de garde ».

par les concepteurs du programme. La fidélité du programme en garderie et du programme des Ateliers familles était très bonne. Le programme en garderie se distinguait même favorablement sur le plan de la qualité par rapport aux programmes typiquement offerts en garderie au sein de milieu minoritaire linguistique. Que ce soit du point de vue des éducatrices, des formatrices ou des parents, le programme testé était perçu par les participants comme étant favorable à la préparation à la scolarisation des enfants.

Les résultats d'analyses quantitatives permettent de constater que : (a) le programme testé a un effet positif sur l'habilité de communiquer en français ainsi que sur le développement du vocabulaire des enfants, (b) la fidélité/qualité supérieur du programme en garderie testé explique très bien cet effet (résultats provenant des analyses médiatrices), (c) le changement des attitudes et des comportements parentaux ne semble pas être associé aux gains observés auprès des enfants, (d) plus un enfant était exposé au programme lors de sa prestation, plus l'effet du programme sur son développement est élevé et (e) l'exposition à la langue française détermine en partie, tel que prévu, la nature de l'effet du programme : les enfants à forte exposition au commencement de l'étude ont bénéficié davantage sur le plan du développement cognitif tandis que les enfants à faible exposition semblent bénéficier davantage par rapport au développement de leurs compétences langagières. Deux conclusions principales découlent de ces résultats. Dans un premier temps, notons que le moteur principal de l'effet du programme semble être le programme en garderie. D'ailleurs, tous les enfants semblent tirer un bénéfice du programme testé, bien que la nature de l'effet observé semble dépendre dans une certaine mesure du profil linguistique de l'enfant et de sa famille. L'image dégagée de l'ensemble des analyses permet de conclure que le programme testé a un impact modeste sur le niveau de préparation à la scolarisation ainsi que sur le développement langagier des enfants du groupe programme. Au niveau des politiques sociales, les résultats de cette recherche s'adressent aux priorités identifiées par le gouvernement fédéral dans le cadre de la *Feuille de route pour la dualité linguistique canadienne de 2008–2013*.

Le lecteur est prié de noter que ce rapport s'insère dans une série de rapports préparés par la SRSA. Il fait suite au rapport *Capacité d'apprentissage dans les communautés francophones en situation minoritaire : Rapport de référence* (Legault, Mák, Verstraete et Bérubé, 2014) dont la version finale a été remise à RHDCC le 13 octobre 2009. Ce premier rapport dressait le profil des enfants, de leur famille et des communautés participant au projet Capacité d'apprentissage. De plus, le présent rapport se veut complémentaire au rapport *Capacité d'apprentissage dans les communautés francophones en situation minoritaire : Rapport de mise en œuvre du projet* (Bérubé, Legault, Janisse, Carson, Saucier et Lefebvre, 2014), soumis à RHDCC le 31 mai 2010 et au rapport *Capacité d'apprentissage dans les communautés francophones en situation minoritaire : Rapport des résultats de la première cohorte* (Thompson, Legault, Lalonde et Bérubé, 2014) soumis à RHDCC le 31 juillet 2010. Deux autres rapports sont prévus, chacun portant respectivement sur les impacts à un an suivant la fin de l'intervention et les impacts deux ans après la fin de l'intervention.

1. Présentation du projet Capacité d'apprentissage dans les communautés francophones en situation minoritaire

1.1. GENÈSE DU PROJET

Le projet Capacité d'apprentissage dans les communautés francophones en situation minoritaire (projet Capacité d'apprentissage) fait partie du *Plan d'action pour les langues officielles* 2003–2008 du gouvernement du Canada et se poursuit sous la bannière de la *Feuille de route pour la dualité linguistique canadienne* de 2008–2013. Les principes directeurs qui orientent le projet comprennent la volonté de fournir un appui aux communautés francophones en situation minoritaire pour assurer aux enfants un bon départ dans la vie et d'encourager les parents à participer activement aux apprentissages de leur enfant (Ressources humaines et Développement des compétences Canada, 2006). Au niveau communautaire, le projet Capacité d'apprentissage se veut une évaluation rigoureuse d'une intervention prometteuse ayant comme objectif le maintien, voire le renouvellement, de la vitalité ethno-linguistique de la communauté francophone minoritaire.²

Ce projet reconnaît que le développement des dimensions linguistique et identitaire prend une place plus importante en contexte linguistique minoritaire que dans un contexte majoritaire. Alors que les membres de la majorité linguistique peuvent prendre ce processus développemental pour acquis, toute la complexité de ce processus ressort en situation linguistique minoritaire. Le développement des dimensions linguistique et identitaire est l'aboutissement du processus de socialisation vécu dans le milieu familial, le milieu scolaire ou préscolaire et le milieu socio-institutionnel (Landry et Allard, 1997). Les études pionnières quant à l'importance de la culture sur le développement de l'enfant nous viennent de Vygotsky (1978). La culture dans laquelle un enfant grandit a une influence sur le développement de ses habiletés langagières et sur ses apprentissages en général via l'intégration des symboles sociaux auxquels il est exposé. Ainsi, l'environnement social dans lequel baigne l'enfant est indissociable de la construction de son identité culturelle et linguistique, de même que de son développement global.

Dans la sphère publique, la réalité d'un contexte minoritaire fait en sorte que les petits francophones sont exposés à deux cultures différentes à un moment où leur identité se définit. Selon Gilbert (2003), l'exposition au français dans l'ensemble des contextes sociaux est d'autant plus importante pour l'enfant grandissant dans un milieu francophone très minoritaire où, par le simple poids démographique, l'anglais prédomine dans tous les aspects de la vie quotidienne. Au niveau de l'enfance, plusieurs auteurs préconisent la disponibilité de services de garde et de scolarisation en français comme principaux vecteurs de la vitalité communautaire (Commission nationale des parents francophones, 2005; Landry et Allard, 1997; Gilbert, 2003). L'idéal, selon les défenseurs de la vitalité de la communauté francophone, serait donc que les parents ayants droit inscrivent leurs enfants dans des services de garde de qualité et dans des écoles de langue française. La réalité est toutefois toute autre.

Les enfants de parents ayants droit sont souvent inscrits dans des écoles anglophones. Selon les résultats de l'Enquête sur la vitalité des minorités de langue officielle de 2006 (EVMLO; Corbeil, Grenier, et Lafrenière, 2007), seulement 56 % des enfants de parents ayants droit

² Voir Guimond (2003) pour un survol des études portant sur la vitalité ethno-linguistique en milieu minoritaire.

fréquentent des écoles primaires de langue française. Cette proportion diminue à 44 % à l'adolescence. Ainsi, une proportion non négligeable de parents ayants droit choisissent d'inscrire leurs enfants dans des écoles d'immersion ou dans des écoles de langue anglaise pour plusieurs raisons, dont: l'anglais est la langue maternelle ou la langue la mieux connue de l'enfant, la proximité de l'école, la non-disponibilité d'une école de langue française et la qualité du programme ou de l'école.

Quant aux jeunes enfants francophones inscrits dans les écoles de langue française, plusieurs études démontrent que ces derniers obtiennent des résultats en littératie et en numératie inférieurs aux enfants des groupes linguistiques majoritaires. Cette disparité entre les deux groupes est observée, entre autres, dans les résultats des tests internationaux tels que le Programme international pour le suivi des acquis des élèves (PISA) où les enfants francophones inscrits à l'école de langue française en milieu linguistique minoritaire obtiennent des résultats inférieurs en lecture par rapport à leurs pairs anglophones canadiens (Bussière, et al., 2001; Conseil canadien sur l'apprentissage, 2008). Les quelques études portant sur les petits francophones minoritaires font ressortir que cet écart apparaît dès le jeune âge. Une étude récente auprès d'enfants franco-manitobains âgés de quatre à six ans conclut que ces derniers obtiennent des scores plus faibles aux tests de vocabulaire (Échelle de vocabulaire en images de Peabody – Révisé, ou ÉVIP-R, et échelles de communication et de connaissances générales de l'Instrument de mesure du développement de la petite enfance, ou IMDPE) et ce d'autant plus pour les enfants grandissant dans un environnement quotidien majoritairement anglophone. Cette tendance se poursuit lorsque ces enfants atteignent la troisième année du primaire. Les enfants ayant grandi dans un environnement familial-pré-scolaire francophone obtiennent de meilleurs résultats en lecture en troisième année que les enfants francophones ayant vécu dans un environnement linguistique majoritairement anglophone (Chartier, Dumaine, et Sabourin, 2011).

L'exposition restreinte de ces enfants à la langue française à la maison, limitant ainsi le développement de leurs habiletés langagières en français, serait vraisemblablement à l'origine de ces difficultés. De fait, près des deux tiers des petits francophones minoritaires sont issus de foyers exogames (67 %) et la plupart adoptent l'anglais comme langue du foyer (Landry, 2010).³ Seulement 20 % des couples exogames choisissent le français pour élever leurs enfants de 0 à 4 ans (Martel, 2001). En outre, les dernières données tirées du Recensement de 2006 indiquent que près de 39 % des francophones vivant hors Québec parlent plutôt l'anglais à la maison, bien que le français demeure une langue utilisée (Corbeil et Blaser, 2007).

Cette plus grande utilisation de la langue anglaise au quotidien explique en partie pourquoi 62 % des adultes francophones hors Québec passant un test de littératie en français (plutôt qu'anglais) n'obtiennent pas le niveau de maîtrise de la littératie jugé nécessaire pour bien fonctionner en société (c.-à-d. un niveau de littératie de plus que 3 sur une échelle de 5; Statistique Canada et RHDSC, 2005, Tableau 3.24). Cette proportion serait sans aucun doute plus élevée si tous les adultes francophones hors Québec avaient fait le test en français (65 % de ces derniers choisissent de remplir le test en anglais bien qu'ils identifient le français comme langue maternelle; Statistique Canada et RHDSC, 2005, p. 54). Selon ces résultats, il y a tout lieu

³ Landry, R. (2010). *Petite enfance et autonomie culturelle : Là où le nombre le justifie... V*. Rapport de recherche réalisé pour la Commission nationale des parents francophones. Institut canadien de recherche sur les minorités linguistiques. Moncton, Nouveau-Brunswick.

de consolider l'acquisition de la langue auprès des populations en contexte minoritaire et d'éveiller l'engagement des parents envers la préservation de la communauté francophone.

Pour les petits francophones minoritaires, l'ensemble de ces constatations sont inquiétants lorsque l'on considère que les compétences langagières demeurent le *pivot central* de la réussite scolaire, qui elle influence la réussite professionnelle et sociale. Une intervention précoce auprès des petits francophones minoritaires ciblant autant le milieu familial que préscolaire pourrait affecter positivement leur trajectoire linguistique favorisant ainsi leur succès scolaire et leur intégration au sein de la communauté francophone. Cette intervention devrait, entre autres, comprendre une composante ciblant les parents afin de les conscientiser aux défis de vivre en milieu minoritaire et aux actions qu'ils peuvent prendre pour transmettre ce riche héritage culturel à leurs enfants. Le projet Capacité d'apprentissage vise à évaluer l'impact d'une telle intervention.

1.2. LE PROJET CAPACITÉ D'APPRENTISSAGE

Le projet Capacité d'apprentissage met à l'essai un programme préscolaire de service de garde et vise à mettre en évidence les avantages qu'apporte ce programme préscolaire aux enfants vivant en milieu minoritaire francophone. Le programme préscolaire testé jumelle un volet de service de garde spécifiquement développé pour satisfaire aux besoins des enfants francophones en milieu minoritaire à un volet d'alphabétisation familiale ciblant les parents de ces enfants. Ce dernier volet cherche à stimuler la participation active des parents dans le développement de leur enfant et dans sa préparation à l'école, ainsi que la transmission de la langue et de la culture françaises. Le projet comme tel s'inscrit dans la foulée des nombreuses études portant sur le développement des enfants d'âge préscolaire et sur la vitalité de la langue française dans des milieux minoritaires.

L'effet du programme sur le développement des enfants est examiné en comparant un groupe de participants ayant été exposé au nouveau programme (connu sous le nom de groupe programme) à des groupes témoins comprenant des participants n'ayant *pas été* exposé au nouveau programme (nommé groupe témoin en garderie et groupe témoin hors garderie). Un comité consultatif de recherche, composé d'universitaires spécialisés en petite enfance francophone et de représentants des communautés francophones, a été créé afin d'aider à concevoir, mettre en œuvre, faire le suivi et évaluer le projet pilote. La Société de recherche sociale appliquée gère, met en œuvre et évalue le programme, en plus de participer activement au développement du projet.

La prestation du programme s'est faite pour deux cohortes de participants. La première cohorte a débuté le programme à l'automne 2007 dans six communautés francophones en milieu minoritaire (Saint-Jean et Edmundston au Nouveau-Brunswick; Orléans, Cornwall et Durham en Ontario; et Edmonton en Alberta). La prestation du programme à la deuxième cohorte a commencé à l'automne 2008 dans deux communautés (Orléans et Cornwall en Ontario). Les enfants sont suivis dès l'âge de trois ans jusqu'à sept ans, alors qu'ils seront au début de la deuxième année du primaire. La durée de l'étude nous permet de suivre le développement des petits francophones minoritaires du préscolaire au début de la scolarisation. Les dernières collectes de données se feront à l'automne 2011 pour les familles de la première cohorte et à l'automne 2012 pour les familles de la deuxième cohorte.

Dans une première phase, le projet Capacité d'apprentissage vise d'abord à répondre à la question de recherche: *Le nouveau programme préscolaire, comprenant un volet en garderie et un volet d'ateliers parents/enfants, a-t-il un impact significatif sur les habiletés langagières, l'identité culturelle francophone et la préparation à la scolarisation des enfants au-delà du développement qui prendrait place en l'absence de ce programme, et indépendamment d'autres facteurs externes qui pourraient entrer en jeu?* Des questions connexes sont également investiguées, dont : *Pour qui ce programme est-il le plus bénéfique? Ce programme est-il rentable? Le nouveau programme peut-il être répété?* Dans une deuxième phase, le projet Capacité d'apprentissage aborde une nouvelle question de recherche: *Le nouveau programme préscolaire permet-il aux enfants francophones grandissant en milieu minoritaire d'être mieux outillés pour réussir dans les tâches essentielles à la réussite scolaire que sont la lecture et les mathématiques?* Alors que la première question de recherche se concentre sur la période préscolaire, la deuxième question de recherche se penche sur la période de scolarisation englobant la première et deuxième année scolaire, alors que les enfants sont âgés de 6 et 7 ans. Le présent rapport porte sur la première phase du projet.

1.3. L'APPORT DU PROJET

L'apport de ce projet est qu'il tient compte de l'environnement particulier des communautés minoritaires francophones en ayant une composante de francisation, une composante d'identité culturelle et une composante de sensibilisation et d'implication des parents. Le projet s'articule autour des jeunes enfants francophones et s'inspire du modèle écologique de Bronfenbrenner (1979) dont l'une des prémisses est que les enfants sont influencés par l'ensemble des environnements dans lesquels ils baignent. Le programme testé vise à renforcer les liens entre les principaux milieux dans lesquels l'enfant grandit — les actions du milieu en garderie appuyant ceux du milieu familial et vice versa — de façon à optimiser les apprentissages de l'enfant tant au niveau de son développement général et de sa préparation à la scolarisation qu'au niveau de son développement langagier et culturel francophones.

Les multiples bienfaits des programmes qui modifient à la fois l'environnement de l'enfant à la garderie et à la maison ont été établis dans le cadre de plusieurs études portant sur d'autres populations dites vulnérables (voir la revue de littérature de Reese, Sparks, et Leyva, 2010; Engle, et al., 2007). La maximisation des effets de ces programmes, croit-on, passe par l'adoption des mêmes approches par le parent et l'éducateur avec l'enfant. C'est ce que démontre une étude menée par Pelletier et Corter (2005). Les résultats ont démontré qu'un programme à deux volets (parent et éducatrice⁴), axé sur les activités de littératie précoce, favorisait davantage les acquisitions des enfants dans ce domaine qu'un programme semblable comprenant seulement un des deux volets. Les enfants dont l'environnement à la garderie et celui à la maison avaient changés, étaient plus avancés dans le développement du vocabulaire, de la pré-lecture et de la pré-numératie.

En somme, l'ensemble du programme préscolaire s'appuie sur une perspective intégrative et tente par différentes approches d'influer sur les principaux vecteurs qui modèlent le développement de l'enfant. Les sections qui suivent présentent d'abord les différentes sources d'influence du développement des tout-petits en tenant compte des particularités de grandir en

⁴ Le féminin est privilégié en reconnaissance de la très forte proportion de femmes dans cette profession.

milieu minoritaire francophone pour ensuite se concentrer sur une définition de la préparation à la scolarisation. Le chapitre se termine avec une section détaillant les objectifs et le contenu du présent rapport.

1.4. LE MODELAGE DU DÉVELOPPEMENT DES ENFANTS EN MILIEU MINORITAIRE

Le projet Capacité d'apprentissage est basé sur une vision écologique selon laquelle des facteurs à divers niveaux influencent le développement de l'enfant. Cette vision est clairement représentée dans le modèle écologique de Bronfenbrenner (1979). Ce chercheur fut le premier à mettre en mots et en images tout le système d'influences qui agit sur le développement de l'enfant. Son modèle s'articule autour de trois prémisses:

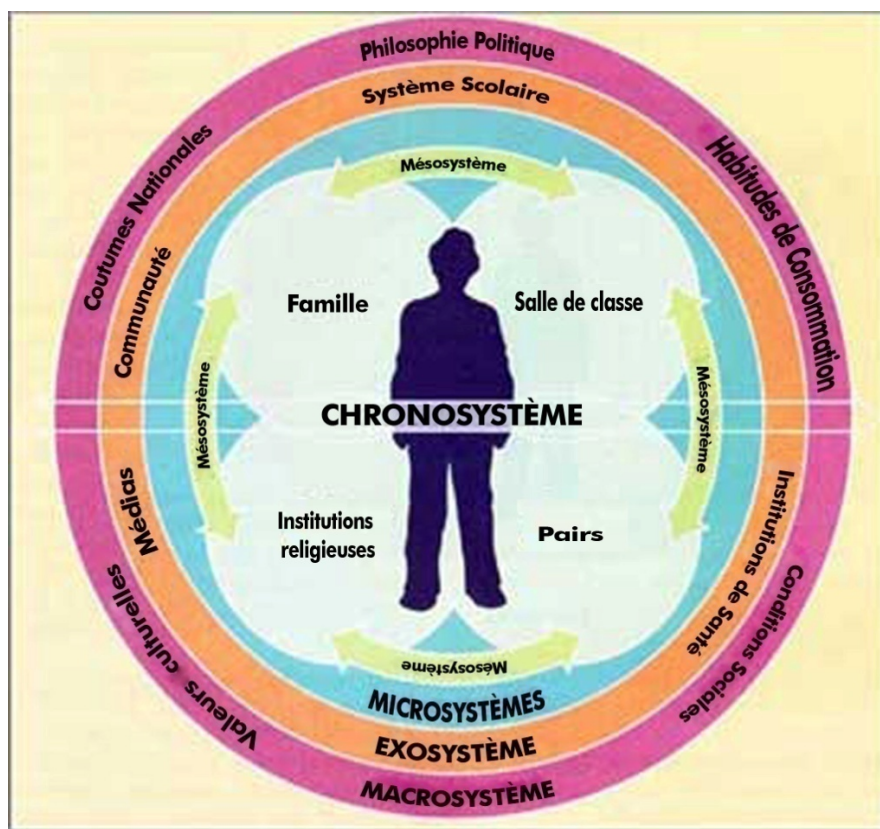
- a) l'enfant est au centre du modèle;
- b) la place centrale qu'occupent les expériences vécues par l'enfant (ces expériences sont considérées comme des « moteurs » du développement); et
- c) la nature des relations entre les différents milieux de l'enfant.

Le modèle de Bronfenbrenner comprend cinq systèmes (figure 1.1) :

- **Microsystème** : Environnement immédiat (famille, école, mode de garde, pairs, voisinage).
- **Mésosystème** : Interactions entre les environnements immédiats (p. ex., entre la maison et l'école).
- **Exosystème** : Environnement externe qui affecte indirectement l'enfant (tel que le travail des parents).
- **Macrosystème** : Contexte culturel plus large (culture occidentale par rapport à culture orientale, économie nationale, culture politique, sous-culture).
- **Chronosystème** : Structure des événements affectant l'environnement et les transitions au cours de l'existence.

Dans le contexte du projet Capacité d'apprentissage où le centre d'intérêt est les jeunes enfants en milieu linguistique minoritaire, trois systèmes du modèle de Bronfenbrenner prennent une place particulièrement importante. Tout d'abord, le microsystème, via les caractéristiques familiales et les caractéristiques du milieu de garde de l'enfant, influe sur le développement des enfants. Au niveau des caractéristiques familiales, nous distinguons les variables contextuelles des processus familiaux (cette division se base sur l'Enquête longitudinale nationale sur les enfants et les jeunes, ou ELNEJ; Statistique Canada et RHDSC, 2006). Les variables contextuelles font référence à des données « factuelles » reconnues comme étant importantes pour le développement des enfants (p. ex. la composition de la famille). Ensuite le mésosystème, tel que les liens qu'entretiennent la famille et le milieu de garde, joue aussi un rôle dans le développement des tout-petits. Enfin, le macrosystème, soit la communauté dans laquelle l'enfant grandit, et tout particulièrement ses caractéristiques linguistiques, fait aussi partie des influences d'intérêt pour le projet.

Figure 1.1 : Modèle écologique complet de Bronfenbrenner (1979)



Source : Dans *UW-Extension ABC Project*, Annexe B (novembre 2004), traduction libre.

1.4.1. Le microsysteme : environnement familial (variables contextuelles)

Le développement des enfants subit l'influence de plusieurs facteurs intrinsèques à l'enfant comme son historique fœtal, son poids à la naissance ou la durée de sa gestation. Outre les caractéristiques inhérentes à l'enfant, les recherches ont identifié au cours des années une série de facteurs environnementaux pouvant affecter son développement. Les variables contextuelles de l'environnement familial font partie des facteurs contribuant le plus au développement de l'enfant (Sanders et Morawska, 2006). Ces dernières comprennent la composition de la famille, le revenu, le niveau de scolarité des parents, ainsi que les langues utilisées à la maison.

Composition de la famille

Dans le contexte du projet Capacité d'apprentissage, nous nous intéressons aux variables familiales influençant le développement de l'enfant, plus particulièrement à celles qui influencent son parcours langagier. Ainsi, le rang de naissance s'avère être une variable d'importance. En effet, les études démontrent que les aînés des familles ont un vocabulaire plus étendu en moyenne que leur fratrie (Tamis-LeMonda et Rodriguez, 2008).

Le type de famille (biparentale intacte, monoparentale, recomposée, etc.) est également une variable à considérer. Une étude effectuée avec les données longitudinales de l'ELNEJ a montré que le type de famille (biparentale ou monoparentale) avait une influence concrète sur plusieurs aspects du développement de l'enfant. En effet, pour les enfants de six à onze ans, les auteurs

concluent que « les enfants des familles biparentales étaient moins hyperactifs, avaient de meilleures aptitudes scolaires, étaient moins anxieux ou déprimés et ont reçu de leurs enseignants une bonne évaluation quant à leur niveau de connaissances » (Adams et Ryan, 2000, p. iii). Cependant, ce n'est pas tant la composition de la famille que les facteurs associés à la monoparentalité, comme le stress maternel ou la diminution du revenu familial, qui affectent l'enfant *a priori*, d'où l'importance d'interpréter les résultats avec précaution.

Revenu familial brut

Des centaines d'études ont déjà démontré une association entre la pauvreté familiale et la santé de l'enfant, sa préparation à l'école, sa réussite scolaire et son comportement. Toutefois, peu de ces études se sont penchées sur l'effet du moment, de la durée et de l'intensité de la pauvreté. Un revenu familial insuffisant affecte le développement de l'enfant via de nombreux canaux : nutrition inadéquate, diminution des situations d'apprentissage, instabilité du lieu de résidence, inscription à des écoles ayant moins de ressources, violence familiale, etc. (voir Duncan et Brooks-Gunn, 1997). Certaines études, comme celle de Berger et de ses collègues (2005), ont confirmé l'hypothèse qu'un faible revenu influence les résultantes du développement de l'enfant par le biais de ses effets multiples sur l'environnement à la maison. Cependant, bien que la relation causale entre un faible revenu et les résultantes chez les enfants (résultantes intellectuelles et comportementales) soit bien claire, l'interprétation de ces associations reste discutable, de même que les implications en termes de politiques.⁵ Les études récentes sur le sujet tendent à démontrer deux principales « routes » par lesquelles un faible revenu affecte potentiellement l'enfant : l'environnement physique et la qualité des pratiques parentales (Berger, Paxson, et Waldfogel, 2005).⁶ Autrement dit, la pauvreté affectera l'achat de ressources matérielles destinées à la famille, en plus d'affecter le niveau de stress familial, lesquels à leur tour influencent le développement de l'enfant. Plusieurs études basées sur des données nationales ont démontré que le niveau de stimulation cognitive amené par l'environnement à la maison (mesuré par le matériel d'apprentissage et les pratiques parentales relatives à l'apprentissage) explique de 33 à 50 % l'association entre le revenu et différentes résultantes sur le développement cognitif et du langage de l'enfant (Dearing, Berry, et Zaslow, 2006).

Les études tendent aussi à démontrer que les effets néfastes de la pauvreté sur le développement cognitif et la réussite scolaire de l'enfant sont plus grands lors de la période préscolaire qu'à toute autre période (Dearing, et al, 2006). La prise en compte de cette donnée devient donc incontournable pour évaluer l'effet du programme dans le cadre du projet Capacité d'apprentissage.

Niveau de scolarité des parents

Tout comme le revenu familial, le niveau de scolarité des parents est un facteur important de la réussite d'un enfant (voir, entre autres, Haveman et Wolfe, 1995). Toutefois, le mécanisme par lequel l'éducation des parents influence le développement de l'enfant a été moins bien étudié. Klebanov et ses collègues (1994) ont démontré que l'éducation de la mère et le revenu familial sont d'importants facteurs reliés à l'aménagement d'un environnement physique propice à

⁵ Par exemple, les transferts monétaires directs vers les familles seraient adéquats si une relation causale claire existait entre le revenu, la qualité de l'environnement et les résultantes des enfants (Berger, et al., 2005).

⁶ La première théorie a été développée initialement dans la littérature économique (voir Becker, 1993) et la seconde dans la littérature en psychologie du développement (p.ex. Dearing, et al., 2006).

l'apprentissage, mais que seule l'éducation s'avère un facteur d'importance dans les pratiques parentales « chaleureuses ». Un ensemble d'études de Davis-Kean (2005) conclut que l'éducation des parents influence le développement de l'enfant non seulement via la réussite sociale des parents, mais aussi via leurs croyances et leurs comportements envers leur enfant. On peut donc voir que le niveau d'éducation est important pour le développement des enfants, au-delà du contexte socio-économique auquel il est souvent associé.

Âge de la mère à la naissance du premier enfant

L'âge de la mère à la naissance du premier enfant est un autre facteur qui a de multiples effets sur le développement des enfants. Les études montrent que l'âge de la mère à la naissance de son premier enfant a un impact sur le développement de l'enfant et des enfants subséquents, même une fois les pratiques parentales et le fonctionnement de la famille pris en compte (Tremblay, et al., 2004). Dans le cas qui nous préoccupe, retenons que les très jeunes mères offrent une stimulation langagière moins complexe à leur enfant. Elles utilisent un éventail de vocabulaire moins riche et elles stimulent moins verbalement leurs enfants (Tamis-LeMonda et Rodriguez, 2008). Ce comportement se répercute chez l'enfant qui est alors à risque de présenter un développement langagier insuffisant pour lui assurer une entrée à l'école réussie.

Langue(s) parlée(s) à la maison

En contexte minoritaire, les langues les plus souvent parlées à la maison par les parents de l'enfant cible sont reliées à la transmission de la langue et à la vitalité du français à la maison (Forgues et Landry, 2006). Ainsi, l'environnement linguistique familial est associé au niveau de préparation scolaire et à la réussite scolaire de l'enfant (Chartier, et al., 2011). Afin de trouver cette information, il faut savoir quelles sont les langues habituellement parlées dans l'environnement de l'enfant, en particulier par ses parents, ses frères et sœurs aînés à la maison et à l'extérieur de la maison. Il est aussi intéressant de connaître les langues habituellement parlées par l'enfant. Ce comportement linguistique est une indication concrète des langues avec lesquelles l'enfant se sent plus à l'aise de s'exprimer. L'ensemble de ces informations permettra d'établir un portrait global des influences linguistiques auxquelles l'enfant est exposé.

1.4.2. Le microsystème : environnement familial (processus familiaux)

Outre les caractéristiques descriptives des familles, les processus familiaux ont une influence certaine sur le développement des enfants. Parmi les principaux éléments affectant le développement de l'enfant, notons le fonctionnement familial, les pratiques parentales et les activités d'alphabetisation de la famille.

Fonctionnement familial

Au-delà des difficultés psychosociales de la mère, l'environnement global dans le ménage est aussi un élément que le chercheur voudra mettre en relation avec le développement de l'enfant. Le fonctionnement de la famille se définit comme la qualité des liens dans la famille, tant au niveau de la qualité de la communication, de l'entente entre les membres que du soutien disponible au sein de la famille. D'un point de vue social, un fonctionnement familial déficient est largement associé à des trajectoires de développement agressives chez les enfants (Tremblay, et al., 2004). Du côté du développement langagier, l'acquisition du vocabulaire chez l'enfant est positivement liée au niveau de fonctionnement de sa famille, soit le niveau d'entente et de

communication entre les membres (Desrosiers et Ducharme, 2006). Le fonctionnement familial a donc des répercussions importantes et multiples sur le parcours de vie des enfants.

Pratiques parentales

En tant que premiers éducateurs de l'enfant, les parents jouent un rôle clé dans le développement et le fonctionnement de leur enfant. Le style parental aura une influence sur le développement social, intellectuel, moral et affectif de l'enfant (Bornstein et Bornstein, 2007). Les styles parentaux sont composés de deux dimensions : la *sensibilité*, qui mesure à quel point le parent est à l'écoute de l'enfant et est capable de répondre à ses besoins et intérêts; et le *contrôle* (ou l'*exigence*) qui réfère au degré de supervision et de discipline, ainsi qu'à quel point le parent exige de son enfant qu'il fasse preuve d'obéissance et de maîtrise de soi (Conseil canadien sur l'apprentissage, 2007). Les études sur le sujet indiquent que les enfants démontrent de meilleures habiletés langagières et ont des scores plus élevés aux tests d'intelligence lorsque leurs parents sont plus encourageants et moins contrôlants (Sanders et Morawska, 2006). Il est donc important de mesurer ces deux dimensions lorsque vient le temps de suivre le développement langagier et cognitif de l'enfant.

Activités d'alphabétisation

Les expériences vécues durant la petite enfance se répercutent sur les habiletés langagières de l'enfant dès l'entrée à l'école (Doherty, 1997). Selon Desrosiers et Ducharme (2006), les enfants à qui les parents ont fait la lecture régulièrement sont moins susceptibles de présenter un retard sur le plan du vocabulaire. La lecture à la maison en bas âge (avant trois ans et demi) serait même associée à une amélioration des capacités verbales chez les enfants qui montraient des difficultés langagières. De la même façon, la participation des parents à des activités d'apprentissage à la maison permet de prédire l'adaptation sociale et scolaire des enfants à long terme (Izzo, Weissberg, Kaspro, et Fendrich, 1999). D'intérêt particulier pour le projet Capacité d'apprentissage, l'exposition au français oral et écrit durant le préscolaire est cruciale pour que l'enfant développe une forte compétence linguistique. Les activités telles que les heures du conte et l'emprunt de livres à la bibliothèque permettent aux éléments uniques de la culture francophone d'être transmis aux enfants (Salerno dans Lafrance, 1993).

1.4.3. Le microsystème : milieu de garde

Les caractéristiques du milieu de garde de l'enfant influent sur de nombreuses facettes du développement de l'enfant. Depuis plusieurs années, psychologues et éducateurs ont conclu, sur la base d'observations et d'expériences, que les soins non parentaux affectent aussi bien le développement cognitif que langagier de l'enfant (National Institute of Child Health and Human Development Early Child Care Research Network et Duncan, 2003). De façon plus précise, la qualité d'un milieu de garde se répercute sur le développement cognitif et langagier des enfants, sur leur préparation à l'école, ainsi que sur leurs comportements (Cleveland, et al., 2006). Par ailleurs, cet impact varie grandement selon certains facteurs familiaux. Par exemple, les effets d'un milieu de garde de qualité sont plus importants auprès des enfants grandissant dans un contexte socio-économique défavorisé (Burchinal, et al., 2000), pour lequel le statut de minorité linguistique est parfois considéré un exemple (Maltais, 2007).

La garderie peut être un lieu d'apprentissage menant à de meilleures habiletés scolaires. Par exemple, dans des études réalisées aux États-Unis, le fait d'avoir fréquenté un service de garde a

été associé à des scores plus élevés en lecture et en mathématiques à l'entrée à la maternelle à cinq ans (voir, entre autres, Howes, et al., 2008). Un matériel adéquat comprenant des jeux et des livres de qualité, un milieu physique approprié et des éducatrices affectueuses soutenant le développement des enfants peuvent s'avérer très bénéfiques pour tous les enfants, et ce, particulièrement pour des enfants vivant en situation familiale difficile. La qualité dans les services de garde peut prendre deux formes : la qualité structurelle, qui prend en compte les facteurs pouvant être modifiés par le biais de la législation (diplôme de l'éducatrice, heures d'ouverture, taille du groupe, etc.), et la qualité des processus qui réfère aux expériences de l'enfant par rapport à son environnement en garderie (qualité des activités offertes et du lien avec l'éducatrice) (Burchinal, et al., 2000). Dans le cadre du projet Capacité d'apprentissage, ces informations sont prises en compte par le biais d'observations. Ces informations permettront de comparer le programme testé offert dans le volet garderie aux activités offertes dans les garderies témoins. C'est à ce niveau que nous observerons si le programme testé a véritablement changé l'environnement de la garderie.

Dans un contexte minoritaire francophone, l'aspect linguistique prend une place particulièrement importante dans l'étude de l'impact d'un programme en garderie. En effet, le fait d'exposer l'enfant à un milieu de garde en français au préscolaire a un impact positif sur sa réussite à l'école. Chartier et ses collègues (2011) ont démontré cette affirmation dans une étude utilisant des données longitudinales portant sur 217 enfants de la communauté francophone du Manitoba.⁷ Les enfants exposés au français dans leur environnement familial et leur milieu de garde ont obtenu des scores supérieurs à l'ÉVIP-R (qui mesure le vocabulaire réceptif) et aux tests de communication et de connaissances générales de l'IMDPE comparativement aux enfants ayant été exposés au français à la maison seulement.

1.4.4. Le mésosystème : liens entre la famille et le milieu de garde

Le mésosystème est composé des liens entre les différents systèmes que l'enfant fréquente. Ces liens contribueraient de façon unique au développement de l'enfant. Plusieurs études appuient l'idée selon laquelle le lien entre le milieu scolaire et la famille contribue au développement de l'enfant. En effet, les enfants dont les parents s'impliquent dans le cheminement scolaire présentent une meilleure adaptation sociale et scolaire que les autres enfants, en plus de posséder des attitudes plus positives envers l'école et de démontrer de meilleures aspirations quant à leur futur, et ce, indépendamment du revenu familial et du niveau d'éducation des parents (Connors et Epstein, 1995). Le même parallèle a été établi entre l'implication des parents en milieu préscolaire et les habiletés de pré-littérature des enfants. Les parents qui s'impliquent en discutant avec l'éducatrice, en lui posant des questions sur le déroulement de la journée de leur enfant et en participant aux activités en garderie, ont des enfants qui démontrent un vocabulaire plus étendu, une meilleure conscience phonologique et de meilleures habiletés de pré-écriture (Arnold, Zeljo, Doctoroff, et Ortiz, 2008). La valeur ajoutée d'un programme combinant les services de garderie et des Ateliers familles est la mise en relief de l'importance d'une bonne collaboration entre les parents et les éducatrices et de l'adoption d'approches communes et complémentaires (Cortier et Pelletier, 2005).

⁷ L'étude se nomme *Parlons petite enfance*, appelée aussi *Étude manitobaine de la cohorte des naissances de 1997*.

1.4.5. Le macrosystème : variables communautaires

De plus en plus d'auteurs reconnaissent l'influence importante des caractéristiques de la communauté sur le développement de l'enfant (Hertzman et Kohen, 2003; Moore, 2005). L'environnement communautaire favorise autant le développement cognitif des enfants que leur état de santé physique et émotionnel (Willms, 2007). Un des facteurs clés influençant le développement des enfants est la disponibilité des ressources et des services offerts seulement en langue française aux familles.

Disponibilité et utilisation des ressources

La disponibilité et l'utilisation des ressources d'une communauté ciblant les jeunes familles d'une communauté s'avèrent des facteurs importants pour le développement des enfants. Selon Connor et Brink (1999), certaines catégories de ressources communautaires sont particulièrement importantes pour le développement des enfants, notamment le système d'éducation et de santé, les divertissements et la culture, les programmes sociétaux, les programmes axés sur les besoins spéciaux, et les sports et les loisirs. Dans le contexte de la minorité francophone, les services et ressources en petite enfance offerts en français se définissent davantage comme un élément protecteur de l'identité francophone et du maintien de la langue française (Commission nationale des parents francophones, 2005). Cette notion d'élément protecteur s'inspire du construit d'infrastructure institutionnelle complète proposée par Breton (1964). À son extrême, une communauté ayant une *infrastructure institutionnelle complète* offrirait à sa population francophone la possibilité de mener toutes ses activités quotidiennes en français. Il s'ensuit que la présence d'institutions de langue française au sein d'une communauté favorise la création de réseaux sociaux et l'accroissement de la cohésion sociale au sein de la communauté. Landry (1994) a repris ce concept d'infrastructure institutionnelle complète dans ses écrits, qui portent sur les quatre formes de capital — démographique, politique, économique et culturel — ayant une énorme influence sur le développement, le maintien (voire même le renouvellement) de la vitalité ethnolinguistique⁸ d'une communauté. C'est par l'entremise de cette vitalité ethnolinguistique que les communautés minoritaires préservent leur sentiment d'appartenance et de fierté, appuyant ainsi l'intégration, et non l'assimilation, de la langue et de la culture francophones au sein de la communauté majoritaire. L'une de ces formes de capital, le capital culturel, « ...se réfère aux ressources et à l'information qui agissent comme agents de transmission de la culture » (Landry, 1994, p. 18, cité dans Guimond, 2003). L'auteur avance le postulat qu'une évaluation de ce capital culturel se voit possible en partie par la diversité des institutions éducatives et l'accès aux ressources culturelles dans la communauté. Landry, Allard, et Deveau (2007b) avancent que les écoles francophones — et particulièrement la présence d'un système scolaire permettant aux jeunes francophones de poursuivre leurs études du préscolaire au postsecondaire en français — ont un rôle crucial à jouer dans l'élaboration, le maintien et la croissance d'une infrastructure institutionnelle complète. Au palier préscolaire, se sont les garderies, les prématernelles, les services de garde en milieu familial, les programmes parascolaires, les centres de ressources et les groupes de jeu qui agissent à titre de porte d'entrée à l'école en français (Gilbert, 2003).

⁸ Voir Guimond (2003) pour un survol des études portant sur la vitalité ethnolinguistique en milieu minoritaire.

1.5. UNE DÉFINITION DE LA PRÉPARATION À LA SCOLARISATION

Comme tous les enfants d'âge scolaire, les petits francophones en milieu minoritaire doivent se préparer à leur entrée à l'école. Cette transition est importante, comme l'indiquent les études dans le domaine. En effet, il existe une forte association entre le niveau de préparation des enfants à leur entrée à l'école et leur réussite scolaire (Lemelin et Boivin, 2007). Certains indicateurs permettent de prédire les dispositions d'un enfant à apprendre à l'école, et ce, dès l'âge de trois ans (Thomas, 2006).

Le terme « préparation à l'école » (« school readiness ») est un concept multidimensionnel, mais qui selon Doherty (1997, p. 25) réfère surtout à la capacité de l'enfant à faire face aux tâches typiques demandées à l'école, telles que de rester assis et d'assimiler la matière. Il englobe aussi les compétences devant être mises en place de la naissance à l'âge de six ans afin d'assurer non seulement la réussite scolaire mais aussi « le succès dans tous les aspects de la vie adulte, notamment sur le marché du travail... ». Cinq aspects de la préparation de l'enfant à l'école sont communs à un ensemble de recherches (Meisels et Atkins-Burnett, 2006) :

- 1) le bien-être physique et le développement moteur approprié;
- 2) la santé émotive et une approche positive face à des expériences nouvelles;
- 3) la connaissance et la compétence sociale appropriées selon l'âge;
- 4) les habiletés linguistiques appropriées selon l'âge; et
- 5) les connaissances générales et aptitudes cognitives appropriées selon l'âge.

Bien que les facteurs 1, 4 et 5 soient les aspects généralement associés à la réussite de tâches précises nécessaires au succès scolaire, les points 2 et 3 permettent aussi de prédire ce succès. En effet, la gestion des émotions, l'attitude générale de l'enfant en classe (l'acte de rester assis toute la journée ou la curiosité face aux apprentissages, par exemple) et les compétences sociales de ce dernier sont autant d'éléments essentiels à la réussite scolaire. Le *National Education Goals Panel* (Meisels et Atkins-Burnett, 2006) reconnaît que ces cinq facteurs ne peuvent être dissociés de la famille, de l'école et de la communauté, et que la préparation à la scolarisation doit être mise en relation avec ces divers niveaux d'influence.

1.6. LE PRÉSENT RAPPORT

Le présent rapport comporte deux études portant sur les deux cohortes. Une première étude a pour objectif l'examen de la fidélité et de la qualité de la mise en œuvre du programme en garderie ainsi que l'examen de la fidélité et de la qualité de la mise en œuvre du programme d'alphabétisation familiale connu sous le nom d'Ateliers familles. Une deuxième étude a pour objectif d'établir les impacts à court terme du nouveau programme préscolaire du volet en garderie sur le développement des enfants notamment au niveau de leur capacité langagière ainsi que sur le niveau de préparation à l'école. Elle vise également à déterminer les impacts du programme des Ateliers familles sur les attitudes et comportements des parents. Précisons que ces analyses se fondent sur les données collectées de mai 2007 à octobre 2009 auprès de la première cohorte de participants et des données collectées de septembre 2008 à octobre 2010 auprès de la deuxième cohorte de participants.

Le chapitre 2 détaille les activités de mise en œuvre des deux volets du programme testé. Le chapitre 3 porte plutôt sur les aspects méthodologiques du projet notamment le devis expérimental adopté, l'échantillon des études, les outils de mesure et les hypothèses des études. Le chapitre 4 détaille la méthodologie employée pour l'étude de la mise en œuvre des deux volets du programme testés. Les chapitres 5 et 6 rapportent, respectivement, les résultats de la mise en œuvre du programme en garderie et du programme d'Ateliers familles. Le chapitre 7 concerne les stratégies mises en place pour assurer la qualité et la fiabilité des données collectées ainsi que le résultat d'analyses préliminaires portant sur l'imputation des données manquantes, la spécification des termes d'erreurs, et l'identification des variables confusionnelles. Le chapitre conclut avec les résultats d'analyses visant à établir le niveau de représentativité de l'échantillon du projet Capacité d'apprentissage relatif à un échantillon de la population générale francophone minoritaire de l'Enquête sur la vitalité des minorités de langue officielle (2006). Le rapport continue avec le détail de l'approche adoptée pour les analyses d'impact contenu dans le chapitre 8. Le chapitre 9 couvre les résultats des analyses d'impact sur les résultantes développementales des enfants de la première année et de la deuxième année de la mise en œuvre du programme. Les résultats d'analyses d'impact du volet Ateliers familles sur différentes dimensions des parents participant à ces derniers se retrouvent au chapitre 10. Nous concluons au chapitre 11 avec un retour sur les principaux constats et une discussion sur l'ensemble des résultats rapportés.

Le lecteur est prié de noter que ce rapport s'insère dans une série de rapports. Il fait suite au rapport *Capacité d'apprentissage dans les communautés francophones en situation minoritaire : Rapport de référence* (Legault, Mák, Verstraete et Bérubé, 2014) dont la version finale a été remise à RHDCC le 13 octobre 2009. Ce premier rapport établissait le profil des enfants, de leur famille et des communautés participant au projet Capacité d'apprentissage. Le présent rapport fait également suite au rapport *Capacité d'apprentissage dans les communautés francophones en situation minoritaire : Rapport de mise en œuvre du projet* (Bérubé, Legault, Janisse, Carson, Saucier et Lefebvre, 2014) et au rapport *Capacité d'apprentissage dans les communautés francophones en situation minoritaire : Rapport des résultats de la première cohorte* (Thompson, Legault, Lalonde et Bérubé, 2014). Le premier rapport fait état des résultats d'analyses effectuées sur les données collectées dans le cadre de l'étude sur la mise en œuvre du programme auprès de la première cohorte alors que le deuxième rapport contient les résultats d'analyses d'impacts du programme sur le développement des enfants de la première cohorte ainsi que sur les attitudes et comportements de leurs parents. Deux autres rapports sont prévus, chacun portant respectivement sur les impacts un an suivant la fin de l'intervention et sur les impacts deux ans après la fin de l'intervention.

2. Activités de mise en œuvre

Ce chapitre décrit les éléments nécessaires à la mise en place et au déroulement des deux volets du programme. Le chapitre débute avec les informations portant sur les démarches entreprises pour le recrutement des communautés et des familles des deux cohortes. Vient ensuite une description des effectifs humains requis pour la mise en œuvre du projet. Le chapitre se poursuit avec des informations au sujet des formations spécifiques offertes aux éducatrices et aux intervenantes en alphabétisation familiale. Il se termine par une description de chacun des volets et de son déroulement.

2.1. RECRUTEMENT DES COMMUNAUTÉS

Première cohorte : La première étape nécessaire à la mise sur pied du projet consistait à rejoindre des communautés intéressées et offrant le potentiel nécessaire pour soutenir le projet durant les années au cours desquelles la prestation du programme testé prendrait place. En vue d'assurer le succès du projet pilote, RHDCC a établi une liste des critères auxquels l'organisme demandeur devait satisfaire pour participer au projet Capacité d'apprentissage. Premièrement, l'organisme demandeur devait représenter une communauté locale et mener des activités dans le domaine de la petite enfance en milieu minoritaire francophone. Deuxièmement, l'organisme demandeur devait être un pourvoyeur de service de garde à but non lucratif ou devait être associé à celui-ci. Troisièmement, la communauté devait être caractérisée par un bassin d'ayants droit francophones suffisant pour permettre le recrutement d'un minimum de 40 enfants âgés de trois ans et fournir une description du profil linguistique de ceux-ci. Quatrièmement, pour se qualifier, une communauté devait identifier des personnes clés prêtes à s'impliquer dans le programme. Enfin, l'organisme demandeur devait démontrer que la communauté possédait les infrastructures et le matériel permettant d'assurer la prestation du programme tout au long du projet pilote, ainsi que le personnel nécessaire pour livrer les services de garde d'enfants dans les garderies.

RHDCC a procédé par appel d'offres afin de recruter les communautés intéressées. Des sept communautés ayant soumis leur candidature, deux n'ont pas été retenues. La première possédait un très bon potentiel à première vue, mais des informations subséquentes sur les caractéristiques des infrastructures en place ont révélé certaines carences. La deuxième offrait déjà un programme de prématernelle dans ses garderies, programme dont le nouveau programme préscolaire à l'étude s'est inspiré. Tôt dans le projet, il est devenu évident qu'il faudrait ajouter une sixième communauté afin d'obtenir la puissance statistique requise pour dégager des résultats fiables. Il a été décidé de développer un nouveau site, mais en utilisant les liens formés avec le conseil scolaire d'une des communautés déjà recrutées, permettant ainsi d'accélérer l'implantation afin qu'elle se déroule en parallèle avec celle des autres communautés du projet Capacité d'apprentissage. Cette communauté avait également l'avantage d'être familière avec le projet Capacité d'apprentissage, ayant participé à une étude pilote plus tôt dans l'année.

Deuxième cohorte : En août 2008, les coordonnatrices des deux communautés ont débuté les activités de pré-implantation. Elles ont d'abord contacté les champions afin d'obtenir leur aval qu'une deuxième cohorte ait lieu dans leur communauté. L'étape suivante était de contacter les

pourvoyeurs de service de garde et les coordinatrices des garderies participant déjà au projet (première cohorte) dans le but d'obtenir leur collaboration lors d'une éventuelle mise en œuvre et évaluation du programme auprès d'une deuxième cohorte.

Dès l'autorisation de RHDCC reçue, les coordonnatrices ont approché toutes les éducatrices et les coordonnatrices des garderies pour bien expliquer les buts et les modalités de leur participation au projet Capacité d'apprentissage et surtout, pour obtenir leur précieuse collaboration au recrutement des familles. À Orléans, pour garantir le nombre voulu de participants à la deuxième cohorte, deux garderies se sont ajoutées au groupe témoin, pour un total de quatre garderies tandis qu'à Cornwall deux nouvelles éducatrices étaient affectées à la garderie programme.

2.2. RECRUTEMENT DES FAMILLES PARTICIPANTES

Plusieurs moyens ou stratégies de communication ont été utilisés pour rejoindre la population dans chacune des communautés choisies. Dans quatre des six communautés composant la première cohorte, les participants ont été recrutés pour composer trois groupes expérimentaux : le groupe programme, le groupe témoin en garderie et le groupe témoin hors garderie. Il n'y a pas eu de groupe témoin en garderie dans les deux autres communautés, faute d'un deuxième milieu de garde francophone pouvant servir de comparaison. Dans les deux communautés de la deuxième cohorte, les participants ont été recrutés dans les trois groupes expérimentaux.

Afin de rejoindre les parents du groupe programme et ceux du groupe témoin en garderie, des démarches ont été faites auprès des garderies francophones de chacune des communautés. Les coordonnatrices communautaires ont approché directement les parents d'enfants admissibles inscrits à la garderie afin de leur expliquer le projet de recherche et les avantages d'y participer. Cette approche a permis de recruter le plus grand nombre d'enfants possible dans plusieurs des communautés. Ensuite, la coordonnatrice prenait rendez-vous avec le parent afin de lui présenter le projet Capacité d'apprentissage d'une manière plus complète, l'inviter à signer le consentement éclairé et remplir le sondage de base. Les coordonnatrices communautaires ont également organisé des séances d'information dans certaines garderies des groupes programme et témoins dans le but de transmettre toutes les informations nécessaires sur le projet Capacité d'apprentissage et de répondre aux questions des parents.

Le recrutement des parents du groupe témoin hors garderie a demandé, quant à lui, une plus grande diversité de démarches. Entre autres, la SRSA a eu recours à des reportages à la télévision et à la radio, la distribution de feuillets d'information dans différents milieux (p. ex., école, cliniques médicales, Centres de la petite enfance, etc.) ainsi qu'à des annonces dans les journaux locaux et régionaux. Dans les faits, c'est la méthode du « bouche-à-oreille » qui a été, de loin, la plus efficace pour recruter les membres de ce groupe témoin.

Ces différentes stratégies ont permis de combler le nombre requis de participants pour les trois groupes expérimentaux, c.-à-d. le groupe programme, le groupe témoin en garderie et le groupe témoin hors garderie et ce, autant pour la première que pour la deuxième cohorte. Ajoutons que la SRSA a profité des réseaux établis au cours des deux premières années du projet pour accélérer le recrutement de la deuxième cohorte. Par exemple, l'aide des parents et des éducatrices de la première cohorte a été sollicitée pour recruter des parents dans la deuxième cohorte.

2.3. EFFECTIFS HUMAINS REQUIS POUR LA MISE EN ŒUVRE

La mise en œuvre du projet a exigé de nombreux effectifs, tant humains que matériels. Ces effectifs ont permis à la fois la mise en place du programme et son évaluation. Les sections qui suivent nous renseignent sur le rôle joué par chacun des acteurs impliqués.

2.3.1. Les champions

Dans chaque communauté, un champion a été identifié pour parrainer le projet. Ce champion pouvait être un individu ou une organisation locale. Il devait avoir une connaissance et une compréhension approfondies du contexte dans lequel vivait la communauté francophone en milieu minoritaire et des défis qu'elle avait à relever. Le champion se devait aussi d'être un bon communicateur et posséder la confiance des parents. En terme concret, il agissait comme ambassadeur du projet auprès des partenaires communautaires et des familles. Son soutien au projet a contribué à établir la crédibilité du projet Capacité d'apprentissage et a servi à promouvoir ce dernier en tant que projet important et bénéfique pour la communauté.

2.3.2. Les coordonnatrices communautaires⁹

Les coordonnatrices communautaires ont d'abord assumé la responsabilité du recrutement des enfants et de leur famille. Par la suite, elles se sont souciées d'entretenir la bonne relation qui s'était établie avec les parents lors du recrutement. Elles se sont assurées de tenir les parents informés des nouveaux développements du projet. Il est à noter que c'est cette bonne relation qui a permis d'assurer un taux élevé de rétention. Les coordonnatrices avaient également la responsabilité de garder les partenaires communautaires informés du déroulement du projet dont les collectes de données et les suivis. Cet échange d'information a contribué à maintenir l'intérêt des partenaires dans le projet et assurer leur appui pour mieux rencontrer les quelques défis.

Au niveau de la mise en œuvre, les coordonnatrices ont été responsables de la logistique rattachée aux formations de base et de suivi. Elles ont également été appelées à jouer un rôle actif dans le cadre du volet Ateliers familles. Elles ont aidé au recrutement des intervenantes en alphabétisation familiale, participé à la préparation du matériel nécessaire à la prestation des ateliers et, au besoin, à son achat, et ont été responsables de l'organisation du repas, du service de halte-garderie et du Centre des ressources.

Enfin, les coordonnatrices communautaires avaient la responsabilité de coordonner les différentes collectes de données dont l'évaluation trimestrielle des enfants. Elles étaient responsables des sondages effectués trois fois par année auprès des parents. Les coordonnatrices ont contribué à la formation des évaluatrices et effectué les suivis nécessaires pour assurer la qualité des données collectées auprès des enfants. Elles ont également assuré la collecte des données sur l'assiduité en garderie et aux Ateliers familles, les observations des activités en garderie et les observations des Ateliers familles. Finalement, elles étaient responsables des entrevues auprès des éducatrices et des représentants communautaires.

2.3.3. Les éducatrices

L'engagement dans le projet et l'enthousiasme des éducatrices pour le nouveau programme se retrouvaient sans contredit parmi les éléments cruciaux qui permettaient d'assurer le succès du

⁹ Le genre féminin est utilisé en reconnaissance de la composition de ce groupe de personnes.

projet. En effet, les responsables de la prestation du programme du volet en garderie étaient sans conteste les éducatrices. Un niveau élevé d'engagement dans le projet Capacité d'apprentissage de leur part contribuait inévitablement à assurer une meilleure qualité du programme livré. En outre, l'engagement des éducatrices facilitait le recrutement et la rétention des parents et de leur enfant au projet Capacité d'apprentissage. L'éducatrice a souvent encouragé le parent qui hésitait à inscrire son enfant ou qui se questionnait sur les modalités du programme et les implications de ce dernier. Les explications complémentaires et les réponses données ont souvent fait une différence dans le choix du parent.

Dans le cadre du programme en garderie, les éducatrices devaient avoir une connaissance approfondie du programme et des résultats d'apprentissage visés. Elles devaient s'approprier le contenu du programme et l'appliquer de façon à respecter l'approche et la philosophie de ce dernier. Ainsi, l'éducatrice se devait de connaître les étapes du développement du langage de l'enfant, de même que les activités permettant à l'enfant d'enrichir son vocabulaire, tout en consolidant son sentiment d'appartenance à la culture et à la communauté francophones. L'éducatrice en chef était responsable de la planification des activités qui étaient en lien avec le thème du mois. Ces activités devaient cibler, par le jeu, la stimulation des cinq sens de l'enfant et le développement des intelligences multiples. Les techniques de francisation utilisées offraient à l'enfant la possibilité d'améliorer et de maîtriser son vocabulaire, de même que de développer son intérêt pour la lecture et la pré-écriture. Le programme mettait l'accent sur des activités de littératie, dont la lecture interactive et l'apprentissage de comptines. L'éducatrice voyait à ce que l'enfant développe une autonomie tant au niveau du contact social avec ses pairs qu'au niveau du choix des activités qui l'intéressaient.

Les éducatrices ont également été invitées à livrer le programme dans le cadre des Ateliers familles, et ce, afin de créer chez l'enfant un sentiment de sécurité. En effet, le déroulement des ateliers prévoyait que les parents et les enfants étaient dans une pièce commune au début et à la fin de l'atelier, mais qu'ils seraient séparés lors de la période intermédiaire. Lors de cette période, les parents se dirigeaient dans un autre local afin de suivre un atelier qui leur était spécialement destiné. La présence de l'éducatrice auprès de l'enfant contribuait à le sécuriser et à optimiser ses apprentissages.

De plus, le soutien que l'éducatrice apportait aux parents participant aux Ateliers familles était un atout important pour le maintien de l'intérêt de ceux-ci à participer aux ateliers. Le rappel régulier des dates de rencontre, du thème qui serait discuté et des activités qui y seraient pratiquées avec les enfants était un moyen qui encourageait la participation régulière des parents. Le projet pilote insistait d'ailleurs sur l'importance du contact journalier et de la communication entre le parent et l'éducatrice de l'enfant. Les échanges et la communication prenant place lors du volet Ateliers familles permettaient d'établir un lien entre parents et éducatrices, et de partager les observations et les découvertes au sujet du cheminement évolutif de leur enfant.

2.3.4. Les intervenantes en alphabétisation familiale

Le volet parent des Ateliers familles était animé par des intervenantes expérimentées dans la prestation d'ateliers auprès d'adultes ou provenant du milieu de l'éducation aux adultes. Elles devaient bien connaître le contenu du programme et les résultats d'apprentissage visés. Elles devaient utiliser, lors de la prestation des ateliers, une approche andragogique et respecter la philosophie du programme. Les intervenantes devaient créer un lien avec les parents afin de

comprendre les difficultés qu'ils rencontraient tant au niveau familial que dans leur milieu communautaire. Elles stimulaient les échanges afin que les parents participent activement et discutent ouvertement des sujets présentés. Elles encourageaient les parents à participer au développement du langage et du vocabulaire de leur enfant en leur présentant des activités à faire à la maison.

De plus, le projet Capacité d'apprentissage demandait aux intervenantes d'assurer une présence aux repas, dans le but de soutenir et d'échanger avec les parents qui avaient des questions supplémentaires, avant ou après l'atelier. L'écoute et le soutien régulier apportés aux parents permettaient de conserver la régularité de leur présence aux Ateliers familles.

2.3.5. Les formatrices

La SRSA a engagé les services de la consultante en petite enfance ayant adapté le programme en garderie afin d'assurer la formation et le suivi des éducatrices en garderie. La principale responsabilité de la formatrice était d'apporter les modifications et ajustements nécessaires afin que les éducatrices puissent livrer le programme fidèlement. Elle a également assuré la formation continue des éducatrices en les conseillant et en leur fournissant des ressources ciblées, et ce, en fonction des besoins qu'elle avait observés ou encore ceux identifiés par les éducatrices.

La SRSA a également engagé les services des conceptrices du programme des Ateliers familles pour veiller à la formation de base et au suivi des intervenantes en alphabétisation familiale. Les formatrices avaient comme principale responsabilité de s'assurer que les intervenantes comprennent bien l'approche, les valeurs et les visées du programme. Les intervenantes pouvaient communiquer, au besoin, avec les formatrices pour obtenir des clarifications sur le contenu du programme et des conseils sur sa prestation.

2.3.6. Les évaluatrices

Les critères d'embauche des évaluatrices exigeaient qu'elles aient la capacité d'établir des contacts et de communiquer facilement avec les enfants. Les évaluatrices devaient en premier lieu suivre une formation de base servant à assurer qu'elles aient les compétences requises pour la mise en application des évaluations auprès des enfants. Dans un deuxième temps, les évaluatrices étaient jumelées deux par deux pour les quatre premiers enfants évalués, afin de leur permettre de bien maîtriser l'outil et d'assurer la validité des données collectées.

Les évaluatrices devaient également connaître et mettre en application les exigences en matière de confidentialité et de protection des renseignements, de manipulation des données collectées ainsi que du transfert des évaluations et des sondages remplis. Elles devaient participer à des formations de suivi avant chaque période d'évaluation. Ces formations de suivi avaient pour but de rappeler aux évaluatrices les modalités d'administration de l'outil de mesure et, au besoin, de les introduire aux modalités d'administration de nouveaux outils de mesure.

2.4. FORMATION

2.4.1. Formation de base des éducatrices pour le programme en garderie

Une formation de base de trois jours a été offerte aux éducatrices, dans chacune des communautés, au cours des mois d'avril et de mai 2007 pour la première cohorte. Pour la deuxième cohorte, la formatrice a offert une mini formation aux éducatrices lors d'un suivi prévu en septembre 2008. Cette décision a été prise parce que la majorité des éducatrices en poste avaient eu une formation de base à l'été 2007, en prévision de la prestation du programme en garderie auprès de la première cohorte. La formation présentait l'aspect théorique ainsi que l'aspect pratique des notions présentées. Parmi les aspects théoriques présentés, notons les objectifs et caractéristiques du projet pilote, le contenu du programme, ses valeurs et ses principes de base, les techniques de francisation, et la communication avec les parents. Les notions pratiques présentées concernaient l'animation d'un livre et d'une chanson ainsi que la planification thématique. Chaque participante a reçu un manuel détaillant le contenu du programme en garderie.

Les échanges entre les participantes ont permis de souligner les différences et similitudes entre le programme en vigueur dans les garderies et le programme testé. Il a aussi permis de recenser les aspects du nouveau programme qui présenteraient des défis particuliers, d'identifier les résistances des éducatrices, et de planifier les séances de suivi en fonction de ces éléments. Cette formation a aussi permis à la formatrice de définir les besoins plus particuliers de chaque éducatrice.

De façon générale, les formations se sont très bien déroulées, les éducatrices se montrant réceptives et enthousiastes vis-à-vis du programme. L'engagement des éducatrices envers le projet Capacité d'apprentissage a également été plus facile dans les communautés où le personnel était déjà familier avec la coordonnatrice communautaire ou lorsque le champion de la communauté était directement impliqué dans le milieu de garde.

2.4.2. Formations de suivi des éducatrices pour le programme en garderie

Suivant la formation de base, des formations de suivi ont eu lieu aux deux mois dans les six communautés, débutant en août 2007 pour la première cohorte et en septembre 2008 pour la deuxième cohorte. Ces formations de suivi ont eu lieu régulièrement, sur une période d'un à deux ans, et ce tant qu'il y avait des enfants participants au projet Capacité d'apprentissage inscrits à la garderie. Aucune formation de suivi n'était offerte durant la période estivale en raison d'une interruption de la programmation régulière dans l'ensemble des garderies.

Dans le cadre de la première cohorte, la première formation de suivi avait pour but de faciliter la réorganisation spatiale des salles en centres d'apprentissage et d'aider les éducatrices à planifier les activités les mois suivants. Cette réorganisation de la salle n'a pas été nécessaire pour la deuxième cohorte puisque la prestation du programme se poursuivait dans les mêmes locaux que ceux utilisés pour la première cohorte. Par la suite, la formatrice observait les activités en classe, offrait une rétroaction aux éducatrices et donnait des formations ciblées. Parmi les sujets abordés durant les suivis, on note la planification thématique, les portfolios des enfants, le développement d'un Centre de littératie, l'intégration des activités de littératie et de numératie dans les activités journalières et le développement du dessin créatif chez l'enfant. Lors de sa dernière visite, la formatrice a invité les éducatrices à donner leurs impressions sur le

programme et à nommer les points forts et les éléments à améliorer dans le programme. Elle a également remis un certificat de participation aux éducatrices qui ont donné le programme en garderie dans le cadre du projet Capacité d'apprentissage. Précisons qu'il y a eu un échec à la mise en œuvre du programme en garderie dans la communauté d'Edmonton.¹⁰ La dernière visite de la formatrice dans cette communauté a eu lieu en septembre 2008.

2.4.3. Formation de base pour le programme des Ateliers familles

Une formation de base d'une durée de trois jours a été offerte, au cours du mois d'octobre 2007. Chaque participante a reçu un manuel détaillant la programmation du volet d'alphabétisation familiale à livrer aux communautés du projet Capacité d'apprentissage. Ce guide se voulait une carte de route et a offert un appui continu important au déroulement des Ateliers familles pour les intervenantes. La formation a permis aux participantes de mieux comprendre la Francophonie minoritaire, de s'initier aux fondements de l'alphabétisation familiale en contexte francophone minoritaire et d'explorer les particularités des familles vivant dans ce contexte. Les formatrices ont fait un survol des dix ateliers et ont présenté les modalités des volets enfant, parent et conjoint. Elles ont fait appel à la participation active des intervenantes pour illustrer les activités obligatoires de chaque atelier. Enfin, les formatrices ont longuement discuté avec les participantes sur la façon d'ajuster le contenu des ateliers pour répondre aux particularités de chaque communauté. Par exemple, certaines intervenantes provenant des communautés qui accueillent plusieurs familles exogames ont exprimé leurs inquiétudes et craintes par rapport au fait que les ateliers ne se dérouleraient qu'en français. Les échanges ont permis de définir des stratégies permettant aux intervenantes de répondre aux besoins des parents anglophones tout en s'assurant que les ateliers demeurent un endroit où le français est privilégié et valorisé.

Une formation de base a également été offerte aux nouvelles intervenantes auprès de la deuxième cohorte en octobre 2008. Un livret de révision a été remis aux anciennes intervenantes (celles impliquées dans la prestation des ateliers auprès de la première cohorte) appelées à offrir les ateliers à la deuxième cohorte. Ce livret visait à rappeler aux anciennes intervenantes les aspects importants du programme des Ateliers familles dont le contenu des ateliers ainsi que l'approche auprès des enfants et des parents vivant en milieu francophone minoritaire.

Notons que les coordonnatrices communautaires ont aussi participé à la formation de base (en octobre 2007 et octobre 2008) afin de se familiariser avec les objectifs et les principes du programme ainsi qu'avec le contenu des ateliers. La formation a permis d'amorcer le travail d'équipe entre la coordonnatrice, les intervenantes auprès des parents et les intervenantes auprès des enfants.

2.4.4. Formation de suivi pour les Ateliers familles

Peu après la fin de la première série d'ateliers, les formatrices ont organisé deux réunions téléphoniques. Ces séances avaient pour but de faire une mise au point sur le déroulement des ateliers, de répondre aux questions des intervenantes et d'offrir des solutions aux difficultés rencontrées. Les formatrices se sont d'abord penchées sur ce qui avait été le plus difficile dans le déroulement des ateliers, et elles ont offert des solutions possibles aux obstacles recensés. La

¹⁰ L'analyse de cet échec se retrouve dans le rapport *Capacité d'apprentissage dans les communautés francophones en situation minoritaire : Rapport de mise en œuvre du projet* remis à RHDCC en août 2010.

problématique d'accommoder les parents anglophones a été abordée. L'une des intervenantes auprès des parents a offert quelques idées pour faciliter l'intégration d'un parent anglophone. Les formatrices sont demeurées disponibles tout au long de la prestation des ateliers afin de répondre ponctuellement aux questions des intervenantes sur le déroulement ou le contenu du programme ainsi que pour offrir des solutions aux difficultés rencontrés.

2.5. DESCRIPTION DU VOLET EN GARDERIE

Le programme livré dans le cadre du projet Capacité d'apprentissage adopte une approche préscolaire ou de préparation à l'école dans lequel le développement des enfants est considéré comme un processus holistique; plusieurs aspects du développement sont donc ciblés. Dans ce type de programme, les enfants sont amenés à atteindre des objectifs de développement précis directement liés à la préparation à l'école et ayant pour but de faciliter leur réussite scolaire. Cette approche s'oppose à une pédagogie dite « sociale » où le programme prévoit des orientations générales et où chaque milieu adopte des éléments en fonction des spécificités de leur communauté. Le but du projet Capacité d'apprentissage étant d'optimiser le développement langagier et global des enfants vivant en contexte francophone minoritaire, le volet en garderie met une emphase particulière sur l'exposition des enfants à la communication verbale par les livres et les chansons, tout en leur offrant de nombreuses occasions de s'exprimer et de développer leur pensée en français.

L'approche du volet en garderie du projet Capacité d'apprentissage préconise la pédagogie par le jeu. Selon cette approche, « les enfants sont vus comme des êtres indépendants qui peuvent contribuer activement à façonner leurs environnements d'apprentissage. » (Bertrand, 2007, p. 4). Enfin, l'approche adoptée encourage la créativité ce qui favorise le processus, l'exploration et l'expérimentation. Par le modèle adopté, l'enfant apprend non seulement des activités de littératie, mais aussi de diverses activités ludiques. L'enfant est encouragé à explorer et à apprendre par le jeu avec le soutien continu de l'éducatrice qui l'encourage et qui observe de près son développement. Le programme vise à influencer la socialisation, la francisation et le développement global de l'enfant (Programme des prématernelles en garderies, *Manuel de formation*, 2007). Les principes fondamentaux du programme offert en garderie incluent:

- Placer les besoins de l'enfant au centre du programme et des activités;
- Stimuler les cinq sens et les intelligences multiples de l'enfant;
- Fournir un environnement riche en langage oral et écrit;
- Utiliser et mettre en application les techniques de francisation dans le but de promouvoir l'utilisation du français;
- Promouvoir l'autonomie;
- Mettre à la disposition de l'enfant des centres d'apprentissage intéressants pour lui et qui l'encouragent à faire des choix;
- Offrir du matériel pédagogique adéquat pour l'enfant et l'installer à sa hauteur;
- Favoriser des interactions positives avec l'enfant et avec les parents;
- Compléter le programme en garderie par un programme d'alphabetisation familiale.

La programmation du volet en garderie s'inspire du Programme fransaskois de la prématernelle développé par le ministère de l'Éducation de la Saskatchewan (2001) pour des enfants de quatre ans. Embauchée par RHDCC, une consultante en petite enfance a adapté le programme pour la population plus jeune du projet Capacité d'apprentissage, soit des enfants de trois ans. En outre, elle a apporté plusieurs modifications au programme en garderie en ajoutant des éléments du programme *Jouer c'est magique*, notamment au niveau de l'horaire quotidien et de la durée de la journée en milieu de garde. Enfin, avec le concept de l'approche créative à l'apprentissage, la consultante a :

- créé ou modifié les outils d'évaluation, tels que les grilles d'observation et le portfolio de l'enfant;
- précisé des thèmes mensuels et simplifié les grilles de planification thématique;
- précisé la méthode d'animation d'un livre et d'une chanson; et
- simplifié le nombre et le contenu des centres d'apprentissage; et
- dressé un liste de ressources et de matériels nécessaire à la prestation du programme.

L'uniformité de la prestation du programme est assurée au moyen de 15 trousse comprenant les ressources et le matériel nécessaire. Le contenu de 12 des 15 trousse se rattachent à un thème particulier mensuel. Il y a également une trousse d'art, une trousse de science et une trousse de musique. Le contenu des trousse inclut des jouets, des casse-tête, des livres, des disques compacts, des instruments de musique et du matériel d'art plastique. Tout le matériel était en français.

2.5.1. Déroulement du programme en garderie

Première cohorte : La première année du programme en garderie regroupe les mois de septembre 2007 à août 2008. Le programme a été inauguré le 1er septembre 2007 dans les communautés de Cornwall, Edmonton, Edmundston et Saint-Jean. Il a été lancé en octobre 2007 dans les communautés de Durham et d'Orléans. La deuxième année du programme regroupe les mois de septembre 2008 à août 2009, et a eu lieu dans quatre des six communautés seulement. Les enfants des communautés d'Edmonton, d'Edmundston et de Saint-Jean ont été exposés au programme en garderie à temps plein alors que les enfants de la communauté de Cornwall y ont été exposés à temps partiel. Les enfants des communautés de Durham et d'Orléans n'ont pas été exposés au programme en garderie durant la deuxième année. Dans ces communautés, les parents ont inscrit leur enfant à la maternelle de quatre ans qui était offerte gratuitement à temps plein. Notons que le programme testé n'était pas fidèlement appliqué lors de la saison estivale, et ce, autant dans les garderies livrant le programme du projet Capacité d'apprentissage que les garderies témoins livrant un autre programme.

Deuxième cohorte : La première année du programme en garderie pour la deuxième cohorte regroupe les mois de septembre 2008 à août 2009. Le programme a simplement continué d'être appliqué dans les communautés de Cornwall et d'Orléans pour le bénéfice d'une deuxième cohorte. Comme pour la première année, nous avons observé que le programme testé n'était pas fidèlement appliqué lors de la saison estivale dans les garderies livrant le programme du projet Capacité d'apprentissage et les garderies témoins. Le calendrier de prestation du programme est présenté au tableau 2.1.

Tableau 2.1 : Calendrier de prestation du programme préscolaire

Communauté	Date d'inauguration	Date de fermeture	Degré d'exposition
Cornwall – Cohorte 1	28 août 2007	24 août 2009	À temps plein la première année et à temps partiel la deuxième année
Cornwall – Cohorte 2	1 septembre 2008	21 juin 2010	À temps plein la première année et à temps partiel la deuxième année
Durham	24 septembre 2007	31 juillet 2009	Un an à temps plein pour la majorité des enfants (2007 à 2008) et deux ans à temps plein pour quelques enfants
Edmonton	17 septembre 2007		Échec de la mise en œuvre
Edmundston	4 septembre 2007	19 juin 2009	Deux ans à temps plein
Orléans – Cohorte 1	1 octobre 2007	29 août 2008	Un an à temps plein
Orléans – Cohorte 2	1 septembre 2008	28 août 2009	Un an à temps plein
Saint-Jean	4 septembre 2007	27 août 2009	Deux ans à temps plein

2.6. DESCRIPTION DU VOLET ATELIERS FAMILLES

Le programme Ateliers familles a été développé spécifiquement pour le projet pilote par Éduk, en collaboration avec RHDCC et la SRSA. Le programme a été conçu pour répondre aux objectifs du projet Capacité d'apprentissage et aux besoins particuliers des francophones en situation minoritaire. Ainsi, il vise le développement de l'enfant et non l'amélioration du niveau d'alphabétisme du parent. L'emphase est mise sur le développement des connaissances, les divers vécus et périodes de socialisation des enfants. Le programme cherche à outiller les parents afin qu'ils puissent appuyer le développement langagier, culturel et identitaire de leur enfant en français, selon qu'ils vivent en contexte familial unilingue, bilingue, trilingue ou multiculturel. Il vise aussi à conscientiser les parents au travail de l'éducatrice et à l'importance des rôles complémentaires parent-éducatrice pour soutenir les apprentissages de l'enfant.

Le contenu s'inspire des forces de grands programmes canadiens reconnus en littératie, dont :

- *Grandir avec mon enfant* (2002) et ses adaptations notamment *J'apprends en famille* de la Nouvelle-Écosse, en particulier en ce qui concerne les activités en lien avec les compétences parentales et les besoins des enfants;
- *Chansons, contes et comptines*, de même que *Grandir avec des livres* compte tenu de l'accent qui est mis sur l'éveil à l'écrit et la composante culturelle francophone;
- le programme anglophone « *Learning Together* » qui a fait l'objet d'une étude longitudinale en Alberta (2001 à 2005) et dont les résultats auprès des enfants et des familles sont bien documentés (Phillips, Hayden et Norris, 2006); et
- le *Programme fransaskois de la prématernelle* (2001).

2.6.1. L'arrimage des deux volets du programme

Pour maximiser l'impact du projet Capacité d'apprentissage sur les familles, il est indispensable d'arrimer le programme en garderie et le programme d'alphabétisation familiale en termes de valeurs, de principes fondamentaux et d'approches adoptés dans les deux programmes. Le programme en garderie met l'accent sur la francisation et l'éveil à l'écrit chez l'enfant d'âge préscolaire (en incluant des sous-thèmes de lecture et d'écriture). Les Ateliers familles complètent le programme en garderie au moyen d'échanges visant à conscientiser les parents sur leur rôle comme premier éducateur auprès de leur enfant et sur les particularités de la vie en milieu linguistique minoritaire. Les principaux thèmes abordés portent sur la façon d'appuyer le développement de leur enfant, de stimuler ses apprentissages et de lui transmettre la langue et la culture françaises. Le contenu vise aussi à conscientiser le parent sur le travail de l'éducatrice en plus de favoriser un partenariat entre les différents intervenants auprès de l'enfant (éducatrice, parents et autres).

Pour lier ces volets, les conceptrices des Ateliers familles ont travaillé en étroite collaboration avec la consultante en petite enfance ayant adapté le programme en garderie. Cette stratégie avait comme avantages d'assurer une cohérence entre les différents aspects du projet Capacité d'apprentissage et une continuité au niveau des apprentissages de l'enfant et de son parent. Les conceptrices ont tenu compte:

- de l'approche et des valeurs qui sous-tendent le programme mis en place à la garderie;
- des stratégies de francisation à mettre en pratique auprès de l'enfant;
- des thèmes traités chaque mois à la garderie;
- de la liste des ressources — jeux, jouets et livres achetés pour le volet en garderie; les ressources proposées pour les Ateliers familles complétaient celles du volet en garderie; et
- du stade développemental des enfants d'âge préscolaire.

L'arrimage des deux volets est, en soi, une importante contribution au développement de l'enfant. Les parents qui s'impliquent en discutant avec l'éducatrice, en lui posant des questions sur le déroulement de la journée de leur enfant et en participant aux activités en garderie ont des enfants qui démontrent un vocabulaire plus étendu, une meilleure conscience phonologique et de meilleures habiletés de pré-écriture (Arnold, et al., 2008). Le programme mis en place en garderie et celui des Ateliers familles mettent en lumière l'importance d'une bonne collaboration entre les parents et les éducatrices. L'arrimage des deux volets favorise un partenariat entre les différentes personnes intervenant auprès de l'enfant (éducatrice, parents et autres) et vise à renforcer les apprentissages de l'enfant vécus dans différentes sphères de sa vie, telles que la garderie de jour, son foyer familial, et la communauté.

2.6.2. Déroulement du programme Ateliers familles

La prestation des Ateliers familles s'est faite uniquement la première année de la mise en œuvre du programme testé. Le programme des Ateliers familles comprenait dix ateliers présentés en deux séries. Pour la première cohorte, une première série de quatre ateliers a été offerte en novembre et décembre 2007 et une deuxième série de six ateliers a été offerte aux mois de

janvier et février 2008. Pour la deuxième cohorte, une première série de cinq ateliers a été offerte en novembre et décembre 2008 tandis que la deuxième série de cinq ateliers a été présentée de janvier à février 2009. Seules les familles du groupe programme étaient appelées à participer au volet d'alphabétisation familiale.

Les ateliers sont offerts dans les locaux de la garderie dans le but d'assurer aux enfants un environnement sécuritaire et adapté à leurs besoins (p. ex., les tables et les chaises sont de taille appropriée). De plus, la garderie constitue un endroit connu par les enfants et, par conséquent, favorise chez eux le sentiment de sécurité, un facteur indispensable lorsque l'on veut optimiser leurs apprentissages. Chaque atelier est d'une durée totale de deux heures et comporte cinq volets : le volet d'accueil, le volet parent, le volet enfant, le volet conjoint et le volet de clôture. Les parents et les enfants sont ensemble, dans le même local, durant le volet accueil, le volet conjoint et le volet de clôture. Le volet parent et le volet enfant prennent place en parallèle dans des locaux distincts. Le cœur du programme prend place durant le volet parent, moment où les informations sont transmises aux parents suivies de discussions et d'échanges par l'entremise desquelles les parents partagent leur vécu afin d'apprendre les uns des autres. Cette approche permet de bâtir sur les forces des familles et de valoriser les pratiques existantes des parents. L'intervenante agit à titre de personne-ressource pour encourager les discussions, faire ressortir les points importants, et diriger les familles vers les informations dont elles ont besoin.

L'emprunt d'une trousse éducative est prévu lors du volet de clôture. Il s'agit de sacs de tissu à apporter à la maison et à rapporter à l'atelier suivant. Chaque trousse comprend deux des ressources en français suivantes : un livre pour le parent ainsi qu'un CD de musique, un DVD ou un jeu éducatif. Les familles rapportaient la trousse à l'atelier suivant, pour en recevoir une nouvelle. La trousse contenait également les matériaux et la démarche écrite pour faire une activité créatrice à la maison avec l'enfant. On compte dix activités créatrices différentes, soit une à insérer dans la trousse à la fin de chaque atelier. Les enfants conservent le matériel. L'activité créatrice vise à renforcer ce que les enfants ont appris à la garderie et durant l'atelier. Les familles sont aussi invitées à visiter le Centre de ressources établi spécifiquement pour les ateliers. Ce centre comprend des ressources de langue française dont des livres pour enfants et pour les parents, des CD, des DVD, des livres/disques et des jeux. Chaque enfant peut choisir un livre ou une autre ressource et l'ajouter à sa trousse.

Plusieurs éléments ont été mis en place pour diminuer les obstacles à la participation des familles aux ateliers notamment l'offre d'un repas avant ou suivant l'atelier. La possibilité de prendre le repas avant que ne débute l'atelier facilite surtout la participation des familles qui n'ont pas à se rendre à la maison pour souper et revenir à la garderie lorsque les ateliers sont offerts en fin de journée. Similairement, lorsque les ateliers sont livrés en matinée, la possibilité de dîner en famille avant le retour à la maison dérange moins l'horaire des jeunes enfants qui souvent font la sieste dans l'après-midi. Là où le nombre le justifie, un service de halte-garderie est offert pour la fratrie des enfants du groupe programme âgés de deux ans et demi ou plus. Les familles peuvent également obtenir un dédommagement fixe pour les frais de gardiennage encourus pour les enfants âgés de moins de deux ans et demi. Enfin, les coordonnatrices communiquaient régulièrement avec les familles pour confirmer leur participation à l'atelier. Par exemple, les coordonnatrices ont effectué des appels, envoyé des courriels ou distribué des lettres à tous les parents inscrits aux Ateliers familles pour leur rappeler la reprise des ateliers au mois de janvier.

2.6.3. Centre de ressources — achat de matériel

L'uniformité de la prestation du programme est assurée par l'achat de plus de 300 ressources de langue française. Les familles des première et deuxième cohortes sont invitées à emprunter ces ressources durant et suivant la fin des ateliers familles. Notons que durant l'année 2008–2009, les parents et les enfants des deux cohortes pouvaient emprunter les livres au Centre de ressources. Environ 60 ressources additionnelles de langue française ont donc été achetées en septembre 2008 pour desservir convenablement le plus grand nombre de familles visitant le Centre de ressources. La majorité des ressources acquises visaient à remplacer le matériel manquant ou abimé suite à la prestation des Ateliers familles auprès de la première cohorte. La SRSA a également augmenté le nombre de livres pour les parents.

Suivant la prestation des Ateliers familles, les coordonnatrices communautaires ont établi un système de prêt de livres comprenant l'ensemble des ressources utilisées dans le cadre des ateliers. L'emprunt de ces ressources se faisait à deux moments au cours d'une semaine. Le jour, les coordonnatrices visitaient les salles de classes des enfants du groupe programme afin de leur permettre d'emprunter un livre au Centre de ressources. En fin de journée, elles se présentaient à la garderie programme afin de permettre aux parents de visiter le Centre de ressources en compagnie de leur enfant. C'était l'occasion pour les familles d'emprunter une ressource pour les enfants et les parents. Précisons que le Centre de ressources était fermé lors de la période estivale.

Les activités du Centre de ressources se sont poursuivies au cours de la deuxième année dans chacune des communautés du projet Capacité d'apprentissage, que les enfants du groupe programme fréquentent ou non la garderie. Ainsi, les familles des communautés de Edmonton, Durham, Edmundston, et Saint-Jean ont pu emprunter des ressources jusqu'en juin 2009. Les familles des communautés de Cornwall et d'Orléans (où se retrouvait une deuxième cohorte) ont pu emprunter des ressources jusqu'en juin 2010. Le Centre de ressources était, par ailleurs, fermé durant la période estivale. Les coordonnatrices ont fait l'inventaire du Centre de ressources à plusieurs reprises pendant et suivant la prestation du programme Ateliers familles et l'établissement du Centre de ressources. À la fin du projet Capacité d'apprentissage, le contenu du Centre de ressources a été remis au pourvoyeur responsable des services offerts à la garderie programme.

3. Méthodologie

Ce chapitre porte sur les aspects méthodologiques utilisés dans le cadre du projet Capacité d'apprentissage. Une première section décrit les critères d'admissibilité et la population ciblée tandis que la seconde présente le devis expérimental du projet. Dans une troisième section, on explore certaines des menaces à la validité interne et les stratégies mises en place pour les contrer. La quatrième section détaille les mesures utilisées dans les analyses d'impact du programme testé. La dernière section concerne la série d'hypothèses testées dans le cadre des analyses d'impact.

3.1. POPULATION CIBLÉE

Le recrutement des parents et des enfants s'est appuyé sur des critères d'admissibilité spécifiques. Le *premier critère* était que l'enfant soit né entre le 1^{er} janvier 2004 et le 31 janvier 2005 pour la première cohorte, et en 2005 pour la deuxième cohorte. Ce critère correspondait au critère d'admissibilité à l'inscription à la maternelle en septembre 2009 pour la première cohorte et en septembre 2010 pour la deuxième cohorte des ministères de l'Éducation de l'Ontario et du Nouveau-Brunswick. Le *deuxième critère* établissait que l'un des parents devait être un ayant droit, tel que défini par l'article 23 de la *Charte canadienne des droits et libertés*.¹¹ Puisque le programme testé visait à augmenter les compétences langagières et le niveau de préparation à la scolarisation des enfants, ce critère visait à assurer de rejoindre la population cible, c'est-à-dire les enfants ayant le droit de fréquenter l'école en français.

Le *troisième critère* concernait l'intention du parent d'envoyer son enfant dans une école de langue française. Ce critère a été en réalité peu appliqué puisque les parents d'enfants aussi jeunes n'ont généralement pas pris leur décision quant au choix de l'école. Par contre, si les parents répondaient que l'école anglaise était leur choix final, la SRSA prenait la décision de ne pas poursuivre avec l'obtention du consentement éclairé étant donné que le nouveau programme préscolaire vise en partie à mieux préparer les enfants à l'école en français.

3.2. DEVIS EXPÉRIMENTAL

L'évaluation du programme se fait au moyen d'un devis quasi expérimental avec groupes témoins non équivalents. Tout comme les études expérimentales effectuées sur le terrain, un devis quasi expérimental a pour but de tester, au moyen de groupes témoins et de mesures pré-intervention (c.-à-d. mesures prises avant l'intervention), l'hypothèse causale selon laquelle une intervention a un effet significatif sur les variables à l'étude au-delà de ce qui prend place en l'absence de ce programme et indépendamment d'autres facteurs externes pouvant entrer en jeu.

La méthodologie prévoit trois groupes expérimentaux : un groupe programme composé d'enfants inscrits dans des garderies francophones offrant le nouveau programme préscolaire; un groupe témoin en garderie comprenant des enfants inscrits dans des garderies francophones n'offrant pas le nouveau programme; et un groupe témoin hors garderie regroupant des enfants dont la garde de jour a lieu à la maison ou en garderie familiale non réglementée. Le premier

¹¹ *Charte canadienne des droits et libertés* disponible au <http://laws.justice.gc.ca/fra/charte/1.html>.

groupe témoin vise à tenir compte de l'influence d'une garderie en installation sur le développement des enfants, ce qui constitue en soi un traitement. Le deuxième groupe témoin vise à tenir compte de l'influence d'un milieu de garde informel sur le développement de l'enfant, surtout au niveau de la langue française. Les enfants inscrits à une garderie de langue anglaise ont été ajoutés au groupe témoin hors garderie pour les analyses d'impact. Bien que ces enfants soient exposés à un environnement de garde formel, ils ne sont pas exposés au français. En ne classifiant pas ces enfants dans le groupe témoin en garderie, nous préservons l'homogénéité de la langue dans le groupe témoin en garderie et l'exposition à un programme en garderie offert en français.

3.3. VALIDITÉ INTERNE

En l'absence d'une assignation aléatoire, il est probable qu'il existe au préalable des différences inhérentes entre le groupe programme et les groupes témoins. La mise en place de conditions visant à assurer la validité interne de l'étude devient encore plus importante afin d'éliminer d'emblée les explications alternatives *plausibles* des résultats. En effet, l'utilisation de techniques de contrôle, souvent statistiques, une fois l'étude terminée est moins avantageuse.

Dans le cadre du projet Capacité d'apprentissage, les conditions nécessaires pour assurer la validité interne sont :

- la taille de l'échantillon;
- une stratégie d'échantillonnage assurant une distribution équitable entre les groupes;
- des mesures pré-intervention prises sur les effets escomptés du programme (ci-après appelées « résultantes ») et les facteurs qui leur sont associés;
- une vérification de l'effet d'attrition de l'échantillon sur la distribution des groupes;
- une attention particulière portée aux situations pouvant mener à la contamination des groupes témoins.

3.3.1. La taille de l'échantillon

Pour s'assurer de la validité interne de l'étude d'impact avec trois groupes, on doit former un échantillon composé au minimum de 156 enfants répartis également entre le groupe programme ($n = 52$), le groupe témoin en garderie ($n = 52$) et le groupe témoin hors garderie ($n = 52$). Ce nombre d'enfants par groupe permet d'obtenir la puissance statistique nécessaire pour détecter un impact moyen, avec un très bon niveau de confiance que la valeur réelle de l'impact estimé se situe à l'intérieur d'un intervalle spécifique (c.-à-d. 19 fois sur 20, nous obtiendrions les mêmes résultats auprès d'autres échantillons).

3.3.2. Stratégie d'échantillonnage ciblé

Une stratégie d'échantillonnage ciblé a été choisie dans le but de créer des groupes témoins relativement homogènes au groupe programme de façon à éliminer l'influence de facteurs autres que ceux déjà mesurés sur les résultantes à l'étude. Des études antérieures ont démontré que certaines caractéristiques, comme le revenu familial et le niveau d'éducation des parents, expliquaient une partie de la trajectoire développementale de l'enfant. Il s'ensuit que si le groupe

programme diverge considérablement du groupe témoin sur l'une de ces caractéristiques, les différences observées entre les deux groupes pourraient bien être expliquées par ces dernières et non par le programme. Ainsi, une attention particulière a été apportée afin de cibler des participants des groupes témoins présentant un profil sociodémographique similaire à celui du groupe programme (p. ex., le niveau socio-économique) et vivant dans le même arrondissement, assurant ainsi qu'ils aient accès aux mêmes ressources et services francophones que le groupe programme.¹² Un premier contrôle sur la composition des groupes a été fait lors de l'inscription des familles au projet, notamment au niveau de l'emplacement du domicile des membres potentiels des groupes témoins. Un deuxième contrôle fut l'utilisation des mesures pré-intervention. C'est de cela dont il sera question dans la prochaine section.

3.3.3. Les mesures pré-intervention

L'inférence causale découlant d'une méthodologie quasi expérimentale est facilitée par l'utilisation de mesures pré-intervention prises sur les résultantes et les facteurs qui leur sont associés. Dans le cas du projet Capacité d'apprentissage, la résultante à l'étude est le niveau de préparation à la scolarisation des enfants mesuré au moyen de l'Évaluation de la petite enfance — Appréciation directe (ÉPE-AD) et de ses sous-échelles de vocabulaire expressif et réceptif. Notons que l'ensemble des résultats présentés dans ce rapport découlent d'analyses menées exclusivement sur la version française de l'outil ÉPE-AD.

Des mesures pré-intervention permettent de mieux saisir l'impact du programme sur le développement de l'enfant puisqu'on pourra effectuer une comparaison du niveau de préparation à la scolarisation de tous les enfants au moment précédent la mise en œuvre du programme avec le niveau observé suite à cette mise en œuvre. Ces mesures permettent également de vérifier si les enfants du groupe programme et ceux des groupes témoins présentent un profil développemental similaire avant l'intervention. Cette information est utile puisque les estimés de l'impact du programme se trouveraient biaisés si l'un des groupes expérimentaux se composait d'enfants démontrant un développement plus avancé comparativement à un autre groupe avant que ne débute l'intervention.

Des mesures pré-intervention des variables sociodémographiques et socio-économiques visent également à établir si les enfants des trois groupes expérimentaux vivent des conditions similaires avant l'intervention. Le choix des variables repose sur les résultats d'études antérieures démontrant que ces dernières influençaient significativement le niveau de préparation à la scolarisation des enfants. Le contrôle statistique de ces variables dans les analyses permettra de distinguer les impacts du nouveau programme préscolaire en garderie sur le développement des enfants des impacts d'autres variables pour lesquels les évidences empiriques ont déjà été établies.

Le profil initial des enfants, de leur famille et des communautés participant au projet Capacité d'apprentissage a été rapporté dans un document intitulé *Capacité d'apprentissage dans les communautés francophones en situation minoritaire : Rapport de référence* (Legault et al., 2014). Ce rapport contient également le résultat d'analyses établissant l'homogénéité entre les groupes expérimentaux avant que l'intervention n'ait lieu. Notons que les résultats rapportés dans le rapport découlaient d'analyses portant sur les six communautés. Ces analyses ont dû être

¹² Le lecteur intéressé à en apprendre plus est invité à lire le Rapport sur les plans de travail et de méthodologie révisés remis à RHDCC le 30 mars 2007.

refaites pour le présent rapport pour inclure uniquement les quatre communautés incluses dans les analyses d'impact. La liste des variables identifiées se retrouve à la Section 7.3 de ce rapport alors que les résultats des analyses comparatives sont compris dans la Section 7.4 de ce rapport.

3.3.4. Attrition

La perte de participants au cours de l'étude entraîne une menace à la validité externe de l'étude (c.-à-d. jusqu'à quel point l'échantillon représente la population qu'il est censé représenter) ainsi qu'à la validité interne (c.-à-d. en changeant la composition des groupes expérimentaux). Elle peut être due à plusieurs raisons dont un déménagement ou encore une perte d'intérêt des participants. Peu importe la raison du retrait, il importe d'établir si le groupe de participants qui se sont retirés de l'étude se distingue du groupe de participants demeurant dans l'étude quant à certaines dimensions sociodémographiques. Des différences entre les deux groupes pourraient signaler la présence d'un sous-groupe de la population cible possédant des caractéristiques particulières (ce qui constitue une menace à la validité externe). En outre, un changement important dans la composition des groupes expérimentaux pourrait affecter la validité des estimés d'impact (c.-à-d. affecter la validité interne). Un contrôle des raisons du retrait au projet a donc été instauré dès le début du projet afin de mieux évaluer la menace à la validité interne de l'étude découlant des circonstances rattachées aux retraits.

3.3.5. Prévention de la contamination des groupes témoins

Il y a contamination lorsque des changements dans les scores obtenus sur les résultantes des groupes témoins sont dus à la mise en application des conditions et modalités du programme testé. En d'autres mots, il y a contamination lorsque les membres des groupes témoins sont exposés aux composantes du programme en garderie et/ou ceux du programme Ateliers familles.

La contamination du groupe témoin en garderie est plus probable dans les communautés où plus d'une des garderies francophones de la région participent au projet, rendant ainsi possible la circulation des informations entre la garderie programme et la garderie témoin. Afin de minimiser les occasions de transfert d'information, plusieurs stratégies ont été mises en place (une liste est fournie ci-dessous). En outre, les pourvoyeurs des services de garde ont été sensibilisés aux situations propices à une contamination des groupes témoins et à l'importance pour l'étude d'éviter ou de minimiser ces occasions. Cette sensibilisation était surtout ciblée dans les communautés comptant plusieurs garderies francophones sous l'égide d'un même pourvoyeur de services de garde. Ces discussions ont résulté en une collaboration étroite entre la SRSA et les pourvoyeurs de services de garde.

Les stratégies mises en place pour minimiser les occasions de contamination étaient :

Au niveau du volet en garderie :

- Des formations de base et de suivi étaient offertes seulement aux éducatrices et aide-éducatrices impliquées dans la livraison du programme en garderie.
- Le mouvement des éducatrices formées au programme en garderie à une garderie témoin était limité. Dans les faits, le mouvement du personnel d'une garderie programme à une garderie témoin a eu lieu seulement dans une communauté. Dans un cas, l'éducatrice a été affectée à un groupe d'enfants plus jeunes que les enfants du projet Capacité d'apprentissage. Dans les autres cas, les aides-éducatrices ont été

affectées à des garderies ne participant pas au projet Capacité d'apprentissage. Tout personnel quittant la garderie programme était sensibilisé à l'importance de ne pas partager les connaissances et les façons de faire avec les autres éducatrices et était tenu de remettre à la coordonnatrice le manuel de formation. Enfin, la coordonnatrice surveillait de près la situation dans la garderie témoin.

- Les classes du groupe témoin en garderie se situaient dans un édifice distinct de celui dans lequel se trouvaient les classes du groupe programme.
- Les ressources et le matériel du programme en garderie devaient rester dans la garderie programme jusqu'à trois mois suivant la fin de la prestation du programme testé.

Au niveau du volet Ateliers familles :

- La formation de base et le suivi étaient offerts seulement aux intervenantes impliquées dans la livraison des Ateliers familles.
- À la fin de la prestation du volet Ateliers familles, les intervenantes devaient remettre à la coordonnatrice le manuel de formation. Elles étaient également sensibilisées à l'importance de ne pas partager les connaissances et les façons de faire avec d'autres personnes.
- Les pourvoyeurs de services de garde et les conseils scolaires étaient tenus de ne pas livrer le programme Ateliers familles ou un programme similaire dans la communauté à l'étude.
- Les ressources et le matériel du Centre de ressources devaient rester dans la garderie programme jusqu'à trois mois suivant la fin de la prestation du programme testé.

3.4. ÉCHANTILLON PAR COMMUNAUTÉS ET PAR GROUPES EXPÉRIMENTAUX

La première cohorte de participants au projet est issue des communautés minoritaires francophones de Cornwall, Orléans et Durham en Ontario, et d'Edmundston au Nouveau-Brunswick.¹³ La période de recrutement de cette première cohorte s'est étendue de mai à octobre 2007. Une deuxième cohorte de participants au projet réside dans les communautés minoritaires francophones de Cornwall et d'Orléans en Ontario. La période de recrutement de cette deuxième cohorte a eu lieu à l'automne 2008. Le lecteur est prié de noter que les statistiques décrivant l'échantillon rapportées dans cette section concernent les participants des deux cohortes combinés.

Échantillon global : Au moment de l'inscription, le projet comptait 356 enfants provenant de 352 familles. Tel qu'observé au tableau 3.1, les communautés d'Edmundston et de Cornwall (première cohorte) comptaient les plus grandes proportions de participants (23,9 % et 20,2 %, respectivement), suivies des communautés d'Orléans (15,5 % pour la première cohorte et 15,7 % pour la deuxième cohorte), Cornwall (12,9 % pour la deuxième cohorte) et Durham (11,8 %).

¹³ L'absence de garderies francophones pouvant servir de contrefactuel rendait impossible de bien mesurer l'impact de ce programme dans les communautés d'Edmonton et de Saint-Jean.

L'échantillon comprenait 341 enfants à la fin de la première année d'intervention (les deux cohortes combinées), dont 167 garçons et 174 filles. La taille de l'échantillon a diminué légèrement à 338 enfants à la fin de la deuxième année d'intervention, dont 166 garçons et 172 filles.

L'âge moyen des enfants à l'inscription était de 38 mois. À ce moment, le projet Capacité d'apprentissage comptait un nombre équilibré de garçons (49,4 %) et filles (50,6 %). La langue maternelle des enfants de l'échantillon (selon le répondant qui connaît le mieux l'enfant) était en majorité le français (71,1 %) suivi par l'anglais ou une autre langue (19,4 %).

D'après le sondage de base, l'âge de la mère à la naissance de l'enfant ciblé était en moyenne de 27,5 ans. De plus, 80 % de ces mères ont obtenu au moins un diplôme collégial, et la moitié de ces dernières ont un diplôme universitaire. La taille du ménage se chiffrait à quatre membres en moyenne. On observait 8,4 % de familles monoparentales. Plus de la moitié des familles participantes (59,1 %) avaient un revenu annuel de plus de 70,000\$; le revenu annuel médian se retrouvait entre 80,000\$ et 99,999\$.

Au niveau du profil linguistique des familles participantes, plus de la moitié des mères (61,7 %) et des pères (55,0 %) parlaient seulement français à leur enfant. La majorité des enfants provenaient de foyers endogames francophones (49,3 %), suivi de foyers exogames (39,2 %).¹⁴

Tableau 3.1 : Répartition des participants par communauté depuis l'inscription

	Inscription	1 ^{ère} année	2 ^e année
Cornwall - Cohorte 1	72 (20,22 %)	69 (20,23 %)	68 (20,12 %)
Cornwall - Cohorte 2	46 (12,92 %)	45 (13,2 %)	45 (13,31 %)
Durham	42 (11,8 %)	36 (10,56 %)	35 (10,36 %)
Edmundston	85 (23,88 %)	83 (24,34 %)	83 (24,56 %)
Orléans - Cohorte 1	55 (15,45 %)	55 (16,13 %)	54 (15,98 %)
Orléans - Cohorte 2	56 (15,73 %)	53 (15,54 %)	53 (15,68 %)
Total	356 (100 %)	341 (100 %)	338 (100 %)

Échantillon par groupes expérimentaux : Au moment de l'inscription, le projet comptait 110 enfants inscrits au groupe programme (G1), 135 enfants au groupe témoin en garderie (G2) et 111 enfants au groupe témoin hors garderie (G3). L'examen du tableau 3.2 révèle une diminution du G1 et du G3 accompagnée d'une augmentation du G2 seize mois après l'inscription (c.à.d. +16 mois).¹⁵ Ces changements dans la composition des groupes expérimentaux s'expliquent par le fait que les enfants de l'Ontario étaient inscrits à l'école à partir du début de la deuxième année du projet. En outre, ces fluctuations sont plus importantes à Cornwall où plusieurs enfants du G1 et du G3 ont été inscrits à une école abritant une garderie

¹⁴ La catégorisation des foyers s'est faite à partir de la combinaison de la Première langue officielle parlée (PLOP) de la mère et de la PLOP du père.

¹⁵ Bien que les enfants soient inscrits à l'école dès douze mois après l'inscription (+12 mois), les changements dans la composition des groupes expérimentaux ne paraissent pas immédiatement dans la collecte de données en raison des règles d'attribution. Ces dernières sont détaillées à la section 8.4.

témoin. Les changements dans la composition des groupes expérimentaux ont donc été pris en considération dans les analyses de la deuxième année du programme.

Tableau 3.2 : Répartition des participants par groupe expérimental

	Inscription	+ 4 mois	+ 8 mois	+ 12 mois	+ 16 mois	+ 20 mois	+ 24 mois
G1	110 (30,9 %)	104 (29,7 %)	97 (28,0 %)	97 (28,4 %)	94 (27,6 %)	91 (26,9 %)	91 (26,9 %)
G2	135 (37,9 %)	132 (37,7 %)	133 (38,3 %)	132 (38,0 %)	159 (46,6 %)	160 (47,3 %)	160 (47,3 %)
G3	111 (31,2 %)	114 (32,6 %)	117 (33,7 %)	113 (33,0 %)	88 (25,8 %)	87 (25,7 %)	87 (25,7 %)
Total	356 (100 %)	350 (100 %)	347 (100 %)	342 (100 %)	341 (100 %)	338 (100 %)	338 (100 %)

3.4.1. Taux de rétention

Le projet Capacité d'apprentissage a un excellent taux de rétention, comptant seulement 18 retraits (5,1 %) depuis le début du projet en 2007. La raison principale pour le retrait d'un enfant du projet est le déménagement de la famille à l'extérieur de la communauté. Notons que certains enfants ont continué à être suivis dans les cas où le déménagement s'est fait vers une autre communauté participant au projet. Ces enfants étaient désormais considérés comme des participants de la communauté d'accueil. Le tableau 3.3 rapporte le décompte des raisons de retrait du projet.

Tableau 3.3 : Raisons justifiant le retrait de l'enfant du projet Capacité d'apprentissage

Raison pour le retrait	1 ^{ère} année	2 ^e année	Total
Déménagement à l'extérieur d'une communauté participante	6	2	8
L'enfant passe à une garderie anglophone/pas assez d'anglais à la garderie programme	3	0	3
Manque de disponibilité des familles	2	2	4
Dérangé par des questions du sondage de base	1	0	1
Perte de contact avec la famille	2	0	2
Total des retraits	14	4	18

3.5. MESURES

Cette section présente le plan des collectes de données effectuées auprès des parents et des enfants, ainsi que pour les Ateliers familles. Elle se poursuit avec une description des outils de mesures quantitatifs utilisés depuis le début du projet jusqu'à la fin de la deuxième année, soit de l'été 2007 à octobre 2009 pour la première cohorte et de octobre 2008 à octobre 2010 pour la deuxième cohorte.

3.5.1. Plan des collectes de données

Les collectes de données auprès des parents et des enfants se sont déroulées sur une base trimestrielle. La collecte des données de base pour la première cohorte a été effectuée de mai à octobre 2007. Les collectes suivantes ont eu lieu en février, juin et octobre 2008, ainsi qu'en

février, juin et octobre 2009. La collecte de données pour la deuxième cohorte a suivi un patron similaire, mais un an plus tard. Plus précisément, la collecte de base a été effectuée en octobre 2008, et les collectes suivantes ont eu lieu en février, juin et octobre 2009, ainsi qu'en février, juin et octobre 2010. La dernière collecte de données au sein des deux cohortes a été effectuée post-programme, c'est-à-dire au moment où tous les enfants étaient inscrits à temps plein dans une école.

Auprès des parents

Au total, on compte sept sondages effectués auprès des parents des deux cohortes depuis le début du projet Capacité d'apprentissage. Le sondage de base a été réalisé lors de l'inscription de l'enfant au projet, soit du 1^{er} mai au 31 octobre 2007 pour les participants de la première cohorte et au courant du mois d'octobre 2008 pour ceux de la deuxième cohorte. D'une durée d'environ une heure, le sondage s'est fait en face-à-face par la coordonnatrice communautaire avec le parent connaissant le mieux l'enfant (PCM).¹⁶ Les six sondages suivants étaient de courte durée, variant entre 10 et 30 minutes. Ils ont été administrés par téléphone par les coordonnatrices ou par un membre de la SRSA. Le taux de réponse moyen des sondages de suivi était de 94,7 % (voir le tableau 3.4).

Tableau 3.4 : Taux de réponses des sondages auprès des parents

	Sondage de base	+ 4 mois	+ 8 mois	+ 12 mois	+ 16 mois	+ 20 mois	+ 24 mois
Sondages administrés	356	341	343	335	338	335	331
Absents	0	9	5	7	3	3	7
Retraits	0	6	8	14	15	18	18
Taux de réponse	100 %	95,8 %	96,3 %	94,1 %	94,9 %	94,1 %	93,0 %

Pour les Ateliers familles

Puisque le nouveau programme testé comprenait un volet Ateliers familles ciblant les parents du groupe programme, nous avons prévu une évaluation de l'impact de ce volet sur les attitudes et comportements des parents. Deux sondages ont été développés afin de recueillir des informations avant et après la mise en œuvre des Ateliers familles. Le sondage pré-intervention a été administré par téléphone auprès de tous les parents du groupe programme en novembre 2007 pour la première cohorte et en novembre 2008 pour la deuxième cohorte. Ce sondage a permis de mesurer les attentes, les opinions et certains comportements des parents du groupe programme au début de l'intervention des Ateliers familles. Un sondage post-atelier fut administré deux à trois semaines après le dernier atelier afin de mesurer l'évolution de ces résultantes. Son administration fut également faite par téléphone. Le lecteur est prié de noter que certaines mesures pré-intervention ont été collectées de façon rétrospective durant le sondage post-atelier.

¹⁶ Un deuxième sondage de base fut administré seulement aux parents de la première cohorte s'ayant inscrit au projet avant le 1^{er} septembre 2007. Ce sondage servait à actualiser l'information relative au type et à la fréquence des activités de littératie et des activités hors garderie faites avec l'enfant au moment où l'enfant faisait ses débuts en garderie. Pour les analyses d'impact, les données du deuxième sondage de base ont remplacé les données du premier sondage de base lorsqu'elles étaient présentes. Ces deux sondages sont donc considérés comme étant un seul sondage dans le cadre des analyses.

Un troisième sondage anonyme a été administré lors du dernier atelier pour solliciter l'opinion des parents sur les aspects logistiques des ateliers, notamment au niveau de l'environnement physique, de la qualité des échanges, et de la qualité de l'animation des intervenantes. L'ensemble de ces données a servi à évaluer la qualité de la mise en œuvre du volet Ateliers familles dont les résultats se retrouvent au chapitre 6 du présent document. Le taux de réponse était excellent pour les sondages pré- et post-atelier, mais relativement bas pour le sondage de logistique en raison de la faible participation à cet atelier (voir le tableau 3.5).

Tableau 3.5 : Taux de réponses des sondages liés aux Ateliers familles

	<i>Pré-intervention</i>	<i>Sondage de logistique</i>	<i>Post-intervention</i>
Sondages administrés	135	92	134
Absents	8	50	8
Retraits	6	7	7
Taux de réponse	90,6 %	61,7 %	89,9 %

Note : Les taux incluent les communautés d'Edmonton et Saint-Jean.

Auprès des enfants

Chaque période de sondage auprès des parents était jumelée avec une évaluation auprès des enfants. En tout, sept évaluations ont été administrées tous les quatre mois auprès des enfants pendant les deux premières années du projet Capacité d'apprentissage avec un taux de réponse moyen de 95,5 % (voir le tableau 3.6).

Tableau 3.6 : Taux de réponses des évaluations auprès des enfants

	Période d'évaluations						
	Évaluation de base	+ 4 mois	+ 8 mois	+ 12 mois	+ 16 mois	+ 20 mois	+ 24 mois
Enfants évalués	346	347	342	337	339	334	334
Absents	4	3	6	5	2	4	4
Retraits	6	6	8	14	15	18	18
Taux de réponse	97,2 %	97,5 %	96,1 %	94,7 %	95,2 %	93,8 %	93,8 %

3.5.2. Mesures auprès des enfants

Au cours des différentes vagues de collectes de données, divers outils ont été utilisés pour suivre le développement des enfants sur plusieurs dimensions importantes liées à la préparation à la scolarisation. Le calendrier et les outils de chaque période d'évaluation se retrouvent à l'annexe A.

Évaluation de la petite enfance — Appréciation directe (Willms, 2007)

L'ÉPE-AD est un outil qui mesure quatre domaines de la préparation à la scolarisation et un cinquième domaine conçu spécifiquement aux fins du projet Capacité d'apprentissage :

- (Domaine A) Conscience de soi et de l'environnement;
- (Domaine B) Habilités cognitives;
- (Domaine C) Langage et communication;
- (Domaine D) Développement physique et moteur; et
- (Domaine E) Conscience et engagement envers la culture francophone.

Chaque domaine comporte un certain nombre de questions présentées en ordre croissant de difficulté. Deux versions de L'ÉPE-AD ont été utilisées dans le cadre de ce projet. L'ÉPE-AD initiale a été l'outil utilisé jusqu'à la 5^e évaluation (+16 mois) tandis que l'« ÉPE-AD modifiée », la version de l'outil retravaillée par la SRSA, fut utilisée lors des 6^e et 7^e évaluations (+20 et +24 mois). Une comparaison du contenu des trois versions de l'ÉPE-AD se retrouve à l'annexe B.

Le domaine A, Conscience de soi et de l'environnement, est composé de 16 questions permettant de savoir jusqu'à quel point l'enfant peut reconnaître et identifier des éléments présents dans son environnement. Par exemple, on demande à l'enfant de nommer des couleurs, des parties du corps, certains professionnels et leur rôle, en plus de sa date de naissance.

Le domaine B, Habiletés cognitives, est composé de 17 questions mesurant différents aspects logico-mathématiques. Par exemple, on demande à l'enfant de compter, de former des groupes à partir de différents objets, de comparer différentes formes et de distinguer la forme la plus grande de la plus petite. De plus, on demande à l'enfant de nommer quelques lettres de l'alphabet, d'identifier leur son, et de lire 8 mots.

Le domaine C, Langage et communication, comprend 14 items mesurant les capacités de communication et de compréhension de l'enfant. L'évaluatrice demande à l'enfant de lui montrer du doigt des images représentant un mot qu'elle lui dit, de répéter des phrases de sept syllabes, de répondre à des questions avec des phrases complètes, d'utiliser des images pour raconter une histoire et de démontrer le sens de quatre mots d'action. C'est le seul domaine administré à tous les enfants en français puisque ce domaine fait partie de l'arbre décisionnel utilisé pour décider de la langue d'administration de l'évaluation.¹⁷

Le domaine D, Développement physique et moteur, est composé de 16 items mesurant la capacité de l'enfant à accomplir certaines activités de motricité fine et de motricité globale. La motricité fine englobe plusieurs éléments de la pré-écriture comme, par exemple, pouvoir tracer des lettres, dessiner une ligne droite ou colorier entre les lignes en tenant son crayon correctement. La motricité globale concerne les habiletés de l'enfant à bouger son corps, par exemple, en sautant au-dessus d'un objet, en se tenant en équilibre sur un pied ou en sautillant sur un pied trois fois de suite. Il est à noter qu'une étude canadienne basée sur les données de l'ELNEJ indique qu'après l'âge de trois ans, on observe peu de différence entre les enfants ayant un développement normal à ce domaine. Ces résultats ont été confirmés lors de la 5^e évaluation (+16 mois). Pour cette raison, ce domaine n'a pas été mesuré lors des 6^e et 7^e évaluations effectuées à +20 mois et +24 mois.

¹⁷ La procédure d'administration de l'ÉPE-AD ainsi que l'arbre décisionnel pour déterminer la langue du test sont détaillés à l'annexe D.

Les items sont notés sur une échelle de 5 points pour les domaines A, B et C et de 4 points pour le domaine D. Les trois premiers domaines requièrent l'utilisation de planches d'évaluation ou d'objets que l'enfant doit montrer du doigt ou manipuler. Durant l'administration du test, il est possible de voir un enfant atteindre un « plateau » dans un domaine donné lorsque les questions deviennent trop difficiles. En effet, l'évaluation d'un domaine est arrêtée lorsque l'enfant obtient un score de « 0 » ou « 1 » à trois items consécutifs. L'évaluatrice passe alors au domaine suivant. L'administration du test dure environ de 45 minutes à une heure.

Jusqu'à l'évaluation +16 mois, l'enfant pouvait être évalué en anglais ou en français selon le score obtenu au domaine E et celui obtenu au domaine C. Le domaine E, conçu spécifiquement pour le projet Capacité d'apprentissage, mesurant la conscience et l'engagement envers la culture francophone de l'enfant, servait davantage à établir un rapport amical avec l'enfant et déterminait la langue du test.¹⁸ Son utilisation fut discontinuée à la 5^e évaluation (+16 mois) puisque les enfants étaient trop jeunes pour répondre clairement aux questions, et par conséquent, ce domaine s'est avéré non fiable.¹⁹

L'ÉPE-AD présente de bonnes propriétés psychométriques. Le résultat des analyses factorielles de l'ÉPE-AD a confirmé l'unidimensionnalité de chacun des domaines de la version française du test, tel que postulée théoriquement.²⁰ La cohérence interne de chacune des dimensions s'est avérée acceptable à très bonne avec des alphas de Cronbach variant de 0,61 à 0,94 pour la version française du test et de 0,79 à 0,94 pour la version anglaise (voir le tableau 3.7).

Tableau 3.7 : Alpha de Cronbach des domaines de l'ÉPE-AD par vague d'évaluation

	Alpha de Cronbach [α (n)]						
	Base eval.	+ 4 mois	+ 8 mois	+ 12 mois	+ 16 mois	+ 20 mois	+ 24 mois
Domaines administrés en français							
A	0,92 (301)	0,91 (300)	0,91 (301)	0,90 (311)	0,90 (329)	0,89 (334)	0,89 (334)
B	0,86 (301)	0,87 (300)	0,86 (301)	0,84 (311)	0,85 (329)	0,82 (334)	0,85 (334)
C	0,92 (346)	0,93 (347)	0,94 (342)	0,93 (337)	0,92 (339)	0,63 (334)	0,61 (334)
D	0,89 (301)	0,92 (300)	0,93 (301)	0,93 (311)	0,93 (329)	s.o.	s.o.
Domaines administrés en anglais							
A	0,85 (45)	0,89 (47)	0,88 (41)	0,91 (26)	0,86 (10)	s.o.	s.o.

¹⁸ Le domaine E contient des questions sur les préférences des enfants en matière de livres, d'émissions de télévision et de chansons, sur la langue de ces ressources (anglais ou français), ainsi que sur les langues parlées avec leurs parents et ami(e)s.

¹⁹ Le lecteur est référé au rapport *Capacité d'apprentissage dans les communautés francophones en situation minoritaire : Rapport de référence* (Legault et al., 2014) pour connaître les raisons sous-jacentes à cette décision.

²⁰ Il a été impossible d'effectuer une analyse factorielle sur la version anglaise du test en raison du trop petit nombre d'individus ayant répondu dans cette langue. Un minimum de 160–170 cas aurait été souhaitable par domaine afin d'assurer la validité statistique des calculs (Tabachnick et Fidell, 2006).

	Alpha de Cronbach [α (n)]						
	Base eval.	+ 4 mois	+ 8 mois	+ 12 mois	+ 16 mois	+ 20 mois	+ 24 mois
B	0,79 (45)	0,89 (47)	0,89 (41)	0,89 (26)	0,86 (10)	s.o.	s.o.
C	0,87 (45)	0,85 (47)	0,89 (41)	0,90 (26)	0,85 (10)	s.o.	s.o.
D	0,83 (45)	0,91 (47)	0,88 (41)	0,94 (26)	0,94 (10)	s.o.	s.o.

Épreuve de dénomination de Gardner (Ska, 1995)

L'Épreuve de dénomination de Gardner (1979) est un outil qui mesure le vocabulaire expressif de l'enfant. L'outil est la traduction validée et normalisée en français pour les enfants de la maternelle à la deuxième année de l'outil *Expressive One-Word Picture Vocabulary Test* (EOWPVT). Ainsi, les consignes et l'ordre de difficulté des items ont été ajustés pour une population francophone. L'exercice consiste à présenter une série d'images à l'enfant et de lui demander de nommer le mot associé à l'image. Après six erreurs consécutives, le test est arrêté, résultant en un temps d'administration variant entre 10 et 15 minutes. L'avantage de ce test est qu'il est très sensible aux différences dans le niveau de maîtrise de la langue française. Le score total représente la somme des images correctement identifiées en français par l'enfant. On tient également compte de certaines réponses incorrectes, soit des mots anglais ou des expressions régionales, offertes par l'enfant. Ces réponses incorrectes sont récoltées à titre d'informations seulement, et ne servent pas au calcul d'un score. Cet outil fut utilisé lors de la 6^e évaluation (+20 mois).

Échelle de vocabulaire en images Peabody — Révisé (ÉVIP-R; Dunn, Thériault-Whalen, et Dunn, 1993)

L'ÉVIP-R est un outil qui mesure le vocabulaire réceptif de l'enfant. Son administration ressemble à un jeu : l'enfant entend un mot prononcé à voix haute et doit désigner, parmi quatre images, celle qui représente le mot entendu. Le test s'arrête après six erreurs parmi huit essais. Selon les règles du test, le point de départ où l'enfant commence le test est déterminé selon son âge et sa performance. Pour les besoins du projet Capacité d'apprentissage, un point de départ commun pour tous les enfants (soit la 15^e question qui correspond au point de départ des enfants âgés de 3 ans et demi) a été établi peu importe leur âge ou leur niveau de performance. Le but du présent projet n'étant pas de comparer la performance des enfants à une « norme », ²¹ l'utilisation d'un point de départ commun suffisamment facile, nous assure de capter le niveau de vocabulaire réceptif auprès des enfants issus de famille exogame.

Le score de l'ÉVIP-R représente la somme des bonnes réponses données par l'enfant à compter de la 15^e question. Cette façon de calculer le score de l'ÉVIP-R diffère légèrement de celle préconisée dans le manuel du test (Dunn, et al., 1993, p. 13–15). Nous observons une corrélation supérieure à .99 entre la méthode adoptée par la SRSA et celle mise de l'avant dans le manuel. C'est donc dire qu'il n'y a aucune perte d'information. La méthode utilisée par la SRSA

²¹ Précisons que les normes actuelles pour la version française du test ont été établies en 1990 et n'ont pas fait l'objet d'une mise à jour. En conséquence, leur validité est inconnue.

pour calculer le score de l'enfant à l'ÉVIP-R comporte plusieurs avantages notamment, elle: a) maximise la variance; et b) permet d'inclure les enfants qui n'ont pas pu faire le test parce qu'ils ont échoué aux essais d'entraînement (ils ont obtenu un score de zéro). Cet outil fut utilisé lors de la 7^e évaluation (+24 mois).

Sous-échelles de vocabulaire (ÉPE-AD)

Au cours de l'hiver 2009, la SRSA a remanié la composition des échelles de l'ÉPE-AD pour créer deux nouvelles sous-échelles de vocabulaire et ce, dans le but d'améliorer la capacité de mesurer de fines différences dans les compétences langagières des enfants. La première sous-échelle mesure le vocabulaire expressif, soit la capacité de l'enfant de dire le mot associé à l'image qui lui est présenté. Cette échelle contient six items du domaine A et deux items du domaine C. La cohérence interne est très bonne pour les items du vocabulaire expressif avec des alphas de Cronbach variant entre 0,77 et 0,88 pour les versions anglaise et française.

La deuxième sous-échelle mesure le vocabulaire réceptif et se compose de deux items du domaine A et trois items du domaine C. On présente à l'enfant une série d'images et celui-ci doit identifier l'image associée au mot prononcé par une évaluatrice. La cohérence interne des items mesurant le vocabulaire réceptif dans la version française est acceptable avec des alphas de Cronbach variant de 0,61 à 0,74. Par contre, la cohérence interne de la version anglaise de la sous-échelle de vocabulaire réceptif est moins acceptable puisqu'elle comporte plusieurs alphas de moins que 0,50 (voir le tableau 3.8).

Tableau 3.8 : Alpha de Cronbach des sous-échelles de vocabulaire de l'ÉPE-AD par vague d'évaluation

	Alpha de Cronbach [α (n)]						
	Base eval.	+ 4 mois	+ 8 mois	+ 12 mois	+ 16 mois	+ 20 mois	+ 24 mois
Domaines administrés en français							
Vocabulaire expressif	0,86 (301)	0,87 (300)	0,88 (301)	0,87 (311)	0,86 (329)	0,86 (334)	0,84 (334)
Vocabulaire réceptif	0,74 (301)	0,68 (300)	0,68 (301)	0,61 (311)	0,64 (329)	s.o.	s.o.
Domaines administrés en anglais							
Vocabulaire expressif	0,81 (45)	0,85 (47)	0,86 (41)	0,84 (26)	0,77 (10)	s.o.	s.o.
Vocabulaire réceptif	0,41 (45)	0,73 (47)	0,38 (41)	0,54 (26)	0,46 (10)	s.o.	s.o.

Évaluation de la petite enfance — Appréciation directe, version modifiée (SRSA, 2009).

Au cours de l'hiver 2009, la SRSA a constaté un potentiel élevé d'effet plafond suite aux premières analyses des résultats de l'ÉPE-AD de la première cohorte et à des projections de scores pour la 5^e évaluation. Un effet plafond des scores entraverait la capacité de continuer à suivre l'évolution du développement des enfants et, par extension, il deviendrait impossible de tester pour des différences entre les enfants des trois groupes expérimentaux. RHDCC a donc

retenu les services du concepteur de l'outil pour créer de nouvelles questions plus difficiles permettant ainsi de poursuivre la trajectoire développementale des enfants aux domaines A, B, et C jusqu'à la fin de la deuxième année du projet (la 7^e évaluation). Le concepteur a également éliminé les questions les plus faciles du test afin de ne pas allonger le temps d'administration. La « version allongée de l'ÉPE-AD » (Willms, 2009) fut prête en mai 2009.

En parallèle, la SRSA a procédé à une restructuration des questions composant la version allongée de l'ÉPE-AD. Cette restructuration avait comme objectif de mesurer plus finement des habiletés langagières des enfants tout en conservant la capacité de suivre la trajectoire développementale des enfants selon trois des quatre domaines de l'ÉPE-AD, soit la conscience de soi et de l'environnement (domaine A), les habiletés cognitives (domaine B), et langage et communication (domaine C).²² Le nouvel outil, nommé « l'ÉPE-AD modifiée », a aussi donné lieu à des sous-échelles mesurant le vocabulaire expressif et réceptif présentées plus tôt. L'ÉPE-AD modifiée comprend:

- 15 questions au domaine A concernant surtout le vocabulaire expressif et les connaissances générales;
- 12 questions au domaine B mesurant la conscience phonologique (soit l'habileté des enfants à jouer avec les sons qui forment les mots) et les concepts de la numératie (soit l'habileté des enfants à jouer avec les chiffres). Deux autres items du domaine B ont été administrés de façon légèrement différente de l'ÉPE-AD allongée. Ainsi, la connaissance du son de toutes les lettres de l'alphabet a été mesurée au moyen d'un livret où les lettres majuscules sont présentées en ordre croissant de difficulté pour les petits francophones. La connaissance du nom de toutes les lettres de l'alphabet a été mesurée au moyen d'un livret où les lettres minuscules sont présentées en ordre croissant de difficulté pour les petits francophones.
- Enfin, deux questions du domaine C ont été conservées pour poursuivre l'échelle de vocabulaire expressif et deux nouvelles questions plus difficiles provenant de l'ÉPE-AD allongée ont été ajoutées pour mesurer la conscience phonologique des enfants.

L'ÉPE-AD initiale a été l'outil utilisé jusqu'à la 5^e évaluation (+16 mois) tandis que l'« ÉPE-AD modifiée », la version de l'outil retravaillée par la SRSA, fut utilisée lors des 6^e et 7^e évaluations (+20 et +24 mois). Une comparaison du contenu des trois versions de l'ÉPE-AD se retrouve à l'annexe B.

Dans le cadre du projet Capacité d'apprentissage, il est important de suivre la trajectoire développementale des enfants lors des deux années du programme. L'établissement d'une continuité dans les dimensions mesurées par deux outils passe par l'examen du patron de corrélations entre les sous-échelles mesurant la même dimension dans chacun des outils. Une forte corrélation (soit de 0,75 et plus) observée entre les sous-échelles signifie qu'ils mesurent respectivement le même concept. L'examen du patron de corrélations au tableau 3.9 confirme la présence de fortes corrélations entre les domaines A et B de l'ÉPE-AD initiale (5^e évaluation) et ceux de l'ÉPE-AD modifiée (6^e évaluation), soit 0,83 et 0,75 respectivement. L'exception est le domaine de la communication (domaine C) où l'on observe une corrélation de 0,63. Ce

²² Le lecteur intéressé à connaître davantage sur le remaniement de l'ÉPE-AD et des connaissances théoriques l'informant est prié de consulter le *Rapport présentant une revue des outils de mesures directes des dimensions développementales des enfants* remis à RHDCC en juillet 2009.

résultat plus faible qu'espéré est toutefois pallié par de bonnes corrélations retrouvées entre le domaine C de l'ÉPE-AD initiale et la sous-échelle du vocabulaire expressif créé à partir de l'ÉPE-AD (0,82), ainsi qu'entre le domaine C de l'ÉPE-AD initiale et l'échelle « de vocabulaire expressif » de Gardner (0,77). On note également une forte corrélation entre le domaine C de l'ÉPE-AD initiale et l'échelle du domaine A de l'ÉPE-AD modifiée. L'ensemble de ces résultats permet de conclure que l'utilisation de la version modifiée entraîne une perte d'information négligeable. Par conséquent, nous pouvons suivre la trajectoire développementale des enfants lors des deux années du programme.

Tableau 3.9 : Coefficients de corrélation Pearson des échelles de l'ÉPE-AD et des échelles de vocabulaire expressif et réceptif

		<i>ÉPE-AD initiale : 5^e évaluation (+16 mois)</i>				
		A	B	C	Sous-échelle vocabulaire expressif (ÉPE-AD)	Sous-échelle vocabulaire réceptif (ÉPE-AD)
Échelles						
ÉPE-AD modifiée : 6^e évaluation (+20 mois)	A	0.83 (309)	0.63 (322)	0.83 (330)	0.80 (311)	0.59 (312)
	B	0.51 (309)	0.75 (322)	0.47 (330)	0.45 (311)	0.39 (312)
	C	0.61 (309)	0.61 (322)	0.63 (330)	0.60 (311)	0.47 (312)
	Sous-échelle vocabulaire expressif (ÉPE-AD)	0.79 (309)	0.49 (322)	0.82 (330)	0.81 (311)	0.54 (312)
	Vocabulaire expressif (Gardner)	0.75 (306)	0.52 (319)	0.77 (327)	0.76 (308)	0.49 (309)
	Vocabulaire réceptif (ÉVIP-R)	0.68 (310)	0.53 (323)	0.71 (330)	0.68 (312)	0.43 (313)

Note : L'ÉPE-AD modifiée ne comprend pas une sous-échelle de vocabulaire réceptif. Ces dernières ont été éliminées par le concepteur de l'outil dans sa version allongée de l'ÉPE-AD puisque faciles. L'analyse des données du projet Capacité d'apprentissage confirme que ces données étaient trop faciles et par conséquent, ne permettaient plus de discriminer entre les enfants.

3.5.3. Mesures auprès des parents

Les sondages de suivi auprès des parents permettent de déterminer à quel point l'environnement de l'enfant s'est modifié au cours des quatre mois précédant le sondage. Dans un premier temps, ils permettent de mesurer la fréquence des activités de littératie, la langue de ces activités et d'autres facteurs qui affecteront la préparation à la scolarisation des enfants et le développement de leur identité francophone. Dans un deuxième temps, ils permettent de vérifier de façon systématique les changements quant au mode de garde de l'enfant et la langue utilisée dans ce nouveau mode de garde pour que les analyses puissent en tenir compte. Dans un troisième temps, ils visent à prendre en note les changements dans la composition de la famille afin de déterminer si l'environnement langagier habituel de l'enfant à la maison en est affecté.

Les échelles utilisées dans les sondages des parents participant au projet Capacité d'apprentissage proviennent d'études effectuées auprès de populations francophones au Canada, telles que l'ELNEJ, l'Étude longitudinale du développement des enfants du Québec (ELDEQ; Institut de la statistique du Québec, 2003) et l'Enquête sur la vitalité des minorités de langue officielle (EVMLO, Statistique Canada, 2006). Les questions qui ont été sélectionnées pour les sondages du projet Capacité d'apprentissage sont celles pertinentes au groupe d'âge des enfants.

Le lecteur est prié de noter que seules les échelles incluses dans les analyses d'impact sont détaillées dans cette section. Il n'est donc pas étonnant que cette liste d'échelles diffère quelque peu de celles rapportées dans le rapport *Capacité d'apprentissage dans les communautés francophones en situation minoritaire : Rapport de référence* (Legault et al., 2014). Plusieurs autres concepts ont été mesurés sans toutefois être retenus dans les analyses d'impact. La décision de les inclure ou non est fondée sur les analyses préliminaires dont les résultats sont rapportés au chapitre 5.

Caractéristiques sociodémographiques

Les caractéristiques sociodémographiques des enfants et de leurs parents proviennent de questions posées dans le formulaire de consentement des parents (par exemple, le sexe de l'enfant et sa date de naissance) et du sondage de base. Les sondages de suivi ont permis d'apporter des modifications, au besoin, à ce portrait initial. Cette section identifie les variables sociodémographiques utilisées pour les analyses d'impact.

Des informations telles que le niveau d'éducation de la mère et celui du père, l'âge de la mère à la naissance de l'enfant et le revenu familial ont été recueillies auprès des parents lors du sondage de base. Le revenu de la famille a été demandé, par la suite, à deux reprises. Une question ouverte portant sur le revenu fut d'abord posée à la 5^e évaluation (+16 mois). Devant le refus de plusieurs parents de partager cette information, la même question a été à nouveau posée à la 7^e évaluation (+24 mois), et ce, spécifiquement aux parents ayant refusé de répondre, mais cette fois avec des catégories de revenu incluses dans les choix de réponses.

Plusieurs questions visaient à établir le profil linguistique des familles participantes. Entre autres, le répondant devait identifier ses langues maternelles, sa connaissance des langues officielles ainsi que les langues parlées avec l'enfant et à la maison. Les mêmes questions étaient posées au conjoint ou à la conjointe. Il existe deux définitions courantes de la population francophone canadienne. Selon Statistique Canada,²³ on peut tout d'abord calculer la population qui déclare avoir le français comme langue maternelle, c'est-à-dire que le français est la première langue apprise et encore comprise. La deuxième façon est de calculer la « première langue officielle parlée » (PLOP), un score prenant en compte la connaissance des deux langues officielles, la langue maternelle et la langue la plus souvent parlée à la maison (Forgues et Landry, 2006).

Chaque sondage de suivi commence par une série de cinq questions adressées seulement aux parents ayant subi un changement dans la composition de leur famille depuis le dernier sondage complété. Les changements dans la composition de la famille par rapport aux réponses du sondage de base (p.ex. la taille du ménage, les familles monoparentales, les enfants plus vieux,

²³ Définition de la population francophone disponible sur le site Web de Statistique Canada à http://www.statcan.ca/francais/concepts/definitions/language_f.htm.

les enfants jumeaux, les enfants plus jeunes) sont pris en considération dans les analyses d'impact.

Le statut d'immigrant des parents a été établi au moyen d'une question dans la 7^e évaluation (+24 mois). Ce sondage a également permis de faire une mise à jour des données sociodémographiques du conjoint ou de la conjointe du répondant. Notons qu'au sondage de base, l'information obtenue définissait les caractéristiques du conjoint ou de la conjointe qui était le parent biologique de l'enfant, même si ce parent n'avait pas de contact avec l'enfant. Afin de clarifier les résultats, ces questions ont été reposerées à la 7^e évaluation (+24 mois) pour obtenir de l'information sur le deuxième adulte qui habite dans la maisonnée avec l'enfant et qui est impliqué de façon active dans les soins de l'enfant.

Habitudes langagières

Plusieurs questions au sujet des habitudes langagières sont incluses dans le sondage de base et les sondages de suivi. Les répondants choisissent parmi plusieurs options la catégorie la plus représentative de leur vécu (par ex. s'ils parlent anglais seulement, anglais et français, français plus que l'anglais, etc.). Ces catégories ont été combinées pour créer une échelle en cinq points représentant un continuum linguistique où un score de 1 signifie que le répondant parle « anglais seulement », un score de 3 signifie qu'il/elle parle le « français et l'anglais et/ou une autre langue », et un score de 5 signifie qu'il/elle parle le « français seulement ». Les analyses d'impact ont fait usage de ce continuum linguistique basé sur une échelle de cinq points. À des fins d'analyses comparatives (projet Capacité d'apprentissage comparé à l'EVMLO), les habitudes langagières des parents ont été regroupées sur un continuum de trois points en raison du petit nombre de cas dans certaines cellules. Dans ces analyses, un score de 1 correspond au « français seulement », un score de 2 signifie le « français et l'anglais également OU le français et une autre langue », et un score de 3 signifie « l'anglais seulement OU l'anglais et une autre langue OU une autre langue ».

Un premier continuum linguistique a été généré représentant la langue habituellement parlée par la mère à l'enfant. Un deuxième continuum linguistique a été calculé en prenant la moyenne des scores de quatre questions portant sur les langues parlées par l'enfant avec sa mère, avec son père, avec ses amis et sa fratrie à la maison, ainsi qu'à l'extérieur de la maison (alpha de Cronbach de 0,95).

Deux indices représentant la langue habituellement parlée au foyer ont été dérivés. Un premier indice fut dérivé en mariant la PLOP de la mère avec celle du père. Un deuxième indice fut créé en combinant la langue habituellement parlée par la mère à l'enfant avec la langue habituellement parlée par le père à son enfant. Ces deux indices ont servi à déterminer l'environnement langagier dans lequel baigne l'enfant à la maison. En outre, la PLOP a servi à déterminer si l'enfant habite un foyer endogame francophone, endogame anglophone ou exogame.

Processus familiaux

Les informations recueillies sur les processus familiaux proviennent du sondage de base. Certains de ces processus sont mesurés à nouveau dans les sondages de suivi. En tout, cinq processus familiaux ont été mesurés au moyen de quatre échelles, notamment les pratiques parentales, le fonctionnement familial, la dépression et les activités d'alphabétisation.

L'échelle des *pratiques parentales positives* comprend cinq items mesurant la fréquence de contacts positifs entre le parent et l'enfant (par ex., combien de fois le parent félicite l'enfant, rit avec lui ou lui donne de l'affection). Les items sont notés sur une échelle en cinq points dont 1 indique « jamais » et 5 indique « plusieurs fois durant la journée ». La cohérence interne de l'échelle est acceptable à 0,62.

L'échelle des *pratiques parentales autoritaires* comprend quatre items mesurant le degré de supervision et de discipline. Elle nous permet de découvrir si, entre autres, l'enfant respecte ou non les punitions qu'on lui impose ou si ces punitions varient selon l'humeur du parent. Les items sont notés sur une échelle en cinq points où 1 indique « jamais » et 5 indique « presque toujours ». La cohérence interne s'avère acceptable avec un alpha de Cronbach de 0,61.

L'échelle de *fonctionnement familial*, composée de huit items, évalue la qualité des liens dans la famille. Le contenu des items porte sur l'ouverture émotionnelle, l'expression des sentiments et la convivialité à l'intérieur de la famille. Les choix de réponses pour les items de l'échelle passent de 1, soit « absolument en désaccord » à 4, soit « entièrement en accord ». L'échelle du fonctionnement familial démontre une excellente cohérence interne avec un alpha de Cronbach de 0,81.

L'échelle de *dépression* mesure, par exemple, le nombre de fois que le répondant s'est senti déprimé ou malheureux, qu'il a pleuré ou s'est senti seul, ou qu'il ne jouissait pas de la vie. Les huit items de l'échelle s'étalent sur un score de quatre points, où 1 indique que le répondant s'est senti ainsi « rarement ou jamais » et 4 signifiant qu'il s'est senti ainsi « la plupart du temps ou tout le temps ». L'échelle de la dépression a une cohérence interne très élevée avec un alpha de Cronbach de 0,80.

L'échelle des *activités d'alphabétisation* comprend cinq items mesurant la fréquence des activités de lecture, d'écriture et de numératie. Plus spécifiquement, les items mesurent la fréquence à laquelle les parents racontent des histoires sans livre à leur enfant, chantent des chansons avec lui, lui enseignent à écrire ou à tracer des lettres ou des chiffres, lui enseignent à lire des mots, et l'encouragent à utiliser des nombres dans ses activités quotidiennes. Ces items sont notés sur une échelle de 5 points, où 1 correspond à une activité qui n'a jamais été faite et 5 correspond à une activité faite plusieurs fois par jour.

Chacune des questions portant sur les activités d'alphabétisation est suivie d'une question vérifiant la langue utilisée lors de cette activité. Les répondants choisissent parmi plusieurs options la catégorie la plus représentative de leur vécu (par ex. s'ils parlent anglais seulement, anglais et français, français plus que l'anglais, etc.). Ces catégories ont été combinées pour créer un score en cinq points représentant un continuum linguistique *pour cette activité* où un score de 1 signifie que l'activité se fait en « anglais seulement », un score de 3 signifie que l'activité se fait en « français et anglais et/ou une autre langue », et un score de 5 signifie que l'activité se fait en « français seulement ». Une échelle des *langues des activités d'alphabétisation* est obtenue en faisant la moyenne des scores obtenus à chaque question. La cohérence interne de cette échelle est très élevée avec un alpha de Cronbach variant entre 0,88 et 0,94.

Les questions permettant de construire l'échelle des activités de littératie ainsi que l'échelle des langues associées à ces activités ont été posées à chaque sondage, incluant le sondage de base. La cohérence interne pour ces échelles s'avère assez acceptable à chacune des périodes de

sondage pour permettre d'avoir confiance dans l'homogénéité de ces échelles (voir le tableau 3.10).

Tableau 3.10 : Alpha de Cronbach de l'échelle des activités de littératie et de l'échelle des langues des activités d'alphabétisation à chaque période de sondage

	<i>Alpha de Cronbach [α (n)]</i>						
	Base	+4 mois	+ 8 mois	+12 mois	+ 16 mois	+20 mois	+24 mois
Activités de littératie	0,52	0,51	0,61	0,72	0,73	0,72	0,69
Langues utilisées lors des activités	0,94	0,93	0,93	0,92	0,89	0,91	0,88

Dimension communautaire

L'échelle du *capital social* provient de questions posées lors du sondage de base. Composée de cinq items, cette échelle permet de mesurer le niveau d'entraide entre voisins dans la communauté du participant. Ces items cherchent à comprendre jusqu'à quel point les membres du voisinage sont prêts à aider leurs voisins, à se rassembler pour régler un problème commun et à assurer la sécurité des enfants. Les scores des items s'étendent de 1, « Absolument en désaccord », à 4, « Absolument d'accord ». La cohérence interne de l'échelle est excellente avec un alpha de Cronbach de 0,80.

Dimension identitaire

Le sondage de base comprenait plusieurs questions portant sur la perception de la vitalité de la communauté francophone, l'engagement envers la culture francophone et le *sentiment d'appartenance à la culture francophone*. Parmi toutes ces dimensions identitaires, seul le sentiment d'appartenance à la culture francophone s'est avéré intéressant dans les analyses d'impact. Il s'agit d'une question visant à déterminer le groupe culturel auquel les parents s'identifient, soit au groupe francophone, anglophone, aux deux ou à un autre groupe.

3.5.4. Connaissances, attitudes et croyances des parents

Les sondages pré- et post-Ateliers familles contiennent une série de questions visant à faire ressortir la nature et la portée des changements au niveau des perceptions et des comportements des parents du groupe programme. Pour les besoins de l'évaluation du programme Ateliers familles, plusieurs mesures ont été créées à partir des questions utilisées dans des études portant sur l'alphabétisation familiale. Certaines de ces études ont été menées par le Centre interdisciplinaire de recherche sur la citoyenneté et les minorités (CIRCEM; LeTouzé, 2006). Une autre source d'inspiration provient de l'étude sur l'impact de l'alphabétisation familiale sur les familles vivant en milieu minoritaire menée par la Coalition francophone pour l'alphabétisation et la formation de base en Ontario (Benoît, s.d.). Des mesures ont aussi été créées à partir du questionnaire post-intervention du programme de transition à l'école provenant du Conseil du district scolaire 3 à Edmundston (Gauthier St-Onge, s.d.) et à partir de l'ELNEJ (Statistique Canada et RHDSC, 2006). Enfin, l'approche rétrospective utilisée pour créer les questions a été inspirée d'un sondage rétrospectif créé par Lamb et Tschillard (2005).

Le choix des construits pour étudier les impacts du programme des Ateliers familles s'inspire en partie du modèle de l'action raisonnée de Fishbein et Ajzen (pour une introduction au modèle voir Brigham, 1991). Ce modèle met en relation les croyances et les attitudes avec la prédiction des comportements. Selon le modèle, les croyances des gens relatives à un comportement influencent à la fois leurs attitudes envers ce comportement et leurs perceptions de contrôle (il s'agit d'un concept similaire à celui des sentiments d'efficacité). Les attitudes et les perceptions de contrôle influencent, à leur tour, l'intention d'effectuer le comportement. Enfin, l'intention détermine la probabilité que le comportement sera fait. Dans le projet Capacité d'apprentissage, nous avons mesuré les croyances et les attitudes des participants face aux dimensions ciblées par les Ateliers familles ainsi que leurs sentiments d'efficacité. Nous avons également mesuré les connaissances des participants envers les principaux thèmes abordés dans les Ateliers familles. Bien qu'il ne soit pas inclus dans le modèle de l'action raisonnée, le niveau de connaissance affecte également les attitudes des gens face à une action. Enfin, une échelle de modelage a été ajoutée pour capter l'importance des activités d'alphabétisation effectuées par les parents sur les attitudes et comportements des enfants. En effet, plusieurs études ont démontré l'effet positif d'un environnement riche en littérature familiale, tel que la présence de matériel écrit ou pour écrire, ainsi que des comportements parentaux valorisant la lecture ou l'écriture, comme de lire ou d'écrire devant les enfants (Dionne, 2007; Sénéchal et LeFevre, 2002; Jordan, Snow, et Porche, 2000). Les quatre concepts propres aux sondages pré- et post-ateliers sont explicités dans les paragraphes suivants.

Quatre questions mesurent certaines *croyances et attitudes* des parents qui sont abordées au cours des Ateliers familles. Ces croyances correspondent à des thèmes majeurs du programme, notamment: le parent comme premier éducateur de l'enfant, la coopération entre l'éducatrice et les parents, l'environnement culturel francophone à la maison, et l'âge propice pour lire à l'enfant. Le répondant indique jusqu'à quel point il est en accord avec chacun des énoncés en utilisant une échelle en 4 points où 1 signifie « absolument en désaccord » et 4 signifie « absolument en accord ». Ces questions ont été posées dans les sondages pré- et post-ateliers. Notons toutefois que ces questions ne forment pas, à proprement parler, une échelle et que, par conséquent, les informations recueillies sont de nature descriptive.

L'échelle de *perception d'auto-efficacité* évalue la perception qu'ont les parents de leur capacité à effectuer certaines actions discutées lors des Ateliers familles. Ainsi, on demande aux parents à quel point ils sont confiants quant à leur capacité de : gérer les émotions et les comportements de leur enfant, parler en français avec lui, l'aider dans ses apprentissages, et l'aider dans le développement d'une identité culturelle francophone. Chacun de ces aspects fait l'objet de deux questions. La première se veut une question de nature rétrospective demandant aux parents à quel point ils se sentaient confiants *avant* les Ateliers familles. La deuxième mesure plutôt l'état de confiance des parents *après* la fin des Ateliers familles (soit au moment du sondage). Ces items sont notés sur une échelle en 5 points, où 1 indique « peu ou pas de confiance » et 5 indique « une confiance complète ». La cohérence interne s'avère bonne avec un alpha de Cronbach de 0,70 pour la sous-échelle mesurant la perception rétrospective du sentiment d'auto-efficacité et de 0,60 pour la sous-échelle mesurant la perception du sentiment d'auto-efficacité suivant les ateliers.

L'*échelle des connaissances* évalue le niveau de connaissance des parents par rapport à cinq des aspects abordés lors des Ateliers familles, dont le développement des enfants, les façons de faire pour aider à la maison le développement et les apprentissages de l'enfant, les moyens

d'aider l'enfant à être prêt pour l'école, le travail d'une éducatrice en garderie, et les services ou ressources offerts en français dans la communauté. Comme pour l'échelle de perception d'auto-efficacité, chacun de ces aspects a fait l'objet de deux questions : l'une rétrospective (*Comment décrivez-vous votre niveau de connaissance des sujets avant les Ateliers familles?*) et l'autre suivant les Ateliers familles (*Comment décrivez-vous votre niveau de connaissance des sujets maintenant, après les Ateliers familles?*).²⁴ Ces items sont notés sur une échelle en 5 points, où 1 indique « une absence de connaissance » et 5 indique « un excellent niveau de connaissance ». La cohérence interne s'avère acceptable avec un alpha de Cronbach de 0,79 pour la sous-échelle mesurant la perception du niveau de connaissance avant les ateliers et de 0,67 pour la sous-échelle mesurant la perception du niveau de connaissance suivant les ateliers.

L'échelle *de modelage des activités de littératie* est construite avec cinq questions basées sur le modèle des activités de littératie. Elle vise à mesurer l'exemple donné par le parent à son enfant face à l'importance de la lecture et de l'écriture. Le contenu des items porte, entre autres, sur la fréquence à laquelle les parents lisent et écrivent devant leur enfant, ainsi que la fréquence à laquelle les parents laissent à la disposition de leur enfant du papier et des crayons. Les réponses aux questions sont notées sur une échelle en 4 points où 1 indique « jamais » et 4 indique « plusieurs fois par jour ». Cette échelle fut remplie par les parents du groupe programme lors du sondage pré-intervention, du sondage post-intervention administré 3 semaines suivant la fin des ateliers (soit en mars 2008 pour la première cohorte et mars 2009 pour la deuxième cohorte) et un dernier sondage administré sept mois suivant la fin des ateliers (soit en octobre 2008 pour la première cohorte et octobre 2009 pour la deuxième cohorte). La cohérence interne s'avère acceptable avec un alpha de Cronbach de 0,75, 0,65 et de 0,64 observé respectivement lors de chacune de ces administrations.

3.5.5. Dosage

Le nombre d'heures de garde par semaine est une variable importante à prendre en compte pour mesurer le « dosage » du programme que reçoivent les enfants en garderie. Pour recueillir cette information, la SRSA a utilisé le registre des présences et absences que les éducatrices remplissaient déjà dans le cadre de leurs tâches, incluant l'heure d'arrivée et de départ des enfants. Le nombre d'heures que l'enfant a passé en garderie est ensuite compilé de façon hebdomadaire et envoyé de façon mensuelle au bureau d'Ottawa de la SRSA. Pour les fins des analyses d'impact d'un an, les heures passées en garderie ont été analysées en prenant la moyenne d'heures passées en garderie par semaine durant les quatre mois précédant immédiatement l'évaluation des enfants. Pour les analyses d'impact de la deuxième année, la moyenne d'heures passées par semaine durant les huit premiers mois post-intervention figure dans les analyses (voir la section 8.3).

En ce qui a trait aux Ateliers familles, l'assiduité des parents, des enfants et des autres adultes présents a été compilée lors de chaque atelier. Le dosage aux ateliers est calculé en prenant la somme d'ateliers suivis par au moins un adulte (voir la section 6.3 pour les résultats du dosage des Ateliers familles).

²⁴ Pour une discussion éclairée sur les avantages d'utiliser la rétrospective dans le cadre de formations, voir Lamb et Tschillard (2005).

3.6. HYPOTHÈSES À L'ÉTUDE

3.6.1. L'impact du volet en garderie sur les dimensions linguistiques des enfants

Le programme en garderie vise particulièrement la préparation à la scolarisation, l'enrichissement des habiletés langagières des enfants, et le développement d'une identité culturelle francophone. C'est donc par rapport à ces dimensions que nous anticipons observer des effets plus prononcés du programme. Les hypothèses émises à ce sujet sont :

1. Les enfants du groupe programme auront des scores plus élevés dans trois des quatre domaines mesurant la préparation à la scolarisation, dont en langue et communication, en habiletés cognitives, et en conscience de soi et de l'environnement, comparativement aux enfants des groupes témoins. Aucune différence entre le groupe programme et les groupes témoins n'est anticipée au niveau du développement physique et moteur puisque le programme testé ne cible pas cette dimension.
 - Cette hypothèse vise à tester directement la question de recherche principale. La validité interne de l'étude est assurée par l'inclusion du groupe témoin en garderie puisque l'influence d'une garderie en installation sur le développement des enfants constitue en soi un traitement. Elle est également assurée par l'ajout du deuxième groupe témoin qui vise à tenir compte de l'influence d'un milieu de garde informel sur le développement des habiletés langagières en français de l'enfant.
 - Enfin, cette hypothèse tient compte des autres conditions existantes dans la communauté pouvant affecter le développement des enfants. Par exemple, l'environnement communautaire joue un rôle important dans le développement de l'enfant dans la mesure où il rend disponible des ressources (p. ex., parcs, patageuses, pistes cyclables) et des services (p. ex., librairies, bibliothèques, cours de natation, etc.).
2. Les enfants du groupe programme auront une meilleure performance sur les mesures d'habiletés langagières comparativement aux enfants des groupes témoins.
3. Le programme testé aura un impact plus prononcé sur les habiletés langagières des enfants du groupe programme issus de foyers exogames ou de foyers endogames anglophones.
4. L'ampleur de l'impact du programme en garderie sur les dimensions de la préparation à la scolarisation des enfants sera influencée par le niveau d'exposition de ces derniers au programme. Ainsi, une plus grande exposition au programme testé sera associée à une meilleure performance des enfants du groupe programme dans trois des quatre domaines mesurant la préparation à la scolarisation, dont en langue et communication, en habiletés cognitives, et en conscience de soi et de l'environnement, comparativement aux enfants des groupes témoins. Aucune différence entre le groupe programme et les groupes témoins n'est anticipée au niveau du développement physique et moteur puisque le programme testé ne cible pas cette dimension.
5. Les garderies du groupe programme seront caractérisées par des indices de fidélité et de qualité du programme supérieurs à ceux des garderies témoins.

6. L'ampleur de l'impact du programme en garderie sur les dimensions de la préparation à la scolarisation et sur les habiletés langagières variera en fonction du degré de fidélité et de qualité du programme testé.

3.6.2. L'impact du volet en garderie sur les dimensions identitaire et culturelle des enfants

Le programme testé vise à influencer les divers contextes nécessaires au développement des dimensions identitaire et culturelle des jeunes enfants. L'importance de mesurer l'identité culturelle se fonde sur des recherches soulignant que l'exposition des enfants au français dans plusieurs sphères de leur vie fortifie leur sentiment d'identité et d'appartenance à la communauté francophone (Landry et Allard, 2000). Cependant, le concept d'identité culturelle est mesuré à l'adolescence dans ces études. La mesure de ce concept dans ces études correspond bien à la théorie sur le soi d'Erik Erikson (1994) selon laquelle la formation de l'identité débute vers l'âge de 12 ans pour se terminer autour de 25 ans. Ainsi, une mesure directe de l'identité culturelle auprès d'enfants en très bas âge s'avère impossible.

Par contre, une mesure du degré *d'exposition* des enfants au français est possible via les parents. En effet, les parents sont les mieux placés pour nous informer de l'environnement langagier des milieux que l'enfant fréquente et qui sont d'intérêt pour le projet Capacité d'apprentissage, soit la famille, le milieu de garde de jour et l'accès aux services et ressources de langue française. Les informations collectées serviront de variables contrôles dans les analyses d'impact afin de mieux dégager les effets du programme testé sur le développement des habiletés langagières des jeunes enfants.

3.6.3. L'impact du volet Ateliers familles sur les attitudes et comportements des parents

Le contenu du programme des Ateliers familles met particulièrement l'accent sur quatre aspects :

- le parent comme premier éducateur de son enfant;
- les activités de pré-lecture, pré-écriture et pré-numératie susceptibles de favoriser la préparation à la scolarisation de l'enfant;
- l'importance de maximiser l'exposition de l'enfant au français; et
- la collaboration parent-éducatrice dans les apprentissages de l'enfant.

C'est donc sur ces aspects que nous anticipons observer des effets plus prononcés du programme. Les hypothèses émises sont les suivantes. Suite aux Ateliers familles et comparativement aux parents des groupes témoins, les parents du groupe programme rapporteront :

- effectuer significativement plus fréquemment différentes activités de littératie avec leur enfant;
- parler significativement plus souvent en français avec leur enfant durant les activités de littératie.

De plus, nous anticipons que les parents du groupe programme rapporteront, suite aux ateliers :

- un niveau de connaissance significativement plus élevé des dimensions développementales des enfants, des moyens pour les aider à être prêt pour l'école, et des services ou ressources communautaires de langue française;
- des sentiments d'efficacité significativement supérieurs, notamment au niveau des moyens pour aider leur enfant à être prêt pour l'école;
- une fréquence significativement accrue d'occasions de modelage des activités de littératie;
- un changement significatif dans leurs croyances, particulièrement pour les éléments visés par le volet Ateliers familles, notamment : le parent comme premier éducateur de son enfant, la coopération entre éducatrice et parent dans les apprentissages de son enfant, l'importance de l'environnement culturel francophone à la maison, et l'âge propice pour lire à son enfant.

4. Méthodologie employée pour l'étude de la mise en œuvre

Ce chapitre contient la matrice d'évaluation utilisée pour étudier l'intégrité de la mise en œuvre du programme en garderie et le programme des Ateliers familles. L'étude de la mise en œuvre d'un programme sert plusieurs objectifs. Elle permet d'abord de bien comprendre comment un programme élaboré sur papier s'intègre dans un milieu réel. Typiquement, l'évaluation de la mise en œuvre est utilisée pour comprendre quels sont les obstacles, les facilitateurs et les ajustements nécessaires pour faciliter l'appropriation du programme par le milieu. Elle sert également d'outil important, pour mieux comprendre et nuancer les effets du programme. Ainsi, de plus en plus d'études démontrent que le degré de mise en œuvre d'une intervention est en lien avec les effets du programme observés sur les participants (Charlebois, Brendgen, Vitaro, Normandeau, et Bourdreau, 2004; Conduct Problem Prevention Research Group, 1999; Dane et Schneider, 1998; Durlak et DuPre, 2008). Le programme n'aura pas le même impact si plusieurs éléments n'ont pas été implantés et si la qualité de la mise en œuvre est médiocre. Sur ce point, Durlak et DuPre (2008) ont conclu à partir des données de cinq méta-analyses que la taille de l'effet d'une intervention est de deux à trois fois plus grande lorsque le programme est mis en œuvre tel que prévu. Ces résultats soutiennent l'importance de bien documenter et de mesurer le degré de mise en œuvre d'un programme qu'on évalue. Une évaluation rigoureuse de l'implantation d'un programme devient primordiale dans le contexte où l'intervention à l'étude informera des pratiques, voire des politiques à venir. Les résultats alors obtenus doivent être mis en relation étroite tant avec la quantité que la qualité des services obtenus par la population, en plus de prendre en compte d'autres facteurs comme les ajustements qui ont été nécessaires afin de s'adapter au milieu évalué.

Outre l'approche adoptée dans l'étude, un portrait global de la mise en œuvre du projet Capacité d'apprentissage a été dégagé de la triangulation de données provenant de sources variées. L'adoption d'une approche méthodologique à trois niveaux (Patton, 1990) a permis cette triangulation. C'est donc dire que nous avons tiré profit de différents outils de mesure tant quantitatifs (comme dans le cas des échelles de fidélité et de qualité) que qualitatifs (comme pour les entrevues auprès des participants). Cette façon de faire, connue aussi sous le nom de *mixed-method model*, offre l'avantage de documenter de plusieurs façons les effets d'un programme, et permet une évaluation plus approfondie d'un programme (Patton, 2008). La méthodologie employée a également donné la parole à différents acteurs, offrant ainsi différentes perspectives sur la mise en œuvre du programme. Ensuite, les données ont été recueillies par différents instruments de mesure pour évaluer le même phénomène. Cette diversité de données renforce la validité et la fiabilité des résultats en raison de la complémentarité des données collectées.

4.1. MODÈLE D'ANALYSE DES DONNÉES QUALITATIVES POUR L'ÉVALUATION DE LA MISE EN ŒUVRE

Il existe un grand nombre de facteurs qui qualifient le degré de mise en œuvre d'un programme. Le premier modèle à avoir structuré ces éléments est celui de Dane et Schneider (1998). Ce modèle permet de vérifier l'intégrité d'un programme, selon les cinq dimensions

suivantes : la **fidélité**, c'est-à-dire la correspondance entre l'intervention appliquée et le programme planifié; la **qualité de l'implantation** du programme, par exemple le degré de préparation et d'enthousiasme de l'animateur; le **dosage**, soit l'exposition des participants au programme; la **participation**, c'est-à-dire la réponse des participants au programme, mesurée par leur degré de participation et leur enthousiasme; et finalement, la **différenciation entre le programme implanté et les autres interventions** (soit le nouvel apport du programme par rapport à ce qui se fait déjà).

Deux autres études ont suggéré d'ajouter des éléments supplémentaires à ce modèle. Carroll et ses collègues (2007) proposent d'inclure la complexité du programme et les **stratégies qui facilitent la mise en œuvre**. Une dernière dimension ajoutée dans le modèle concerne les **retombées perçues auprès des participants**. Nous avons inclus cette dimension étant donné qu'elle nous renseigne sur les bénéfices perçus par les acteurs impliqués dans le programme. Cette perception aura des répercussions sur la volonté d'assurer la pérennité de l'intervention.

4.2. OUTILS DE COLLECTE DE DONNÉES

Dans le cadre du projet Capacité d'apprentissage, une panoplie d'outils de nature autant quantitative que qualitative ont permis de recueillir de l'information sur la mise en œuvre du programme testé auprès d'une grande diversité de gens.

4.2.1. Étude de la mise en œuvre du volet en garderie

L'étude de la mise en œuvre permet de repérer les écarts entre la théorie et la pratique, les goulots d'étranglement ou autres problèmes dans la mise en œuvre nécessitant d'être rectifiés, les éléments sous-utilisés du programme ainsi que le mouvement des participants entre les éléments du programme. Elle permet également de relever les différences et les similitudes dans la prestation du programme dans les six garderies programme de la première cohorte et les deux garderies programme de la deuxième cohorte en plus de faire des analyses comparatives entre le programme testé et ceux livrés dans les garderies témoins²⁵. Plusieurs mesures sont utilisées, chacune complétant l'autre au niveau du type de renseignement recherché, du niveau de détails et de la source de l'information. Cette façon de faire permet d'arriver à une triangulation de l'information provenant de diverses perspectives, tantôt celle de l'éducatrice, tantôt celle de l'observateur, ou encore celle de la formatrice. Les sections ci-dessous rapportent, respectivement, le détail des instruments de mesure développés pour évaluer la mise en œuvre du programme en garderie, dont la grille des observations en garderie, le protocole d'entrevue en profondeur des éducatrices, le journal de bord des éducatrices et l'assiduité des enfants à la garderie.

Grille d'observations en garderie

Le but des observations en garderie était de recueillir des données pertinentes sur la prestation du programme (telle que définie par les objectifs du projet Capacité d'apprentissage). Selon Durlak et DuPre (2008), les mesures d'observations sont plus susceptibles d'être reliées aux résultats de l'intervention que les mesures auto-rapportées. Un des premiers éléments à

²⁵ Le lecteur est prié de noter que les données d'une des six garderies de la première cohorte n'ont pas été incluses dans les analyses de mise en œuvre puisque le programme n'a pas été implanté dans cette communauté.

observer concerne la fidélité de la mise en œuvre du programme. Il s'agit de s'assurer que l'ensemble des éléments essentiels du programme a été mis en place dans les garderies. Le protocole d'observation demandait donc aux observatrices de vérifier la présence de certains éléments propres au programme testé (comme la présence d'une grille de planification visible pour les parents ou encore la présence d'une routine imagée décrivant aux enfants le déroulement de la journée). Les observatrices devaient également décrire le déroulement des activités, en plus de nous informer sur la réaction des enfants lors des différentes activités.

Les observations ont également permis d'obtenir des renseignements sur la qualité des différents milieux de garde des enfants. À cette fin, certaines sous-échelles ont été empruntées de l'échelle la plus utilisée dans les études passées pour évaluer la qualité des milieux de garde en Amérique du Nord, soit la *Early Childhood Environment Rating Scale*, version révisée, ou l'*ECERS-R* (Harms, Clifford, et Cryier, 1998). Sa version française, l'Échelle d'évaluation de l'environnement préscolaire — Révisée (ÉÉEP-R) est un instrument validé pour évaluer la qualité des services de garde telle que définie par les spécialistes en éducation à la petite enfance de la *National Association for the Education of Young Children* (NAEYC). Les critères d'accréditation définis dans les années 80 par la NAEYC sont basés sur les connaissances scientifiques des facteurs influant sur le développement physique, social, intellectuel et émotionnel des jeunes enfants (National Association for the Education of Young Children, 1984; Bredekamp et Copple, 1999). L'ÉÉEP-R a été utilisée à plusieurs reprises dans des projets francophones pour évaluer la qualité des services de garde. Cette échelle, dont la fiabilité mesurée en termes de validité interne et de fidélité est bien établie, mesure les multiples dimensions d'une garderie qui contribuent à sa qualité globale. Les dimensions couvertes par l'ÉÉEP-R incluent à la fois des aspects structurels et de processus présents dans la garderie.

La qualité est habituellement mesurée selon deux dimensions : la qualité structurelle et la qualité des processus. La qualité structurelle fait référence à un ensemble de règlements mis en place par les lois en vigueur et inclut des éléments comme le ratio enfants/éducatrice, la taille du groupe, le niveau d'éducation du personnel, le taux de roulement du personnel ainsi que la formation continue des éducatrices et leur salaire.²⁶

La qualité des processus renvoie plutôt aux expériences sociales et éducatives des enfants découlant de leurs interactions avec les éducatrices. Cette dimension reconnaît explicitement le rôle important de l'éducatrice dans la création d'un environnement riche et stimulant où les enfants sont incités à apprendre. Certaines études affirment que les interactions impliquant les éducatrices sont l'aspect le plus important d'une bonne qualité en garderie (Committee for Economic Development Research and Policy Committee, 1993).

Dans le cadre du projet Capacité d'apprentissage, l'étude de la mise en œuvre du programme en garderie a exigé l'utilisation de certains indicateurs représentatifs de la qualité structurelle et des processus. Puisque l'intérêt de l'étude était l'évaluation de processus précis inhérents à la prestation du programme, les observatrices ont rempli uniquement les sous-échelles les plus en lien avec les visées du programme. Comme indicateur général de qualité, les observatrices ont recueilli des informations sur la qualité des espaces intérieurs (éclairage, ventilation et espace pour chaque enfant), sur les éléments en place pour assurer la santé et la sécurité des enfants, sur

²⁶ Plusieurs éléments de ce type de qualité comme le salaire, la formation et le niveau d'éducation sont difficiles à obtenir par observation. Ces informations sont plutôt obtenues par les entrevues en profondeur ou en se référant aux données des éducatrices récoltées lors de la signature du formulaire de consentement.

la qualité de l'accueil des enfants à leur arrivée, de même que sur la qualité de la relation entre l'éducatrice et les enfants.

Afin de mesurer la qualité éducative du milieu de garde, les observatrices ont dû décrire les activités observées, à savoir quelles composantes du développement des enfants y étaient ciblées et quelle était la réaction des enfants lors de l'activité. Une attention particulière a été portée sur l'utilisation de la communication pour favoriser le développement linguistique des enfants. Ainsi, les observations ont permis de remplir les sous-échelles suivantes de l'ÉÉEP-R : Utilisation du langage spontané, Soutien à la communication chez les enfants, et Utilisation du langage pour développer les habiletés de raisonnement ou Raisonnement. Étant donné la place importante que le programme faisait à l'exposition des enfants à la lecture, la sous-échelle Livres et images a également été remplie.

Il a été décidé au début du projet Capacité d'apprentissage de ne pas utiliser la grille structurée de l'ÉÉEP-R, puisque ce type d'observation pouvait être perçu comme trop intrusif et nuire à l'obtention de la pleine collaboration des garderies programme et témoins. Rappelons à cet égard les difficultés de recruter les garderies témoins dans certaines communautés (voir la section 5.6 dans le *Rapport de mise en œuvre du projet*, Bérubé et al., 2014). Il a donc été décidé que les observatrices prendraient des notes de façon moins formelle sur les différents éléments de l'ÉÉEP-R pour les sous-catégories ciblées. Préalablement aux observations, chaque observatrice a reçu une formation sur les éléments à observer, ainsi qu'un protocole d'observation détaillé leur rappelant les objectifs de l'observation, les éléments à prendre en note et des exemples concrets de pratiques excellentes, bonnes ou inadéquates. Par la suite, les observations ont été validées en comparant les notes des observatrices à celles de la formatrice et en s'assurant qu'il y avait concordance dans la description de chaque milieu. Les résultats des analyses ont également été présentés aux observatrices afin qu'elles confirment que le portrait de chaque milieu de garde était représentatif.

Afin de compléter le portrait des activités en garderie, plusieurs informations ont été récoltées sur le Cercle d'amis, période durant laquelle les éducatrices du groupe programme devaient lire une histoire aux enfants. Les observations nous renseignent sur le style de lecture pratiqué, à savoir une lecture interactive ou traditionnelle, ainsi que sur l'utilisation de chansons et de comptines en français. Des observations ont également été faites durant les périodes d'activités dirigées et de jeux libres. Les observatrices se sont alors concentrées sur la place laissée aux enfants pour développer leur autonomie et leur créativité.

Les observations ont eu lieu dans les garderies du groupe programme, ainsi que dans les garderies témoins. Ces données nous permettaient de mieux comprendre ce que le programme préscolaire apportait comme nouveauté dans les milieux de garde. Il se pourrait que certaines pratiques soient déjà en place dans les milieux de garde et que le programme ne propose pas de grands changements comparativement à ce qui était déjà en place. Si c'était le cas, il se pourrait que les enfants des deux conditions présentent un développement similaire. La comparaison entre ce qui se fait dans le groupe programme et le groupe témoin en garderie nous permet de vérifier en quoi le programme se démarque des activités prenant place naturellement dans les milieux de garde.

Protocole d'entrevue des éducatrices

Des entrevues ont été effectuées auprès des éducatrices afin de recueillir leurs opinions et leurs expériences sur la mise en œuvre du programme, à savoir sa composante en garderie, de même que son volet d'Ateliers familles pour les éducatrices y ayant participé. Le protocole d'entrevue utilisé se fonde sur une revue de la littérature existante concernant les facteurs pouvant influencer sur la mise en œuvre d'un programme et qui ne sont pas observables ou mesurables de manière quantitative. Ainsi, les questions d'entrevue visaient à prendre en compte les difficultés encourues par les éducatrices dans l'application des différentes composantes du programme, les ajustements faits, ainsi que l'assimilation et l'intégration dans le quotidien de nouvelles pratiques découlant du programme. Les questions d'entrevue visaient également à obtenir les rétroactions des éducatrices sur la formation et les suivis. Enfin, quelques questions portaient sur l'environnement organisationnel, un élément important à considérer dans l'explication du succès ou de l'échec d'un nouveau programme. En effet, il est bien connu que sans appui ni soutien administratif, l'implantation d'un nouveau programme est souvent vouée à l'échec (Chen, 2005).

Les sous-sections ci-dessous détaillent les modalités d'administration de l'entrevue en profondeur et son contenu.

Entrevue en profondeur : modalités d'administration

Les entrevues auprès des éducatrices responsables d'appliquer le programme ont eu lieu deux fois au cours de la durée du projet Capacité d'apprentissage, et ce, pour les six sites de la première cohorte et les deux sites de la deuxième cohorte. Les entrevues ont été enregistrées sous la forme de fichiers audio en format MP3. Les entrevues étaient d'une durée moyenne de 41 minutes, avec un intervalle allant de 22 minutes à 65 minutes. Elles étaient ensuite retranscrites et la transcription a été validée par un juge externe.

Contenu de la grille d'entrevue des éducatrices

La SRSA a construit la grille d'entrevue des éducatrices à partir de la documentation existante concernant les facteurs pouvant nuire à la mise en œuvre d'un programme et qui ne sont pas observables ni mesurables de manière quantitative. Elle permet d'examiner l'appréciation de la formation de base et des visites de suivi, l'expérience des éducatrices quant à l'apprentissage et à la mise en œuvre du nouveau programme, les retombées perçues sur les enfants et sur les parents, de même que la perception des éducatrices quant à leur rôle auprès des enfants. Une emphase est également mise sur l'environnement général de travail. Il est clair que la qualité du programme est influencée par divers facteurs associés à la gestion même de la garderie, et ces questions ont pour but de donner une voix aux éducatrices à ce sujet. Ainsi, deux questions principales ont été posées à cet égard, à savoir la description des conditions de travail actuelles et celle du soutien des supérieurs immédiats dans la garderie face au nouveau programme.

Journal de bord des éducatrices

Le but du journal de bord était de fournir aux éducatrices un moyen de partager ce qu'elles pensent et ressentent face à la mise en place du programme (les belles découvertes, les frustrations ressenties, etc.). Le journal a été rendu disponible dès le début de la mise en œuvre

du volet en garderie. Les informations ont été recueillies de façon ponctuelle, c'est-à-dire que les éducatrices étaient invitées à écrire leurs idées quelques fois par mois. Le journal était confidentiel et rempli de façon anonyme. Pour encourager la participation, les coordonnatrices ont rappelé régulièrement aux éducatrices l'existence du journal de bord. Cette activité facultative a pris la forme de feuilles agrafées, de calepins et de cahiers de notes, selon les communautés. Le journal de bord invitait les éducatrices à partager leurs idées sur quatre thèmes formulés sous forme d'énoncés :

- Ce que j'ai appris et ce que j'ai aimé jusqu'à maintenant du programme en garderie;
- Ce qui a été le plus difficile à mettre en place ou à faire jusqu'à maintenant;
- Mes suggestions pour améliorer le programme;
- J'aimerais ajouter autre chose.

Les informations ont permis d'enrichir les données récoltées par le biais des entrevues auprès des éducatrices.

Assiduité à la garderie

Le nombre d'heures de garde par semaine est une variable importante à prendre en considération si l'on veut mesurer le « dosage » ou le « temps d'exposition » des enfants au programme en vigueur dans les garderies programme et témoins. Cette information était déjà recueillie au moyen du registre des présences et absences utilisé par les garderies. Ainsi, les éducatrices devaient remplir une feuille de présence (avec heures d'arrivée et de départ de chaque enfant) chaque semaine. La SRSA aurait pu utiliser cette feuille pour obtenir l'information sur les enfants participant au projet. Toutefois, le risque que des informations sur des enfants ne participant pas au projet Capacité d'apprentissage soient notées (les parents des enfants ne participant pas au projet Capacité d'apprentissage n'ont jamais donné leur accord pour que cette information soit partagée avec la SRSA) rendait l'utilisation d'une telle feuille inappropriée. Dans le but d'être vigilant à ce sujet, une nouvelle feuille permettant la comptabilisation des présences (moment et durée) a été créée spécifiquement pour le projet Capacité d'apprentissage. Cette feuille de présence incluait seulement le numéro d'identification personnel (NIP) des enfants participant au projet Capacité d'apprentissage et l'espace pour y inscrire les dates et les heures d'arrivée et de départ.

Les coordonnatrices communautaires étaient responsables de se rendre à la garderie pour effectuer la comptabilisation des présences des enfants participant au projet Capacité d'apprentissage deux ou trois fois par mois. Elles transcrivaient les présences enregistrées dans les documents de la garderie sur la feuille de présence du projet Capacité d'apprentissage. Elles acheminaient par la suite les feuilles de présence dûment remplies au bureau d'Ottawa sur une base mensuelle.

4.2.2. Étude de la mise en œuvre des Ateliers familles

L'étude de la mise en œuvre des Ateliers familles comprend plusieurs mesures, les unes complétant les autres sur la méthode de collecte des données et la source d'information. Cette façon de faire permet d'arriver à une triangulation de l'information provenant de différentes perspectives, dont celle de l'intervenante auprès des parents, celle de l'intervenante auprès des enfants, celle de l'observateur ou encore celle des parents participant aux ateliers. Les sections

ci-dessous rapportent, respectivement, le détail des instruments de mesure développés pour le projet Capacité d'apprentissage, soit la grille d'observation des Ateliers familles, le journal de bord des intervenantes, la grille d'entrevue en profondeur auprès des intervenantes, l'évaluation ponctuelle des parents de chaque atelier, le sondage logistique rempli par les parents, le sondage post-intervention des parents, ainsi que l'assiduité des familles et du personnel aux ateliers.

Grille d'observation des Ateliers familles

Les observations des Ateliers familles ont permis de recueillir des données pertinentes sur la prestation du programme (telle que définie par les objectifs du projet Capacité d'apprentissage et ceux du programme Ateliers familles). Le travail a commencé par une revue de la documentation afin de recenser des outils validés permettant de mesurer le degré de la mise en œuvre du programme Ateliers familles et ses impacts sur les participants.

Pour atteindre les buts du projet Capacité d'apprentissage, une grille semi-structurée a été créée pour guider les caractéristiques à observer, bien que les observations elles-mêmes aient pris la forme de notes informelles. Les renseignements recueillis ont permis d'examiner quels thèmes ont été abordés durant les ateliers, quelles étaient les aptitudes de l'intervenante lors de l'animation des ateliers, et quelles étaient les réactions des parents quant aux différents thèmes abordés durant les ateliers.

Les coordonnatrices communautaires et les membres de l'équipe de recherche de la SRSA étaient responsables d'effectuer les observations lors des ateliers. En reconnaissance de la nouveauté du programme, tous les ateliers ont été observés. Le protocole d'observation se concentrait davantage sur le volet parent que sur les volets enfant ou conjoint.

Journal de bord des intervenantes

Le but du journal de bord des intervenantes s'inscrivait dans la même optique que celui des éducatrices, c'est-à-dire qu'il a servi d'outil de communication entre les intervenantes et l'équipe de recherche de la SRSA. Bien que cette activité soit facultative, les intervenantes étaient fortement encouragées à y inscrire de façon ponctuelle leurs observations et réactions face au déroulement des ateliers. Le journal de bord invitait les intervenantes à partager leurs réactions suite aux ateliers, de même qu'à y inscrire leurs observations quant aux réactions des parents face aux ateliers. Il comportait quatre thèmes formulés sous la forme de sept énoncés :

1. Qu'est-ce qui va bien jusqu'à maintenant dans les Ateliers familles?
2. Qu'est-ce qui est le plus difficile à mettre en place et à faire jusqu'à maintenant dans les Ateliers familles?
3. Qu'est-ce que les parents ont aimé ou apprécié?
4. Qu'est-ce que les parents ont moins aimé?
5. Si j'avais à livrer l'atelier une autre fois, qu'est-ce que je ferais différemment?
6. Mes suggestions pour améliorer le programme (en général)...
7. J'aimerais ajouter autre chose.

Le journal était confidentiel et rempli de façon anonyme. L'intervenante était simplement tenue d'inscrire le nom de la communauté, la date de l'atelier pour lequel le journal était rempli et le groupe qu'elle animait (parents ou enfants).

Grille d'entrevue des intervenantes

Le but de l'entrevue des intervenantes était de recueillir leurs opinions et suggestions afin d'améliorer le programme et de voir les aspects qui ont bien fonctionné et ceux qui ont été plus difficiles à mettre en pratique. En l'absence de grilles d'entrevue valides et/ou publiées, la SRSA a développé une grille d'entrevue en grande partie fondée sur la grille d'entrevue utilisée pour les éducatrices. Les questions incluses dans la grille proviennent également d'outils élaborés dans le cadre d'autres projets de recherche effectués à la SRSA et à l'Université d'Ottawa. Enfin, certaines questions s'inspiraient du manuel de formation du programme Ateliers familles.

Entrevue en profondeur : modalités d'administration

Toutes les intervenantes²⁷ auprès des parents, de même que quelques intervenantes auprès des enfants, ont été invitées à prendre part à une entrevue une ou deux semaines suivant la fin des ateliers. Cette décision a été prise en reconnaissance du fait que ce programme est nouveau et vise principalement les parents, d'où l'importance d'interviewer toutes les intervenantes auprès des parents. Les entrevues ont été enregistrées sous la forme de fichiers audio en format MP3. Elles ont toutes été menées soit en personne ou au téléphone. Les entrevues avec les intervenantes auprès des parents ont duré en moyenne 1 heure 8 minutes (allant de 47 minutes à 1 heure 35 minutes). Les entrevues avec les intervenantes auprès des enfants étaient, quant à elles, d'une durée moyenne de 48 minutes (allant de 19 minutes à 1 heure 9 minutes). Par la suite, elles ont été enregistrées et retranscrites, et ce, avec le consentement des répondantes. La transcription a été, par la suite, validée par un juge externe.

Contenu de la grille d'entrevue des intervenantes

La SRSA a construit la grille d'entrevue des intervenantes de façon à capter les facteurs pouvant influencer sur la mise en œuvre d'un programme qui ne sont pas observables ni quantifiables. La grille d'entrevue permettait d'examiner les points suivants :

- Les connaissances au sujet du projet Capacité d'apprentissage, des programmes d'alphabétisation familiale et du programme des Ateliers familles;
- Les réactions initiales face au programme des Ateliers familles;
- La mise en œuvre et la logistique des ateliers (c.-à-d., l'organisation, la durée et le format de l'atelier; les obstacles et défis rencontrés; les avantages et inconvénients relatifs à la prestation de chaque volet — parent, enfant et conjoint; la pertinence du matériel fourni);
- Le contenu du programme (c.-à-d., la pertinence pour les parents; la pertinence du contenu pour les enfants; l'utilité du matériel distribué aux parents);

²⁷ Le féminin est privilégié en reconnaissance du genre des personnes responsables de la prestation des ateliers.

- La formation et le soutien (c.-à-d., la pertinence de la formation de base; les besoins en matière de consultations et l'identification des ressources supplémentaires consultées);
- Les impacts du programme sur l'intervenante au niveau professionnel; et
- Les impacts du programme, observés par les intervenantes ou rapportés par les parents, sur les croyances, attitudes et habitudes parentales.

Évaluation ponctuelle des parents de chaque atelier

Les parents étaient encouragés à remplir une courte appréciation du contenu de l'atelier. L'évaluation ponctuelle recueillait les réactions des parents sur ce qu'ils avaient aimé dans chaque atelier, ce qu'ils avaient moins aimé ou trouvé utile, et ce sur quoi ils auraient aimé en apprendre davantage. Le but de cette évaluation ponctuelle était d'informer les intervenantes et l'équipe de recherche de la SRSA des réactions des parents permettant ainsi de faire des ajustements appropriés lors du prochain atelier. Par exemple, si plusieurs parents signalaient vouloir en connaître davantage sur un certain thème, l'intervenante pouvait fournir des références ou des ressources lors du prochain atelier. L'évaluation ponctuelle était administrée à la fin de chaque atelier et exigeait environ 5 minutes pour répondre.

Sondage logistique rempli par les parents

Lors du dernier atelier, les parents ont rempli un court sondage logistique portant sur le calendrier des ateliers et leur déroulement, l'animation de l'intervenante, les aspects appréciés, le contenu des ateliers, et les réactions des parents face aux ateliers. Le but de ce sondage logistique était d'informer la prestation de nouvelles séries d'ateliers au niveau de la logistique, de leur déroulement et du contenu abordé. Notons que ce sondage était rempli de façon anonyme afin d'encourager les parents à être honnêtes dans leurs rétroactions. Ce sondage, d'une quinzaine de questions, prenait environ 15 minutes à remplir. Il était administré au début du dernier atelier aux parents présents. Il s'ensuit que les parents absents au dixième atelier n'ont pas rempli ce sondage. Précisons que certaines des questions du sondage logistique ont été reprises dans le sondage post-intervention en raison de leur importance pour connaître l'impact des ateliers sur les parents et pour améliorer le programme des Ateliers familles.

Sondage post-intervention des parents

Les parents ont été invités à répondre à des questions suite aux ateliers lors d'un sondage post-intervention. Pour les besoins de l'étude de la mise en œuvre, deux versions du sondage post-intervention ont été créées : l'une adressée aux parents participants (soit ceux ayant assisté à trois ateliers ou plus) et l'autre ciblant les parents non participants (soit ceux ayant participé à deux ateliers ou moins). La règle de trois ateliers ou plus pour créer un premier groupe identifié comme participant aux ateliers a été établie de façon arbitraire. On avait estimé qu'à compter de trois ateliers, les parents auraient une assez bonne idée du déroulement d'un atelier et de son contenu pour être à même de donner des rétroactions éclairées sur ces aspects.

D'intérêt particulier pour l'étude de la mise en œuvre des ateliers familles, le sondage post-intervention comprenait des questions sur l'expérience des parents participant aux Ateliers familles (p. ex., retombées des ateliers au niveau des habitudes quotidiennes des parents; suggestions ou commentaires pour améliorer le programme; facilitateurs ou obstacles à la

participation) tandis que les non-participants aux ateliers devaient répondre à des questions visant à cerner les thèmes d'intérêt ainsi que les raisons sous-jacentes à leur non-participation (p. ex., quels sujets du développement de l'enfant et de la préparation à l'école les parents non participants auraient-ils aimé aborder; suggestions ou commentaires pour améliorer le programme; facilitateurs ou obstacles à la participation). La mise en œuvre du programme a donc été documentée par ces réponses, qui ont fourni des explications sur les obstacles ou les motifs favorisant ou non la participation des parents aux ateliers.

Assiduité aux Ateliers familles

Le nombre de séances auxquelles chacune des familles a assisté est une variable importante à prendre en considération si l'on veut mesurer le « dosage » ou le « temps d'exposition » des parents au programme. Les coordonnatrices communautaires étaient responsables de recueillir cette information. L'assiduité était notée pour les parents, les enfants, les intervenantes auprès des parents, les intervenantes auprès des enfants, les aides intervenantes et toute autre personne impliquée dans la prestation des services. L'assiduité des personnes responsables de la prestation des services était importante à considérer puisque leur présence assurait une continuité dans les relations intervenantes – participants et une certaine stabilité dans la prestation des services.

4.3. MÉTHODE D'ANALYSE DES DONNÉES

4.3.1. Stratégies d'analyse

L'analyse des données s'inspire de la méthodologie de la théorisation ancrée, telle que présentée par Strauss et Corbin (1998). Cette méthodologie permet d'étudier un objet complexe et de comprendre comment il s'insère dans une réalité donnée. Il est ainsi possible d'organiser les données et d'en dégager un sens, tout en restant très près des propos recueillis lors des entrevues ou des réponses à des questions ouvertes. Par ailleurs, la théorisation ancrée accorde une importance particulière au contexte social dans lequel s'inscrit l'objet étudié (Laperrière, 1997).

Le logiciel N-Vivo a été utilisé afin de faciliter la gestion du matériel recueilli. Les données ont été analysées en fonction d'une codification par axe, telle que définie par Strauss et Corbin (1998). Une première analyse verticale a été effectuée, soit l'analyse du contenu provenant de chaque instrument de mesure. L'analyse du contenu de chaque instrument a guidé la création de dimensions. Une analyse transversale a ensuite été réalisée, soit une comparaison entre les résultats des analyses réalisées pour chaque dimension de chacun des instruments de mesure. Cette analyse a permis de définir davantage les particularités de chaque regroupement de données selon la mesure et de former les grands thèmes abordés dans l'ensemble des données provenant de différents instruments.

Par exemple, chaque observation du programme en milieu de garde a été lue en entier et les éléments de l'observation se rapportant à une des dimensions de la mise en œuvre à l'étude ont été regroupés (p. ex., tous les passages indiquant que l'éducatrice a affiché les bricolages des enfants sur les murs de la garderie ont été placés dans la catégorie *fidélité structurelle/bricolage*). Une fois toutes les observations analysées, une matrice a été réalisée à l'aide du logiciel N-Vivo. Chaque rangée de la matrice représente les observations d'une classe en milieu de garde en particulier. Les colonnes, quant à elles, représentent chaque catégorie pour la dimension à l'étude

(p. ex., pour la matrice sur la fidélité structurelle en garderie, les colonnes représentent chaque élément de fidélité retenu pour l'étude). Les résultats de la matrice nous renseignent sur la présence ou l'absence d'éléments dans chaque classe (score 0/1). Finalement, les scores obtenus peuvent être agglomérés afin d'établir un portrait pour les milieux de garde du groupe programme versus ceux du groupe témoin. Le tableau 4.1 illustre par des résultats fictifs une matrice obtenue pour la fidélité structurelle.

Tableau 4.1 : Résultats fictifs pour la matrice sur la fidélité structurelle

Fidélité structurelle dans les milieux de garde, Période 2 : novembre 2007 – février 2008							
Classe	Centres bien identifiés	Affiches image-mot présentes	Grille thématique présente	Activités centrées sur un thème	Routine imagée présente	Routine imagée utilisée	Bricolage
1	0	1	0	0	1	1	1
2	1	1	1	0	1	1	0
3	1	0	1	0	1	1	1
4	1	1	1	1	1	1	1
5	0	1	1	1	1	1	1
6	1	1	1	0	1	1	1
7	1	0	1	0	1	1	1

Au total, les analyses ont permis d'examiner sept dimensions dans le cadre de l'évaluation de la mise en œuvre du programme :

1. la fidélité;
2. la qualité de l'intervention;
3. la différenciation entre le programme testé et les autres programmes;
4. la participation (soit la réponse des participants);
5. le dosage;
6. les obstacles et facilitateurs rencontrés;
7. les retombées perçues auprès des participants.

L'ensemble des analyses a été validé à plusieurs niveaux, d'abord afin de s'assurer que le contenu regroupé dans une même catégorie abordait bien le même thème et que le titre donné à la catégorie était représentatif du contenu. Le contenu de chaque catégorie et son titre ont été par la suite validés par une tierce personne. Enfin, les résultats des matrices ont été validés en comparant les résultats obtenus concernant la présence ou l'absence de différents éléments dans

les milieux de garde aux notes d'observation de la consultante en petite enfance pour s'assurer qu'il y avait concordance entre les portraits d'un même milieu établis selon deux sources différentes. Dans le cas des éléments à mettre en place dans l'environnement, les photos prises dans les milieux de garde ont été consultées, afin de s'assurer qu'un élément considéré comme absent ne se retrouvait pas sur les photos. Finalement, les résultats des analyses ont été présentés aux observatrices afin qu'elles confirment que le portrait de chaque milieu de garde était représentatif.

Afin d'obtenir un portrait global de la mise en œuvre du programme testé, nous avons procédé par la triangulation de données provenant de sources variées. Pour effectuer la triangulation, nous avons utilisé une méthodologie mixte à trois niveaux (Patton, 1990). D'abord, les données sur un même objet d'étude provenaient de diverses perspectives. En effet, nous avons demandé aux participants, aux intervenantes et à des observateurs de nous renseigner sur leur expérience du programme au moyen du même instrument. Ensuite, les données ont été recueillies par différents instruments de mesure pour évaluer le même phénomène. Enfin, l'étude a permis une triangulation des méthodologies. En effet, les renseignements ont été récoltés en utilisant des médiums différents, autant de nature quantitative que qualitative.

5. Résultats de la mise en œuvre du programme en garderie

Ce chapitre présente les résultats de l'évaluation de la mise en œuvre du programme en garderie. Le chapitre débute par une section détaillant la *stratégie d'analyse des observations en garderie*, y inclus le traitement des données manquantes. La section 5.2 détaille les résultats pour la *qualité structurelle* des garderies observées. Cet élément concerne la qualité globale de l'environnement des garderies.

Les prochaines sections portent sur les éléments directement liés au programme en garderie. Plus précisément, la section 5.3 rapporte les résultats concernant la fidélité de la mise en œuvre du programme en garderie. Les résultats rapportés concernent la *fidélité structurelle*, soit la mise en place des éléments du programme dans l'environnement. La section 5.4 porte sur la *fidélité du contenu éducatif*, soit les éléments qui doivent être présents dans les activités auxquelles les enfants sont exposés. Afin de faciliter la lecture, les résultats concernant la qualité de la mise en œuvre des éléments de contenu du programme, notamment les éléments de *qualité éducative* (c.-à-d., le degré de préparation de l'éducatrice), sont présentés en même temps que la fidélité. Les données de ces sections proviennent toutes des notes d'observations faites en garderie. Puisque des observations ont été effectuées dans les garderies participant au programme et dans les garderies du groupe témoin, les résultats nous renseignent également sur la *différenciation*, c'est-à-dire la différence entre ce que le programme apporte aux milieux et ce qui se fait naturellement dans les milieux sans le programme.

La section 5.5 concerne la *qualité des interactions entre les éducatrices et les enfants*. La section 5.6 porte sur le *dosage*, soit la fréquence ou la quantité du programme reçue dans les différentes garderies. La section 5.7 détaille les *réactions des éducatrices* au programme. Les données sont tirées des entrevues auprès des éducatrices. La section 5.8 présente les *facilitateurs et les défis* rencontrés lors de la mise en œuvre du programme. La section 5.9 porte sur les *retombées perçues* du programme sur les enfants et les éducatrices. Les données de la section 5.8 et 5.9 sont principalement tirées des entrevues auprès des éducatrices et de la formatrice, qui ont par la suite été validées par les observations dans les garderies. Les résultats de ces sections découlent également des notes d'observations de la formatrice lors de ses visites dans les garderies du groupe programme. Enfin, un sommaire des résultats est présenté à la section 5.10.

Le lecteur est prié de noter les points suivants :

1. L'étude de la mise en œuvre du programme en garderie se fonde sur les données collectées dans cinq des six communautés de la première cohorte et les deux communautés composant la deuxième cohorte. Les données d'une des communautés de la première cohorte n'ont pas été considérées parce que le programme n'a pas été mis en œuvre dans cette communauté.
2. Les analyses ont révélé que le déroulement de la prestation du programme était similaire entre la première et la deuxième cohorte. Aucune différence marquée n'a été notée au niveau des indices de qualité et fidélité. Les résultats sont donc présentés pour les deux cohortes combinées. Cette approche permet de préserver l'anonymat

des éducatrices de la deuxième cohorte qui étaient peu nombreuses au cours de la deuxième année du programme (c.-à-d. moins de cinq éducatrices).

3. Bien que la deuxième cohorte soit composée de deux des communautés de la première cohorte (c.-à-d., Cornwall et Orléans), la présentation des résultats est facilitée en traitant chacune des cinq communautés de la première cohorte et les deux communautés de la deuxième cohorte de façon indépendante, pour un total de sept sites.

5.1. STRATÉGIE D'ANALYSE DES OBSERVATIONS EN GARDERIE

Durant la première année du projet Capacité d'apprentissage, les différentes observations dans les garderies du groupe programme ont été récoltées entre août 2007 et juin 2008 pour la première cohorte et entre octobre 2008 et juin 2009 pour la deuxième cohorte. Ces observations nous permettent d'obtenir un aperçu du déroulement au cours d'une année type. Les observations ont été regroupées en trois périodes de quatre mois chacune, afin de correspondre aux moments des évaluations des enfants. Dans les faits, certaines garderies du groupe programme ont fait l'objet de plus d'une observation au cours d'une période donnée. La **première période** correspond aux mois d'*août à octobre*. La **deuxième période** s'étale du mois de *novembre au mois de février* et constitue le cœur du programme. Finalement, la **troisième période** rassemble les mois de *mars à juin* (voir tableau 5.1 pour un résumé des périodes d'observations).

En ce qui concerne les observations dans les garderies témoins, elles ont eu lieu moins fréquemment étant donné qu'aucune intervention particulière en classe n'était prévue. Les garderies témoins ont toutes été observées durant la **deuxième période** (de novembre à février) et la majorité d'entre elles (13 des 19 classes) ont été observées une deuxième fois durant la **troisième période**. Il s'ensuit que le nombre d'observations par garderie témoin varie d'un site à l'autre.

La prestation du programme en garderie s'est poursuivie pour une deuxième année dans cinq des sept sites à l'étude. Dans deux des sites (Orléans/1^e cohorte, Orléans/2^e cohorte), le programme s'est arrêté puisque tous les enfants avaient commencé l'école à temps plein.²⁸ Devant le nombre important de changements d'éducatrices offrant le programme testé, il a été décidé de continuer d'effectuer des observations sur une base mensuelle lors de la deuxième année du programme. Comme pour la première année, les observations ont été regroupées en trois périodes de quatre mois chacune correspondant aux évaluations des enfants. Ainsi, la **quatrième période** correspond aux mois d'*août à octobre*, la **cinquième période** s'étend de *novembre à février* et enfin, la **sixième période** s'étale de *mars à juin* (voir tableau 5.1 pour un résumé des périodes d'observations). Lors de la deuxième année, la majorité des garderies témoins ont été observées à une seule reprise, soit lors de la période 5 ou de la période 6.

²⁸ Dans la communauté de Durham, seulement quelques enfants ont participé au programme de garderie lors de la deuxième année parce que la plupart des enfants avaient commencé l'école. Des observations en garderie ont tout de même eu lieu durant cette période et les données ont été incluses dans les analyses.

Tableau 5.1 : Périodes d'observation par année

	1 ^e cohorte	2 ^e cohorte
Première année		
Période 1	août à octobre 2007	octobre 2008
Période 2	novembre 2007 à février 2008	novembre 2008 à février 2009
Période 3	mars à juin 2008	mars à juin 2009
Deuxième année		
Période 4	août à octobre 2008	août à octobre 2009
Période 5	novembre 2008 à février 2009	novembre 2009 à février 2010
Période 6	mars à juin 2009	mars à juin 2010

Durant la première année, les observations ont été effectuées par les coordonnatrices communautaires dans 18 classes formant le groupe programme (12 classes de la première cohorte et 6 classes de la deuxième cohorte) et 19 classes formant le groupe témoin (12 classes de la première cohorte et 7 classes de la deuxième cohorte). Le programme s'est poursuivi pour une deuxième année à Cornwall (deux cohortes), Durham, Edmundston et St-Jean. Durant cette deuxième année du programme, les observations par les coordonnatrices communautaires ont été effectuées dans 10 classes formant le groupe programme (huit classes de la première cohorte et deux classes de la deuxième cohorte) et 14 classes formant le groupe témoin (10 classes de la première cohorte et 4 classes de la deuxième cohorte; voir tableau 5.2).

Tableau 5.2 : Nombre de classes faisant partie du projet selon le groupe expérimental, l'année d'implémentation et la cohorte

	1 ^{re} année		2 ^e année	
	Groupe programme	Groupe témoin	Groupe programme	Groupe témoin
Cohorte 1	12	12	8	10
Cohorte 2	6	7	2	4
Total	18	19	10	14

Traitement des observations manquantes en garderie

Il a été impossible d'effectuer toutes les observations planifiées dans les classes des garderies du groupe programme et témoin au cours des deux années du programme en raison des conflits d'horaire et de certaines réserves exprimées par le personnel des garderies. En conséquence, le nombre d'observations varie d'un site à l'autre au cours d'une période donnée (p. ex., de novembre à février), ce qui a exigé un traitement selon le contexte. Les pourcentages ont été calculés selon les observations effectuées. Les données manquantes sont identifiées dans le texte s'il y a lieu.

Gestion des observations multiples

Plus d'une stratégie ont été utilisées afin de gérer les observations multiples, notamment :

- Prendre **la moyenne** des observations (p. ex., si un élément est présent pour une évaluation et non pour l'autre, nous attribuons une valeur de 0,5 à la classe). L'avantage de cette stratégie est qu'elle représente pleinement les fluctuations observées dans l'application de chacun des éléments.
- Appliquer **une règle flexible** selon laquelle une cote de « 1 » est attribuée si l'élément est présent lors d'au moins une des observations pendant la période spécifiée. Cette approche permet de rapporter l'information de façon à exprimer les résultats en ce qui concerne *des classes* ayant satisfait à un critère donné (p. ex., 11 des 18 classes ont satisfait au critère X).

Le lecteur est prié de noter les points suivants :

1. Dans les tableaux, nous rapportons les résultats en calculant la moyenne des observations. D'une part, la fidélité structurelle (tableau 5.4) et la fidélité du contenu éducatif des éléments du programme (tableau 5.6) sont présentées en pourcentage. D'autre part, la qualité structurelle (tableau 5.3) et la qualité du contenu éducatif (tableau 5.5) sont présentées sous forme de scores moyens.
2. Dans le texte, nous présentons généralement les résultats selon la règle flexible, permettant ainsi de décrire le nombre de classes ayant satisfait à un critère donné lors de la première et la deuxième année de la prestation du programme en garderie.
3. Puisque les pourcentages présentés dans le texte et dans les tableaux sont calculés à partir de différentes méthodes de calcul, il va de soi que les chiffres présentés dans les tableaux ne concorderont pas nécessairement avec ceux présentés dans le texte. Les deux sources d'informations reflètent des aspects de la mise en œuvre distincts bien que complémentaires.

5.2. QUALITÉ STRUCTURELLE ET DIFFÉRENTIATION DES GARDERIES OBSERVÉES

Afin de déterminer si les garderies programme et témoins des deux cohortes offrent un niveau de qualité générale équivalent selon certains aspects mesurables, des indices de qualité ont été calculés pour chacune des classes dans chaque garderie. La qualité générale mesurée dans le cadre de cette étude se penche sur l'aspect structurel des milieux. Compris dans ces indices sont la qualité de l'espace physique, les mesures de santé et sécurité, ainsi que la mesure d'accueil et départ des enfants et des parents par les éducatrices. Ces indices de la qualité ont été mesurés par les sous-échelles de l'ÉÉEP-R (Harms, Clifford, et Cryer, 2010).

Les résultats de chaque sous-échelle de l'ÉÉEP-R sont compilés sur une échelle à sept points définis comme suit : la cote **inadéquate** (1) correspond à des conditions ou soins qui se situent bien en deçà des exigences élémentaires d'un service de garde; la cote **minimale** (3) correspond à des conditions ou soins qui satisfont aux besoins élémentaires et, dans une faible mesure, aux besoins de base au plan du développement; la cote **bonne** (5) correspond à des conditions ou soins qui stimulent, dans une certaine mesure, le développement des enfants; et la cote **excellente**

(7) correspond à des conditions ou soins personnalisés et de qualité supérieure (Harms et Clifford, 1990). Les résultats de cette section sont présentés au tableau 5.3.

Pour la sous-échelle « *Espace intérieur* » de l'ÉÉEP-R, la moyenne pour le groupe programme était de 6,89 sur 7 (É.-T. = 0,32) et celle du groupe témoin était de 5,89 sur 7 (É.-T. = 2,28). La majorité des éducatrices des deux groupes expérimentaux (16 du groupe programme et 15 du groupe témoin) ont obtenu un score de 7 (excellent). Dans l'ensemble, les scores obtenus par les deux groupes expérimentaux se sont avérés entre bons et excellents sur cette dimension de qualité.

Quant à la qualité des « *Mesures d'hygiène* », le groupe programme avait une moyenne de 6,22 sur 7 (É.-T. = 1,17). La qualité du groupe témoin était de 6,00 sur 7 (É.-T. = 1,25). Plusieurs éducatrices ont obtenu un score de 7 (12 du groupe programme et 11 du groupe témoin). Quelques éducatrices ont obtenu un score de 5 (quatre du groupe programme et cinq du groupe témoin) ou 4 (deux du groupe programme et trois du groupe témoin). En général, toutes les garderies du projet avaient un environnement hygiénique de haute qualité.

Finalement, pour l'évaluation de « *Accueil et départ* », la moyenne pour le groupe programme était de 6,11 sur 7 (É.-T. = 1,53) et pour le groupe témoin, la moyenne se situait à 5,68 sur 7 (É.-T. = 0,95). Plusieurs éducatrices ont obtenu un score de 7 (10 du groupe programme et 3 du groupe témoin) ainsi qu'un score de 6 (5 du groupe programme et 10 du groupe témoin), montrant que la qualité d'accueil des garderies du projet Capacité d'apprentissage est de bonne qualité.

En somme, les notes d'observations ont révélé qu'il existe une très légère différence au niveau de la qualité globale des garderies, et ce, à l'avantage des garderies du groupe programme. Ces différences sont toutefois mineures et les deux groupes montrent une bonne qualité (de 5 et plus sur une échelle de 7) pour chacune des dimensions évaluées. Par conséquent, nous pouvons dire qu'il n'y a pas de différenciation entre les garderies du groupe programme et témoin au niveau de qualité structurelle. Pour mettre les résultats suivants en contexte, nous présentons également au tableau 5.3 les moyennes d'une étude effectuée par Japel et ses collègues (2005). Cette étude a évalué le niveau de qualité dans 296 garderies dans la province de Québec en lien avec l'Étude longitudinale du développement des enfants du Québec (l'ÉLDEQ). Les garderies observées étaient à but lucratif, à but non lucratif, et des garderies en milieu familial accréditées. Les résultats de l'étude de Japel, Tremplay, et Côté (2005) montrent que sur chaque sous-échelle de l'ÉÉEP-R utilisée dans cette étude, les garderies participant au projet Capacité d'apprentissage sont supérieures aux garderies québécoises.

Tableau 5.3 : Score moyen de la qualité structurelle des milieux observés selon les dimensions évaluées et le groupe expérimental

Qualité globale des milieux de garde (selon les sous-échelles de l'ÉÉEP-R)	Classes du groupe programme n = 18	Classes du groupe témoin n = 19	Garderies de l'étude de Japel et al. (2005) n = 296
Espace intérieur	6,89 (0,32)	5,89 (2,28)	2,80 (2,09)
Mesures d'hygiène	6,22 (1,17)	6,00 (1,25)	3,06 (2,23)
Accueil et départ	6,11 (1,53)	5,68 (0,95)	5,14 (2,14)

Note : Les indices sont calculés à partir des observations de la première année de mise en œuvre pour la première et la deuxième cohorte combinée. Le nombre d'observations lors de la deuxième année était insuffisant pour calculer ces indices.

5.3. ÉLÉMENTS STRUCTURELS EN PLACE DANS LES MILIEUX OBSERVÉS : FIDÉLITÉ ET DIFFÉRENCIATION

Les observations effectuées par les coordonnatrices communautaires ont permis de vérifier dans quelle mesure les éléments du programme étaient mis en place dans toutes les classes pendant les deux années du projet. Dans cette section, nous rapportons les résultats des analyses de la fidélité structurelle par année pour le groupe programme et le groupe témoin.

La section commence par un portrait des divers éléments mis en place au cours de la première et de la deuxième année de la prestation du programme testé. Les pourcentages rapportés représentent la présence ou l'absence des éléments évalués (c.-à-d., selon la règle flexible). Dans la mesure du possible, les résultats sont rapportés par période d'observation permettant ainsi de mieux apprécier l'évolution de la mise en œuvre dans le temps. Nous terminons par un portrait global des classes à l'étude découpée selon l'année de prestation du programme et des groupes expérimentaux. Ce portrait permet d'évaluer le niveau de fidélité et de stabilité des éléments structurels mis en place dans les garderies du groupe programme et celles du groupe témoin. Dans tous les cas, nous traitons les classes comme unité d'analyse et nous faisons la distinction entre celles provenant des garderies du groupe programme et celles qui proviennent des garderies du groupe témoin. Les résultats de la fidélité des éléments structurels sont présentés selon la règle flexible dans le texte et en fonction de la moyenne des observations (en pourcentages) au tableau 5.4.

L'évaluation de la fidélité structurelle inclut sept éléments, notamment : (1) des centres clairement identifiés, (2) des affiches montrant une image et un mot, (3) des grilles thématiques, (4) des activités centrées autour d'un thème, (5) une routine imagée présente, (6) une routine imagée utilisée et (7) des bricolages affichés.

Fidélité structurelle du programme

Un des premiers éléments structurels du programme était de mettre en place, dans les locaux des garderies du groupe programme, des centres ciblant différents apprentissages comme, par exemple, un centre de blocs, un centre d'art, un centre imaginaire, un centre de jeux de table, etc. La présence de centres d'apprentissage est assez répandue dans les milieux de garde en général. Or, la particularité du programme en garderie était son exigence voulant que les **centres soient clairement identifiés à l'aide d'une image accompagnée d'un mot**. Cette exigence obligeait les éducatrices à clairement définir chacun des centres et permettait d'exposer les enfants à l'univers de l'écrit en appuyant les symboles écrits par des images. Il s'agit ici d'un élément du programme qui a été mis en œuvre dans toutes les garderies du groupe participant à l'intervention, et ce, dès le début du programme (11 des 18 classes ont été observées durant la première période) sauf dans un site (quatre classes). Dans ce site, les éducatrices ont bien identifié les centres seulement dans la troisième période (mars à juin). À l'opposé, les données indiquent que les classes formant le groupe témoin en garderie avaient moins l'habitude de clairement identifier leurs centres d'apprentissage. Bien que toutes les classes du groupe témoin possèdent des centres d'apprentissage, les centres étaient identifiés à l'aide d'affiches présentant une image ou un mot au moins une fois durant la première année dans seulement 8 des 19 classes (42 %). La deuxième année, les centres ont été clairement identifiés au moins une fois dans toutes les classes du groupe programme (100 %), alors que la pratique était observée dans 8 des 14 classes (57 %) du groupe témoin en garderie.

Un deuxième élément du programme concernait l'installation de différentes **affiches montrant une image et un mot** dans les locaux des garderies du groupe programme. Ces affiches permettaient d'exposer quotidiennement les enfants au langage écrit. Huit des 11 classes du groupe programme (73 %) observées ont mis en place cet élément du programme en garderie dans la première période. Cette proportion a augmenté au cours de l'année, à 78 % et 82 % dans la deuxième et troisième période respectivement. En comparaison, cette pratique était présente dans sept classes du groupe témoin en garderie (37 %) au cours de la deuxième période. Lors de la deuxième année, 7 des 10 classes du groupe programme (70 %) ont implanté les affiches image-mot. On retrouvait des affiches image-mot dans 10 des 14 classes des garderies témoins (71 %), réduisant la différence entre les deux groupes à cet égard.

Un troisième élément du programme testé en garderie était que les éducatrices devaient utiliser une **grille thématique** présentant les activités qui seraient faites au cours de la semaine avec les enfants. Cette grille devrait être toujours visible pour les parents à l'entrée de la classe et à jour. Dans 17 des 18 classes (94 %) la grille de planification a été utilisée de façon systématique. En ce qui concerne les classes des garderies témoins, 14 des 19 classes (74 %) utilisaient une grille de planification régulièrement. Quant à la deuxième année, on observe que 8 des 10 classes du groupe programme (80 %) et la moitié des 14 classes du groupe témoin (50 %) ont utilisé ces grilles au moins une fois.

Quatrièmement, en plus d'utiliser les grilles thématiques, les éducatrices devaient **centrer leurs activités autour du thème** choisi pour le mois. Les données révèlent que 16 des 18 classes du groupe programme (89 %) centraient leurs activités autour d'un thème au cours de la première année. Cette pratique se retrouvait dans 14 classes sur 19 des garderies témoins (74 %). La deuxième année, toutes les classes des garderies du groupe programme (100 %) et la moitié des 14 classes du groupe témoin en garderie (50 %) avaient relié leurs activités à un thème.

Un cinquième élément du programme était l'**affichage de la routine imagée**. La routine présentait aux enfants les différents moments de la journée dans l'ordre prévu à l'horaire. Les éducatrices devaient afficher la routine imagée au niveau des enfants dans un endroit non obstrué. La routine imagée a été affichée dans 15 des 18 classes des garderies du groupe programme (83 %) dès la deuxième période d'observation. Il est à noter qu'à partir de la troisième période, toutes les classes du groupe programme (100 %) ont affiché la routine imagée. Par contre, aucune des classes du groupe témoin en garderie n'affichait une routine imagée. Lors de la deuxième année d'activités, 9 des 10 classes du groupe programme (90 %) ont affiché la routine imagée. Dans les classes du groupe témoin une routine illustrée était affichée au moins une fois dans 4 des 14 classes (29 %).

Le sixième élément du programme concernait l'**utilisation de la routine imagée avec les enfants**. Cette pratique permettait aux enfants de comprendre le déroulement de la journée, d'anticiper les activités à venir, et surtout, de se sentir sécurisés. Cet élément a été mis en application dans seulement 11 des 18 classes du groupe programme (61 %) lors de la première année. Aucune des classes du groupe témoin n'a utilisé une routine imagée à ce moment. Pour la deuxième année, seulement 4 des 10 classes du groupe programme (40 %) ont utilisé la routine imagée. Dans les classes du groupe témoin une routine imagée a été utilisée dans une seule classe (7 %).

Le dernier élément à mettre en place dans le cadre du programme en garderie était l'**affichage des bricolages** des enfants dans les locaux de la garderie. L'affichage des bricolages

permettait aux enfants de développer un sentiment d'appartenance à la garderie, en plus de valoriser leurs réalisations. Selon les observations, les bricolages étaient affichés au moins une fois dans toutes les classes du groupe programme (100 %) et dans 16 des 19 classes du groupe témoin (84 %). Les bricolages devaient aussi être placés *au niveau du regard des enfants*. Quinze classes du groupe programme (83 %) ont mis en place cette pratique durant la première année, alors que dans les garderies témoins, cette pratique se retrouvait dans sept des classes observées (37 %). Enfin, les éducatrices devaient *écrire le nom de l'enfant* à l'avant des bricolages, afin d'habituer les enfants à reconnaître leur nom et les lettres qui le forment. Le nom des enfants a été inscrit sur les bricolages au moins une fois durant l'année dans toutes les classes du groupe programme (100 %). Cette pratique était en place dans 16 des 19 classes du groupe témoin (84 %). La deuxième année, toutes les classes du groupe programme sauf une (90 %) ont affiché les bricolages selon les demandes du programme alors que cette pratique se retrouvait dans 5 des 14 classes du groupe témoin (36 %).

L'ensemble des observations a été aggloméré par année afin de dresser un portrait global des deux groupes à l'étude. L'examen du tableau 5.4 révèle qu'en moyenne, pendant la première année, les classes des garderies du groupe programme ont mis en place 78 % des éléments du programme en garderie (étendue de 46 % à 91 %), alors que l'ensemble de ces éléments se retrouvait dans moins de la moitié des classes des garderies témoins (moyenne de 40 %; étendue de 0 % à 71 %). Au cours de la deuxième année, le niveau d'implantation dans les garderies du groupe programme était très similaire, soit une moyenne de 73 % (étendue de 40 % à 90 %). En ce qui concerne les garderies témoins, la moyenne s'élevait à 42 %, soit semblable à la première année (étendue de 4 % à 71 %).

En somme, dans la première année, tous les éléments du programme ont été présents et pratiqués plus souvent dans les garderies du groupe programme comparés à celles du groupe témoin. Plus particulièrement, les centres étaient plus souvent identifiés, les affiches montrant une image et un mot étaient plus présentes, et il y avait une routine imagée présente et utilisée durant la journée. Lors de la deuxième année du programme, tous les éléments ont été mis en place plus souvent par les classes des garderies du groupe programme que celles du groupe témoin, sauf pour la présence des affiches présentant une image et un mot, où on observe des taux similaires entre les deux groupes expérimentaux (70 % versus 71 %). L'ensemble de ces résultats suggère que les éléments du programme ont été majoritairement mis en place dans les classes du groupe programme, et ce, de façon assez stable.

Tableau 5.4 : Pourcentage moyen de fidélité des éléments structurels selon l'année de la prestation du programme et le groupe expérimental

Fidélité des éléments structurels	1 ^{ère} année		2 ^e année	
	Classes du groupe programme n = 18	Classes du groupe témoin n = 19	Classes du groupe programme n = 10	Classes du groupe témoin n = 14
Centres clairement identifiés	89 %	34 %	78 %	45 %
Affiches montrant une image et un mot	79 %	39 %	70 %	71 %
Grilles thématiques	91 %	71 %	73 %	46 %
Activités centrées autour d'un thème	89 %	71 %	90 %	46 %
Routine imagée présente	75 %	0 %	80 %	21 %
Routine imagée utilisée	0 %	0 %	40 %	4 %
Bricolages affichés	79 %	65 %	81 %	58 %
Moyenne des 7 éléments	78 %	40 %	73 %	42 %

5.4. CONTENU ÉDUCATIF : FIDÉLITÉ, QUALITÉ ET DIFFÉRENTIATION DU CONTENU DES ACTIVITÉS

Étant donné l'importance du développement des compétences langagières des enfants qui vivent dans un contexte francophone minoritaire, le programme en garderie mettait une emphase particulière sur l'exposition des enfants à la communication orale par les livres et les chansons, tout en leur offrant de nombreuses opportunités de s'exprimer et de développer leur pensée. Les observations ont permis de vérifier dans quelle mesure le contenu éducatif du programme était mis en place durant les deux années du projet. Dans les paragraphes qui suivent, nous examinons chacun des éléments évalués et les résultats s'y rattachant. Trois principaux éléments sont abordés : (1) *la lecture*, (2) *les conditions propices à la communication en français* et (3) *la pré-écriture*. Trois autres éléments sont également évalués, notamment la *créativité*, *l'autonomie* et les *transitions*. Nous terminons par une présentation synthétique des résultats. Nous rappelons au lecteur que les résultats des éléments de cette section sont rapportés principalement selon la règle flexible dans le texte, en scores moyens au tableau 5.5 pour représenter la qualité du contenu éducatif, et en pourcentages au tableau 5.6 (c.-à-d., selon la moyenne des notes d'observation) pour représenter la fidélité du contenu éducatif.

La lecture

Premièrement, l'emphase a été mise sur l'importance de faire la lecture aux enfants et de mettre des livres de qualité à la disposition des enfants. En effet, plusieurs études ont démontré l'importance de la lecture pour les jeunes enfants. D'ailleurs, comme l'expliquent Snow et ses collègues (1998) l'accès aux livres permet d'exposer les enfants à du nouveau vocabulaire, en plus de permettre un dialogue entre l'adulte et l'enfant. Dans le but d'assurer un environnement stimulant en communication orale, les garderies du groupe programme ont reçu près de 300 livres en plus d'autres ressources de langue française (p. ex., CD, DVD). L'ajout de ces ressources, de même que l'importance mise sur la lecture aux enfants dans le cadre du programme se reflète dans la grande différence de score de qualité entre les classes du groupe

programme et celles du groupe témoin, à la sous-échelle « *Livres et images* » de l'ÉÉEP-R. Cette sous-échelle mesure la quantité et l'accessibilité des livres dans la classe ainsi que la qualité des activités de littérature dirigées par les éducatrices.

Les classes du groupe programme ont obtenu un score moyen de 5,0 (É.-T. = 2,06) sur cette dimension, une cote qualifiant de « bonnes » les activités de littérature dans les garderies du groupe programme. Huit classes ont obtenu un score de 7, huit autres ont obtenu un score de 4, et deux classes ont obtenu un score de 1, indiquant une qualité inadéquate. En comparaison, les classes du groupe témoin en garderie ont obtenu un score de 2,42 (É.-T. = 1,98) signifiant que leur milieu était tout juste au-dessus d'inadéquat pour cette dimension. Sept classes ont obtenu un score entre 4 et 6, alors que 12 classes avaient un score de 1. La principale lacune des éducatrices du groupe programme était qu'elles n'initiaient pas la lecture de façon spontanée durant la journée. Pour leur part, les éducatrices du groupe témoin ne lisaient pas souvent des livres aux enfants et elles n'avaient pas à leur disposition assez de livres pour le nombre d'enfants présents dans la classe.

Cependant, il n'est pas suffisant d'assurer l'accès à du matériel, encore faut-il que les éducatrices l'utilisent. Ainsi, les observations des coordonnatrices communautaires ont confirmé que, durant la première année, les éducatrices de 17 des 18 classes du groupe programme (94 %) *lisaient des livres durant l'activité du cercle d'amis*. Les observations ont aussi souligné qu'au cours de la deuxième période d'observation, 5 des 18 classes du groupe programme (28 %) *lisaient des livres aux enfants à d'autres moments* que lors du cercle d'amis. Cette proportion a augmenté au cours de l'année de sorte que durant la troisième période, les éducatrices de 8 des 18 classes du groupe programme (44 %) ont pris l'habitude de lire des livres aux enfants à d'autres moments dans la journée. De plus, à partir de la troisième période, les éducatrices de sept classes du groupe programme (39 %) ont exposé les enfants à la lecture dans au moins deux contextes différents. Quant aux éducatrices des classes du groupe témoin, les notes d'observations des coordonnatrices pour la première année ont révélé que les éducatrices de 11 classes sur 19 (58 %) ont fait la lecture aux enfants durant le cercle d'amis et que les éducatrices de cinq classes (26 %) ont lu des livres aux enfants dans le cadre d'une activité dirigée. Cela signifie donc que dans toutes les classes des garderies du groupe programme (100 %), les enfants ont été exposés à la lecture au moins une fois au cours de la journée alors que cela n'a été le cas que pour 6 des 19 classes dans les garderies témoins (32 %).

Lors de la deuxième année, les éducatrices de toutes les classes du groupe programme (100 %) ont fait la lecture aux enfants durant le cercle d'amis, et dans six de ces classes (60 %), les éducatrices ont également lu des livres aux enfants dans le cadre d'activités dirigées. Ainsi, les éducatrices de six classes sur neuf du groupe programme (67 %; l'observation pour une classe est manquante) ont exposé les enfants à la lecture dans plus d'un contexte. En ce qui concerne le groupe témoin en garderie, les éducatrices de 5 des 14 classes (36 %) ont pris avantage du cercle d'amis pour faire la lecture aux enfants et les éducatrices de neuf classes (64 %) ont fait la lecture lors des activités dirigées. Les éducatrices de 4 sur 12 classes du groupe témoin (33 %; les observations pour deux classes sont manquantes) ont donc exposé les enfants à la lecture dans au moins deux contextes différents.

Selon les recherches sur le développement des enfants, il semble que la lecture comme telle ne soit pas toujours suffisante pour améliorer la capacité langagière des enfants. En effet, comme l'ont démontré Hargrave et Sénéchal (2000), les enfants profiteraient davantage d'une *lecture*

interactive où l'adulte implique l'enfant dans la discussion sur les éléments de l'histoire et où plusieurs questions sont posées par l'adulte afin d'inciter l'enfant à approfondir sur l'histoire et sur la compréhension qu'il en a. Les enfants exposés à ce type de lecture auraient, d'après les résultats des chercheurs, une étendue de vocabulaire beaucoup plus grande que ceux exposés à une lecture traditionnelle. Dans une étude effectuée par Wasik et ses collègues (2006), l'intervention a été poussée plus loin. En effet, des livres ont été fournis dans les garderies dans des boîtes thématiques contenant également du matériel permettant d'élaborer des activités autour du thème abordé dans les livres. Les éducatrices ont été formées à réutiliser le vocabulaire rencontré durant la lecture lors des activités thématiques. Elles ont également été sensibilisées à l'importance de poser des questions ouvertes et d'expliquer aux enfants le sens des mots ciblés avant la lecture. Les enfants ayant participé à ce programme ont acquis davantage de vocabulaire que les enfants exposés à une lecture de style traditionnel.

Le programme en garderie du projet Capacité d'apprentissage préconisait également la méthode de la lecture interactive. Cette méthode faisait d'ailleurs partie de la formation des éducatrices du programme en garderie. Comme dans l'étude de Wasik et ses collègues (2006), les livres ont été distribués dans les garderies dans des trousseaux thématiques contenant également du matériel destiné à développer des activités autour du thème choisi. Cet arrangement permettait aux éducatrices et aux enfants de réutiliser et de consolider le vocabulaire auquel ils avaient été exposés durant la lecture. L'animation de la lecture devait comprendre la présence de questions ouvertes et de discussions autour du livre. L'éducatrice devait également utiliser du matériel (p. ex., marionnettes, images, objets, accessoires, etc.) pour soutenir la lecture, permettant aux enfants de mieux suivre le déroulement de l'histoire, tout en conservant leur intérêt.

Les données ont révélé que, durant la première année, toutes les éducatrices des classes du groupe programme (100 %) ont animé au moins une fois un livre en respectant les exigences du programme testé. Notons qu'il était impossible de comparer la mise en œuvre de cet élément dans le groupe témoin en raison de données incomplètes; plusieurs éducatrices dans les garderies témoins n'ont pas lu un livre aux enfants durant les observations. La deuxième année, le style de lecture interactif a été retrouvé dans 9 des 10 classes du groupe programme (90 %). Dans le groupe témoin, ce style de lecture a été observé dans une seule classe sur 14 (7 %).

Les conditions propices à la communication en français

Un deuxième aspect important du programme en garderie consistait à donner des opportunités multiples et différentes aux enfants afin de les encourager à communiquer en français. Cet objectif a été atteint dans différents contextes formels comme informels. Il faut mentionner que toutes les éducatrices devaient utiliser des techniques de francisation, entre autres parler uniquement en français aux enfants et répéter à l'enfant en français lorsque celui-ci parlait en anglais. Selon les notes d'observations, les éducatrices des deux groupes expérimentaux ont souvent posé des questions qui encourageaient les enfants à utiliser le français. Elles ont également souvent simplifié leurs demandes plutôt qu'utiliser l'anglais lorsque les enfants ne comprenaient pas les mots français. Enfin, toutes les éducatrices des deux groupes expérimentaux ont parlé uniquement en français avec les autres adultes dans leur classe.

Afin d'examiner davantage les communications entre l'éducatrice et l'enfant, nous avons observé des moments favorisant la communication des enfants et employé les sous-échelles de l'ÉÉEP-R pour l'évaluer. Afin de stimuler la communication chez les enfants, les éducatrices du

groupe programme étaient encouragées à animer des cercles d'amis, c'est-à-dire de regrouper les enfants afin que ceux-ci discutent de différents sujets choisis en fonction des événements de la journée et portant sur le thème en cours. Dès la première année, le *cercle d'amis* a été mis en place dans toutes les garderies du groupe programme. En ce qui concerne les garderies du groupe témoin, 17 classes (89 %) animaient un cercle d'amis lors de la deuxième période d'observation, mais trois classes ont arrêté cette pratique lors de la troisième période. La deuxième année, le cercle d'amis a été observé dans toutes les classes du groupe programme et celles du groupe témoin (les observations pour trois classes du groupe témoin sont manquantes).

La qualité de la communication verbale entre les éducatrices et les enfants dans un contexte informel a été mesurée au moyen de la sous-échelle « *Utilisation du langage spontané* » de l'ÉÉEP-R. Les critères de cette sous-échelle mesurent à quel point les éducatrices encouragent les enfants à s'exprimer en initiant des conversations avec les enfants et en encourageant les enfants à parler entre eux. Il s'avère que les garderies du groupe programme encourageaient davantage la communication spontanée que les garderies témoins. La moyenne des classes des garderies du groupe programme est de 5,33 sur 7 (É.-T. = 1,53) et la moyenne des classes des garderies témoins est de 4,68 sur 7 (É.-T. = 1,67). Il est à noter que 8 des 18 classes des garderies du groupe programme ont obtenu le score maximal de 7, ce qui qualifie leur pratique d'excellente, comparativement à cinq classes des garderies du groupe témoin. Les autres classes des garderies du groupe programme, de même que 11 classes des garderies du groupe témoin, ont obtenu un score de 4, ce qui signifie que leurs pratiques se trouvent entre les cotes minimales (3 sur 7) et bonne (5 sur 7). La principale lacune de ces éducatrices était qu'elles encourageaient peu les enfants à s'engager davantage aux discussions, soit en ajoutant de nouveaux éléments à leurs propos, ou encore, en posant des questions pour approfondir le sujet.

Les observations ont également porté sur le soutien à la communication par des techniques plus formelles. Par exemple, les éducatrices faisant partie du groupe programme étaient encouragées à utiliser du matériel visuel afin d'aider les enfants à communiquer (comme l'utilisation de marionnettes ou de petits personnages permettant de reconstituer une histoire ou d'en élaborer une nouvelle). De plus, elles étaient encouragées à former de petits groupes de discussion autour de sujets précis. Cette dimension a été mesurée à l'aide de la sous-échelle « *Soutien à la communication chez les enfants* » de l'ÉÉEP-R. Quinze des 18 classes des garderies du groupe programme ont obtenu la note maximale à cette sous-échelle contre 11 des 19 classes des garderies témoins. Les autres éducatrices des deux groupes expérimentaux ont obtenu un score de 4 ou plus, indiquant un niveau de qualité au-delà de la cote minimale (3), sauf une éducatrice du groupe programme. La moyenne des deux groupes expérimentaux à l'étude est de 6,39 (É.-T. = 1,58) pour les garderies du groupe programme et de 5,84 (É.-T. = 1,42) pour les garderies témoins. La principale lacune de ces éducatrices était qu'elles introduisaient peu les conversations spontanées durant les activités et jeux libres hors du cercle d'amis pour encourager les enfants à communiquer leurs idées.

Les *chansons et les comptines* présentaient une autre opportunité unique d'exposer les enfants à du nouveau vocabulaire et de les conscientiser aux sons qui forment les mots, comme dans le cas des chansons et des comptines qui utilisent la rime. Les données de la première année ont révélé que les éducatrices de toutes les classes des garderies du groupe programme (100 %) utilisaient les chansons lors du cercle d'amis. Dans sept classes (39 %), les éducatrices faisaient chanter les enfants dans le cadre d'une activité dirigée et dans deux tiers des classes (67 %), elles utilisaient les chansons durant les transitions. La moyenne générale pour cette dimension à

travers les différents moments se situe à 58 %. Par ailleurs, lors de la première année, les éducatrices dans 15 des 19 classes du groupe témoin en garderie (79 %) faisaient chanter les enfants lors du cercle d'amis, dans quatre des classes (21 %), les éducatrices encourageaient les enfants à chanter lors des activités dirigées, et dans une forte proportion des classes (79 %), les éducatrices utilisaient la chanson lors des transitions. La moyenne générale pour cette dimension à travers ces différents moments se situe à 52 %. Durant la deuxième année de mise en œuvre, toutes les éducatrices du groupe programme (100 %) ont fait appel aux chansons lors des cercles d'amis; dans quatre classes (40 %), les enfants chantaient lors des activités dirigées et dans 8 des 10 classes (80 %), les éducatrices chantaient systématiquement avec les enfants lors des transitions. La moyenne générale s'élève à 68 %. En ce qui concerne le groupe témoin, toutes les éducatrices observées chantaient durant le cercle d'amis (100 %; les observations pour quatre classes sont manquantes). Dans une classe (8 %), l'éducatrice chantait avec les enfants lors des activités dirigées (les observations pour deux classes sont manquantes) et dans la moitié des 14 classes (50 %), les éducatrices animaient les transitions à l'aide de chansons. La moyenne générale des classes du groupe témoin pour l'ensemble de ces moments se situe à 50 %. En somme, les résultats révèlent que les deux groupes expérimentaux à l'étude utilisaient les chansons dans une proportion assez semblable.

Les activités qui encouragent le *raisonnement* chez les enfants sont liées au développement langagier et celui de la pensée logique chez les enfants. Ce type d'activité était promu dans les garderies programme et témoins, puisqu'il représentait une opportunité d'encourager l'enfant à s'exprimer et à élaborer ses idées en parlant. Les observations ont d'abord permis de connaître la fréquence à laquelle les garderies faisaient des activités visant à promouvoir le raisonnement chez les enfants, comme des activités de classification et des activités ayant trait aux couleurs ou aux formes, sans égard à l'utilisation du langage par les enfants lors de ces activités. Les éducatrices tiraient avantage de différents moments de la journée pour développer cette habileté chez les enfants. Les observations de la deuxième période (durant la première année) révèlent que les éducatrices de 15 des classes du groupe programme (83 %) profitaient du cercle d'amis pour exposer les enfants à des activités de raisonnement, contre 14 classes du groupe témoin en garderie (74 %). Les éducatrices de 11 des classes du groupe programme (61 %) profitaient des activités dirigées pour le faire, contre 14 classes pour le groupe témoin (74 %). Sept éducatrices (41 %) tiraient avantage du moment de la collation (l'observation pour une classe est manquante), contre deux classes du groupe témoin (13 %; les observations pour quatre classes sont manquantes). De plus, les éducatrices de deux classes du groupe programme (12 %) introduisaient différents concepts aux enfants lors des transitions (l'observation pour une classe est manquante), contre trois éducatrices pour le groupe témoin (16 %). En somme, toutes les éducatrices du groupe programme (100 %) et 14 des éducatrices du groupe témoin en garderie (74 %) utilisaient le raisonnement dans plus d'un contexte. Les deux groupes expérimentaux deviennent équivalents durant la deuxième année, où les enfants font des activités de raisonnement à plus d'une occasion dans à peu près les deux tiers des groupes observés autant dans les garderies du groupe programme (7 classes sur 10, soit 70 %) que dans les garderies du groupe témoin (10 classes sur 14, soit 71 %).

Les observations faites en garderie ont permis d'étudier la qualité des activités de raisonnement en reconstituant la sous-échelle « *Utilisation du langage pour développer les habiletés de raisonnement* » de l'ÉÉEP-R. Cette sous-échelle évalue la façon dont les éducatrices discutent de logique avec les enfants lorsqu'elles jouent avec eux et leurs fournissent

du matériel pour développer leur raisonnement. L'échelle évalue aussi à quel point les éducatrices encouragent les enfants à expliquer le raisonnement qu'ils suivent pendant une activité. Sur cette sous-dimension de l'ÉÉEP-R, les classes du groupe programme ont obtenu une moyenne de 4,67 sur 7 (É.-T. = 1,68), donc une note signifiant que leurs pratiques se trouvent entre la cote minimale et bonne. Quatre éducatrices ont reçu un score de 7 (excellent), la plupart (huit) ont obtenu un score de 4, et une éducatrice a obtenu un score de 1 (inadéquat). Les classes du groupe témoin ont obtenu un score moyen de 3,37 (É.-T. = 2,06), soit un indice de qualité légèrement supérieur à la cote minimale. Seulement deux éducatrices des 19 ont obtenu un score de 7, neuf éducatrices ont reçu un score entre 3 et 4, et cinq éducatrices ont obtenu un score de 1. Presque toutes les éducatrices ont éprouvé des difficultés à plusieurs niveaux sur cette sous-échelle. En effet, elles n'ont pas initié ou encouragé les conversations avec les enfants dans le contexte de jeux de raisonnement, ou raisonné avec les enfants lorsqu'un problème se présentait.

Le tableau 5.5 présente la qualité moyenne du contenu éducatif des milieux de garde selon qu'ils appartiennent au groupe programme ou au groupe témoin. Ce tableau présente également les résultats obtenus par Japel et ses collègues (2005) sur la qualité moyenne du contenu éducatif dans 296 garderies au Québec. Les résultats présentés dans ce tableau suggèrent que la qualité est supérieure dans les garderies du groupe programme pour toutes les sous-échelles mesurées que dans les garderies du groupe témoin ou les garderies au Québec.

Tableau 5.5 : Scores moyens pour la qualité du contenu éducatif selon les dimensions évaluées et le groupe expérimental

Qualité du contenu éducatif (selon les sous-échelles de l'ÉÉEP-R)	Classes du groupe programme n = 18	Classes du groupe témoin n = 19	Garderies de l'étude de Japel et al. (2005) n = 296
Livres et images	5,00 (2,06)	2,42 (1,98)	3,71 (2,14)
Utilisation du langage spontané	5,33 (1,53)	4,68 (1,67)	4,03 (1,95)
Soutien à la communication chez les enfants	6,39 (1,58)	5,84 (1,42)	4,47 (1,96)
Utilisation du langage pour développer les habiletés de raisonnement	4,67 (1,68)	3,37 (2,06)	3,74 (2,00)

Note : Les indices sont calculés à partir des observations de la première année de mise en œuvre pour la première et la deuxième cohorte combinées. Le nombre d'observations lors de la deuxième année était insuffisant pour calculer ces indices.

Pré-écriture

Une troisième dimension observée concerne l'exposition des enfants à la *pré-écriture*. Cette dimension mesure la fréquence à laquelle les enfants tracent des lettres ou écrivent leur nom. Ces exercices aident les enfants à raffiner leurs habiletés de motricité fine, en plus de les aider dans la reconnaissance des lettres de l'alphabet. Cette habileté est associée à la capacité des enfants d'âge scolaire de décoder un texte, une habileté nécessaire à l'apprentissage de la lecture (Lonigan, Schatschneider, Westberg, et The National Early Literacy Panel, 2008). Lors de la deuxième période d'observation, les enfants de 11 des 18 classes du groupe programme (61 %) ont été observés alors qu'ils traçaient ou écrivaient leur nom. On observe une petite augmentation durant la troisième période d'observation (de mars à juin), alors que les éducatrices de 12 classes du groupe programme (71 %) encourageait les enfants à écrire des lettres (l'observation pour une

classe est manquante). À l'opposé, 5 des 19 classes du groupe témoin en garderie (33 %) ont incité les enfants à écrire (les observations pour quatre classes sont manquantes). Lors de la deuxième année, toutes les classes observées des deux groupes expérimentaux sauf une classe du groupe témoin (92 %) ont fait écrire les enfants lors de cette période (les observations pour deux classes du groupe témoin sont manquantes). Il faut dire que les enfants étaient plus âgés, donc plus habiles et plus intéressés à faire cette tâche.

Créativité, autonomie et transitions

Trois derniers éléments du programme devaient être mis en place afin d'implanter le curriculum complet du programme testé. Premièrement, le programme fait une place importante à la ***créativité*** chez les enfants. Dans ce contexte, les enfants étaient encouragés à explorer et à construire eux-mêmes en fonction de leur compréhension de l'environnement. La créativité des enfants était particulièrement encouragée lors des bricolages où les enfants étaient libres de choisir parmi différents matériels et où le produit fini était différent d'un enfant à l'autre. Cette façon encourageait l'autonomie chez l'enfant tout en le valorisant. Les observations ont révélé que les éducatrices de toutes les classes du groupe programme (100 %) ont permis aux enfants d'explorer et d'être créatifs dans leurs réalisations lors de la première année, comparativement à 14 des 19 classes (78 %) du groupe témoin (l'observation pour une classe est manquante). Lors de la deuxième année, toutes les classes du groupe programme, sauf une (90 %), permettaient aux enfants d'être créatifs, alors que cette situation se retrouvait dans sept classes du groupe témoin (58 %; les observations pour deux classes sont manquantes).

Deuxièmement, le programme fait une place importante à l'***autonomie*** chez les enfants. Les observations indiquent toutefois que les instances d'encouragement du développement de l'autonomie ont été tout aussi importantes dans les garderies du groupe programme que dans les garderies du groupe témoin. En effet, toutes les éducatrices observées lors des deux années encourageaient les enfants à être autonomes, par exemple en les incitant à s'habiller ou à se laver les mains seuls, et en leur permettant de choisir leur activité lors des jeux libres.

Enfin, certaines des éducatrices d'un groupe comme de l'autre ont mis des stratégies en place afin de faciliter les ***transitions*** entre les activités, un autre élément jugé important dans le programme testé. Au cours de la première année, au moins 75 % des transitions se sont bien déroulées dans 15 des 18 classes du groupe programme et dans 14 des 19 classes du groupe témoins. Lors de la deuxième année, bien que certaines éducatrices éprouvaient de la difficulté lors des transitions, toutes les transitions observées (100 %) se sont bien déroulées dans 7 des 10 (70 %) classes du groupe programme et dans 12 des 14 classes du groupe témoin (86 %).

En somme, tel que présenté dans le tableau 5.6, le contenu éducatif du programme a été respecté à 74 % par les classes du groupe programme au cours de la première année de la mise en œuvre (avec une étendue de 29 % à 100 %). On observe une mise en application des éléments centraux du programme dans les garderies témoins avec une moyenne de 54 % (étendue de 5 % à 97 %). Tous les éléments du programme ont été présents à un niveau équivalent ou plus élevé dans les classes du groupe programme que celles du groupe témoin pendant la première année. Néanmoins, quelques éléments ont été mis en place beaucoup plus souvent dans les garderies du groupe programme. En particulier, la lecture durant le cercle d'amis était plus fréquente, le style de lecture interactive était plus souvent identifié, et il y avait plus d'emphase sur la pré-écriture. Au cours de la deuxième année, le niveau d'application s'est élevé pour les deux groupes. Pour

les garderies du groupe programme, la moyenne a augmenté à 83 % (étendue de 50 % à 100 %) alors que pour les garderies témoins, elle s'est élevée à 63 % (étendue de 9 % à 100 %). C'est donc dire que les deux milieux partageaient certains points en commun, mais que le groupe programme conservait certaines particularités. En effet, lors de la deuxième année du programme, les garderies du groupe programme ont mis en place tous les éléments aussi souvent ou plus souvent que les garderies du groupe témoin, sauf un élément (c.-à-d., le déroulement efficace des transitions).

Tableau 5.6 : Pourcentages moyens de la fidélité du contenu éducatif selon l'année de la prestation du programme et selon le groupe expérimental

Fidélité du contenu éducatif	1 ^{ère} année		2 ^e année	
	Classes du groupe programme n = 18	Classes du groupe témoin n = 19	Classes du groupe programme n = 10	Classes du groupe témoin n = 14
La lecture				
Lecture lors du cercle d'amis	93 %	53 %	100 %	45 %
Lecture plus d'une fois lors de l'observation	29 %	5 %	65 %	35 %
Lecture interactive	87 %	24 %	78 %	9 %
Les conditions propices à la communication en français				
Cercle d'amis	100 %	84 %	100 %	100 %
Chansons et comptines	58 %	52 %	68 %	50 %
Raisonnement	44 %	45 %	50 %	47 %
La pré-écriture				
Pré-écriture	63 %	33 %	100 %	91 %
Trois autres éléments				
Créativité	84 %	69 %	90 %	64 %
Autonomie	97 %	97 %	100 %	100 %
Transitions	84 %	80 %	75 %	85 %
Moyenne des 10 éléments	74 %	54 %	83 %	63 %

5.5. LA QUALITÉ DES INTERACTIONS ENTRE LES ÉDUCATRICES ET LES ENFANTS

Les observations nous ont permis de comparer la qualité de la relation entre l'éducatrice et les enfants (c.-à-d., le niveau de sensibilité des éducatrices) entre les garderies du groupe programme et celles du groupe témoin. Une sous-échelle de l'ÉÉEP-R « Interactions personnel-enfant » a été utilisée pour déterminer la sensibilité des éducatrices auprès des enfants. Cette sous-échelle mesure à quel point le personnel répond aux besoins des enfants, est chaleureux, respecte les enfants, et aime être avec les enfants.

L'échelle a révélé que les éducatrices du groupe programme et celles du groupe témoin étaient très sensibles aux besoins des enfants. La moyenne du groupe programme était de 6,22 sur 7 (É.-T. = 1,63) alors que celle du groupe témoin était de 5,26 sur 7 (É.-T. = 2,31). Pour le groupe programme, 13 des 18 classes ont obtenu un score de 7 (excellent), deux ont obtenu un score de 6 (entre bon et excellent), et une a obtenu un score de 1 (inadéquat). Quant au groupe témoin, 10 classes ont obtenu un score de 7, deux ont obtenu un score de 6, et trois ont reçu un score de 1. Ces observations démontrent qu'il existe des différences mineures entre les groupes expérimentaux. Ajoutons que les deux groupes montrent une bonne qualité d'interaction entre les éducatrices et les enfants dans les garderies.

5.6. DOSAGE

Comme il a été mentionné précédemment, le programme s'est déroulé pendant une année seulement dans certaines communautés alors qu'il s'est poursuivi pour une deuxième année dans d'autres communautés (voir tableau 2.1 au chapitre 2 du présent rapport pour un calendrier de la prestation du programme dans chacune des communautés pour les deux cohortes).

Le dosage, calculé selon la moyenne du nombre d'heures que les enfants ont fréquenté la garderie par semaine, est décrit en fonction du groupe expérimental et de la cohorte pour les deux années du programme. En moyenne, les enfants des deux groupes expérimentaux ont fréquenté la garderie pendant un nombre similaire d'heures par semaine au cours des deux années (voir tableau 5.7). Spécifiquement, durant la première année, les enfants de la première cohorte du groupe programme ont fréquenté la garderie en moyenne 29,1 heures par semaine, alors que les enfants de la deuxième cohorte l'ont fréquentée en moyenne 26,5 heures par semaine (moyenne pour les deux cohortes combinées = 27,8 heures par semaine). Les enfants du groupe témoin de la première cohorte ont, quant à eux, fréquenté la garderie en moyenne 28,2 heures par semaine, alors que ceux de la deuxième cohorte l'ont fréquenté en moyenne 31,5 heures par semaine (moyenne pour les deux cohortes combinées = 29,8 heures par semaine).

Durant la deuxième année du programme, les enfants du groupe programme de la première cohorte ont fréquenté la garderie en moyenne 24,8 heures par semaine, tandis que les enfants de la deuxième cohorte l'ont fréquentée en moyenne 11,1 heures par semaine (moyenne pour les deux cohortes combinées = 18,0 heures par semaine). Pour leur part, lors de la deuxième année, les enfants dans le groupe témoin de la première cohorte ont fréquenté la garderie en moyenne 26,9 heures par semaine, alors que les enfants de la deuxième cohorte l'ont fréquentée en moyenne 11,2 heures par semaine (moyenne pour les deux cohortes combinées = 19,0). Tel que démontré dans le tableau 5.7, durant la deuxième année du programme, les enfants de la deuxième cohorte ont passé moins de temps en garderie que ceux de la première cohorte, et ce, parce qu'ils fréquentaient la maternelle soit à temps partiel ou à temps plein. Il est à noter toutefois que le nombre d'heures passées en garderies par semaine durant la deuxième année est similaire entre les deux groupes expérimentaux, soit 11,1 heures pour les enfants du groupe programme et 11,2 heures pour ceux du groupe témoin.

Tableau 5.7 : Dosage en heures par semaine que les enfants ont fréquenté la garderie selon le groupe expérimental et cohorte

	1 ^e année		2 ^e année	
	Groupe programme	Groupe témoin	Groupe programme	Groupe témoin
Cohorte 1	29,1	28,2	24,8	26,9
Cohorte 2	26,5	31,5	11,1	11,2
Total	27,8	29,8	18,0	19,0

5.7. RÉACTIONS DES ÉDUCATRICES DU GROUPE PROGRAMME

Les réactions des éducatrices au programme ont été mesurées par le biais des entrevues de mi-parcours et de sortie auprès des éducatrices qui ont participé dans le groupe programme. Un échantillon ciblé d'éducatrices chefs ou celles faisant la programmation en classe ont été interviewées par les coordonnatrices communautaires. Les entrevues ont été réalisées avec les éducatrices soit six mois après le début de leur participation (entrevue mi-parcours) ou à la fin de la prestation du programme (entrevue de sortie). Au total, 20 éducatrices des deux cohortes ont accepté d'être interviewées sur leur participation. Quatre d'entre elles sont passées en entrevue deux fois, c'est-à-dire qu'elles ont participé à l'entrevue mi-parcours et l'entrevue de sortie. Les résultats présentent les réactions positives, mixtes et négatives rapportées par les éducatrices des deux cohortes combinées. Nous rappelons au lecteur que les cinq communautés de la première cohorte et les deux communautés de la deuxième cohorte sont traités de façon indépendante, pour une somme de sept sites.

Réactions positives au programme

Deux éléments sont ressortis comme étant les plus aimés par les éducatrices participant au projet, notamment, le cercle d'amis et la routine imagée. D'abord, bien que le cercle d'amis ne soit pas une nouvelle pratique, les éducatrices dans quatre des sept sites ont rapporté avoir beaucoup aimé cet élément du programme.

« Moi j'ai beaucoup aimé le cercle d'amis, la façon de le faire, comme la façon que la formatrice nous a enseignée, appris à le faire, avec beaucoup de chansons et la lecture du conte, j'ai bien aimé ça moi. » (Éducatrice)

Deuxièmement, la routine imagée est une composante unique au programme testé. La fonction de la routine imagée est décrite à la section 4.2. Dans trois des sept sites, les éducatrices ont soulevé la routine imagée comme étant un élément fantastique du programme. Il faut aussi souligner que des impacts associés à l'utilisation de la routine imagée sont mentionnés dans la section 5.9 sur les retombées pour les enfants.

« L'horaire imagé, c'est excellent, s'il y a des changements dans une journée tu peux le faire avec la routine. Tu peux l'expliquer qu'on va changer ça avec ça, puis ce n'est pas comme « Oh! Qu'est-ce qui arrive? » s'il y a quelque chose qui est différent parce que je sais qu'il y en a qui ont besoin de la stabilité, de la structure pendant la journée. » (Éducatrice)

Réactions mixtes au programme

Certains éléments du programme ont entraîné des réactions mixtes chez les éducatrices. Le principal élément ayant suscité ce type de réactions est lié à l'encouragement à la créativité chez les enfants dans leurs jeux et bricolages. Alors que certaines éducatrices ont déclaré avoir bien apprécié cet élément, d'autres éducatrices ont mentionné que cet élément était difficile à mettre en place.

Tel que souligné par la formatrice, la créativité était une des composantes principales du programme en garderie. Les éducatrices ont dû laisser les enfants créer des bricolages libres et les encourager à être créatifs dans tous leurs jeux. Les éducatrices dans quatre des sept sites ont mentionné avoir vraiment apprécié le volet créativité puisque ce dernier leur avait ouvert les yeux quant aux capacités des enfants.

« C'est le bricolage que j'ai apprécié le plus (...). C'est le temps de t'asseoir avec les enfants puis communiquer avec eux, et en même temps, c'est de voir leur développement, qu'est-ce qu'ils font, puis leur originalité. (...) Quand je vois les bricolages sur le mur, je peux voir comment que l'enfant voit les choses, de la façon que lui il les voit et non pas ce que moi j'ai demandé de faire. » (Éducatrice)

Malgré les réactions positives, les éducatrices de deux des sept sites ont déclaré que les bricolages libres n'offraient pas assez d'encadrement aux enfants et que ces derniers auraient profité d'instructions supplémentaires sur la façon de s'y prendre pour découper, dessiner, ou d'utiliser le matériel de bricolage. Cette difficulté avec le volet créativité a également été confirmée par la formatrice lors de ses visites aux garderies programmes et dans ses réponses aux entrevues.

« C'était toujours la même chose que les enfants faisaient parce que (...) c'était du bricolage libre. Il n'y avait pas d'amis qui allaient là. Donc, c'était un peu une perte d'espace ou une perte de temps, tandis que moi j'aurais préféré faire un bricolage nouveau ou leur expliquer, « c'est comme ça qu'on fait ça », pour vraiment qu'ils connaissent comment faire leur bricolage. » (Éducatrice)

Réactions négatives au programme

L'élément du programme qui a suscité le plus de réactions négatives concernait le fait de devoir toujours communiquer en français, même avec les enfants qui ne comprennent que l'anglais. Cet aspect du programme a été mentionné dans trois des sept sites.

« Avec mes anglophones, parler seulement en français est un désavantage, parce qu'il y a des choses qui se règlent si tu les dis en anglais, parce que l'enfant va comprendre. » (Éducatrice)

5.8. FACILITATEURS ET DÉFIS

Des entrevues ont été effectuées auprès des éducatrices, de même qu'auprès de la formatrice, afin de recueillir leurs impressions sur le déroulement du programme en garderie. Certaines des questions posées lors des entrevues portaient sur les éléments qui ont facilité ou rendu difficile la mise en œuvre du programme. Les informations ont aussi parfois été recueillies par le biais d'observations notés par la formatrice durant ses formations et ses visites dans les garderies du

groupe programme. Dans les propos des éducatrices et de la formatrice, il est rapidement devenu évident que la présence d'un élément était identifiée comme un facilitateur alors que son absence était perçue comme un obstacle à la bonne démarche du programme. Par exemple, les éducatrices ont trouvé facilitant d'avoir du temps pour planifier leurs activités. Les éducatrices qui n'ont pas eu cette opportunité ont mentionné qu'un obstacle était de ne pas avoir de temps pour planifier leurs activités. Afin de faciliter la lecture, nous avons écrit le texte en présentant les éléments de façon positive, c'est-à-dire en parlant de facilitateurs plutôt que d'obstacles. Lors de la lecture, il faut comprendre que l'absence d'un facilitateur représente un obstacle à la mise en œuvre.

Les résultats de cette section sont regroupés en trois parties, à l'intérieur desquelles les éléments sont présentés en ordre d'importance selon les analyses.

Les facilitateurs principaux selon les éducatrices et la formatrice :

- Le soutien du programme par la direction;
- Le travail d'équipe entre éducatrices;
- Le matériel fourni aux garderies;
- La flexibilité du programme;
- Les suivis de la formatrice;
- Le soutien des coordonnatrices communautaires.

Les facilitateurs supplémentaires selon la formatrice :

- L'ouverture des éducatrices aux rétroactions et à de nouvelles idées;
- L'expérience de travail auprès des enfants.

Les défis particuliers selon les éducatrices et la formatrice :

- L'aménagement du temps relié à la prestation du programme;
- Thèmes difficiles à mettre en place;
- Roulement du personnel.

Facilitateurs principaux selon les éducatrices et la formatrice

Premièrement, presque toutes les éducatrices (dans six des sept sites) ont mentionné l'importance d'avoir un bon soutien de la part de la direction du service de garde afin de bien appliquer le programme. Le programme testé nécessitait un investissement supplémentaire de temps comparativement au programme que les éducatrices avaient suivi avant leur participation au projet Capacité d'apprentissage. Les éducatrices devaient planifier leurs activités, et certaines activités, comme l'animation de livres, exigeaient que les éducatrices préparent leur matériel à l'avance. Le soutien de la direction du service de garde devenait alors essentiel, car il permettait de dégager le temps requis pour la planification des activités demandées par le programme. Le soutien de la direction du service de garde a également été mentionné par la formatrice du

programme comme étant un élément important qui faisait une véritable différence dans la mise en œuvre.

« Elle [la directrice] a suivi les sessions de quatre jours aussi. Oui, je suis sûre que si elle trouvait que ce n'était pas la peine de le faire, elle ne l'aurait pas fait. Elle nous a encouragées beaucoup (...) encouragée d'aller dans les boîtes; elle nous donnait du temps pour faire nos planifications, pour faire certaine qu'elles étaient faites. Si l'on avait besoin de n'importe quoi, elle nous aidait à faire les recherches aussi sur l'ordi. » (Éducatrice)

À l'opposé, l'absence d'un bon soutien de la part de la direction du service de garde a été considérée comme un obstacle à l'application du programme. Cet élément a été mentionné dans cinq des sept sites. En effet, le manque de soutien représentait un grand défi pour les éducatrices. Cette lacune au niveau de l'appui de la direction se manifestait, entre autres, par un manque de temps accordé aux éducatrices pour la planification des activités demandées par le programme. Cet élément a aussi été mentionné par la formatrice comme un obstacle qui empêchait les éducatrices de réaliser leur plein potentiel.

Un deuxième élément qui a facilité la mise en œuvre du programme a été le soutien que les éducatrices se donnaient mutuellement. Dans la plupart des classes, deux éducatrices s'occupaient ensemble d'un groupe variant entre 10 et 16 enfants. La bonne entente entre les deux éducatrices influençait grandement la qualité de la mise en œuvre. Elle permettait une meilleure planification des activités, ainsi qu'une meilleure harmonisation de l'application du programme au quotidien. Les éducatrices dans six des sept sites, ainsi que la formatrice, ont rapporté qu'un bon travail d'équipe facilitait la livraison du programme.

« J'ai une bonne relation avec [les autres éducatrices], et on s'aide pas mal toutes. Si j'ai besoin d'aide pour quelque chose, je sais que je peux aller à côté de moi. Vraiment pour toutes les éducatrices, si je leur demandais d'aider pour de quoi je sais qu'elles seraient là. » (Éducatrice)

Certaines éducatrices ont rapporté avoir connu un manque de soutien de la part de leurs pairs. En effet, dans trois des sept sites, les éducatrices ont indiqué que le travail d'équipe avec une éducatrice irrespectueuse ou déplaisante rendait la mise en place du programme plus difficile. Ce manque de collaboration entre éducatrices réduisait la qualité de l'application du programme dans certaines classes. La formatrice a aussi rapporté que les conflits entre les éducatrices avaient un impact sur leurs capacités à bien livrer le programme.

Troisièmement, l'apport de plusieurs nouvelles ressources à la garderie a été un élément qui a favorisé l'application du programme. En effet, les activités de mise en œuvre prévoyaient fournir à chacune des garderies du groupe programme une grande quantité de matériel afin d'assurer que toutes avaient les outils nécessaires pour appliquer l'ensemble des composantes du programme. Cet apport en matériel a représenté un atout important dans presque tous les sites (six des sept sites). La formatrice a d'ailleurs indiqué qu'à plusieurs des garderies, ces ressources ont permis de remplacer le vieux matériel.

« Le programme nous a donné beaucoup du matériel. À chaque thème on avait du matériel, du neuf en plus, et ça nous aidait. (...) Voilà! On ne manquait de rien. » (Éducatrice)

Malgré la quantité considérable de matériel offerte aux garderies, les éducatrices de trois des sept sites ont mentionné qu'elles auraient profité de ressources supplémentaires pour présenter les thèmes en profondeur.

Quatrièmement, les éducatrices dans quatre des sept sites ont mentionné l'importance des visites de suivi de la formatrice tout au long du projet. Les éducatrices ont particulièrement aimé l'approche positive de la formatrice et ses encouragements. En outre, la formatrice a apporté beaucoup de matériel et de documentation dans les garderies afin de faciliter l'application du programme. Il s'agit d'un autre aspect grandement apprécié par les éducatrices.

« J'ai bien aimé [les rétroactions de la formatrice]. Elle nous disait : OK, je vois le progrès qu'il y a dans la salle. Je vois la différence. Puis elle disait : vous avez fait ça. Je vais montrer aux autres garderies comment c'est possible. (...) On n'a pas souvent des encouragements. Moi, c'est les observations puis les rétroactions que j'ai adorées. » (Éducatrice)

Cinquièmement, les éducatrices dans quatre des sept sites ont souligné avoir apprécié la flexibilité offerte par le programme. La structure du programme et sa flexibilité donnent aux éducatrices l'inspiration qui anime le choix d'activités à faire avec les enfants, adaptés selon les besoins des enfants dans leur classe. Cet aspect du programme permettait aux éducatrices de changer leurs routines pour répondre aux demandes de chaque journée.

« Comme différentes activités, différents centres et quoi mettre dans les centres selon le groupe d'âge, selon le thème et puis, tu voyais comment le programme avançait. Je trouvais que ça avançait très bien, parfait. » (Éducatrice)

Un sixième facilitateur selon les éducatrices dans deux des sept sites et la formatrice est le soutien de la coordonnatrice communautaire. La coordonnatrice communautaire était la personne responsable du projet Capacité d'apprentissage au sein de la communauté. Selon les notes d'observations de la formatrice, le soutien de la coordonnatrice était présent dans la plupart des communautés des deux cohortes. Les coordonnatrices communautaires témoignaient de ce soutien en étant à l'écoute des besoins du personnel et elles possédaient une bonne maîtrise du domaine de la petite enfance.

Facilitateurs supplémentaires selon la formatrice

Un premier élément noté par la formatrice ayant un impact sur la mise en œuvre du programme concerne l'ouverture des éducatrices aux rétroactions et à. Dans le cadre d'un programme qui comprenait des visites régulières de suivi, cet élément prenait toute son importance. Les éducatrices ouvertes aux rétroactions ont pu bénéficier de chacune des visites de la formatrice et enrichir la façon dont elles mettaient en œuvre le programme.

« À noter que [l'éducatrice] a pris mes commentaires et mes suggestions à cœur et au sérieux. Le lendemain, elle a réaménagé son local. (...) Elle a avoué que déjà les enfants sont plus calmes et elle a plus de contrôle. » (Formatrice)

Les éducatrices qui n'étaient pas ouvertes aux rétroactions n'ont pas tiré profit des visites de la formatrice et n'ont pas changé la façon dont elles mettaient en œuvre le programme. Certaines éducatrices ne prenaient pas en considération les changements suggérés par la formatrice tandis que d'autres ne semblaient pas bien comprendre les particularités du programme.

Le deuxième élément facilitateur mentionné par la formatrice, concernait le niveau d'expérience des éducatrices. En effet, selon la formatrice, le programme semblait plus facile à mettre en place lorsque les éducatrices avaient déjà accumulé une certaine expérience de travail auprès des enfants. Elles étaient alors plus en mesure de répondre aux demandes du programme en raison d'une bonne connaissance en matière de besoins et de gestion des enfants. Certaines éducatrices montraient aussi un talent naturel envers les enfants, ce qui favorisait à la fois l'application et la qualité du programme. Dans la même veine, les éducatrices qui ont participé au projet pendant deux années consécutives ont maîtrisé davantage les éléments du programme durant la deuxième année.

« C'était beaucoup plus facile pour celles qui avaient déjà compris. Elles ont continué à s'améliorer. Celles qui avaient quand des difficultés la première année, oui ça s'est amélioré la deuxième année. » (Formatrice)

À l'opposé, un manque d'expérience de travail et de formation de base représentaient des défis à la mise en œuvre du programme. En effet, les conduites de certaines éducatrices mettaient clairement en évidence leur manque d'expérience. Par exemple, elles n'étaient pas confiantes dans leur application du programme et dans leurs interactions avec les enfants.

Défis particuliers selon les éducatrices et la formatrice

Un premier défi concerne l'aménagement du temps relié à la prestation du programme testé. En particulier, les éducatrices de cinq des sept sites ont mentionné avoir manqué de temps pour la planification du programme. De plus, les éducatrices dans trois des sept sites auraient désiré avoir plus de temps pour préparer la classe selon les exigences du programme. Enfin, les éducatrices de trois des sept sites ont jugé difficile d'insérer le programme dans une journée qui était déjà bien remplie. Il a donc été parfois difficile de trouver du temps pour intégrer certaines activités propres au programme.

« Des fois, on voudrait faire des activités et on manque de temps. On se rend compte, ah, il est déjà cette heure-là. C'est le temps qui passe, qui est un gros défi. » (Éducatrice)

Un deuxième défi du programme testé touche aux thèmes à aborder. Le programme comprend un thème général pour chaque mois avec des sous-thèmes pour chaque semaine. Dans trois des sept sites, les éducatrices ont rapporté que certains thèmes étaient plus difficiles à aborder en raison du manque de matériel éducatif fourni aux classes (p. ex., la cabane à sucre, les formes).

Troisièmement, la stabilité du personnel s'est avérée essentielle à la réussite du programme. Le programme comprenait un certain niveau de complexité, exigeant une période d'adaptation plus ou moins longue selon les éducatrices. Par conséquent, la mise en œuvre du programme s'est avérée plus difficile dans les garderies où le personnel était souvent en rotation. Ce défi a été mentionné par les éducatrices de deux sites et par la formatrice.

5.9. RETOMBÉES PERÇUES

Les entrevues des 20 éducatrices nous renseignent sur ce qu'elles perçoivent comme étant des retombées positives du programme auprès des enfants. Les éducatrices ont également fait

part des changements qu'elles avaient apportés à leur propre pratique avec les enfants. Les entrevues avec la formatrice du programme en garderie servent comme une source d'information supplémentaire aux réponses des éducatrices.

Retombées perçues pour les enfants

D'abord, plusieurs effets positifs chez les enfants découlent de la routine imagée. Selon les éducatrices (dans six des sept sites), la routine imagée encourageait les enfants à être autonomes, calmes et sécurisés, à mieux comprendre la routine et à développer davantage une référence dans l'espace-temps.

Deuxièmement, les éducatrices (dans cinq des sept sites) ont rapporté que le programme permettait de bien préparer les enfants à la scolarisation. Cette préparation renvoie à plusieurs dimensions, notamment, la capacité des enfants à suivre la routine, à répondre aux attentes du milieu scolaire et à développer leur degré d'autonomie. Plusieurs éducatrices parlaient également de l'habileté des enfants à reconnaître les lettres de l'alphabet, et à reconnaître leur nom et celui de leurs amis. Les enfants commençaient aussi à écrire les noms. Les éducatrices ont été surprises de la facilité avec laquelle les enfants avaient maîtrisé cette habileté. La connaissance des lettres de l'alphabet figure d'ailleurs parmi les meilleurs prédicteurs de la littératie chez les enfants (Lonigan et al., 2008). Il est donc encourageant que les éducatrices aient perçu ce changement chez les enfants ayant participé au programme. En lien avec la littératie, certaines éducatrices ont souligné que les enfants exprimaient un plus grand amour pour les livres et pour la lecture grâce à leur exposition régulière dans le cadre du programme.

« J'ai été surprise de ça, parce que je ne pensais pas que ça pouvait être instantané à ce point-là (...) en un mois un enfant était capable de reconnaître tous les noms de son groupe. (...) J'ai trouvé cela assez impressionnant, je ne pensais pas que ça pouvait être aussi efficace. » (Éducatrice)

Troisièmement, dans plus de la moitié des sites, les éducatrices étaient d'avis que le programme avait permis aux enfants d'améliorer leur niveau de compréhension du français et d'enrichir leur vocabulaire. L'effet était particulièrement marqué pour les petits anglophones. La formatrice a aussi indiqué que les stratégies de francisation du programme ont eu un impact considérable sur le développement du français chez les enfants. Enfin, certaines éducatrices ont soulevé d'autres impacts positifs chez les enfants, notamment, une amélioration de leur comportement, socialisation et bien-être.

Retombées perçues pour les éducatrices

D'abord, pour les éducatrices dans tous les sites, le programme et les formations de suivi ont permis de rafraîchir leurs connaissances sur le développement des enfants. Quelques-unes ont aussi indiqué que les formations leur avaient permis de se sentir plus rassurées quant à leurs compétences.

« Des fois on doute de nos compétences d'éducatrice parce qu'on ne sait jamais la réaction des parents. On a de la difficulté de se situer par rapport aux lois et règlements, par rapport à plein de choses, mais d'entendre quelqu'un nous dire que, qu'est-ce qu'on fait c'est adéquat, et qu'on est sur la bonne route, cela aide beaucoup à évoluer, à se donner de nouveaux défis, à continuer notre métier qui n'est pas toujours facile. » (Éducatrice)

Deuxièmement, les éducatrices dans tous les sites ont déclaré que leur participation au programme les avait amenées à être plus centrées sur les besoins des enfants. Cette approche se manifestait en étant plus à l'écoute des enfants et en créant des activités en lien avec leurs intérêts.

« C'est sûr qu'on se concentre sur l'enfant. On va essayer d'écouter l'enfant, ces conversations, c'est de là qu'on va partir. (...) Au lieu de faire un programme puis d'aller chercher le thème, on va aller chercher le thème dans la conversation de l'enfant. Parce que c'est quelque chose qui l'intéresse, alors pourquoi ne pas essayer de l'inclure dans la salle? » (Éducatrice)

Troisièmement, les éducatrices dans quatre des sept sites ont choisi d'adopter la grille thématique utilisée dans le cadre du programme en garderie. Cette grille permettait aux éducatrices de partager avec les parents les informations des activités au sein de la garderie, par exemple, le thème du mois et les activités qui vont avoir lieu.

Quatrièmement, les éducatrices dans quatre des sept sites ont indiqué qu'elles avaient appris à inciter les enfants à être plus créatifs dans leurs activités (ex., bricolages). Ce changement de pratique chez les éducatrices a aussi été observé par la formatrice. Cet impact est important puisque, selon la formatrice, le volet créativité est un élément central du programme.

« Y a des choses que je ne faisais pas avant (...) comme les laisser faire leur brico. (...) Maintenant, je les laisse faire, j'aime vraiment ça. Ça m'a ouvert les yeux à wow! Regarde qu'est-ce qu'ils peuvent faire. Ce n'est pas tout pareil, c'est ce qui est le fun sur les tableaux et ça m'a beaucoup ouvert les yeux. » (Éducatrice)

Cinquièmement, certaines éducatrices ont mentionné que le programme leur a permis de perfectionner certaines de leurs habiletés et stratégies dans leur travail. Par exemple, elles ont amélioré la routine en garderie, alterné le matériel dans les différents centres, et appuyé davantage l'autonomie des enfants.

Finalement, dans plus de la moitié des sites, les éducatrices ont indiqué que, grâce au programme, elles ont amélioré leurs relations avec les parents. Néanmoins, quelques éducatrices de la première cohorte ont concédé ne pas avoir travaillé cet aspect. Dans les faits, dans la plupart des sites, les éducatrices ont mentionné que le programme n'avait pas influencé leur relation avec les parents parce que celle-ci était déjà respectable.

5.10. SOMMAIRE

Le chapitre 5 rapporte les résultats de l'évaluation de la mise en œuvre du volet en garderie. Différents aspects du programme ont été évalués, dont la qualité structurelle (qualité générale) des milieux de garde observés, la fidélité structurelle, la fidélité du contenu éducatif, la qualité éducative, la qualité des interactions entre les éducatrices et les enfants, le dosage, les réactions des éducatrices au programme, les facilitateurs et les défis, ainsi que les retombées perçues. Le présent sommaire résume les résultats pour chacune de ces dimensions.

D'abord, les notes d'observations des coordonnatrices communautaires ont révélé qu'il existe une très légère différence au niveau de la **qualité structurelle** entre les garderies du groupe programme et du groupe témoin, et ce, à l'avantage des garderies du groupe programme. Ces différences sont toutefois mineures et les deux groupes montrent une bonne qualité (de 5 et plus

sur une échelle de 7) pour chacune des trois dimensions évaluées, notamment, la qualité de l'espace physique, la santé et sécurité, ainsi que l'accueil et le départ. Par conséquent, nous pouvons dire qu'il n'y a pas de différenciation entre les garderies du groupe programme et celles du groupe témoin au niveau de la qualité générale. Ces résultats sont présentés en parallèle avec ceux de l'étude de Japel et al. (2005). Cette étude a évalué le niveau de qualité dans 296 garderies dans la province de Québec dans le cadre de l'Étude longitudinale du développement des enfants du Québec. Les résultats indiquent que sur chaque sous-échelle de l'ÉÉEP-R, les garderies qui participent au projet Capacité d'apprentissage offrent une meilleure qualité structurelle que les garderies québécoises.

En second lieu, les résultats de l'évaluation de la *fidélité structurelle du programme* indiquent qu'en moyenne, pendant la première année, les classes des garderies du groupe programme ont mis en place 78 % des éléments du programme en garderie, alors que l'ensemble de ces éléments se retrouvait dans moins de la moitié des classes des garderies témoins (moyenne de 40 %). Au cours de la deuxième année, le niveau d'implantation en garderie était très similaire, soit une moyenne de 73 % pour les garderies du groupe programme, comparativement à 42 % pour les garderies du groupe témoin. Dans la première année, tous les éléments du programme ont été présents et pratiqués plus souvent dans les garderies du groupe programme comparés aux garderies témoins. Plus particulièrement, les centres étaient plus souvent identifiés, les affiches montrant une image et un mot étaient plus présentes, et il y avait une routine imagée présente et utilisée durant la journée. Lors de la deuxième année du programme, tous les éléments ont été mis en place plus souvent par les classes des garderies du groupe programme, sauf pour la présence des affiches présentant une image et un mot, où les taux étaient similaires entre les deux groupes expérimentaux (70 % versus 71 %). L'ensemble de ces résultats suggère que les éléments du programme ont été majoritairement mis en place dans les classes du groupe programme, et ce, de façon assez stable.

Troisièmement, les résultats concernant la *fidélité et la qualité du contenu éducatif* du programme ont révélé que le contenu du programme a été respecté à 74 % par les classes du groupe programme au cours de la première année de la mise en œuvre et à 54 % par les classes du groupe témoin. Tous les éléments du programme ont été présents aussi souvent ou plus souvent dans les classes du groupe programme que celles du groupe témoin pendant la première année. En particulier, la lecture durant le cercle d'amis était plus fréquente, le style de lecture interactive était plus souvent observé, et il y avait plus d'emphase sur la pré-écriture. Au cours de la deuxième année, le niveau d'application a connu une hausse pour les deux groupes. Pour les garderies du groupe programme, la moyenne a augmenté à 83 %, alors que pour les garderies témoins, la moyenne s'élevait à 63 %. C'est donc dire que les deux milieux partageaient certains points en commun, mais que le groupe programme conservait certaines particularités. Lors de la deuxième année du programme, les garderies du groupe programme ont mis en place tous les éléments aussi souvent ou plus souvent que les garderies du groupe témoin sauf un (c.-à-d., le déroulement efficace des transitions). Quant à la qualité du contenu éducatif, les résultats indiquent que la qualité éducative était supérieure dans les garderies du groupe programme que dans celles du groupe témoins, ainsi que dans les garderies au Québec.

Les observations nous ont aussi permis de comparer la qualité de la *relation entre l'éducatrice et les enfants* (c.-à-d., le niveau de sensibilité des éducatrices) entre les garderies du groupe programme et celles du groupe témoin. Afin de mesurer cette dimension, nous avons utilisé la sous-échelle de l'ÉÉEP-R « interactions personnel-enfant » qui mesure à quel point le

personnel répond aux enfants, est chaleureux, respecte les enfants, et aime être avec eux. L'échelle a révélé que les éducatrices du groupe programme et celles du groupe témoin étaient très sensibles aux besoins des enfants. La moyenne du groupe programme était de 6,22 sur 7 (É.-T. = 1,63), alors que celle du groupe témoin était de 5,26 sur 7 (É.-T. = 2,31). Ces observations montrent que les garderies des deux groupes expérimentaux offrent une bonne qualité d'interaction entre les éducatrices et les enfants.

Par ailleurs, les observations sur le *dosage* montrent qu'en moyenne, les enfants des deux groupes expérimentaux ont fréquenté la garderie pendant un nombre similaire d'heures par semaine au cours des deux années. Spécifiquement, durant la première année, les enfants du groupe programme ont fréquenté la garderie en moyenne 27,8 heures par semaine, alors que les enfants du groupe témoin l'ont fréquentée en moyenne 29,8 heures par semaine. Durant la deuxième année du programme, les enfants du groupe programme ont fréquenté la garderie en moyenne 18,0 heures par semaine, tandis que les enfants du groupe témoin l'ont fréquentée en moyenne 19,0 heures par semaine. Les moyennes du dosage durant la deuxième année du programme sont plus faibles, car durant cette période, plusieurs des enfants fréquentaient la garderie seulement à temps partiel.

Ensuite, les *réactions des éducatrices au programme en garderie* ont été recueillies par le biais d'entrevues avec 20 éducatrices ciblées du groupe programme. Deux éléments du programme ont suscité des réactions positives chez les éducatrices. D'abord, les éducatrices ont tiré profit du contenu du cercle d'amis, qui à leur avis, répondait bien aux besoins des enfants en milieu minoritaire. Deuxièmement, elles ont tiré profit de la routine imagée qui présentait aux enfants l'ordre prévu à l'horaire des différents moments de la journée. Par ailleurs, l'aspect lié à la créativité a suscité des réactions mixtes chez les éducatrices. Alors que cet élément était apprécié des éducatrices, certaines ont déclaré avoir de la difficulté à le mettre en place. Enfin, l'élément du programme le moins apprécié des éducatrices était la nécessité de toujours parler en français, même aux enfants anglophones.

Afin d'évaluer les *facilitateurs et les défis*, des entrevues ont été effectuées auprès des éducatrices et de la formatrice afin de recueillir leurs impressions sur le déroulement du programme en garderie. Les informations ont aussi parfois été recueillies par le biais des observations et formations effectuées par la formatrice dans les garderies du groupe programme. Les éducatrices et la formatrice ont identifié plusieurs facilitateurs de la mise en œuvre du programme testé. Les principaux facilitateurs mentionnés par les éducatrices et la formatrice incluent le soutien du programme par la direction, le travail d'équipe entre éducatrices, le matériel fourni aux garderies, la flexibilité du programme, les suivis de la formatrice et le soutien des coordonnatrices communautaires. À l'inverse, l'absence de ces facteurs représentait un obstacle à la mise en œuvre du programme. Quelques facilitateurs supplémentaires ont été mentionnés par la formatrice, notamment, l'ouverture des éducatrices aux rétroactions et aux nouvelles idées, ainsi que leur expérience de travail auprès des enfants. Les éducatrices et la formatrice ont également soulevé trois défis, dont l'aménagement du temps relié à la prestation du programme, la difficulté d'aborder certains thèmes et le roulement des éducatrices dans les garderies du groupe programme.

Enfin, les informations recueillies auprès des éducatrices et de la formatrice ont permis d'évaluer les *retombées perçues* du programme pour les enfants et pour les éducatrices vis-à-vis de leurs pratiques en classe. Les éducatrices et la formatrice ont perçu plusieurs retombées du

programme chez les enfants. Entre autres, elles ont mentionné que grâce à certains éléments du programme (p. ex., routine imagée), les enfants se comportaient mieux, étaient plus calmes, plus autonomes et se sentaient plus sécurisés. De plus, le programme préparait davantage les enfants à la scolarisation et favorisait le développement des compétences en français et en littératie. Pour leur part, les éducatrices ont acquis des connaissances sur le développement des enfants par le biais des formations. Elles ont également modifié certaines de leurs pratiques en garderie, les principales étant de se centrer davantage sur les besoins de l'enfant, d'adopter la grille thématique et d'encourager la créativité des enfants. Enfin, pour certaines éducatrices, le programme a favorisé leurs relations avec les parents.

6. Étude de la mise en œuvre du programme d'Ateliers familles

Ce chapitre rapporte les résultats de l'étude de la mise en œuvre du volet des Ateliers familles pour les deux cohortes de l'étude. Le programme d'Ateliers familles comprend une série de 10 ateliers, chacun abordant plusieurs thèmes dont certains étaient obligatoires et d'autres facultatifs. Le sommaire des thèmes abordés dans chacun des dix ateliers se retrouve à l'annexe C. Dans le cadre de chaque atelier, l'intervenante agissait à titre de personne-ressource, transmettant l'information pour ensuite encourager les discussions dans le groupe de parents. Les parents avaient également l'occasion de partager leur vécu afin d'apprendre les uns des autres.

Le chapitre débute par une section (6.1) portant sur la **fidélité** du contenu, à savoir la proportion des thèmes composant chaque atelier qui a été abordée. La section 6.2 porte sur le **dosage**, soit le taux de participation des parents aux ateliers. La section 6.3 détaille les résultats concernant la **qualité** de l'animation des ateliers tandis que la section 6.4 présente les **réactions** des parents aux ateliers. On retrouve à la section suivante (6.5) les **facilitateurs et les obstacles** rencontrés lors de la mise en œuvre du programme. Enfin, la dernière section (6.6) porte sur les **retombées perçues** du programme sur les parents.

Les analyses pour chacune des dimensions de la mise en œuvre du programme évaluées ont utilisé deux approches. D'abord, les analyses ont été menées sur la série de 10 ateliers offerte à 15 groupes distincts de participants au programme d'Ateliers familles, soit 11 groupes de participants répartis dans six communautés pour la première cohorte et quatre groupes de participants répartis dans deux communautés pour la deuxième cohorte. Chaque « groupe » comprend des parents et leur intervenante au cours des dix ateliers, soit 11 groupes pour la première cohorte et 4 groupes pour la deuxième cohorte. Cette première série d'analyses nous permet de mieux comprendre les expériences relatives à un groupe de parents. Ensuite, les analyses ont été menées par atelier, donc en regroupant les informations de l'ensemble des 15 groupes de parents pour chacun des dix ateliers présentés. Le résultat de ces analyses nous donne un aperçu des expériences rattachées au contenu de chacun des ateliers.

Le lecteur est prié de noter les points suivants :

1. L'évaluation des Ateliers familles est principalement centrée sur le volet parent puisque ce dernier vise à outiller les parents afin qu'ils puissent mieux appuyer le développement et la préparation à la scolarisation de leur enfant.
2. Les analyses comprennent les données collectées dans les six communautés de la première cohorte et dans les deux communautés composant la deuxième cohorte, pour un total de huit sites.²⁹
3. Seuls six des 10 ateliers offerts à la *deuxième cohorte* ont fait l'objet d'observations tandis que les 10 ateliers offerts à la *première cohorte* ont tous été observés.

²⁹ Rappelons que l'étude de la mise en œuvre du programme en garderie se fonde sur les données collectées dans les deux communautés composant la deuxième cohorte et cinq des six communautés de la première cohorte. Les données de cette sixième communauté n'ont pas été considérées dans les analyses en raison du fait que la mise en œuvre de ce programme en garderie n'ait pas pris place.

Spécifiquement, pour la deuxième cohorte, des observations ont eu lieu pour les ateliers jugés comme étant les plus pertinents au programme notamment, les ateliers 1, 2, 5, 6, 9 et 10. Pour faciliter le rassemblement des données des deux cohortes, nous rapportons uniquement les notes d'observations recueillies pour ces six ateliers dans la présente étude de mise en œuvre du programme Ateliers familles.³⁰

4. Outre les notes d'observations, l'évaluation de la mise en œuvre du programme repose sur les trois sources suivantes: (1) l'évaluation des ateliers par les parents au moyen de sondages ponctuels répondus à la fin des neuf premiers ateliers, d'un sondage logistique répondu à l'atelier 10 et d'un sondage post-intervention complété deux semaines après le dernier atelier; (2) l'évaluation des intervenantes ayant animé les ateliers, soit par le biais d'entrevues ou de rétroactions notées dans un journal de bord; et (3) l'assiduité des parents aux ateliers.
5. Les résultats par cohorte ne sont pas rapportés dans le texte. De même, aucune comparaison directe du degré de mise en œuvre entre les deux cohortes n'est rapportée dans le texte. Cette décision a été prise en raison du fait que le nombre d'intervenantes impliquées dans la prestation des Ateliers familles auprès de la deuxième cohorte était inférieur à 5, et par conséquent, la présentation des résultats par cohorte ne permettrait pas d'assurer l'anonymat de ces intervenantes.

6.1. LA FIDÉLITÉ DU CONTENU ABORDÉ DANS LES ATELIERS FAMILLES

La fidélité réfère à l'intégrité du programme appliqué : c'est-à-dire la correspondance entre l'intervention appliquée et le programme planifié. Dans le cadre des Ateliers familles, le programme prévoyait certains thèmes qui devaient obligatoirement être abordés par les intervenantes, en plus de suggérer des thèmes optionnels. Les notes d'observations nous renseignent sur la couverture des thèmes obligatoires abordés dans chaque atelier. Ces informations sont complétées par l'analyse des feuilles mobiles (flip chart) utilisées par l'intervenante durant les ateliers, de sorte qu'un thème est rapporté comme non abordé s'il ne figure ni dans les observations, ni sur les feuilles mobiles.

Afin de comparer la fidélité d'une série d'ateliers à l'autre, le pourcentage des thèmes obligatoires abordés dans les six ateliers observés a été calculé en fonction des groupes de parents. Les résultats indiquent une excellente couverture des thèmes dans les 15 groupes de parents (couverture moyenne de 98 %). On observe au tableau 6.1 qu'au moins 92 % de la matière obligatoire a été abordée dans tous les groupes de parents.

³⁰ Cette décision a été prise en raison du taux très élevé de fidélité obtenue dans les 11 séries d'ateliers de la première cohorte (couverture moyenne de 96 % pour toutes les séries d'ateliers; *Capacité d'apprentissage dans les communautés francophones en situation minoritaire : Rapport de mise en œuvre du projet* (Bérubé et al., 2014). Le terme « série d'ateliers » dans le *Rapport de mise en œuvre du projet* fait référence aux différents « groupes de parents » ayant participé aux ateliers dans le présent rapport.

Tableau 6.1 : Contenu abordé selon le groupe de parents

Groupe de parents	% du contenu abordé
1	92 %
2	93 %
3	97 %
4	97 %
5	100 %
6	93 %
7	100 %
8	100 %
9	100 %
10	96 %
11	100 %
12	97 %
13	97 %
14	100 %
15	100 %
Moyenne pour les deux cohortes	98 %

Le degré de fidélité par atelier a été calculé à partir du contenu couvert dans chacun des six ateliers observés, tous les groupes de parents confondus. On note au tableau 6.2 qu'entre 87 % à 100 % des thèmes obligatoires ont été abordés dans chaque atelier pour une moyenne de 98 %, indiquant une excellente couverture. C'est l'atelier 9 qui se distingue des autres ateliers par un degré de couverture des thèmes obligatoires de 87 %. Selon les notes d'observations, la couverture de la matière dans cet atelier a été complète dans six des 15 groupes de parents tandis que pour huit autres groupes, l'activité obligatoire « C'est ma communauté » n'a pas été faite. Il faut dire que l'ensemble des ateliers avait présenté aux parents plusieurs opportunités de discuter des services offerts dans leur communauté (soit le thème de l'activité omise) et les intervenantes avaient apporté aux parents à diverses occasions des informations et des brochures sur les services offerts dans la communauté. Pour cette raison, les intervenantes ont peut-être jugé qu'il n'était pas pertinent de discuter à nouveau de ce sujet.

En somme, le contenu des ateliers observés pour les deux cohortes a été très bien couvert et le seul thème qui a été moins bien abordé avait été présenté dans le cadre d'autres ateliers. La présentation de ce thème pourrait donc être planifiée différemment si le programme des Ateliers familles devait être donné à nouveau.

Tableau 6.2 : Contenu abordé selon l'atelier observé

Numéro de l'atelier	% du contenu abordé pour les deux cohortes
1	100 %
2	98 %
5	100 %
6	100 %
9	87 %
10	100 %
Total	98 %

6.2. LA QUALITÉ DE L'ANIMATION DES ATELIERS FAMILLES

« Le personnel nous ont très bien reçus, c'est-à-dire qu'elles nous ont mis à l'aise dès notre arrivée et nous ont guidés à travers la soirée. » (Parent)

« Tout était très bien expliqué, très bon atelier et sujet très intéressant. » (Parent)

Les observations des coordonnatrices nous renseignent sur la qualité de l'animation des Ateliers familles offerts aux 15 groupes de parents. Les observatrices ont noté les aptitudes des intervenantes à animer les ateliers et les difficultés qu'elles ont rencontrées lors de leur animation. D'abord, les observatrices ont rapporté que dans la majorité des cas, les intervenantes possédaient des bonnes aptitudes pour animer les ateliers. En particulier, les intervenantes réussissaient à bien transmettre leur message et étaient perçues comme étant accueillantes. Elles encourageaient également la participation des parents sans porter de jugement sur leurs commentaires, questions et réponses, ce qui créait un environnement propice à la discussion et à l'apprentissage.

Les observatrices ont aussi noté que certaines intervenantes ont affiché quelques difficultés ou défis qui risquaient de nuire à la bonne transmission du message aux parents. Les principales difficultés notées se résumaient au style d'animation des intervenantes. Par exemple, certaines intervenantes imposaient les réponses lors des discussions, coupaient la parole aux participants, avaient de la difficulté à transmettre le message et abordaient le contenu trop rapidement. Les observations ont révélé que trois intervenantes ont éprouvé des difficultés de façon répétitive lors des ateliers. Deux d'entre elles ont, de manière récurrente, abordé le contenu de manière trop magistrale et trop rapidement. La troisième avait tendance à être désorganisée et éprouvait de la difficulté à animer en général. C'est donc dire que dans ces trois groupes, les parents n'ont pas reçu le programme des Ateliers familles avec un niveau optimal de qualité. À ce niveau, les analyses quantitatives présentées dans l'étude d'impact du présent rapport nous renseignent davantage sur les répercussions de la qualité de l'animation par l'intervenante des Ateliers familles sur les connaissances, croyances et habitudes des parents.

Un autre indice de la qualité de l'animation noté par les observatrices était la capacité de respecter le temps alloué pour donner l'atelier. En général, les intervenantes ont eu de la

difficulté à respecter le temps qui leur était accordé. La plupart avaient tendance à dépasser le temps alloué pour chaque atelier. L'animation du volet parent de l'atelier devait se faire en 55 minutes. Plusieurs des intervenantes ont pris de 60 à 80 minutes pour transmettre le contenu aux parents. À cet égard, les observatrices ont noté que les ateliers contenaient trop de matière à couvrir pour le temps alloué au volet parents. Il faut également mentionner que les parents ont beaucoup aimé discuter entre eux. Les ateliers devraient donc prévoir des périodes de discussion plus longues, même si les parents sont exposés à un corpus de matériel moins important.

Le dernier indice de qualité sur lequel les observations nous renseignent a trait à l'environnement spatial dans lequel les ateliers se sont déroulés. Dans la plupart des sites, la disposition des participants favorisait la discussion entre eux. Il n'y a eu qu'un seul site où, à deux reprises, deux groupes de parents ont été regroupés pour ne former qu'un seul groupe animé par les deux intervenantes. À ces deux occasions, les parents ont dû être placés en deux rangées, ce qui a rendu difficile la discussion entre les participants.

La qualité de l'animation a également été évaluée par les parents participant aux ateliers par le biais des sondages ponctuels remplis à la fin de chaque atelier. D'abord, les résultats indiquent qu'à l'exception d'un groupe de parents, les participants ont jugé l'intervenante comme étant excellente. De plus, les résultats révèlent que, pour plusieurs groupes de parents, l'intervenante comptait parmi les points forts des ateliers. Plus précisément, ces parents ont mentionné que leur intervenante était dynamique et présentait les informations de façon intéressante. La plupart des groupes de parents rapportent avoir particulièrement tiré profit des exemples concrets, trucs et anecdotes que l'intervenante partageait avec eux. Ils ont également rapporté avoir bénéficié de la riche information présentée lors des ateliers. Certains parents ont aussi mentionné avoir apprécié l'approche interactive et informelle que l'intervenante adoptait pour animer les ateliers. Enfin, tous les groupes de parents ont mentionné que les ateliers auraient été plus agréables s'ils avaient été plus espacés, par exemple, toutes les deux semaines plutôt qu'à chaque semaine.

Les résultats du sondage logistique, répondu au début du 10^e atelier, nous renseignent davantage sur les aptitudes des intervenantes. Plus précisément, la majorité des parents ont jugé l'intervenante comme étant excellente à plusieurs niveaux, soit, leur connaissance des sujets traités, leur préparation, leur organisation générale, leur style et la prestation des ateliers, leur sensibilité aux besoins du groupe, ainsi que de leur habileté à créer un environnement propice aux discussions et à l'apprentissage.

En somme, la qualité de l'animation a été jugée favorablement par les observatrices et les parents. En effet, la plupart des intervenantes semblaient disposées à animer les Ateliers familles de façon appropriée et elles étaient considérées comme étant engageantes et respectueuses. Toutefois, une pratique répandue chez les intervenantes était de dépasser le temps alloué pour chaque atelier, ce qui en retour avait un effet néfaste sur la qualité de l'animation. Cette lacune au niveau de la gestion de temps peut être expliquée par la quantité considérable de contenu à aborder dans chacun des ateliers du volet parents. Ajoutons les difficultés éprouvées par quelques intervenantes au niveau de l'animation des ateliers. Enfin, sauf pour deux occasions, l'environnement spatial du volet parent des Ateliers familles était propice à la participation.

6.3. DOSAGE

Taux de participation pour chacun des sites des deux cohortes

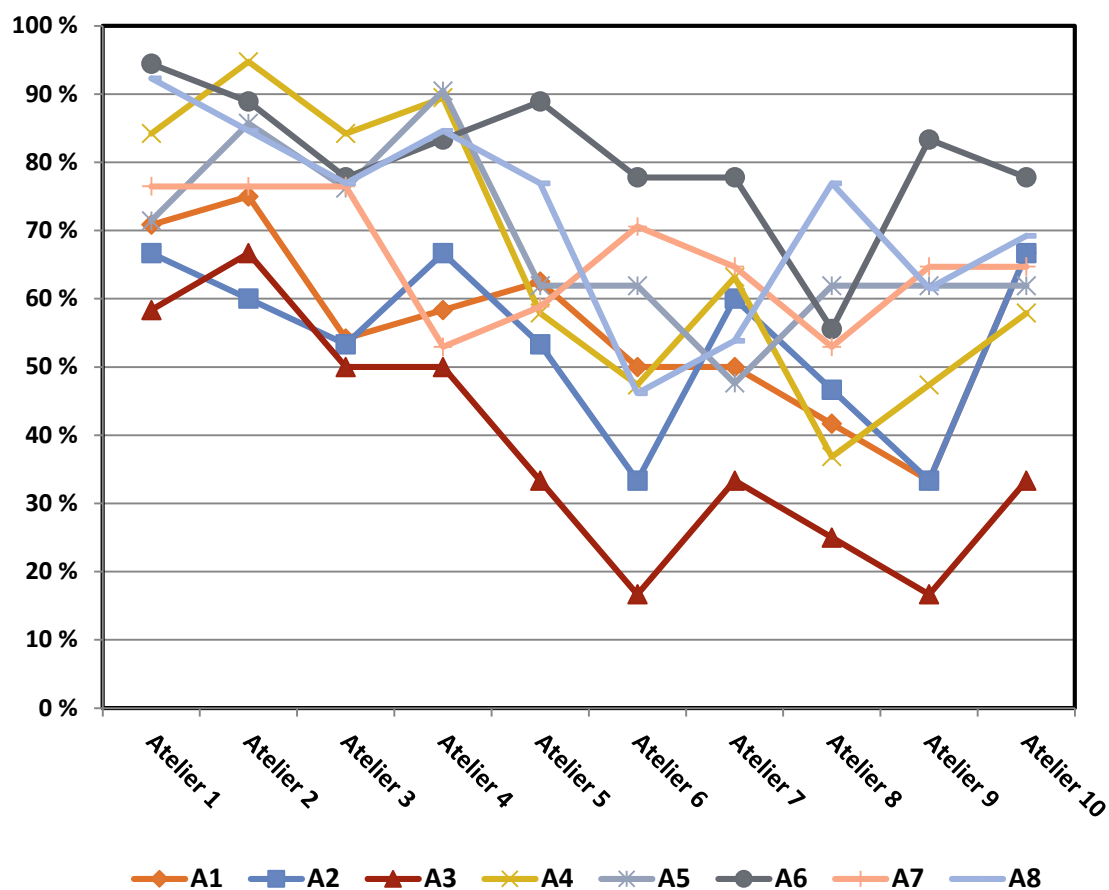
Le taux de participation moyen pour les participants des deux cohortes s'élève à 63.5 % et varie de 38 % à 81 %. Par souci de clarté, les résultats sont rapportés dans la figure 6.1 au niveau du site et non au niveau des groupes de parents. Pour les besoins de cet exposé, les communautés de la deuxième cohorte sont comptées avec les communautés de la première cohorte si bien que nous considérons avoir au total 8 sites participants (les six sites de la première cohorte + les deux sites de la deuxième cohorte)

Le premier site (A1) a connu un taux de participation aux ateliers variant entre 33 % et 75 %, avec une moyenne de 56 %. On observe également que deux des 23 familles (9 %) ont participé à deux ateliers ou moins. Dans le deuxième site (A2), le taux de participation aux ateliers a varié entre 33 % et 67 %, avec une moyenne de 54 %. Dans ce site, quatre des 15 familles (27 %) ont participé à deux ateliers ou moins, et de ce nombre, on compte trois familles n'ayant participé à aucun des ateliers. Dans le troisième site (A3), les familles participaient en moins grand nombre aux ateliers que dans les autres sites ayant un nombre comparable de parents inscrits au groupe programme. Un total de trois intervenantes auprès des parents a été requis pour assurer la prestation du programme Ateliers familles dans ce site. Cette situation était due à un roulement de personnel important. Le taux de participation aux ateliers de ce site a varié entre 17 % et 67 %, avec une moyenne de 38 %, ce qui est considérablement plus faible que dans les autres sites où les ateliers ont été fréquentés par au moins la moitié des parents. De plus, quatre des 12 familles (33 %) ont assisté à deux ateliers ou moins.

Les parents du quatrième site (A4) ont participé en grand nombre aux Ateliers familles. En effet, 19 des 20 familles inscrites au groupe programme ont participé aux ateliers. Seule une famille (5 %) n'a participé à aucun atelier. Le taux de participation aux ateliers a varié entre 37 % et 95 %, avec une moyenne de 66 %. Dans le cinquième site (A5), le taux de participation aux ateliers a varié entre 48 % et 90 %, avec une moyenne de 68 %. Une seule famille (5 %) a participé à deux ateliers ou moins. Enfin, le sixième site (A6) a connu le meilleur taux de participation des familles. On y observe un taux de participation aux ateliers variant entre 56 % et 94 %, avec une moyenne de 81 %. Seule une famille (6 %) a participé à deux ateliers ou moins.

Dans le septième site (A7), le taux de participation aux ateliers a varié entre 53 % et 76 %, avec une moyenne de 66 %. Dans ce site, une seule famille (6 %) n'a participé à aucun atelier (6 %) et trois autres familles ont assisté à deux ateliers ou moins. Toutefois, parmi les 18 familles de ce site, 10 d'entre elles (56%) ont participé à au moins huit ateliers. Enfin, le huitième site (A8) a connu le deuxième meilleur taux de participation des familles, avec un taux de participation variant entre 46 % et 92 % et une moyenne de 72 %. Dans ce site, toutes les familles ont participé aux ateliers et une seule famille a participé à deux ateliers ou moins. On observe également que neuf des familles sur 13 (69 %) ont participé à au moins huit des 10 ateliers.

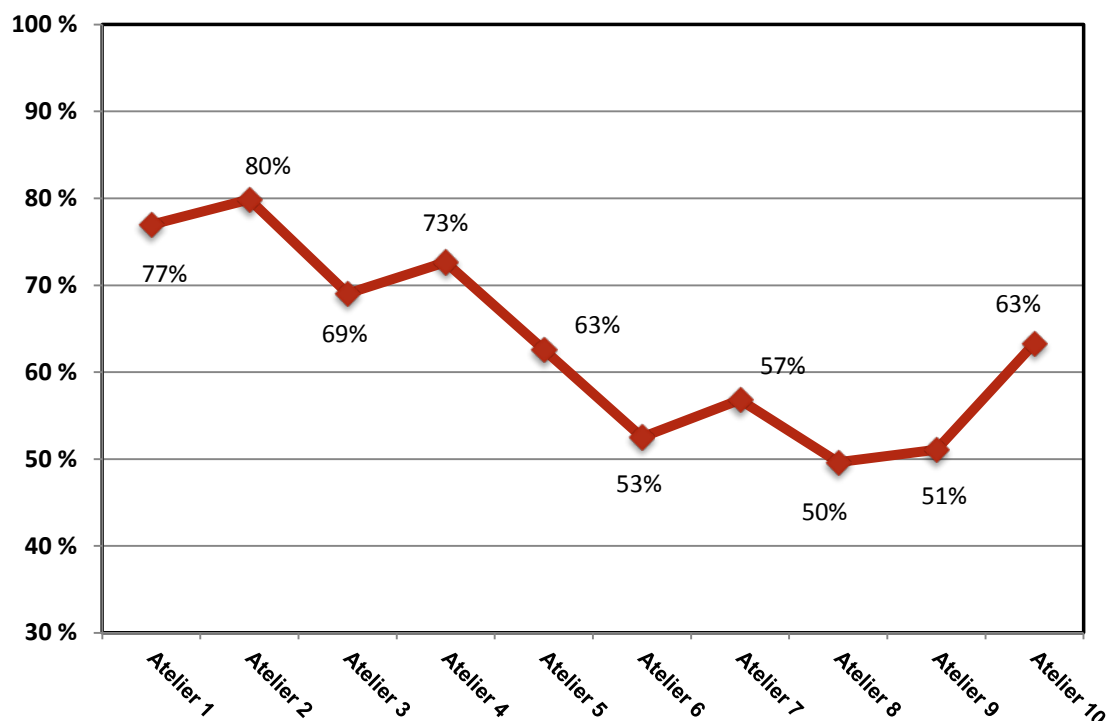
Figure 6.1 : Présence des parents selon l'atelier et le site



Taux de participation par atelier

Le taux de participation moyen par atelier était de 64 %. Selon le patron observé à la figure 6.2, le taux de participation était le plus élevé lors des quatre premiers ateliers (taux moyen de 75 %) comparativement aux six derniers ateliers (taux moyen de 56 %). Tel qu'anticipé, les taux de participation les plus élevés se retrouvent aux deux premiers ateliers, moment où les participants considèrent la pertinence des ateliers pour eux. Les ateliers 6, 8 et 9 se démarquent des autres par un faible taux de participation variant de 50 % à 53 %.

Figure 6.2 : Taux de participation des parents selon l'atelier



6.4. RÉACTIONS DES PARENTS LORS DES ATELIERS

« Mon enfant a aimé le temps qu'on a passé ensemble à faire une activité et manger avec ses amis. » (Parent)

« Comme d'habitude, j'ai adoré les discussions avec les autres parents sur divers sujets. » (Parent)

« J'ai passé de belles expériences. On a eu des parents fantastiques. Les groupes, on parlait en même temps, on s'entendait tellement bien... » (Intervenante)

Les réactions des parents lors des ateliers ont d'abord été captées au moyen des notes d'observation. La plupart des thèmes des ateliers ont suscité des réactions positives chez les parents. Ainsi, les observatrices ont décrit les parents comme étant soit participatifs, attentifs, détendus ou enthousiastes dans la plupart des ateliers. Parmi les six ateliers observés, les ateliers 1, 2 et 5 ont suscité le plus de réactions chez les parents. Les thèmes abordés dans ces ateliers ont provoqué à la fois des réactions négatives chez la majorité des parents (incompréhension, réticence, ou désaccord) et des réactions positives (niveau élevé de participation et d'intérêt). Notons que les deux premiers ateliers servaient à présenter le projet Capacité d'apprentissage, les modalités des Ateliers familles et la notion du parent comme premier éducateur de son enfant, tandis que l'atelier 5 portait sur le développement de l'autodiscipline chez l'enfant. En ce qui concerne les deux premiers ateliers, les facteurs de nouveauté et de gêne pourraient en partie expliquer la réaction des parents. Pour sa part, l'atelier 5 était plus susceptible de provoquer des réactions fortes puisqu'il s'agissait d'un sujet

pertinent pour les participants, mais également central à l'identité parentale. Enfin, les observatrices ont rapporté que les participants avaient bien profité de l'atelier 9, qui lui, portait sur la vie au sein de la communauté francophone. Cet atelier a suscité des réactions positives considérables chez la majorité des parents avec très peu de réactions négatives.

Les réactions des parents ont aussi été captées par l'entremise du sondage ponctuel à la fin de chaque atelier. Ce sondage comportait des questions permettant de connaître l'opinion des parents quant à chaque atelier auquel il venait de participer. Selon les résultats de ces sondages ponctuels, les parents ont en général apprécié le contenu des ateliers. Trois thèmes préférés ressortent des commentaires de plus de la moitié des groupes de parents. D'abord, les parents ont particulièrement aimé le thème sur la communication parent-enfant qui abordait les stratégies efficaces pour communiquer un message à l'enfant et l'importance d'avoir une écoute active lorsqu'on parle avec l'enfant (un thème de l'atelier 7). Ensuite, les parents ont aimé apprendre sur les intelligences multiples, une approche permettant de tenir compte de tous les aspects du potentiel cognitif unique à l'enfant (un thème de l'atelier 2). Enfin, les parents ont apprécié le thème sur l'éveil à l'écrit qui incluait des stratégies sur le développement des compétences en littératie (un thème de l'atelier 3).

Les parents devaient également répondre au sondage logistique administré au début du dernier atelier. Spécifiquement, les parents étaient invités à répondre à quelques questions à choix multiples sur l'appréciation des ateliers et à exprimer ouvertement leurs opinions sur ce qu'ils avaient apprécié durant l'atelier, ce qu'ils avaient appris et ce qu'ils auraient aimé apprendre. Les résultats des questions fermées révèlent que la majorité des parents ont évalué favorablement le contenu et le déroulement des ateliers. Ils ont jugé les buts des ateliers comme étant souvent clairs (90 %) et l'organisation de chacun des ateliers comme étant bonne ou excellente (94 %). Concernant le contenu, les résultats révèlent que les thèmes discutés durant les ateliers étaient considérés par les parents comme étant souvent intéressants (84 %), parfois ou souvent utiles (92 %) et souvent pertinents (70 %). Les réponses aux questions ouvertes du sondage indiquent que tous les groupes de parents ont apprécié les ateliers 2, 5 et 7. Les thèmes les plus aimés de l'atelier 2 portaient sur la notion du parent en tant que premier éducateur de l'enfant et celui des intelligences multiples. Quant à l'atelier 5, les parents ont particulièrement aimé le thème sur le développement de l'autonomie de l'enfant. Enfin, à l'atelier 7, les parents ont profité du contenu sur la communication.

Les résultats du sondage logistique nous renseignent non seulement sur l'évaluation du contenu proprement dit des ateliers, mais nous donnent également un aperçu plus général des rencontres. On note par exemple que la majorité des parents étaient satisfaits ou très satisfaits de leur expérience des Ateliers familles (91 %). La plupart ont également rapporté que leur enfant avait assez ou beaucoup aimé participer aux Ateliers familles (95 %). Ce résultat est également reflété dans les sondages ponctuels, où de nombreux parents ont mentionné que leur enfant avait hâte d'assister aux ateliers. Les parents ont particulièrement aimé discuter entre eux. Il s'agit de l'élément mentionné par le plus grand nombre de parents, autant dans les sondages ponctuels que dans le sondage logistique. Le deuxième élément le plus mentionné dans les deux sondages était l'activité avec l'enfant dans le cadre du volet conjoint. En effet, les parents ont beaucoup aimé ce moment privilégié avec leur enfant et plusieurs d'entre eux ont mentionné qu'ils auraient apprécié avoir plus de temps dédié au volet conjoint. Enfin, plusieurs parents ont rapporté que l'accès aux ressources en français représentait un atout concret du programme d'Ateliers familles.

Pour leur part, la plupart des intervenantes ont indiqué, par le biais d'entrevues et des rétroactions dans leur journal de bord, que les parents semblaient intéressés aux ateliers de façon générale et y participaient activement. Selon ces dernières, l'atelier 2 (« Je suis le premier éducateur de mon enfant ») et l'atelier 6 (« Le développement langagier, culturel, et identitaire de mon enfant ») ressortent comme étant ceux qui ont généré le plus d'intérêt et de participation des parents. De plus, certaines des intervenantes ont mentionné que les parents avaient apprécié les ressources en français qui leur avaient été offertes dans le cadre des ateliers.

En somme, les informations collectées auprès de différentes sources suggèrent que, dans l'ensemble, les ateliers ont suscité des réactions positives des parents. Plusieurs thèmes préférés ont été identifiés parmi ceux abordés dans les ateliers 2, 3, 5, 6, 7 et 9. Notons que l'atelier 2 a été mentionné par le plus grand nombre de sources, soit par les observatrices, les parents et les intervenantes. Il s'agit de l'atelier portant sur le parent en tant que premier éducateur de son enfant, y compris les sous-thèmes sur l'étayage, les styles de parentage et les intelligences multiples. Les ateliers 3, 5 et 7 semblent également compter des thèmes préférés puisqu'ils ont été mentionnés par plus d'une source. Enfin, les trois autres éléments qui ressortent des analyses sont l'appréciation des discussions entre parents, l'activité avec l'enfant dans le cadre du volet conjoint et les ressources en français.

6.5. FACILITATEURS ET OBSTACLES

La programmation des Ateliers familles a été conçue en fonction de certaines pratiques exemplaires déjà identifiées dans des études antérieures portant sur la littératie familiale (LeTouzé, 2005, 2006, 2007; Phillips, et al., 2006; Raikes, Summers, et Roggman, 2005). Par exemple, les études montrent que l'assiduité aux ateliers est favorisée lorsque la programmation inclut la présence d'enfants, l'horaire proposé convient le mieux aux participants, les intervenantes sont qualifiées, la présentation de thèmes est adaptée à la clientèle et l'accès à des trousseaux éducatifs est assuré. Les modèles d'intervention en alphabétisation familiale efficaces privilégient également une approche andragogique qui permet de bâtir sur les connaissances et compétences préalables des participants.

Une revue globale de l'ensemble des données récoltées dans le cadre des Ateliers familles a permis de dégager les principaux facilitateurs et obstacles du présent programme. Ces constats proviennent des données récoltées auprès des intervenantes qui ont animé les Ateliers familles (par le biais d'entrevues ou de journaux de bord), auprès des parents (au moyen des sondages ponctuels, du sondage logistique et du sondage post-intervention), de même que des observations des coordonnatrices communautaires.

Facilitateurs

« Les enfants ont du plaisir et c'est une belle occasion de rencontrer d'autres parents. » (Parent)

« Les parents adorent le Centre de ressource que nous avons pour eux. » (Intervenante)

Dans le cadre des ateliers, quelques stratégies ont porté fruit lorsqu'il s'agissait d'attirer et de maintenir la participation des parents aux ateliers. Les résultats des analyses confirment l'appui considérable des pratiques mises en place lors de la programmation à ce niveau. Un premier

élément gagnant a été d'offrir un repas aux parents. Ainsi, pour les ateliers ayant lieu en soirée, les parents arrivaient à la garderie après leur journée de travail et un repas leur était servi. Par ailleurs, pour les ateliers offerts les samedis, un repas était servi à l'heure du midi. La période du repas permettait aux parents de passer du temps avec leur enfant, d'échanger des informations avec l'éducatrice de leur enfant qui était aussi présente, et de discuter de façon informelle avec les autres parents.

Un deuxième facilitateur a été la présence du volet conjoint. En effet, les enfants étaient très contents d'avoir un moment où ils pouvaient présenter à leurs parents les réalisations effectuées durant le volet enfant des ateliers et de faire une nouvelle activité avec eux. La présence des enfants a d'ailleurs incité certains parents à rester à la fin de la journée, puisque les enfants les attendaient avec enthousiasme pour partager le repas.

Un troisième facilitateur avait trait à la présence des éducatrices durant les ateliers. Par le fait de leur présence, les parents et les éducatrices avaient la chance de partager de l'information sur l'enfant, de discuter de son cheminement à la garderie et à la maison et d'apprendre à mieux se connaître, ce qui avait d'ailleurs comme conséquence d'améliorer leurs relations.

Un quatrième facilitateur a été la qualité des ateliers. Plusieurs parents ont commenté sur cet aspect des ateliers, autant en matière de thèmes pertinents et intéressants pour des parents ayant de jeunes enfants que d'aptitudes de l'intervenante à bien animer les ateliers. Tel que suggéré par LeTouzé (2005), la qualification des intervenantes est un élément indispensable au succès des programmes d'alphabétisation familiale. Rappelons que toutes les intervenantes ont assisté à une formation. Une des formatrices a souligné que les intervenantes en formation semblaient intéressées, motivées et compétentes. Pour leur part, plusieurs des intervenantes en formation ont affirmé avoir bénéficié d'un approfondissement global de leurs connaissances et d'avoir appris des stratégies d'animation à l'aide d'activités pratiques, d'exemples concrets et de discussions. Elles ont également rapporté avoir été très satisfaites de la formation au niveau du contenu, de la pertinence, du déroulement et de l'animation. La formation auprès des intervenantes a donc permis de favoriser la qualité des ateliers. En retour, la qualité des ateliers encourageait la participation des parents aux discussions et exercices lors du volet parent. D'ailleurs, les parents ont noté qu'ils avaient beaucoup appris des discussions et des échanges avec les autres parents présents aux ateliers.

Enfin, les parents ont grandement aimé le fait que les ateliers leur permettaient d'avoir accès à des ressources grâce aux trousseaux éducatives. De plus, après les ateliers, les enfants et les parents pouvaient emprunter une ressource en français pour la semaine, soit un livre, un CD, un DVD ou un jeu de société. Tel que noté par les intervenantes, ces ressources étaient appréciées non seulement par les parents, mais aussi par les enfants qui manifestaient beaucoup d'enthousiasme face à l'emprunt du matériel pédagogique.

Obstacles

« Considérer l'étalement des leçons aux deux semaines. » (Parent)

*« Beaucoup trop de matériel; j'ai débordé de beaucoup dans le temps. »
(Intervenante)*

Le principal obstacle à la participation des parents a été l'exigence que représentait leur participation à un atelier hebdomadaire dans un contexte chargé de conciliation travail-famille. À

cet effet, plusieurs parents ont suggéré que les ateliers soient plus espacés dans le temps. Certains parents ont indiqué qu'ils auraient préféré que les ateliers soient étalés sur une période de deux ans, leur permettant ainsi de participer plus longtemps, mais moins intensément. Toujours concernant les horaires, certains parents ont trouvé que leur enfant était trop fatigué lors des ateliers, surtout lorsque ceux-ci étaient offerts en fin de journée. Dans les faits, certains ateliers se sont étirés jusqu'à 20 h et les parents ont mentionné que leur enfant était alors trop fatigué pour profiter pleinement des ateliers. Enfin, certains parents ont souligné qu'il y avait parfois trop de contenu abordé dans les ateliers.

Par ailleurs, les entrevues auprès des intervenantes et les rétroactions qu'elles ont notées dans leur journal de bord nous informent sur les défis qu'elles ont dû surmonter et sur les choses qu'elles feraient différemment pour améliorer l'animation des Ateliers familles. Comme les parents, la plupart des intervenantes du volet parents ont affirmé qu'elles avaient manqué de temps pour couvrir tout le contenu prévu pour les ateliers, ce qui se traduisait parfois à un manque d'approfondissement de la matière. De plus, certaines d'entre elles ont mentionné qu'elles modifieraient quelques aspects du contenu des ateliers. Plus précisément, elles changeraient l'ordre des thèmes (p. ex., bilinguisme comme un des premiers ateliers), présenteraient moins de théorie et répartiraient les sujets autrement dans les différents ateliers de sorte qu'ils soient plus en lien avec les activités du volet conjoint. Pour leur part, les intervenantes du volet enfants ont mentionné que la routine et le déroulement des ateliers fonctionnaient bien. Toutefois, certaines ont rapporté qu'elles auraient profité d'une meilleure préparation avant d'offrir les ateliers, par exemple, l'apprentissage des chansons à l'avance.

6.6. RETOMBÉES PERÇUES

« J'ai utilisé les choses que nous avons vues dans les ateliers à la maison. Je compte les utiliser dans le futur. » (Parent)

« Nous sommes les modèles de nos enfants et ces derniers tenteront de les reproduire au fil des ans. Même si mes enfants vivent dans un milieu où la langue française est de moins en moins valorisée, je vais continuer à promouvoir cette langue et à faire vivre à mes enfants des activités en français. » (Parent)

« On a trouvé que c'est des parents qui font plus d'effort maintenant pour parler à leurs enfants en français, du moins, on le voyait aux ateliers. » (Intervenante)

Les sondages aux parents (c.-à-d. les sondages ponctuels, le sondage logistique et le sondage post-intervention) avaient des questions portant sur ce que ces derniers avaient appris durant les ateliers et ce qu'ils avaient mis en pratique suite aux ateliers. Plusieurs parents ont rapporté avoir acquis plusieurs nouvelles connaissances lors des ateliers, en particulier, sur la communication (atelier 7), le français à la maison (atelier 6), ainsi que sur l'autonomie et la discipline (atelier 5). Les parents ont d'ailleurs affirmé avoir mis en pratique différentes stratégies apprises lors des ateliers. Une réponse commune des sondages logistique et post-intervention des parents se rapportait à la mise en pratique des différentes stratégies apprises pour communiquer avec leur enfant. Ces pratiques ont été discutées dans le cadre de l'atelier portant sur la communication (atelier 7) dont l'une des activités consistait en une mise en situation qui avait pour but de conscientiser les parents à l'importance de se mettre au niveau de l'enfant pour s'assurer qu'il

comprenne bien le message qui lui est communiqué. Plusieurs parents ont adapté leur façon de communiquer avec leur enfant suivant cet atelier.

Une deuxième réponse commune tirée des deux sondages (logistique et post-intervention) concerne la prise de conscience rapportée par les parents de l'importance d'exposer leur enfant au français. Introduite lors de l'atelier 6, plusieurs parents ont affirmé que cette prise de conscience a augmenté l'usage du français à la maison et le nombre d'activités qu'ils faisaient en français (p. ex., lecture, films, émissions de télévision). En troisième lieu, les parents ont souligné que les ateliers les ont sensibilisés par rapport au rôle parental dans le développement de l'enfant. En effet, les parents ont rapporté avoir appris et mis en application les notions apprises sur la gestion des émotions, de l'autonomie et de la discipline chez les enfants (atelier 5). Il est à noter que selon les observatrices, les parents ont démontré non seulement des réactions immédiates positives, mais aussi des réactions négatives lors de cet atelier. Cet atelier a donc fait réagir les parents sur le coup, mais ces derniers semblent en avoir retiré des bénéfices à plus long terme.

Par ailleurs, de nombreux parents ont mentionné que les Ateliers familles les avaient incités à dévouer plus de temps en tête-à-tête avec leur enfant à travers le jeu et les activités de littératie. Ainsi, la majorité des parents ont rapporté que la participation aux Ateliers familles a eu un effet positif sur leur relation parent-enfant. Enfin, pour plusieurs parents, les ateliers sont devenus une façon de chercher des solutions concrètes aux défis que posent la gestion des enfants, la conciliation travail-famille et l'apprentissage du français en contexte minoritaire. À cet égard, les parents ont dit avoir appris beaucoup des discussions avec d'autres parents. La possibilité d'échanger avec d'autres parents représentait un des grands atouts du volet des Ateliers familles.

Les observations menées par les coordonnatrices communautaires ont permis de confirmer les retombées du programme d'Ateliers familles rapportées par les parents. Les observatrices ont noté d'importants changements chez les parents ayant participé aux ateliers. Elles ont observé que les parents étaient plus à l'écoute des enfants, communiquaient plus souvent au niveau de l'enfant, faisaient plus d'activités en français avec leur enfant et avaient adopté de nouvelles pratiques parentales reliées au développement de l'autonomie, de l'autodiscipline et de l'estime de soi.

Pour leur part, les intervenantes ont noté que les ateliers ont eu pour effet de conscientiser les parents au rôle parental dans le développement de leur enfant en leur offrant de nouvelles informations et en leur permettant de revoir et consolider des notions déjà connues. Les intervenantes ont aussi mentionné que les ateliers offraient une occasion aux parents de passer du temps privilégié avec leur enfant, et d'un même coup, d'améliorer la relation parent-enfant. Selon quelques intervenantes du volet parent, les ateliers permettaient également un échange entre familles, ce qui a eu pour effet d'augmenter le sentiment d'appartenance à la communauté. Enfin, certaines intervenantes du volet enfant ont déclaré que le contact avec les éducatrices a été bénéfique pour les parents puisqu'il a permis d'élucider le rôle de l'éducatrice et les buts du programme en garderie, notamment celui de la préparation à la scolarisation.

De toutes évidences, les Ateliers familles semblent avoir eu des effets positifs sur les parents. Selon différentes sources d'information, un des effets du programme est de conscientiser les parents par rapport à leur rôle dans le développement de leur enfant et à l'importance de l'exposition au français en milieu minoritaire francophone. Les résultats des sondages de parents et des observations suggèrent que les parents ont acquis de nouvelles connaissances, mis en

application de nouvelles pratiques parentales apprises dans le cadre des ateliers et augmenté le nombre d'activités en français qu'ils faisaient avec leur enfant. Enfin, plus d'une source mentionne l'effet positif des ateliers sur la relation parent-enfant et l'apport des échanges entre familles sur l'apprentissage.

6.7. SOMMAIRE

Le présent chapitre rapporte les résultats de la mise en œuvre du programme d'Ateliers familles pour les deux cohortes de l'étude. Ce programme comprend une série de 10 ateliers, chacun abordant plusieurs thèmes dont certains étaient obligatoires et d'autres facultatifs. On compte 15 groupes distincts de participants au programme, soit 11 groupes de participants répartis dans six communautés pour la première cohorte et quatre groupes de participants répartis dans deux communautés pour la deuxième cohorte.

L'évaluation de l'intégrité du programme des Ateliers familles s'est faite à partir de six dimensions, notamment : la fidélité du contenu, à savoir la proportion du programme abordé durant les ateliers; la qualité de l'animation des ateliers par les intervenantes; le dosage, soit le taux de participation des parents aux ateliers; la réaction des parents aux ateliers; les facilitateurs et les obstacles rencontrés lors de la mise en œuvre du programme; et les retombées du programme au niveau des parents. Les paragraphes qui suivent présentent tour à tour les résultats des analyses portant sur chacune de ces dimensions. Les résultats de la fidélité du contenu découlent des notes d'observations, alors que les résultats des autres indicateurs reposent sur plusieurs sources d'information, dont les notes d'observations, les sondages auprès des parents et les rétroactions des intervenantes obtenues par le biais d'entrevues ou d'un journal de bord.

L'évaluation de la **fidélité** des Ateliers familles révèle une couverture à 98 % du contenu obligatoire. Quant à la **qualité** de l'animation des Ateliers familles, les observatrices et les parents ont généralement jugé favorablement la qualité de l'animation durant les Ateliers familles. La plupart des intervenantes possèdent les qualités nécessaires pour bien animer les ateliers. Seul point négatif, elles avaient tendance à dépasser le temps alloué pour chaque atelier. Ensuite, les thèmes abordés durant les ateliers ont suscité des **réactions** positives chez les parents. Parmi les dix ateliers, des thèmes préférés ont été présentés dans six ateliers. De plus, les parents identifient trois autres éléments dont l'appréciation des discussions avec d'autres parents, l'activité avec leur enfant dans le cadre du volet conjoint et l'accès à des ressources en français.

Le **taux de participation** moyen s'élève à 63,5 %. La participation varie considérablement selon les familles. Certaines familles ont participé de façon régulière aux ateliers alors que d'autres y ont peu assisté. Les analyses par site montrent qu'une des communautés de la première cohorte a connu un taux de participation plus faible que les autres sites.³¹ Le taux de participation moyen à chaque atelier souligne une plus grande participation aux quatre premiers ateliers comparativement aux six derniers.

Selon les parents, les observatrices et les intervenantes, les principaux **facilitateurs** de la participation des parents sont le repas offert aux familles, l'atelier conjoint où les parents passaient du temps avec leur enfant, et la présence des éducatrices durant les ateliers. La qualité

³¹ Rappelons que les deux communautés de la seconde cohorte ont été additionnées aux six de la première cohorte pour obtenir une somme de huit sites.

des ateliers, et l'accès à des ressources ont également encouragé la participation des parents. À l'inverse, certains **obstacles** ont nui à la participation des parents. En particulier, les parents ont déclaré que l'horaire hebdomadaire des ateliers était trop exigeant dans un contexte chargé de conciliation travail-famille. Enfin, les intervenantes ont soulevé le manque de temps pour couvrir tout le contenu prévu dans chaque atelier, ce qui avait souvent comme conséquence d'allonger la séance.

Deux des principales **retombées** des Ateliers familles sont la sensibilisation des parents vis-à-vis de leur rôle dans le développement de leur enfant et de l'importance du français en milieu minoritaire francophone. Ajoutons que les parents semblent avoir acquis de nouvelles connaissances et adopté de nouvelles pratiques parentales. Enfin, les parents ont davantage utilisé le français dans diverses activités faites avec leur enfant.

7. Analyses préliminaires

Ce chapitre décrit les étapes suivies lors de la validation des données quantitatives provenant des deux cohortes du projet Capacité d'apprentissage. Ces analyses préliminaires visent deux buts principaux : l'optimisation de la base de données et l'évaluation des limites imposées par les données sur la validité externe³² et la validité interne³³ des résultats.

Les sections suivantes présentent respectivement : le processus du contrôle de la qualité (section 7.1), la gestion des valeurs manquantes (section 7.2), l'identification de variables confusionnelles (section 7.3) et l'évaluation de la représentativité de l'échantillon du projet (section 7.4). Le chapitre termine par un sommaire de l'analyse préliminaire ainsi que les conséquences qui en découlent pour l'évaluation de l'impact du programme testé (section 7.5).

Le lecteur est prié de noter les points suivants :

1. À moins d'avis contraire dans le texte, les résultats rapportés dans ce chapitre portent sur les deux cohortes et non sur une cohorte en particulier.
2. Les analyses préliminaires rapportées dans ce chapitre portent sur quatre communautés, soit Orléans (ON), Cornwall (ON), Durham (ON) et Edmundston (NB). Les communautés d'Edmonton (AL) et Saint-Jean (NB) sont exclues de l'analyse à cause de considérations méthodologiques : dans le cas d'Edmonton, le programme en garderie n'a pas été mis en place; le groupe témoin en garderie n'était pas représenté à Saint-Jean et à Edmonton. L'exclusion de ces communautés permet de rapporter un ensemble de résultats valide et plus cohérent.
3. La première cohorte de participants comprend des familles provenant des quatre communautés, tandis que la deuxième cohorte comprend exclusivement des familles d'Orléans et de Cornwall.

7.1. PROCESSUS DU CONTRÔLE DE LA QUALITÉ

La SRSA a appliqué un processus rigoureux pour effectuer le contrôle de la qualité des données faisant l'objet des analyses préliminaires et des analyses d'impact. Une description sommaire des étapes suivies est fournie à la section 4.1 du rapport *Capacité d'apprentissage dans les communautés francophones en situation minoritaire : Rapport des résultats de la première cohorte* (Thompson, Legault, Lalonde et Bérubé, 2014; titre abrégé : *Rapport des résultats de la première cohorte*). Ces précautions ont éliminé d'emblée les erreurs et optimisé la base de données conformément aux postulats de base de chaque analyse (p. ex. absence de données aberrantes).

³² Portant sur la question « Est-il raisonnable de penser que l'effet serait obtenu avec un échantillon représentatif de la population ciblée? »

³³ Portant sur la question « Dans quelle mesure pouvons-nous affirmer que les effets rapportés sont dus uniquement au programme testé? »

7.2. ANALYSE DES VALEURS MANQUANTES

La présence de valeurs manquantes dans une base de données peut compromettre la validité d'une analyse. Cette menace provient de deux sources principales. D'abord, les cas avec des valeurs manquantes sur une variable sont habituellement exclus d'une analyse, ce qui peut être problématique lorsque ceux-ci ont des caractéristiques différentes des cas non-exclus. Si la composition de l'échantillon change suite à l'exclusion d'un sous-échantillon d'enfants, ceci met en cause la validité externe des résultats. Par exemple, si les valeurs manquantes surviennent de façon disproportionnée pour les filles, les résultats obtenus pourraient ne pas être généralisables à cette population. De façon similaire, la validité interne de l'étude est remise en question si la composition des groupes à l'étude est instable sur le plan longitudinal puisque des perturbations marquées dans la composition des groupes d'une évaluation à l'autre peuvent biaiser l'estimation de l'effet du programme. Nous présentons maintenant une analyse des valeurs manquantes évaluant le potentiel pour les deux types de biais. Les stratégies utilisées pour annuler ces effets sont également identifiées.

7.2.1. Répartition des valeurs manquantes

Les valeurs manquantes dans une base de données peuvent être divisées en deux types principaux: planifiées et non planifiées. Dans le cadre du projet Capacité d'apprentissage, les valeurs manquantes planifiées sont générées par l'utilisation de règles de départ dans l'administration de l'ÉPE-AD. Pour leur part, les valeurs manquantes non planifiées sont ubiquistes, étant générées par des mécanismes agissant à deux niveaux : a) l'incapacité d'administrer un instrument de mesure lors d'une période de test donnée (manquant complètement); et b) l'échec d'une observation en raison d'un refus de la part du participant ou d'une particularité de l'administration de l'instrument de mesure. La fréquence et le traitement de ces deux types de données manquantes sont maintenant rapportés.

Valeurs manquantes planifiées

Les valeurs manquantes planifiées sont générées par l'utilisation d'une règle de départ dans l'administration de l'ÉPE-AD aux enfants de la première cohorte. La fréquence et le traitement de ces valeurs manquantes sont décrits en détail respectivement dans les sections 4.2.1 et 4.2.3 du *Rapport des résultats de la première cohorte* (2014). Il suffit ici de rappeler que la règle de départ faisait en sorte que certains enfants sautaient des items qu'ils auraient pu échouer, nécessitant l'utilisation d'un algorithme d'imputation pour estimer les valeurs manquantes.

Suite à l'expérience acquise avec la première cohorte, une stratégie différente a été adoptée pour la deuxième cohorte. Ainsi, aucune règle de départ n'a été appliquée avant la cinquième évaluation alors qu'il est devenu évident que la majorité des enfants pouvaient aisément répondre aux premiers items de chaque échelle de l'ÉPE-AD. De façon pratique, nous avons opté pour l'utilisation d'une règle de départ individualisée « sur mesure, »³⁴ basée sur la performance de

³⁴ En principe, les items de chaque échelle de l'ÉPE-AD sont ordonnés selon leur degré de difficulté. Par conséquent, la probabilité qu'un enfant obtienne le score maximum pour un item est plus forte au début qu'à la fin d'une échelle. Typiquement, un enfant âgé de plus de quatre ans réussit parfaitement les premiers x items d'une échelle pour ensuite échouer aux items plus difficiles. D'où la règle de départ générique adoptée pour la première cohorte (voir la section 4.2.1 du *Rapport des résultats de la première cohorte*, Thompson et al., 2014) et la règle individualisée appliquée aux enfants de la deuxième cohorte lors de leur cinquième évaluation. Spécifiquement, la règle de départ utilisée pour la deuxième cohorte précisait que l'administration d'une échelle s'amorcerait avec un item situé vers la fin d'une série de réussites consécutives observée lors de l'évaluation précédente [item de départ = item numéro ($x - 2$)]. Selon cette règle, un enfant qui obtient le score maximum sur les cinq premiers items

l'enfant lors de l'évaluation précédente. Cette façon de procéder faisait en sorte qu'une simple imputation de la valeur maximale des items sautés pouvait être utilisée sans perte d'information.

Valeurs manquantes non planifiées

Les valeurs manquantes non planifiées sont présentes dans les données issues des évaluations et sondages. Ces valeurs manquantes peuvent provenir d'un sondage ou d'une évaluation non-reçu(e). De plus, les données provenant des sondages sont parfois incomplètes, ce qui signifie que des données valides sont obtenues pour certaines questions et non pour d'autres, soit en raison d'une erreur commise par l'interviewer, soit parce que le participant a refusé de répondre.³⁵ Des deux types de valeurs manquantes, la seconde est possiblement plus problématique puisque le mécanisme qui l'a générée peut être une fonction des caractéristiques de la question elle-même.³⁶ Les prochains paragraphes portent sur les sources de valeurs manquantes ainsi que leurs conséquences pour la sélection d'une stratégie d'imputation.

À chaque période d'évaluation, on observe un faible pourcentage de données manquantes complètement. Celles-ci sont dues soit à l'impossibilité d'administrer le sondage, soit à l'attrition des participants (voir le Chapitre 3). Les données manquantes en raison de l'attrition des participants sont problématiques puisqu'elles limitent les options d'imputation des données dans un devis longitudinal. Par exemple, dans le cadre d'un devis longitudinal, il est possible de poser la même question à plusieurs reprises afin d'accroître la possibilité d'obtenir l'information recherchée auprès de tous les participants. Pour illustrer ceci dans le contexte du présent projet, nous avons posé des questions sur le revenu familial à trois reprises durant les deux premières années de l'étude.³⁷ Pour les participants qui sont demeurés dans le projet jusqu'au huitième sondage, il est possible d'exploiter entièrement une telle redondance en imputant les valeurs manquantes pour ces questions lorsqu'elles surviennent.³⁸

Une autre source de valeurs manquantes dans la base de données est les sondages contenant des données incomplètes. Comme nous l'avons noté ci-dessus, ce type de données manquantes peut causer des problèmes pour l'analyse des données. Le principal problème est de savoir si une quelconque propriété de la question posée augmente systématiquement la possibilité qu'un participant ne réponde pas à la question. Par exemple, il est bien connu que les répondants situés aux deux extrêmes de la fourchette de revenu sont plus susceptibles de ne pas vouloir répondre aux questions liées au revenu. On voit clairement dans cet exemple que le processus produisant les valeurs manquantes est systématique. Si le mécanisme produisant les valeurs manquantes est systématique (non aléatoire), l'exclusion de cas comportant des valeurs manquantes risque de modifier les caractéristiques de l'échantillon global de façon significative.

lors de la quatrième évaluation débiterait l'échelle au troisième item lors de la cinquième évaluation à condition que la langue d'administration du test soit demeurée constante. Advenant un changement de la langue d'administration d'une évaluation à l'autre, l'administration de l'échelle débiterait avec le premier item.

³⁵ Ce type de valeur manquante partielle n'a pas été observé pour la version française des échelles de l'ÉPE-AD. Toutes les analyses rapportées dans ce document étant fondées sur la version française de l'ÉPE-AD, la question des valeurs manquantes partiellement n'entre plus en considération en ce qui a trait à ces mesures de résultantes.

³⁶ L'exemple classique servant habituellement à illustrer ce point est le cas où l'on pose des questions sur le revenu familial.

³⁷ Cette question a été posée dans le sondage de base et dans les sondages des parents administrés lors des cinquième et septième périodes d'évaluation.

³⁸ Nous devons aussi présumer qu'il est peu probable que la variable mesurée change systématiquement avec le temps ou qu'elle soit affectée par le traitement. Nous considérons que le revenu familial est un excellent exemple d'une variable pouvant évoluer au fil du temps, mais sans vraisemblablement montrer de relation systématique avec le temps (dans l'intervalle de 1-2 ans considérée) ou avec l'assignation des participants aux groupes à l'étude.

La règle utilisée pour décider de la langue d'administration de l'ÉPE-AD exemplifie ce genre de mécanisme (détaillée à l'annexe D). Pour les versions françaises de toutes les échelles (sauf pour celle mesurant la Communication), cette règle génère des valeurs manquantes chez les enfants ayant les compétences les plus faibles en langue française puisque la version anglaise de l'outil est administrée à ces derniers. Cette attrition a pour effet de réduire l'étendue de la variable (et sa variabilité), atténuant potentiellement les impacts du programme qui dépendent de la maîtrise du français. Tous mécanismes semblables peuvent similairement biaiser les résultats. La nature du mécanisme ayant généré les données manquantes du projet est élucidée par l'analyse rapportée dans la section 7.2.2 et les stratégies d'imputation employées pour minimiser l'influence des valeurs manquantes dans les analyses d'impact sont décrites dans la section 7.2.3.

7.2.2. Patron des données manquantes

Le processus engendrant les données manquantes peut, comme tout autre résultat, être modélisé. Les conclusions tirées de cet exercice de modélisation déterminent les mesures à prendre pour préserver la validité interne et la validité externe de l'étude. Si le processus qui a engendré les données manquantes est aléatoire, la validité n'est pas menacée. Par contre, s'il n'est pas aléatoire, il faut alors prendre des mesures pour éviter d'introduire un biais en excluant des cas auto-sélectionnés. Dans les paragraphes suivants, nous décrivons la stratégie qui a servi à discerner la nature du patron des données manquantes dans la base de données.

Selon Little et Rubin (1987), on peut classer les valeurs manquantes non planifiées selon trois types: a) manquant complètement aléatoire (*missing completely at random*; MCAR); b) manquant aléatoire (*missing at random*; MAR); et c) manquant non aléatoire (*missing not at random*; MNAR non-négligeable). Le scénario le plus souhaitable est une situation où un petit nombre de valeurs manquantes est distribué aléatoirement dans l'ensemble de données (c.-à-d. MCAR). En revanche, le pire scénario est d'avoir un grand nombre de données manquantes distribuées de façon non-aléatoire. Dans le premier cas, le problème des valeurs manquantes peut être résolu par l'application d'une « *listwise deletion* » sans risque de biaiser les résultats d'une analyse (c.-à-d. suppression des cas avec valeurs manquantes). Cette solution n'est pas à conseiller dans le deuxième scénario. Traiter ainsi les valeurs manquantes distribuées de façon non-aléatoire risque de biaiser les estimations des effets du traitement. Ce biais peut être évité soit: a) en excluant la ou les variables irrégulières de l'analyse; b) en employant une stratégie d'imputation des données; ou c) en acceptant le biais dans l'analyse et en tenant compte de sa nature dans l'interprétation des résultats (p. ex. les valeurs manquantes étaient plus courantes dans la population X, par conséquent les résultats fondés sur les cas ayant les données complètes peuvent ne pas être généralisables à la population X).

La première étape dans ce processus de décision est d'évaluer la prévalence des valeurs manquantes dans la base de données (c.-à-d. le pourcentage de valeurs manquantes associé à chaque variable) et le patron de ces valeurs manquantes. La prévalence des valeurs manquantes est simplement évaluée au moyen d'analyses descriptives de base. La détermination du patron des données manquantes, quant à lui, exige une analyse plus approfondie.

La question la plus fondamentale est de savoir si les valeurs manquantes sont prévisibles ou non. Comme nous l'avons déjà vu, les valeurs manquantes imprévisibles ou « aléatoires » sont classées comme MCAR. Une condition requise pour démontrer MCAR est d'établir que la relation entre l'occurrence des valeurs manquantes et celle des valeurs observées des variables

dans la base de données est statistiquement nulle. On peut vérifier cette propriété en faisant des tests d'association indépendants pour chacune des variables dans la base de données. Dans le cas des variables continues, on peut également recourir au test MCAR de Little. Si le test MCAR ou un autre test d'association fait état de valeurs manquantes systématiquement, on peut alors rejeter l'hypothèse d'une distribution des valeurs manquantes de type MCAR et opter pour une hypothèse moins restrictive: la distribution des valeurs manquantes est en réalité de type MAR.

Pour qu'un patron de données manquantes soit considéré comme MAR, deux conditions doivent être respectées: a) les valeurs manquantes doivent être distribuées de façon non aléatoire et b) il doit être possible de prédire la valeur à attribuer à la valeur manquante. La première condition est remplie lorsque la ou les variables de la base de données permettent de prédire l'incidence de valeurs manquantes; la deuxième condition est remplie lorsque de forts prédicteurs de la variable mesurée permettent d'anticiper avec justesse la valeur qu'aurait prise l'observation manquante. Lorsque la première condition est remplie, mais non la deuxième, la distribution de données manquantes est de type MNAR. Dans ce cas, nous n'avons pas d'autres options que de décrire le patron des données manquantes de façon suffisamment détaillée pour bien comprendre les limites relatives à la validité de l'étude.

Les problématiques que pose l'évaluation des valeurs manquantes ayant été examinées, nous présentons à la section suivante les résultats de l'analyse des valeurs manquantes effectuée en prévision des analyses d'impact du projet Capacité d'apprentissage. Nous rapportons d'abord les résultats de l'analyse quantitative de l'attrition des participants. Ensuite nous présentons le résultat d'une analyse des valeurs manquantes complètement en raison d'une incapacité à évaluer un enfant ou à administrer un sondage. Nous terminons en détaillant le résultat de l'analyse des valeurs manquantes causées par la langue d'administration de l'ÉPE-AD. Pour chaque analyse, nous décrivons les valeurs manquantes sous deux aspects soient en fonction de leur répartition à travers les groupes expérimentaux et de leur relation avec d'autres covariables dans la base de données.

Valeurs manquantes dues à l'attrition

Le taux d'attrition est faible pour les deux cohortes de participants. On compte 14 familles de la première cohorte qui se sont retirées de l'étude à la fin de la deuxième année. Spécifiquement, 10 familles ont quitté l'étude à la fin de la première année soit un taux d'attrition de 3,9 % des 254 familles de l'échantillon recrutées au sein des quatre communautés retenues pour les analyses d'impact. La deuxième année, le taux d'attrition a grimpé à 5,5 % de l'échantillon des quatre communautés. Pour la deuxième cohorte, seulement quatre familles se sont retirées de l'étude lors de la première année et aucune famille lors de la deuxième année, soit un taux d'attrition de 3,9 % des 102 familles de l'échantillon d'inscription. En somme, les deux cohortes combinées constituent un échantillon d'inscription de 356 enfants avec un taux de rétention de 95 % lors des deux premières années de l'étude.

Les faibles fréquences d'attrition n'ont pas permis de mener un test formel pour déterminer si la rétention des participants était étroitement liée à l'assignation aux groupes à l'étude. Toutefois, une analyse descriptive des fréquences ne révèle pas l'existence d'une relation systématique. En tout, six familles du groupe programme, deux familles du groupe témoin en garderie et cinq familles du groupe témoin hors garderie ont quitté l'étude. De plus, le retrait de l'étude n'est pas significativement associé à aucune échelle de l'ÉPE-AD de la première évaluation. La plus

forte corrélation observée est avec le vocabulaire expressif, $r = 0,10$. Notons que ces analyses par groupe expérimental et par échelle de l'ÉPE-AD n'incluent pas quatre participants ayant quitté l'étude avant la première évaluation.

Le taux d'attrition observé dans le projet Capacité d'apprentissage se compare favorablement à ceux d'autres études d'évaluation (p. ex., Rogers, Fernandez, Thurber, et Smitley, 2004). À moins de cinq pourcent, il s'avère être une source négligeable de données manquantes (Tabachnick et Fidell, 2006, p. 63). De plus, le taux d'attrition ne semble pas être lié à l'assignation aux groupes expérimentaux. En conclusion, l'attrition ne risque pas de biaiser l'estimation de l'effet du programme testé, que ce soit par rapport à la généralisation des résultats ou à la comparabilité des groupes à l'étude.

Valeurs manquantes dues aux données manquantes complètement (évaluation ou sondage)

On observe au tableau 7.1 un faible pourcentage de données manquantes pour chacune des évaluations ou sondages administrés dans la première année de l'étude. L'incidence des données manquantes est insuffisante pour tester son association avec l'assignation aux groupes. Ajoutons que le pourcentage de données manquantes observé pour chaque groupe expérimental est comparable.

Les cas pour lesquels un sondage auprès des parents ou plus est manquant ont tendance à moins bien performer à l'évaluation initiale. Une corrélation négative est observée entre le nombre de sondages manquants et le score obtenu dans les domaines de la Communication ($r = -0,21$), la Conscience de soi ($r = -0,13$), la Cognition ($r = -0,11$), Physique ($r = -0,11$), du Vocabulaire expressif ($r = -0,14$) et du Vocabulaire réceptif ($r = -0,13$). Nous observons aussi une corrélation négative entre la probabilité d'un sondage manquant et la Langue des activités de littératie ($r = -0,19$) en ce sens que les participants utilisant plutôt l'anglais lors des activités de littératie avaient une plus grande probabilité d'un sondage manquant. De même, les cas pour lesquels une évaluation ou plus est manquante ont tendance à moins bien performer à l'évaluation initiale. Des corrélations négatives sont observées entre le nombre d'évaluations manquantes et le score obtenu aux domaines de la Communication ($r = -0,16$), du Vocabulaire expressif ($r = -0,12$) et du Vocabulaire réceptif ($r = -0,11$). Encore une fois, la Langue des activités de littératie est aussi négativement associée à la présence de données manquantes ($r = -0,20$). Une imputation des valeurs manquantes est donc jugée nécessaire afin de pallier au biais que pourrait engendrer l'exclusion des cas avec valeurs manquantes à la généralisation des résultats.

En somme, les valeurs manquantes sont peu fréquentes, distribuées équitablement entre les groupes et leurs valeurs réelles peuvent être estimées avec précision en raison du grand nombre de prédicteurs et de l'administration répétée de plusieurs instruments de mesures. Nonobstant la quantité négligeable de valeurs manquantes, le fait que certaines des résultantes à l'étude soient négativement et significativement associées avec la présence de valeurs manquantes fait en sorte qu'une simple exclusion de ces cas rehausserait artificiellement le niveau de compétence de l'échantillon, ce qui limiterait la généralisation des résultats. C'est notamment le cas pour la présente étude puisque l'impact du programme est présumé varier en fonction des compétences en langue française. Le problème posé par les valeurs manquantes doit être géré en utilisant une

stratégie plus sophistiquée : le remplacement des valeurs manquantes par imputation (voir la section 7.2.3).

Tableau 7.1 : Pourcentage de cas pour lesquels il manque au moins un sondage ou une évaluation lors de la première année de l'étude (quatre périodes d'évaluation)

	Échantillon au complet (%)	Groupe programme (%)	Groupe témoin en garderie (%)	Groupe témoins hors garderie (%)
Un sondage manquant ou plus	4,68	4,90	2,26	7,48
Une évaluation manquante ou plus	4,09	2,94	5,26	3,74

Valeurs manquantes dues à la langue d'administration (évaluation seulement)

L'incidence de valeurs manquantes dues à l'administration de l'ÉPE-AD en anglais est rapportée pour l'échantillon au complet et par groupe expérimental au tableau 7.2. D'après les tests de χ^2 , l'appartenance à un groupe expérimental est associée au taux d'évaluations manquantes seulement à la quatrième évaluation ou on observe un plus fort pourcentage d'administrations en anglais pour le groupe témoin hors garderie. Dans l'ensemble, les résultats ne donnent pas à penser que la langue d'administration risque de biaiser les comparaisons planifiées. Une imputation des valeurs manquantes est néanmoins requise car les données manquantes sont nombreuses (plus de 10 % des observations) et systématiques (c.-à-d. reliées à la maîtrise de la langue). Dans la mesure du possible, les analyses d'impact devraient représenter la pleine étendue de maîtrise de la langue présente dans la population. Le succès de la stratégie d'imputation utilisée dans le remplacement de ces valeurs manquantes est détaillé dans la section 7.2.3.

Tableau 7.2 : Enfants évalués en anglais lors des cinq premières évaluations en fonction du groupe

	Intra-groupe				Test $\chi^2(2)$
	Échantillon au complet	Groupe programme	Groupe témoin en garderie	Groupe témoin hors garderie	
	Fréquence (%)	Fréquence (%)	Fréquence (%)	Fréquence (%)	
Niveau de base	43 (12,8)	17 (17,0)	15 (11,5)	11 (10,4)	2,32
4 mois	46 (13,5)	14 (13,9)	14 (10,8)	18 (16,4)	1,61
8 mois	40 (11,8)	12 (12,5)	11 (8,5)	17 (15,2)	2,66
12 mois	26 (7,7)	6 (6,3)	5 (3,9)	15 (13,3)	7,84**
16 mois	30 (12,6)	9 (14,1)	11 (9,8)	10 (16,1)	1,61

Note : N = 342, Seuils de signification : *** < 1 %; ** 5 %; * < 10 %.

Valeurs manquantes dues aux données manquant partiellement (sondage seulement)

L'analyse des valeurs manquantes dues aux données manquant partiellement porte d'abord sur les variables confusionnelles devant être incluses dans les analyses de régression (pour la liste complète, voir la section 7.3). Pour certaines de ces variables, il existe suffisamment de redondance dans la base de données (p. ex. la question sur le revenu familial est posée trois fois au cours des trois premières années) pour effectuer, dans un premier temps, une imputation simple des valeurs manquantes.³⁹ Suite à cette première imputation, la liste de covariables est soumise à une analyse de valeurs manquantes formelle (cette section) et à une imputation (voir la section 7.2.3).

Toutes les covariables incluses dans les analyses d'impact ont moins de 1% de données manquantes sauf la Vitalité avec un taux de 14,9 %. La langue parlée par la mère à l'enfant, la langue parlée par le père à l'enfant et l'Autorité parentale ont chacune seulement 0,9 % de données manquantes. Le test MCAR de Little n'étant pas significatif, $\chi^2(46) = 51.87, p = 0,26$, nous retenons donc l'hypothèse que les valeurs manquantes de ces covariables sont distribuées aléatoirement. Concernant la Vitalité, le résultat des analyses préliminaires rapportées au tableau B.4 indique que cette dernière est associée aux groupes expérimentaux et doit par conséquent être contrôlée dans les analyses d'impact. Puisque le taux de valeurs manquantes associées à cette variable est supérieur à 5 % (Tabachnick et Fidell, 2006, p. 63), une imputation est effectuée dans le but d'inclure les cas avec des données manquantes dans les analyses d'impact.⁴⁰

7.2.3. Stratégie d'imputation des données

La stratégie servant à traiter les valeurs manquantes change en fonction de la variable traitée. Dans tous les cas, nous misons sur l'utilisation de mesures répétées dans l'étude. Comme nous l'avons noté plus haut, le fait que nous mesurons les mêmes variables (ou des variables très semblables) auprès des mêmes personnes de façon répétée au cours de l'étude nous permet d'estimer avec précision la valeur qui aurait été obtenue si la variable avait été observée. Nous décrivons dans des sections séparées ci-dessous, la stratégie d'imputation employée pour les mesures des évaluations (c.-à-d. les résultantes pour les enfants) et pour les mesures des sondages (c.-à-d. les covariables et les résultantes pour les parents).

Imputation des résultantes des enfants

Dans un rapport précédent (*Rapport des résultats de la première cohorte*, Thompson et al., 2014), nous avons décrit une stratégie d'imputation pour les échelles de l'ÉPE-AD qui tirait profit de la nature longitudinale de l'étude (voir la section 4.2.3 du *Rapport des résultats de la*

³⁹ Lorsqu'une même question/échelle est administrée à plusieurs reprises, nous pouvons combiner les différentes instances afin de créer une seule variable pour les analyses d'impact contenant le maximum d'information et le minimum de valeurs manquantes. Dans certains cas (p. ex., le revenu), une différence d'échelle de mesure a nécessité une imputation par régression pour faire le pont entre les deux instances de mesure de cette variable, créant ainsi une seule variable qui contient toute l'information disponible dans la base de données sur le concept mesuré. Une description détaillée de la stratégie adoptée pour ce type d'imputation est disponible à la section 4.2.3 du *Rapport des résultats de la première cohorte* (Thompson et al., 2014). Les chiffres exacts ne sont pas rapportés pour l'analyse combinée des deux cohortes puisqu'ils ajoutent peu de valeur à l'interprétation des résultats. Notez que contrairement au *Rapport des résultats de la première cohorte*, les caractéristiques de la mère ont comblé les données manquantes des caractéristiques du père quand il s'agissait d'une famille monoparentale.

⁴⁰ Dans le *Rapport des résultats de la première cohorte* (Thompson et al., 2014) l'imputation a été évitée en éliminant la Vitalité de la liste de covariables. Cette option n'est pas possible dans la présente analyse en raison de l'émergence de la Vitalité comme variable confusionnelle importante pour le présent échantillon.

première cohorte). L'utilisation de cette stratégie a produit une base de données imputées qui a fait l'objet des analyses d'impact du *Rapport des résultats de la première cohorte*. La même stratégie est utilisée dans le présent rapport.

Les analyses d'impact portant sur les deux cohortes combinées incluent les données de la première cohorte utilisées dans les analyses d'impact du *Rapport des résultats de la première cohorte*. L'analyse intégrale des données provenant des deux cohortes nécessite la création d'une base de données des scores ÉPE-AD équivalente pour la deuxième cohorte. En conséquence, la même stratégie d'imputation doit être utilisée pour les données de la deuxième cohorte.

Nous rapportons dans le présent rapport seulement l'imputation des données de la deuxième cohorte puisque ces informations ont déjà été rapportées pour la première cohorte (voir la section 4.2.3 du *Rapport des résultats de la première cohorte*, 2014). De plus, nous rapportons seulement l'imputation des cinq premières évaluations puisque ce sont celles affectées par la source principale de valeurs manquantes : l'administration de l'ÉPE-AD en anglais.

L'algorithme servant à imputer les scores des échelles manquantes utilise l'information du participant tirée uniquement de la version française de l'ÉPE-AD.⁴¹ D'abord, nous avons estimé un score de gain « moyen » pour chaque combinaison enfant-échelle. Le score de gain moyen d'un enfant représente le changement moyen observé chez l'enfant entre deux évaluations consécutives pour une échelle donnée.⁴² Dans l'exécution de cette stratégie, les valeurs observées servent de points d'ancrage pour la procédure d'estimation. Par exemple, les valeurs manquantes de la deuxième évaluation ont été imputées par la somme du score obtenu lors de la première évaluation et le score de gain moyen (score imputé = score précédent + score de gain). Si des valeurs manquaient dans la première évaluation, la procédure estimait alors « à rebours » en soustrayant du score observé à la troisième évaluation le score de gain moyen (score imputé = score subséquent – score de gain moyen). Cette stratégie est limitée par le fait qu'il faut au moins deux observations valides (c.-à-d. deux évaluations) pour compléter la série d'observations d'un enfant.

Les résultats de l'imputation sont rapportés par groupe expérimental pour la deuxième cohorte au tableau 7.3. Nous détaillons séparément l'échelle de Communication, qui était administrée en français à chaque enfant, et les autres échelles de l'ÉPE-AD. Notez que pour l'échelle de la Communication, le nombre de valeurs manquantes est très faible, surtout comparativement au taux observé avec les autres échelles. Peu importe, l'algorithme d'imputation a pu estimer la valeur manquante dans la majorité des cas,⁴³ ce qui nous permet de

⁴¹ L'emploi de l'information du participant fait en sorte que les attributs uniques du participant sont correctement représentés par la stratégie d'imputation. La stratégie employée n'exagère donc pas la congruence entre les scores d'un enfant donné et le reste du groupe. Par contre, il est possible que la variabilité intra-individuelle soit sous-estimée. L'envergure du biais potentiel est déterminée par la pondération des valeurs imputées.

⁴² Les scores de gain moyen ont été calculés en fonction de la différence moyenne prévue entre deux périodes d'évaluation consécutives. Si l'enfant a été testé deux fois au cours de la première et quatrième périodes d'évaluation et a obtenu des scores respectifs de 4 et 14, le score de changement moyen serait la différence observée entre les évaluations ($14 - 4 = 10$), divisée par le nombre d'intervalles ou d'étapes entre les évaluations, soit, dans le cas qui nous intéresse, trois (score de changement moyen de $10/3 = 3,33$). Toutes les différences entre les évaluations ont contribué au calcul, pourvu qu'elles se situent à l'intérieur de trois étapes l'une de l'autre. De plus, aucune imputation n'était effectuée si les deux scores observés pour l'enfant sur une échelle donnée provenaient de la première et cinquième évaluation (c.-à-d. quatre étapes).

⁴³ D'autres stratégies d'imputation reposent sur l'hypothèse que les paramètres de l'échantillon global permettraient d'estimer adéquatement la valeur manquante d'un individu. Puisque la majorité des valeurs manquantes appartiennent à des enfants qui sont atypiques (c.-à-d. ils n'ont pas pu compléter le test en français à au moins une occasion), cette hypothèse serait infirmée a priori. La stratégie adoptée dans les présentes analyses d'impact utilise uniquement des informations intra-individuelles.

conserver dans les analyses d'impact plusieurs enfants ayant une maîtrise plus faible du français. La justesse de cette méthode d'imputation est attestée par le résultat d'une analyse visant à valider la stratégie d'imputation des données manquantes qui est rapportée à l'annexe E.

Tableau 7.3 : Distribution des valeurs manquantes et des imputations de l'ÉPE-AD selon les groupes expérimentaux (2^e Cohorte)

	<i>Périodes d'évaluation ÉPE-AD</i>				
	1^{ère}	2^e	3^e	4^e	5^e
Échelle de Communication	# imputé / n (non-imputé)	# imputé / n (non-imputé)	# imputé / n (non-imputé)	# imputé / n (non-imputé)	# imputé / n (non-imputé)
Groupe des garderies programme	0/31	0/31 (1)	0/31 (1)	1/31 (1)	0/31 (1)
Groupe des garderies comparaison	0/37	0/37	0/37	0/37	0/37
Groupe de comparaison hors garderie	0/33	0/33	1/33	2/33	2/33
Total	0/101	0/101	1/101 (1)	3/101 (1)	2/101 (1)
Total (% imputé)	0,0%	0,0%	1,0% / 1,0%	3,0% / 1,0%	2,0% / 1,0%
Autres échelles de l'ÉPE-AD	# imputé / n (non-imputé)	# imputé / n (non-imputé)	# imputé / n (non-imputé)	# imputé / n (non-imputé)	# imputé / n (non-imputé)
Groupe des garderies programme	5/31 (2)	2/31 (3)	1/31 (3)	1/31 (3)	0/31 (1)
Groupe des garderies comparaison	4/37 (1)	3/37 (1)	2/37 (1)	0/37 (1)	0/37
Groupe de comparaison hors garderie	3/33 (1)	5/33 (2)	3/33 (2)	3/33 (2)	3/33 (2)
Total	12/101 (4)	10/101 (6)	9/101 (6)	4/101 (6)	3/101 (3)
Total (% imputé / % non-imputé manquant)	10,9% / 4,0%	9,9% / 5,9%	8,9% / 5,9%	4,0% / 5,9%	3,0% / 3,0%

Note : Les chiffres exacts pour la catégorie « Autres échelles de l'ÉPE-AD » proviennent de l'échelle Conscience de soi. Les résultats reflètent approximativement ce qui est observé pour les autres échelles (à 1-2 près).

Imputation des données de sondage

Seulement quatre des covariables retenues dans les analyses d'impact ont des données manquantes : la Vitalité, la langue parlée par la mère à l'enfant, la langue parlée par le père à l'enfant et l'Autorité parentale. La valeur que devaient prendre ces valeurs manquantes est estimée au moyen d'une régression multiple, une méthode qui ne réduit pas de façon artificielle la variabilité.⁴⁴ Le modèle de régression est construit à partir des variables continues provenant des sondages. Cette analyse inclut toutes les covariables (non redondantes) disponibles dans la base de données et non uniquement la liste réduite de variables retenues dans les analyses d'impact. Cependant, les résultantes provenant de l'ÉPE-AD n'ont pas été utilisées à cette fin.

⁴⁴ Les résidus d'un cas échantillonné aléatoirement sont ajoutés à la valeur estimée par le modèle de régression dans le but d'imiter la variabilité « aléatoire » typique d'une observation réelle.

7.3. VARIABLES CONFUSIONNELLES

Cette section vise à résumer la pertinence pour les analyses d'impact des variables mesurées dans le cadre du projet Capacité d'apprentissage. Cet exercice permet de dresser une liste de variables à retenir comme covariables dans les analyses d'impact. Les critères de sélection se résument comment suit : une variable est jugée utile comme covariable si elle est significativement associée aux résultantes et si elle varie significativement en fonction des groupes expérimentaux. L'importance de ces deux critères est facile à comprendre. Pour influencer les résultats (p. ex., augmenter la puissance statistique; corriger un biais dans la composition des groupes), une variable doit être significativement associée aux résultantes. Pour qu'une variable puisse biaiser l'estimation de l'effet du programme, la variable doit être significativement associée aux groupes expérimentaux. Dans l'absence d'une de ces conditions, le contrôle statistique ou non de la variable lors des analyses d'impact ne changera pas dramatiquement l'estimation de l'impact du programme. En particulier, le contrôle des biais liés à la composition des groupes est essentiel pour maximiser la validité interne des comparaisons d'intérêts dans le cadre d'une étude quasi-expérimentale.⁴⁵

En conséquence, nous avons privilégié les variables qui sont à la fois associées aux résultantes à l'étude et à l'appartenance aux groupes expérimentaux. Il est à noter que les variables redondantes et les variables mesurées uniquement au sein du groupe programme (p. ex. ceux qui proviennent du sondage post-programme) ont été exclues de ces analyses préliminaires. Par contre, nous avons examiné certaines variables particulières au mode de garde en garderie (p. ex., les indices de qualité, le dosage).

Nous avons d'abord vérifié l'association d'une variable avec les conditions de traitement et avec les résultantes des enfants pour la période de niveau de base dans une série d'analyses préliminaires rapportée à la section 7.3.1. Ensuite, les variables qui étaient associées significativement aux groupes expérimentaux ont été examinées de plus près. L'objectif de cette analyse supplémentaire est d'identifier un sous-ensemble de variables prédictives du développement des enfants depuis la première évaluation (voir la section 7.3.2). Ces variables seront contrôlées dans le cadre des analyses d'impact.

7.3.1. Associations avec les groupes expérimentaux et les résultantes des enfants

La stratégie suivante a servi à dégager les variables à inclure comme covariables lors des analyses de régression.

- 1) Nous avons vérifié si la relation entre la variable et l'appartenance à un groupe expérimental était significative au niveau de base. Le test utilisé dépendait de la nature de la variable examinée : le test F fut utilisé pour les variables continues, le test du khi-carré pour les variables catégorielles.

⁴⁵ Le lecteur est prié de noter qu'une variable mesurée peut seulement corriger pour des différences observées entre les groupes expérimentaux. Puisque la variable mesurée est imparfaite (elle contient de l'erreur de mesure), il se peut fort bien qu'une différence réelle non-détectée existe entre les groupes. Une différence réelle mais non-détectée ne peut pas être corrigée par l'inclusion de la variable mesurée dans une analyse de régression. C'est un exemple d'une différence non-observée. Des différences non-observées peuvent être annulées uniquement via une assignation aléatoire aux groupes expérimentaux. Par contre, tel que discuté au Chapitre 8, l'effet statique des différences non-observées sur les variables résultantes peut être annulé en utilisant l'estimateur DD.

- 2) Nous avons vérifié si la relation était significative entre la variable et au moins une des résultantes observées au niveau de base.
- 3) Les variables satisfaisant ces deux critères sont incluses dans la spécification des modèles de régression pour corriger le biais potentiel qu'elles représentent. Elles ont aussi fait l'objet d'une analyse supplémentaire (voir la section 7.3.2). Les résultats de ces analyses préliminaires sont rapportés à l'annexe F. Les variables présentées en caractères gras représentent celles retenues comme covariables. Pour plus de détails, le lecteur est prié de consulter l'annexe.

Les covariables retenues dans les analyses d'impact sont : communauté, cohorte, fratrie plus jeune, taille du ménage, Fréquence des activités de littératie (niveau de base), Langue des activités de littératie (niveau de base), langue parlée à l'enfant par la mère, langue parlée à l'enfant par le père, Continuum de français parlé par l'enfant, âge de l'enfant (en mois), Vitalité et Style parental autoritaire. Lors de la deuxième année, on ajoute à cette liste l'inscription de l'enfant à l'école (à temps plein, à temps partiel, non-inscrit).⁴⁶

En somme, un processus de contrôle des variables observées et de leurs associations avec le groupe expérimental et les résultantes des enfants est rigoureusement effectué. Les variables observées non-retenues dans les analyses d'impact se répartissent similairement dans les groupes expérimentaux et par conséquent elles ne contribuent pas significativement à l'estimation de l'impact du programme testé (c.-à-d. les groupes sont appariés pour ces caractéristiques). Pour leur part, les variables retenues servent à effectuer un appariement « statistique » dans le cadre des analyses d'impact. Une source de biais potentielle dans la composition des groupes reste à vérifier, soit la constance dans le temps de l'effet des variables confusionnelles. C'est le sujet de la prochaine section.

7.3.2. Effet instable des covariables dans le temps

À la section précédente, nous avons considéré seulement l'effet des variables observées sur les résultantes des enfants mesurées au niveau de base. Il reste à vérifier si la taille de ces effets est constante pour l'ensemble des évaluations. Les effets qui varient significativement d'une évaluation à l'autre doivent être traités de façon spéciale dans la spécification des analyses de régression.⁴⁷ Les analyses rapportées dans cette section portent seulement sur la première année de l'étude et concernent seulement les variables identifiées à l'annexe F qui étaient significativement associées aux conditions de traitement.⁴⁸

Uniquement l'effet de la taille du ménage est constant dans le temps pour toutes les variables résultantes (voir le tableau F.8 à l'annexe F). Par exemple, les trajectoires développementales des enfants varient significativement en fonction des communautés quand on considère la Communication, le Vocabulaire réceptif et la Fréquence des activités de littératie. Des paramètres représentant cette variabilité seront donc insérés dans les analyses d'impact afin de

⁴⁶ La langue d'enseignement de l'école (français versus immersion en français et anglais) ne varie pas systématiquement en fonction des groupes expérimentaux.

⁴⁷ Des paramètres sont ajoutés à la spécification dans le but de correctement représenter les changements dans la taille de l'effet de la covariable au fil du temps. L'utilité de cette technique pour la correction du biais que peuvent engendrer les variables observées à effet instable est commentée ailleurs par Abadie (2005). Comme toutes estimations fondées sur une correction statistique, les analyses d'impact sont valides dans la mesure où les postulats de bases sont valides (p. ex. le modèle de régression est bien spécifié).

⁴⁸ La « cohorte » de l'enfant était aussi considérée.

corriger le biais qu'elle pourrait engendrer dans l'estimation des impacts du programme. Des paramètres équivalents sont également ajoutés pour les autres variables identifiées aux tableaux F1 à F8, retrouvés à l'annexe F, y inclus la taille du ménage puisque cette variable approche le seuil de signification minimal pour la Fréquence des activités de littératie. Nous faisons l'hypothèse selon laquelle, suite à l'appariement statistique des groupes expérimentaux, ceux-ci se seraient développés au même rythme en l'absence du programme testé.

7.4. REPRÉSENTATIVITÉ DE L'ÉCHANTILLON (PROJET CAPACITÉ D'APPRENTISSAGE VERSUS EVMLO)

La section précédente concerne principalement la comparabilité des groupes pour les fins des analyses d'impact, un critère affectant la validité interne de l'étude. Cette section porte plutôt sur la validité externe de l'étude. Cette dernière a déjà été abordée dans le chapitre 5 du rapport *Capacité d'apprentissage dans les communautés francophones en situation minoritaire : Rapport de référence* (Legault, Mák, Verstraete et Bérubé, 2014; titre abrégé : *Rapport de référence*), où une série systématique de comparaisons a été effectuée qui mettaient en cause l'échantillon du projet Capacité d'apprentissage (1^{ère} cohorte seulement) et celui de l'Enquête sur la vitalité des minorités de langue officielle. La même série d'analyses a été répétée au chapitre 4 du *Rapport des résultats de la première cohorte* pour un échantillon plus restreint : l'échantillon d'enfants provenant des quatre communautés retenues pour les analyses d'impact de la première cohorte. La présente section se veut une actualisation de cette deuxième analyse, prenant en compte cette fois les enfants des deux cohortes.

7.4.1. Problématique

L'échantillon de l'EVMLO sert ici de point de référence pour l'échantillon du projet Capacité d'apprentissage. L'objectif des analyses en question est tout simplement de répondre à la question suivante : si le programme de garderie et les Ateliers familles étaient étendus à toute la minorité francophone dans les communautés participant au projet, est-ce que les effets observés seraient semblables à ceux obtenus dans le projet Capacité d'apprentissage? Afin de répondre à cette question, il est essentiel que les deux échantillons soient tirés de la même population de minorités francophones. Ceci n'est toutefois pas le cas.

L'interprétation des résultats doit être nuancée par le fait que l'EVMLO et le projet Capacité d'apprentissage ont employé des stratégies d'échantillonnage différentes; leurs populations sont différentes et ceci par conséquent limite la comparabilité des échantillons de plusieurs points de vue. Ces différences découlent en partie de la façon dont la population francophone minoritaire est définie dans les deux projets. Comme on l'a expliqué dans le *Rapport de référence*, la procédure d'échantillonnage employée par l'EVMLO est moins restrictive. Des différences démographiques importantes entre les deux échantillons peuvent alors exister uniquement à cause d'une divergence dans la définition de francophonie minoritaire utilisée.

Ensuite, l'échantillonnage du projet Capacité d'apprentissage et celui de l'EVMLO se distinguent au niveau des données démographiques. Afin de veiller à ce qu'un échantillon suffisamment grand soit extrait de la base de données EVMLO, les données provenant d'enfants de 3 à 5 ans ont été prises en considération dans les analyses comparatives. Ce fait contraste avec l'âge moyen de 3 ans observé pour l'échantillon à l'étude lors de l'administration du sondage de

base. Enfin, les participants du projet ont été sélectionnés d'une manière non probabiliste en fonction de la fréquentation d'une garderie, ce qui signifie que la distribution géographique de l'échantillon est localisée au sein de communautés données. Par contraste, la distribution de l'échantillon EVMLO est plus diffuse du point de vue géographique en raison du fait que l'enquête dont il est question ici a recours à l'échantillonnage aléatoire.

Concluons en précisant que les analyses rapportées dans le *Rapport de référence* et dans le *Rapport des résultats de la première cohorte* évaluant la représentativité de l'échantillon à l'étude se sont basées sur des échantillons différents, les premières incluant les communautés d'Edmonton et de Saint-Jean, les deuxièmes les excluant. Certaines différences entre les résultats provenant de ces deux séries d'analyses sont notées dans le *Rapport des résultats de la première cohorte*. Il reste à voir si le même patron de résultats s'obtient lorsqu'on inclut les deux cohortes de l'échantillon projet Capacité d'apprentissage. Nous anticipons que les conclusions tirées de ces nouvelles analyses seront semblables à celles tirées du *Rapport des résultats de la première cohorte* car l'ajout des enfants de la deuxième cohorte (un échantillon nettement plus petit que celui de la première cohorte) ne devrait pas perturber considérablement la composition de l'échantillon.

Le lecteur est prié de noter que dans la mesure du possible, nous avons utilisé les données disponibles pour réévaluer les statistiques de l'EVMLO en fonction de l'échantillon réduit des quatre communautés de sorte à maximiser la validité de la comparaison avec l'échantillon projet Capacité d'apprentissage employée dans les analyses d'impact.⁴⁹ Lorsque cette stratégie était irréalisable pour des raisons pratiques, nous avons effectué des comparaisons à l'échantillon EVMLO glanées à partir des six régions géographiques (*Rapport de référence*). Ces dernières sont reconnues dans le corps du texte.

7.4.2. Statut d'immigrant et profil linguistique

Les procédures d'échantillonnage du projet Capacité d'apprentissage et de l'EVMLO ont été soigneusement détaillées et mises en contraste dans le *Rapport de référence*. La citation qui suit résume les conclusions de cette analyse :

« Selon Forgues et Landry (2006), une population francophone (comme celle qui est utilisée dans le projet Capacité d'apprentissage) qui est définie selon le critère « ayant droit » aboutirait à un bassin bien plus restrictif, tandis qu'une population francophone (comme celle qui est utilisée dans l'EVMLO) qui est définie à l'aide de plusieurs critères (p. ex., langue maternelle, connaissance des langues officielles et langues parlées à la maison) aboutirait à un nombre supérieur de personnes admissibles. »

Deux prédictions ont été avancées à la lumière de cette analyse: a) l'échantillon EVMLO devrait comprendre une proportion plus élevée d'immigrants que le projet Capacité d'apprentissage et b) relativement moins d'enfants devraient déclarer le français comme langue maternelle dans l'EVMLO. Les résultats de l'analyse des informations recueillies sur le statut

⁴⁹ La SRSA a actuellement accès aux données de fréquence groupées par communauté, ce qui permettait le calcul d'estimations appropriées pour l'échantillon global comprenant les quatre communautés. Toutefois, ce calcul n'était pas possible pour certaines variables où l'analyse par communauté a abouti à des tailles d'échantillon trop petites pour être extraites de Statistique Canada. Les six communautés dans l'échantillon EVMLO représentent néanmoins un groupe comparatif intéressant aux fins d'établir la validité externe de l'échantillon projet Capacité d'apprentissage des quatre communautés.

d’immigrant sont présentés dans les prochaines sections. Les comparaisons fondées sur la langue maternelle des enfants et des parents sont aussi rapportées.⁵⁰ Le patron des résultats rapportés ici pour l’analyse par langue maternelle est équivalent à celui qui a été communiqué dans le *Rapport de référence* et dans le *Rapport des résultats de la première cohorte*.

Répondant né au Canada

Le statut d’immigrant des répondants des échantillons projet Capacité d’apprentissage et EVMLO (quatre communautés) est rapporté dans le tableau 7.4. La première rangée rapporte la fréquence des répondants qui signalent être nés au Canada et la deuxième, celle des personnes nées hors du Canada. Comme on pouvait s’y attendre, l’examen de la distribution des réponses pour les deux enquêtes indique clairement que le projet Capacité d’apprentissage comprend une proportion plus grande de répondants nés au Canada. Environ 92 % des membres de l’échantillon à l’étude sont nés au Canada alors que ce fait n’était vrai que pour 75,7 % des répondants de l’échantillon EVMLO. Cette différence apparente a été confirmée par l’application d’un test statistique de khi carré qui s’est révélé significatif [$\chi^2(1, N = 1\ 120) = 36,90, p < 0,001$].

Tableau 7.4 : Comparaison entre le projet Capacité d’apprentissage et l’EVMLO

Statut d’immigrant	Projet Capacité d’apprentissage	EVMLO	Différences significatives entre les deux échantillons? <i>Khi carré</i>
	N (%)	N (%)	
Né au Canada	302 (91,5)	598 (75,7)	Oui***
Né hors du Canada	28 (8,4)	192 (24,3)	

Note : L’échantillon de l’EVMLO provient de quatre régions. Échantillon du projet fondé seulement sur les familles qui ont répondu au huitième sondage. Seuils de signification : *** $p < 0,1\%$; ** $p < 1\%$; * $p < 5\%$

Première langue apprise et toujours comprise — Enfants

La langue maternelle des enfants du projet Capacité d’apprentissage est recueillie à partir du formulaire de consentement rempli par les parents. Pour l’EVMLO, la langue maternelle de l’enfant est déduite de la question suivante: (Statistique Canada, 2006, p. 35): « Quelle est la langue que [nom de l’enfant] a appris en premier à la maison et qu’il comprend encore? ».

Le tableau 7.5 montre que l’échantillon du projet Capacité d’apprentissage est plus francophone que celui de l’EVMLO. En effet, un pourcentage plus important d’enfants du projet déclare seulement le français comme langue maternelle (première rangée du tableau). Le pourcentage d’enfants dont la langue maternelle est l’anglais seulement ou l’anglais et une autre langue est plus grand dans l’échantillon EVMLO (troisième rangée du tableau). La représentation des bilingues dans les deux échantillons est pratiquement identique (deuxième rangée du tableau).

L’application d’un test statistique de khi carré confirme que la distribution des enfants du projet (quatre communautés) parmi les différentes catégories de langue maternelle n’est pas représentative de la population minoritaire francophone dans les six régions géographiques selon les données de l’EVMLO [$\chi^2(2, N = 1\ 103) = 103,73, p < 0,001$]. Ces résultats ne sont pas

⁵⁰ La PLOP (première langue officielle parlée) n’a pas été utilisée ici pour comparer les échantillons du fait que cette donnée a été recueillie dans l’EVMLO de telle sorte que toutes les comparaisons soient invalidées (voir le *Rapport de référence*).

surprenants, étant donné les différences signalées précédemment au sujet des stratégies d'échantillonnage pour les deux projets de recherche.

Tableau 7.5 : Comparaison entre le projet Capacité d'apprentissage et l'EVML0 — Enfants groupés par langue maternelle

Langue maternelle	Projet Capacité d'apprentissage	EVML0	Différences significatives entre les deux échantillons? <i>Khi carré</i>
	N (%)	N (%)	
Français seulement	247 (72,2)	306 (40,2)	
Anglais et français également OU français et une autre langue	33 (9,7)	89 (11,7)	Oui***
Anglais seulement OU anglais et une autre langue OU autre(s) langue(s)	62 (18,1)	366 (48,1)	

Note : L'échantillon de l'EVML0 provient de six régions. Seuils de signification : *** $p < 0,1\%$; ** $p < 1\%$; * $p < 5\%$.

Première langue apprise et toujours comprise — Mères

Le tableau 7.6 indique que la plupart des mères des échantillons du projet (67,0 %) et de l'EVML0 (58,3 %) déclarent le français comme seule langue maternelle (première rangée du tableau), bien que la proportion à cet égard soit légèrement plus élevée dans l'échantillon du projet Capacité d'apprentissage. De plus, une proportion moindre de mères du projet déclarent la catégorie « anglais seulement OU anglais et une autre langue OU autre(s) langue(s) » (23,1 % dans la troisième rangée). L'application du test de khi carré suggère que la distribution des mères du projet (quatre communautés) entre les différentes catégories de langue maternelle n'est pas représentative de la population minoritaire francophone dans les six régions géographiques fondées sur les données de l'EVML0 [$X^2(2, N = 1\ 131) = 9,68, p < 0,01$].

Tableau 7.6 : Comparaison entre les mères du projet Capacité d'apprentissage et de l'EVML0 — Mères groupées par langue maternelle

Langue maternelle	Projet Capacité d'apprentissage	EVML0	Différences significatives entre les deux échantillons? <i>Khi carré</i>
	N (%)	N (%)	
Français seulement	229 (67,0)	460 (58,3)	
Anglais et français également OU français et une autre langue	34 (9,9)	75 (9,5)	Oui**
Anglais seulement OU anglais et une autre langue OU autre(s) langue(s)	79 (23,1)	254 (32,2)	

Note : L'échantillon de l'EVML0 provient de six régions. Seuils de signification : *** $p < 0,1\%$; ** $p < 1\%$; * $p < 5\%$.

Première langue apprise et toujours comprise — Pères

Le tableau 7.7 montre le profil linguistique des pères du projet Capacité d'apprentissage et de l'EVML0 en fonction de leur langue maternelle. Le patron apparaît au premier coup d'œil semblable à celui observé chez les mères. La principale différence réside dans le nombre

important de pères de l'EVMLLO qui ont déclaré « anglais seulement OU anglais et une autre langue OU autre(s) langue(s) », comme langue maternelle. La proportion de pères de l'EVMLLO aux deux extrémités de la distribution dans le tableau 7.7 est virtuellement identique (46,7 % contre 47,7 %). Par contre, les pères de l'échantillon à l'étude sont plus massivement représentés dans la catégorie « français seulement » (56,1 % dans la première rangée) que dans la catégorie « anglais seulement OU anglais et une autre langue OU autre(s) langue(s)» (35,5 % dans la troisième rangée). Ce dernier patron ressemble à celui observé dans les deux échantillons pour la langue maternelle de la mère.

L'application du test de khi carré confirme que la distribution des pères du projet (quatre communautés) entre les différentes catégories de langue maternelle n'est pas représentative de la population minoritaire francophone dans les six régions géographiques fondées sur les données de l'EVMLLO [$X^2(2, N = 1\ 116) = 15,04, p < 0,001$].

Tableau 7.7 : Comparaison entre les pères du projet Capacité d'apprentissage et de l'EVMLLO — Pères groupés par langue maternelle

Langue maternelle	Projet Capacité d'apprentissage	EVMLLO	Différences significatives entre les deux échantillons? <i>Khi carré</i>
	N (%)	N (%)	
Français seulement	185 (56,1)	367 (46,7)	
Anglais et français également OU français et une autre langue	28 (8,5)	44 (5,6)	Oui***
Anglais seulement OU anglais et une autre langue OU autre(s) langue(s)	117 (35,4)	375 (47,7)	

Note : L'échantillon de l'EVMLLO provient de six régions. Seuils de signification : *** $p < 0,1\%$; ** $p < 1\%$; * $p < 5\%$.

7.4.3. Caractéristiques sociodémographiques

Dans le *Rapport de référence* et le *Rapport des résultats de la première cohorte*, l'échantillon projet Capacité d'apprentissage a été comparé avec l'échantillon EVMLLO en fonction du sexe de l'enfant, de la composition de la famille (taille et structure de la famille) et de la situation socio-économique (scolarité parentale, revenu familial). Les rapports concluent que les échantillons des deux enquêtes sont comparables en ce qui concerne les variables sexe de l'enfant (vérifiée pour le *Rapport de référence* seulement), structure familiale, taille de la famille et revenu familial. Par contre, on constate que la distribution des réponses varie entre les échantillons pour la scolarité parentale (père et mère) et le nombre de frères et sœurs. Dans tous les cas, le patron général de ces résultats s'est reproduit dans l'analyse fondée sur les deux cohortes dont les résultats sont présentés ci-dessous.

Revenu familial total

Le tableau 7.8 montre que les familles du projet Capacité d'apprentissage et celles de l'EVMLLO sont distribuées d'une façon similaire entre les catégories de revenu considérées ici. Dans les deux cas, la catégorie modale et médiane pour les deux échantillons est de 60 000 \$ ou plus par année. L'application du test de khi carré confirme que la distribution des parents du projet (quatre communautés) au sein des différentes classes de revenu est statistiquement

équivalente à celle observée auprès de la population minoritaire francophone des six régions géographiques fondées sur les données de l'EMVLO [$X^2(5, N = 1\ 131) = 5,58, p > 0,05$]. Les résultats suggèrent que la plupart des enfants dans les deux échantillons jouissent d'une bonne qualité et d'une bonne quantité de ressources matérielles pour leur développement.

Tableau 7.8 : Comparaison entre le projet Capacité d'apprentissage et l'EMVLO — Familles par classe de revenu

Classe de revenu	Projet Capacité d'apprentissage	EMVLO	Différences significatives entre les deux échantillons? <i>Khi carré</i>
	N (%)	N (%)	
10 000 \$ ou moins	20 (5,8)	54 (6,9)	Non
Entre 20 000 \$ et 29 999 \$	14 (4,1)	23 (2,9)	
Entre 30 000 \$ et 39 999 \$	19 (5,6)	64 (8,1)	
Entre 40 000 \$ et 49 999 \$	18 (5,3)	57 (7,2)	
Entre 50 000 \$ et 59 999 \$	46 (13,4)	95 (12,0)	
60 000 \$ et plus	225 (65,8)	496 (62,9)	

Note : L'échantillon de l'EMVLO provient de six régions. Seuils de signification : *** $p < 0,1\%$; ** $p < 1\%$; * $p < 5\%$.

Degré de scolarité des mères

Le tableau 7.9 révèle trois points principaux. Premièrement, les mères du projet Capacité d'apprentissage ont en moyenne plus de scolarité que les mères de l'échantillon de l'EMVLO. En fait, près de 80 % des mères du projet détiennent un diplôme d'études collégiales (DEC) ou un diplôme universitaire comparativement à environ 70 % des mères de l'EMVLO. Cette différence est principalement attribuable au petit nombre de mères ayant au moins un diplôme ou un certificat d'études collégiales dans l'EMVLO en comparaison à celles du projet (deuxième rangée). Deuxièmement, il y a autant de mères avec un diplôme d'études collégiales (38,9 %) qu'il y a de mères avec un diplôme universitaire (40,9 %) dans le projet. Troisièmement, il y a plus de mères dans l'EMVLO ayant fréquenté l'université (42,7 %) que dans le projet (40,9 %) bien que cette différence soit négligeable.

L'application du test de khi carré confirme que le niveau de scolarité des mères du projet Capacité d'apprentissage n'est pas représentatif de la population minoritaire francophone dans les quatre régions géographiques fondées sur les données de l'EMVLO [$X^2(2, N = 883) = 22,74, p < 0,01$].

Tableau 7.9 : Comparaison des degrés de scolarité des mères du projet Capacité d'apprentissage et de l'EMVLO

Degré de scolarité	Projet Capacité d'apprentissage	EMVLO	Différences significatives entre les deux échantillons? <i>Khi carré</i>
	N (%)	N (%)	
Diplôme d'études secondaires ou moins OU quelques cours postsecondaires	69 (20,2)	172 (31,8)	
Diplôme ou certificat collégial (p. ex. école de métier)	133 (38,9)	138 (25,5)	Oui***
Diplôme universitaire (baccalauréat, maîtrise ou doctorat)	140 (40,9)	231 (42,7)	

Note : L'échantillon de l'EMVLO provient de quatre régions. Seuils de signification : *** $p < 0,1\%$; ** $p < 1\%$; * $p < 5\%$.

Degré de scolarité des pères

Le tableau 7.10 indique qu'il y a plus de pères qui ont fréquenté l'université dans l'EMVLO (35,9 %) que dans le projet Capacité d'apprentissage (34,6 %). Toutefois, les pères du projet sont généralement plus scolarisés que les pères de l'EMVLO. En effet, près des deux tiers d'entre eux ont un diplôme d'études collégiales (DEC) ou un diplôme universitaire, tandis que 60 % des pères de l'EMVLO ont un degré de scolarité équivalent. Enfin, l'application du test khi carré suggère que le niveau de scolarité des pères du projet Capacité d'apprentissage n'est pas représentatif de la population minoritaire francophone dans les quatre régions géographiques fondées sur les données de l'EMVLO [$\chi^2(2, N = 873) = 7,48, p < .05$].

Tableau 7.10 : Comparaison entre les degrés de scolarité des pères du projet Capacité d'apprentissage et de l'EMVLO

Degré de scolarité	Projet Capacité d'apprentissage	EMVLO	Différences significatives entre les deux échantillons? <i>Khi carré</i>
	N (%)	N (%)	
Diplôme d'études secondaires ou moins OU quelques cours postsecondaires	110 (33,1)	216 (39,9)	
Diplôme ou certificat d'études collégiales (p. ex., école de métier)	107 (32,2)	131 (24,2)	Oui**
Diplôme universitaire (baccalauréat, maîtrise ou doctorat)	115 (34,7)	194 (35,9)	

Note : L'échantillon de l'EMVLO provient de quatre régions. Seuils de signification : *** $p < 0,1\%$; ** $p < 1\%$; * $p < 5\%$.

Taille de la famille

Selon le tableau 7.11, la taille modale et médiane de la famille est de quatre personnes pour les deux échantillons (quatre communautés). Dans les deux enquêtes, des familles composées de

quatre personnes représentent environ la moitié de l'échantillon. Les familles restantes sont distribuées assez également entre les familles de trois personnes ou moins et les familles de cinq personnes ou plus. L'application du test de khi carré suggère l'absence de différence significative dans la distribution de la taille des familles participant au projet Capacité d'apprentissage et celle des familles provenant des quatre régions géographiques de l'EVMLO [$X^2(2, N = 880) = 2,48, p > 0,05$].

Tableau 7.11 : Comparaison de la taille des familles¹ dans le projet Capacité d'apprentissage et l'EVMLO

Nombre de personnes	Projet Capacité d'apprentissage	EVMLO	Différences significatives entre les deux échantillons? <i>Khi carré</i>
	N (%)	N (%)	
Trois personnes ou moins	78 (22,8)	143 (26,6)	Non
Quatre personnes	196 (57,3)	280 (52,0)	
Cinq personnes ou plus	68 (19,9)	115 (21,4)	

Note : L'échantillon de l'EVMLO provient de quatre régions. Seuils de signification : *** $p < 0,1\%$; ** $p < 1\%$; * $p < 5\%$. ¹ Le nombre de personnes dans une famille comprend seulement le nombre de parents et d'enfants.

Fratrie

Le tableau 7.12 indique que le nombre modal et médian d'enfants par répondant (famille) est de deux pour le projet Capacité d'apprentissage et pour l'EVMLO. Toutefois, il y a un peu plus de familles avec deux enfants dans le projet (60,8 %) que dans l'EVMLO (50,4 %). De plus, il y a plus de familles avec trois enfants ou plus dans l'EVMLO (30,1 %) que dans le projet (20,4 %). En revanche, le nombre de familles avec un enfant seulement, environ 20 %, est à peu près le même dans les deux échantillons. L'application du test de khi carré confirme que la distribution du nombre d'enfants par répondant du projet Capacité d'apprentissage n'est pas représentative de la population minoritaire francophone dans les quatre régions géographiques fondées sur les données de l'EVMLO [$X^2(2, N = 1\,128) = 12,97, p < 0,01$].

Tableau 7.12 : Comparaison entre le projet Capacité d'apprentissage et l'EVMLO — Nombre d'enfants par répondant

Nombre d'enfants	Projet Capacité d'apprentissage	EVMLO	Différences significative entre les deux échantillons? <i>Khi carré</i>
	N (%)	N (%)	
Un enfant	64 (18,7)	154 (19,6)	Oui**
Deux enfants	208 (60,8)	396 (50,4)	
Trois enfants ou plus	70 (20,5)	236 (30,0)	

Note : L'échantillon de l'EVMLO provient de quatre régions. Seuils de signification : *** $p < 0,1\%$; ** $p < 1\%$; * $p < 5\%$.

Structure familiale

Nous avons eu à redéfinir les familles du projet Capacité d'apprentissage comme monoparentales ou biparentales afin de pouvoir comparer la structure familiale de l'échantillon du projet avec celle de l'EVMLO (voir tableau 7.13). Notons que la catégorie biparentale

regroupe les familles intactes et les familles reconstituées où les deux parents (ou un parent et son conjoint) vivent avec l'enfant. La catégorie monoparentale comprend les familles où seul un parent vit dans la maison avec l'enfant.

Précisons que le père ou la mère de l'enfant peut être soit le parent biologique, soit le parent adoptif. Enfin, les couples de même sexe ont été exclus de l'analyse ainsi que les enfants élevés par quelqu'un d'autre que la mère ou le père biologique ou adoptif de l'enfant. L'application du test de khi carré confirme que la distribution des enfants dans les foyers monoparentaux ou biparentaux dans le projet est représentative de la population minoritaire francophone dans les quatre régions géographiques fondées sur les données de l'EMVLO [$X^2(1, N = 1\ 131) = 1,68, p > 0,05$].

Tableau 7.13 : Comparaison entre le projet Capacité d'apprentissage et l'EMVLO — Nombre de familles monoparentales et biparentales

Structure familiale	Projet Capacité d'apprentissage	EMVLO	Différences significatives entre les deux échantillons?
	N (%)	N (%)	Khi carré
Monoparentale	29 (8,5)	87 (11,0)	Non
Biparentale	313 (91,5)	702 (89,0)	

Note : L'échantillon de l'EMVLO provient de quatre régions. Seuils de signification : *** $p < 0,1\%$; ** $p < 1\%$; * $p < 5\%$.

7.5. SOMMAIRE DES IMPLICATIONS POUR LES ANALYSES D'IMPACT

Dans ce chapitre, nous avons abordé les diverses questions méthodologiques touchant la validité interne et externe des analyses d'impact à suivre. Les questions techniques liées aux procédures de contrôle de la qualité, aux valeurs manquantes, à l'imputation, aux variables confusionnelles et à la validité externe ont été examinées.

Le chapitre commence en répertoriant les procédures de vérification et de contrôle de la qualité des données (section 7.1). Ces procédures ont été conçues pour minimiser les problèmes (p. ex. les erreurs de mesure) posés par le processus de collecte de données et pour identifier et corriger les problèmes présents dans les bases de données électroniques. Ces procédures et l'utilisation d'une approche par méthodes mixtes (ou « opérations convergentes ») dans la recherche assurent la validité des résultats découlant des analyses d'impact.

Le chapitre se poursuit avec l'analyse des valeurs manquantes (section 7.2) dont le résultat montre que, dans l'ensemble, les valeurs manquantes se répartissent également à travers les groupes expérimentaux et constituent un problème d'une ampleur très limitée. Le taux d'attrition est très faible durant les deux premières années du projet, alors que le taux de réponses est très élevé. Les analyses indiquent également que la performance sur les résultantes des enfants ne permet pas de prédire l'attrition des participants. Bien que ces analyses soulignent que les participants se retirant de l'étude ont des caractéristiques particulières (p. ex. ils sont moins « francophones », voir le *Rapport des résultats de la première cohorte*), la généralisation des résultats à la population minoritaire francophone n'en est pas affectée. Les tests statistiques révèlent que les autres valeurs manquantes associées aux données des sondages sont produites

aléatoirement. Ces dernières ont été imputées pour maximiser les effectifs en vue des analyses d'impact. Par contre, le traitement des données manquantes non aléatoires a été problématique avec toutes les échelles de l'ÉPE-AD, sauf pour la Communication. Ces données manquantes sont causées par l'administration de l'ÉPE-AD en anglais plutôt qu'en français. La stratégie d'imputation adoptée ici a permis d'imputer certaines des données manquantes comme MAR (c.-à-d. manquante non aléatoire qui peut être imputée de façon fiable selon d'autres variables dans la base de données), mais pas toutes (MNAR, c.-à-d. manquante non aléatoire qui ne peut pas être imputée). En conséquence, les analyses d'impact reposant sur ces échelles ne peuvent être généralisées qu'à une population de francophones minoritaires dont les enfants sont susceptibles de rencontrer les critères nécessaires pour compléter l'ÉPE-AD en français (c.-à-d. les enfants dont le score est relativement élevé sur l'échelle Langage et communication en français, domaine C, et l'échelle Conscience et engagement envers la culture francophone, domaine E) ou, plus exactement, à une population d'enfants qui aurait pu compléter l'ÉPE-AD en français deux fois ou plus au cours de cinq évaluations alors qu'ils sont à peu près âgés entre trois et quatre ans.

Ensuite, nous rapportons dans le chapitre le résultat de l'analyse des variables confusionnelles (section 7.3). Le but de cette analyse est d'identifier des variables permettant de: a) faire des ajustements statistiques compensant pour les changements survenant dans la composition des groupes au fil du temps; et b) faire des ajustements statistiques compensant pour les différences dans la trajectoire développementale qui se seraient manifestées entre les groupes en l'absence d'un traitement. Au terme de cette analyse, plusieurs covariables à inclure dans les analyses d'impact sont identifiées (voir l'annexe F pour une liste exhaustive des variables examinées et le détail des résultats). L'insertion de ces covariables dans les analyses d'impact servira à maximiser la validité interne de cette étude quasi-expérimentale à groupe non-équivalents.

Enfin, nous avons examiné la question de la validité externe de l'étude. Des travaux antérieurs ont conclu que l'échantillon comprenait plus de francophones dans le projet Capacité d'apprentissage comparativement à l'échantillon de l'EVMLO. Ce résultat est reproduit ici avec l'échantillon des deux cohortes du projet. Dans le *Rapport de référence*, nous avons formulé l'hypothèse selon laquelle ce biais apparent découlerait de la population d'enfants typiquement desservie dans des garderies francophones. Dans la mesure où cet argument est fondé, tous les résultats obtenus dans la présente étude s'appliqueraient uniquement à une population d'enfants francophones inscrits en garderie. De prime abord, cette limite semble aller de soi en ce sens que, cette population d'enfants est celle qui serait touchée par une intervention en garderie. Par contre, si les différences dans la langue maternelle observées entre le projet Capacité d'apprentissage et l'EVMLO reflètent un manque réel de représentativité, ce qui revient à dire que l'échantillon du projet n'est pas représentatif de la population ciblée, alors les résultats des analyses d'impact figurant aux chapitres 9 et 10 du présent rapport peuvent en réalité sous-estimer l'importance de l'effet réel du programme pour une population caractérisée par une plus grande diversité linguistique. La possibilité que les variables linguistiques accentuent ou atténuent l'effet du programme testé peut être examinée en réévaluant les effets du programme séparément pour les enfants utilisant surtout la langue française et ceux l'utilisant moins (voir les sections 9.2.5 et 9.3.5). L'information ressortant de ces analyses peut aider à déterminer si l'intervention serait plus efficace si elle ciblait des sous-populations données, dont celle composée d'enfants surtout exposés à des langues autres que le français dans leur environnement familial.

Le fait que les échantillons du projet Capacité d'apprentissage et de l'EVMLO soient généralement similaires sur des caractéristiques non linguistiques soutient l'argument selon lequel l'échantillon du projet est représentatif des francophones en milieu minoritaire. La seule différence méritant d'être mentionnée tient au fait que les parents du projet Capacité d'apprentissage sont légèrement plus susceptibles d'atteindre un niveau de scolarité supérieur au niveau secondaire comparativement à l'échantillon de l'EVMLO.

En somme, plusieurs précautions ont été prises afin d'assurer la validité interne des effets estimés du programme testé et d'en estimer le degré de validité externe. D'autres vérifications font l'objet de discussions à mesure qu'elles deviennent pertinentes à l'interprétation des résultats dans les sections suivantes du présent rapport.

8. Approche adoptée pour les analyses

Le présent chapitre traite de la logique sous-jacente aux analyses décrites aux chapitres 9 et 10. L'approche économétrique adoptée prend compte de la structure longitudinale et multi-niveau des données et prévoit une vérification systématique de la robustesse de résultats (voir la section 8.1). Enfin, la méthode utilisée pour identifier l'effet du programme est convenable pour une étude quasi-expérimentale à groupes non-équivalents (voir la section 8.2).

Cette approche générale a servi à la spécification de plusieurs modèles empiriques des données. Plusieurs d'entre eux sont présentés dans les chapitres à suivre, chacun représentant par sa spécification une façon différente de concevoir l'exposition des familles et des enfants au programme testé. Ici, l'exposition au programme testé est conçue comme un continuum allant de faible à forte intensité (voir la section 8.3). La plus forte intensité d'exposition possible serait reçue par un enfant qui fréquente à temps plein une garderie qui met en œuvre le programme précisément tel qu'il a été conçu et avec le plus haut degré de qualité. Les modèles complémentaires considérés dans le présent rapport sont introduites dans les sections 8.3 et 8.4. Finalement, l'organisation de la présentation des résultats dans les chapitres 9 et 10 est annoncée à la section 8.5.

8.1. MODÉLISATION LINÉAIRE HIÉRARCHIQUE (MLH)

Les données ont été analysées à l'aide de modèles de régression linéaires qui font le postulat fondamental voulant que chaque point d'observation ou point de données présent dans l'analyse ait été observé de façon indépendante. Ce postulat n'est pas respecté lorsque les unités d'échantillonnage (p. ex. les garderies ou les enfants) contribuent à de multiples observations d'un ensemble de données. Dans ce cas, on dit des observations faites auprès d'une même unité d'échantillonnage qu'elles sont imbriquées ou groupées. Dans un devis longitudinal, les observations sont groupées par participant (c.-à-d. que chaque participant contribue à plusieurs observations) et parfois en fonction d'un type supplémentaire d'unités d'analyse. Le projet Capacité d'apprentissage en particulier présente des données qui sont imbriquées par garderie et par participant. Cette structure à multiples niveaux, souvent analysée par MLH, doit être prise en compte dans les analyses pour éviter de surestimer la signification statistique des résultats (Hox, 2002; Moulton, 1990). Le traitement de l'aspect longitudinal de l'étude est détaillé plus bas à la section 8.2, tandis que le traitement de l'effet des « garderies » est abordé maintenant.

Pour maximiser la robustesse des résultats, nous utilisons l'estimateur de cohérence de l'hétérogénéité de Huber-White (White, 1980) avec une modification qui le rend robuste au groupement (Williams, 2000). Selon Woodbridge (2002), cette méthode possède des propriétés satisfaisantes pour l'analyse d'une base de données comme celles du projet Capacité d'apprentissage (c.à.d. le ratio observé du nombre de groupes par rapport aux observations par groupe) quand le nombre de groupes détermine les degrés de liberté des tests de signification.

La précision des effets rapportés est alors robuste à l'hétérogénéité et au groupement par milieu de garde.⁵¹

8.2. LA MÉTHODE DIFFÉRENCE DES DIFFÉRENCES

Une stratégie courante pour estimer les effets programme dans la littérature en économétrie consiste à employer un estimateur différence des différences (DD) (Abadie, 2005; Bertrand, Duflos, et Mulliainathan, 2004). L'estimateur DD s'applique bien aux données recueillies au moyen d'un panel et est approprié aux devis de recherche non expérimental à mesures répétées employant une mesure pré-test « niveau de base » et un groupe témoin. Les lecteurs qui connaissent moins bien la littérature en économétrie seront peut-être plus familiers avec l'idée d'une interaction ou d'un effet modérateur pour laquelle l'estimateur DD est un cas spécial. Le terme estimateur DD fait référence à l'interaction impliquant deux variables fictives, l'une représentant les périodes de test pré- et post-traitement (pré-test vs. post-test) et l'autre représentant les deux groupes comparés (groupe d'intervention vs. groupe témoin).

Cet estimateur peut être généralisé à des scénarios plus complexes en spécifiant plusieurs termes d'interaction de ce type dans le même modèle de régression (p. ex., pour comparer le pré-test à de multiples mesures post-test dans le même modèle de régression) ou pour comparer des groupes multiples. Les analyses d'impact rapportées dans le présent document ont fait appel à des estimateurs DD multiples pour représenter les comparaisons des trois groupes à l'étude en fonction de plusieurs évaluations post-intervention. Les trois groupes ont été inclus dans la même analyse afin de maximiser la stabilité des tests statistiques, ce qui augmente les chances de déceler les effets réels du programme.

Comme son nom le suggère, l'estimateur DD possède deux composantes de base. La première est une estimation du changement (Δ) de l'évaluation pré-test à une évaluation post-test pour chaque groupe ($\Delta_{\text{groupe d'intervention}} = \text{post-test} - \text{pré-test}$; $\Delta_{\text{groupe témoin}} = \text{post-test} - \text{pré-test}$). Ici, la mesure pré-test est la norme à laquelle toutes les évaluations subséquentes sont comparées. En d'autres termes, les mesures post-test sont toujours comparées à la mesure pré-test lors de l'estimation des effets du programme. Ces scores de changement (c.-à-d. de différences) ne sont toutefois pas suffisants pour isoler l'effet traitement, car le changement faisant l'objet d'une estimation pourrait être survenu suite aux processus développementaux naturels (p. ex., la maturation).

La deuxième composante de l'estimateur DD élimine cette ambiguïté en enregistrant la différence entre les scores de changement du groupe d'intervention et ceux du groupe témoin ($DD = \Delta_{\text{intervention}} - \Delta_{\text{témoin}}$). Ici, l'estimation du changement fournie par le groupe témoin sert à ajuster les estimations de changement pour le groupe d'intervention. La validité de l'estimateur DD part de la prémisse selon laquelle une fois que vous éliminez l'estimation contrefactuelle du changement fournie par le groupe témoin, tout ce qui reste est le changement provoqué par l'effet traitement. Comme avec la plupart des statistiques, les postulats soutenant cette prémisse

⁵¹ Puisque cette méthode est parfois critiquée (Donald et Lang, 2007), la robustesse des résultats a été vérifiée en répétant les analyses à plusieurs reprises. À chaque itération, l'effet des « garderies » a été traité d'une façon différente (p. ex., non-spécifié; comme effet aléatoire; comme effet fixe). Cette vérification révèle que la méthode utilisée n'affecte pas dramatiquement les résultats obtenus et que la méthode adoptée est typiquement la plus conservatrice pour la présente base de données. Notre décision a été d'être conservateur dans nos choix méthodologiques, préférant la « cohérence » à l'« efficacité » (Hayes & Cai, 2007).

sont susceptibles d'être violés dans la pratique. Cela est particulièrement vrai dans le contexte d'une étude non expérimentale effectuée sur le terrain, dont le projet Capacité d'apprentissage est un exemple.

La validité de l'estimateur DD dépend d'au moins deux postulats. En ce qui concerne les scores de changement, nous supposons que la composition des groupes demeure constante pour toutes les périodes d'évaluation soumises à la comparaison, ce qui n'est peut-être pas le cas si les participants migrent d'un groupe à l'autre au cours de l'étude. Les changements de groupe posent donc un problème pour le projet Capacité d'apprentissage qui exige un contrôle dans les analyses. Pour ce qui est de la comparaison des scores de changement, nous supposons qu'en l'absence du programme testé, les groupes que l'on compare entre eux auraient suivi une évolution identique. Il n'y a évidemment aucune façon de vérifier ce postulat directement, mais les différences initiales entre les groupes quant aux variables associées à la mesure dépendante soulèvent la possibilité de pentes non parallèles.

Comme pour le *Rapport des résultats de la première cohorte*, nous avons minimisé ces menaces à la validité de l'estimateur DD en incluant dans nos analyses des covariables qui ont permis d'effectuer des ajustements qui tenaient compte des différences entre les groupes à leur état initial. L'effet statique et l'effet dynamique de ces covariables sont représentés dans les modèles empiriques respectivement par l'effet direct de ces variables et leurs interactions avec le facteur longitudinal. Les détails concernant ce processus de sélection ont déjà été présentés au chapitre 7 et à l'annexe F. Une discussion plus approfondie des mesures prises pour assurer la validité des résultats est présentée à la section 8.4.

8.3. CONCEPTUALISATION DU TRAITEMENT ET DU DOSAGE

Le modèle le plus simple possible pour capter un effet du traitement contient deux groupes de participants: le groupe « traité » et le groupe « non traité ». Ce type de modèle est le plus valide lorsque la distinction entre un groupe traité et un groupe non traité est absolue, c'est-à-dire que le groupe non traité ne reçoit absolument aucun traitement, alors que tous les membres du groupe traité reçoivent un traitement équivalent (p. ex., un dosage ou une « exposition » équivalent(e)). Dans le cas d'une étude effectuée sur le terrain, des distinctions aussi nettes sont plutôt rares, et le projet Capacité d'apprentissage ne fait pas exception à cette règle. L'intensité du traitement reçu par les participants variait en fonction d'au moins deux dimensions: le degré d'exposition (dosage) au traitement et la qualité du traitement. Nous détaillons dans les prochaines sections ces deux sources de variation dans l'intensité du traitement.

8.3.1. Exposition hétérogène au traitement

Lors de leur inscription au projet, les participants faisaient partie d'un des trois groupes: le groupe programme en garderie, le groupe témoin en garderie ou le groupe témoin hors garderie. Il n'est peut-être pas étonnant de réaliser que cette appartenance ait parfois varié au fil du temps suite aux choix des parents quant au mode de garde de leur enfant. Par exemple, un pourcentage relativement faible de parents (voir le tableau 3.2) a décidé de changer de garderie entre les évaluations. Par conséquent, un enfant qui faisait partie du groupe programme en garderie lors des deux premières évaluations pouvait se trouver dans le groupe témoin en garderie ou dans le groupe témoin hors garderie pour les évaluations subséquentes.

Même si les participants n'ont pas migré d'un groupe à un autre, le temps qu'ils ont passé en garderie pouvait varier au fil du temps. Les résultats rapportés à l'annexe F attestent que les deux groupes en garderie ne différaient pas quant à leur exposition moyenne au milieu de garde. Une conséquence pratique de cette équivalence est que cette variable ne constitue pas une menace évidente à la validité de nos estimations sur les effets du programme testé. Néanmoins, il est intéressant de se demander si le degré d'exposition à une condition de traitement a de l'importance et si les effets du programme, le cas échéant, ont une interaction synergétique avec le degré d'exposition. En d'autres mots, de poser la question : est-ce qu'une heure passée dans une garderie programme produit de meilleurs résultats qu'une période de temps équivalente passée dans une garderie témoin? Nous avons traité cette question en insérant les variables liées au degré d'exposition au traitement dans les analyses d'impact.

8.3.2. Évaluation de l'effet du dosage ou de l'exposition

Deux stratégies ont été adoptées pour gérer la migration des participants au fil du temps. La première est brute, mais sa simplicité lui confère l'avantage de faciliter la comparaison des groupes (c.-à-d. traités versus non traités). La seconde stratégie définit le traitement avec plus de précision en indiquant les heures d'exposition à un milieu de garde donné (c.-à-d. la moyenne d'heures passées en garderie par semaine). Cette dernière stratégie est un peu plus complexe, mais elle possède l'avantage de représenter le degré d'exposition aux deux conditions de traitement en garderie. Ces deux techniques servant à représenter l'hétérogénéité de l'exposition au traitement font l'objet de deux séries d'analyses indépendantes, dont nous traitons plus en détails ci-dessous.

La première série d'analyses repose sur la définition brute de l'exposition au traitement, où des codes fictifs ont servi à représenter l'appartenance à un groupe expérimental. Pour toute période de temps donnée, un participant codé comme appartenant à un certain groupe expérimental était considéré comme ayant reçu entièrement le traitement associé à ce groupe. L'appartenance à un groupe pouvait changer au fil du temps, mais nous avons maximisé l'intégrité des groupes expérimentaux en exigeant que les participants soient exposés à leur nouveau mode de garde durant une période de temps minimale avant de reconnaître ce changement dans les analyses. Plus précisément, les deux conditions suivantes ont été posées. Premièrement, on considérait qu'un enfant avait changé de groupe uniquement si le changement était survenu plus d'un mois avant l'évaluation. Deuxièmement, les changements de modalités de garde effectués au cours des mois d'été ont été considérés valides que si ces nouveaux arrangements avaient persisté pendant quelques mois après la rentrée scolaire. Un participant qui rapportait son départ d'une garderie programme et son inscription subséquente dans une garderie témoin pendant l'été, par exemple, était considéré comme membre du groupe programme en garderie pour l'évaluation automnale. Pour les fins d'analyses, le changement de groupe n'entrait en vigueur que pour l'évaluation hivernale puisque les enfants avaient alors été exposés au milieu des garderies pendant quelques mois durant lesquels le programme habituel était en place. Les deux mesures de vocabulaire normalisées sont exceptionnelles puisqu'elles sont administrées une seule fois chacun. Le regroupement des participants qui était en vigueur lors de l'administration de ces mesures était utilisé pour les fins des analyses. Peu importe les résultantes mesurées, nous évaluons dans tous les cas le biais dans l'estimation de l'effet du programme pouvant résulter des changements de groupe.

Dans la seconde série d'analyses visant les données de la première année, le traitement a été défini en utilisant les heures d'exposition à un milieu de garde, conjointement avec les caractéristiques de ce milieu (c.-à-d. le groupe programme en garderie ou le groupe témoin en garderie). Nous avons défini les heures d'exposition comme étant la moyenne d'heures qu'un enfant passe en garderie chaque semaine durant la période de quatre mois précédant une évaluation particulière. L'exception était la période de base pour laquelle on utilisait seulement deux mois (septembre et octobre 2007 et 2008, respectivement, pour les deux premières cohortes). L'inclusion de cette variable dans les analyses a permis de définir l'exposition à notre programme en garderie de manière plus précise. En croisant la variable des heures d'exposition avec notre variable de groupement (c.-à-d. en spécifiant un terme d'interaction), nous avons pu: a) estimer l'effet moyen du traitement associé à un nombre donné d'heures d'exposition par semaine; et b) tester si l'effet du degré d'exposition à la garderie varie en fonction du type de programme. Ce dernier test est simplement une généralisation de notre hypothèse de recherche fondamentale selon laquelle, en présence d'un degré d'exposition équivalent, les participants du groupe programme seront avantagés quant aux trois domaines de préparation à la scolarisation. Cette conceptualisation plus précise du dosage accroît la puissance statistique de nos analyses de détecter les effets du programme. Elle nous offre aussi la possibilité de découvrir le degré d'exposition au programme requis pour provoquer l'effet souhaité.

Lors de l'analyse de suivi des résultantes de la deuxième année, nous avons utilisé les données sur les heures d'exposition de façon quelque peu différente. À la fin de la première année, la plupart des enfants de l'étude (c.-à-d., les enfants vivant dans les trois communautés ontariennes) étaient inscrits à la maternelle à temps partiel ou à temps plein. Pour cette raison, les données sur la fréquentation des garderies recueillies pendant la deuxième année du projet sont difficiles à interpréter en tant que prédicteurs des résultantes puisqu'elles ne reflètent pas adéquatement le temps passé dans un milieu de garde de qualité. Les enfants inscrits à temps plein ou à temps partiel dans une école sont exposés à un environnement de qualité. Par conséquent, l'enfant qui passe quelques heures par semaine dans une garderie programme peut en réalité passer beaucoup d'heures à l'extérieur de la maison et vivre l'expérience d'un traitement d'une intensité comparable.

Il est indubitable que les « heures d'exposition en garderie » ne constituent pas une variable importante dans la deuxième année du projet. Il serait techniquement possible d'effectuer une analyse de ce type à l'aide seulement de la communauté d'Edmundston (c.-à-d. qu'aucun enfant de cette communauté n'était inscrit à l'école), mais le nombre d'enfants inclus dans une telle analyse serait trop petit pour fournir de l'information utile.⁵² Nous avons plutôt traité les évaluations de la deuxième année en tant que tests de suivi quant à l'effet des heures d'exposition en garderie de la première année. L'investissement nécessaire pour fournir des services de garde au cours de la deuxième année du projet était néanmoins important afin d'assurer le maintien des effets du programme pour les enfants qui n'étaient pas inscrits à l'école à plein temps. Le rendement de ces enfants ne fait pas l'objet d'une analyse ciblée, mais contribue néanmoins aux effets du programme rapportés dans les chapitres 9 et 10. Autrement dit, les effets persistants des heures d'exposition au traitement ont été testés dans cette analyse.

⁵² Si l'étude était répétée avec cent échantillons différents, un effet réel de 0,30 écart-type serait détecté dans moins de 50 % des répétitions, présumant un échantillon de 90, des covariables pouvant expliquer 50 % de la variance résiduelle, et un seuil de signification de 0,05. De plus, la stabilité des paramètres estimée serait douteuse dans une telle analyse.

Pour représenter cet aspect de l'exposition des enfants au milieu de garde, nous avons calculé une variable représentant le nombre moyen d'heures par semaine qu'un enfant passe soit dans une garderie programme ou dans une garderie témoin pendant les huit premiers mois de prestation du programme.⁵³ Cette variable a été croisée avec le groupe expérimental dans un modèle de régression prédisant les résultantes des enfants au bout de la deuxième année. En d'autres mots, le groupe expérimental de la première année a été croisé avec le dosage en garderie de la première année pour tenter de prédire les résultantes des enfants pour la deuxième année. Si une année passée dans une garderie programme au cours de la première année se traduit, en moyenne, par de meilleures résultantes qu'une année passée dans une garderie témoin, alors on peut s'attendre à observer un effet bien plus significatif du dosage pour le groupe fréquentant une garderie programme que pour le groupe fréquentant une garderie témoin au cours de la deuxième année, peu importe que l'effet du programme soit persistant ou retardé.

8.3.3. Évaluation de l'effet de la fidélité/qualité du programme en garderie

Le mécanisme par lequel l'intervention influence les résultantes en matière de développement des enfants est postulé se faire par l'entremise de la qualité du programme et de la fidélité de sa prestation. Ainsi, l'appartenance à un groupe expérimental peut se concevoir comme une approximation de la qualité et de la fidélité. La validité interne de l'étude repose fondamentalement sur la véracité de cet énoncé d'où l'examen de la différenciation présenté à la section 9.2.4. Il s'ensuit qu'une condition requise pour observer un effet programme sur les résultantes ciblées est que les enfants inscrits dans les garderies où l'intervention a été mise en œuvre (groupe programme en garderie) vivent des expériences qui se comparent favorablement à celles des enfants faisant partie du groupe témoin (groupe témoin en garderie), qui sont aussi exposés à un programme en mode de garde. De même, à l'intérieur de chaque groupe, la nature du programme offert d'une garderie à l'autre doit être la plus semblable que possible, c'est-à-dire qu'elle doit être cohérente au sein de toutes les garderies du même groupe. En d'autres mots, les garderies doivent être groupées de façon cohérente. Ces deux conditions ont été vérifiées au moyen d'analyses qualitatives présentées au Chapitre 5. Nous rapportons aussi les résultats de tests quantitatifs formels de la différenciation des programmes en garderie au chapitre 9.

L'utilisation des groupes expérimentaux pour estimer l'ampleur des effets du traitement constitue une simplification utile où les dimensions potentiellement continues (p. ex., la fidélité et la qualité) sont réduites à des catégories (c.-à-d. groupe programme en garderie et groupe témoin en garderie), dans le but de faire des comparaisons. Cette simplification est la méthode la plus appropriée lorsque les membres de chaque groupe sont très semblables sur des dimensions continues. Bien entendu, la nature du programme offert dans une garderie varie toujours un peu d'une garderie à l'autre, ce qui signifie que le recours aux groupes expérimentaux aboutit à une perte d'information (c.-à-d. la variabilité intra-groupe). En effet, la distribution des scores de fidélité/qualité des deux groupes peuvent, en principe, se chevaucher.

Nous avons vérifié si cette perte d'information, engendrée par l'emploi des groupes expérimentaux, était importante. Pour ce faire, nous avons procédé à une série d'analyses où la présence d'un intermédiaire, soit le groupe expérimental, a été éliminée et remplacée par des

⁵³ Nous avons exclu les données relatives aux quatre mois précédant la quatrième évaluation puisqu'elles portaient sur des variations attribuables aux perturbations de la période estivale. La mesure la plus pure de l'effet moyen d'exposition au traitement repose sur l'information collectée au cours de l'année scolaire.

indicateurs continus de fidélité et de qualité en tant que prédicteurs des résultantes des enfants. Cette méthode optimise l'utilisation de l'information disponible et elle offre de meilleures possibilités de déceler les effets véritables du programme.

Nous voulions aussi déterminer si certains effets observés du groupe expérimental étaient attribuables au programme lui-même ou à d'autres caractéristiques des garderies programme. Nous avons testé cette idée en estimant les effets du groupe expérimental après avoir contrôlé statistiquement la fidélité et la qualité. Si c'est spécifiquement le programme testé qui est responsable des effets observés, alors les effets ajustés du groupe expérimental résultant de cette analyse ne devraient pas s'avérer statistiquement significatifs.⁵⁴ La logique de cette analyse repose sur un test de médiation (Cohen, Cohen, West, et Aiken, 2003, p. 457) où l'effet de l'appartenance à un groupe expérimental sur les résultantes est présumé s'exercer indirectement par la qualité ou la fidélité du service.⁵⁵ Tout effet résiduel lié au groupe expérimental serait nécessairement attribuable à un autre facteur (p. ex., des différences non contrôlées entre les groupes au début du projet, une autre source de biais, un aspect du programme qui n'est pas capté par les indicateurs de fidélité ou de qualité). L'inclusion des analyses en question dans le présent rapport en augmente substantiellement la longueur, mais elle améliore en revanche notre compréhension des résultats rapportés et notre confiance à leur égard.

Dans les analyses rapportées aux sections 9.2.4 et 9.3.4, chaque type d'indicateur a été opérationnalisé à un niveau général et à un niveau plus détaillé. Deux estimations détaillées de fidélité ont été calculées selon l'adhérence aux éléments spécifiques à la structure du programme et au contenu du programme, respectivement. L'estimation générale de fidélité a été calculée simplement en faisant la moyenne de ces deux indicateurs. En ce qui concerne la qualité proprement dite, nous avons calculé les indices représentant les dimensions structurelles, éducatives, et de sensibilité des éducatrices. Encore une fois, un indice de qualité global a été calculé en faisant la moyenne de ces trois indicateurs de qualité. Un quatrième indicateur a saisi la qualité des activités de littératie en salle de classe (qualité de lecture). Cet indicateur a été traité séparément des autres, en raison de son importance théorique et empirique. Une description plus détaillée de ces indicateurs est donnée à la section 9.1.2.

Les indices de fidélité et de qualité ont été insérés dans les analyses de la façon suivante. Nous avons d'abord mené une série d'analyses où les indices de fidélité et de qualité, plutôt que l'appartenance au groupe expérimental, ont servi d'indicateurs de l'exposition au traitement. Le but de cette analyse était de vérifier si: a) les résultats obtenus en employant la simple définition du traitement seraient reproduits; et b) l'analyse plus sensible révélerait des effets plus marqués, plus persistants, ou les deux. Dans une seconde série d'analyses, nous avons réintroduit l'appartenance au groupe expérimental comme indicateur de l'exposition au traitement, tout en effectuant un contrôle statistique des indices de fidélité et de qualité. Le but de ce deuxième test

⁵⁴ Cet énoncé doit, cependant, être nuancé. La fidélité et la qualité pourraient expliquer la plus grande partie de l'effet du programme sans complètement l'éliminer. En effet, la taille du coefficient représentant l'effet du programme pourrait même augmenter si un effet de suppression est présent (Cohen, et al., 2003, pp. 457–458). L'importance du rôle de la fidélité et de la qualité peut néanmoins être établi avec certitude par le biais d'un test formel de médiation (Frazier, Tix, et Barron, 2004).

⁵⁵ Un test de médiation approprié comporte normalement une série d'analyses de régression (Baron & Kenny, 1986; Frazier, et al., 2004). L'une des conditions essentielles à la démonstration d'une médiation complète est que l'effet direct de la variable X (dans ce cas, le traitement) est éliminé lors du contrôle du médiateur M (dans ce cas, la qualité/fidélité). Une médiation partielle est observée si l'effet direct est réduit mais non complètement éliminé. Notez qu'un effet de médiation peut exister sans une réduction de l'effet direct en raison d'un phénomène nommé la « suppression » (Cohen, et al., 2003, pp. 457–458). L'important est de montrer que l'effet indirect est significatif.

était de vérifier si la façon plus simple de définir le traitement est redondante avec la fidélité/qualité pour expliquer la fluctuation des résultantes pour les enfants. Si les indices de fidélité et de qualité permettent de rendre compte adéquatement du cheminement par lequel l'appartenance au groupe expérimental exerce ses effets sur les résultantes des enfants, nous nous attendrions alors à ce que l'ampleur des estimations des effets du traitement diminue considérablement lorsque la qualité et la fidélité des garderies sont contrôlées dans les analyses ou qu'en présence d'un effet de suppression, l'effet médiateur (effet indirect) de la qualité et la fidélité est significatif. Si l'on obtient un tel résultat, cela fournirait une justification de plus pour attester que les effets du traitement estimés rapportés ici ne sont pas qu'un artefact de la méthodologie, mais plutôt qu'ils reflètent véritablement l'impact du programme en garderie testé.

8.4. CONTRÔLE DE BIAIS DANS LE CADRE D'UNE ÉTUDE LONGITUDINALE

L'estimateur DD neutralise efficacement l'effet statique des différences observées et non observées au sein des groupes expérimentaux à la période de base.⁵⁶ Par définition, ces différences sont éliminées en distinguant initialement le post-test du pré-test. Néanmoins, comme discuté à la section précédente, l'estimateur DD peut être biaisé lorsque la composition des groupes change au fil du temps et quand on soupçonne des différences au pré-test dans les trajectoires développementales des groupes comparés. Nous abordons à tour de rôle les stratégies servant à gérer ces deux types de biais.

8.4.1. Changements dans la composition des groupes

Les changements survenant dans la composition des groupes peuvent biaiser l'estimation DD. Un tel biais peut se produire lorsqu'une variable liée à une résultante varie avec le temps pour un groupe en moyenne. Par exemple, on sait que le sexe est associé à plusieurs résultantes dans la littérature sur le développement. Si la proportion de filles augmente soudainement dans un groupe avant une mesure post-test, les calculs des différences requis pour le calcul de l'estimateur DD seront biaisés.

Cette menace à la validité statistique a été contrôlée de deux manières. D'abord, nous avons surveillé les changements de composition des groupes au fil du temps. Par exemple, nous avons vérifié si le fait de changer de groupe ou de se retirer de l'étude était associé aux variables dépendantes ou à la condition de traitement. Nous n'avons pas observé de telle association (voir l'annexe F). Tout compte fait, nos analyses préliminaires n'ont pas réussi à dégager des problèmes importants concernant les changements de composition des groupes. Nous avons quand même eu recours à une deuxième stratégie consistant à inclure comme covariables toutes les variables de base non redondantes associées à une résultante de façon significative. Cette inclusion des covariables visait à maximiser la validité de l'estimateur DD.⁵⁷ Nous présumons

⁵⁶ L'effet statique d'une variable de base fait référence à une association avec des résultantes qui est stable dans le temps. Un exemple serait le fait que l'effet moyen du sexe est d'une ampleur comparable à toutes les évaluations durant la première année. Si l'effet du sexe était dynamique, sa force d'association avec une résultante varierait considérablement en fonction de l'évaluation considérée.

⁵⁷ Dans ce cas, l'estimateur DD représente l'effet traitement moyen estimé de notre intervention, sous réserve des covariables qui sont incluses dans le modèle. Nous présumons que la condition de traitement ne tempère pas l'effet des covariables.

que l'ensemble des covariables inclus dans notre modèle a contrebalancé tout biais lié à la sélection de participants pour les groupes expérimentaux.

8.4.2. Pentas développementales non-parallèles

Même si la composition des groupes demeure constante à travers le temps, des différences initiales entre les groupes expérimentaux peuvent être associées à des tendances développementales non parallèles. Lorsque présentes, des tendances développementales non parallèles (en l'absence de traitement) compromettent la validité de l'estimateur DD. Si la cause de ce non-parallélisme est observée (c.-à-d. que nous avons collecté les données d'une covariable appropriée), il est possible d'apporter les ajustements statistiques qui corrigeront un tel biais (Abadie, 2005). Un exemple de cet ajustement serait d'inclure, en tant que covariable, l'interaction entre le temps et la variable pertinente. Les variables qu'on pourrait songer à employer pour apporter cet ajustement comprennent les covariables exogènes de base qui sont: a) associées à la mesure de résultantes; et b) réparties différemment dans les groupes comparés. Il serait idéal de traiter toutes les covariables identifiées de cette façon afin de maximiser le contrôle du biais. Cependant, les limites pratiques imposées par la taille de notre échantillon nous obligeaient à être sélectifs dans nos choix des covariables utilisées pour apporter un tel ajustement, afin d'éviter le surajustement des données. Ainsi, nous avons seulement choisi d'estimer des paramètres d'interaction entre le temps et une covariable seulement si l'effet était significatif pour au moins une variable résultante (n'importe quelle résultante mesurée à plusieurs reprises) d'après les données provenant de notre échantillon (voir les tests de stabilité rapportés à l'annexe F au tableau F.8). Comme nous l'avons mentionné au Chapitre 7, les prédicteurs non-significatifs ne requièrent pas nécessairement de contrôle supplémentaire.

8.4.3. Validité des impacts basés sur une variable autre que le groupe

Dans la discussion précédente, l'exposition au traitement ou « dosage » ainsi que la fidélité/qualité de la prestation du programme ont été présentées comme alternatives à la seule appartenance au groupe expérimental pour classer les participants de non traités à traités le long du continuum. Les estimations des effets reposant sur ces définitions plus précises peuvent être plus efficaces, mais elles sont assujetties aux mêmes limites que les estimations reposant seulement sur l'appartenance au groupe expérimental. Nous invoquons ici les postulats qui étayaient la validité de nos estimations de l'effet traitement : que les conditions du traitement, en l'absence de tout nouvel investissement de la part des autres paliers de gouvernement, auraient été comparables.

L'analyse préliminaire visait principalement à établir la comparabilité des groupes expérimentaux. Si nous devons comparer des enfants aux degrés de dosage variables à l'intérieur des groupes, et d'un groupe à l'autre, (p. ex., pour estimer l'effet des heures passées en garderie), alors nous devons présumer que ces enfants sont par ailleurs comparables. Par exemple, nous présumons que les enfants qui passent en moyenne 40 heures par semaine en garderie sont équivalents sur toute autre caractéristique pertinente aux enfants qui passent seulement en moyenne 10 heures par semaine en garderie. Ce postulat a peu de chances d'être respecté, étant donné qu'une diversité de variables sociodémographiques sont potentiellement associées aussi bien au temps passé en garderie qu'aux résultantes. Nous avons aussi présumé que les garderies programme et les garderies témoin auraient été équivalentes au niveau de la fidélité et de la qualité en l'absence d'intervention. Ce postulat n'a pas été vérifié empiriquement (c.-à-d.

qu'aucune mesure pré-test n'a été prise de ces dimensions) et il est peu probable qu'il soit (complètement) respecté étant donné le petit nombre de garderies et l'impossibilité de la mise en correspondance des variables en question.

Nous avons utilisé deux stratégies pour pallier à ces deux sources potentielles de biais dans l'estimation de l'effet du dosage ainsi que de la fidélité et de la qualité du programme. En premier lieu, le devis longitudinal de la présente étude permet de recourir à l'estimateur DD, qui neutralise l'effet statique des variables de base (tant observées que non observées). En second lieu, la spécification finale qui a servi à estimer les effets du traitement comprenait toutes les covariables non redondantes associées à au moins une des mesures des résultantes. En présumant que toutes les variables associées à différents degrés d'exposition à la garderie (ou à la fidélité/qualité) ont été adéquatement prises en compte dans le cadre de cette stratégie (c.-à-d. que le modèle a été correctement spécifié), nous pouvons conclure que les estimations conditionnelles résultant des effets du dosage ne sont pas biaisées. Les mêmes stratégies et postulats ont été retenus pour estimer l'effet de la fidélité au programme et de sa qualité.⁵⁸

8.4.4. Ateliers familles : Un cas spécial

L'analyse des données sur les ateliers d'alphabétisation familiale doit être traitée séparément, car il s'agit d'un cas spécial dans le contexte plus large des analyses qui figurent dans ce rapport. Seulement deux échelles parentales ont servi dans chaque sondage de suivi: la Fréquence des activités de littératie; et la Langue des activités de littératie. Ces échelles ont fait l'objet d'une analyse comparable à celle de l'ÉPE-AD (c.-à-d. par l'entremise de l'estimateur DD). Pour les autres résultantes (c.-à-d. échelles des connaissances, d'auto-efficacité et de modélisation), les données étaient disponibles que pour les parents dont les enfants étaient inscrits dans une garderie programme. Nous appellerons dorénavant ce groupe d'échelles « les échelles des ateliers pour parents ». Pour l'analyse de ces échelles, nous avons adopté une stratégie quelque peu différente que nous décrivons dans les sections ci-dessous.

Échelles des ateliers pour parents : Choix de l'échantillon

L'une des premières distinctions de l'analyse de ces sous-échelles parentales est qu'elle repose sur l'échantillon total de participants recrutés pour le projet et dont les enfants étaient inscrits dans une garderie programme au moment où les ateliers ont été offerts (N = 116, provenant de six communautés et deux cohortes). Contrairement à l'analyse des résultantes portant sur les enfants, les communautés d'Edmonton et de Saint-Jean n'ont pas été exclues de cette analyse. Cette décision s'appuyait sur trois raisons. D'abord, contrairement à l'intervention en garderie, les Ateliers familles ont été correctement mis en œuvre dans toutes les communautés (voir le *Rapport de mise en œuvre du projet*). Ensuite, étant donné que l'analyse vise seulement les parents dont les enfants étaient inscrits à l'une des garderies programme, la disponibilité d'un groupe témoin en garderie dans toutes les communautés ne pose pas de problème. Enfin, l'exclusion des communautés de Saint-Jean et Edmonton aurait réduit l'échantillon à un niveau ne permettant pas d'évaluer correctement les effets des ateliers. Les caractéristiques de l'échantillon total des familles du groupe programme en garderie ne sont pas rapportées ici (voir

⁵⁸ Toutes les covariables sont au niveau du « participant » plutôt que de la garderie, sauf pour la variable de la communauté. Le nombre de garderies était insuffisant pour permettre d'inclure une large palette de caractéristiques à ce niveau d'analyse.

le Chapitre 3 pour les taux de réponses), mais elles font l'objet d'une description détaillée dans le *Rapport de référence*.

Échelles des ateliers pour parents : Stratégie d'analyse

L'analyse des échelles des ateliers pour parents se distingue également de part l'absence d'un groupe témoin planifié. Les mesures ont été administrées qu'aux parents dont les enfants étaient inscrits à l'une des garderies offrant le programme testé et il n'existe pas de données pour certaines mesures dans le cas des familles ayant participé à moins de trois ateliers ($n = 14$). Un devis équilibré serait doté d'un ensemble complet de mesures pré-test et post-test pour toutes les variables et pour tous les participants, mais ce n'est pas le cas ici. Par souci d'efficacité, un devis non équilibré a été employé, ce qui a nécessité une méthode sélective de collecte de données.

Spécifiquement, les familles classées comme « non participantes » n'ont eu qu'à fournir des estimations post-test de leurs attitudes, ce qui signifie que les estimations rétrospectives (Lamb & Tschillard, 2005; Rockwell & Kohn, 1989) ne sont pas disponibles pour ce groupe. Un ensemble complet de données post-test et de données rétrospectives de pré-test étaient disponibles pour les échelles des connaissances et d'auto-efficacité des familles classées comme « participantes ». Les estimations réelles pré-test et post-test ont été recueillies pour l'échelle de modélisation de toutes les familles, peu importe la participation. L'approche analytique utilisée dans les analyses a été adaptée aux complications présentées par ce devis.

Étant donné la structure de l'ensemble des données, nous disposons d'au moins deux façons d'estimer l'impact des ateliers. La première méthode estime l'impact des ateliers en fonction de la différence entre les mesures pré-test et post-test. La deuxième façon d'estimer l'impact des ateliers consiste à comparer les parents classés comme participants ($n=116$) à ceux qui ne l'étaient pas ($n=14$). Chaque type de comparaison en lui-même manque de validité, mais le patron global produit par une série de tests peut être informatif.

La validité des scores de changement (différence entre les mesures pré-test et post-test pour les participants uniquement) est douteuse, non parce que ces derniers sont rétrospectifs (pour une discussion sur les avantages des estimations rétrospectives par rapport aux pré-tests traditionnels, voir Lamb et Tschillard, 2003; 2005), mais plutôt en raison de l'absence d'un groupe témoin servant à contrôler pour les effets de maturation et historiques. Ce problème de validité est exacerbé par le fait que les participants, conscients d'être traités, pourraient produire des estimations de leur état pré-test et post-test conformes à leurs attentes de l'effet positif des ateliers.⁵⁹ Dans le même ordre d'idées, la comparaison de participants et de non-participants est invalidée par la possibilité évidente d'une sélection biaisée qui ne peut être contrôlée, ni statistiquement ni au moyen de l'appariement en raison du nombre trop faible de participants au sein du groupe de non-participants.

Individuellement, les deux estimations de l'effet traitement sont biaisées, mais ensemble elles peuvent produire des résultats informatifs sans écarter complètement toutes les sources de biais.

⁵⁹ Ce type de biais peut prendre plusieurs formes et n'est pas propre aux pré-tests rétrospectifs. Les parents peuvent exagérer l'effet programme en sous-estimant leur état pré-test et en surestimant leur état post-test, ou en produisant des estimations contaminées des deux manières. L'expérience a montré qu'une sous-estimation de l'état pré-test est caractéristique des mesures rétrospectives (Taylor, Russ-Eft & Taylor, 2009). Les mesures pré-test prospectives traditionnelles sont également biaisées d'autres façons (p. ex. biais de réponses améliorées; Howard, 1980), ce qui entraîne parfois la surestimation de l'état pré-test (Moore & Tananis, 2009).

Par exemple, nous soutenons que le patron de résultats suivant correspond à un véritable effet traitement: a) l'effet traitement basé sur le score de changement pour les participants est statistiquement significatif et positif; b) l'estimation pré-test rétrospective pour les participants est équivalente à l'estimation « post-test » fournie par les non-participants; et c) les participants rapportent des estimations post-ateliers supérieures à celle des non-participants en moyenne. Nous présumons qu'en l'absence d'une participation au programme, l'estimation « post-test » des non-participants est aussi valide qu'une estimation de leur état pré-test. En d'autres mots, nous postulons que leur état véritable, en moyenne, n'a pas changé durant la période où les ateliers étaient offerts. Si le patron de résultats escompté est obtenu, cela permet de supposer que les ateliers ont eu un effet positif. Si seulement le résultat (b) est obtenu, alors on peut supposer que l'effet du traitement est nul (ou négatif). Si les résultats (a) et (c) sont obtenus sans le résultat (b), alors on peut fortement supposer que les estimations sont biaisées et que l'interprétation des effets positifs du programme doit être nuancée en conséquence. Autrement dit, il y a un patron spécifique de résultats qui appuie fortement l'impact positif réel des ateliers d'alphabétisation familiale, et bon nombre d'autres patrons qui ne procurent pas cet appui. Si notre prédiction « risquée » se confirme, elle prête de la crédibilité à la validité de notre interprétation quant aux effets du traitement.

Les limites imposées par les données disponibles signifient qu'un estimateur comme l'estimateur DD doit être écarté pour les variables n'ayant pas de mesure pré-test pour les deux groupes. Les trois hypothèses formulées ci-dessus peuvent plutôt être testées à l'aide d'une série de tests indépendants, un pour chaque hypothèse. Les covariables n'ont pas pu servir à tester les différences entre les groupes, car la trop petite taille d'échantillon du groupe de non-participants rendait la chose impossible ($n = 15$). De plus, les covariables n'ont pas servi à estimer les scores de changement des participants parce que: a) l'effet statique des caractéristiques des participants est neutralisé (« c'est-à-dire éliminé ou absent des différences ») au cours de cette estimation; et b) les changements dans la composition des groupes ne posent pas un problème.

Les estimations de fidélité, de la qualité et du dosage du programme (proportion d'ateliers fréquentés) étaient disponibles pour l'échantillon de parents qui ont assisté aux ateliers ($n = 114$). Nous avons évalué l'impact de ces facteurs en utilisant un estimateur DD comme pour d'autres analyses rapportées ici. Les estimations produites représentent les impacts des ateliers quant à l'intensité du traitement reçu. Encore une fois, comme les changements dans la composition des groupes ne constituaient pas un problème pour cette analyse, nous avons compté sur le fait que l'estimateur DD n'était pas biaisé par les caractéristiques de base observées et non observées.

8.4.5. Tests de modération et de médiation

Le lecteur notera que la stratégie d'analyse adoptée pour ce rapport s'appuie sur plusieurs analyses secondaires. Nous testons si le programme a eu un impact, mais aussi si cet impact dépend de la qualité/fidélité, le dosage et du profil linguistique. L'apport de chaque analyse est limité par la taille modeste de l'échantillon, mais ensemble elles peuvent, si elles produisent un patron de résultats cohérent, écarter plusieurs explications alternatives des résultats. Elles peuvent attester que le mécanisme ayant généré les résultats obtenus est bel et bien compris, ce qui nous permet d'anticiper certaines contingences pour la réalisation de l'effet du programme.

8.5. PLAN DES CHAPITRES À SUIVRE

Les chapitres à suivre sont organisés de la façon suivante. Une évaluation de l'impact du programme testé sur les enfants est présentée au chapitre 9. Ce chapitre présente une comparaison des groupes expérimentaux ainsi qu'une série d'analyses plus approfondies. Nous rapportons des analyses examinant comment l'effet du programme testé dépend de facteurs comme le dosage, la fidélité et la qualité de la mise en œuvre du programme en garderie et du profil linguistique des enfants. Au chapitre 10, l'impact du volet Ateliers familles sur le comportement et les attitudes des parents est évalué. Nous rapportons les résultats d'analyses examinant comment l'effet des ateliers familles sur les parents dépend de facteurs comme le dosage et la fidélité/qualité de sa mise en œuvre. De plus, ce chapitre présente le résultat d'une analyse examinant l'effet indirect des ateliers sur les enfants. Cette dernière analyse fait le lien entre les changements observés chez les parents et le développement des enfants.

9. Impact du programme testé sur les enfants

Les deux volets centraux de l'intervention étaient le nouveau programme en garderie et les Ateliers familles. Le but principal de la première composante était d'influencer *directement* les résultantes des enfants, tandis que celui de la seconde composante était d'influencer *indirectement* les résultantes des enfants en modifiant les attitudes et les comportements des parents. Les analyses rapportées dans ce chapitre ne peuvent pas nettement distinguer l'effet d'un volet de l'autre. Un plan expérimental plus complexe aurait été nécessaire pour permettre de faire cette distinction. En conséquence, les analyses principales comparant les groupes expérimentaux sont des tests de l'effet du programme à deux volets, soit l'*effet combiné* des deux volets du programme sur le développement des enfants. C'est le sujet principal du présent chapitre.

Il existe néanmoins la possibilité d'effectuer des analyses donnant une idée de l'importance relative des deux volets. Pour le volet garderie, les analyses portant sur le dosage et la qualité/fidélité de la prestation du programme permet de quantifier l'importance de sa contribution aux effets programme. Plus l'effet de ces variables sur le développement des enfants est grand, plus la contribution du volet programme en garderie sur les résultantes des enfants lors de la comparaison des groupes expérimentaux est importante. Les résultats de ces analyses secondaires sont présentés dans le présent chapitre. Pour le volet Atelier familles, le développement des enfants est mis en relief avec les changements rapportés par les parents. Cette deuxième série d'analyses est examinée au Chapitre 10 suite aux résultats des analyses d'impact des Ateliers familles sur les parents.

Pour ce qui est du présent chapitre, les données provenant des enfants lors de la première et de la deuxième année de la mise en œuvre du programme ont été analysées séparément. Nous avons effectué les analyses ainsi en raison d'une insuffisance de degrés de liberté pour tenir compte de toutes les évaluations au sein d'un même modèle de régression.⁶⁰ De plus, l'entrée de la plupart des enfants en maternelle au début de la deuxième année du programme constituait un point de rupture conceptuel naturel. Le lecteur intéressé trouvera l'ensemble des résultats pour les deux années présenté graphiquement à la section 9.3.

Une description des variables utilisées aux fins des analyses d'impact est donnée à la section 9.1. Les analyses portant sur la première année sont présentées à la section 9.2 et les analyses portant sur de la deuxième année sont présentées à la section 9.3. Les détails techniques propre à chaque analyse sont présentés aux sections 9.2.1 et 9.3.1. Pour les deux années du projet, nous rapportons une série d'analyses complémentaires.

L'impact du programme testé est évalué: (a) en comparant les trois groupes expérimentaux (voir les sections 9.2.2 et 9.3.2); (b) en évaluant si l'effet du programme varie en fonction du dosage (voir les sections 9.2.3 et 9.3.3); (c) en testant l'effet direct de la fidélité du programme et de sa qualité sur le développement des enfants fréquentant une garderie (voir les sections 9.2.4 et

⁶⁰Le nombre de garderies détermine le nombre de degrés de liberté (environ 20) pour les tests de signification. L'inclusion de toutes les mesures répétées dans la même analyse porterait les degrés de liberté sous la barre des dix degrés pour certains tests de signification, ce qui veut dire qu'un test équitable du programme ne serait pas possible.

9.3.4); (d) en testant si la fidélité du programme et sa qualité expliquent en tout ou en partie les différences observées entre le groupe programme et le groupe témoin en garderie (voir les sections 9.2.4 et 9.3.4); et e) en réévaluant les différences entre les groupes expérimentaux, d'une part, pour les enfants dont l'exposition au français est élevée et, d'autre part, pour ceux dont elle est faible (voir les sections 9.2.5 et 9.3.5). L'objet de ces dernières analyses est de déterminer les sous-populations pour lesquelles le programme de garderie testé semble particulièrement efficace.

Le lecteur est prié de noter les points suivants :

- Les analyses d'impact rapportées dans ce chapitre portent sur quatre communautés, soit Orléans (ON), Cornwall (ON), Durham (ON) et Edmundston (NB). Les communautés d'Edmonton (AL) et de Saint-Jean (NB) sont exclues de l'analyse à cause de considérations méthodologiques : dans le cas d'Edmonton, le programme en garderie n'a presque pas été bien mis en place; le groupe témoin en garderie n'était pas représenté à Saint-Jean. L'exclusion de ces communautés permet de rapporter un ensemble de résultats valide et cohérent.
- La première cohorte de participants comprend des familles provenant des quatre communautés, tandis que la deuxième cohorte comprend exclusivement des familles d'Orléans et de Cornwall.

9.1. TRAITEMENT DES VARIABLES

Parmi les nombreuses variables mesurées au cours de l'étude, les suivantes ont été retenues pour les analyses d'impact de la première année et de la deuxième année. Ces variables peuvent être divisées en deux grandes catégories, soit les mesures de résultantes et de prédicteurs. Nous décrivons maintenant comment ces deux types de variables étaient traités dans le cadre des analyses d'impact.

9.1.1. Mesures de résultantes

Première année

Ce rapport d'impact concerne la préparation à la scolarisation en langue française. Pour cette raison, nous avons mesuré les domaines relatifs à la préparation à la scolarisation au moyen des échelles de l'ÉPE-AD (version française seulement) suivantes : Langage et communication (Communication), Conscience de soi et de l'environnement (Conscience de soi), Habiletés cognitives (Cognition) et Physique/moteur (Physique). Les données complètes étaient seulement disponibles pour l'échelle Communication.⁶¹ Nous avons aussi créé des sous-échelles de vocabulaire plus précises en choisissant et en combinant des items pertinents compris dans les

⁶¹ Pour les trois autres échelles, les données n'étaient disponibles que pour les participants qui avaient complété l'ÉPE-AD en français au moins deux fois au cours des cinq premières évaluations. Pour des raisons évidentes, cela signifie que les analyses menées avec les scores des échelles Conscience de soi et de l'environnement, Habiletés cognitives, Physique/moteur, Vocabulaire réceptif et Vocabulaire expressif sont des analyses de sous-échantillons efficaces fondées sur un sous-groupe excluant les enfants dont les compétences langagières en français sont les plus faibles. Cette exclusion n'était pas fortement liée aux groupes, ce qui signifie que la validité interne des estimations sur les effets programme n'est pas compromise. En quelque sorte, cette question affecte la validité externe en limitant l'application générale des résultats aux enfants qui répondaient aux critères établis pour la prise du test en français.

quatre échelles. En créant ces échelles, nous avons fait la distinction bien documentée entre le vocabulaire réceptif (échelle de Vocabulaire réceptif) et expressif (échelle de Vocabulaire expressif). En tout, six résultantes des enfants mesurées par l'ÉPE-AD ont fait l'objet des analyses qui suivent. Notez que les enfants qui se sont retirés du projet avant la fin de la première année sont exclus de ces analyses.

Deuxième année

En tout, six résultantes des enfants ont été prises en compte pour les analyses d'impact de la deuxième année notamment: Communication, Conscience de soi, Cognition et Vocabulaire expressif mesurées par l'ÉPE-AD. De plus les analyses ont porté sur les résultantes couramment utilisées dans les études sur la préparation à la scolarisation dont le vocabulaire réceptif (mesuré au moyen de l'Échelle de vocabulaire en images de Peabody ou ÉVIP-R) et le vocabulaire expressif (mesuré au moyen de l'Épreuve de dénomination de Gardner, la traduction validée et normalisée en français pour les enfants de l'outil *Expressive One-Word Picture Vocabulary Test* ou EOWPVT). Les deux dernières mesures ont été administrées respectivement à la sixième et à la septième évaluation. Notez que les échelles Physique et Vocabulaire réceptif ne figurent pas dans les analyses de la deuxième année⁶² et que les enfants qui se sont retirés du projet avant la fin de la deuxième année sont exclus des analyses.

Standardisation des résultantes

Avant de procéder aux analyses, nous avons recodé les résultantes sur une même échelle de scores afin qu'elles soient standardisées pour chaque période de temps, faisant en sorte que chaque résultante ait une moyenne globale de 0 et un écart-type de 1 pour l'échantillon total de participants. Une conséquence pratique de cette transformation est que les scores pour chaque période de temps peuvent être interprétés en termes d'échelle commune. La décision de standardiser les scores s'appuyait sur le fait que la nature des scores d'échelle bruts a changé sur le plan qualitatif durant la deuxième année de l'étude. Plus précisément, la composition des échelles a été altérée (c.-à-d. que des items ont été éliminés) et par conséquent le nombre total d'items a été modifié. La forte corrélation entre les échelles originales et les échelles révisées appuie notre thèse voulant que très peu d'informations aient été perdues suite à ces modifications. Néanmoins, les différences d'échelle (p. ex. score maximal) de nos mesures au fil du temps posaient un obstacle technique aux analyses statistiques : la signification des valeurs absolues des scores varie avec le temps, ce qui invalide toute comparaison effectuée d'une évaluation à l'autre. Nous avons relevé ce défi technique mineur en recourant à la procédure de standardisation discutée ci-dessus. Cette dernière assure que les scores des résultantes ont la même signification au fil du temps.

Interprétation des scores standardisés

Les scores standardisés employés dans les analyses d'impact sont interprétés comme suit. Le score standardisé de chaque participant représente la différence entre le score du participant et le score moyen de l'échantillon. De façon éclairée, cette différence est exprimée en unités d'écart-type. Par exemple, un score de 1,11 signifie que les participants ont obtenu un score qui était à

⁶² Les données pour les échelles Physique et la sous-échelle de Vocabulaire réceptif de l'ÉPE-AD étaient disponibles pour les quatrième et cinquième évaluations, mais les analyses reposant sur ces mesures sont rapportées uniquement dans le cadre des analyses des données collectées lors de la première année (voir la section 9.2).

1,11 écart-type au-dessus du score moyen obtenu par les participants de l'échantillon du projet Capacité d'apprentissage pour l'évaluation correspondante. Les estimations DD des effets programme sont interprétées de manière semblable comme des différences entre les groupes en unités d'écart-type.

Une différence « standardisée » entre groupes est la façon la plus couramment utilisée d'exprimer la taille d'un effet. Suivant Cohen (1988), nous appelons cette statistique d . Cohen fournit des références conventionnelles pour interpréter l'ampleur des effets exprimés dans une échelle standardisée. Une différence standardisée entre les groupes de $d = 0,20$ est considérée comme petite, une différence de $d = 0,50$ est considérée comme moyenne et une différence de $d = 0,80$ est considérée grande. Ces points de références concordent avec les résultats d'une méta-analyse examinant la distribution des tailles d'effets standardisés de diverses études d'intervention (Lipsey & Wilson, 1993). Ils doivent cependant seulement servir de guide général pour juger de l'importance d'un effet. Un effet peut être considéré plus ou moins important selon le contexte de recherche (Kane, 2004; Hill, Bloom, Black, & Lipsey, 2008).

Selon Kane (2004) et Hill et ses collègues (2008), une façon utile de comprendre l'importance de l'effet d'une intervention visant la petite enfance est de le comparer avec l'effet du développement normal, c'est-à-dire de poser la question : comment la taille de l'effet se compare-t-elle avec les gains normaux observés lors d'un an de développement? Selon Hill et ses collègues (2008; tableau 1), le gain moyen attendu dans le développement de la littératie et de la numératie pour la période de la maternelle à la première année est d'environ $d = 1,33$. En d'autres mots, un effet de programme de $d = 1,33$ représenterait un an de développement et un effet de $0,67$ représenterait un gain d'environ six mois de développement.

Évaluation des conséquences de la standardisation

L'utilisation de scores standardisés ne nous permettra pas de : a) comparer les scores bruts de notre échantillon à ceux d'une population normative; et b) évaluer la trajectoire développementale de notre échantillon au fil du temps.⁶³ Nous faisons valoir que les conséquences de la standardisation sont négligeables parce que: a) aucun échantillon normatif n'existe pour la version de l'ÉPE-AD utilisée dans le cadre du projet Capacité d'apprentissage; et b) les enfants de l'étude ont été évalués à plusieurs reprises, ce qui a sans doute induit des effets d'entraînement massifs (c.-à-d. de meilleurs scores suite à la répétition d'une tâche).⁶⁴ Ces effets d'entraînement sont nécessairement confondus avec toute tentative d'estimer le développement de notre échantillon au fil du temps. Autrement dit, la signification intrinsèque de la valeur absolue des scores bruts observés est limitée, alors que les changements de scores au fil du temps (c.-à-d. la trajectoire développementale de notre échantillon) ne peuvent être différenciés dans un cas ou l'autre des effets d'entraînement. Fait important : l'effet d'entraînement est équivalent pour les trois groupes expérimentaux puisque le nombre d'évaluations y est constant. Ceci nous permet d'analyser les données normalisées selon la

⁶³ La standardisation des scores donne une moyenne de 0. La standardisation à l'intérieur de chaque période de temps donne une moyenne de 0 à chaque période de temps. Si le score moyen est de 0 à chaque période de temps, alors la moyenne est constante à travers le temps et l'effet du « Temps » dans un modèle de régression sera nécessairement nul. Lorsque l'ensemble des gains développementaux pour l'échantillon de participants n'est pas d'intérêt direct, comme c'est le cas pour la présente étude, cela représente une perte d'information négligeable.

⁶⁴ S'il est impossible d'évaluer la trajectoire globale de l'échantillon, il demeure tout de même possible de comparer les trajectoires des groupes dans le cadre d'une analyse DD.

méthode DD, comparant les trajectoires développementales relatives de nos trois groupes expérimentaux.^{65, 66}

9.1.2. Variables cibles et covariables

Plusieurs variables ont été incluses dans les équations de régression servant à modéliser les résultantes pour les enfants en matière de préparation à la scolarisation. La plus grande distinction que l'on puisse faire ici est celle qui existe entre les variables cibles et les variables qui servent uniquement de covariables. Les variables cibles font l'objet d'un intérêt particulier en matière de recherche, alors que les covariables servent simplement à améliorer la validité interne des tests plus importants. Dans ce qui suit, nous présentons un survol des variables cibles incluses dans les analyses d'impact ainsi que les covariables utilisées dans le cadre des modèles ajustés.

Variables cibles

La liste de variables cibles inclut le temps (période d'évaluation), le groupe auquel appartient l'enfant, la moyenne d'heures que l'enfant passe en garderie à chaque semaine (le dosage), la qualité/fidélité des services de garde, et la composition linguistique du foyer. Chacune de ces variables sert à dégager un aspect ou un autre de l'effet du programme testé, soit individuellement ou en combinaison. La spécification exacte utilisée afin de modéliser l'effet de ces variables peut varier. Ces détails sont propres à chaque analyse et sont spécifiés dans le texte.

Le temps

Le temps est une variable catégorielle qui représente la période d'évaluation à laquelle une observation a été collectée. Cette variable sert à identifier l'effet dynamique des covariables ainsi que l'effet du programme testé par le biais de l'estimateur DD.

Le groupe

Le groupe est une variable catégorielle représentant l'appartenance des enfants et de leur famille à l'un des groupes expérimentaux. Ils appartiennent soit au groupe programme, au groupe témoin en garderie ou au groupe témoin hors garderie. Dans les tableaux, ces groupes sont parfois désignés respectivement par les étiquettes G1, G2 et G3. Les règles utilisées pour déterminer l'appartenance à un groupe d'une période d'évaluation à l'autre sont décrites à la section 8.3.2.

Cette variable est essentielle pour dégager les retombées du programme. Elle est souvent combinée avec d'autres variables pour rehausser la validité de l'estimation (p. ex. dans le cas de l'estimateur DD) ou pour tester l'effet d'un modérateur potentiel (p. ex. le dosage).

⁶⁵ Même si la moyenne de l'échantillon est de zéro, les moyennes des groupes ne le sont pas. Elles varient librement d'une évaluation à l'autre. Il est donc possible de les comparer en utilisant des analyses statistiques.

⁶⁶ La standardisation rend obscure certaines caractéristiques de la distribution des scores bruts. Nous tenons à souligner que les analyses préliminaires indiquaient que toutes les mesures des résultantes pour les enfants étaient sensibles au changement au fil du temps. De plus, aucune des mesures n'avait plafonné à la 5^e évaluation.

Le dosage

Le dosage est une variable à échelle continue représentant le nombre d'heures moyen par semaine qu'un enfant passe en garderie. La valeur prise par cette variable peut varier au fil du temps (les moyennes pour l'échantillon à différentes périodes sont rapportées à l'annexe F au tableau F.6) ainsi que la méthode utilisée pour la calculer. Lors de la première année, le dosage est calculé à partir de l'assiduité des enfants à la garderie dans les mois qui précèdent l'évaluation. Lors de la deuxième année, le dosage représente la moyenne du dosage lors de la première année du projet. Plus de détails à ce sujet sont fournis à la section 8.3.2. L'intérêt d'inclure le dosage dans les analyses est de vérifier si cette variable exerce un contrôle sur la taille de l'effet du programme.

La fidélité et la qualité

L'étude de la mise en œuvre a fourni des informations portant sur les programmes offerts en garderie lorsque les enfants complétaient la première année du projet. Ces informations ont servi au calcul d'indices de la fidélité et de la qualité du programme pour les garderies du groupe programme et pour celles du groupe témoin. On a attribué au programme offert dans chaque garderie des scores sur de multiples indices de fidélité et de qualité. Deux séries de scores distinctes sont attribuées aux garderies qui ont accueilli les deux cohortes d'enfants. Chaque série représente le programme tel qu'offert aux enfants lors de leur première année de participation.

Il est à noter que pour les fins des analyses d'impact, nous ne distinguons pas entre les éducatrices d'une même garderie. On attribue aux garderies ayant plus d'une éducatrice un score reflétant la moyenne des scores attribués aux éducatrices. Ce regroupement des données simplifie la structure de la base de données pour les analyses d'impact.

Deux indices de fidélité sont retenus comme descripteurs des garderies, l'un portant sur la *fidélité structurelle* de la mise en œuvre et l'autre concernant la *fidélité de contenu*. Les deux indices expriment la proportion d'éléments du programme testé en place dans chaque garderie. La fidélité structurelle reflète la présence d'éléments dans l'environnement comme les affiches image-mot ou encore les routines imagées. Pour sa part, la fidélité de contenu indique le niveau d'intégration des éléments du programme dans sa programmation en milieu de garde. Par exemple, le cercle d'amis peut intégrer des activités de lecture ou de raisonnement. La procédure utilisée pour calculer ces deux indices est détaillée à la section 8.3.3

Quatre indices de qualité sont également retenus comme descripteurs du programme et de l'environnement dans lequel il prend place. Ces indices expriment le niveau de qualité du programme offert dans les garderies sur une échelle de 7 points où « 1 » correspond à des soins se situant bien en deçà des exigences élémentaires d'un service de garde et « 7 » à des soins personnalisés et de qualité supérieure (Harms et al., 1998). Ces mesures de la qualité proviennent de l'Échelle d'évaluation de l'environnement préscolaire — Révisée (ÉÉEP-R). Aux fins des analyses d'impact, certaines sous-échelles de l'ÉÉEP-R ont été regroupées pour former quatre indices de qualité d'intérêt. Le premier indice, *qualité structurelle*, mesure la qualité globale de l'environnement des milieux de garde. Il regroupe les sous-échelles Espace intérieur, Mesure d'hygiène et Accueil et départ. Le deuxième indice, *qualité éducative*, porte une attention particulière aux activités permettant d'encourager la communication chez les enfants et d'enrichir leur vocabulaire. Cet indice regroupe les sous-échelles de l'ÉÉEP-R Utilisation du langage spontané, Soutien à la communication chez les enfants, et Utilisation du langage pour

développer les habiletés de raisonnement. Le troisième indice, la sous-échelle Interaction personnel-enfant, a été utilisé comme indice de *sensibilité* de l'éducatrice afin de tenir compte de cette influence particulièrement importante pour le développement de l'enfant. Enfin, *qualité de lecture* isole la sous-échelle Livres et images, étant donnée l'importance toute particulière de la lecture pour l'acquisition du vocabulaire chez les enfants.

En tout, six indices ont servi à estimer le rôle que joue le programme en garderie dans le développement des enfants. Dépendamment des objectifs d'une analyse en particulier, ces indices sont traités individuellement ou ils sont combinés pour former des indexes globaux de fidélité et de qualité. Les indexes globaux sont calculés en prenant une simple moyenne des indices afin d'attribuer une pondération égale à chaque indice.

Dans tout les cas, ces variables ont servi à vérifier l'hypothèse selon laquelle les indices de qualité et de fidélité captent le mécanisme par lequel le programme en garderie exerce son effet sur le développement des enfants. L'effet de ces indices est estimé de la même façon que l'effet du groupe : par le biais de l'estimateur DD.

Le type de foyer

La question de comment définir le profil linguistique des enfants est traitée en détails à la section 5.1.3 du *Rapport des résultats de la première cohorte* (2014). La stratégie adoptée dans le *Rapport des résultats de la première cohorte* était de rapporter des résultats pour plusieurs indicateurs du profil linguistique des enfants. Ces indicateurs variaient selon l'unité de mesure (continue ou catégorielle) et leur définition de l'exposition au français (exposition active versus exposition passive).

Le présent rapport comprend les résultats provenant d'un seul indicateur du profil linguistique : le type de foyer. Cette variable dichotomique contraste les familles à forte et à faible exposition au français. Nous définissons l'exposition au français sur la base de la langue parlée à l'enfant par la mère croisée avec la langue parlée par le père.⁶⁷ Ce choix est justifié par le fait qu'une variable catégorielle génère des résultats plus faciles à comprendre dans le cadre d'une analyse de modération qu'une variable continue.⁶⁸ Les familles à faible et à forte exposition peuvent se concevoir comme deux points sur un continuum.

Les analyses rapportées dans le *Rapport des résultats de la première cohorte* (2014) tendaient à confirmer l'hypothèse selon laquelle les bénéficiaires du programme testé se faisaient ressentir surtout au niveau des résultantes langagières (p. ex. le vocabulaire) pour les enfants caractérisés par une plus faible exposition au français tandis que les enfants caractérisés par une plus forte exposition au français bénéficiaient principalement d'un développement accéléré de leurs compétences cognitives (l'échelle de Cognition). Cette hypothèse se base sur un modèle du développement reconnaissant que la maîtrise de la langue d'enseignement facilite l'acquisition de compétences académiques telles que ciblées par l'échelle Cognition. Selon cette

⁶⁷ Le *Rapport des résultats de la première cohorte* (2014) considère aussi une définition alternative du type de foyer se basant sur la première langue officielle parlée des parents. Nous avons argumenté que cette définition ne captait pas adéquatement l'expérience langagière de l'enfant au foyer.

⁶⁸ Il est vrai que les variables continues comme le Continuum de français parlé par l'enfant sont susceptibles d'être plus efficaces pour détecter les effets réels du programme. Ainsi, l'effet modérateur du profil linguistique est parfois significatif avec cette variable continue et non avec le type de foyer (analyses non-rapportées). Ceci dit, l'ensemble des résultats n'est pas clarifié lorsqu'on utilise cette variable continue. L'analyse plus simple se basant sur le type de foyer communique essentiellement les mêmes informations plus limpiment.

conceptualisation, le programme testé a un impact positif sur le développement de tous les enfants, mais cet impact varie en fonction de leur état de préparation individuel.

Cette analyse est répétée dans le présent rapport afin de vérifier si l'inclusion de la deuxième cohorte affecte en quelque sorte les résultats antérieurs. L'hypothèse d'un effet distinct du programme en fonction de l'état de préparation des enfants est testée en juxtaposant les facteurs de temps, de groupe et de type de foyer pour déterminer si leur interaction est significative (par le biais de l'estimateur DDD). Cette façon de procéder permet d'estimer l'effet du programme indépendamment pour les enfants issus de familles à forte exposition et ceux issus de familles à faible exposition.

Covariables de la première année

Pour les analyses de la première année, les effets statique et dynamique de 12 covariables sont représentés dans la spécification des modèles ajustés. L'effet statique est capté par l'effet direct des variables tandis que l'effet dynamique est capté par leur interaction avec le facteur de temps (c.-à-d. les variables fictives représentant chaque période d'évaluation). Pour plus de détails, voir la section 8.4 du présent rapport.

Covariables de la deuxième année

Les mêmes covariables figurent dans les modèles ajustés de la deuxième année. À cette liste, une 13^{ème} variable est ajoutée pour représenter l'inscription à l'école. Rappelons qu'au début de la deuxième année du projet, certains enfants ont été inscrits à l'école à temps plein, d'autres ont été inscrits à temps partiel, et enfin, d'autres n'étaient pas inscrits à l'école. Une variable catégorique représentant ces différents degrés d'exposition à l'école a été insérée comme covariable dans nos analyses de régression. Cette variable est codée comme 0 pour la période de référence et comme 0 ou 1 à compter de la quatrième évaluation pour les enfants inscrits à l'école; les valeurs attribuées à chaque enfant sont fixes pour toute la deuxième année du projet, soit de la quatrième à la septième période d'évaluation.

9.1.3. Contre-validation non-paramétrique

Les résultats rapportés dans le présent rapport sont issus d'analyses paramétriques ayant comme objectif de tester des différences moyennes entre les groupes. Les analyses paramétriques sont très puissantes pour détecter des effets (un attribut très attrayant pour l'analyse de données provenant d'un petit échantillon). Par contre, elles sont sujettes aux critiques puisque leur validité repose sur plusieurs postulats plus ou moins plausibles et difficiles à établir empiriquement. Pour répondre à ces critiques possibles, nous avons vérifié la robustesse des résultats en estimant des modèles de régression ordinaire logistique (utilisant la fonction « logit ») avec une erreur type robuste pour toutes les résultantes considérées dans ce chapitre. Afin d'alléger la présentation, nous présentons les résultats des analyses non-paramétriques seulement pour les modèles ajustés (voir section 9.2.2).

Les analyses non-paramétriques servent à confirmer si oui ou non l'effet du programme est significatif pour un effet DD donné. Cette interprétation des analyses est donnée dans le texte. Le lecteur est prié de noter les détails techniques suivants. L'analyse par régression logistique porte sur une version simplifiée des résultantes créées en les transformant en variables ordinales (c.-à-d. les scores ont été classifiés en cinq intervalles, des quintiles). Le programme estime quatre

paramètres représentant les chances de se trouver dans un quintile supérieur : i) supérieur au premier quintile, ii) supérieur au deuxième quintile, iii) supérieur au troisième quintile, iv) supérieur au quatrième quintile.⁶⁹ Ces paramètres sont agglomérés pour former un seul indice global représentant les chances de se trouver dans le quintile supérieur. Cet indice global (une cote) est estimé pour les trois groupes expérimentaux, et ce, pour chaque évaluation.

Une analyse DD produit des statistiques de « rapport des cotes » qui servent à évaluer l'impact du programme sur l'indice global représentant les chances de se trouver dans un quintile supérieur. Si aucune différence n'existe entre les groupes, le rapport des cotes prend la valeur de 1. Si l'intervalle de confiance pour le rapport des cotes ne chevauche pas la valeur de 1, nous pouvons conclure que l'effet est significatif. Les intervalles de confiance à 90 % sont rapportés, ce qui correspond à un seuil de signification de 10 %. En d'autres mots, on postule que l'intervalle estimée comprendra la vraie valeur populationnelle du rapport des cotes dans la majorité des cas (9 échantillons indépendants sur 10).

9.2. LORS DE LA PREMIÈRE ANNÉE

Cette section présente les résultats des analyses d'impact de la première année. Les estimations de l'effet traitement reposent sur la trajectoire développementale relative des trois groupes à l'étude au cours des quatre premières évaluations. La première évaluation est considérée comme une évaluation pré-intervention (niveau de base) et les trois évaluations subséquentes comme des évaluations post-intervention. Tel que mentionné, l'effet traitement est mesuré par l'estimateur DD permettant ainsi de capter les différences entre les trajectoires développementales des groupes expérimentaux.

Au total, six résultantes des enfants ont fait l'objet d'analyses basées sur les données collectées de la version française de l'ÉPE-AD. Un effet positif du programme testé est anticipé sur toutes les mesures sauf l'échelle Physique. Nous décrivons maintenant comment les mesures de résultantes ont été traitées dans les analyses d'impact.

9.2.1. Détails techniques

Il est important de déterminer dans quelle mesure les résultats obtenus dépendent de la spécification utilisée pour la modélisation des données. La stratégie générale adoptée ici est de présenter deux spécifications alternatives pour chaque analyse. La première, le modèle initial, considère uniquement les prédicteurs importants servant à identifier l'effet du programme. La deuxième spécification est plus complexe puisqu'elle inclut en plus les covariables. Les deux spécifications produisent des résultats qui, lorsque comparés, permettent de diagnostiquer des problèmes techniques ou de nuancer l'interprétation des données. Pour cet exercice de comparaison, le modèle initial sert principalement à enrichir l'interprétation des résultats définitifs provenant du modèle ajusté. Nous présentons également pour le modèle ajusté seulement, les résultats d'une analyse supplémentaire utilisant une approche non-paramétrique.

⁶⁹ Il serait redondant d'inclure un cinquième paramètre.

Détails de la spécification de modèles

Modèle initial

Le modèle initial de l'analyse par groupe de traitement comprend une série de variables fictives représentant la période d'évaluation, l'appartenance au groupe et un terme représentant les interactions entre les indicateurs de ces variables fictives (c.-à-d. les estimateurs DD). Le facteur temps ou « période d'évaluation » est représenté par trois variables fictives comparant les évaluations de suivi à la première évaluation (c.-à-d. la période pré-intervention). En d'autres mots, la période pré-intervention est placée en « référence ». Le facteur « groupe » est représenté par deux variables fictives servant à comparer les groupes témoins au groupe programme. En d'autres mots, le groupe programme est placée en « référence ». *Il est important de noter que la décision de placer le groupe programme en référence signifie que les valeurs négatives des estimations DD pour toutes les comparaisons entre les groupes représentent un effet traitement positif (c.-à-d. un avantage pour le groupe programme).* Le modèle initial est modifié afin de tester des hypothèses différentes (p. ex. l'ajout du dosage ou la qualité au modèle). Les changements apportés au modèle initial sont notés dans le texte au début de chaque sous-section.

Modèle ajusté

Le modèle initial est imbriqué dans le modèle ajusté, qui lui inclut en plus toutes les covariables identifiées au chapitre 7 ainsi que leur interaction avec le facteur temps. La liste complète des effets inclus dans le modèle ajusté est fournie au chapitre 7 (section 7.3).

Échantillons qui font l'objet de l'analyse

Échantillons de base

L'échantillon faisant l'objet des analyses varie en fonction de la variable résultante analysée. L'échantillon **complet** de participants contribue aux analyses portant sur l'échelle Communication. Pour les autres échelles, les analyses portent sur un **sous-échantillon** dans lequel sont exclus les cas ayant des données manquantes dues à l'administration de l'ÉPE-AD en anglais et pour qui les valeurs ne pouvaient pas être imputées.

Filtrage des données

Les enfants qui se sont retirés de l'étude avant l'évaluation de 12 mois sont exclus des analyses afin de maximiser la validité interne des estimateurs DD. Cet estimateur est aussi susceptible d'être biaisé par la présence de scores aberrants. Ce type de scores peut avoir une influence démesurée sur la valeur que prennent les paramètres de régression. Une règle de filtrage des données est typiquement imposée afin d'éliminer les observations les plus extrêmes. La règle adoptée dans le contexte des analyses rapportées touche moins de 0,01 % des observations. D'après cette règle, une observation est valide si elle se situe entre -3,5 et +3,5 écarts-type de la moyenne du groupe pour une évaluation donnée. Les observations affectées par la règle varient d'une analyse à l'autre par ce que la règle fait référence à des paramètres qui sont particuliers à la variable qui fait l'objet d'analyse.

Certains cas peuvent aussi être aberrants en raison d'une combinaison atypique de caractéristiques mesurées par les covariables utilisées dans le modèle de régression. Ces cas

peuvent nuire aux résultats des analyses parce qu'ils influencent démesurément l'estimation des paramètres de régression. Des analyses préliminaires (non-rapportées ici) n'ont pas décelé de cas problématiques ayant une combinaison aberrante de scores sur les covariables,⁷⁰ ce qui suggère qu'il n'y a pas de cas particulièrement influant nuisant à la généralisation des résultats découlant du modèle ajusté. C'est une considération particulièrement importante pour une analyse portant sur un échantillon aussi modeste.

9.2.2. Analyses par groupes

Cette sous-section présente les analyses par groupe de traitement. Les résultats pour le modèle initial sont présentés au tableau 9.1 et ceux du modèle ajusté sont présentés au tableau 9.2. Les résultats pour chaque échelle sont décrits en premier pour le modèle initial et ensuite, pour le modèle ajusté. Une contre-validation non-paramétrique des résultats est alors présentée pour le modèle ajusté. La section se termine par un sommaire des résultats des analyses par groupe.

Échelle Communication

Modèle initial

L'effet au niveau de base rapporté dans le tableau 9.1 représente la position relative des groupes pour la période pré-intervention (première rangée). Une valeur positive indique un déficit du groupe programme tandis qu'une valeur négative indique un avantage. Pour la Communication, un déficit non-significatif est observé de 0,45 et de 0,23 respectivement relatif au groupe témoin en garderie et au groupe témoin hors garderie. Ces différences non-significatives sont annulées dans l'estimation des effets DD selon lesquels nous observons un effet constant dans le temps du programme sur la Communication relativement aux deux groupes témoins. De plus, les gains des enfants du groupe programme pour chaque évaluation post-intervention sont significativement plus importants que ceux observés pour les enfants des groupes témoins. La taille de cet effet est de l'ordre de 0,30 à 0,40 écart-type. Notons qu'une simple comparaison des valeurs indiquées dans le tableau suffit pour montrer que les effets positifs du programme ne sont pas suffisants pour combler le déficit initial du groupe programme sur ce domaine.

Modèle ajusté

L'effet au niveau de base rapporté dans le tableau 9.2 représente la position relative des groupes pour la période pré-intervention après avoir ajusté pour les covariables. Comme pour le modèle initial, une valeur positive indique un déficit du groupe programme tandis qu'une valeur négative indique un avantage. Le modèle de régression avec contrôle statistique des variables confusionnelles apporte une correction partielle du désavantage non-significatif observé au niveau de base pour le groupe programme relativement au groupe témoin en garderie. La taille du déficit relativement au groupe témoin en garderie est de seulement 0,33 écart-type dans cette analyse. Pour sa part, le déficit non significatif du groupe programme par rapport au groupe témoin hors garderie est carrément annulé.

⁷⁰ Nous avons évalué la similitude des cas par rapport à leur combinaison de caractéristiques en estimant leur distance de la moyenne multivariée, c'est-à-dire leur valeur de « distance de Mahalanobis » (de l'anglais *Mahalanobis distance*). Aucun cas ne se distinguait de la distribution empirique de cette statistique.

Les estimations DD de l'effet du programme se voient augmentées de façon générale pour l'échelle de Communication par l'inclusion des covariables. C'est notamment le cas pour tous les tests DD pour la comparaison des groupes en garderie et pour la comparaison avec le groupe témoin hors garderie. Les effets DD pour la comparaison de groupes en garderie sont d'environ 0,40 écart-type tout au long de la première année de l'étude. L'effet correspondant pour la comparaison au groupe témoin hors garderie semble émerger plus tardivement puisqu'un effet de 0,40 est observé seulement à l'évaluation de 12 mois.

Contre-validation non-paramétrique

Dans cette analyse non-paramétrique, la performance observée des enfants sur l'échelle de Communication est exprimée en termes de la probabilité d'être classifiée dans un quintile supérieur : la cote de réussite. Les résultats de l'analyse reproduisent le degré de signification des effets DD pour la comparaison des groupes en garderie. Le rapport des cotes est de 0,31 (IC 90% de 0,20 à 0,47), de 0,44 (IC 90 % de 0,25 à 0,80), et de 0,41 (IC 90 % de 0,22 à 0,78) en faveur du groupe programme respectivement pour les effets DD à quatre, huit et 12 mois. En d'autres mots, les enfants du groupe témoin en garderie sont 69 %, 56 % et 59 % moins susceptibles de se retrouver dans le quintile supérieur respectivement à quatre, huit et 12 mois comparativement aux enfants du groupe programme.

L'analyse non-paramétrique révèle également que l'impact du programme relatif au groupe témoin hors garderie est moins robuste. Un effet significatif de l'estimateur DD est observé seulement à quatre mois alors que l'examen des IC 90% révèle que l'effet à 12 mois s'approche du seuil de signification de 10 %. Le rapport des cotes est de 0,69 (IC 90 % de 0,50 à 0,95), de 0,83 (IC 90 % de 0,51 à 1,35) et de 0,64 (IC 90 % de 0,40 à 1,02) en faveur du groupe programme respectivement pour les effets DD à quatre, huit et 12 mois. En d'autres mots, les enfants témoin hors garderie sont 31 % moins susceptibles de se retrouver dans le quintile supérieur à 4 mois comparativement aux enfants du groupe programme.

Échelle Conscience de soi

Modèle initial

L'effet au niveau de base rapporté dans le tableau 9.1 représente la position relative des groupes pour la période pré-intervention. Une valeur positive indique un déficit du groupe programme tandis qu'une valeur négative indique un avantage. On observe un déficit non-significatif de 0,36 et 0,25 relatif aux groupes témoins en garderie et hors garderie respectivement. Pour ce qui est des estimations DD, nous observons un effet du programme qui émerge relatif au groupe témoin en garderie à quatre mois et se maintient jusqu'à la fin de la première année. L'importance de l'effet diminue d'environ 0,32 à 4 mois jusqu'à 0,21 écart-type à 12 mois. Comme pour l'échelle de la Communication, l'effet du programme testé, bien que significatif, ne réussit pas à compenser entièrement le déficit observé pour le groupe programme au niveau de base. En comparaison des enfants du groupe témoin hors garderie, un effet significatif est détecté seulement à huit mois pour le groupe programme. Toutefois, l'effet observé est peu crédible en raison de son instabilité dans le temps.

Modèle ajusté

L'effet au niveau de base rapporté dans le tableau 9.2 représente la position relative des groupes pour la période pré-intervention après avoir ajusté pour les covariables. Comme pour le modèle initial, une valeur positive indique un déficit du groupe programme tandis qu'une valeur négative indique un avantage. Le modèle de régression avec contrôle statistique de variables confesionnelles révèle un désavantage réduit du groupe programme relativement aux groupes témoin. Comparativement au groupe témoin en garderie, le déficit observé est réduit à 0,25 écart-type tandis qu'il est presque entièrement annulé par rapport au groupe témoin hors garderie.

Pour ce qui est des effets DD, nous observons un effet constant dans le temps du programme indiquant un avantage significatif en faveur du groupe programme comparativement au groupe témoin en garderie même après avoir effectué un contrôle statistique de variables confusionnelles. Dans ce modèle, la taille de l'effet est relativement stable dans le temps, se fixant à peu près à 0,30 écart-type. Par contre, l'estimation DD en faveur du groupe programme comparativement au groupe témoin hors garderie est significatif à huit mois seulement. Comme pour le modèle initial, ce dernier peut être jugé comme étant peu crédible en raison de son instabilité dans le temps.

Contre-validation non-paramétrique

Pour l'analyse non-paramétrique, la performance des enfants observée avec l'échelle de Conscience de soi est exprimée en termes de la probabilité d'être classifiée dans un quintile supérieur : la cote de réussite. Les résultats de l'analyse reproduisent la signification des effets DD pour la comparaison du groupe programme au groupe témoin en garderie. Le rapport des cotes se chiffre à 0,49 (IC 90 % de 0,36 à 0,66), 0,58 (IC 90 % de 0,37 à 0,90) et 0,55 (IC 90 % de 0,39 à 0,79) en faveur du groupe programme respectivement pour les effets DD à quatre, huit et 12 mois. En d'autres mots, les enfants du groupe témoin en garderie sont 51%, 42% et 45% moins susceptibles de se retrouver dans le quintile supérieur respectivement à quatre, huit et 12 mois comparativement aux enfants du groupe programme.

L'analyse non-paramétrique reproduit également les impacts observés du programme relatif au groupe témoin hors garderie; le seul effet DD significatif se retrouve à l'évaluation de huit mois. Le rapport des cotes pour cet effet est de 0,55 (IC 90 % de 0,36 à 0,83) en faveur du groupe programme, indiquant que les enfants du groupe programme sont 45 % plus susceptible de se retrouver dans le quintile supérieur (c. à d. la catégorie réussite).

Échelle Cognition

Modèle initial

L'effet au niveau de base rapporté dans le tableau 9.1 représente la position relative des groupes pour la période pré-intervention. Une valeur positive indique un déficit du groupe programme tandis qu'une valeur négative indique un avantage. On observe un déficit significatif du groupe programme relativement aux groupes témoin au niveau de base. D'ailleurs, l'analyse DD ne montre aucun effet significatif du programme sur les gains développementaux subséquents des enfants sur cette dimension de la préparation à la scolarisation : l'effet global testé par la statistique de Wald et les estimations DD ne sont pas significatifs.

Modèle ajusté

L'effet au niveau de base rapporté dans le tableau 9.2 représente la position relative des groupes pour la période pré-intervention après avoir ajusté pour les covariables. Comme pour le modèle initial, une valeur positive indique un déficit du groupe programme tandis qu'une valeur négative indique un avantage. Le modèle de régression avec contrôle statistique de variables confusionnelles n'a pas corrigé le déficit de 0,20 écart-type du groupe programme par rapport au groupe témoin en garderie au niveau de base. Par contre, le déficit au niveau de base relatif au groupe témoin hors garderie est complètement annulé.

L'analyse DD ajustée décèle des impacts positifs significatifs du programme de 0,20 écart-type observés à 12 mois relatif aux deux groupes témoins. Une interprétation possible de ce résultat est que c'est un effet valide qui émerge tardivement pour après se maintenir lors de la deuxième année du projet. Il reste à voir si cette hypothèse sera confirmée lors des analyses d'impact de la deuxième année.

Contre-validation non-paramétrique

Pour l'analyse non-paramétrique, la performance des enfants observée avec l'échelle de Cognition était exprimée en termes de la probabilité d'être classifiée dans un quintile supérieur : la cote de réussite. Les résultats de l'analyse reproduisent partiellement les résultats du modèle ajusté. Ainsi, l'effet à 12 mois pour la comparaison des groupes en garderie n'est pas significatif avec un rapport des cotes de 0,76 (IC 90 % de 0,48 à 1,19) tandis que celui pour la comparaison au groupe témoin hors garderie est significatif avec un rapport de cotes en faveur du groupe programme de 0,57 (IC 90 % de 0,39 à 0,82).

Tableau 9.1 : Impact du programme non-ajusté sur la préparation à la scolarisation lors de la première année — Scores standardisés

Type de différence	Indicateur de la préparation à la scolarisation											
	Communication		Conscience de soi		Cognition		Physique		Vocabulaire réceptif		Vocabulaire expressif	
	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.
G1 vs G2												
Niveau de base	0,45	0,32	0,36	0,25	0,27**	0,12	0,30	0,19	0,32	0,33	0,43	0,26
4 mois (DD)	-0,43***	0,10	-0,32***	0,04	-0,13	0,13	-0,10	0,13	-0,20	0,30	-0,44***	0,12
8 mois (DD)	-0,27**	0,10	-0,24**	0,11	-0,01	0,16	-0,32**	0,12	-0,41	0,25	-0,20	0,13
12 mois (DD)	-0,30***	0,10	-0,21*	0,11	-0,13	0,15	-0,22	0,15	-0,34	0,21	-0,22**	0,10
G1 vs G3												
Niveau de base	0,23	0,33	0,25	0,28	0,13	0,09	-0,09	0,16	0,13	0,35	0,30	0,25
4 mois (DD)	-0,24**	0,09	-0,01	0,10	0,00	0,10	0,13	0,11	-0,11	0,29	-0,01	0,16
8 mois (DD)	-0,24*	0,13	-0,26*	0,13	-0,01	0,15	0,01	0,13	-0,04	0,23	-0,10	0,18
12 mois (DD)	-0,33**	0,14	-0,10	0,08	-0,10	0,14	0,14	0,13	-0,23	0,21	-0,05	-0,05
Groupe x Temps Wald F	4,18***		9,34***		1,11		2,11		3,00**		3,59**	

Note : Degrés de liberté du test Wald F sont 6 et 17. Les erreurs types robuste de Huber-White sont rapportées avec groupement des erreurs par garderie. Les valeurs négatives des estimations DD pour toutes les comparaisons entre les groupes représentent un effet traitement positif (c.-à-d. un avantage pour le groupe programme). Le seuil de signification est indiqué à *** $p < 1\%$, ** $p < 5\%$; * $p < 10\%$. Diff = différence, E.T. = erreur type.

**Tableau 9.2 : Impact du programme (après ajustement pour covariables) sur la préparation à la scolarisation lors de la première année
— Scores standardisés**

Type de différence	Indicateur de la préparation à la scolarisation											
	Communication		Conscience de soi		Cognition		Physique		Vocabulaire réceptif		Vocabulaire expressif	
	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.
G1 vs G2												
Niveau de base	0,33***	0,08	0,26***	0,07	0,17**	0,07	0,28**	0,10	0,19*	0,10	0,32***	0,08
4 mois (DD)	-0,46***	0,10	-0,36***	0,05	-0,15	0,10	-0,13	0,14	-0,16	0,22	-0,43***	0,12
8 mois (DD)	-0,39***	0,12	-0,32**	0,11	0,01	0,12	-0,35***	0,11	-0,33**	0,15	-0,25*	0,13
12 mois (DD)	-0,43***	0,13	-0,24**	0,09	-0,16*	0,08	-0,23**	0,10	-0,27**	0,12	-0,26**	0,12
G1 vs G3												
Niveau de base	0,01	0,06	0,03	0,07	0,04	0,08	-0,04	0,11	-0,07	0,09	0,05	0,07
4 mois (DD)	0,10**	0,09	-0,06	0,06	-0,03	0,12	0,03	0,12	-0,03	0,20	-0,01	0,11
8 mois (DD)	-0,27**	0,11	-0,37***	0,11	-0,08	0,12	0,01	0,07	0,05	0,17	-0,17	0,14
12 mois (DD)	-0,40***	0,11	-0,12	0,08	-0,21***	0,06	0,08	0,14	-0,20*	0,11	-0,10	0,09
Groupe x Temps Wald F	4,74***		11,20***		4,89***		3,33**		2,07		4,56***	

Note : Degrés de liberté du test Wald F sont 6 et 17. Les erreurs-types robuste de Huber-White sont rapportées avec groupement des erreurs par garderie. *Les valeurs négatives des estimations DD pour toutes les comparaisons entre les groupes représentent un effet traitement positif (c.-à-d. un avantage pour le groupe programme).* Le seuil de signification est indiqué à *** $p < 1\%$, ** $p < 5\%$; * $p < 10\%$. Diff = différence, E.T. = erreur type.

Échelle Physique

Modèle initial

L'effet au niveau de base rapporté dans le tableau 9.1 représente la position relative des groupes pour la période pré-intervention. Une valeur positive indique un déficit du groupe programme tandis qu'une valeur négative indique un avantage. La statistique de Wald testant l'effet global de l'interaction entre le groupe et la période d'évaluation n'est pas significative, indiquant que les groupes expérimentaux suivent des trajectoires qui sont globalement parallèles au cours de la première année de l'étude. La seule perturbation observée parmi les estimations DD est l'effet significatif à huit mois pour la comparaison des groupes en garderie. Nous accordons à cet effet peu de crédibilité puisque il n'est pas prédit a priori et qu'il s'avère instable dans le temps.

Modèle ajusté

L'effet au niveau de base rapporté dans le tableau 9.2 représente la position relative des groupes pour la période pré-intervention après avoir ajusté pour les covariables. Comme pour le modèle initial, une valeur positive indique un déficit du groupe programme tandis qu'une valeur négative indique un avantage. Le modèle de régression avec contrôle statistique de variables confusives n'a pas affecté la taille du déficit du groupe programme relativement aux groupes témoins. Toutefois, la statistique de Wald est significative après l'ajustement statistique, ce qui indique que les groupes expérimentaux suivent des trajectoires développementales qui sont globalement non parallèles au cours de la première année de l'étude. Notons que deux effets DD sont significatifs notamment à huit mois et à 12 mois pour la comparaison des groupes en garderie. Cet effet n'était pas anticipé puisque le domaine physique et moteur n'était pas ciblé par le programme testé.

Contre-validation non-paramétrique

Pour l'analyse non-paramétrique, la performance des enfants observée avec l'échelle Physique est exprimée en termes de la probabilité d'être classifiée dans un quintile supérieur : la cote de réussite. Les résultats de l'analyse reproduisent partiellement les effets DD observés dans le modèle ajusté pour la comparaison des groupes en garderies. Le rapport des cotes est de 0,76 (IC 90 % de 0,49 à 1,20), 0,53 (IC 90 % de 0,33 à 0,84), et 0,52 (IC 90 % de 0,31 à 0,87) en faveur du groupe programme respectivement pour les effets DD à quatre, huit et 12 mois.

Échelle Vocabulaire réceptif

Modèle initial

L'effet au niveau de base rapporté dans le tableau 9.1 représente la position relative des groupes pour la période pré-intervention. Selon la statistique de Wald, la trajectoire des trois groupes de traitement diverge significativement lors de la première année de l'étude. Toutefois, cette divergence n'est pas observée au niveau de base (voir les effets DD non-significatifs). L'absence de significativité observée pour toutes les périodes d'évaluation suggère que l'effet global testé par la statistique de Wald n'est pas attribuable à un effet du programme testé. Il est à

noter que le Vocabulaire réceptif est l'échelle parmi les six avec la cote de fidélité de mesure la moins élevée (voir le tableau 3.8).

Modèle ajusté

L'effet au niveau de base rapporté dans le tableau 9.2 représente la position relative des groupes pour la période pré-intervention après avoir ajusté pour les covariables. D'après la statistique de Wald, la trajectoire des trois groupes de traitement est globalement parallèle après ajustement statistique. Les estimations DD calculées en fonction des deux groupes témoins indiquent un effet significatif de 0,20 et de 0,33 écart-type à la fin de la première année de l'étude. C'est un effet qui émerge peut-être tardivement, comme celui observé pour l'échelle Cognition. Cette constatation est à vérifier dans le cadre des analyses de la deuxième année du projet.

Contre-validation non-paramétrique

Pour l'analyse non-paramétrique, la performance des enfants observée avec l'échelle Vocabulaire réceptif est exprimée en termes de la probabilité d'être classifiée dans un quintile supérieur : la cote de réussite. Les résultats de l'analyse reproduisent la signification des effets DD pour la comparaison du groupe programme au groupe témoin hors garderie à l'évaluation 12 mois avec un rapport de cotes de 0,65 (IC 90 % de 0,44 à 0,95). Pour la comparaison des groupes en garderie, le rapport des cotes est de 0,66 (IC 90 % de 0,27 à 1,62), 0,40 (IC 90 % de 0,21 à 0,76), et 0,43 (IC 90 % de 0,28 à 0,66) respectivement pour les effets DD à quatre, huit et 12 mois. Comme pour l'analyse paramétrique, seulement les effets à huit et à 12 mois sont significatifs.

Échelle Vocabulaire expressif

Modèle initial

L'effet au niveau de base rapporté dans le tableau 9.1 représente la position relative des groupes pour la période pré-intervention. Selon la statistique de Wald, les groupes de traitement divergent significativement au fil du temps. L'examen des estimations DD révèle un effet de programme significatif relativement au groupe témoin en garderie pour les périodes d'évaluation à quatre et à 12 mois. L'effet paraît crédible puisque sa taille est relativement constante d'une évaluation à l'autre, variant entre 0,20 et 0,45 écart-type. L'effet DD à quatre mois est plus important que le déficit initial, ce qui indique un renversement de la position relative des groupes par rapport à leur score absolu sur cette échelle. Aucun effet significatif du programme n'est observé relativement au groupe témoin hors garderie.

Modèle ajusté

L'effet au niveau de base rapporté dans le tableau 9.2 représente la position relative des groupes pour la période pré-intervention après avoir ajusté pour les covariables. Comme pour le modèle initial, une valeur positive indique un déficit du groupe programme tandis qu'une valeur négative indique un avantage. Suite à l'ajustement apporté par l'inclusion des covariables, la statistique de Wald testant la non-équivalence de la trajectoire des groupes de traitement demeure significative. Le modèle ajusté produit des estimations DD significatives pour toutes les

évaluations de la première année, soit à quatre, huit et 12 mois relativement au groupe témoin en garderie. L'inclusion des covariables semble avoir eu pour effet d'augmenter la précision des estimations DD. La taille des effets n'est toutefois pas affectée, variant encore entre 0,25 et 0,43 écart-type environ. Aucun effet significatif du programme n'est observé relativement au groupe témoin hors garderie.

Contre-validation non-paramétrique

Pour l'analyse non-paramétrique, la performance des enfants observée avec l'échelle Vocabulaire expressif est exprimée en termes de la probabilité d'être classifiée dans un quintile supérieur : la cote de réussite. Les résultats de l'analyse reproduisent partiellement la signification des effets DD pour la comparaison des groupes en garderie. Le rapport des cotes est de 0,43 (IC 90 % de 0,28 à 0,68), 0,76 (IC 90 % de 0,41 à 1,42) et 0,60 (IC 90 % de 0,35 à 1,03) en faveur du groupe programme respectivement à quatre, huit et 12 mois. En d'autres mots, l'effet à quatre mois est confirmé, mais celui de 12 mois ne l'est pas, bien que ce dernier approche le seuil de signification. Enfin, comme pour l'analyse paramétrique, aucun effet n'est décelé pour la comparaison au groupe témoin hors garderie.

Sommaire : Analyses par groupe

Les résultats de l'analyse par groupes de traitement indiquent que le programme testé a eu un effet positif sur le développement de la préparation à la scolarisation des enfants du groupe programme. Cette conclusion découle des résultats de l'analyse paramétrique des données et de la reproduction des résultats dans l'analyse non-paramétrique.

Par rapport au niveau de base, les enfants inscrits dans le groupe programme connaissent une croissance plus rapide sur plusieurs dimensions de la préparation à la scolarisation comparativement à leurs pairs dans les deux groupes témoins. Ce fait est attesté par les estimations DD rapportés aux tableaux 9.1 et 9.2. L'effet du programme est plus marqué relativement au groupe témoin en garderie, peut-être en raison de la composition relativement homogène de ce groupe témoin.

De façon générale, il est à noter que l'ajout de la deuxième cohorte a pour effet de rendre les résultats plus stable sur la dimension du temps comparativement aux résultats rapportés dans le *Rapport des résultats de la première cohorte* (2014). Pour la Communication, la Conscience de soi et le Vocabulaire expressif, l'effet du programme testé relatif au groupe témoin en garderie est significatif à quatre mois et le demeure par la suite, avec des effets de l'ordre de 0,25 à 0,43 écart-type. En se basant sur les normes développementales rapportées par Hill et collègues (2008), c'est un gain dans le développement des enfants du groupe programme d'environ 2,3 à 3,9 mois comparativement au groupe témoin en garderie.

D'autre part, les analyses d'impact figurant dans le *Rapport des résultats de la première cohorte* (2014) n'ont pas décelé d'effets du programme testé relativement au groupe témoin hors garderie. Ce n'est pas le cas dans les présentes analyses. Lors de l'entrée à la maternelle de quatre ans, nous notons un effet programme de 0,30, 0,20 et 0,20 écart-type pour la Communication, la Cognition, et le Vocabulaire réceptif, respectivement. En d'autres mots, c'est un gain dans le développement des enfants du groupe programme d'environ 1,8 à 2,7 mois relativement au groupe témoin hors garderie.

9.2.3. Analyses par dosage

Cette section présente les analyses tenant compte du dosage. Le lecteur est prié de noter que ces analyses portent uniquement sur les groupes expérimentaux qui sont en garderie. Les résultats pour le modèle initial sont présentés au tableau 9.3 et ceux du modèle ajusté, au tableau 9.4. La variable de dosage (moyenne des heures en garderie par semaine) a été standardisée pour les fins des analyses. Ainsi, une valeur de zéro représente la moyenne de l'échantillon (c.-à-d. centrée). Les modèles rapportés comprennent uniquement l'effet linéaire du dosage.⁷¹

Les résultats pour chaque échelle sont décrits en premier pour le modèle initial et ensuite, pour le modèle ajusté. La reproduction non-paramétrique des résultats du modèle ajusté est par la suite présentée. Enfin, la section se termine par un sommaire des résultats des analyses par dosage. Le lecteur est prié de noter que 98,3% de l'échantillon des groupes en garderie était en garde plus de dix heures par semaine en garderie.

Échelle Communication

Modèle initial

Pour l'échantillon entier, le nombre d'heures passées en garderie est positivement associé au rendement sur l'échelle Communication (voir le tableau 9.3). L'effet estimé du dosage au niveau de base indique qu'une augmentation de 10 heures par semaine correspond pour le présent échantillon à une augmentation moyenne d'environ 0,20 écart-type pour la Communication.

Nous avons estimé l'instabilité de cet effet au fil du temps (voir les effets dosage par temps) et nous trouvons que la valeur de base de l'effet du dosage s'applique aussi bien à huit et à 12 mois. Exceptionnellement, nous trouvons que l'effet du dosage est plus fort à quatre mois pour le présent échantillon (voir les effets dosage par temps).

Le modèle nous permet également d'estimer l'impact du programme selon le niveau de dosage. Les résultats rapportés au tableau 9.3 indiquent que le dosage influence l'impact du programme à quatre mois seulement. Pour cette évaluation, l'impact du programme serait d'environ 0,10 écart-type (non-significatif) pour un enfant qui passe une moyenne de 17 heures par semaines en garderie, d'environ 0,30 écart-type pour un dosage de 27 heures par semaines (significatif) et d'environ 0,50 écart-type quand le dosage s'élève à 37 heures par semaines (significatif).^{72,73} En raison du petit nombre d'enfants participant à l'étude, cette estimation du

⁷¹ La relation entre le dosage et le développement peut être linéaire ou non-linéaire. Une relation linéaire tient si l'effet du dosage est constant peu importe l'intensité d'exposition au milieu de garde (p. ex., 10 heures semaine contre 40 heures semaine). Des analyses préliminaires (p. ex., en utilisant des termes quadratiques; inspection des résiduels) n'ont pas décelé de relation non-linéaire dans le modèle ajusté. La conclusion à tirer de ce résultat est que l'effet linéaire du dosage suffit pour les fins de l'analyse rapportée dans le présent document, ce qui est cohérent avec la littérature indiquant que les effets du dosage sont linéaires (National Institute of Child Health and Human Development Early Child Care Research Network et Duncan, 2003).

⁷² La moyenne de l'échantillon est d'environ 27 heures par semaine. Un dosage de 10 heures de moins de cette moyenne hebdomadaire serait de 17 heures par semaine. Un dosage de 10 heures de plus de cette moyenne hebdomadaire serait de 37 heures par semaine.

⁷³ L'effet du groupe n'est pas rapporté dans le tableau 9.3, mais était néanmoins estimé dans le cadre du modèle du dosage. La valeur de 0,30 provient du modèle de régression initiale du dosage et elle représente l'impact du programme relatif au groupe témoin en garderie à quatre mois pour un enfant de dosage moyen (27 heures par semaine). Les estimations rapportées dans le texte sont calculées en additionnant l'effet de l'interaction triple significative entre le temps, le groupe et le dosage rapportée au tableau 9.3 à l'effet DD du programme correspondant (non-rapporté). L'insertion du dosage dans le modèle peut affecter l'estimation de l'effet du programme, donc les valeurs estimées pourraient fort bien ne pas se concorder avec des estimations provenant d'un modèle sans le dosage.

seuil critique de dosage pour l'obtention d'un impact significatif du programme n'est pas très précise.

Modèle ajusté

Le patron de résultats produit par le modèle ajusté se distingue de celui du modèle initial de par l'absence de significativité de l'effet du dosage au niveau de base (voir le tableau 9.4). L'effet du dosage pour le groupe programme émerge à quatre mois, demeurant stable pour les périodes d'évaluation à huit et 12 mois (voir les effets dosage par temps au tableau 9.4). Selon le modèle, un enfant passant 10 heures de plus en garderie connaîtrait une augmentation de 0,20 écart-type ou un gain de 1,8 mois dans son développement. Les effets DDD à quatre et à 12 mois indiquent que cet effet DD positif du dosage est observé uniquement chez les enfants du groupe programme.

En d'autres mots, le modèle indique que la taille de l'effet du programme pour la Communication varie en fonction du dosage. À l'entrée à la maternelle de quatre ans (à 12 mois), le modèle prédit un impact du programme de 0,20 écart-type pour un enfant dont le dosage est de 20 heures par semaine en garderie, de 0,40 écart-type pour un enfant dont le dosage est 30 heures par semaine et de 0,60 écart-type pour un enfant dont le dosage est de 40 heures par semaine (valeurs approximatives). Il importe de noter que ces estimations reflètent bien ce qui est observé en moyenne pour le présent échantillon, mais leur précision est trop faible pour conclure qu'elles reflètent justement les paramètres de la population des enfants francophones minoritaires inscrits à une garderie francophone.

Contre-validation non-paramétrique

Pour l'analyse non-paramétrique, la performance des enfants observée avec l'échelle Communication est exprimée en termes de la probabilité d'être classifiée dans un quintile supérieur : la cote de réussite. Les effets à reproduire sont les estimations DDD testant la différenciation de l'effet du dosage selon le groupe de traitement.

Relativement au groupe témoin en garderie, le rapport des cotes pour l'effet DDD est de 0,96 (IC 90 % de 0,91 à 1,01), 1,00 (IC 90 % de 0,94 à 1,06) et 0,94 (IC 90 % de 0,89 à 0,99) en faveur du groupe programme respectivement pour les périodes d'évaluation quatre, huit et 12 mois. Spécifiquement, l'effet du dosage pour les enfants du groupe programme est significativement plus prononcé que pour les enfants du groupe témoin en garderie pour l'évaluation à 12 mois. L'effet à quatre mois approche le seuil de signification de 10 %. En somme, les résultats de l'analyse non-paramétrique reproduisent l'effet différenciation de l'effet du dosage à l'entrée des enfants à la maternelle de quatre ans (c.-à-d. l'évaluation à 12 mois). Selon le modèle, le rapport des cotes de l'effet programme augmente de 6 % pour chaque heure de plus passée en garderie par rapport à la moyenne.

Tableau 9.3 : Différentiation non-ajustée de l'impact des heures en garderie (Dosage) selon le programme reçu (1^{ière} année) — Scores standardisés

Type de différence	Indicateur de la préparation à la scolarisation											
	Communication		Conscience de soi		Cognition		Physique		Vocabulaire réceptif		Vocabulaire expressif	
	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.
Dosage												
Niveau de base	0,019*	0,010	0,009	0,011	-0,010	0,013	-0,001	0,017	0,016*	0,008	0,004	0,009
Dosage x groupe												
G1 vs G2 (DD)	0,000	0,015	0,004	0,016	0,020	0,017	0,002	0,018	-0,008	0,011	0,008	0,014
Dosage x temps												
4 mois (DD)	0,020***	0,006	0,012	0,002	0,020***	0,004	0,002	0,013	0,010	0,006	0,018***	0,004
8 mois (DD)	0,004	0,018	0,002	0,011	0,003	0,012	-0,014	0,019	0,009	0,008	-0,003	0,013
12 mois (DD)	0,006	0,009	0,017	0,009	0,027*	0,009	-0,007	0,015	0,006	0,010	0,017**	0,008
Dosage x temps x groupe												
4 mois (DDD)	-0,021*	0,011	-0,015	0,010	-0,026*	0,011	-0,010	0,016	-0,003	0,012	-0,023**	0,010
8 mois (DDD)	-0,007	0,022	-0,014	0,018	-0,015	0,018	0,001	0,023	-0,011	0,014	-0,003	0,021
12 mois (DDD)	-0,013	0,018	-0,031	0,019	-0,054***	0,017	-0,026	0,020	-0,014	0,015	-0,029	0,019
Dosage x temps Wald F	1,38		0,57		0,56		2,39		1,43		0,78	
Dosage x temps x groupe Wald F	1,19		1,36		5,14*		2,97*		0,33		2,07	

Note : Degrés de liberté des tests Wald F sont 6 et 19. Les erreurs-types robuste de Huber-White sont rapportées avec groupement des erreurs par garderie. *Les valeurs négatives des estimations DD pour toutes les comparaisons entre les groupes représentent un effet traitement positif (c.-à-d. un avantage pour le groupe programme).* Le seuil de signification est indiqué à *** $p < 1\%$, ** $p < 5\%$; * $p < 10\%$. Diff = différence, E.T. = erreur type.

Tableau 9.4 : Différentiation ajustée de l'impact des heures en garderie (Dosage) selon le programme reçu (1^{ière} année) — Scores standardisés

Type de différence	Indicateur de la préparation à la scolarisation											
	Communication		Conscience de soi		Cognition		Physique		Vocabulaire réceptif		Vocabulaire expressif	
	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.
Dosage												
Niveau de base	-0,001	0,008	-0,007	0,008	-0,006	0,008	0,004	0,010	0,005	0,007	-0,011	0,007
Dosage x groupe												
Niveau de base	0,001	0,009	0,005	0,009	0,012	0,012	-0,002	0,010	-0,012	0,008	0,009	0,009
Dosage x temps												
4 mois (DD)	0,019***	0,006	0,015***	0,004	0,011**	0,004	0,005	0,015	0,018***	0,003	0,017*	0,008
8 mois (DD)	0,019**	0,009	0,015**	0,007	-0,002	0,007	-0,009	0,019	0,021***	0,006	0,002	0,008
12 mois (DD)	0,024***	0,008	0,029***	0,009	0,036***	0,010	0,003	0,011	0,017**	0,007	0,031***	0,010
Dosage x temps x groupe												
4 mois (DDD)	-0,025***	0,009	-0,022**	0,010	-0,025**	0,012	-0,005	0,017	-0,018**	0,008	-0,025**	0,010
8 mois (DDD)	-0,013	0,013	-0,018	0,013	-0,009	0,016	0,003	0,020	-0,016	0,011	0,000	0,014
12 mois (DDD)	-0,019*	0,011	-0,033**	0,013	-0,054***	0,015	-0,030**	0,014	-0,009	0,011	-0,033**	0,013
Heures par temps Wald F (3, 16)	1,73		0,99		2,06		1,05		2,59*		1,49	
Heures par groupe par temps Wald F (3, 16)	2,83*		2,96*		19,93***		6,24***		2,03		4,06**	

Note : Les erreurs-types robuste de Huber-White sont rapportées avec groupement des erreurs par garderie. Les valeurs négatives des estimations DD pour toutes les comparaisons entre les groupes représentent un effet traitement positif (c.-à-d. un avantage pour le groupe programme). Le seuil de signification est indiqué à *** $p < 1\%$, ** $p < 5\%$; * $p < 10\%$. Diff = différence, E.T. = erreur type.

Échelle Conscience de soi

Modèle initial

L'effet dosage n'est pas significatif au niveau de base, ni aux évaluations subséquentes et il ne varie pas significativement selon le groupe expérimental.

Modèle ajusté

Le modèle ajusté ne produit pas le même patron de résultats que le modèle initial. L'effet de dosage, nul au niveau de base, augmente significativement pour l'ensemble des enfants au fil du temps selon les estimations DD (voir dosage par temps). Il faut préciser que selon les estimations DDD l'effet du dosage est significativement plus fort pour le groupe programme à quatre mois et à 12 mois pour le groupe programme. À 12 mois, l'effet est de l'ordre de 0,30 écart-type par 10 heures en garderie, et ce, indépendamment des covariables.

Contre-validation non-paramétrique

Pour l'analyse non-paramétrique, la performance des enfants observée avec l'échelle Conscience de soi est exprimée en termes de la probabilité d'être classifiée dans un quintile supérieur : la cote de réussite. Les effets à reproduire sont les estimations DDD testant la différenciation de l'effet du dosage selon le groupe.

Relativement au groupe témoin en garderie, le rapport des cotes pour l'effet DDD est de 0,93 (IC 90 % de 0,88 à 0,98), 0,95 (IC 90 % de 0,90 à 1,01) et 0,93 (IC 90 % de 0,86 à 0,99) en faveur du groupe programme respectivement pour les périodes d'évaluation à quatre, huit et 12 mois. Les résultats de l'analyse reproduisent partiellement la signification des effets DDD à quatre et 12 mois. À l'entrée des enfants à la maternelle de quatre ans, le modèle estime que le rapport des cotes pour l'impact du programme s'intensifie en faveur du groupe programme de 7 % pour chaque heure de plus passée en garderie que la moyenne hebdomadaire. À l'inverse, l'impact du programme s'atténue de 7 % pour chaque heure de moins passée en garderie que la moyenne hebdomadaire.

Échelle Cognition

Modèle initial

L'effet direct d'heures en garderie n'est pas significatif au niveau de base mais augmente significativement et positivement pour l'ensemble des participants à quatre mois et à 12 mois selon l'estimateur DD. L'estimateur DDD sert à préciser qu'à quatre et à 12 mois l'effet positif du dosage est observé uniquement pour le groupe programme.

Modèle ajusté

Le modèle ajusté produit essentiellement le même patron de résultats que le modèle initial. L'effet direct d'heures en garderie n'est pas significatif au niveau de base mais augmente significativement et positivement pour le groupe programme à quatre et à 12 mois. À 12 mois, l'effet est très prononcé, représentant un avantage de 0,50 écart-type (4,5 mois de développement environ) par 10 heures de plus en garderie programme.

Contre-validation non-paramétrique

Pour l'analyse non-paramétrique, la performance des enfants observée à l'échelle Cognition est exprimée en termes de la probabilité d'être classifiée dans un quintile supérieur : la cote de réussite. Les effets à reproduire sont les estimations DDD testant la différenciation de l'effet du dosage selon le groupe.

Relativement au groupe témoin en garderie, le rapport des cotes pour l'effet DDD est de 0,96 (IC 90 % de 0,92 à 1,01), 1,00 (IC 90 % de 0,94 à 1,06) et 0,90 (IC 90 % de 0,86 à 0,96) respectivement pour les évaluations à quatre, huit et 12 mois. Les résultats de l'analyse reproduisent partiellement la signification des effets DDD observés, et ce, à 12 mois seulement. Le modèle estime que le rapport des cotes associé à l'impact du programme augmente de 10 % pour chaque heure de plus passée en garderie excédant la moyenne.

Échelle Physique

Modèle initial

Aucun effet significatif du dosage n'est observé, soit directement ou en interaction avec le groupe et/ou le temps.

Modèle ajusté

Le modèle ajusté ne produit pas le même patron de résultats que le modèle initial. Le seul effet observé est l'émergence d'un effet négatif du dosage à 12 mois pour le groupe témoin en garderie (voir les effets DDD au tableau 9.4). Nous accordons peu de crédibilité à ce résultat car l'effet est faible, il n'était pas anticipé et il semble être spécifique à la période d'évaluation de 12 mois.

Contre-validation non-paramétrique

Pour l'analyse non-paramétrique, la performance des enfants observée à l'échelle Physique est exprimée en termes de la probabilité d'être classifiée dans un quintile supérieur : la cote de réussite. Les effets à reproduire sont les estimations DDD testant la différenciation de l'effet du dosage selon les groupes de traitement.

Relativement au groupe témoin en garderie, le rapport des cotes pour l'effet DDD est de 0,99 (IC 90 % de 0,94 à 1,05), 0,99 (IC 90 % de 0,91 à 1,08) et 0,90 (IC 90 % de 0,84 à 0,96) respectivement pour les évaluations à quatre, huit et 12 mois. Les résultats de l'analyse reproduisent la signification des effets DDD observés à 12 mois. Le modèle estime que le rapport des cotes associé à l'impact du programme augmente de 10 % pour chaque heure de plus passée en garderie que la moyenne hebdomadaire.

Échelle Vocabulaire réceptif

Modèle initial

Selon l'estimateur DD, un effet positif du dosage est observé au niveau de base. Par ailleurs, cet effet ne varie pas significativement en fonction du temps, du groupe de traitement ou de la combinaison de temps et de groupe de traitement.

Modèle ajusté

Le modèle ajusté produit un effet nul du dosage au niveau de base, émergeant significatif, selon l'estimateur DD, à quatre mois pour ensuite demeurer stable. Selon l'estimateur DDD, l'effet positif du dosage est significatif en faveur du groupe programme seulement pour la période d'évaluation à quatre mois. Il n'y a aucune différence significative de l'effet dosage à huit et 12 mois. Autre que cet effet significatif à quatre mois, l'effet du dosage sur le Vocabulaire réceptif ne varie pas significativement en fonction du type de programme reçu.

Contre-validation non-paramétrique

Pour l'analyse non-paramétrique, la performance des enfants observée à l'échelle du Vocabulaire réceptif est exprimée en termes de la probabilité d'être classifiée dans un quintile supérieur : la cote de réussite. Les effets à reproduire sont les estimations DDD testant la différenciation de l'effet du dosage selon le groupe.

Relativement au groupe témoin en garderie, le rapport des cotes pour l'effet DDD est de 0,91 (IC 90 % de 0,86 à 0,96), 0,93 (IC 90 % de 0,87 à 0,99) et 0,95 (IC 90 % de 0,90 à 0,99) respectivement pour les périodes d'évaluation à quatre, huit et 12 mois. Ces résultats reproduisent partiellement la signification des effets DDD observés à quatre mois avec un nouvel effet significatif détecté à la période d'évaluation de 12 mois. Ce patron de résultats est une indication que l'analyse paramétrique est moins sensible pour cette période d'évaluation. Le modèle estime que le rapport des cotes associé à l'impact du programme augmente de 5 % pour chaque heure de plus passée en garderie que la moyenne du groupe en garderie durant la période précédant l'évaluation à 12 mois. L'inverse est également estimé à partir des cotes associé à l'impact du programme. Les effets du programme diminuent de 5 % pour chaque heure de moins passée en garderie que la moyenne hebdomadaire des enfants en garderie durant la période précédant l'évaluation à 12 mois.

Échelle Vocabulaire expressif

Modèle initial

Les effets observés selon l'estimateur DD indiquent un effet positif du dosage émergeant à quatre et à 12 mois. Toutefois, les effets DDD sont significatifs seulement à quatre mois en faveur des enfants du groupe programme. Ensemble, ces résultats suggèrent que l'effet du dosage est statistiquement équivalent pour les deux groupes à 12 mois.

Modèle ajusté

Le modèle ajusté reproduit les résultats de modèle initial. Selon les effets DDD, nous observons des effets positifs significatifs du dosage émergeant à l'évaluation de quatre et de 12 mois en faveur du groupe programme. En d'autres mots, la taille de l'effet du programme varie en fonction du dosage. Dix heures supplémentaires par semaine en garderie programme entraîneraient, selon le modèle, une augmentation d'environ 0,30 écart-type ou un gain de 2,7 mois dans le développement du vocabulaire expressif (prenant comme référence les données de Hill et collègues, 2008). Le même résultat n'est pas obtenu avec les enfants du groupe témoin en garderie, pour qui le dosage n'a aucun effet significatif décelable.

Contre-validation non-paramétrique

Pour l'analyse non-paramétrique, la performance des enfants observée avec l'échelle Vocabulaire expressif est exprimée en termes de la probabilité d'être classifiée dans un quintile supérieur : la cote de réussite. Les effets à reproduire sont les estimations DDD testant la différenciation de l'effet du dosage selon le groupe.

Relativement au groupe témoin en garderie, le rapport des cotes pour l'effet DDD est de 0,92 (IC 90 % de 0,88 à 0,96), 1,00 (IC 90 % de 0,94 à 1,07) et 0,92 (IC 90 % de 0,86 à 0,99) en faveur du groupe programme respectivement pour les évaluations à quatre, huit et 12 mois. Ces résultats reproduisent partiellement la signification des effets DDD observés dans le modèle ajusté pour les périodes d'évaluation à quatre et à 12 mois. Le modèle indique que pour ces périodes, le rapport de cote de l'effet programme est augmenté de 8 % pour chaque heure de plus passée en garderie que la moyenne des enfants en garderie. L'inverse est également estimé à partir des cotes associées à l'impact du programme. Les effets du programme diminuent de 8 % pour chaque heure de moins passée en garderie que la moyenne des enfants en garderie durant les périodes d'évaluation à quatre et à 12 mois.

Sommaire : Analyses par dosage

L'analyse des données de la première cohorte rapportée dans le *Rapport des résultats de la première cohorte* (2014) révèle un effet différencié du dosage uniquement pour le domaine de la Communication, et ce, seulement pour l'évaluation à quatre mois (« le temps deux » selon la terminologie du *Rapport des résultats de la première cohorte*). L'ajout de la deuxième cohorte à l'échantillon augmente sensiblement les effectifs pour les deux groupes en garderie (se chiffrant maintenant à plus de 100 par groupe), ce qui a pour effet d'augmenter la fidélité des estimations de l'effet du dosage pour chaque groupe. La stabilité conférée par l'ajout de la deuxième cohorte a porté fruit puisque nous détectons maintenant des effets du dosage (tel qu'observé dans la littérature portant sur la petite enfance) et, plus intéressant encore, des effets différenciés selon le groupe.

Pour les domaines de la Communication, la Conscience de soi, la Cognition et le Vocabulaire expressif, on observe un effet accru du programme à l'entrée de la maternelle de quatre ans (c.-à-d. la période d'évaluation à 12 mois) d'environ 0,30 écart-type pour chaque 10 heures de plus passées en garderie que la moyenne hebdomadaire (27 + 10 = 37 heures par semaine). En d'autres mots, cet effet s'additionne à l'effet d'être membre du groupe programme observé dans l'analyse par groupe (voir les effets à 12 mois rapportés au tableau 9.2). Pour la Communication, l'effet du programme pour les enfants à dosage intense serait d'environ 0,70 écart-type, pour la Conscience de soi et 0,50 écart-type pour le Vocabulaire expressif. Notons que l'effet du programme est symétrique. Ainsi, l'effet du programme à 12 mois diminue d'environ 0,30 écart-type pour chaque 10 heures de moins passées en garderie que la moyenne hebdomadaire, 27 – 10 = 17. En d'autres mots, l'effet du programme testé est nul pour les enfants qui reçoivent un dosage relativement faible. Le patron de ces résultats suggère que le programme en garderie est principalement responsable pour l'impact du programme.

En somme, les résultats donnent à penser qu'environ 25 à 30 heures en garderie par semaine sont nécessaires pour obtenir un impact significatif du programme. De plus, un impact accru est observé pour les enfants qui passent plus de temps encore en garderie (environ 40 heures semaine). Selon les normes développementales de Hill et collègues (2008), l'accentuation de

l'effet programme associé à une augmentation de 10 heures semaine par rapport à la moyenne hebdomadaire équivaudrait à peu près à 2,7 mois de développement.

9.2.4. Analyses par qualité et fidélité du programme en garderie

Cette section présente les analyses examinant l'effet de la qualité et de la fidélité du programme en garderie sur le développement des enfants. Les résultats de trois séries d'analyses sont présentés dans les sections suivantes : (a) une vérification quantitative de la différenciation des groupes expérimentaux sur le plan de leur programme en garderie, (b) une vérification de la valeur prédictive des dimensions clés du programme testé sur le plan de la fidélité et de la qualité de la mise en œuvre, et (c) une analyse de médiation vérifiant si la mise en œuvre du programme en garderie est responsable, en tout ou en partie, des effets du programme découlant des analyses par groupes expérimentaux (rapportés au tableau 9.2). L'objectif de ces analyses est d'estimer l'importance relative du programme en garderie dans la détermination des impacts observés du programme testé lors de la première année de sa mise en œuvre. Nous rappelons au lecteur que les données de fidélité/qualité décrivant le programme reçu par les enfants de la deuxième cohorte sont traitées dans ces analyses comme si elles provenaient de garderies différentes.

Différenciation du programme en garderie

Les résultats de l'analyse examinant la différenciation des programmes en garderie pour les garderies programme et les garderies témoins sont maintenant présentés. Le lecteur est prié de noter que la puissance statistique de ces tests n'est pas grande.⁷⁴ Face à une puissance statistique faible, il devient important de tenir compte de la taille des effets observés lors de l'interprétation des résultats. Nous rapportons au tableau 9.5 l'analyse de la qualité et au tableau 9.6, l'analyse de la fidélité. Un test paramétrique (test t) et non paramétrique (test Z de rang Wilcoxon) est rapporté pour chaque indice. Afin de déterminer la contribution unique de ces indices aux analyses à venir, nous rapportons aussi leurs inter-corrélations au tableau 9.7.

Qualité

Les analyses révèlent que les garderies programme se distinguent des garderies témoins sur deux dimensions de la qualité, notamment la qualité structurelle et la qualité de lecture. La taille de ces effets est grande, excédant 0,75 écart-type, et est reproduite par l'analyse non-paramétrique. La qualité éducative se place au troisième rang en ordre d'importance avec une différenciation de 0,68 écart-type. Enfin, la différenciation des garderies programme et des garderies témoins au niveau de la sensibilité des éducatrices est négligeable à 0,26 écart-type.

Fidélité

L'analyse examinant la différenciation de la fidélité de programme indique que les garderies du groupe programme se distinguent principalement par rapport à la fidélité structurelle ($ds > 2$) La différenciation de la fidélité de contenu est néanmoins importante, variant entre 0,61 et 1,20 écart-type. Pour cet indice, la différenciation à 8 mois était la plus robuste.

⁷⁴ La probabilité de détecter un effet de 0,50 écart-type est inférieure à deux fois sur cinq essais pour cette analyse.

Tableau 9.5 : La différenciation de la qualité du programme testé

Indice	Garderies programme (N=6)	Garderies témoins (N=17)	Différenciation (taille d'effet d^a)	Résultats	
	M (É.-T.)	M (É.-T.)		Test t robuste (d.l.)	Test Z de rang
Qualité structurelle	6,51 (0,48)	5,71 (1,02)	+0,80 (0,78)	2,56** (18)	1,83*
Qualité éducative	5,33 (0,88)	4,49 (1,24)	+0,84 (0,68)	1,81* (12)	1,63
Sensibilité de l'éducatrice	5,92 (1,86)	5,29 (2,41)	+0,62 (0,26)	0,65 (11)	0,62
Qualité de lecture	5,50 (1,34)	2,15 (1,46)	+3,35 (2,29)	5,15*** (9)	3,36***

Note : Le seuil de signification est indiqué à *** $p < 1\%$, ** $p < 5\%$; * $p < 10\%$. ^aLa valeur d , la taille standardisée de l'effet, est calculée à partir de l'écart-type du groupe témoin en garderie. M = moyenne, É.-T. = écart-type, d.l. = degrés de liberté.

Tableau 9.6 : La différenciation de la fidélité du programme testé

Fidélité structurelle	Garderies programme (N=6)	Garderies témoins (N=17)	Différenciation (taille d'effet d^a)	Résultats	
	M (É.-T.)	M (É.-T.)		Test t robuste (d.l.)	Test Z de rang
Niveau de base	0,77 (0,13)	0,44 (0,14)	+0,33 (2,36)	5,29*** (9)	3,35***
4 mois	0,87 (0,14)	0,44 (0,14)	+0,43 (3,07)	6,47*** (9)	3,46***
8 mois	0,85 (0,11)	0,45 (0,19)	+0,41 (2,16)	6,56*** (15)	3,37***
Fidélité de contenu	M (É.-T.)	M (É.-T.)	Différenciation (taille d'effet d^a)	Test t robuste (d.l.)	Test Z de rang
Niveau de base	0,75 (0,07)	0,63 (0,18)	+0,11 (0,61)	2,17** (20)	0,84
4 mois	0,79 (0,10)	0,64 (0,18)	+0,15 (0,83)	2,45** (15)	1,54
8 mois	0,83 (0,08)	0,65 (0,15)	+0,18 (1,20)	3,61** (15)	2,42**

Note : Le seuil de signification est indiqué à *** $p < 1\%$, ** $p < 5\%$; * $p < 10\%$. ^aLa valeur d , la taille standardisée de l'effet, est calculée ici à partir de l'écart-type du groupe témoin en garderie. M = moyenne, É.-T. = écart-type, d.l. = degrés de liberté.

Corrélations entre les indices de qualité et de fidélité

On retrouve au tableau 9.7 les statistiques descriptives pour l'échantillon au complet ainsi que les corrélations de Pearson pour chaque paire d'indice de qualité/fidélité considérée dans les analyses d'impact. Notons que la probabilité de détecter une corrélation réelle de 0,20 dans cette analyse est seulement de 0,23, soit une fois sur cinq essais. Par contre, une corrélation réelle de 0,50 serait détectée quatre fois sur cinq, ce qui correspond à une probabilité de 0,80.

L'échantillon de 23 garderies nous permet de présenter une analyse descriptive des corrélations observées. Dans un premier temps, il est clair que l'information captée par ces indices ne se résume pas par un seul facteur. Les indices semblent se regrouper en deux facteurs au moins, l'un centré sur la qualité éducative et l'autre sur la qualité de lecture. La qualité éducative est corrélée significativement avec la qualité structurelle, la fidélité structurelle, la fidélité de contenu et la sensibilité des éducatrices. En d'autres mots, chacun de ces indices

partage de la variance avec la qualité éducative, ce qui donne à penser qu'ils ont une source commune. La qualité de lecture est corrélée significativement avec la fidélité structurelle et la fidélité de contenu, ce qui suggère encore une fois une source de variance commune.

Tableau 9.7 : Inter-corrélation des indices de la qualité et de la fidélité de programme en garderie

Indice	Statistiques Descriptives		Corrélations de Pearson				
	M	É.-T.	1	2	3	4	5
1. Qualité structurelle	5,92	0,97					
2. Qualité éducative	4,71	1,19	0,51***				
3. Sensibilité de l'éducatrice	5,46	2,26	0,17	0,53***			
4. Qualité de la lecture	3,02	2,05	0,18	0,16	0,17		
5. Fidélité structurelle ^a	0,55	0,25	0,23	0,45**	0,34	0,68***	
6. Fidélité de contenu ^a	0,70	0,16	0,23	0,50***	0,28	0,76***	0,62***

Note : N = 23 garderies. Le seuil de signification est indiqué à *** $p < 1\%$, ** $p < 5\%$; * $p < 10\%$. ^a Indice provenant des observations prises durant la période précédant la période de l'évaluation de 8 mois. M = moyenne, É.-T. = écart-type.

Impact de la fidélité et de la qualité

Les résultats des analyses par qualité et par fidélité sont présentés respectivement aux tableaux 9.8 et 9.9. Nous décrivons les résultats des analyses pour chaque indice individuellement dans les paragraphes qui suivent.

Qualité structurelle

Au niveau de base, la relation est significativement négative entre la qualité structurelle et deux des résultantes, soit la Cognition et le Vocabulaire réceptif. Ensuite, nous observons à l'entrée à la maternelle de quatre ans (c.-à-d. la période d'évaluation à 12 mois) des gains supérieurs (DD) pour les garderies caractérisées par une meilleure qualité structurelle pour la Communication, la Cognition, le Vocabulaire réceptif et le Vocabulaire expressif. La taille de cet effet DD est d'à peu près 0,20 écart-type pour chaque 'point' de qualité, ce qui est plus que suffisant pour renverser la relation négative observée au niveau de base pour certaines des résultantes.

Qualité éducative

Nous observons une relation négative au niveau de base entre la qualité éducative et le score moyen obtenu aux échelles de Communication, Conscience de soi, et Vocabulaire expressif. Ceci indique que les garderies de haute qualité ont tendance à desservir des enfants présentant des difficultés sur ces résultantes. Les effets DD indiquent que cette association se corrige puisque les gains relatifs au niveau de base pour la Communication et le Vocabulaire expressif sont positivement reliés à la qualité éducative à la période d'évaluation de 12 mois. Des effets sont observés plus tôt pour la Conscience de soi (période d'évaluation à 4 mois) et la Cognition (périodes d'évaluation à 4 et à 8 mois). Notez que les effets DD sont non-significatifs pour les échelles Physique et Vocabulaire réceptif. Selon le modèle rapporté, une différence de qualité éducative de un point au préscolaire est associée à un gain au niveau de la Communication et du

Vocabulaire expressif de 0,12 écart-type à l'entrée à la maternelle de quatre ans (période d'évaluation à 12 mois).

Sensibilité des éducatrices

La sensibilité des éducatrices n'est pas associée significativement à aucune des résultantes au niveau de base. Par ailleurs, les effets DD indiquent que les gains développementaux observés aux évaluations suivantes sont associés à la sensibilité des éducatrices seulement pour le Vocabulaire expressif et ce, uniquement à la période d'évaluation de 12 mois. Selon le modèle rapporté, une augmentation de sensibilité au préscolaire de un point est associée à un gain de 0,17 écart-type au niveau du Vocabulaire expressif à l'entrée à la maternelle de quatre ans.

Qualité de la lecture

Au niveau de base, nous observons une relation négative significative entre la qualité de la lecture et le score moyen obtenu aux échelles de Communication, Physique, et Vocabulaire expressif. Lors de la rentrée à la maternelle de quatre ans (période d'évaluation à 12 mois), nous notons un effet positif pour le domaine Physique seulement. Pour les résultantes Communication, Conscience de soi, Cognition et Vocabulaire expressif, une influence positive et significative sur les gains développementaux est décelée lors de la période d'évaluation à quatre mois bien que la significativité de cet effet disparaît pour les autres périodes d'évaluation.

Qualité Globale

Au niveau de base, nous observons une relation négative entre la qualité globale et tous les indicateurs de préparation à la scolarisation sauf le Vocabulaire réceptif. Les estimations DD indiquent qu'à l'entrée à la maternelle quatre ans (période d'évaluation à 12 mois), la qualité globale du programme est positivement associée aux gains développementaux des enfants en garderie au niveau de la Communication et du Vocabulaire expressif. Selon le modèle présenté, une augmentation de un point sur l'échelle de qualité représente une augmentation d'environ 0,13 écart-type sur ces deux résultantes.

Fidélité structurelle

Au niveau de base, nous observons une association négative entre la fidélité structurelle et le score moyen obtenu aux échelles de Communication, Conscience de soi, Cognition, et Vocabulaire expressif. Des gains significatifs sont associés à la fidélité structurelle du programme pour la Communication et ce, à chaque période d'évaluation de la première année (voir les effets DD). À l'entrée de la maternelle de quatre ans (période d'évaluation à 12 mois) et selon le modèle proposé, nous estimons une différence de 0,30 écart-type au score de Communication entre une garderie dont la cote de fidélité est de 40% et une garderie dont la cote de fidélité est de 80%.⁷⁵ Des effets DD significatifs associés aux autres indicateurs de la préparation à la scolarisation ne sont pas observés à toutes les périodes d'évaluation.

⁷⁵Dans les tableaux, les valeurs rapportées sont des proportions. Ici, la même variable est exprimée en pourcentage. Il est à noter qu'une cote de fidélité de 0,80 est équivalente à une cote de fidélité de 80%.

Fidélité du contenu

Au niveau de base, nous observons une association négative significative entre la fidélité du contenu et le score moyen obtenu à l'échelle de Communication. Les estimateurs DD révèlent un seul effet positif significatif du programme, soit à la période d'évaluation de quatre mois pour le Vocabulaire expressif. Dans l'ensemble, les analyses d'impact de la fidélité du contenu tendent à infirmer l'hypothèse selon laquelle le contenu du programme est un déterminant important du développement des enfants francophones d'âge préscolaire. Si le programme a un effet significatif sur le développement, c'est par le biais d'un autre élément du programme.

Fidélité globale

Au niveau de base, nous observons des associations négatives significatives entre la fidélité globale et le score moyen obtenu aux échelles de Communication, Conscience de soi, et Vocabulaire expressif. Les estimateurs DD révèlent un effet significatif et stable de la fidélité globale sur les gains développementaux en Communication. Selon le modèle proposé, une augmentation de 40 % à 80 % de la fidélité globale est associée à un accroissement d'environ 0,32 écart-type en Communication.

Contre-validation non-paramétrique

Les analyses des effets de la qualité et de la fidélité sont d'importance secondaire, donc un compte rendu détaillé de la contre-validation de ces analyses n'est pas présenté. De façon générale, les effets associés aux indices de fidélité sont très robustes, étant confirmés par l'analyse non-paramétrique. Par contre, les effets associés aux indices de qualité se sont avérés moins robustes. Parmi ces derniers, l'indice de qualité de la lecture était le prédicteur le plus robuste et le Vocabulaire réceptif était la résultante la plus souvent associée aux indices de qualité.

Tableau 9.8 : L'effet de la qualité du programme en garderie avec ajustement (1^{ère} année) — Scores standardisés

Type de différence	Indicateur de la préparation à la scolarisation											
	Communication		Conscience de soi		Cognition		Physique		Vocabulaire réceptif		Vocabulaire expressif	
	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.
Qualité structurelle												
Niveau de base	-0,16	0,09	-0,07	0,08	-0,13*	0,06	-0,10	0,08	-0,19***	0,05	-0,11	0,10
4 mois (DD)	0,18	0,11	0,16***	0,05	0,12	0,09	0,17**	0,06	0,16*	0,08	0,24**	0,10
8 mois (DD)	0,13*	0,07	0,06	0,07	0,06	0,10	0,06	0,06	0,17***	0,05	0,10*	0,05
12 mois (DD)	0,24*	0,12	0,12	0,07	0,15**	0,06	0,11	0,07	0,23***	0,07	0,22*	0,11
<i>Groupe x Temps Wald F</i>	1,33		2,98*		3,86**		2,73*		3,63**		1,92	
Qualité éducative												
Niveau de base	-0,11**	0,04	-0,09*	0,05	-0,13***	0,04	-0,02	0,05	-0,07	0,05	-0,12**	0,04
4 mois (DD)	0,09*	0,04	0,11***	0,03	0,08*	0,04	0,05	0,07	-0,05	0,06	0,22***	0,04
8 mois (DD)	0,06	0,06	0,01	0,06	0,11*	0,06	0,05	0,08	0,02	0,06	0,02	0,07
12 mois (DD)	0,12*	0,06	0,06	0,06	0,06	0,04	0,05	0,06	0,07	0,07	0,12*	0,06
<i>Groupe x Temps Wald F</i>	1,38		6,32***		1,14		0,38		1,47		9,71***	
Sensibilité des éducatrices												
Niveau de base	-0,03	0,03	-0,01	0,02	-0,03	0,02	-0,01	0,03	-0,01	0,02	-0,04	0,03
4 mois (DD)	0,05*	0,02	0,04**	0,01	0,00	0,02	0,02	0,04	0,01	0,04	0,09**	0,03
8 mois (DD)	0,02	0,05	-0,02	0,02	-0,03	0,02	0,01	0,03	-0,02	0,03	0,01	0,03
12 mois (DD)	0,05	0,05	0,01	0,02	-0,03	0,02	-0,03	0,03	0,01	0,03	0,07*	0,03
<i>Groupe x Temps Wald F</i>	5,00**		2,91*		0,61		0,76		0,50		3,82**	
Qualité de lecture												
Niveau de base	-0,03*	0,01	-0,02	0,01	-0,02	0,02	-0,05*	0,02	0,01	0,02	-0,04**	0,01
4 mois (DD)	0,04**	0,02	0,06***	0,01	0,03***	0,01	-0,01	0,04	-0,03	0,04	0,09***	0,02
8 mois (DD)	0,01	0,02	0,03	0,02	-0,01	0,02	0,05	0,03	0,02	0,02	0,01	0,03
12 mois (DD)	0,01	0,02	0,01	0,02	0,01	0,01	0,07***	0,01	0,01	0,02	0,01	0,02
<i>Groupe x Temps Wald F</i>	2,34		11,12***		7,75***		7,80***		1,18		5,68***	
Qualité globale												
Niveau de base	-0,13***	0,04	-0,10*	0,05	-0,12**	0,04	-0,09*	0,04	-0,04	0,04	-0,15**	0,06
4 mois (DD)	0,17***	0,04	0,20***	0,03	0,10**	0,04	0,04	0,09	-0,03	0,07	0,32***	0,05
8 mois (DD)	0,07	0,07	0,06	0,06	0,01	0,05	0,10*	0,05	0,04	0,05	0,04	0,08
12 mois (DD)	0,14*	0,07	0,09	0,05	0,02	0,04	0,10	0,06	0,08	0,05	0,13*	0,07
<i>Groupe x Temps Wald F</i>	6,45***		12,51***		2,99*		1,56		2,14		11,00***	

Note : Contrairement aux analyses par groupes, les valeurs positives pour les effets DD indiquent une relation positive entre l'élément de qualité et l'indicateur de la préparation à la scolarisation. Le niveau de base sert de référence dans les analyses. Les degrés de liberté du test Wald F sont 3 et 15. La qualité globale est la moyenne des quatre indicateurs de qualité. Les erreurs types robuste de Huber-White sont rapportées avec regroupement par garderie. Le seuil de signification est indiqué à *** $p < 1\%$, ** $p < 5\%$; * $p < 10\%$. Diff = différence, E.T. = erreur type.

Tableau 9.9 : L'effet de la fidélité du programme en garderie avec ajustement (1^{ière} année) – Scores standardisés

Type de différence	Indicateur de la préparation à la scolarisation											
	Communication		Conscience de soi		Cognition		Physique		Vocabulaire réceptif		Vocabulaire expressif	
Fidélité structurelle	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.
Niveau de base	-0,64***	0,18	-0,50***	0,16	-0,40*	0,21	-0,40	0,30	-0,12	0,22	-0,71***	0,18
4 mois (DD)	0,77***	0,22	0,76***	0,13	0,35*	0,18	0,01	0,39	-0,26	-0,26	1,13***	0,21
8 mois (DD)	0,74**	0,27	0,60**	0,28	0,19	0,23	0,66**	0,25	0,39	0,26	0,57	0,33
12 mois (DD)	0,74**	0,30	0,32	0,20	0,24	0,19	0,49*	0,24	0,29	0,25	0,40	0,29
<i>Groupe x Temps Wald F</i>	3,62**		14,15***		1,41		3,52**		3,02*		11,62***	
Fidélité du contenu	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.
Niveau de base	-0,53**	0,23	-0,28	0,29	-0,17	0,30	0,15	0,26	0,04	0,32	-0,58	0,40
4 mois (DD)	0,40	0,27	0,59	0,37	0,16	0,38	-0,36	0,79	-10,05*	0,52	10,63***	0,52
8 mois (DD)	0,03	0,27	-0,06	0,33	-0,40	0,51	0,01	0,29	-0,05	0,25	-0,29	0,45
12 mois (DD)	0,09	0,28	-0,23	0,35	-0,29	0,45	0,35	0,55	-0,01	0,38	-0,24	0,36
<i>Groupe x Temps Wald F</i>	0,78		2,34		0,68		0,24		1,52		8,07***	
Fidélité globale	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.
Niveau de base	-0,83***	0,27	-0,63**	0,24	-0,46	0,30	-0,43	0,34	-0,11	0,30	-0,91***	0,29
4 mois (DD)	0,92**	0,33	0,99***	0,21	0,41	0,28	-0,03	0,59	-0,56	0,49	1,65***	0,31
8 mois (DD)	0,74*	0,38	0,60	0,40	0,08	0,38	0,70**	0,31	0,34	0,28	0,49	0,49
12 mois (DD)	0,77*	0,37	0,30	0,29	0,17	0,29	0,68**	0,31	0,27	0,30	0,36	0,37
<i>Groupe x Temps Wald F</i>	2,37		9,39***		1,15		2,75*		3,02*		8,81***	

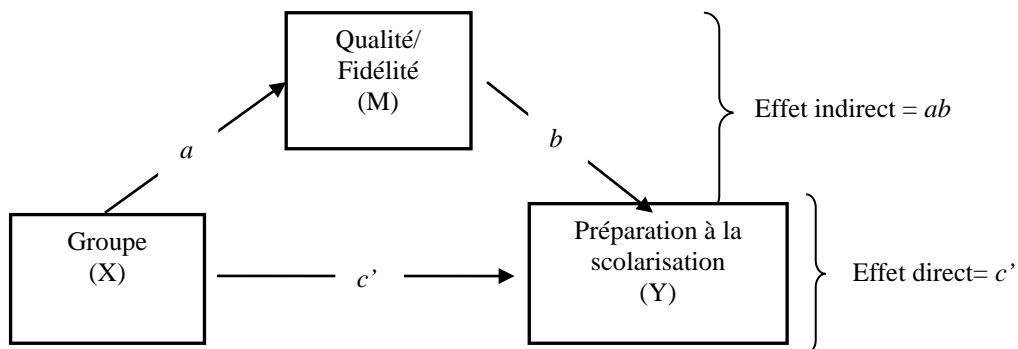
Note : Contrairement aux analyses par groupes, les valeurs positives pour les effets DD indiquent une relation positive entre l'élément de fidélité et l'indicateur de la préparation à la scolarisation. Le niveau de base sert de référence dans les analyses. Les degrés de liberté du test Wald F sont 3 et 15. La fidélité globale est la moyenne des scores de fidélité structurelle et de fidélité du contenu. Les erreurs types robuste de Huber-White sont rapportées avec regroupement par garderie. Le seuil de signification est indiqué à *** $p < 1$ %, ** $p < 5$ %; * $p < 10$ %. Diff = différence, E.T. = erreur type.

Effet médiateur de la fidélité et de la qualité

Cette section présente les résultats d'une analyse de médiation dont l'objectif est de déterminer dans quelle mesure la fidélité et la qualité du programme en garderie sont responsables des effets du programme testé. Les analyses effectuées examinent si l'effet du programme peut être expliqué complètement ou partiellement par le biais des indices de qualité/fidélité. Dans la négative, nous devons conclure que d'autres aspects du programme seraient responsables des effets observés sur les indicateurs de préparation à la scolarisation (p. ex. l'atelier famille).

Nous prenons la variable de groupe comme facteur explicatif exogène (X), la qualité et la fidélité comme variables médiatrices endogènes (M) et les variables de préparation à la scolarisation comme résultantes (Y). L'analyse de médiation se base sur trois paramètres et leurs erreurs type : le coefficient a captant la relation entre X et M, le coefficient b captant la relation entre M et Y et le coefficient c' captant l'effet résiduel de X sur Y après avoir éliminé la variance associée à la variable M. Un diagramme illustrant la logique qui sous-tend cette analyse est présenté à la figure 9.1 (voir aussi Krull et MacKinnon, 1999).

Figure 9.1 : Diagramme illustrant les deux aspects par l'entremise desquelles le programme exerce son effet sur la préparation à la scolarisation : l'aspect attribuable à la qualité et la fidélité du programme en garderie (a, b) et celui qui ne l'est pas (c').



À la figure 9.1, deux effets d'intérêts sont distingués. L'*effet indirect* du programme capte la partie de l'effet du programme testé qui est attribuable à la qualité/fidélité du programme en garderie. Cet effet est estimé en introduisant la variable 'fidélité' dans le modèle de régression. La partie de l'effet du groupe qui est éliminé dans le modèle ajusté par l'inclusion de la fidélité est attribué à cet élément du programme testé. L'effet résiduel du groupe est l'*effet direct* du programme, c'est à dire la partie de l'effet qui est attribuable à d'autres aspects du programme. Si ce dernier effet est non-significatif, alors nous présumons que tout l'effet du programme passe par sa qualité/fidélité. Par contre, si cet effet est significatif, alors nous supposons qu'il reste une partie réelle de l'effet programme à expliquer (p. ex. par l'effet de l'Atelier famille). Pour simplifier la présentation, nous rapportons uniquement l'analyse de médiation à 12 mois lors de l'entrée des enfants à la maternelle de quatre ans et uniquement pour les indices de

qualité/fidélité les plus fortement prédictifs du développement des enfants. Les résultats sont rapportés aux tableaux 9.10 et 9.11 pour la qualité et la fidélité respectivement.⁷⁶

Qualité

L'indice de qualité structurelle était le plus performant dans la prédiction du développement des enfants (voir le tableau 9.8). Cet indice est utilisé comme médiateur dans les analyses rapportées au tableau 9.10. Des effets indirects significatifs sont observés pour toutes les résultantes ayant été influencées significativement par le programme testé lors des analyses par groupe de traitement. En effet, la médiation est complète dans tous les cas sauf pour la Communication (voir les effets directs). Ce résultat signifie que la qualité structurelle est suffisante pour expliquer l'effet du programme testé sur toutes les résultantes sauf la Communication. Pour la Communication seulement, il reste une partie de l'effet du programme à expliquer.

Tableau 9.10 : Test de médiation prenant la qualité comme médiateur de l'effet programme à 12 mois

Résultante	Effet direct			Effet indirect	
	Effet DD à 12 mois	E.T.	% réduction	Test t	E.T. Sobel
Communication ^a	-0,30**	0,12	-30.2 %	1,57*	0,11
Conscience de soi ^b	SO	SO	SO	SO	SO
Cognition ^a	-0,03	0,08	-81.3 %	1,77**	0,07
Vocabulaire réceptif ^a	-0,36	0,43	Supp.	2,00**	0,08
Vocabulaire expressif ^a	-0,17	0,11	-34.6 %	1,56*	0,11

Note : Les valeurs négatives des estimations DD pour toutes les comparaisons entre les groupes représentent un effet traitement positif (c.-à-d. un avantage pour le groupe programme). L'effet DD est relatif au groupe témoin en garderie. Supp. = Effet de suppression voulant que l'effet du programme soit plus fort après avoir contrôlé pour la qualité. Ceci veut dire qu'il existe une faible corrélation négative entre la qualité et un autre élément utile du programme (p. ex., l'impact de l'atelier famille sur les parents). L'effet de la qualité avait pour effet de masquer l'effet positif de cet aspect inconnu du programme testé (à 12 mois). ^a Médiateur est la qualité structurelle. ^b Aucun indicateur de qualité n'était prédictif du développement de la Conscience de soi à 12 mois. Le seuil de signification est indiqué à *** $p < 1\%$, ** $p < 5\%$; * $p < 10\%$. Diff = différence, E.T. = erreur type, SO = sans objet.

Fidélité

La fidélité structurelle était l'indice de fidélité la plus performante dans la prédiction du développement des enfants (voir le tableau 9.9). Cet indice est donc utilisé comme variable médiatrice pour les analyses rapportées au tableau 9.11. Le lecteur est prié de noter que seuls les résultats pour la Communication sont présentés en raison de l'absence d'une association significative entre la fidélité et les autres résultantes à 12 mois. Les résultats présentés au tableau 9.11 indiquent que la fidélité structurelle explique une part significative de l'effet du programme (voir l'effet indirect) bien qu'une partie non-négligeable de l'impact du programme reste à expliquer (voir l'effet direct). C'est essentiellement le même patron de résultats obtenus avec la qualité.

⁷⁶ Une analyse considérant simultanément ces deux dimensions de la mise-en-œuvre n'est pas entreprise en raison du nombre relativement restreint de garderies. Suivant les recommandations de Krull et MacKinnon (1999), nous utilisons la formule de Sobel pour calculer l'erreur type de l'effet indirect.

Tableau 9.11 : Test de médiation prenant la fidélité comme médiateur de l'effet programme à 12 mois

Résultante	Effet direct			Effet indirect	
	Effet DD à 12 mois	E.T.	% réduction	test t	E.T. Sobel
Communication ^a	-0,57**	0,26	Supp.	2,43***	0,66

Note : Les valeurs négatives des estimations DD pour toutes les comparaisons entre les groupes représentent un effet traitement positif (c.-à-d. un avantage pour le groupe programme). L'effet DD est relatif au groupe témoin en garderie. Supp. = Effet de suppression voulant que l'effet du programme soit plus fort après avoir contrôlé pour la fidélité. Ceci veut dire qu'il existe une faible corrélation négative entre la fidélité et un autre élément utile du programme (p. ex., l'impact de l'atelier famille sur les parents). L'effet de la qualité avait pour effet de masquer l'effet positif de cet aspect inconnu du programme testé (à 12 mois).^a Le médiateur est la fidélité structurelle. Les autres résultantes ne sont pas présentées puisqu'elles n'étaient pas prédites par les indices de fidélité. Le seuil de signification est indiqué à *** $p < 1\%$, ** $p < 5\%$; * $p < 10\%$. Diff = différence, E.T. = erreur type.

Sommaire : Effet médiateur de la fidélité et de la qualité

La mise en œuvre du programme en garderie explique tous les effets du programme observés à 12 mois sauf celui de la Communication. Pour cette résultante, une partie de l'effet du programme testé reste à expliquer. Il est intéressant de noter que la qualité et la fidélité structurelle étaient les indicateurs de mise en œuvre les plus performants. D'ailleurs, c'est la raison pour laquelle ces indices ont été considérés à titre de médiateurs lors de cette série d'analyses. Par contre, la dimension sur laquelle le programme testé se distinguait le plus (c.-à-d. la qualité de lecture) n'était pas prédictive du rendement en Communication ou en Vocabulaire expressif à 12 mois. Par conséquent, cet indice de qualité ne pouvait pas servir de variable médiatrice.

9.2.5. Analyses par profil linguistique

Le profil linguistique des enfants peut être défini de plusieurs façons. Deux dimensions pertinentes à prendre en considération incluent de bien capter la participation active de l'enfant lors des échanges linguistiques (voir le *Rapport des résultats de la première cohorte*, 2014) et la précision de la mesure de ces échanges (une dichotomie versus une variable continue). Pour le présent rapport, les analyses préliminaires ont établi la similarité dans le patron de résultats obtenus peu importe la variable linguistique utilisée (disponibles sur demande). En conséquence, nous avons opté de seulement rapporter les analyses basées sur un indicateur simple (dichotomique) calculé à partir du croisement de la langue parlée par l'enfant à sa mère et de la langue parlée par l'enfant à son père. Cet indicateur nommé Type de foyer comprend deux catégories : à forte exposition et à faible exposition. Cette dernière catégorie comprend toutes configurations linguistiques autres que deux parents francophones.

Les résultats sont rapportés en deux étapes : a) les résultats des estimations DDD testant l'effet modérateur du type de foyer sur l'impact du programme testé (voir le tableau 9.12) et b) les estimations DD de l'effet du programme calculé séparément pour les deux types de foyer (voir le tableau 9.13).

Tableau 9.12 : L'effet modérateur du Type de foyer sur les indicateurs de la préparation à la scolarisation

Test de modération	Indicateur de la préparation à la scolarisation											
	Communication		Conscience de soi		Cognition		Physique		Vocabulaire réceptif		Vocabulaire expressif	
	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.
G1 vs G2												
4 mois (DDD)	-0,19	0,13	0,11	0,12	0,23	0,18	-0,28	0,19	-0,47*	0,24	0,28	0,21
8 mois (DDD)	0,01	0,13	-0,01	0,17	0,52**	0,19	-0,05	0,13	-0,11	0,21	0,20	0,18
12 mois (DDD)	-0,25	0,14	-0,04	0,20	0,48**	0,19	0,30	0,27	-0,28	0,18	-0,08	0,24
G1 vs G3												
4 mois (DDD)	-0,29	0,20	0,05	0,09	0,59***	0,20	-0,29	0,27	-0,47*	0,27	0,11	0,20
8 mois (DDD)	-0,39**	0,18	0,10	0,17	0,51**	0,21	-0,07	0,25	-0,74***	0,22	-0,01	0,27
12 mois (DDD)	-0,59***	0,17	0,28**	0,12	0,54***	0,19	0,02	0,32	-0,63***	0,18	0,04	0,24

Note : Les effets DDD représente la différence dans l'impact du programme pour les enfants à faible et à forte exposition. La catégorie de référence est « faible exposition », ce qui veut dire que l'effet DDD doit être ajoutés à l'impact du programme pour ce sous-groupe afin d'obtenir l'effet du programme pour les enfants à forte exposition. *En d'autres mots, les effets DDD sont interprétés de la façon suivantes : (a) les effets positifs indiquent que l'effet du programme testé diminue avec une exposition forte au français (b) les effets négatifs indiquent que l'effet du programme testé est accentué avec une exposition plus forte au français. Comme les autres analyses, les valeurs négatives des estimations DD pour toutes les comparaisons entre les groupes représentent un effet traitement positif (c.-à-d. un avantage pour le groupe programme).* Les erreurs types robustes de Huber-White sont rapportées avec groupement d'erreurs par garderie. Le seuil de signification est indiqué à *** $p < 1\%$, ** $p < 5\%$; * $p < 10\%$. Diff = différence, E.T. = erreur type.

Tableau 9.13 : Impact du programme testé selon le Type de foyer : à forte ou à faible exposition au français — (1^{ère} année) — Scores ajustés

Indicateur de la préparation à la scolarisation												
<i>Familles à faible exposition</i>												
Type de différence	Communication		Conscience de soi		Cognition		Physique		Vocabulaire réceptif		Vocabulaire expressif	
	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.
G1 vs G2												
Niveau de base	0,48***	0,11	0,38**	0,14	0,21*	0,11	0,32**	0,12	0,34**	0,15	0,43***	0,13
4 mois (DD)	-0,55***	0,08	-0,30***	0,07	-0,03	0,11	-0,27*	0,14	-0,40**	0,17	-0,28*	0,15
8 mois (DD)	-0,37**	0,15	-0,33**	0,14	0,29**	0,13	-0,37**	0,15	-0,42*	0,21	-0,14	0,16
12 mois (DD)	-0,54***	0,12	-0,26*	0,14	0,09	0,12	-0,08	0,19	-0,42**	0,15	-0,29**	0,13
G1 vs G3												
Niveau de base	-0,08	0,14	-0,08	0,18	-0,18	0,12	-0,11	0,15	0,03	0,20	0,03	0,24
4 mois (DD)	-0,39***	0,13	-0,06	0,06	0,36**	0,16	-0,13	0,17	-0,31**	0,13	0,03	0,15
8 mois (DD)	-0,49***	0,16	-0,30**	0,14	0,22	0,17	-0,01	0,15	-0,48*	0,23	-0,17	0,18
12 mois (DD)	-0,71***	0,12	0,05	0,11	0,09	0,12	0,13	0,25	-0,59***	0,18	-0,04	0,12
<i>Familles à forte exposition</i>												
Type de différence	Communication		Conscience de soi		Cognition		Physique		Vocabulaire réceptif		Vocabulaire expressif	
	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.
G1 vs G2												
Niveau de base	0,16	0,12	0,14	0,12	0,13	0,12	0,24*	0,14	0,05	0,12	0,21	0,14
4 mois (DD)	-0,36**	0,15	-0,41***	0,09	-0,26*	0,15	0,01	0,18	0,07	0,27	-0,57***	0,16
8 mois (DD)	-0,38***	0,13	-0,32**	0,16	-0,23	0,14	-0,31**	0,12	-0,30**	0,14	-0,34*	0,18
12 mois (DD)	-0,29*	0,15	-0,22	0,15	-0,40***	0,12	-0,39**	0,14	-0,13	0,13	-0,21	0,20
G1 vs G3												
Niveau de base	-0,02	0,10	0,03	0,13	0,14	0,11	-0,08	0,13	-0,20*	0,11	-0,01	0,16
4 mois (DD)	-0,10	0,13	-0,11	0,09	-0,24	0,15	0,16	0,19	0,17	0,27	-0,08	0,13
8 mois (DD)	-0,09	0,10	-0,40**	0,16	-0,30**	0,13	0,07	0,15	0,27*	0,15	-0,16	0,21
12 mois (DD)	-0,12	0,12	-0,24*	0,11	-0,46***	0,09	0,11	0,15	0,05	0,07	-0,09	0,18

Note : Les valeurs négatives des estimations DD pour toutes les comparaisons entre les groupes représentent un effet traitement positif (c.-à-d. un avantage pour le groupe programme). L'exposition à la langue française est définie d'après le type de foyer, un indicateur croisant la langue parlée par l'enfant à son père avec la langue parlée par l'enfant à sa mère. Deux catégories sont distinguées : faible et forte exposition au français. Les erreurs types robuste de Huber-White sont rapportées avec groupement des erreurs par garderie. Le seuil de signification est indiqué à *** $p < 1\%$, ** $p < 5\%$; * $p < 10\%$. Diff = différence, E.T. = erreur type.

Communication

Pour la communication, les estimés DDD sont significatifs en faveur du groupe programme lorsque comparés au groupe témoin hors garderie (voir tableau 9.12). Un examen des effets du programme selon le Type de foyer (voir le tableau 9.13) indique que le programme testé a un effet marqué sur le développement de la Communication, mais seulement pour les enfants membres de familles à faible exposition. Pour la comparaison des groupes en garderie, les estimés DD (tableau 9.13) indiquent un effet positif du programme, peu importe le Type de foyer (voir aussi les effets DDD non-significatifs au tableau 9.12). L'examen approfondi du tableau 9.13 souligne un effet accru du programme pour les enfants à faible exposition. Dans l'ensemble, les résultats tendent à confirmer nos hypothèses : l'effet du programme testé est particulièrement prononcé chez les enfants à faible exposition au français dans le foyer.

Conscience de soi

L'association anticipée entre le Type de foyer et l'effet du programme testé sur la Conscience de soi n'est pas particulièrement évidente dans les résultats rapportés aux tableaux 9.12 et 9.13. L'absence de cette association est peut-être attribuable au fait que la Conscience de soi comprend des mesures de compétences langagières et cognitives. Les interactions prédites pour ces deux dimensions sont incompatibles, comme le suggère le résultat d'analyses indépendantes portant sur les dimensions Communication et Cognition. Il s'en suit que les effets du programme auraient tendance à s'annuler pour une échelle réunissant ces deux dimensions.

Cognition

Pour le domaine de la Cognition, les estimés DDD rapportés dans le tableau 9.12 indiquent clairement que l'effet du programme testé varie significativement selon le Type de foyer. Tel que prédit, les enfants issus de familles à forte exposition retirent des bénéfices sur le plan cognitif. La taille de l'effet est de l'ordre de 0,40 écart-type à l'entrée à la maternelle de quatre ans (période d'évaluation à 12 mois) peu importe le groupe témoin utilisé en référence.

Par contre, le programme testé semble nuire au développement cognitif d'enfants issus de familles à faible exposition. Quelques effets DD rapportés au tableau 9.13 indiquent un effet du programme significativement négatif pour ce sous-échantillon. Cependant, nous tenons à souligner que ces effets semblent transitoires et pour cette raison peu crédibles. Par exemple, l'effet du programme pour ces enfants à l'entrée à la maternelle de quatre ans s'approche de zéro, indiquant un effet nul du programme testé. Cette constatation est à confirmer lors des analyses d'impact de la deuxième année.

Vocabulaire réceptif

Selon les estimés DDD rapportés au tableau 9.12, l'effet du programme sur le Vocabulaire réceptif varie significativement selon le type de foyer. Cette contingence est particulièrement évidente pour la comparaison du groupe programme avec le groupe témoin hors garderie. Les estimés DD rapportés au tableau 9.13 confirment que ce sont les enfants issus de familles à faible exposition qui tirent profit du programme testé sur le plan du développement de leur vocabulaire.

Vocabulaire expressif

L'examen des estimés DDD rapportés au tableau 9.12 souligne l'absence d'une association entre le Type de foyer et l'impact du programme testé sur le Vocabulaire expressif. Les estimés DD rapportés au tableau 9.13 sont similaires pour les deux types de foyer, ce qui tend à confirmer l'absence de l'effet modérateur du Type de foyer. Il reste à voir si l'effet anticipé émergera lors des analyses des données de la deuxième année du projet.

Sommaire : Analyses par profil linguistique

Les analyses par profil linguistique des enfants confirment que l'impact du programme avait tendance à varier selon l'environnement langagier des enfants au sein de leur famille. L'impact du programme sur certaines résultantes langagières telles que la Communication et le Vocabulaire réceptif était plus fort chez les enfants issus de familles à faible exposition. L'effet modérateur du Type de foyer était significatif pour ces deux variables vers la fin de l'année (à 8 et à 12 mois) relativement au groupe témoin hors garderie seulement. Par contre, tel que prévu, l'impact du programme sur la Cognition était plus fort chez les enfants issus de famille à forte exposition. Encore une fois, l'effet modérateur du Type de foyer était significatif vers la fin de l'année (à 8 et à 12 mois), mais cette fois relativement aux deux groupes témoins. En somme, tous les enfants bénéficient du programme, mais le type de bénéfice semble dépendre du profil linguistique des enfants.

9.3. LORS DE LA DEUXIÈME ANNÉE

Cette section présente les résultats des analyses d'impact de la deuxième année. Les estimés de l'effet traitement reposent sur la trajectoire développementale relative des trois groupes à l'étude au cours des quatre évaluations de la deuxième année. L'évaluation de référence demeure la première évaluation (c. à d. la période pré-intervention) et les évaluations post-intervention sont celles effectuées à 12, 16, 20 et 24 mois.

Au total, quatre mesures de résultantes provenant de la version française de l'ÉPE-AD ont fait l'objet d'analyses : la Communication, la Conscience de soi, la Cognition et le Vocabulaire expressif. Comme pour les analyses portant sur la première année du programme (voir la section 9.2), l'effet traitement relatif aux groupes témoins est mesuré pour ces résultantes par l'estimateur DD, qui capte les différences entre les trajectoires développementales des trois groupes de traitement. Deux autres échelles de vocabulaire sont ajoutées à la batterie: l'EOWPVT et l'ÉVIP-R. Ces échelles sont analysées en utilisant un estimateur « ANCOVA » (c.-à-d. un estimateur découlant de l'analyse de covariance), prenant le score de Communication au niveau de base comme covariable pour annuler statistiquement les différences entre les enfants qui peuvent exister à la période pré-intervention.⁷⁷ En tout, six résultantes sont examinées pour évaluer l'effet du programme au cours de la deuxième année de sa prestation.

9.3.1. Détails techniques

Comme pour les analyses de la première année, la stratégie générale adoptée ici est de présenter deux spécifications alternatives pour chaque analyse. La première, le modèle initial,

⁷⁷ L'estimateur ANCOVA est moins valide que l'estimateur DD quand deux groupes non-équivalents sont comparés (pour une discussion, voir Jamieson, 2003).

considère uniquement les prédicteurs importants servant à expliquer l'effet du programme. La spécification du modèle initial peut varier dépendamment si l'estimateur DD ou l'estimateur ANCOVA est utilisé. La deuxième spécification est une complexification de la première incluant les covariables en plus des prédicteurs importants. Les deux spécifications produisent des résultats qui peuvent être comparés afin de diagnostiquer des problèmes techniques ou de nuancer l'interprétation des résultats. Pour cet exercice de comparaison, le modèle initial sert principalement à enrichir l'interprétation des résultats définitifs provenant du modèle ajusté. Le lecteur est prié de noter qu'il est spécifié dans le texte lorsque les résultats du modèle initial sont rapportés en annexe ou qu'ils sont disponibles sur demande. Ceci dit, les résultats du modèle initial ont été considérés dans l'interprétation préliminaire des impacts ajustés dans tous les cas.

Détails de la spécification de modèles

Modèle initial

Le modèle initial de l'analyse DD par groupe de traitement comprend une série de variables fictives représentant la période d'évaluation, l'appartenance au groupe et un terme représentant l'interaction entre ces indicateurs de variables fictives (c.-à-d. les estimateurs DD). Le facteur temps ou « période d'évaluation » est représenté par quatre variables fictives comparant les évaluations de suivi à la première évaluation (ou évaluation de base). En d'autres mots, la période d'évaluation de base (ou pré-intervention) est placée en « référence ». Le facteur « groupe » est représenté par deux variables fictives servant à comparer les groupes témoins au groupe programme. Le groupe programme est donc placé en « référence ». Le modèle initial DD est modifié légèrement à chaque section selon la question de recherche examinée. Les changements apportés au modèle initial sont notés dans le texte. Notez que les variables de dosage, de fidélité et de qualité sont interprétées de la façon habituelle (c.-à-d. des effets positifs indiquent une augmentation dans la résultante à l'étude en fonction d'une augmentation dans la variable associée).

Le modèle initial de l'analyse de covariance (ANCOVA) par groupe de traitement comprend une série de variables fictives représentant l'appartenance au groupe de traitement ainsi que le score sur la dimension Communication provenant de la première évaluation. Le groupe programme demeure la catégorie de référence pour ces analyses, donc une valeur négative indique un effet positif du programme testé, tout comme pour le modèle DD. Le modèle initial ANCOVA est modifié légèrement à chaque section selon la question de recherche examinée. Les changements apportés au modèle initial sont notés dans le texte pour chaque analyse. Notez que les variables de dosage, de fidélité et de qualité sont interprétées de la façon habituelle (c.-à-d. des effets positifs indiquent une augmentation dans la résultante en fonction d'une augmentation dans la variable associée). Il n'était pas nécessaire de représenter la dimension « temps » dans le modèle car il n'y a qu'une seule observation par enfant pour chaque résultante.

Modèle ajusté

Le modèle initial (DD ou ANCOVA) est imbriqué dans le modèle ajusté. Le modèle ajusté comprend toutes les covariables identifiées à la section 7.3 du chapitre 7 ainsi que leur interaction avec le facteur temps.

Échantillons qui font l'objet des analyses

Échantillons de base

L'échantillon faisant l'objet des analyses varie selon la résultante à l'étude. L'échantillon total de participants sert aux analyses de l'échelle de Communication, l'EOWPVT et l'ÉVIP-R. Pour les autres échelles, l'échantillon faisant l'objet des analyses est celui qui exclut les enfants ayant des données manquantes dues à l'administration de l'ÉPE-AD en anglais et pour qui les valeurs manquantes à la version française de l'ÉPE-AD ne pouvaient pas être imputées. Enfin, les enfants qui se sont retirés de l'étude avant l'évaluation de 24 mois sont exclus des analyses pour maximiser la validité interne de l'estimateur DD. Les autres précautions mises en place pour la préparation de données sont les mêmes que celles des analyses de la première année.

9.3.2. Analyses par groupe

Cette section présente les analyses par groupe de traitement. Le lecteur intéressé trouvera des tableaux sommaires pour l'analyse du modèle initial et l'analyse du modèle ajusté portant sur chacune des résultantes à l'annexe H. Dans la présente section, uniquement les résultats de l'analyse par groupe provenant du modèle ajusté sont rapportés dans des figures. Nous rappelons au lecteur que le groupe programme, le groupe témoin en garderie et le groupe témoin hors garderie sont identifiés respectivement par les étiquettes G1, G2 et G3. *Contrairement aux analyses rapportées dans les tableaux, un effet positif du programme est indiqué dans les figures par une valeur positive.*

Dans ce qui suit, une figure est présentée pour chaque résultante avec des estimés découlant du modèle ajusté. L'effet du programme est présenté sous forme d'intervalles de confiance à 90 %. Les intervalles de confiance représentent une étendue de scores comprenant probablement la valeur réelle de l'impact du programme (c.-à-d. neuf fois sur dix dans le présent cas). *Si l'intervalle ne comprend pas la valeur 0, l'effet est considéré significatif à $p = 10\%$.*⁷⁸ Comme pour les tableaux, nous rapportons des effets pour chaque croisement des variables temps et groupe témoin. Pour faciliter l'interprétation longitudinale des résultats, nous présentons les estimations DD de la deuxième année avec celles des analyses de la première année dans les figures.⁷⁹

Échelle Communication

Modèle initial

Les résultats du modèle initial sont rapportés à l'annexe H (tableau H.1).

Modèle ajusté

Les résultats du modèle ajusté sont présentés à la figure 9.2 pour les deux années de prestation du programme. Notons dans un premier temps que l'effet programme n'infirmé jamais notre hypothèse de base selon laquelle le programme a un effet positif sur le développement : les intervalles de confiance incluent tous des valeurs positives.

⁷⁸ Les résultats complets présentés à l'annexe C distinguent entre les seuils de signification de $p = 10\%$, 5% et 1% .

⁷⁹ Le lecteur est prié de noter que les résultats provenant d'évaluations de la première année sont identiques aux résultats présentés à la section 9.2 lors de l'analyse de la première année.

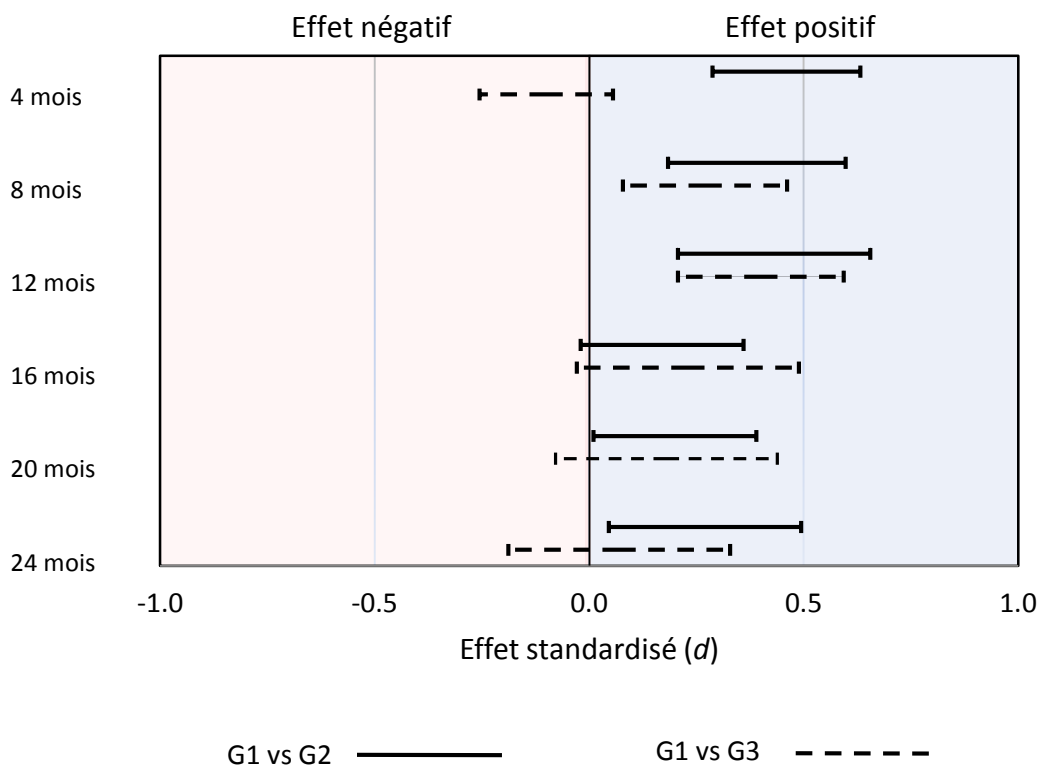
L'effet programme relatif au groupe témoin hors garderie semble s'atténuer avec le temps. Les intervalles de confiance pour cet effet chevauchent la valeur zéro à 16, 20 et 24 mois indiquant l'absence d'effets significatifs. Pour sa part, l'effet programme relatif au groupe témoin en garderie est non-significatif à 16 mois seulement.

Il serait utile toutefois de nuancer l'interprétation de cette absence de signification à certaines périodes d'évaluation. Chaque effet DD pourrait être considéré une reproduction du test de l'effet du programme. Selon cette perspective, chaque estimation DD tente d'estimer la même valeur réelle de l'effet du programme.⁸⁰ Si nous acceptons ce postulat, il importe d'interpréter non seulement la signification de chaque intervalle de confiance, mais aussi leur chevauchement. On pourrait donc poser la question : Quelles sont les valeurs que pourraient prendre l'effet du programme lorsqu'il y a chevauchement des intervalles de confiance observés à chaque période d'évaluation?

Dans les faits, il y a une région de chevauchement des intervalles de confiance s'étendant de 0,26 à 0,36 écart-type pour les effets DD représentant la comparaison du groupe programme et le groupe témoin en garderie. En d'autres mots, la taille de l'effet varie d'une évaluation à l'autre (et est parfois non-significative, comme pour la période d'évaluation à 16 mois) mais les estimations n'infirment jamais l'hypothèse d'un effet positif du programme de l'ordre d'environ 0,30 écart-type. Excluant l'effet à quatre mois, le chevauchement des intervalles de confiance pour les effets DD représentant la comparaison du groupe programme et du groupe témoin hors garderie suggère un effet d'à peu près la même envergure. La possibilité d'interpréter les résultats de cette façon est un des avantages principaux de travailler avec des intervalles de confiance au lieu d'un simple test de signification.

⁸⁰ Nous reconnaissons que l'effet réel du programme peut varier dans le temps. Nous adoptons la perspective d'une reproduction simple afin d'illustrer la logique de l'interprétation globale des intervalles de confiance. Dans l'absence d'une tendance claire d'une évaluation à l'autre (une accentuation ou une atténuation de l'effet du programme systématique au fil du temps), cette perspective est raisonnable.

Figure 9.2 : Effet du programme (DD) sur la Communication à chaque évaluation



Les barres représentent les intervalles de confiance de 90%. Les intervalles de confiance ne comprenant pas la valeur de zéro indiquent un effet significatif à 10%.

Contre-validation non-paramétrique

Pour l'analyse non-paramétrique, la performance des enfants observée avec l'échelle Communication est exprimée en termes de la probabilité d'être classifiée dans un quintile supérieur : la cote de réussite. L'analyse non-paramétrique confirme partiellement la signification des effets DD pour la comparaison du groupe programme au groupe témoin en garderie. Le rapport des cotes est de 0,76 (IC 90 % de 0,46 à 1,26), 0,60 (IC 90 % de 0,38 à 0,96) et 0,55 (IC 90 % de 0,28 à 1,08) en faveur du groupe programme respectivement pour les effets DD à 16, 20 et 24 mois. Ainsi, l'effet à 20 mois rapporté à la figure 9.2 est le seul à être reproduit. Ajoutons toutefois que l'effet à 24 mois approche le seuil de signification de 10 %. Enfin, comme pour l'analyse paramétrique rapportée à la figure 9.2, aucun des effets DD à 16, 20 et 24 mois pour la comparaison au groupe témoin hors garderie n'est significatif.

Échelle Conscience de soi

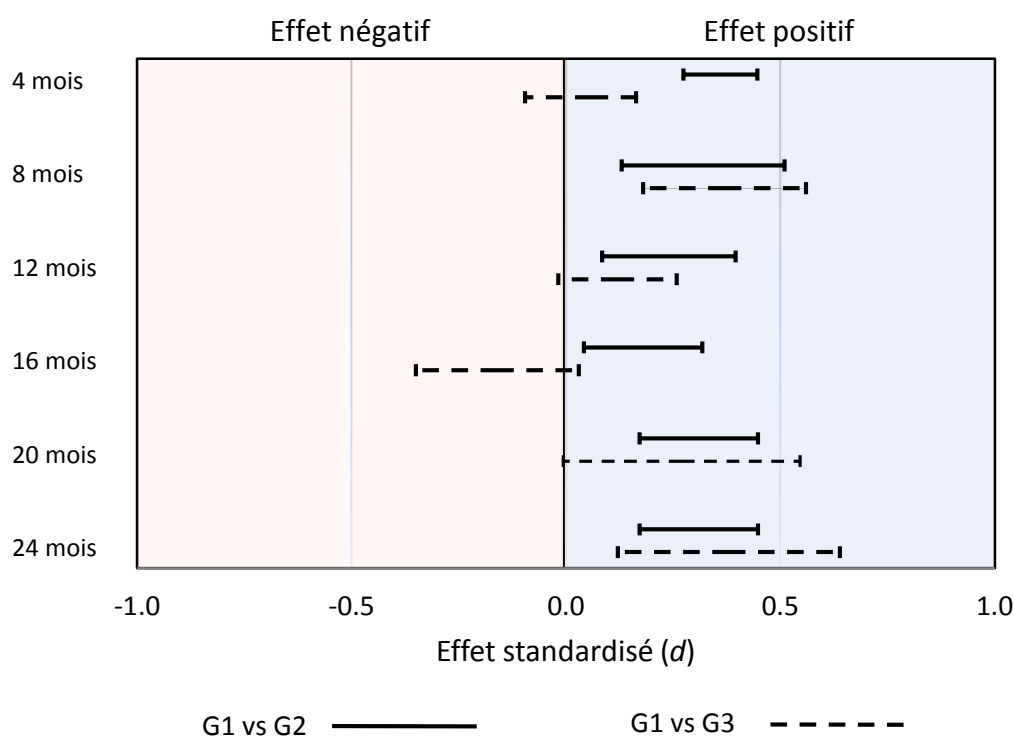
Modèle initial

Les résultats du modèle initial sont rapportés à l'annexe H (tableau H.1).

Modèle ajusté

Les résultats du modèle ajusté sont présentés à la figure 9.3 pour les deux années de prestation du programme. Notons dans un premier temps que les intervalles de confiance n'infirmement jamais notre hypothèse de base voulant que le programme ait un effet positif sur le développement. L'effet programme relatif au groupe témoin hors garderie est le moins précis (intervalles de confiance plus grands) en plus d'être instable d'une évaluation à l'autre en termes de significativité et direction (positif versus négatif). Il est toutefois significatif à 24 mois ($d = 0,38$). La comparaison des groupes en garderie révèle un effet stable et significatif en faveur du programme pour les deux ans de prestation du programme. La région de chevauchement des ICs s'étend de 0,27 à 0,32 écart-type, indiquant que les intervalles de confiances rapportées n'infirmement jamais l'hypothèse que l'effet réel du programme se situe entre ces valeurs (voir l'analyse de la Communication pour des explications supplémentaires).

Figure 9.3 : Effet du programme (DD) sur la Conscience de soi à chaque évaluation



Les barres représentent les intervalles de confiance de 90%. Les intervalles de confiance ne contenant pas la valeur de zéro indiquent un effet significatif à 10%.

Contre-validation non-paramétrique

Pour l'analyse non-paramétrique, la performance des enfants observée avec l'échelle Conscience de soi est exprimée en termes de la probabilité d'être classifiée dans un quintile supérieur : la cote de réussite. L'analyse non-paramétrique reproduit la signification des effets DD observés pour les comparaisons des groupes en garderie rapportées à la figure 9.3. Ainsi, le rapport des cotes est de 0,67 (IC 90 % de 0,46 à 0,99), 0,48 (IC 90 % de 0,34 à 0,69), et 0,59 (IC 90 % de 0,41 à 0,82) respectivement pour les effets DD à 16, 20 et 24 mois. Par contre, l'analyse

non-paramétrique ne reproduit pas la signification des effets DD observés pour les comparaisons relatives au groupe témoin hors garderie. Le rapport des cotes est de 0,68 (IC 90 % de 0,46 à 1,01) pour l'évaluation à 24 mois. Bien que cet effet approche le seuil de signification de 10 %, il est néanmoins non significatif.

Échelle Cognition

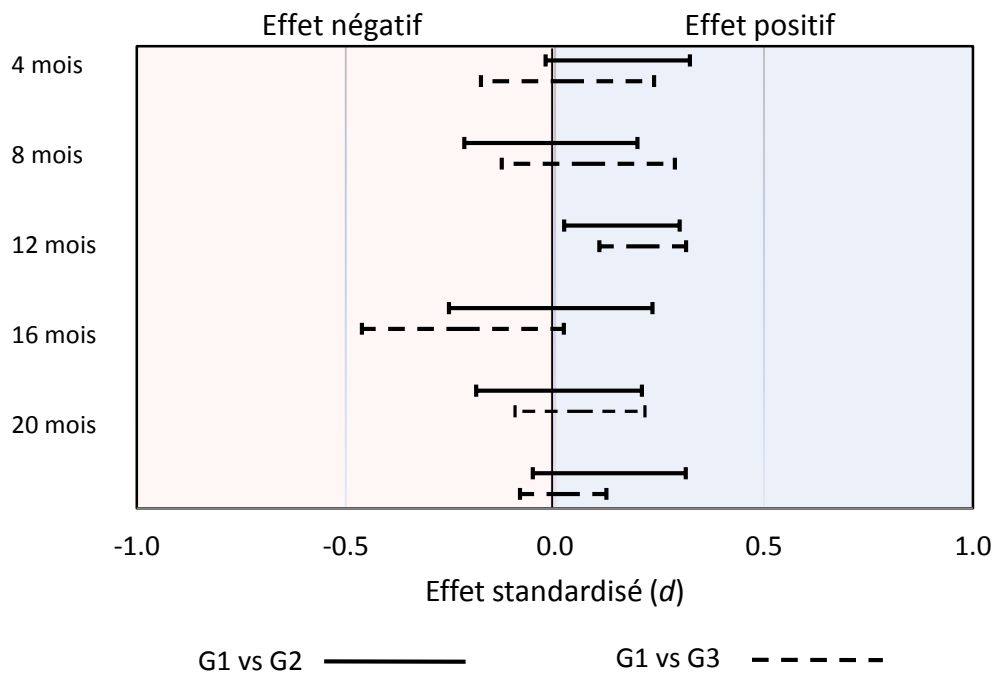
Modèle initial

Les résultats du modèle initial sont rapportés à l'annexe H (tableau H.1).

Modèle ajusté

Les résultats du modèle ajusté sont présentés à la figure 9.4 pour les deux années de prestation du programme. On y observe l'absence de tous effets significatifs en faveur du groupe programme pour les deux groupes témoins.

Figure 9.4 : Effet du programme (DD) sur la Cognition à chaque évaluation



Les barres représentent les intervalles de confiance de 90%. Les intervalles de confiance ne comprenant pas la valeur de zéro indiquent un effet significatif à 10%.

Contre-validation non-paramétrique

Pour l'analyse non-paramétrique, la performance des enfants observée avec l'échelle Cognition est exprimée en termes de la probabilité d'être classifiée dans un quintile supérieur : la cote de réussite. L'analyse non-paramétrique reproduit l'absence de significativité des effets DD observés pour cette résultante dans les comparaisons avec les deux groupes témoins.

Échelle Vocabulaire expressif

Modèle initial

Les résultats du modèle initial sont rapportés à l'annexe H (tableau H.1).

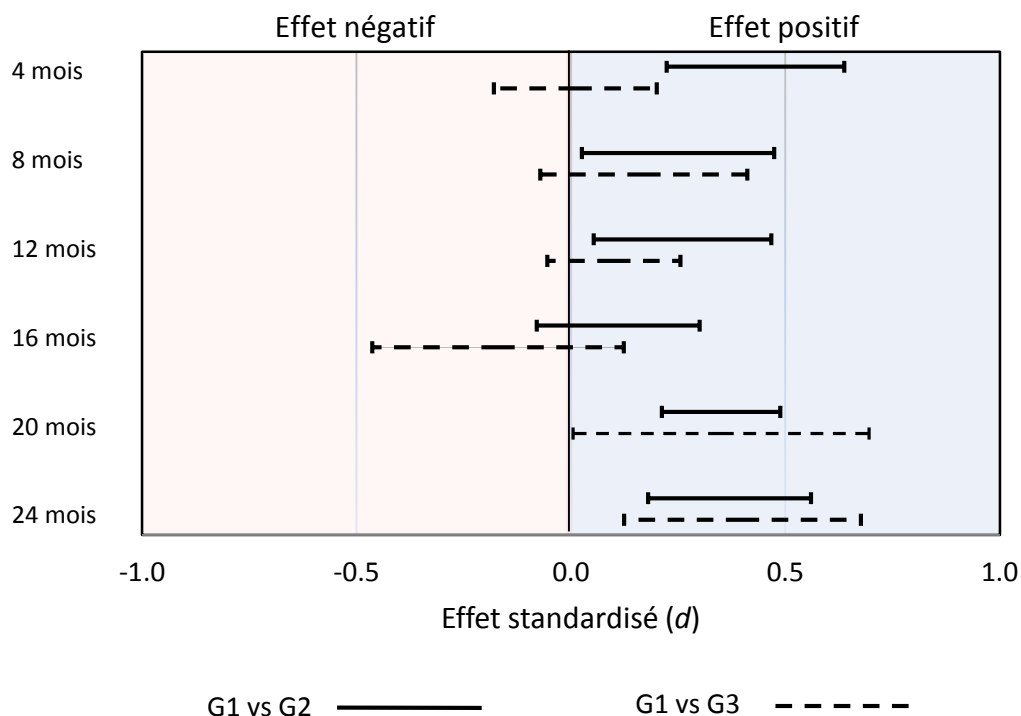
Modèle ajusté

Les résultats du modèle ajusté sont présentés à la figure 9.5 pour les deux années de prestation du programme. Tous les intervalles de confiance incluent des valeurs positives, appuyant l'idée que le programme a un effet positif sur le développement du vocabulaire des enfants.

À l'entrée de la maternelle de cinq ans (période d'évaluation à 24 mois), nous observons un avantage significatif du groupe programme de 0,37 écart-type relatif au groupe témoin en garderie et de 0,41 écart-type relatif au groupe témoin hors garderie. Ces bénéfices excèdent le déficit initial du groupe programme et représentent des gains de 3,3 et 3,7 mois dans le développement du vocabulaire des enfants du groupe programme relatif aux groupes témoins lorsqu'on se réfère aux données de Hill et ses collègues (2007). Cette estimation est en toute probabilité conservatrice si l'on se réfère plutôt aux normes de l'ÉVIP-R (Dunn et al., 1993, tableau 4.4, Forme A) selon lesquelles un gain en vocabulaire de 0,49 écart-type est attendu entre les âges de 3,5 et 4,5 ans. Appliquant la courbe développementale en vocabulaire réceptif capté par l'ÉVIP-R à celle du vocabulaire expressif, l'effet rapporté ici représente un gain d'environ 9,6 mois relatif au groupe témoin en garderie et d'environ 10 mois relatif au groupe témoin hors garderie. Il reste à voir si des résultats comparables sont obtenus avec l'ÉVIP-R dans les analyses à suivre.

Ces résultats confirment un effet positif et stable du programme comparativement au groupe témoin en garderie, et l'effet est persistant jusqu'à l'évaluation de 24 mois. Par ailleurs, la significativité de l'effet du programme relatif au groupe témoin hors garderie est seulement observée lors de la deuxième année, émergeant tardivement.

Figure 9.5 : Effet du programme (DD) sur le Vocabulaire expressif à chaque évaluation



Les barres représentent les intervalles de confiance de 90%. Les intervalles de confiance ne comprenant pas la valeur de zéro indiquent un effet significatif à 10%.

Contre-validation non-paramétrique

Pour l'analyse non-paramétrique, la performance des enfants observée avec l'échelle Vocabulaire expressif est exprimée en termes de la probabilité d'être classifiée dans un quintile supérieur : la cote de réussite. L'analyse non-paramétrique reproduit la signification des effets DD pour la comparaison des groupes en garderie rapportés à la figure 9.5. Le rapport des cotes est de 0,81 (IC 90 % de 0,48 à 1,33), 0,52 (IC 90 % de 0,36 à 0,75), et 0,48 (IC 90 % de 0,26 à 0,87) en faveur du groupe programme respectivement pour les effets DD à 16, 20 et 24 mois. L'effet DD pour la comparaison avec le groupe témoin hors garderie à 24 mois est moins robuste, n'étant pas significatif selon l'analyse non-paramétrique, 0,80 (IC 90 % de 0,48 à 1,31). En effet, cet effet n'approche pas du tout le seuil de signification de 10 %.

Échelles EOWPVT et ÉVIP-R

Les estimations ANCOVA du modèle initial et du modèle ajusté sont rapportées pour l'EOWPVT (vocabulaire expressif) et l'ÉVIP-R (le vocabulaire réceptif) au tableau G.1 (annexe G). Nous rappelons au lecteur que la distinction entre l'estimateur ANCOVA et l'estimateur DD est que dans le cas de l'estimateur ANCOVA, une variable pré-intervention (dans ce cas-ci la Communication) sert de covariable dans le modèle, jouant un rôle conceptuellement semblable à celui de l'évaluation au niveau de base dans le cadre de l'estimateur DD (les deux méthodes ne sont pas équivalentes au niveau des statistiques). Seulement le résultat pour le modèle ANCOVA

ajusté est rapporté graphiquement à la figure 9.6. Nous abordons les résultats pour les deux résultantes parallèlement.

Modèle initial

Les résultats du modèle initial sont rapportés dans un tableau à l'annexe G. Puisque les résultats du modèle ajusté sont interprétés en fonction du modèle initial, le lecteur est prié de noter que les effets « positifs » du programme sur le développement de la préparation à la scolarisation prennent une valeur négative dans les tableaux et une valeur positive dans les figures. La comparaison des deux types de résultats se fait sur la base de la taille absolue de l'effet (c.-à-d. la taille de l'effet en ignorant le signe négatif ou positif puisque dans les deux cas il est question d'un impact positif du programme sur le développement).

Modèle ajusté

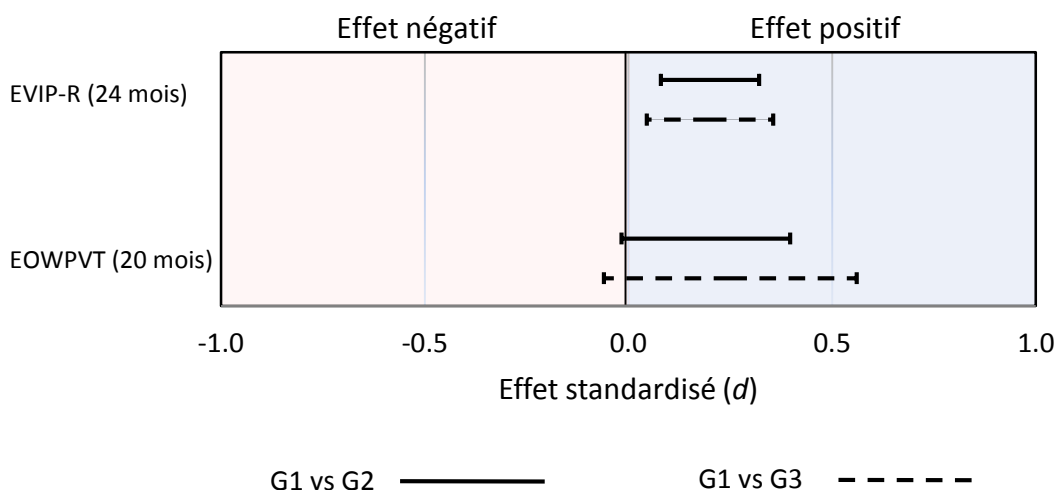
Selon les estimations ANCOVA du modèle ajusté, on observe une différence d'environ 0,20 écart-type entre les groupes en garderie pour l'ÉVIP-R et pour l'EOWPVT. Cette différence est significative pour l'ÉVIP-R et non significative pour l'EOWPVT. Nous notons que la taille de l'effet du programme sur le vocabulaire relative au groupe témoin en garderie est robuste au changement de spécifications et de résultantes (ÉVIP-R versus l'EOWPVT).

Le modèle ajusté révèle une différence significative entre le groupe programme et le groupe témoin hors garderie sur l'ÉVIP-R (voir l'intervalle de confiance rapporté qui ne comprend pas la valeur zéro à la figure 9.6). Un avantage de 0,20 écart-type est observé en faveur du groupe programme pour cette comparaison. Un effet de taille comparable, mais non-significatif, est observé avec l'EOWPVT relativement au groupe témoin hors garderie. En somme, les analyses indiquent que pour le présent échantillon le programme testé a un effet positif sur le développement du vocabulaire réceptif.

Nous utilisons les normes de l'ÉVIP-R pour exprimer les gains observés en termes de « mois » de développement. Selon les normes de l'ÉVIP (Dunn et al., 1993, Tableau 4,4, Forme A), un gain en vocabulaire d'environ 0,38 écart-type est typique entre les âges de quatre et cinq ans. Les effets observés dans le cadre du projet Capacité d'apprentissage représenterait donc environ 6,6 mois de développement en vocabulaire réceptif (ÉVIP-R).⁸¹ La taille des effets non-significatifs rapportés à la figure 9.6 donnent à penser qu'un effet comparable serait observé avec le vocabulaire expressif (EOWPVT), soit un gain d'environ une demi année.

⁸¹ Formule de conversion: [(0,21 d'effet du programme/0,38 de changement typique)*12= 6,6 mois]

Figure 9.6 : Intervalles de confiance de l'impact du programme standardisé (ANCOVA) sur le vocabulaire français : Échelles normalisées de vocabulaire réceptif (ÉVIP-R) et expressif (EOWPVT)



Les barres représentent les intervalles de confiance de 90 %. Les intervalles de confiance ne comprenant pas la valeur de zéro indiquent un effet significatif à 10 %.

Contre-validation non-paramétrique

Pour l'analyse non-paramétrique, la performance des enfants observée avec l'ÉVIP-R et l'EOWPVT est exprimée en termes de la probabilité d'être classifiée dans un quintile supérieur : la cote de réussite. Cette analyse confirme les résultats rapportés à la figure 9.6. L'analyse non-paramétrique reproduit les effets significatifs observés avec l'ÉVIP. On obtient un rapport de cote de 0,64 (IC 90 % de 0,47 à 0,87) en faveur du groupe programme pour la comparaison au groupe témoin en garderie et un rapport de cote de 0,59 (IC 90 % de 0,38 à 0,90) en faveur du groupe programme pour la comparaison au groupe témoin hors garderie. De même, l'analyse non-paramétrique reproduit l'absence de significativité de l'effet du programme sur l'EOWPVT. On observe un rapport des cotes de 0,63 (IC 90 % de 0,34 à 1,18) et de 0,63 (IC 90 % de 0,28 à 1,41) respectivement pour les comparaisons relatives aux groupes témoins en garderie et hors garderie.

Sommaire : Analyses par groupe

Les résultats indiquent que le programme testé semble avoir un effet positif sur la préparation à la scolarisation des enfants. Pour la Communication, la Conscience de soi, le Vocabulaire expressif et l'ÉVIP-R, un effet positif du programme testé est observé lors de l'entrée à la maternelle de cinq ans (24 mois) des enfants participants. Cet effet est observé par rapport aux deux groupes de comparaisons pour chacun de ces résultantes sauf la Communication. Pour la Communication, l'effet est observé uniquement relativement au groupe témoin en garderie. Aucun effet stable du programme n'est observé d'une évaluation à l'autre pour la Cognition dans le cadre de cette analyse. De même, le programme n'a pas eu un effet significatif sur l'EOWPVT lors de cette analyse. Les effets attendus seront observés dans le cadre d'autres modèles qui sont présentés dans ce qui suit.

9.3.3. Analyses par dosage

Les résultats de l'analyse examinant l'effet du dosage en fonction de l'appartenance au groupe programme ou au groupe témoin en garderie sont rapportés au tableau 9.14 pour le modèle initial et au tableau 9.15 pour le modèle ajusté. Puisque les deux modèles ont produit essentiellement le même patron de résultats, nous interprétons uniquement les résultats du modèle ajusté (voir le tableau 9.15). De plus, nous discutons seulement des effets de dosage pour l'échelle Cognition en raison de l'absence d'effets significatifs du dosage pour les autres échelles. *Le lecteur est prié de noter que, contrairement aux figures présentées à la section précédente, les effets « positifs » du programme testé sur le développement prennent des valeurs négatives pour les estimés DD rapportés dans les tableaux.*

Échelle Cognition

Modèle initial

Tel que indiqué plus haut, aucune interprétation n'est faite des résultats provenant de ce modèle. Ils sont rapportés au tableau 9.14.

Modèle ajusté

Les résultats de l'analyse indiquent que le dosage lors de la première année est peu prédictif du développement durant la deuxième année sur les échelles de l'ÉPE-AD. La seule exception observée est pour l'échelle de Cognition. Les estimations DDD pour cette variable indiquent que l'effet du dosage est différencié selon le groupe de traitement lors de toutes les périodes d'évaluations de la deuxième année, sauf celle à 24 mois. L'effet, quand il est significatif, est d'environ 0,30 à 0,40 écart-type par augmentation/réduction de 10 heures par semaine. Il est intéressant de noter que l'effet du programme est complètement absent dans l'analyse par groupes de traitement. Ceci suggère que pour les évaluations à 16 et 20 mois, contrairement à ce qui est présenté à la figure 9.4, un effet du programme est observé pour l'échelle de la Cognition seulement pour les enfants ayant reçu un très fort dosage (au moins 37 heures par semaine).

Contre-validation non-paramétrique

Pour l'analyse non-paramétrique, la performance des enfants observée avec l'échelle Cognition est exprimée en termes de la probabilité d'être classifiée dans un quintile supérieur : la cote de réussite. L'analyse non-paramétrique reproduit seulement les effets DDD significatifs à 16 et à 20 mois. Les rapports de cotes pour les périodes d'évaluation à 16, 20 et 24 mois sont de 0,94 (IC 90 % de 0,89 à 0,98), 0,92 (IC 90 % de 0,86 à 0,99) et 0,97 (IC 90 % de 0,89 à 1,05) en faveur du groupe programme.

Tableau 9.14 : Nombre d'heures passées en garderie lors de la première année (Dosage) et son effet non-ajusté sur le développement de la préparation à la scolarisation lors de la deuxième année selon les deux groupes en garderie — Scores standardisés

Type de différence	Indicateur de la préparation à la scolarisation							
	Communication		Conscience de soi		Cognition		Vocabulaire expressif	
	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.
Dosage								
Niveau de base	0,029***	0,009	0,015	0,017	-0,023	0,016	0,009	0,015
Dosage x groupe								
G1 vs G2 (DD)	-0,016	0,015	-0,005	0,022	0,024	0,020	-0,001	0,020
Dosage x temps								
12 mois (DD)	0,004	0,021	0,021	0,015	0,040**	0,017	0,016	0,015
16 mois (DD)	0,009	0,013	0,002	0,007	0,040***	0,008	0,018	0,012
20 mois (DD)	-0,012	0,017	0,015	0,009	0,029**	0,012	0,021**	0,007
24 mois (DD)	-0,009	0,011	0,015**	0,007	0,011	0,009	0,022***	0,004
Dosage x temps x groupe								
12 mois (DDD)	0,012	0,022	-0,011	0,019	-0,045**	0,019	-0,004	0,019
16 mois (DDD)	0,000	0,012	0,003	0,012	-0,039**	0,015	-0,008	0,014
20 mois (DDD)	0,005	0,020	-0,013	0,013	-0,034*	0,017	-0,014	0,011
24 mois (DDD)	0,013	0,010	-0,009	0,015	-0,008	0,016	-0,009	0,011
Dose x Temps Wald F	2,02		1,29		4,35**		2,32*	
Dose x temps x groupe Wald F	0,74		0,92		5,09***		0,54	

Note : Les valeurs négatives des estimations DD pour toutes les comparaisons entre les groupes représentent un effet traitement positif (c.-à-d. un avantage pour le groupe programme. Degrés de liberté des tests Wald F sont 4 et 19. Les erreurs types robuste de Huber-White sont rapportées avec regroupement par garderie. La différence entre les groupes est significative à *** $p < 1\%$, ** $p < 5\%$; * $p < 10\%$. Diff = différence, E.T. = erreur type.

Tableau 9.15 : Nombre d'heures passées en garderie lors de la première année (Dosage) et son effet ajusté sur le développement de la préparation à la scolarisation lors de la deuxième année selon les deux groupes en garderie — Scores standardisés

Type de différence	Indicateur de la préparation à la scolarisation							
	Communication		Conscience de soi		Cognition		Vocabulaire expressif	
	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.
Dosage								
Niveau de base	0,007	0,008	0,002	0,012	-0,015	0,011	-0,005	0,009
Dosage x groupe								
G1 vs G2 (DD)	-0,017*	0,008	-0,005	0,014	0,016	0,012	-0,002	0,011
Dosage x temps								
12 mois (DD)	0,006	0,021	0,016	0,015	0,028*	0,014	0,010	0,016
16 mois (DD)	0,010	0,014	-0,004	0,011	0,026***	0,008	0,005	0,010
20 mois (DD)	0,020	0,015	0,009	0,015	0,033***	0,011	0,013	0,011
24 mois (DD)	0,000	0,014	0,010	0,015	0,013	0,014	0,009	0,010
Dosage x temps x groupe								
12 mois (DDD)	0,014	0,020	-0,013	0,019	-0,038**	0,017	-0,002	0,018
16 mois (DDD)	0,005	0,012	0,005	0,014	-0,032**	0,012	0,000	0,011
20 mois (DDD)	-0,012	0,015	-0,015	0,016	-0,038**	0,014	-0,014	0,012
24 mois (DDD)	0,014	0,013	-0,012	0,018	-0,012	0,017	-0,002	0,012
Dose x Temps Wald F	5,37***		1,92		2,00		0,32	
Dose x temps x groupe Wald F	6,25***		1,64		5,77***		1,33	

Note : Les valeurs négatives des estimations DD pour toutes les comparaisons entre les groupes représentent un effet traitement positif (c.-à-d. un avantage pour le groupe programme. Degrés de liberté des tests Wald F sont 4 et 19. Les erreurs types robuste de Huber-White sont rapportées avec groupement des erreurs par garderie. Le seuil de signification est indiqué à *** $p < 1\%$, ** $p < 5\%$; * $p < 10\%$. Diff = différence, E.T. = erreur type.

Échelles EOWPVT et ÉVIP-R

Les estimations ANCOVA du modèle initial et du modèle ajusté sont rapportées pour l'EOWPVT (vocabulaire expressif) et l'ÉVIP-R (le vocabulaire réceptif) au tableau 9.16. Nous rappelons au lecteur que la distinction entre l'estimateur ANCOVA et l'estimateur DD est que dans le cas de l'estimateur ANCOVA, une variable pré-intervention (dans ce cas-ci la Communication) sert de covariable dans le modèle, jouant un rôle conceptuellement semblable à celui de l'évaluation au niveau de base dans le cadre de l'estimateur DD (les deux méthodes ne sont pas équivalentes au niveau des statistiques). Nous abordons les résultats pour les deux échelles parallèlement.

Modèle initial

Les résultats provenant du modèle initial (tableau 9.16) indiquent que le dosage a un effet positif sur les vocabulaires réceptif et expressif, tels que mesurés par l'ÉVIP-R et l'EOWPVT.

L'effet du dosage est significatif, mais l'effet d'interaction entre le dosage et le groupe traitement ne l'est pas. Selon ces résultats, les heures passées en garderie apportent des gains positifs en vocabulaire, peu importe le programme offert en garderie (programme ou témoin).

Modèle ajusté

L'effet du dosage observé avec le modèle initial est annulé dans le modèle ajusté (tableau 9.16). Ce patron de résultats suggère que l'effet observé avec le modèle initial est un artéfact associé aux variables confusionnelles. De plus, l'effet du groupe de traitement est significatif dans le modèle ajusté, et ce, pour les deux variables résultantes. La taille des effets observés est semblable à celle observée dans l'analyse par groupe, soit environ 0,20 écart-type (voir la figure 9.6). L'observation de cette différence dans le modèle ajusté est attribuable à une plus grande puissance statistique suite à l'inclusion du dosage dans la spécification.

En somme, les résultats de l'analyse indiquent que le dosage lors de la première année est peu prédictif du développement en vocabulaire expressif et réceptif durant la deuxième année. Ceci dit, la plus grande puissance statistique provenant de l'inclusion du dosage dans le modèle a tout de même permis de détecter statistiquement la différence de 0,20 écart-type en faveur du groupe programme. Enfin, nous estimons à partir des normes de l'ÉVIP-R que cet effet représente un gain de vocabulaire réceptif et expressif d'environ 6,6 mois.

Tableau 9.16 : Impact du dosage en garderie sur le vocabulaire français : Échelles normalisées de vocabulaire réceptif (ÉVIP-R) et expressif (EOWPVT)

Type de différence	Vocabulaire réceptif (ÉVIP-R)				Vocabulaire expressif (EOWPVT)			
	Non-ajusté (N = 217)		Ajusté (N = 212)		Non-ajusté (N = 215)		Ajusté (N = 209)	
	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.
Groupe (G1 vs G2)								
Score brute	-4,029	2,409	-4,648***	1,129	-2,595	1,538	-2,595**	1,076
Score standardisé	-0,224	0,134	-0,258***	0,063	-0,193	0,114	-0,193**	0,080
Dosage								
Score brute	0,236***	0,057	-0,011	0,096	0,231**	0,090	0,072	0,087
Score standardisé	0,013***	0,003	-0,001	0,005	0,017**	0,007	0,005	0,006
Dosage x groupe								
Score brute	-0,138	0,120	-0,112	0,132	-0,076	0,130	-0,068	0,091
Score standardisé	-0,008	0,007	-0,006	0,007	-0,006	0,010	-0,005	0,007

*Note : Les valeurs négatives des estimations DD pour toutes les comparaisons entre les groupes représentent un effet traitement positif (c.-à-d. un avantage pour le groupe programme). Les score standardisés sont des données brutes transformés en score Z ayant une moyenne de zéro et un écart-type de 1. L'affectation aux groupes à l'étude lors de l'administration des échelles est utilisée pour les comparaisons rapportées dans ce tableau. L'ÉVIP-R a été administré à 20 mois et l'EOWPVT à 24 mois. Les erreurs types robuste de Huber-White sont rapportées avec groupement par garderie. Le seuil de signification est indiqué à *** $p < 1\%$, ** $p < 5\%$; * $p < 10\%$. Diff = différence, E.T. = erreur type.*

Contre-validation non-paramétrique

Pour l'analyse non-paramétrique, la performance des enfants observée avec l'ÉVIP-R et l'EOWPVT est exprimée en termes de la probabilité d'être classifiée dans un quintile supérieur :

la cote de réussite. Comme pour les résultats paramétriques présentés au tableau 9.16, aucun effet différencié du dosage (DDD) selon le groupe traitement n'est détecté de cette analyse.

Sommaire : Analyses par dosage

Selon les analyses rapportées, le dosage lors de la première année de l'étude n'est pas très utile pour prédire l'impact du programme lors de la deuxième année de prestation du programme. Exceptionnellement, l'impact du programme sur le développement cognitif semble dépendre du dosage, du moins pour les évaluations à 12, 16 et 20 mois (voir le tableau 9.15). Il semble donc que les impacts à long terme du programme sont plus sensibles au dosage en ce qui concerne le développement cognitif comparativement au développement d'autres compétences. Enfin, le contrôle statistique du dosage révèle un impact positif du programme testé sur le développement du vocabulaire expressif (EOWPVT-R) et réceptif (PPVT-R). Ce dernier résultat démontre l'utilité du dosage en tant que covariable servant à rehausser la précision des tests statistiques.

9.3.4. Analyses par qualité et fidélité du programme en garderie lors de la première année

Cette section présente les analyses examinant l'effet de la qualité et de la fidélité du programme en garderie. Les résultats de deux séries d'analyses sont présentés dans les sections suivantes : (a) une vérification de la valeur prédictive lors de la deuxième année des dimensions clés du programme testé sur le plan de la fidélité et la qualité de la mise en œuvre; et (b) une analyse de médiation vérifiant si la mise en œuvre du programme en garderie est responsable, en tout ou en partie, des effets du programme rapportés à la section 9.3.2. Ces analyses servent à estimer l'importance des éléments du programme en garderie dans l'explication des effets observés du programme testé.

Impact de la fidélité et de la qualité

Les résultats de l'analyse par qualité et par fidélité du programme en garderie sont présentés respectivement aux tableaux 9.17 et 9.18 pour les résultantes provenant de l'ÉPE-AD et pour les échelles de vocabulaire au tableau 9.19. Nous décrivons les résultats pour chaque indicateur de préparation à la scolarisation individuellement.

Qualité structurelle

Au niveau de base, la relation est significativement négative entre la qualité structurelle et deux des résultantes, soit la Communication et la Cognition (voir le tableau 9.17). À l'entrée à la maternelle de cinq ans (période d'évaluation à 24 mois), nous observons des gains supérieurs (voir les effets DD) pour les garderies caractérisées par une meilleure qualité structurelle, et ce, seulement pour la Cognition. La taille de l'effet est de 0,13 écart-type pour chaque 'point' de qualité dans la deuxième année de prestation du programme, ce qui constitue une faible atténuation de l'effet observé à 12 mois. Il est à noter que l'effet de la qualité structurelle est généralement positif et significatif pour certains tests DD à d'autres périodes de temps et pour d'autres résultantes. Le modèle ajusté rapporté au tableau 9.19 indique que la qualité structurelle prédit aussi le développement du vocabulaire expressif tel que mesuré par l'EOWPVT.

Qualité éducative

Nous observons une relation négative au niveau de base entre la qualité éducative et toutes les résultantes qui figurent dans le tableau 9.17. Selon ce résultat, les enfants inscrits à une garderie de haute qualité ont tendance à être défavorisés sur ces résultantes. Toutefois, la qualité éducative reçue durant la première année en garderie a eu un effet positif sur le développement des enfants lors de la deuxième année du projet. Presque tous les effets DD sont significatifs et positifs, et ce, pour toutes les périodes d'évaluation. De façon approximative, le modèle prédit qu'une différence d'un point en qualité éducative entraînerait un gain de 0,09 à 0,17 écart-type en Communication lorsque les enfants commencent la maternelle de cinq ans (période d'évaluation à 24 mois).

Selon les estimés DD du modèle ajusté rapportés au tableau 9.19, la qualité éducative prédit aussi le développement du vocabulaire expressif tel que mesuré par l'EOWPVT. Pour cette résultante, la taille de l'effet est d'environ 0,10 écart-type par point de qualité.

Sensibilité de l'éducatrice

L'examen du tableau 9.17 souligne l'absence d'une association significative entre la sensibilité des éducatrices et chacune des résultantes au niveau de base. Les gains développementaux observés par la suite ne sont pas systématiquement associés à ce facteur, sauf pour ce qui est de quelques effets DD positifs (à 16 mois pour la Conscience de soi; à 12 et à 16 mois pour le vocabulaire expressif). Enfin, lors de l'entrée à la maternelle de cinq ans, aucun effet DD de la sensibilité des éducatrices n'est décelé.

Le modèle ajusté rapporté au tableau 9.19 indique que la sensibilité de l'éducatrice prédit aussi le développement du vocabulaire expressif tel que mesuré par l'EOWPVT. Pour cette résultante, la taille de l'effet est d'environ 0,05 écart-type par point de qualité.

Qualité de lecture

Au niveau de base, nous observons une relation négative entre la qualité de lecture et le vocabulaire expressif (voir le tableau 9.17). À la rentrée à la maternelle de cinq ans (période d'évaluation à 24 mois), nous observons des gains supérieurs relativement au niveau de base pour la Conscience de soi et le Vocabulaire expressif. Selon le modèle présenté, une différence de trois points sur l'échelle de qualité de lecture équivaut à un effet d'environ 0,20 écart-type sur la Conscience de soi et le Vocabulaire expressif.

Le modèle ajusté rapporté au tableau 9.19 indique que la qualité de lecture prédit aussi le développement du vocabulaire expressif tel que mesuré par l'EOWPVT et le vocabulaire réceptif tel que mesuré par l'ÉVIP-R. Pour ces résultantes, la taille de l'effet est d'environ 0,03 écart-type par point de qualité. C'est la seule dimension de qualité significativement associée à ces deux mesures de vocabulaire.

Qualité globale

La Qualité globale (c.-à-d. la moyenne des quatre indices de qualité ci-dessus) est significativement associée à une préparation à la scolarisation déficitaire (voir le tableau 9.17). Les estimations DD indiquent qu'à l'entrée à la maternelle cinq ans (période d'évaluation à 24 mois), la Qualité globale du programme est positivement associée aux gains développementaux

des enfants en garderie pour la Conscience de soi, la Cognition et le Vocabulaire expressif. L'effet significatif le plus stable dans le temps est celui observé pour le Vocabulaire expressif qui se retrouve à toutes les périodes d'évaluation. Selon le modèle présenté, une différence de trois points de qualité (à peu près l'étendue observée) correspondrait à un effet de 0,60 écart-type.

Le modèle ajusté rapporté au tableau 9.19 indique que la qualité globale prédit significativement le développement du vocabulaire expressif tel que mesuré par l'EOWPVT. Pour cette résultante, la taille de l'effet est d'environ 0,13 écart-type par point de qualité. Il n'est pas surprenant que cet index global capte à lui seul l'effet observé compte tenu de tous les éléments qui le composent.

Fidélité structurelle

Au niveau de base, nous observons au tableau 9.18 une association négative entre la fidélité structurelle et la Communication, la Conscience de soi et le Vocabulaire expressif. Il est à noter que cette relation est annulée ou renversée à 20 mois. Par exemple, pour la Communication, le coefficient DD positif de 0,537 est plus important que le déficit de -0,492 observé au niveau de base, $0,537 - (-0,49) = +0,047$. De plus, les coefficients DD pour la période de l'entrée des enfants à la maternelle de cinq ans (période d'évaluation à 24 mois) sont significativement positifs pour la Communication, la Conscience de soi et le Vocabulaire expressif. L'effet le plus important est observé sur le Vocabulaire expressif, pour lequel nous estimons qu'une différence de 50% en fidélité correspond à un gain d'environ 0,50 écart-type.

Le modèle ajusté rapporté au tableau 9.19 indique que la fidélité structurelle est prédictive du développement en vocabulaire tel que mesuré par l'ÉVIP-R et l'EOWPVT. Dans le deux cas, une différence de 40 points de pourcentage de fidélité correspondrait à un effet d'environ 0,15 écart-type.

Fidélité de contenu

Au niveau de base, la fidélité du contenu est négativement associée à la préparation à la scolarisation (voir le tableau 9.18). Cette relation s'expliquerait par une tendance des enfants moins préparés de s'inscrire aux garderies avec une plus grande fidélité de contenu. L'hypothèse par laquelle le contenu du programme serait le plus important déterminant des gains développementaux se voit infirmer par l'absence d'estimées DD significatifs pour cet indicateur. Le modèle ajusté rapporté au tableau 9.18 indique aussi l'absence d'effets positifs associés à cet aspect.

Fidélité globale

Le patron de résultats observé au tableau 9.18 sur la fidélité globale ressemble à celui de la fidélité structurelle. Par conséquent, une description détaillée des résultats serait redondante. Le modèle ajusté rapporté au tableau 9.19 indique une relation positive uniquement pour l'EOWPVT.

Contre-validation non-paramétrique

Les résultats de l'effet de la qualité et de la fidélité du programme sur les résultantes des enfants découlant des analyses DD et de l'ANCOVA (voir tableaux 9.17 et 9.18) sont reproduits, à quelques exceptions près, dans les analyses non-paramétriques.

Sommaire : Analyses par qualité et fidélité du programme en garderie lors de la première année

Cette section contient des résultats indiquant que certains éléments du programme, notamment sa fidélité structurelle, sa qualité structurelle et sa qualité éducative, expliquent les résultantes développementales des enfants dès l'évaluation de base (période d'évaluation pré-intervention), surtout pour la Cognition et le Vocabulaire expressif. Ayant établi que ceci est le cas, nous vérifions à la section suivante si ces mêmes indicateurs permettent d'expliquer l'effet du programme testé (c.-à-d. l'effet du groupe de traitement) rapporté à la section 9.3.2. Cette analyse est un moyen de déterminer dans quelle mesure nous pouvons attribuer l'effet du programme au volet en garderie seul. Il est intéressant de noter que la qualité de lecture, l'indice qui témoigne de la plus importante différence entre le programme en garderie testé et celui offert dans les autres garderies est significativement associé aux gains en Conscience de soi et en Vocabulaire expressif à l'entrée à la maternelle de cinq ans, ainsi qu'aux échelles de vocabulaire ÉVIP-R et EOWPVT.

Tableau 9.17 : La qualité du programme en garderie lors de la première année et son effet ajusté sur le développement de la préparation à la scolarisation lors de la deuxième année — Scores standardisés

Type de différence	Indicateur de la préparation à la scolarisation							
	Communication		Conscience de soi		Cognition		Vocabulaire expressif	
	Diff.	ET	Diff.	ET	Diff.	ET	Diff.	ET
Qualité structurelle								
Niveau de base	-0,17*	0,09	-0,09	0,08	-0,16**	0,07	-0,13	0,10
12 mois (DD)	0,21*	0,11	0,17**	0,06	0,18**	0,07	0,26**	0,11
16 mois (DD)	0,11	0,08	-0,01	0,03	0,20*	0,10	0,08*	0,04
20 mois (DD)	0,24**	0,11	0,16**	0,08	0,13**	0,06	0,18*	0,10
24 mois (DD)	0,10	0,07	0,14	0,08	0,12*	0,07	0,16	0,09
<i>Effet par Temps Wald F</i>	1,13		4,56**		1,56		1,42	
Qualité éducative								
Niveau de base	-0,13***	0,03	-0,11**	0,04	-0,18***	0,03	-0,15***	0,04
12 mois (DD)	0,13**	0,05	0,09*	0,05	0,11***	0,03	0,17**	0,06
16 mois (DD)	0,09**	0,03	0,04	0,04	0,14**	0,05	0,09**	0,04
20 mois (DD)	0,14***	0,04	0,14***	0,04	0,16***	0,04	0,19***	0,05
24 mois (DD)	0,08	0,05	0,08**	0,03	0,13***	0,03	0,17***	0,03
<i>Effet par Temps Wald F</i>	2,47*		3,49**		6,49***		4,68***	
Sensibilité des éducatrices								
Niveau de base	-0,03	0,02	-0,01	0,03	-0,04	0,03	-0,04	0,03
12 mois (DD)	0,03	0,04	0,01	0,02	-0,03	0,03	0,07*	0,03
16 mois (DD)	0,04	0,03	0,04*	0,02	0,01	0,03	0,07**	0,02
20 mois (DD)	0,04	0,03	0,03	0,03	0,03	0,02	0,06	0,04
24 mois (DD)	0,04	0,04	-0,00	0,03	0,01	0,03	0,04	0,03
<i>Effet par Temps Wald F</i>	0,49		2,08		7,47***		2,06	
Qualité de lecture								
Niveau de base	-0,02	0,01	-0,02	0,01	-0,01	0,02	-0,04**	0,01
12 mois (DD)	0,01	0,01	0,01	0,02	0,00	0,02	0,01	0,02
16 mois (DD)	-0,01	0,01	0,01	0,02	-0,04	0,02	0,01	0,02
20 mois (DD)	0,01	0,01	0,05***	0,01	0,00	0,02	0,06***	0,01
24 mois (DD)	0,02	0,02	0,03**	0,01	0,01	0,02	0,06**	0,02
<i>Effet par Temps Wald F</i>	0,36		5,68***		3,80**		3,98**	

Type de différence	Indicateur de la préparation à la scolarisation							
	Communication		Conscience de soi		Cognition		Vocabulaire expressif	
Qualité globale	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.
Niveau de base	-0,13***	0,04	-0,10*	0,05	-0,13**	0,05	-0,16***	0,05
12 mois (DD)	0,13*	0,06	0,08	0,05	0,03	0,04	0,16**	0,06
16 mois (DD)	0,07	0,05	0,06	0,04	0,02	0,06	0,11**	0,04
20 mois (DD)	0,12*	0,06	0,17***	0,05	0,10**	0,03	0,22***	0,07
24 mois (DD)	0,10	0,07	0,10*	0,05	0,09*	0,04	0,20***	0,06
<i>Effet par Temps Wald F</i>	1,15		5,31***		2,64*		3,42**	

Note : Contrairement aux analyses par groupe de traitement, les valeurs positives pour les effets DD indiquent une relation positive entre l'indice de fidélité et les indicateurs de la préparation à la scolarisation. Les degrés de liberté des tests Wald F sont 4 et 18. La qualité globale est la moyenne des quatre indicateurs de qualité. Les erreurs types robuste de Huber-White sont rapportées avec regroupement par garderie. Le seuil de signification est indiqué à *** p < 1 %, ** p < 5 %, * p < 10 %. Diff. = différence, E.T. = erreur type.

Tableau 9.18 : La fidélité du programme en garderie lors de la première année et son effet ajusté sur le développement de la préparation à la scolarisation lors de la deuxième année — Scores standardisés

Type de différence	Indicateur de la préparation à la scolarisation							
	Communication		Conscience de soi		Cognition		Vocabulaire expressif	
Fidélité structurelle	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.
Niveau de base	-0,49***	0,15	-0,33**	0,15	-0,34	0,24	-0,65***	0,15
12 mois (DD)	0,64***	0,21	0,14	0,14	0,22	0,17	0,46*	0,24
16 mois (DD)	0,29	0,22	0,15	0,17	0,00	0,30	0,25	0,23
20 mois (DD)	0,35*	0,17	0,54***	0,14	0,01	0,22	0,88***	0,18
24 mois (DD)	0,53**	0,21	0,49***	0,14	0,34	0,21	0,93***	0,17
<i>Effet par Temps Wald F</i>	2,37*		4,61***		3,37**		8,82***	
Fidélité du contenu	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.
Niveau de base	-0,65***	0,21	-0,51*	0,27	-0,64	0,37	-0,61*	0,33
12 mois (DD)	0,33	0,22	0,10	0,35	0,22	0,34	-0,09	0,39
16 mois (DD)	0,07	0,31	-0,12	0,37	-0,27	0,47	-0,23	0,46
20 mois (DD)	0,12	0,31	0,46	0,45	0,25	0,47	0,58	0,44
24 mois (DD)	0,36	0,47	0,28	0,38	0,20	0,41	0,75	0,48
<i>Effet par Temps Wald F</i>	0,53		1,21		0,60		1,81	
Fidélité globale	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.
Niveau de base	-0,66***	0,18	-0,47**	0,20	-0,52	0,32	-0,79***	0,21
12 mois (DD)	0,69**	0,26	0,17	0,21	0,27	0,24	0,37	0,31
16 mois (DD)	0,29	0,30	0,08	0,25	-0,08	0,41	0,14	0,35
20 mois (DD)	0,36	0,26	0,64**	0,23	0,10	0,31	0,98***	0,27
24 mois (DD)	0,60*	0,34	0,53**	0,20	0,38	0,27	1,08***	0,27
<i>Effet par Temps Wald F</i>	1,71		2,20		2,40*		5,17***	

Note : Contrairement aux analyses par groupe de traitement, les valeurs positives pour les effets DD indiquent une relation positive entre l'indice de fidélité et les indicateurs de la préparation à la scolarisation. Les degrés de liberté des tests Wald F sont 4 et 18. La qualité globale est la moyenne des quatre indicateurs de qualité. Les erreurs types robuste de Huber-White sont rapportées avec regroupement des erreurs par garderie. Le seuil de signification est indiqué à *** $p < 1\%$, ** $p < 5\%$; * $p < 10\%$. Diff = différence, E.T. = erreur type.

Tableau 9.19 : Impact de la qualité et la fidélité de programme sur le vocabulaire français : Échelles normalisées de vocabulaire réceptif (ÉVIP-R) et expressif (EOWPVT)

Type de différence	Qualité de programme							
	Vocabulaire réceptif (ÉVIP-R)				Vocabulaire expressif (EOWPVT)			
	Non-ajusté (N = 210)		Ajusté (N = 207)		Non-ajusté (N = 209)		Ajusté (N = 203)	
	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.
Qualité structurelle								
Score brute	-0,30	1,87	1,55	1,35	-0,10	1,00	1,48**	0,57
Score standardisé	-0,02	0,10	0,09	0,08	-0,01	0,07	0,11**	0,04
Qualité éducative								
Score brute	0,86	0,93	0,66	0,94	0,49	0,92	1,28***	0,45
Score standardisé	0,05	0,05	0,04	0,05	0,04	0,07	0,10***	0,03
Sensibilité de l'éducatrice								
Score brute	0,34	0,49	0,16	0,56	0,01	0,50	0,71**	0,27
Score standardisé	0,02	0,03	0,01	0,03	0,00	0,04	0,05**	0,02
Qualité de lecture								
Score brute	0,65	0,67	0,58*	0,29	0,52	0,44	0,40**	0,15
Score standardisé	0,04	0,04	0,03*	0,02	0,04	0,03	0,03**	0,01
Qualité globale								
Score brute	1,25	1,29	1,40	0,85	0,71	0,79	1,76***	0,41
Score standardisé	0,07	0,07	0,08	0,05	0,05	0,06	0,13***	0,03
Type de différence	Fidélité de programme							
	Vocabulaire réceptif (ÉVIP-R)				Vocabulaire expressif (EOWPVT)			
	Non-ajusté (N = 210)		Ajusté (N = 207)		Non-ajusté (N = 209)		Ajusté (N = 203)	
	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.
Fidélité structurelle								
Score brute	9,12**	3,94	6,99*	3,44	3,76	3,24	5,10***	1,64
Score standardisé	0,51**	0,22	0,39*	0,19	0,28	0,24	0,38***	0,12
Fidélité de contenu								
Score brute	6,63	9,34	3,32	4,92	5,89	6,63	4,33	3,99
Score standardisé	0,37	0,52	0,18	0,27	0,44	0,49	0,32	0,30
Fidélité globale								
Score brute	10,80*	6,22	7,33	4,64	5,47	4,85	5,97**	2,34
Score standardisé	0,60*	0,35	0,41	0,26	0,41	0,36	0,44**	0,17

Note : Contrairement aux analyses par groupe de traitement, les valeurs positives pour les effets DD indiquent une relation positive entre l'élément de fidélité et l'indicateur de vocabulaire. Les score standardisés sont des données brutes transformés en score Z ayant une moyenne de zéro et un écart-type de 1. L'affectation aux groupes à l'étude lors de l'administration des échelles est utilisée pour les comparaisons rapportées dans ce tableau. L'ÉVIP-R a été administré à 20 mois et l'EOWPVT à 24 mois. Les erreurs types robuste de Huber-White sont rapportées avec groupement par garderie. Le seuil de signification est indiqué à *** $p < 1\%$, ** $p < 5\%$; * $p < 10\%$. Diff = différence, E.T. = erreur type.

Effet médiateur de la fidélité et de la qualité lors de la première année

La logique de l'analyse de médiation présentée ici est la même qu'à la section 9.2.4 portant sur l'analyse des données de la première année. Nous retenons comme médiateurs l'indice de fidélité et l'indice de qualité qui prédisent le mieux le développement des enfants : la fidélité et la qualité structurelle. Les deux questions posées sont : (a) la fidélité/qualité de la mise en œuvre du programme explique-elle en partie ou en totalité l'effet du programme sur les résultantes? Dans la négative, nous devons conclure que d'autres aspects du programme seraient responsables des effets observés sur les indicateurs de préparation à la scolarisation (p. ex. l'Atelier famille). Les résultats d'une analyse de médiation ayant comme objectif de répondre à ces questions sont rapportés aux tableaux 9.20 et 9.21.

Fidélité

À l'examen du tableau 9.20, on voit que la fidélité structurelle explique une partie significative de l'effet du programme sur toutes les résultantes des enfants (voir effet indirect). On observe une réduction complète ou très grande de l'effet du programme (entre 23 % à 88.9 %) lorsqu'on contrôle pour la fidélité de la prestation du programme. Le contrôle de la fidélité dans l'analyse réduit la différence entre les groupes à 24 mois (DD) par 88.9 % et 81.1 % respectivement pour la Communication et le Vocabulaire expressif. Le seul effet direct significatif est celui associé à l'ÉVIP-R pour lequel il reste une partie de significative de l'effet programme à expliquer.

Tableau 9.20 : Test de médiation prenant la fidélité comme médiateur de l'effet programme à 24 mois

Résultante	Effet direct			Effet indirect	
	Effet à 24 mois	ET	% Réduction	test t	ET Sobel
Communication ¹	-0,03	0,22	-88.9 %	2,49**	0,41
Conscience de soi ¹	-0,35	0,20	Supp.	3,41***	0,31
Cognition ¹	0,10	0,17	-23 %	1,61*	0,45
Vocabulaire expressif ¹	-0,07	0,24	-81.1 %	5,14***	0,39
EOWPVT (20 mois) ¹	-0,09	0,16	-56.2 %	3,02***	0,26
ÉVIP-R (24 mois) ¹	-0,26**	0,11	Supp.	1,98**	0,41

Note : Les valeurs négatives des effets comparant les groupes (c.à.d. les effets direct) représentent un effet traitement positif (c.-à-d. un avantage pour le groupe programme). % réduction indique la réduction de la différence entre les groupes en garderie relative au estimés rapportés au tableau D2. Supp. = Effet de suppression voulant que l'effet du programme soit plus fort après avoir contrôlé pour la fidélité. Ceci veut dire qu'il existe une faible corrélation négative entre la fidélité et un autre élément utile du programme (p. ex., l'impact de l'atelier famille sur les parents). Cet effet est si faible qu'une interprétation détaillée n'est pas indiquée. ¹Médiateur est la fidélité structurelle. Le seuil de signification est indiqué à *** p < 1 %, ** p < 5 %; * p < 10 %.

Qualité

Dans le cas des résultantes provenant de l'ÉPE-AD, c'est la qualité éducative qui est adoptée comme médiateur. Dans le cas de l'ÉVIP-R et de l'EOWPVT, c'est la qualité structurelle qui était choisie en raison du fait qu'elle était la plus performante lors des analyses présentées à la section précédente. Selon les résultats présentés au tableau 9.21, l'effet du programme sur toutes

les résultantes est en partie attribuable à la qualité de la mise en œuvre, sauf pour la Communication et l'ÉVIP-R car leur effet indirect est non-significatif. Nous constatons que l'effet direct du programme sur la Communication, la Cognition et l'EOWPVT est complètement attribuable à la mise en œuvre du programme en garderie. Pour les autres indicateurs, il reste une partie de l'effet du programme à expliquer car les estimés de l'effet direct sont significatifs pour ces résultantes.

Tableau 9.21 : Test de médiation prenant la qualité comme médiateur de l'effet du programme en garderie à 24 mois

Résultante	Effet direct			Effet indirect	
	Effet à 24 mois	ET	% Réduction	test t	ET Sobel
Communication ¹	-0,16	0,11	-40.7 %	1,13	0,04
Conscience de soi ¹	-0,28***	0,07	-9.8 %	1,38*	0,04
Cognition ¹	-0,05	0,10	-61.5 %	1,51*	0,05
Vocabulaire expressif ¹	-0,32***	0,09	-13.5 %	1,55*	0,07
EOWPVT (20mois) ²	-0,14	0,09	-26.3 %	1,79**	0,04
ÉVIP-R (24 mois) ²	-0,22***	0,06	Supp.	1,08	0,06

Note : Les valeurs négatives des effets comparant les groupes (c.-à-d. les effets directs) représentent un effet traitement positif (c.-à-d. un avantage pour le groupe programme). % réduction indique la réduction de la différence entre les groupes en garderie relative aux estimés rapportés au tableau D2. Supp. = Effet de suppression voulant que l'effet du programme soit plus fort après avoir contrôlé pour la qualité. Ceci veut dire qu'il existe une faible corrélation négative entre la qualité et un autre élément utile du programme (p. ex., l'impact de l'atelier famille sur les parents). Cet effet est si faible qu'une interprétation détaillée n'est pas indiquée. ²Médiateur était la qualité structurelle. Le seuil de signification est indiqué à *** $p < 1\%$, ** $p < 5\%$; * $p < 10\%$.

Sommaire : Effet médiateur de la fidélité et de la qualité

Les analyses montrent que la fidélité et la qualité du programme en garderie sont associées au développement des enfants. La fidélité en particulier est particulièrement utile pour expliquer l'effet du programme testé. Les effets de médiation particulièrement forts observés avec la fidélité peuvent être attribués à la correspondance intime qui existe entre le programme testé et la fidélité. Par contre, les indices de qualité comprennent des aspects généraux qui ne sont pas nécessairement spécifiques au programme testé. Pour l'ÉVIP-R seulement, la fidélité n'explique pas entièrement l'effet du programme testé comme le suggère l'effet direct significatif. Pour cette résultante uniquement, une partie de l'effet du programme testé reste à être expliquée à l'aide d'autres variables (p. ex., captant l'impact des Ateliers familles sur les parents).

9.3.5. Analyses par profil linguistique

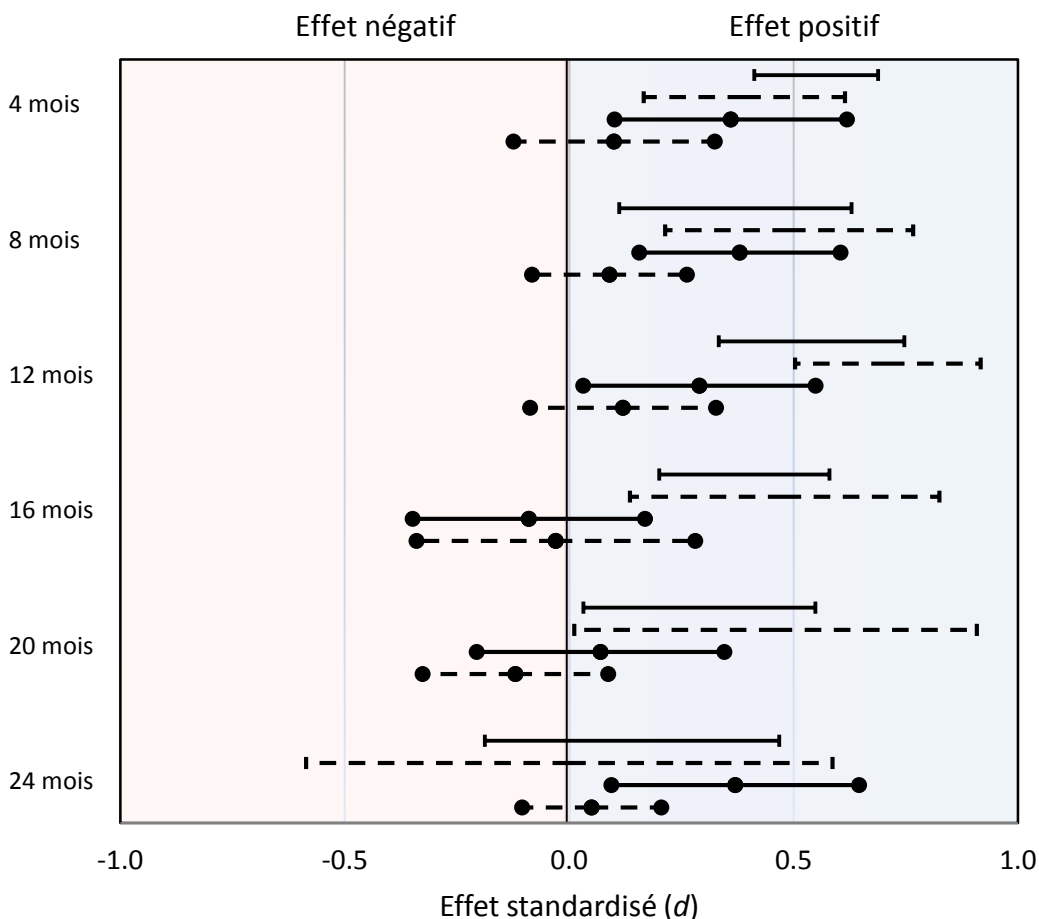
L'analyse par profil linguistique est rapportée dans les figures 9.7 à 9.10 pour les résultantes de l'ÉPE-AD. La figure 9.11 contient les résultats de l'analyse pour les autres mesures de vocabulaire (ÉVIP-R et EOWPVT). Un compte-rendu détaillé de tous ces résultats est présenté à l'annexe I. *Nous rappelons que pour les figures, une valeur positive indique un effet positif du programme testé sur le développement.* Les résultats sont maintenant décrits pour chaque résultante.

Échelle Communication

Les intervalles de confiance pour les effets standardisés du programme testé (DD) à chaque période d'évaluation sont présentés à la figure 9.7. Dans la figure, les foyers sont distingués en fonction du niveau d'exposition (forte ou faible) des enfants au français. Une inspection visuelle de la figure révèle que, tel que prévu, les effets pour cette variable ont tendance à être plus petits pour la condition forte exposition, l'effet modérateur de l'exposition étant le plus marqué à 12, 16 et 20 mois.

Pour les enfants à faible exposition, l'impact du programme est significatif relativement aux deux groupes témoins à 4, 8, 12, 16, et 20 mois. Pour le groupe à forte exposition, l'impact du programme n'est jamais significatif relativement au groupe témoin hors garderie, mais l'est relativement au groupe témoin en garderie à 4, 8, 12 et 24 mois. Dans l'ensemble, l'impact du programme sur l'échelle de la Communication est plus stable pour les enfants à faible exposition, bien que des effets positifs soient aussi observés pour le groupe à forte exposition. Il est intéressant de noter que l'effet du programme testé est supérieur à celui d'un milieu de garde informel par rapport au développement de la Communication seulement dans le contexte d'une faible exposition au français (hors garderie).

Figure 9.7 : Intervalles de confiance pour l'effet standardisé (DD) du programme sur la Communication par type de foyer



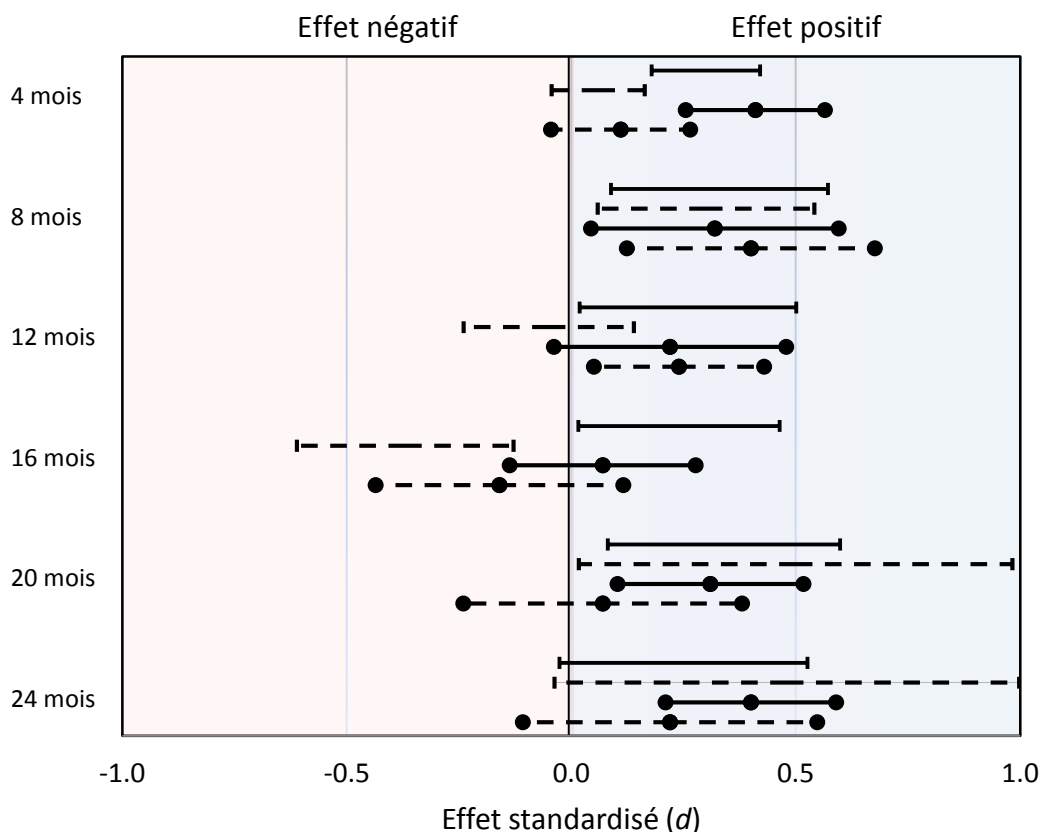
	Condition faible exposition	Condition forte exposition
G1 vs G2	-----	●-----●
G1 vs G3	-----	●-----●

Les barres représentent les intervalles de confiance de 90 %. Les intervalles de confiance ne comprenant pas la valeur de zéro indiquent un effet significatif à 10 %.

Échelle Conscience de soi

Les intervalles de confiance pour les effets standardisés du programme testé (DD) à chaque période d'évaluation sont présentés à la figure 9.8. Dans la figure, les foyers à faible et à forte exposition sont distingués. Une inspection visuelle des résultats ne révèle pas une relation claire entre le profil linguistique de la famille et la taille des effets observés. C'est peut-être dû au fait que cette échelle cible des éléments de la communication et de la cognition, ce qui pourrait avoir comme effet d'annuler les impacts du programme différenciés selon le profil linguistique.

Figure 9.8 : Intervalles de confiance pour l'effet standardisé (DD) du programme sur la Conscience de soi par type de foyer



Condition faible exposition Condition forte exposition

G1 vs G2 |—————| ●—————●

G1 vs G3 |-----| ●-----●

Les barres représentent les intervalles de confiance de 90 %. Les intervalles de confiance ne comprenant pas la valeur de zéro indiquent un effet significatif à 10 %.

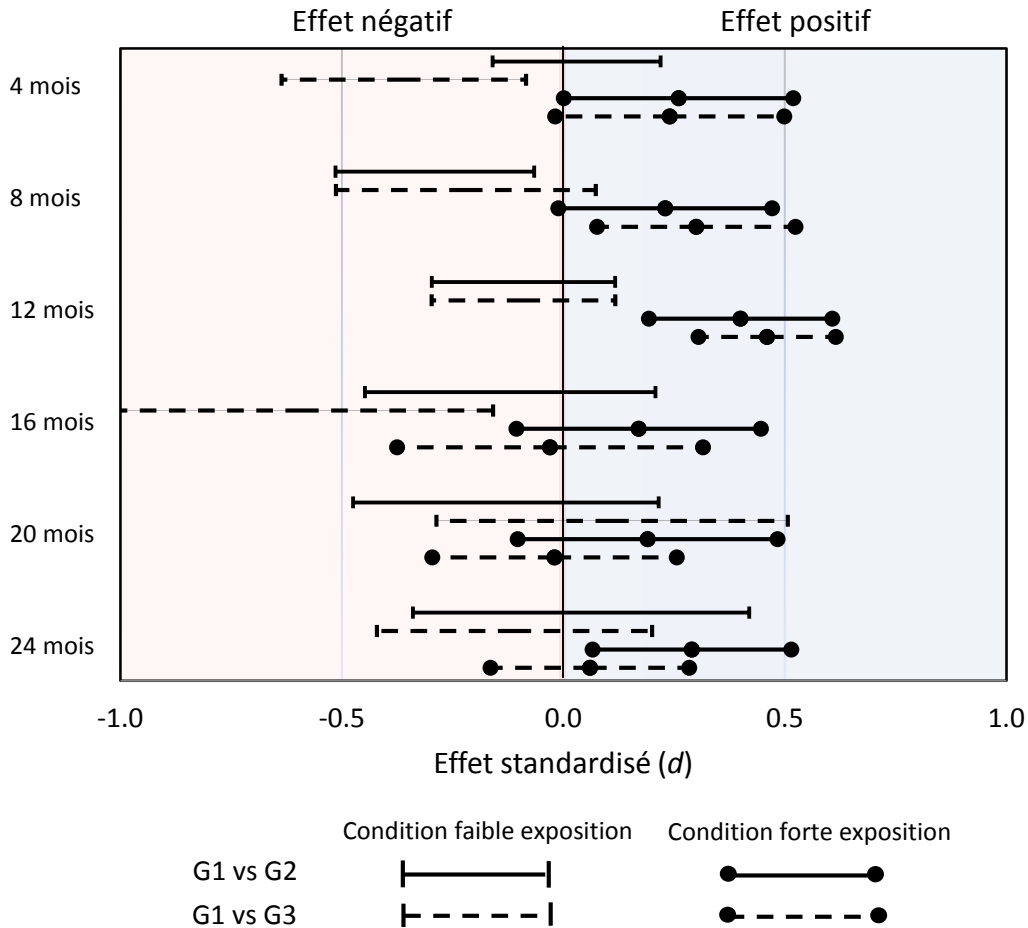
Échelle Cognition

Les intervalles de confiance pour les effets standardisés du programme testé (DD) à chaque période d'évaluation sont présentés à la figure 9.9. Encore une fois, les foyers à faible et à forte exposition au français sont distingués. L'inspection visuelle des résultats permet d'identifier un patron clair : les enfants à forte exposition retirent de plus grands bénéfices sur le plan de leur développement cognitif.

Pour les enfants à forte exposition au français, l'effet relativement au groupe témoin en garderie est significatif à 4 mois, à 12 mois et à 24 mois. Relativement au groupe témoin hors garderie, l'effet pour le même groupe d'enfants est significatif à 8 et à 12 mois. Par contre, un effet positif et significatif du programme n'est jamais observé pour les enfants à faible exposition au français. En effet, pour ces derniers, le programme testé semble avoir un effet négatif sur leur développement cognitif et ce, pour quelques périodes d'évaluation. Heureusement, cet effet

semble temporaire puisqu'il n'est pas observé à 20 et à 24 mois. Comment expliquer cet effet? Il se peut que le fait d'insister pour que la communication se fasse seulement en français dans les garderies programme présente un obstacle pour certains enfants à faible exposition au français, du moins en ce qui concerne l'acquisition des compétences cognitives. Encore une fois, cet effet semble être temporaire et compensé par des acquisitions par rapport aux compétences langagières.

Figure 9.9 : Intervalles de confiance pour l'effet standardisé (DD) du programme sur la Cognition par type de foyer



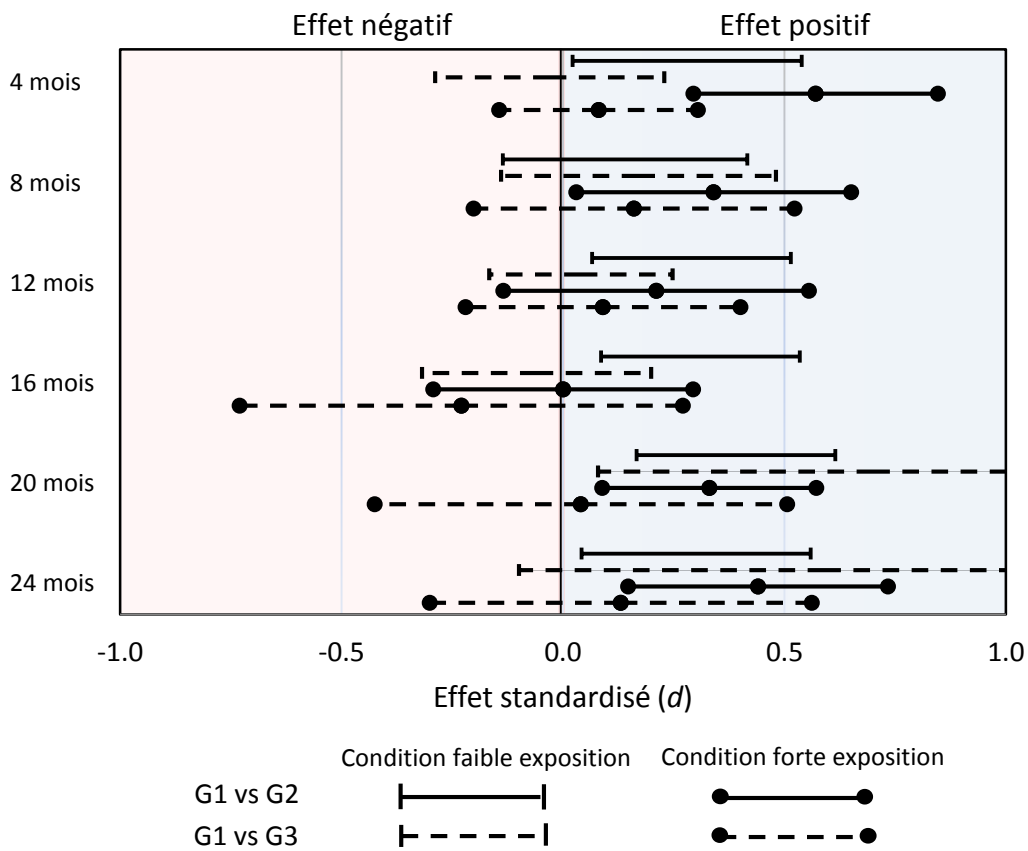
Les barres représentent les intervalles de confiance de 90%. Les intervalles de confiance ne contenant pas la valeur de zéro indiquent un effet significatif à 10%.

Échelle Vocabulaire expressif

Les intervalles de confiance pour les effets standardisés du programme testé (DD) à chaque période d'évaluation sont présentés à la figure 9.10. Encore une fois, les familles à faible et à forte exposition sont opposées. Contrairement à nos attentes, nous observons à 4 mois, un effet plus important du programme chez les enfants à forte exposition. Lors de la deuxième année, l'effet attendu est observé : des effets plus importants du programme dans le développement du vocabulaire expressif des enfants à faible exposition.

Pour les enfants à faible exposition, l'effet du programme relativement au groupe témoin en garderie est significatif à 4, 12, 16, 20 et 24 mois. Pour le même groupe, l'effet du programme est significatif relativement au groupe témoin hors garderie à 20 mois. En contraste, pour les enfants à forte exposition, l'effet du programme est significatif par rapport au groupe témoin en garderie à 4, 8, 20 et 24 mois. Pour le même groupe, l'effet du programme relativement au groupe témoin hors garderie n'est jamais significatif. En somme, l'effet du programme sur le vocabulaire expressif, tel que mesuré par cette échelle, ne semble pas dépendre étroitement de l'exposition à la langue française de l'enfant. Ceci dit, l'impact du programme semble être plus stable pour cette mesure pour les enfants à faible exposition.

Figure 9.10 : Intervalles de confiance pour l'effet standardisé (DD) du programme sur le Vocabulaire expressif par type de foyer



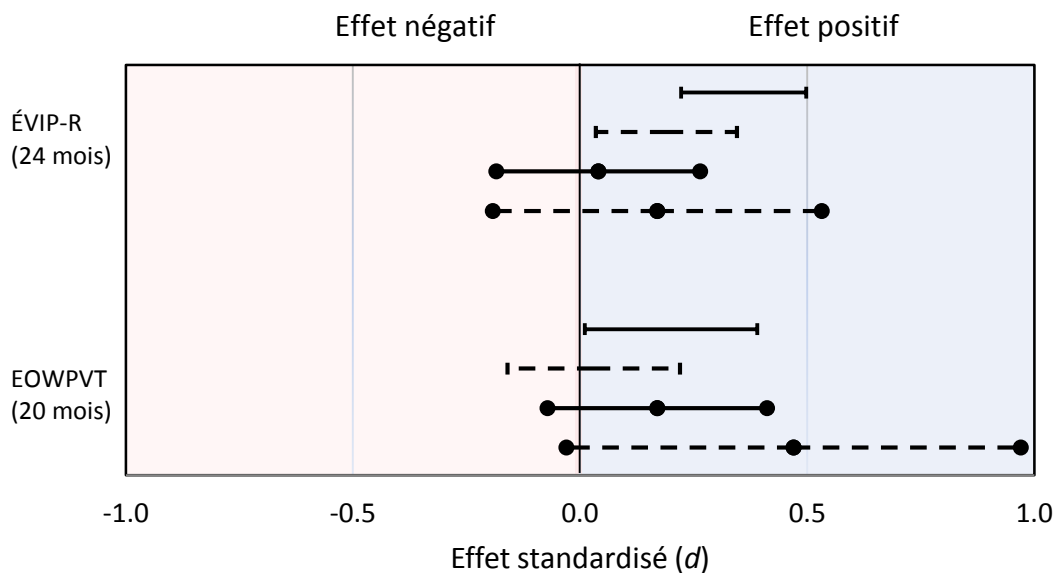
Les barres représentent les intervalles de confiance de 90 %. Les intervalles de confiance ne comprenant pas la valeur de zéro indiquent un effet significatif à 10 %.

Échelles ÉVIP-R et EOWPVT

Les intervalles de confiance pour les effets standardisés du programme testé (ANCOVA) pour l'ÉVIP-R et l'EOWPVT sont présentés à la figure 9.11. Comme pour les résultantes provenant de l'ÉPE-AD, nous distinguons entre les enfants à faible et à forte exposition. Tous les intervalles de confiance comprennent des valeurs positives indiquant que l'hypothèse d'un impact positif du programme n'est jamais infirmée.

Au niveau du vocabulaire réceptif, les résultats révèlent que, tel que prévu, les enfants à faible exposition retirent des bénéfices plus importants du programme comparativement aux enfants à forte exposition. Des effets significatifs (indiqué par l'absence de la valeur zéro dans les intervalles de confiance) sont observés uniquement pour le sous-échantillon d'enfants à faible exposition. Au niveau du vocabulaire expressif, les enfants à faible exposition développent leur vocabulaire plus rapidement que leurs pairs du groupe témoin en garderie. Aucun effet n'est observé pour les enfants à forte exposition. L'hypothèse voulant que l'impact du programme sur le vocabulaire français varie en fonction de l'exposition à la langue semble être confirmée. Toutefois, cette conclusion doit être nuancée par le fait que tous les intervalles de confiances se chevauchent : un effet n'est pas différent de l'autre (statistiquement), même si certains effets sont significativement différents de zéro.

Figure 9.11 : Intervalles de confiance pour l'effet standardisé (ANCOVA) du programme sur l'ÉVIP-R et l'EOWPVT par type de foyer



	Condition faible exposition	Condition forte exposition
G1 vs G2	-----	●-----●
G1 vs G3	-----	●-----●

Les barres représentent les intervalles de confiance de 90 %. Les intervalles de confiance ne comprenant pas la valeur de zéro indiquent un effet significatif à 10 %.

Sommaire : Analyses par profil linguistique

Les résultats tendent à supporter les hypothèses mises de l'avant au début de cette section. Les effets du programme sur les compétences langagières ont tendance à être significativement plus prononcés chez les enfants dont l'exposition à la langue française dans le foyer est faible. Par contre, les effets du programme sur les habiletés cognitives (Cognition) tendent à être

significativement plus prononcés chez les enfants issus de foyer caractérisé par une forte exposition à la langue française. Par contre, pour les autres résultantes, les intervalles de confiance pour les effets différenciés par le niveau d'exposition à la langue se chevauchent habituellement, indiquant que l'effet du programme est quand même similaire peu importe le profil linguistique. Un plus grand échantillon de participants serait nécessaire pour obtenir une réponse définitive, mais les résultats tendent toutefois à supporter l'hypothèse originelle.

9.4. SOMMAIRE

Ce chapitre avait comme objectif d'évaluer l'effet combiné des deux volets du programme testé sur la préparation à la scolarisation des enfants. L'analyse des données collectées au cours des deux années de la prestation du programme a révélé des impacts positifs et significatifs du programme sur diverses dimensions de la préparation à la scolarisation des enfants. La comparaison des groupes expérimentaux souligne un avantage pour les enfants ayant reçu le programme par rapport à leurs pairs en termes des dimensions ciblées par les échelles de l'ÉPE-AD : Communication, Cognition, Vocabulaire réceptif et Vocabulaires expressif. L'effet positif et significatif du programme sur le développement des compétences langagières est reproduit au niveau du vocabulaire réceptif (ÉVIP-R) et expressif (EOWPVT). Peu importe le groupe témoin, les enfants du groupe programme semblent mieux préparés pour un milieu scolaire en langue française. C'est le cas à l'entrée à la maternelle de quatre (à l'évaluation de 12 mois) et à l'entrée à la maternelle de cinq ans (à l'évaluation de 24 mois).

Bien que des effets significatifs soient observés, il importe d'établir s'ils sont importants. La taille des effets est relativement constante au fil des deux ans du projet, se situant selon les points de repère donnés par Cohen (1988) entre un effet petit (0,20) et moyen (0,50). Exprimée d'une autre façon, selon les données de Hill et ses collègues (2008), la taille de ces effets équivaut à une accélération du développement et de l'apprentissage d'environ quelques mois. Il se trouve que le désavantage en matière de littératie observé chez les francophones en milieu minoritaire à l'âge de quinze ans est d'à peu près de la même envergure (Allen, 2004). Cette analogie ne veut pas dire que l'intervention testée serait en soi suffisante pour combler l'écart observé en matière de littératie entre les francophones en milieu minoritaire et leurs pairs anglophones. Elle vise simplement à concrétiser la taille des effets rapportés.

La robustesse des résultats est examinée de plusieurs façons. La validité des analyses est d'abord vérifiée en reproduisant les résultats des analyses paramétriques au moyen d'analyses non-paramétrique. Ensuite, la crédibilité des constats découlant des comparaisons entre les groupes de traitement est rehaussée par une série d'analyses supplémentaires associant l'intensité de l'exposition au programme (dosage, fidélité) aux effets observés. Le résultat selon lequel l'avantage du groupe programme semble dépendre du dosage et de la fidélité de la mise en œuvre augmente notre confiance que les effets observés sont dus au programme testé (plus précisément le volet programme en garderie) et non pas à d'autres facteurs. Si la plus grande partie de l'effet du programme testé peut être expliquée par la mise en œuvre du programme en garderie, il reste à déterminer la contribution du volet Ateliers familles (voir le chapitre 10).

Enfin, le résultat qui prédit un effet différencié du programme selon le type de foyer, à faible ou à forte exposition au français, consolide la validité interne de l'étude. Ainsi, les enfants issus de foyers caractérisés par une faible exposition au français manifestent des gains plus importants

et significatifs dans le développement de leurs compétences langagières. Les enfants provenant de foyers caractérisés par une forte exposition au français bénéficient principalement au niveau de leur développement cognitif. Ces résultats soulignent des effets bénéfiques du programme peu importe le profil linguistique de l'échantillon et celui de la population francophone en milieu minoritaire. Ces points sont repris en plus de détail au Chapitre 11.

10. Impact du volet Ateliers familles

Le chapitre 10 porte sur le résultat des analyses évaluant l'apport du volet Ateliers familles. Deux questions de recherche sont considérées : a) Les Ateliers familles ont-ils eu un effet sur les attitudes et les comportements des parents? et b) Dans quelle mesure les changements observés auprès des parents sont-ils responsables de l'effet du programme testé sur les enfants? Ces questions sont abordées respectivement aux sections 10.1 et 10.2.

Le lecteur est prié de noter les points suivants :

1. Les analyses rapportées à la section 10.1.2 sont basées sur les quatre communautés d'Orléans, Cornwall, Durham et Edmundston. Les trois groupes d'enfants à l'étude figurent dans ces analyses. L'inclusion de seulement ces quatre communautés dans ces analyses a pour but de maximiser la comparabilité des résultats avec ceux découlant des analyses du programme en garderie rapportées au chapitre 9.
2. Les analyses rapportées à la section 10.1.3 sont basées sur les six communautés du projet : soit Edmonton, St-Jean, Orléans, Cornwall, Durham et Edmundston. Les résultantes à l'étude sont basées uniquement sur les enfants du groupe programme.
3. La première cohorte de participants comprend, dans la plupart des analyses, des familles provenant des quatre communautés (six dans le cas des analyses rapportées à la section 10.1.3), tandis que la deuxième cohorte comprend exclusivement des familles d'Orléans et de Cornwall.

10.1. IMPACT DIRECT SUR LES PARENTS

Cette section vise à déterminer si le programme offert dans le cadre des Ateliers familles a un effet positif sur les attitudes et les comportements des parents, membres du groupe programme.

10.1.1. Variables retenues pour l'analyse

Dans les analyses qui suivent, nous examinons deux classes de résultantes. Les données de la première classe de résultantes furent collectées par le biais d'un sondage post programme administré uniquement aux parents du groupe programme. Les résultantes examinées sont : les questions portant sur les opinions des parents quant au développement de leur enfant (Opinions), la fréquence de la modélisation d'activités de littératie (Modélisation), le sentiment d'efficacité des parents (Efficacité), l'auto-évaluation des parents par rapport à leurs propres connaissances sur le développement général des enfants et les ressources à leur disposition dans la communauté (Connaissances) et leur sentiment d'appartenance à la communauté francophone (Sentiment d'appartenance).⁸² La deuxième classe de résultantes fut collectée par le biais du sondage de base et lors de sondages de suivi administrés auprès de l'ensemble des parents du projet. Ces données sont analysées de la même façon que les données provenant des périodes d'évaluation des

⁸² Exceptionnellement, le nombre d'effectifs est maximisé en incluant les enfants provenant des six communautés du projet pour l'analyse de certaines résultantes parentales.

enfants (c.-à-d. estimation DD et analyse non paramétrique). Les résultantes examinées incluent la Fréquence et la Langue des activités de littératie.

10.1.2. Analyses par groupe

Deux résultantes furent collectées auprès des parents des trois groupes expérimentaux, et ce, à chaque sondage de suivi : la Fréquences des activités de littératie et la Langue des activités de littératie. Cette section présente les résultats d'une analyse visant à établir l'effet du programme Ateliers familles sur ces résultantes. Les résultats complets pour le modèle initial et le modèle ajusté sont présentés à l'annexe J. Le résultat des analyses d'impact du programme est présenté à la figure 10.4 pour la Fréquence des activités de littératie et à la figure 10.8 pour la Langue des activités de littératie.

Notons que dans le *Rapport des résultats de la première cohorte* (2014), le résultat des analyses d'impact a souligné l'impossibilité d'une augmentation dans la Fréquence des activités de littératie en raison d'un effet de plafonnement.⁸³ Cette information est donc importante à prendre en considération pour bien interpréter le résultat des analyses d'impact de ces résultantes. C'est ainsi que nous examinons d'abord les distributions empiriques de la Fréquence des activités de littératie et la Langue des activités de littératie pour trois périodes d'évaluation : au niveau de base (figures 10.1 et 10.5), à 12 mois (figures 10.2 et 10.6) et à 24 mois (figures 10.3 et 10.7).

Fréquence des activités de littératie

Le résultat de trois analyses est présenté dans cette section, soit : (1) la distribution empirique de l'échelle Fréquence des activités de littératie; (2) les résultats des analyses d'impact; et (3) le résultat d'une confirmation non-paramétrique des résultats obtenus avec cette échelle.

Distributions empiriques

La distribution empirique de l'échelle Fréquence des activités de littératie est présentée aux figures 10.1, 10.2 et 10.3 pour les périodes d'évaluation de base, à 12 et à 24 mois respectivement. Notons qu'on attribue à chaque choix de réponses une valeur numérique de 1 à 5 (où 1 = Jamais et 5 = Plusieurs fois par jour). Les données rapportées dans les figures reflètent la réponse « moyenne » des parents à chacun des items de l'échelle.⁸⁴ Afin de faciliter l'interprétation des figures, nous fournissons l'étiquette associée à chaque valeur numérique. L'examen des figures permet de conclure à une distribution approximativement normale de cette variable pour l'échantillon à l'étude.

⁸³ Le plafonnement est le résultat d'un choix de réponses irréalistes pour la fourchette supérieure de la distribution. Peu de parents ont le temps de faire cinq activités de littératie distinctes par jour, à une fréquence de trois fois ou plus chacune.

⁸⁴ Les scores moyens sont arrondis au chiffre le plus près dans les figures. Par contre, les analyses d'impact sont basées sur les scores non-arrondis.

L'examen des figures 10.2 et 10.3 montre que la fréquence des activités de littératie augmente d'une année à l'autre. En particulier, on note que la distribution de la figure 10.3 est décalée vers la droite par rapport à celle de la figure 10.2. Ces deux distributions sont décalées vers la droite comparativement à la figure 10.1. Ce décalage est plus marqué pour le groupe programme (G1) en raison d'un plus grand pourcentage de répondants rapportant faire des activités de littératie une fois par jour ou plus.

Figure 10.1 : Distribution de la Fréquence des activités de littératie au niveau de base

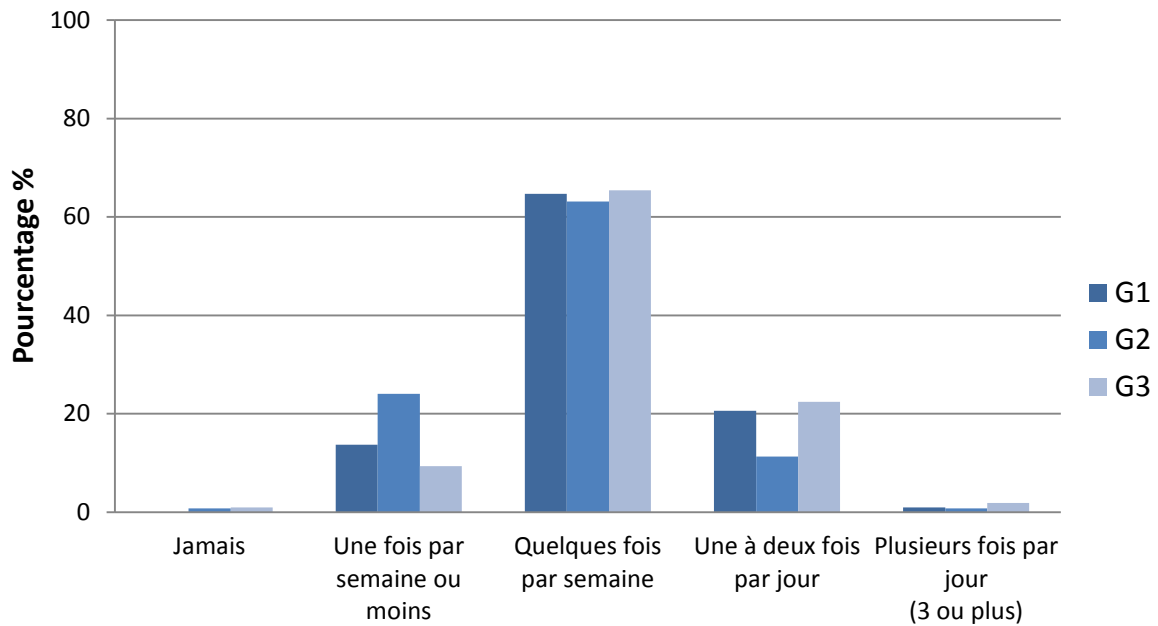


Figure 10.2 : Distribution de la Fréquence des activités de littératie à 12 mois

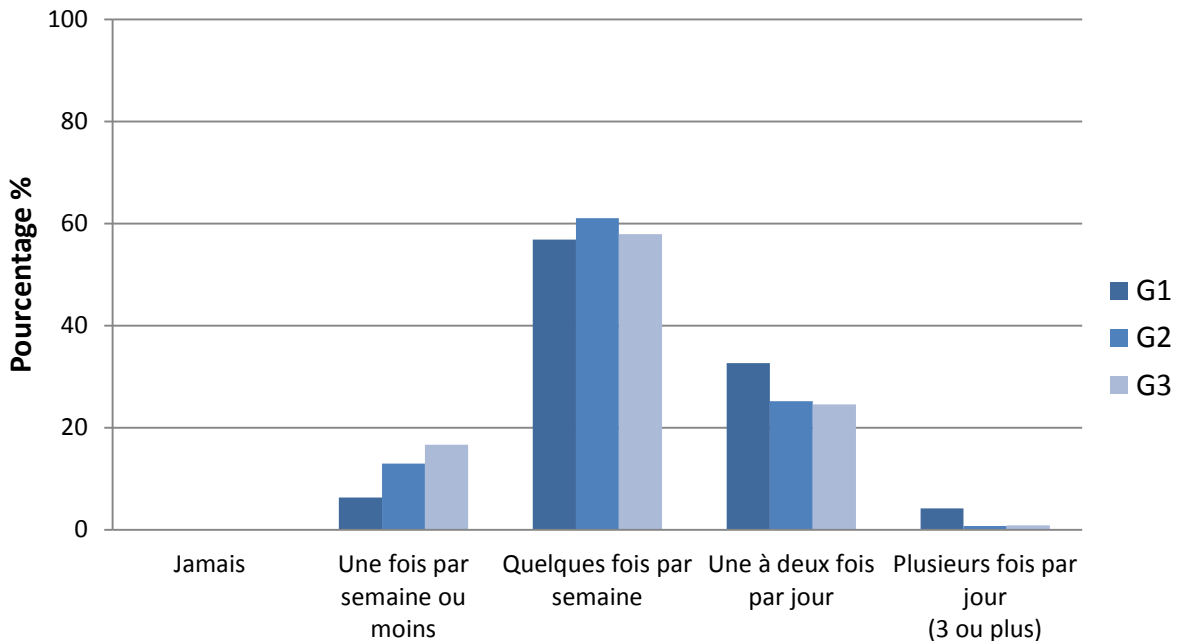
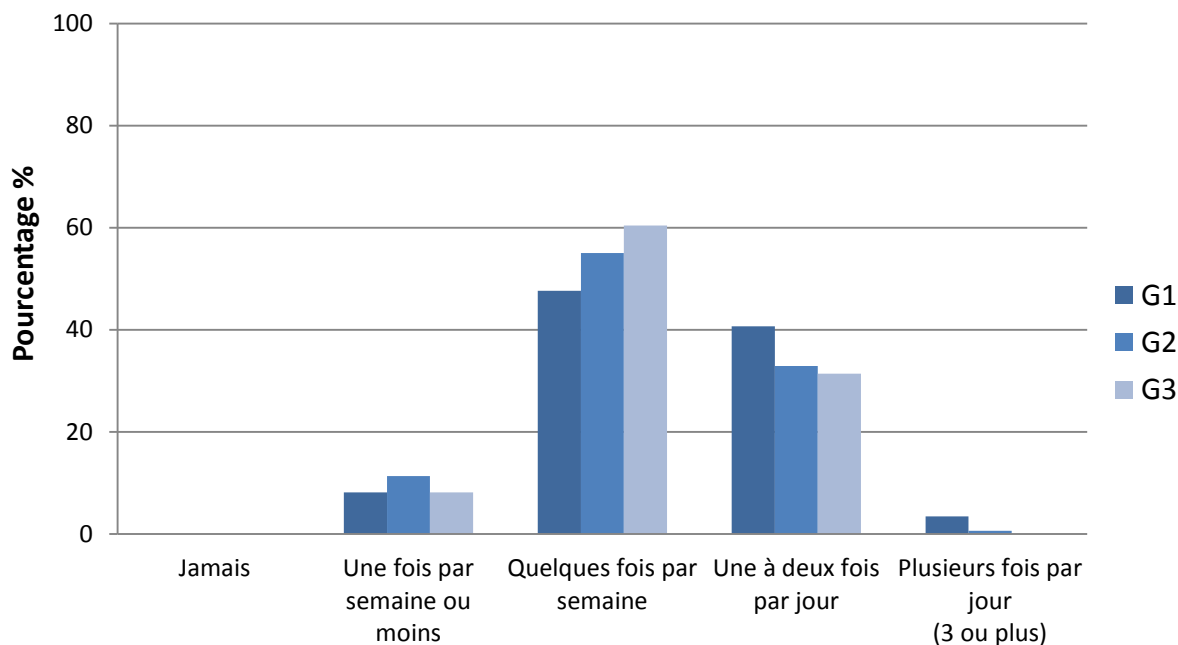


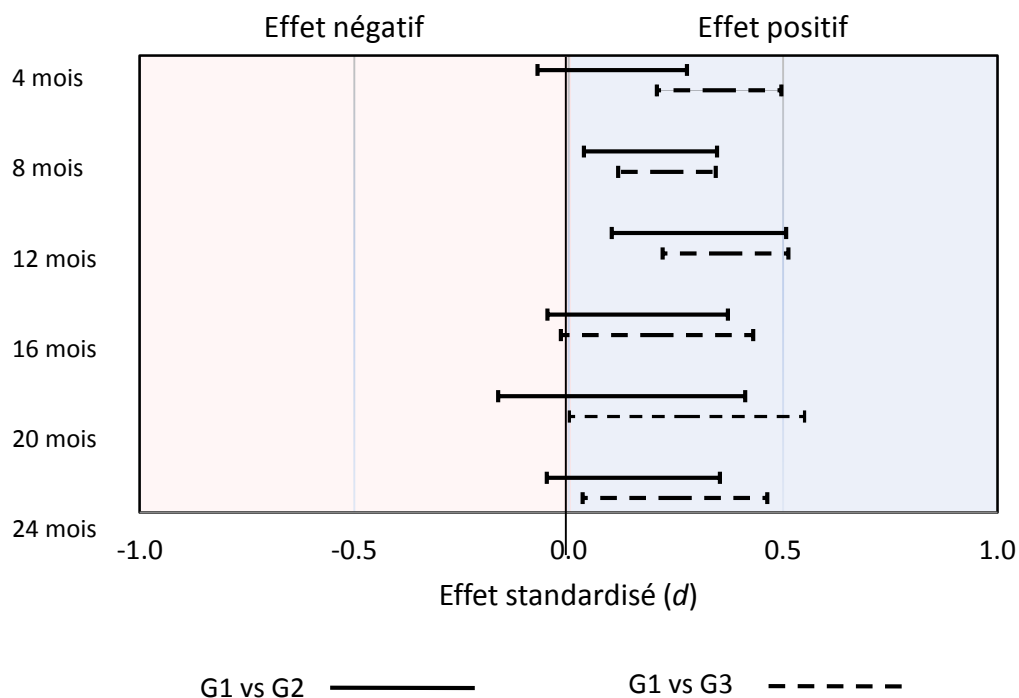
Figure 10.3 : Distribution de la Fréquence des activités de littératie à 24 mois



Modèle ajusté

Les intervalles de confiance pour les effets DD standardisés calculés selon le modèle ajusté sont rapportés à la figure 10.4. On y observe, relativement au group témoin hors garderie, un effet du programme émergeant à quatre mois et persistant jusqu'à 24 mois. Relativement au groupe témoin en garderie, un effet significatif du programme est observé à huit et à 12 mois. Ce résultat apporte un soutien empirique à l'hypothèse voulant que le programme des Ateliers familles influence positivement le comportement des parents. La vérification de cette hypothèse est une condition nécessaire mais non suffisante pour pouvoir affirmer que ce volet du programme testé contribue positivement au développement des enfants. La deuxième condition, que les changements observés chez les parents soient prédictifs du développement des enfants, fait l'objet des analyses rapportées à la section 10.2.

Figure 10.4 : L'effet du programme Ateliers familles sur la Fréquence des activités de littératie



Les barres représentent les intervalles de confiance de 90 %. Les intervalles de confiance ne comprenant pas la valeur de zéro indiquent un effet significatif à 10 %.

Contre-validation non-paramétrique

L'analyse non-paramétrique des données a été effectuée en utilisant la même stratégie qu'au chapitre 9, section 9.31. Le résultat est exprimé en termes de la probabilité d'être classifiée dans un quintile supérieur : la cote de réussite. Le rapport des cotes pour les effets DD correspondant aux effets rapportés à la figure 10.4 est rapporté au tableau 10.1.

L'examen des intervalles de confiance rapportés au tableau 10.1 confirme partiellement la signification des effets DD pour la comparaison du groupe programme au groupe témoin hors garderie. Les estimations DD pour cet effet sont significatives et en faveur du groupe programme au cours de la première année et à 24 mois. Par contre, l'analyse non-paramétrique confirme partiellement la signification des effets DD pour la comparaison du groupe programme au groupe témoin en garderie. Les rapports de cotes sont toujours en faveur du groupe programme, mais seulement l'effet à 12 mois est significatif selon l'analyse non-paramétrique. Un effet représentant un changement dans l'environnement des enfants d'une durée de quelques mois est peut-être suffisant pour entraîner des conséquences positives auprès des enfants.

Tableau 10.1 : Rapport des cotes des effets DD standardisés du modèle ajusté de la Fréquence des activités de littératie

Effet	Rapport des cotes	Intervalles de confiance 90 %	
		Borne inférieure	Borne supérieure
G1 vs G2			
4 mois	0,69	0,40	1,19
8 mois	0,79	0,54	1,14
12 mois	0,55	0,35	0,87
16 mois	0,89	0,41	1,92
20 mois	0,88	0,41	1,88
24 mois	0,72	0,42	1,23
G1 vs G3			
4 mois	0,43	0,27	0,69
8 mois	0,71	0,55	0,91
12 mois	0,55	0,38	0,80
16 mois	0,61	0,29	1,29
20 mois	0,59	0,27	1,30
24 mois	0,52	0,28	0,95

Langue des activités de littératie

Le résultat des trois analyses est présenté dans cette section, soit : (1) la distribution empirique de l'échelle Langue des activités de littératie; (2) les résultats des analyses d'impact; et (3) le résultat d'une confirmation non-paramétrique des résultats obtenus avec cette échelle.

Distributions empiriques

La distribution empirique de l'échelle Langue des activités de littératie est présentée aux figures 10.5, 10.6 et 10.7 pour les périodes d'évaluation de base, 12 et 24 mois respectivement. Notons qu'on attribue à chaque choix de réponses pour cette échelle une valeur numérique de 1 à 5 (où 1 = Jamais en français et 5 = Uniquement en français). Les données rapportées dans les figures reflètent la réponse « moyenne » des parents à chacun des items de l'échelle.⁸⁵ Afin de faciliter l'interprétation des figures, nous fournissons l'étiquette associée à chaque valeur numérique. Un examen des figures permet de conclure à une asymétrie dans la distribution de

⁸⁵ Les scores moyens sont arrondis au chiffre le plus près dans les figures. Par contre, les analyses d'impact sont basées sur les scores non-arrondis.

cette variable, la langue choisie pour les activités de littératie étant le français dans la majorité des cas.

Par ailleurs, une comparaison des distributions ne dégage pas un effet évident du programme. L'effet, s'il est présent, se situe à la fourchette inférieure de la distribution des scores de cette échelle. On y observe le plus petit pourcentage de répondants rapportant utiliser l'anglais et /ou une autre langue lors des activités de littératie. À 24 mois, à l'entrée des enfants à la maternelle de cinq ans (figure 10.7), on observe une légère tendance vers une plus grande utilisation de l'anglais et /ou une autre langue lors des activités de littératie.

Figure 10.5 : Distribution de la Langue des activités de littératie au niveau de base

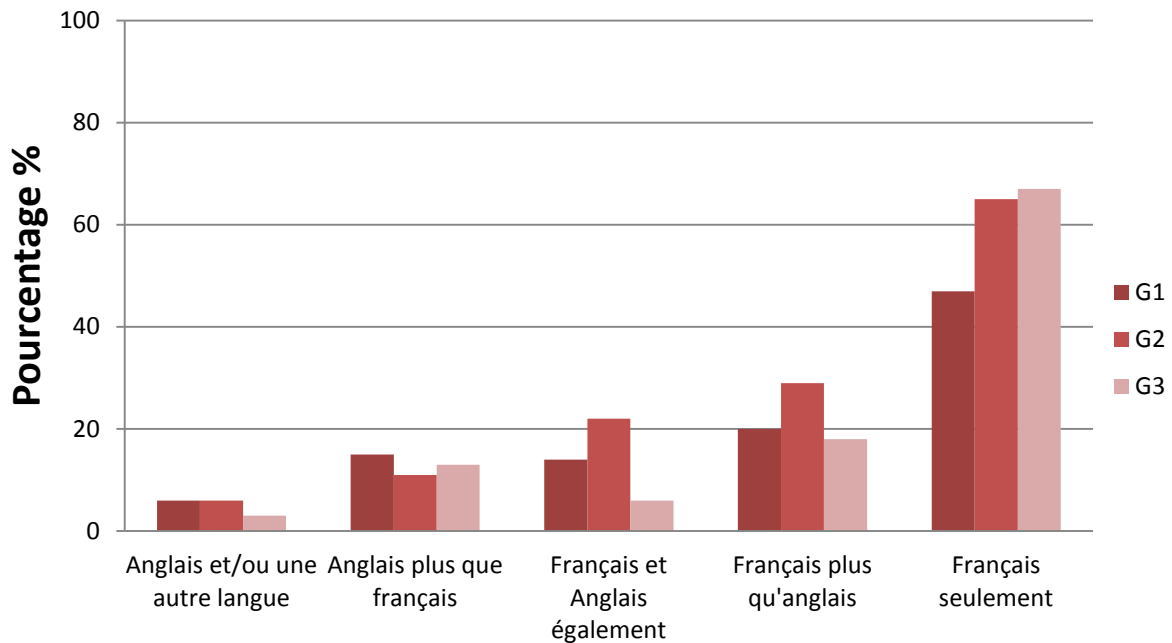


Figure 10.6 : Distribution de la Langue des activités de litt ratie   12 mois

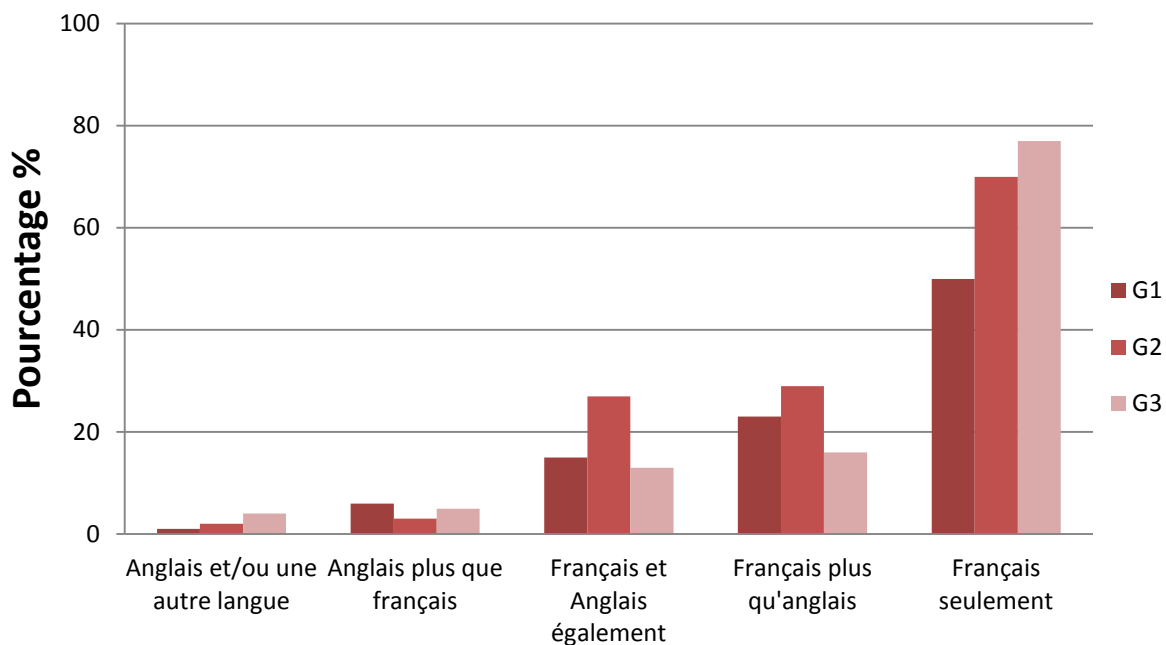
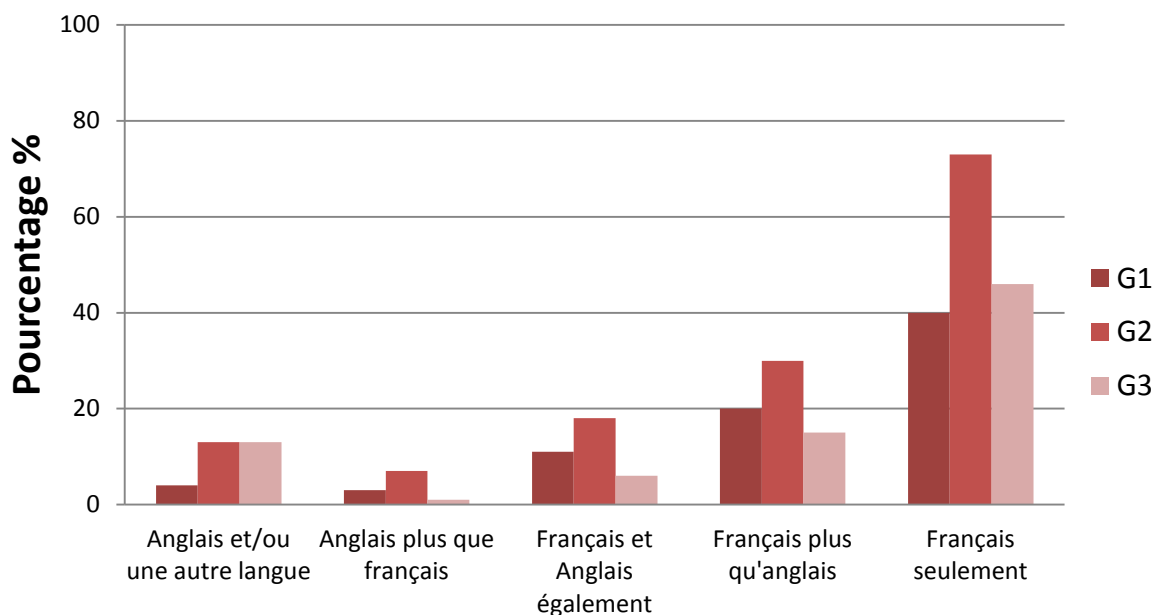


Figure 10.7 : Distribution de la Langue des activit s de litt ratie   24 mois

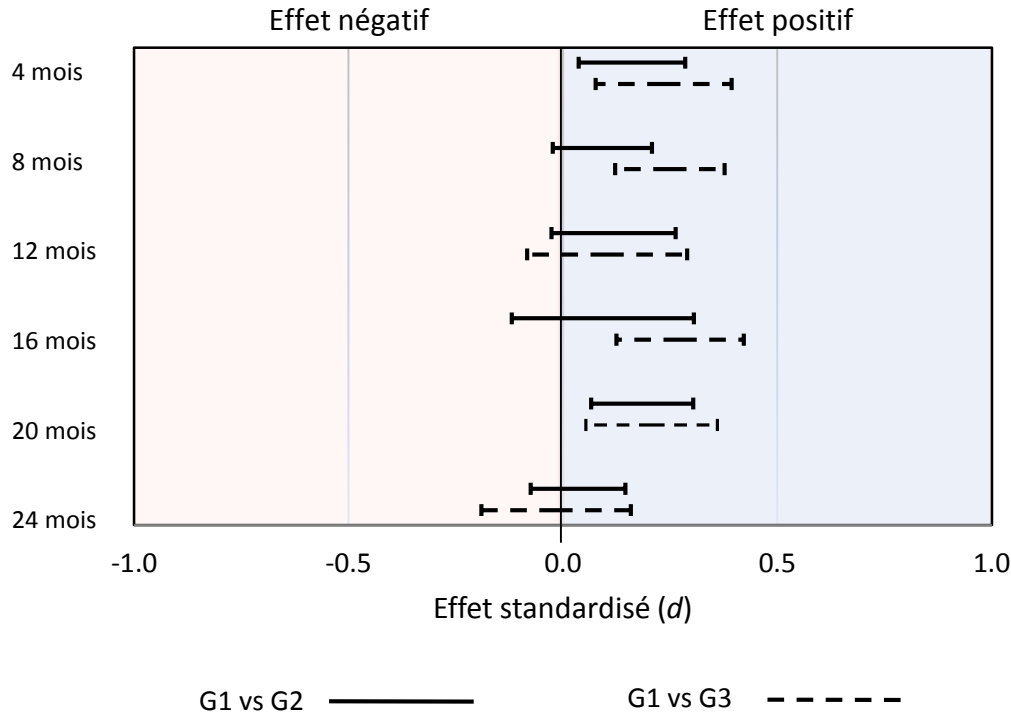


Mod le ajust 

Les intervalles de confiance pour les effets DD standardis s calcul s selon le mod le ajust  sont pr sent s   la figure 10.8. Par rapport au groupe t moin en garderie, on y observe un effet du programme  mergeant   4 mois qui r appara t   20 mois. Par rapport au groupe t moin hors

garderie, un impact du programme est observé sauf pour les évaluations à 12 mois et à 24 mois. Peu importe le groupe témoin qui sert de référence, l'effet du programme à 24 mois semble complètement atténué. Il reste à vérifier si les effets observés entre 4 et 20 mois sont suffisant pour entraîner des conséquences positives chez les enfants.

Figure 10.8 : L'effet du programme testé sur la Langue des activités de littératie



Les barres représentent les intervalles de confiance de 90 %. Les intervalles de confiance ne comprenant pas la valeur de zéro indiquent un effet significatif à 10 %.

Contre-validation non-paramétrique

L'analyse non-paramétrique des données a été effectuée utilisant la même stratégie que celle menée pour l'analyse de la Fréquence des activités de littératie. Le résultat est exprimé en termes de la probabilité d'être classifié dans un quintile supérieur : la cote de réussite. Le rapport des cotes pour les effets DD rapportés à la figure 10.8 est rapporté au tableau 10.2.

L'hypothèse selon laquelle le programme Ateliers familles influence significativement les comportements linguistiques des parents est infirmée par l'analyse non-paramétrique. Tous les intervalles de confiance rapportés au tableau 10.2 comprennent la valeur « 1 » indiquant ainsi l'absence d'un effet significatif du programme. Ces résultats ne reproduisent pas les effets DD rapportés à la figure 10.8, peut-être en raison de la distribution asymétrique de cette variable. Des analyses non-paramétriques complémentaires (non-rapportées) montrent que l'effet du programme se retrouve surtout aux deuxième et troisième quintiles. Nous prenons l'analyse non-

paramétrique comme étant la plus crédible⁸⁶ et concluons que le programme n'a aucun effet significatif sur les comportements linguistiques des parents.

Tableau 10.2 : Rapport des cotes des effets DD standardisés du modèle ajusté de la Langue des activités de littératie

Effet	Rapport des cotes	Intervalles de confiance 90%	
		Borne inférieure	Borne supérieure
G1 vs G2			
4 mois	0,78	0,44	1,38
8 mois	0,91	0,50	1,67
12 mois	0,99	0,43	2,30
16 mois	0,82	0,40	1,67
20 mois	0,85	0,62	1,14
24 mois	0,78	0,53	1,16
G1 vs G3			
4 mois	0,69	0,44	1,08
8 mois	0,98	0,61	1,56
12 mois	1,29	0,70	2,39
16 mois	0,64	0,37	1,11
20 mois	0,80	0,52	1,23
24 mois	0,88	0,57	1,35

10.1.3. Analyses portant sur les parents du groupe programme

Les résultantes qui suivent ont été mesurées uniquement auprès des parents du groupe programme et leur analyse est basée sur les six communautés. Ces données sont analysées au moyen de trois stratégies complémentaires. Dans un premier temps, les parents participants (c.-à-d. ceux ayant participé à trois ateliers ou plus, n = 116) sont comparés aux parents classifiés comme non-participants (c.-à-d. ceux ayant participé à deux ateliers ou moins, n = 14).⁸⁷ Ensuite, nous examinons si l'assiduité des parents « participants » est associée à l'ampleur de l'impact du programme. Enfin, nous rapportons le résultat d'une série d'analyses mettant en lien des indices

⁸⁶ La non-normalité de la distribution de cette variable peut nuire aux résultats des analyses paramétriques.

⁸⁷ Une analyse de puissance statistique conclut que cette comparaison à effectifs non-équilibrés peut détecter des effets réels de 0,20, 0,50 et 0,80 dans respectivement 28 %, 68 % et 94 % du temps. Un seuil de significativité de 10 % est présumé.

de fidélité et de qualité de la mise en œuvre des Ateliers familles et l'impact du programme Ateliers familles sur les résultantes des parents.

Le petit échantillon de non-participants rend impossible une correction « statistique » des différences pré programme existant entre les groupes. Précisons que le résultat de tests t conclut à une différence significative entre les groupes (participants et non-participants) uniquement par rapport à leur Capital social et leur Sentiment d'appartenance au groupe culturel.

Analyses par participation (participants versus non-participants)

Dans les prochaines sous-sections, l'effet de la participation ou de la non-participation des parents aux Ateliers familles est examiné pour chacune des résultantes parentales.

Opinions des parents

Les parents ont été invités à répondre à quatre questions concernant leurs croyances envers des facteurs liés au développement de leur enfant. Les résultats de l'analyse des réponses à trois de ces questions sont présentés au tableau 10.3, 10.4 et 10.5 respectivement. Le quatrième énoncé, « Je crois que mon enfant est trop jeune pour se faire lire des histoires », n'a pas été analysé parce que tous les parents se sont dits absolument en désaccord avec cet énoncé.

On observe au tableau 10.3 aucune différence significative pré programme entre les parents participants et les parents non-participants. La taille de l'effet se situe entre petit et moyen (c.-à-d. entre une valeur absolue de 0,10 et 0,40). Suivant les Ateliers familles, on observe une tendance non significative dans la direction attendue (c.-à-d. une croyance accrue des parents participants d'être premier éducateur de son enfant).

L'estimateur DD révèle un effet moyen des Ateliers familles ($d = 0,54$ écart-type) bien que non significatif selon le test t. Ajoutons que cet effet découle principalement d'un léger changement négatif et non-significatif observé chez les non-participants ($d = -0,38$) en parallèle avec une faible augmentation non significative observée chez les participants ($d = 0,15$).

Tableau 10.3 : Sommaire des impacts du programme Ateliers familles sur l'opinion des parents : le parent comme premier éducateur de l'enfant

	Moyennes (É-T)		Changement		
	Pré	Post	Diff (Post- Pré)	Diff (d)	test-t
Participants (P)	2,91 (0,89)	3,05 (0,88)	0,14 (0,87)	0,15	1,70
Non-participants (NP)	3,21 (0,89)	2,86 (0,95)	-0,36 (1,15)	-0,38	-1,16
Différence (P – NP)	-0,30	0,19	0,49		
Différence standardisé (d)	-0,33	0,21	0,54		
test t	1,19	0,73	1,30		

Note : Le seuil de signification : *** $p < 1\%$; ** $p < 5\%$; * $p < 10\%$.

On n'observe au tableau 10.4 aucune différence significative entre les parents participants et les parents non-participants aux Ateliers familles, et ce, avant et suivant la prestation du programme. De même, l'estimateur DD indique l'absence d'un changement significatif dans le temps autant chez les participants que chez les non-participants selon le test-t.

Tableau 10.4 : Sommaire des impacts du programme Ateliers familles sur l'opinion des parents : l'importance de créer un environnement francophone à la maison

	Moyennes (É-T)		Changement		
	Pré	Post	Diff (Post- Pré)	Diff (<i>d</i>)	test-t
Participants (P)	3,00 (0,82)	3,05 (0,88)	0,05 (0,76)	0,05	0,62
Non-participants (NP)	3,08 (0,49)	3,38 (0,51)	0,31 (0,75)	0,61	1,47
Différence (P – NP)	-0,08	-0,34	-0,26		
Différence standardisé (<i>d</i>)	-0,09	-0,39	-0,34		
test-t	0,48	2,07	1,11		

Note : Le seuil de signification : *** $p < 1\%$, ** $p < 5\%$; * $p < 10\%$.

On observe au tableau 10.5 aucun effet significatif du programme. Les participants ne se distinguent pas significativement des non-participants avant ou suivant la prestation du programme, ni par rapport aux gains observés avant et suivant la prestation du programme.

Tableau 10.5 : Sommaire des impacts du programme Ateliers familles sur l'opinion des parents : le parent et l'éducatrice travaillent ensemble pour préparer l'enfant à l'école

	Moyennes (É-T)		Changement		
	Pré	Post	Diff (Post- Pré)	Diff (<i>d</i>)	test-t
Participants (P)	3,37 (0,78)	3,52 (0,72)	0,15 (0,83)	0,19	1,91
Non-participants (NP)	3,00 (0,91)	3,00 (1,22)	0,00 (0,82)	0,00	0,00
Différence (P – NP)	0,37	0,52	0,14		
Différence standardisé (<i>d</i>)	0,47	0,66	0,17		
test-t	1,42	1,50	0,32		

Note : Le seuil de signification : *** $p < 1\%$, ** $p < 5\%$; * $p < 10\%$.

Modélisation

On observe au tableau 10.6 aucune différence significative entre les groupes avant ou suivant la prestation du programme Ateliers Familles. La fréquence des activités de modélisation ne change pas significativement d'un moment à l'autre pour les deux groupes. On observe toutefois une tendance moyenne à faire plus d'activités de modélisation chez les parents participants comparativement aux parents non-participants, et ce, avant et suivant la prestation du programme.

Tableau 10.6 : Sommaire des impacts du programme des Ateliers familles sur la modélisation (deux semaines post programme)

	Moyennes (É-T)		Changement		
	Pré	Post	Diff (Post - Pré)	Diff (<i>d</i>)	test-t
Participants (P)	18,46 (3,28)	18,80 (2,84)	0,37 (3,07)	0,12	1,28
Non-participants (NP)	17,92 (3,15)	18,62 (3,01)	0,69 (2,87)	0,22	0,87
Différence (P – NP)	0,53	0,19	-0,32		
Différence standardisé (<i>d</i>)	0,16	0,06	-0,10		
test-t	0,57	0,21	0,24		

Note : Le seuil de signification : *** $p < 1\%$, ** $p < 5\%$; * $p < 10\%$.

L'examen du tableau 10.7 ne révèle aucune différence significative entre les parents participants et les parents non-participants avant et suivant la prestation du programme Ateliers familles. En somme, aucun effet du programme ne se manifeste immédiatement après la fin de sa prestation ou à huit mois post-intervention.

Tableau 10.7 : Sommaire des impacts du programme des Ateliers familles sur la modélisation (huit mois post-programme)

	Moyennes (É-T)		Changement		
	Pré	Post	Diff (Post- Pré)	Diff (<i>d</i>)	test-t
Participants (P)	18,46 (3,28)	18,67 (3,15)	0,23 (3,33)	0,07	0,73
Non-participants (NP)	17,92 (3,15)	18,56 (3,38)	0,63 (3,18)	0,19	0,72
Différence (P – NP)	0,53	0,11 (0,98)	-0,40		
Différence standardisé (<i>d</i>)	0,16	0,03	-0,12		
test-t	0,57	0,11	0,29		

Note : Le seuil de signification : *** $p < 1\%$, ** $p < 5\%$; * $p < 10\%$.

Efficacité

Lors du dernier atelier, on a demandé aux parents participants d'estimer leur niveau d'efficacité avant et suivant les Ateliers familles. L'examen du tableau 10.8 révèle une différence significative de taille moyenne en faveur des parents non-participants comparativement aux parents participants avant la prestation du programme. Cette différence est reproduite par un test non-paramétrique (test Z de rang). Toutefois, suivant la fin du programme, les parents participants rapportent un niveau d'efficacité comparable à celui rapporté par les parents non-participants.

Cet effet de rattrapage est attribuable à des gains rapportés par les parents participants aux Ateliers familles. On observe chez les parents participants un changement positif et significatif de 0,92 écart-type avant et suivant la prestation du programme. Ce résultat est reproduit par une analyse non-paramétrique. En d'autres mots, les parents participants rapportent se sentir significativement plus confiants à bien remplir leur rôle parental suite au programme Ateliers familles.

Tableau 10.8 : Sommaire des impacts du programme Ateliers familles sur l'auto-efficacité des parents

	Moyennes (É-T)		Changement		
	Pré	Post	Diff (Post- Pré)	Diff (d)	test-t
Participants (P)	15,48 (2,61)	17,51 (1,69)	2,03 (1,7)	0,92	12,75***
Non-participants (NP)	16,93 (1,64)				
Différence (P – NP)	-1,45	0,58			
Différence standardisé (d)	-0,57	0,34			
test-t	-2,88**	1,24			

Note : Le seuil de signification : *** $p < 1\%$, ** $p < 5\%$; * $p < 10\%$.

Connaissances

Lors du dernier atelier, on a demandé aux parents participants d'estimer l'état de leurs connaissances avant et suivant les Ateliers familles. Les questions portaient sur l'état de leurs connaissances du développement général des enfants, de stratégies pour aider leur enfant à mieux se préparer pour l'école, du travail d'une éducatrice et enfin, de la disponibilité des services et des ressources de langue française dans leur communauté.

Selon le tableau 10.9, les parents participants rapportent un état de connaissances équivalent à celui rapporté par les parents non participants avant la prestation du programme. Suivant le programme, les parents participants rapportent en connaître significativement davantage comparativement aux parents non-participants. La taille de l'effet est grand, se chiffrant à environ 1,0 écart-type, peu importe que le calcul soit fait au moyen d'une comparaison des groupes (DD) ou à partir du changement dans l'état des connaissances rapporté par les participants avant et suivant le programme (test-t). Ces résultats sont reproduits par l'analyse non-paramétrique. En somme, les résultats convergent pour indiquer un effet positif et significatif du programme Ateliers familles sur les connaissances des parents.

Tableau 10.9 : Sommaire des impacts du programme Ateliers familles sur les connaissances des parents

	Moyennes (É-T)		Changement		
	Pré	Post	Diff (Post - Pré)	Diff (d)	test-t
Participants (P)	17,19 (3,32)	20,77 (2,1)	3,58 (2,53)	1,29	15,24***
Non-participants (NP)	18,54 (2,73)				
Différence (P – NP)	-1,35	2,23			
Différence standardisé (d)	-0,41	1,02			
test-t	-1,65	2,85*			

Note : Le seuil de signification : *** $p < 1\%$, ** $p < 5\%$; * $p < 10\%$.

Sentiment d'appartenance

On n'observe au tableau 10.10 aucune différence significative entre les groupes de parents avant ou suivant la prestation du programme. Notons toutefois une différence de taille moyenne entre les groupes suivant le programme bien que celle-ci soit non significative. L'analyse

comparant les scores des participants avant et suivant le programme décèle un changement significatif chez les parents participants. L'effet est de taille moyenne, soit de 0,55 écart-type. Il est reproduit par l'analyse non-paramétrique.

Tableau 10.10 : Sommaire des impacts du programme Ateliers familles sur leur appartenance à la communauté francophone

	Moyennes (É-T)		Changement		
	Pré	Post	Diff (Post - Pré)	Diff (<i>d</i>)	test-t
Participants (P)	2,99 (1,21)	3,60 (1)	0,61 (0,74)	0,55	8,86***
Non-participants (NP)	2,93 (1,14)				
Différence (P – NP)	0,06	0,67			
Différence standardisé (<i>d</i>)	0,05	0,66			
test-t	0,19	2,11			

Note : Le seuil de signification : *** $p < 1\%$, ** $p < 5\%$; * $p < 10\%$.

Analyses par dosage

La fréquence des présences aux ateliers (ou assiduité) des parents participants est variable. Il est donc possible de vérifier l'hypothèse voulant que les parents assidus (c.-à-d. ayant le plus grand dosage) connaissent de plus grands changements suite aux ateliers comparativement aux parents participants moins assidus. Cette hypothèse est mise à l'épreuve aux tableaux 10.11 et 10.12 respectivement pour ce qui est des opinions et des autres résultantes des parents. Dans ces tableaux, nous rapportons l'association entre les résultantes des parents et le dosage. Précisons que ces analyses par dosage sont faites au moyen de divers indicateurs du dosage : présence de l'enfant, présence de la mère, présence du père, et présence d'un parent.

On observe au tableau 10.11 une association positive et significative entre le dosage et l'opinion des parents par rapport à leur rôle comme premier éducateur de leur enfant. Le taux de présence du père est significativement relié à l'opinion ayant trait à l'importance de l'environnement francophone. Il semble donc qu'une participation active des parents aux ateliers soit associée aux impacts escomptés sur ces résultantes.

On observe au tableau 10.12 une association positive et significative entre le dosage et les résultantes Efficacité et Connaissance. Lorsque l'indicateur du dosage considéré est la présence d'un parent, on observe une corrélation de 0,34 avec les scores de changement de la résultante Efficacité et de 0,45 avec les scores de changement de la résultante Connaissance. En d'autres mots, une augmentation de 27 % de l'assiduité (par rapport à la moyenne de 67 %) est associée à un accroissement de 0,34 écart-type en Efficacité et de 0,45 écart-type dans l'état des Connaissances.

Tableau 10.11 : Corrélations entre le dosage et trois questions ciblant l'opinion des parents

Indicateur	Statistiques descriptives		Premier éducateur (Tableau 10.3)			Environnement francophone (Tableau 10.4)			Parent-éducatrice (Tableau 10.5)		
	M	É-T	Pré	Post	Diff	Pré	Post	Diff	Pré	Post	Diff
Présence de l'enfant	0,67	0,27	-0,01	0,19**	0,20**	0,09	0,14*	0,06	0,02	0,04	-0,05
Présence de la mère	0,56	0,3	0,10	0,16**	0,05	0,00	-0,01	-0,01	-0,05	0,08	0,06
Présence du père	0,29	0,31	-0,11	0,10	0,21***	0,09	0,22***	0,15**	0,06	0,06	-0,03
Présence d'un parent	0,67	0,27	-0,02	0,18**	0,20**	0,10	0,15**	0,07	0,02	0,04	-0,05

Note : N = 111. Corrélations de Pearson. Les corrélations Spearman sont comparables (non rapportées). Les résultats clés sont les corrélations avec le Diff (différence post-pré). Le seuil de signification : *** p < 1 %, ** p < 5 %; * p < 10 %.

Tableau 10.12 : Corrélations entre dosage et Modélisation, Efficacité, Connaissances et Appartenance

Indicateur	Modélisation (Tableau 10.6)			Efficacité (Tableau 10.8)			Connaissances (Tableau 10.9)			Appartenance (Tableau 10.10)		
	Pré	Post	Diff	Pré	Post	Diff	Pré	Post	Diff	Pré	Post	Diff
Présence de l'enfant	-0,02	0,04	-0,02	-0,12*	0,15	0,36***	-0,24**	0,14	0,44***	0,07	0,18**	0,13*
Présence de la mère	-0,02	0,00	0,01	-0,10	0,10	0,26***	-0,05	0,21***	0,25***	0,02	0,11	0,12*
Présence du père	0,03	0,07	-0,01	0,01	0,12	0,12*	-0,22**	-0,09	0,20**	0,12*	0,13*	-0,01
Présence d'un parent	-0,02	0,05	-0,03	-0,11	0,15	0,34***	-0,24**	0,15**	0,45***	0,08	0,18**	0,12*

Note : N = 111. Corrélations de Pearson. Les corrélations Spearman sont comparables (non-rapportées). Les résultats clés sont les corrélations avec le Diff (différence post-pré). Spearman sont comparables (non-rapportées). Les résultats clés sont les corrélations avec le Diff (différence post-pré). Le seuil de signification : *** p < 1 %, ** p < 5 %; * p < 10 %.

Analyses par qualité

Dans cette sous-section, nous examinons si la qualité de la prestation des ateliers permet d'expliquer le changement observé chez les parents participants. Les résultats de cette analyse sont présentés au tableau 10.13 pour les opinions des parents et au tableau 10.14 pour les autres résultantes.

On observe au tableau 10.13 que la qualité de l'animation est associée à un changement positif et significatif de l'opinion des parents relativement au travail collaboratif du parent et de l'éducatrice pour favoriser le développement et la préparation à l'école de l'enfant. De même, une association positive et significative est observée entre la quantité du contenu abordée et le changement dans l'opinion des parents participants relativement à leur rôle en tant que premier éducateur de leur enfant. Enfin, une association positive et significative est observée entre les réactions positives des parents (observées durant les ateliers) et un changement vis-à-vis leur rôle en tant que premier éducateur de leur enfant (voir le tableau 10.3).⁸⁸

L'examen du tableau 10.14⁸⁹ indique que le pourcentage de contenu abordé est associé à une augmentation des activités de modélisation. Pour sa part, la qualité de l'animation est liée à des gains observés chez les participants par rapport à l'Efficacité et la Connaissance. La durée des séances est, quant à elle, corrélée positivement avec une augmentation des activités de modélisation et à un accroissement des sentiments d'appartenance.

⁸⁸ La probabilité qu'un des douze tests de corrélations des scores « Diff » soit significatif simplement par chance est élevée.

⁸⁹ La probabilité que deux des 24 tests de corrélations des scores « Diff » soit significatif simplement par chance est élevée.

Tableau 10.13 : Corrélations entre la qualité de la prestation du programme et trois questions ciblant l'opinion des parents

Indicateur	Statistiques Descriptives		Premier éducateur (Tableau 10.3)			Environnement francophone (Tableau 10.4)			Parent-éducatrice (Tableau 10.5)		
	M	É-T	Pré	Post	Diff	Pré	Post	Diff	Pré	Post	Diff
% du contenu abordé	0,97	0,03	-0,11	0,11	0,21**	0,05	0,08	0,04	-0,01	-0,11	-0,13
Qualité de l'animation	0,71	0,30	-0,10	-0,15	-0,05	-0,22***	-0,12*	0,09	0,03	0,11	0,26***
Réactions positives observés chez les parents	0,77	0,18	-0,08	0,09	0,17**	-0,06	-0,05	0,02	0,05	0,03	0,07
Réactions négatives observés chez les parents	0,60	0,18	0,08	0,12*	0,04	-0,02	-0,05	-0,03	-0,12*	-0,15	-0,09

Note : Les résultats clés sont les corrélations avec le Diff (différence post-pré). N = 111. Corrélations de Pearson. Les corrélations Spearman sont comparables (non-rapportées).
Le seuil de signification : *** p < 1 %, ** p < 5 %; * p < 10 %.

Tableau 10.14 : Corrélations entre la qualité de la prestation du programme et la Modélisation, l'Efficacité, Connaissances et l'Appartenance.

Indicateur	Modélisation (Tableau 10.6)			Efficacité (Tableau 10.8)			Connaissances (Tableau 10.9)			Appartenance (Tableau 10.10)		
	Pré	Post	Diff	Pré	Post	Diff	Pré	Post	Diff	Pré	Post	Diff
% du contenu abordé	-0,19	0,21	0,33***	-0,17*	-0,16*	0,08	-0,15	-0,17**	0,04	-0,06	0,01	0,11
Qualité de l'animation	-0,05	-0,08	-0,14*	-0,08	0,15**	0,28***	-0,04	0,21**	0,23***	0,13*	0,06	-0,12*
Réactions positives observés chez les parents	-0,11	0,65	0,10	-0,16**	-0,13	0,11	-0,14*	-0,13	0,07	-0,31***	-0,27***	0,13*
Réactions négatives observées chez les parents	0,04	0,06	0,01	-0,05	-0,06	0,00	-0,02	-0,07	-0,03	-0,16	-0,06	0,16
Durée des séances	-0,13	0,10	0,13*	0,00	-0,18	-0,19**	0,08	-0,08	-0,17*	-0,39***	-0,21**	0,34***
Qualité globale	-0,15	0,34	0,03	-0,20**	-0,09	0,20**	-0,11	0,00	0,14	0,30***	0,11	-0,33***

Note : résultats clés sont les corrélations avec le Diff (différence post-pré). N = 111. Corrélations de Pearson. Les corrélations Spearman sont comparables (non-rapportées). Le seuil de signification : *** p < 1 %, ** p < 5 %; * p < 10 %.

10.1.4. Sommaire des analyses d'impact sur les parents

Le résultat des analyses montre que le programme des Ateliers familles influence positivement la fréquence des activités de littératie menées par les parents, leurs connaissances et leur sentiment d'auto-efficacité. Les analyses subséquentes montrent que le taux de participation (en particulier, le taux de présence d'un parent lors des ateliers), la fidélité du contenu abordé, la qualité de l'animation et la durée des séances sont les principaux déterminants de l'impact des Ateliers familles, peu importe la résultante parentale considérée. En somme, les analyses montrent que ce volet du programme testé atteint un de ces buts: effectuer un changement dans les connaissances, attitudes et comportements des parents. Il reste à déterminer si les changements observés chez les parents sont, à leur tour, associés à des gains développementaux observés chez les enfants du groupe programme.

10.2. IMPACT INDIRECT SUR LES ENFANTS

L'objectif des analyses rapportées dans cette section est de déterminer dans quelle mesure l'impact du programme sur les enfants peut être attribué à un effet du programme des Ateliers familles. Cet objectif est atteint au moyen d'une analyse de médiation vérifiant si la mise en œuvre du programme Ateliers familles est responsable, en tout ou en partie, pour les effets du programme sur les résultantes des enfants. Pour la Fréquence et la Langue des activités de littératie, une analyse de médiation formelle est possible (voir la figure 9.1). Pour les autres variables parentales, la stratégie privilégiée est d'associer les scores de changement des parents (pré- et post-ateliers) avec les scores de changement dans le développement des enfants.⁹⁰ C'est une façon d'estimer le lien « indirect » *ab* illustré à la figure 9.1. Si ce lien est absent, on conclut en l'absence d'une médiation de l'effet du programme.

Dans la section 10.2.1, nous vérifions si les changements observés chez les parents sont associés aux changements observés dans le développement des enfants durant la première année (score de changement = score à la période d'évaluation de 12 mois – score à la période d'évaluation de base). Dans la section 10.2.2, cette analyse est répétée, mais les scores de changement cette fois captent les gains observés entre la période d'évaluation de base et la période d'évaluation à 24 mois.

10.2.1. Lors de la première année

Des scores de changement ont été calculés pour quatre échelles standardisées de l'ÉPE-AD : Communication, Conscience de soi, Cognition et Vocabulaire expressif. Ces scores représentent le changement dans la position relative de chaque enfant dans la distribution des scores observée à la période d'évaluation de base et dans la distribution des scores observée à la période d'évaluation à 12 mois. Des scores de changement ont également été calculés pour les résultantes parentales suivantes : Fréquences des activités de littératie, Langue des activités de littératie, Connaissances, Efficacité, Modélisation, et Appartenance.

⁹⁰ Une analyse de médiation est menée seulement pour la Fréquence et la Langue des activités de littératie. Les données pour les groupes témoins n'ont pas été collectées pour les autres résultantes parentales.

L'analyse avait comme objectif de déterminer si les changements observés auprès des parents du groupe programme permettrait de mieux prédire le développement des enfants. Cette analyse comprenait un modèle initial et un modèle ajusté.⁹¹

Communication

Les résultats de l'analyse démontrent un effet significatif du programme des Ateliers familles favorisant un accroissement du sentiment d'efficacité des parents et l'impact positif subséquent de ce sentiment sur les compétences en communication des enfants du groupe programme, soit un gain de 1.6 %. Aucun autre indicateur parental de l'effet du programme Ateliers familles atteint le seuil de signification de 10 %. Nous attribuons peu de crédibilité à ce résultat puisqu'a priori, les autres indicateurs (p. ex. changements dans la fréquence et la langue utilisée lors des activités de littératie) sont jugés être des déterminants plus importants pour le développement des habiletés en communication en français et qu'aucun effet correspondant n'est observé pour le Vocabulaire expressif.

Conscience de soi

Aucun effet significatif n'est décelé pour l'ensemble des indicateurs parentaux.

Cognition

Aucun effet significatif n'est décelé pour l'ensemble des indicateurs parentaux.

Vocabulaire expressif

Aucun effet significatif n'est décelé pour l'ensemble des indicateurs parentaux.

10.2.2. Lors de la deuxième année

Des scores de changements ont été calculés pour les quatre échelles standardisées de l'ÉPE-AD : Communication, Conscience de soi, Cognition et Vocabulaire expressif. Ces scores représentent le changement dans la position de chaque enfant relativement à la distribution des scores observée à la période d'évaluation de base et à la distribution des scores observée à la période d'évaluation à 24 mois. Des scores de changement ont également été calculés pour les résultantes parentales suivantes : Fréquences des activités de littératie, Langue des activités de littératie, Connaissances, Efficacité, Modélisation, et Appartenance. Pour l'ÉVIP et l'EOWPVT, les scores bruts ont été associés aux scores de changement parentaux.

Communication

Aucun effet significatif n'est décelé pour l'ensemble des indicateurs parentaux.

Conscience de soi

Aucun effet significatif n'est décelé pour l'ensemble des indicateurs parentaux.

⁹¹ Le modèle ajusté ajoute au modèle initial des covariables ainsi que les résultantes parentales mesurées au niveau de base.

Cognition

Une série d'analyses de régression indépendantes montrent que les gains parentaux en Efficacité et en Connaissance expliquent, respectivement, 3.9 % et 6.3 % des gains observés dans le développement cognitif des enfants du groupe programme. Pour sa part, les gains observés dans le sentiment d'appartenance expliquent 8.6 % des gains observés dans le développement cognitif des enfants du groupe programme. L'ensemble de ces résultats suggère un lien significatif entre l'effet des Ateliers familles sur les attitudes des parents et leur impact subséquent sur le développement cognitif des enfants. Il est intéressant de noter que cette relation émerge plus d'un an plus tard.

Vocabulaire expressif

Aucun effet significatif n'est décelé pour l'ensemble des indicateurs parentaux.

L'ÉVIP-R et l'EOWPVT

Aucun effet significatif n'est décelé pour l'ensemble des indicateurs parentaux.

10.3. SOMMAIRE

Ce chapitre visait d'abord à vérifier l'impact des Ateliers familles sur les attitudes et comportements des parents du groupe programme. Le résultat des analyses d'impact confirme un effet positif et significatif du programme testé sur les connaissances, les attitudes et les comportements des parents. Ce constat est appuyé par des analyses selon les groupes expérimentaux, des analyses portant sur les parents du groupe programme (participants vs non-participants), des analyses selon le dosage et des analyses selon la fidélité et la qualité de la mise en œuvre du programme Ateliers familles.

Ensuite, ce chapitre cherchait à vérifier si l'impact du programme sur les parents se transférait en partie pour affecter le développement des enfants. Les résultats tendent à infirmer cette hypothèse. Seul le développement cognitif des enfants lors de la deuxième année du projet semble affecté par le changement dans les attitudes des parents, changement attribué aux effets des Ateliers familles. Ainsi, si les Ateliers familles sont responsables pour une partie de l'impact du programme sur les enfants, cet effet est limité à une dimension de la préparation à la scolarisation, du moins selon les variables considérées dans les analyses menées dans le présent rapport. Dans l'ensemble, le programme en garderie est une source plus crédible et plus importante de l'effet du programme testé en raison des résultats des analyses de médiation rapportés au chapitre 9.

11. Discussion

La présente recherche comportait deux objectifs : mener une étude de la mise en œuvre d'une nouvelle intervention à deux volets et en évaluer ses impacts sur de jeunes enfants de minorité francophone et leurs parents. Le programme testé visait à promouvoir le développement des compétences langagières et la préparation à la scolarisation des enfants francophones vivant dans un contexte de langue minoritaire. Selon certains théoriciens, le moteur de l'acquisition de ces compétences est l'interaction d'un enfant avec son environnement (Bronfenbrenner et Morris, 1998). En s'inspirant de la théorie écologique de Bronfenbrenner (1979) sur le développement de l'enfant, nous avons conçu l'univers de l'enfant comme une série de sphères d'influence qui se chevauchent partiellement et qui comprennent le foyer, la garderie (ou l'école) et la communauté. Comme nous l'avons vu dans l'introduction, la langue majoritaire peut influencer les personnes qui emploient la langue minoritaire dans chacune de ces sphères, y compris les deux milieux primaires pour les jeunes enfants, soit le foyer et la garderie. Par conséquent, une approche à deux volets a été employée pour obtenir un effet maximal sur les résultantes de l'enfant (Reese et al., 2010) : a) des Ateliers d'alphabétisation familiale (Ateliers familles) cherchant à influencer positivement l'environnement au foyer par l'entremise des parents, et b) un programme francophone en garderie de grande qualité (programme en garderie) cherchant à influencer les enfants plus directement.

Les prochaines sections portent sur les résultats de l'analyse de la mise en œuvre (section 11.1) et les résultats des analyses d'impact (section 11.2). La conclusion du rapport est présentée à la section 11.3.

11.1. LA MISE EN OEUVRE DU PROJET

Quelles sont les principales conclusions de l'évaluation de la mise en œuvre du programme testé dans le projet Capacité d'apprentissage? Ce programme à deux volets est basé sur les meilleures pratiques dans le domaine du développement de la petite enfance et celui de la littératie familiale. L'étude de la mise en œuvre de ce programme a servi plusieurs objectifs. Elle a d'abord permis de bien comprendre comment un programme élaboré sur papier s'intègre dans un milieu réel. Typiquement, l'évaluation de la mise en œuvre est utilisée pour comprendre quels sont les obstacles, les facilitateurs et les ajustements nécessaires pour faciliter l'appropriation du programme par le milieu. Elle sert également d'outil important pour mieux comprendre et nuancer les effets du programme. Ainsi, de plus en plus d'études démontrent que le degré de mise en œuvre d'une intervention est en lien avec les effets observés du programme sur les participants (Charlebois, et al., 2004; Conduct Problem Prevention Research Group, 1999; Dane et Schneider, 1998; Durlak et DuPre, 2008). Le programme n'aura pas le même impact si plusieurs éléments ne sont pas implantés et si la qualité de la mise en œuvre est médiocre. Sur ce point, Durlak et DuPre (2008) ont conclu à partir des données de cinq méta-analyses que la taille de l'effet d'une intervention est de deux à trois fois plus grande lorsque le programme est mis en œuvre tel que prévu.

Outre l'approche adoptée dans l'étude, un portrait global de la mise en œuvre du projet Capacité d'apprentissage a été dégagé de la triangulation de données provenant de sources variées. L'adoption d'une approche méthodologique à trois niveaux (Patton, 1990) a permis cette triangulation. C'est donc dire que nous avons tiré avantage de différents outils de mesure tant quantitatifs (comme dans le cas des échelles de fidélité et de qualité) que qualitatifs (comme pour les entrevues auprès des participants). Cette façon de faire, connue aussi sous le nom de *mixed-method model*, offre l'avantage de documenter de plusieurs façons les effets d'un programme, et permet une évaluation plus approfondie d'un programme (Patton, 2008). La méthodologie employée a également donné la parole à différents acteurs, offrant ainsi différentes perspectives sur la mise en œuvre du programme. Ensuite, les données ont été recueillies à l'aide de différents instruments de mesure pour évaluer le même phénomène. Cette diversité de données renforce la validité et la fiabilité des résultats en raison de la complémentarité des données collectées.

Les constats qui suivent tiennent compte de l'intégrité de la mise en œuvre du programme d'abord du volet en garderie et ensuite, du volet d'alphabétisation familiale.

11.1.1. Le programme en garderie

Le programme évalué en garderie est novateur à plusieurs égards. Les résultats des analyses de la mise en œuvre soulignent que le programme testé contribue de nouveaux éléments autant par rapport à son environnement structurel qu'à son contenu. Autrement dit, certains éléments du nouveau programme ne se retrouvaient pas nécessairement dans les milieux de garde du groupe témoin. L'examen de l'intégrité du programme en garderie s'est fait à partir de sept dimensions, notamment la fidélité structurelle, c'est-à-dire la correspondance entre l'intervention appliquée et le programme planifié; la fidélité et la qualité du contenu éducatif du programme, par exemple les éléments qui doivent être mis en place dans les activités et la qualité de la mise en œuvre de ces éléments; le dosage, soit l'exposition des participants au programme; la différenciation entre le programme implanté et les autres interventions, soit le nouvel apport du programme par rapport à ce qui se fait déjà; les stratégies qui facilitent la mise en œuvre; les défis associés à la mise en œuvre du programme; et les retombées perçues auprès des participants, c'est-à-dire les bénéfiques pour les enfants. Bien que toutes ces dimensions aient été examinées, les dimensions comprises dans la fidélité structurelle, ainsi que la fidélité et la qualité du contenu éducatif de la mise en œuvre du programme, sont les éléments centraux de l'étude des impacts du programme testé. Un dernier élément central de l'étude des impacts concerne la vérification empirique servant à établir de façon quantitative dans quelle mesure le programme testé se différencie des programmes offerts dans les garderies témoins. Aux fins des analyses d'impact, cette différenciation entre les deux groupes de garderies peut être conçue comme étant *le programme en garderie*. Si les comparaisons sont valides, la différenciation (observée et non observée) sera considérée comme responsable de l'effet observé du programme.

Les conclusions tirées de l'étude de la mise en œuvre sont présentées au tableau 11.1. On constate que la mise en œuvre du programme est relativement stable au cours des deux années du projet. De plus, le degré d'intégrité de la mise en œuvre s'avère élevé, tant au niveau des éléments structurels que du contenu éducatif, et ce, pour les deux cohortes. Quant à la mesure de différenciation, le programme en garderie testé se distingue des garderies témoins de par sa fidélité et sa qualité, montrant une différenciation par rapport à ce qui existe déjà dans la communauté. Ces résultats trouvent un appui dans l'étude menée par Japel et collègues (2005)

portant sur les normes de la qualité des garderies québécoises (voir tableau 11.1). Par ailleurs, la participation des enfants révèle un niveau de dosage suffisamment élevé pour pouvoir détecter l'effet du programme. Globalement, les éducatrices ont apprécié le nouveau programme. En particulier, elles ont réagi favorablement au cercle d'amis et à la routine imagée, alors qu'elles ont éprouvé de la difficulté à mettre en place l'aspect créativité du programme. Certains facilitateurs ont été identifiés, principalement au niveau du soutien des différents membres de l'équipe et du matériel fourni aux garderies. Les principaux défis associés à la mise en œuvre se rapportent à l'aménagement du temps relié à la prestation du programme, la difficulté d'aborder certains thèmes et le roulement des éducatrices dans les garderies du groupe programme.

En somme, l'analyse de la mise en œuvre du programme en garderie suggère que le nouveau programme fait preuve d'une bonne fidélité et qualité, et devrait donc avoir l'effet escompté sur les résultantes des enfants. L'analyse qualitative des retombées perçues corrobore d'ailleurs ce constat. En effet, les éducatrices et la formatrice ont perçu plusieurs retombées du programme, non seulement sur le comportement et le développement des enfants, mais également sur leurs pratiques en tant qu'éducatrices.

11.1.2. Les Ateliers familles

L'attention individuelle qu'une éducatrice en service de garde peut accorder à un enfant donné est limitée (Wasik, 2008), ce qui explique pourquoi certains ont fait valoir que les parents sont une ressource importante dans toute intervention ciblant les enfants (Reese et al., 2010). Les ateliers pour les parents ont été conçus pour mettre cette ressource au profit du développement de la langue et de la culture francophone.

L'évaluation de l'intégrité du programme des Ateliers familles s'est faite à partir de six dimensions, notamment : la fidélité du contenu, à savoir la proportion du programme abordé durant les ateliers; la qualité de l'animation des ateliers par les intervenantes; le dosage, soit le taux de participation des parents aux ateliers; la réaction des parents aux ateliers; les facilitateurs et les obstacles rencontrés lors de la mise en œuvre du programme; et les retombées perçues du programme par les parents et les éducatrices. Les paragraphes qui suivent présentent tour à tour les résultats des analyses portant sur chacune de ces dimensions.

Les résultats montrent que les Ateliers familles ont été livrés de façon fidèle au programme planifié. Quant à la qualité de l'animation des Ateliers familles, les observatrices et les parents ont généralement jugé favorablement la qualité de l'animation des Ateliers familles et les parents rapportent avoir tiré profit de leurs expériences. Le taux de participation des parents était acceptable, bien qu'il varie considérablement selon les familles. Notons que les analyses d'impact ont établi un lien entre la qualité de la mise en œuvre et l'assiduité, et les impacts observés chez les parents. Certains éléments ont facilité la participation des parents, notamment, le repas offert aux familles, l'atelier conjoint et la présence des éducatrices durant les ateliers. À l'inverse, les exigences associées à l'horaire des ateliers et la densité du contenu ont fait obstacle à la participation des parents. Quelques retombées perçues découlent des Ateliers familles, les principales étant la sensibilisation des parents vis-à-vis de leur rôle parental et l'adoption de nouvelles pratiques parentales. Ajoutons que les parents ont aussi été sensibilisés à l'importance de la langue française en milieu minoritaire francophone, ce qui s'est manifesté par une augmentation de l'utilisation du français dans le cadre d'activités familiales.⁹²

⁹² Il faut reconnaître, par contre, que cet effet n'était pas robuste lors de l'analyse par régression logistique.

Tableau 11.1 : Sommaire des résultats de l'analyse de la mise en œuvre

Volet Programme en garderie		Volet Ateliers familles	
Dimension	Conclusion	Dimension	Conclusion
Qualité structurelle	<ul style="list-style-type: none"> • Qualité est bonne dans les garderies programme et les garderies témoins (voir Japel et coll., 2005) • Qualité des garderies programme est significativement supérieure à celle observée dans les garderies témoins (voir les tableaux 5.3 et 9.5) 	Qualité de l'animation	<ul style="list-style-type: none"> • Évaluation généralement favorable par les parents • Séances ont tendance à dépasser le temps alloué
Fidélité structurelle	<ul style="list-style-type: none"> • Éléments du programme sont plus souvent présents dans les garderies programme, comparativement aux garderies témoin, et la différence est significative (voir les tableaux 5.4 et 9.6) 	Fidélité de contenu	<ul style="list-style-type: none"> • En moyenne, 98% du matériel obligatoire est abordé
Qualité éducative	<ul style="list-style-type: none"> • Qualité est bonne dans les garderies programme et les garderies témoins • Qualité des garderies programme est significativement supérieure à celle observée dans les garderies témoin (voir les tableaux 5.5 et 9.5) 		
Fidélité du contenu éducatif	<ul style="list-style-type: none"> • Fréquence des éléments du programme est plus élevée dans les garderies programme, comparativement aux garderies témoin; différence est significative à 8 mois (voir les tableaux 5.6 et 9.6) 		
Qualité des interactions	<ul style="list-style-type: none"> • Qualité est bonne dans les garderies programme et les garderies témoin • Aucune différence significative n'est observée entre les garderies programme et les garderies témoin (voir le tableau 9.5) 		
Le dosage	<ul style="list-style-type: none"> • Le dosage est comparable pour les deux groupes, se situant en moyenne à 27 - 30 heures (voir le tableau 5.7) 	Le dosage	<ul style="list-style-type: none"> • L'assiduité moyenne est de 63,5%, avec une grande variation
Les réactions des éducatrices	<ul style="list-style-type: none"> • Réactions positives au contenu du cercle d'amis et à la routine imagée • Aspect créatif du programme est plus difficile à mettre en place 	Réactions des parents	<ul style="list-style-type: none"> • Thèmes préférés identifiés pour 6 des 10 séances • Accès au matériel en français, discussions avec d'autres parents et activités avec leur enfant dans le cadre du volet conjoint sont appréciés
Facilitateurs et défis	<ul style="list-style-type: none"> • Facilitateurs : soutien de la direction, travail d'équipe, matériel fourni, la flexibilité du programme, suivis de la formatrice • Obstacles : gestion de temps, thèmes difficiles; roulement des éducatrices 	Facilitateurs et défis	<ul style="list-style-type: none"> • Facilitateurs : repas offerts, atelier conjoint et présence des éducatrices • Obstacles : horaire hebdomadaire trop exigeant, trop de matériel prévu pour le temps de séance donné

Volet Programme en garderie		Volet Ateliers familles	
Dimension	Conclusion	Dimension	Conclusion
Retombées perçues	<ul style="list-style-type: none"> • Meilleure discipline, les enfants sont plus calmes, autonomes et sécurisés • Meilleure préparation à la scolarisation en français et littératie • Éducatrices rapportent de plus grandes connaissances, ont changé leurs pratiques en tant qu'éducatrice (ex., plus centrées sur l'enfant) 	Retombées perçues par les parents	<ul style="list-style-type: none"> • Plus grande sensibilisation vis-à-vis de leur rôle dans le développement de leurs enfants et de l'importance du français en milieu minoritaire francophone • Acquisition de nouvelles connaissances, nouvelles pratiques parentales et une plus grande utilisation du français dans diverses activités faites avec leur enfant

11.1.3 Recommandations pour la mise en œuvre du programme à l'avenir

Le roulement du personnel peut être important dans certaines garderies. C'était notamment un problème important dans le cadre du présent projet dans la communauté d'Edmonton. La mise en œuvre du programme exige un investissement de ressources pour la formation des éducatrices, ce qui veut dire que le programme est mieux adapté pour des garderies avec une certaine stabilité de personnel et dans un contexte où la direction offre un appui actif au programme. En effet, l'exclusion d'Edmonton des analyses d'impact signifie que les résultats rapportés se généralisent seulement à une population de garderies où ces conditions favorables à la mise en œuvre du programme (au minimum) sont en place. L'impact du programme risque d'être atténué si le programme est mené à plus grande échelle sans prendre en considération ces aspects de la mise en œuvre.

11.2. L'IMPACT DU PROGRAMME TESTÉ

L'évaluation de ce programme à deux volets s'est faite au moyen d'une méthodologie quasi expérimentale avec groupes témoins non équivalents. La méthodologie prévoyait trois groupes expérimentaux : le groupe programme composé d'enfants inscrits dans la garderie francophone offrant le nouveau programme préscolaire; le groupe témoin en garderie comprenant les enfants inscrits dans une garderie francophone n'offrant pas le nouveau programme; et le groupe témoin hors garderie regroupant les enfants dont la garde de jour a lieu à la maison ou en garderie familiale non réglementée. Le groupe témoin en garderie vise à tenir compte de l'influence d'une garderie en installation sur le développement des enfants, ce qui constitue en soi un traitement. Le groupe témoin hors garderie vise à tenir compte de l'influence d'un milieu de garde informel sur le développement de l'enfant.

Les deux volets centraux de l'intervention étaient le nouveau programme en garderie et les Ateliers familles. Le but principal du premier volet était d'influencer *directement* les résultantes des enfants, tandis que celui du second volet était d'influencer *indirectement* les résultantes des enfants en modifiant les attitudes et les comportements des parents. Les analyses menées dans le cadre du projet Capacité d'apprentissage ne peuvent pas nettement distinguer l'effet d'un volet de l'autre. Un plan expérimental plus complexe aurait été nécessaire pour permettre de faire cette

distinction. En conséquence, les analyses principales comparant les groupes expérimentaux sont des tests de l'effet du programme à deux volets, soit l'*effet combiné* des deux volets du programme sur le développement des enfants.

11.2.1. Analyse par groupe de l'impact du programme

Les résultats principaux des analyses d'impact sont présentés au tableau 11.2. Ces analyses montrent que le programme testé a un effet positif sur la préparation à la scolarisation (les résultats exacts dépendent de quel groupe témoin sert de référence, voir le tableau 11.2 pour des détails). L'impact s'est manifesté par une croissance accélérée des compétences langagières et des compétences cognitives chez les enfants du groupe programme comparativement aux enfants des groupes témoins. Spécifiquement, on observe à l'entrée des enfants à la maternelle de quatre ans des effets moyens variant de 0,24 à 0,43 écart-type pour les échelles de la Communication, Conscience de soi et Vocabulaire expressif. La taille de ces effets équivaut à une croissance accélérée de 2,3 à 3,9 mois selon les données de Hill et ses collègues (2007). Les effets sont de taille comparable un an plus tard lors de l'entrée des enfants à la maternelle de cinq ans. Pour l'échantillon à l'étude, le programme testé semble donc avoir eu un impact positif non négligeable sur ces dimensions de la préparation à la scolarisation des enfants mesurées par l'ÉPE-AD. Les effets positifs et significatifs du programme sur les compétences langagières sont reproduits dans les analyses menées sur les échelles de vocabulaire: l'ÉVIP-R et l'EOWPVT. Des effets d'environ 0,20 écart-type sont obtenus avec ces outils, ce qui équivaut à un bond d'environ six mois de croissance du vocabulaire selon les normes de l'ÉVIP-R (Dunn et coll., 1993, Tableau 4.4, Forme A). Ces effets rapportés sont d'une importance pratique pour renforcer l'utilisation du français en contexte minoritaire et, ultimement, la préservation de la communauté francophone en situation minoritaire.

Tableau 11.2 : Sommaire des résultats des analyses d'impact du programme testé

	Première année					Deuxième année				
	Effet à 12 mois?	S'intensifie avec un dosage plus grand?	Expliqué par la qualité/fidélité?	Expliqué par l'effet sur les parents?	Plus grand bénéfice pour un sous-échantillon?	Effet à 24 mois?	S'intensifie avec un dosage plus grand?	Expliqué par la qualité/fidélité?	Expliqué par l'effet sur les parents?	Plus grand bénéfice pour un sous-échantillon?
G1 vs G2										
Communication	*	*	*		FE	*		*		FE
Conscience de soi	*	*				*		*		
Cognition	*	*	*		E		2	*	*	E
Physique	*					SO	SO	SO	SO	SO
Vocabulaire réceptif	*	*	*		FE	SO	SO	SO	SO	SO
Vocabulaire expressif	*	*	*		E	*		*		FE
ÉVIP-R (à 24 mois)	SO	SO	SO	SO	SO	*		*		FE

	Première année					Deuxième année				
	Effet à 12 mois?	S'intensifie avec un dosage plus grand?	Expliqué par la qualité/fidélité?	Expliqué par l'effet sur les parents?	Plus grand bénéfice pour un sous-échantillon?	Effet à 24 mois?	S'intensifie avec un dosage plus grand?	Expliqué par la qualité/fidélité?	Expliqué par l'effet sur les parents?	Plus grand bénéfice pour un sous-échantillon?
EOWPVT (à 20 mois)	SO	SO	SO	SO	SO	*1		*3		FE
G1 vs G3										
Communication	*	SO	SO		FE		SO	SO		FE
Conscience de soi		SO	SO			*	SO	SO		
Cognition	*	SO	SO		E		SO	SO	*	E
Physique		SO	SO			SO	SO	SO	SO	SO
Vocabulaire réceptif	*	SO	SO		FE	SO	SO	SO	SO	SO
Vocabulaire expressif		SO	SO		E	*	SO	SO		FE
EVIP-R (à 24 mois)	SO	SO	SO	SO	SO	*	SO	SO	SO	FE
EOWPVT (à 20 mois)	SO	SO	SO	SO	SO		SO	SO	SO	

Note : Le symbole * indique que l'hypothèse en question est vérifiée. ¹Effet détecté suite à l'inclusion du dosage dans le modèle.

²Effets différenciés du dosage à 16 et à 20 mois. ³L'effet n'est pas complètement expliqué. SO = sans objet ; FE = environnement familial à faible exposition au français ; E = environnement familial à forte exposition au français ; G1 = groupe programme, G2 = groupe témoin en garderie ; G3 = groupe témoin hors garderie.

11.2.2. Programme en garderie versus Ateliers familles

Les analyses principales comparant les groupes expérimentaux sont des tests de l'*effet combiné* des deux volets du programme sur le développement des enfants. Il existe néanmoins la possibilité d'effectuer des analyses donnant une idée de la contribution relative des deux volets à l'impact du programme. Ainsi, nous pouvons supposer que si l'impact du programme varie selon l'exposition de l'enfant au programme (le dosage) et si la fidélité/qualité de la prestation du programme semble expliquer l'effet du programme, c'est que le volet programme en garderie est principalement responsable de l'effet observé du programme. Si les changements observés chez les parents suite aux ateliers familles ne sont pas suivis de gains correspondants dans le développement de leurs enfants, c'est que le volet Atelier familles n'a pas d'effet observé sur les enfants du groupe programme.

Une première série d'analyses a cherché à démontrer un lien étroit entre l'impact du programme testé et le programme en garderie reçu par l'enfant. On constate que l'impact moyen du programme est basé sur un dosage moyen d'à peu près 27 heures par semaine en garderie. L'effet à la fin de la première année est accentué de 0,30 écart-type pour les enfants passant environ 37 heures par semaine en garderie. Par contre, l'effet à l'entrée de la maternelle de quatre ans est essentiellement nul pour les enfants passant seulement 17 heures par semaine en garderie. Ceci suggère fortement que les impacts observés dépendent du programme offert en garderie, du moins à l'entrée à la maternelle de quatre ans. Ensuite, nous avons mené une analyse

de médiation afin de vérifier si la mise en œuvre du programme en garderie était responsable, en tout ou en partie, des effets attribués au programme dans le cadre des analyses comparant les groupes à l'étude. On constate que la plus grande partie de l'effet du programme peut être attribuée avec confiance à la différenciation du programme en garderie testé.⁹³ En d'autres mots, le programme en garderie offert par les garderies programme est différent de ce qui est offert dans les garderies témoins, et cette différence explique l'effet du programme observé.

Concernant le programme Ateliers familles, nous concluons que les parents du groupe programme ont connu des bénéfices. Ces effets se dessinent dans les attitudes (c.-à-d. sentiment d'efficacité, connaissances au sujet du développement de l'enfant) et les comportements (c.-à-d. la fréquence des activités de littératie) rapportés par les parents. Ces effets sont aussi reliés au dosage et à la qualité de la mise en œuvre du programme, ce qui rehausse leur crédibilité en tant qu'effets réels du programme testé. Autrement dit, le programme Ateliers familles semble avoir eu un effet sur les parents du groupe programme.

Malheureusement, il y a peu de preuves empiriques que les changements dans les attitudes et les comportements des parents favorisent la préparation à la scolarisation des enfants. Seul le développement cognitif des enfants à l'entrée à la maternelle de cinq ans semble être affecté par le changement dans les attitudes des parents, changement attribué aux effets des Ateliers familles. Ainsi, les effets du programme Ateliers familles sont limités à une dimension de préparation à la scolarisation, du moins selon les variables considérées dans les analyses menées dans le présent rapport. Dans l'ensemble, le programme en garderie est une source plus crédible et plus importante de l'effet du programme testé.

Cela dit, le devis quasi expérimental utilisé ne nous permet pas d'écarter la possibilité que des ateliers d'alphabétisation familiale ne soient pas nécessaires pour obtenir des effets significatifs du programme en garderie. Certaines études montrent que l'effet d'une intervention ciblant les parents et les enfants est plus grand que l'effet d'une intervention ciblant uniquement l'un ou l'autre de ces deux groupes (Brooks-Gunn, Berlin, et Fuligni, 2000; Reese, et al., 2010). La contribution unique du volet d'alphabétisation familiale risque toutefois d'être très petite. L'effet de programmes d'alphabétisation familiale beaucoup plus ciblés est généralement positif, mais de petite taille, $d = 0,18$ (voir la méta-analyse de van Steensel, McElvany, Kurvers, et Herppich, 2011). Selon une méta-analyse effectuée par Sénéchal et Young (2008), les programmes d'alphabétisation familiale ont des effets plus importants lorsqu'on donne aux parents des stratégies concrètes à utiliser avec leurs enfants, et non uniquement des conseils généraux (pour d'autres exemples, voir Reese, et al., 2010). Dans le présent projet, il semble que le contenu du programme Ateliers familles ait été trop général pour avoir un effet important sur le développement. Précisons que les résultats obtenus retrouvent un soutien empirique dans la littérature qui semble appuyer la conclusion que le moteur principal des effets rapportés est le programme en garderie, sans écarter un rôle complémentaire pour les Ateliers familles.

En somme, le portrait dégagé de l'ensemble de ces résultats appuie la conclusion qu'une grande partie de l'effet « combiné » du programme est attribuable à l'effet du programme en

⁹³ Nous tenons à rappeler le lecteur que nous n'avons pas d'observations pré-intervention pour les programmes offerts en garderie. Il est possible que certaines différences au niveau de la programmation existent déjà avant la mise en œuvre du programme en garderie. C'est alors une information de nature corrélationnelle et non expérimentale. Ceci dit, dans certaines communautés les garderies programme et comparaison étaient sous l'égide d'un même pourvoyeur de services de garde, donc il y a de bonnes raisons de penser que les programmes auraient été assez similaires dans l'absence de l'intervention.

garderie. Les Ateliers familles apportent probablement aussi une contribution positive, mais ce lien, s'il existe, ne s'est pas systématiquement manifesté par le biais des variables observées.

11.2.3. Le profil linguistique des enfants

La littérature sur le bilinguisme fait une distinction claire entre les cas où l'acquisition d'une langue seconde profite au développement général d'un enfant et les cas où le développement est affecté (p. ex., Landry, Allard, et Deveau, 2009). La recherche distingue deux types de bilinguisme : additif et soustractif. La forme que prend le bilinguisme dépend de la réponse à la question suivante : La *langue maternelle* est-elle suffisamment développée pour soutenir l'acquisition d'une deuxième langue sans entraîner de retard dans le développement approprié à l'âge des habiletés cognitives ou des compétences dans la langue maternelle (Ball, 2010)? Dans l'affirmative, on parle de bilinguisme additif. Dans la négative, on parle plutôt de bilinguisme soustractif. Cette question est particulièrement importante lorsque la langue maternelle est la langue d'enseignement. C'est le cas, notamment, de plusieurs francophones en milieu minoritaire.

La condition requise pour le développement d'un bilinguisme additif n'est pas remplie pour beaucoup d'enfants francophones vivant en milieu minoritaire, une population chez qui on observe certains retards en littératie par rapport à leurs pairs (Bussière, et al., 2001; Chartier et al., 2011; Conseil canadien sur l'apprentissage, 2008; Education Quality and Accountability Office, 2009; Knighton, Brochu, et Gluszynski, 2010). Les causes sous-jacentes à cet écart de performance sont relativement bien comprises. La littérature sur le bilinguisme fait appel à un mécanisme assez clair: l'exposition à la langue. Pour que le bilinguisme soit additif, un seuil minimal d'exposition ou d'utilisation de la langue maternelle doit être dépassé (pour une revue, voir Pearson 2007). Pour diverses raisons (p. ex., motivationnelles, plus grande exposition à la langue de la majorité dans plusieurs milieux, Landry et al., 2009), le seuil minimal requis est plus élevé lorsque la langue maternelle est une langue minoritaire (Pearson, Fernandez, Lewedag, et Oller, 1997; Vihman, Lum, Thierry, Nakai, et Keren-Portnoy, 2006). En conséquence, les enfants francophones grandissant dans un milieu bilingue ont besoin d'un soutien spécial pour améliorer leurs chances d'atteindre un bilinguisme additif.

C'est dans cet esprit que nous avons étudié la possibilité que les enfants exposés à des langues autres que le français (habituellement l'anglais) profitent le plus du programme préscolaire testé. Cette question de recherche a fait l'objet d'une analyse approfondie dans le *Rapport des résultats de la première cohorte* (2014) et elle est reprise dans le présent rapport. Nous référons ici explicitement à l'effet combiné des deux volets du programme testé (en garderie et d'alphabétisation familiale) comme source d'influence environnementale sur le développement linguistique des enfants. Ainsi, nous anticipions une différenciation des effets selon le profil linguistique des enfants. Les enfants ayant une faible exposition au français à la maison sont plus à risque de développer un bilinguisme de type soustractif. Ce risque est accru par une faible exposition au français à l'extérieur de la maison parce que ces enfants vivent en milieu linguistique minoritaire.

Deux hypothèses furent émises : Les enfants à risque, donc ayant une faible exposition au français en milieu familial, devraient bénéficier davantage du programme par rapport à leurs compétences langagières; Les enfants moins à risque, donc ceux ayant une forte exposition au français en milieu familial, devraient en bénéficier davantage sur le plan de leur développement

cognitif. Ces hypothèses furent investiguées dans une série d'analyses où les effets du programme ont été estimés pour les enfants présentant un faible versus un fort degré d'exposition à la langue française au début de la prestation du programme (période d'évaluation de base). Les résultantes à l'étude comprenaient celles captant les compétences langagières (c.-à-d. les mesures de Communication, Vocabulaire expressif, l'ÉVIP-R et l'EOWPVT) et celles ciblant des compétences cognitives plus générales (c.-à-d. les mesures de Cognition). L'échelle de Conscience de soi ne fut pas retenue puisqu'elle ciblait un ensemble de compétences hétérogènes.

Résultantes ciblant les compétences langagières

Tel qu'anticipé, l'impact du programme testé sur la Communication est plus important pour les enfants ayant une faible exposition au français, et ce, à chaque évaluation sauf la dernière. De façon similaire, l'impact du programme testé sur le Vocabulaire expressif relatif au groupe témoin en garderie est plus prononcé pour les enfants ayant une faible exposition au français. Quand on oppose le groupe programme au groupe témoin en garderie, les analyses sur les échelles de vocabulaire de l'ÉVIP-R et de l'EOWPVT confirment un effet du programme testé chez les enfants ayant une faible exposition au français. Le patron est moins net quand le groupe témoin hors garderie sert de référence, car dans ce cas, un effet du programme est observé uniquement pour l'ÉVIP-R. Aucune différence entre les groupes expérimentaux n'est observée sur ces résultantes pour les enfants ayant une forte exposition au français à la maison.

Ce patron de résultats tend à confirmer nos hypothèses. Notons toutefois que les intervalles de confiance pour les effets estimés d'après les deux types de foyer (exposition faible, exposition forte) se chevauchaient habituellement, indiquant que même si la direction des différences était conforme à nos hypothèses, la taille de l'effet du programme est généralement comparable peu importe le type de foyer. Ceci dit, le patron de résultats rapportés dans le cadre du présent projet reproduit les résultats observés dans d'autres recherches portant sur ce thème (p. ex., Maltais, 2007). De plus, dans certains cas, l'effet du programme était bel et bien significativement plus fort pour les enfants à faible exposition (voir par exemple l'effet à 12 mois pour l'échelle de communication comparativement au groupe témoin en garderie).

Résultantes ciblant d'autres compétences cognitives

Le patron de résultats obtenu avec l'échelle Cognition de l'ÉPE-AD apporte un soutien empirique plus solide à nos hypothèses. L'effet du programme est significativement plus grand pour les enfants ayant une forte exposition au français tout au long de la première année de la prestation du programme. Par rapport à leurs pairs dans les groupes témoins, ces enfants sont avantagés au niveau de leur développement d'habiletés cognitives. L'avantage observé chez ces enfants, si l'on considère la taille de cet effet, persiste jusqu'à l'entrée à la maternelle de cinq ans. À ce moment, le chevauchement des intervalles de confiance entre les groupes expérimentaux indique un effet programme de taille comparable pour les deux types de foyer (exposition faible, exposition forte) lors de cette période d'évaluation, et ce, peu importe le groupe témoin utilisé en tant que référence pour estimer l'effet du programme.

L'hypothèse est-elle vérifiée?

Ensemble, ces résultats suggèrent que les compétences linguistiques de base facilitent le développement des habiletés cognitives (p. ex., littératie). Cette explication trouve un appui

après des théories développementales sur l'acquisition des compétences (Bloom, 1976; Vygotsky, 1978) et avec l'importance que revêt la maîtrise de la langue pour le succès scolaire (Cummins, 1979; Doherty, 1997; Hindman, Skibbe, Miller, et Zimmerman, 2010). Elle est aussi congruente avec les constats de Maltais (2007), selon lesquels l'effet d'un programme de prématernelle à temps plein sur le développement linguistique dépend d'une faible exposition à la langue, tandis que les effets de ce programme sur le développement cognitif (c.-à-d. gains en habiletés de lecture) sont observés chez les enfants dont l'exposition au français est forte, enfants pour qui les effets sont plus importants.

En somme, le haut degré de fidélité et de qualité du programme testé a mené à une transmission plus efficace de compétences et de connaissances. Tous les enfants tirent profit du programme bien que les avantages reposent sur leur niveau de préparation à apprendre certains types de compétences dans un environnement francophone. Ces résultats ont des implications quant aux attentes face à l'impact du programme. Il semble que tous les enfants tirent un bénéfice du programme par rapport à leur préparation à la scolarisation bien que le type de bénéfice retiré varie selon les caractéristiques linguistiques de la clientèle. La crédibilité des résultats repose, en partie, sur des considérations théoriques et les résultats provenant d'autres recherches (Maltais, 2007), puisque la différenciation des effets du programme selon le profil linguistique n'était pas aussi prononcée que ce à quoi on s'attendait.

Implications pour la généralisation des résultats

Le résultat d'analyses comparatives démontre que l'échantillon à l'étude se distingue de l'échantillon de l'EVML0 principalement sur les caractéristiques linguistiques. Nous pensons toutefois que l'écart existant entre le profil linguistique du présent échantillon et celui de la population générale de francophones en situation minoritaire ne présente aucun obstacle à la généralisation des résultats. Si la population est comparativement plus exposée au français à la maison, alors l'impact sera observé principalement sur le développement de compétences cognitives. Si, au contraire, la population est comparativement moins exposée au français, un plus grand impact sur les compétences langagières (p. ex., le vocabulaire) est attendu. Si on doute de la validité des analyses indiquant une différenciation selon le profil linguistique, alors un effet positif généralisé sur la préparation à la scolarisation est à tout le moins attendu sur la base des résultats principaux comparant les groupes expérimentaux. Il n'y a donc aucune raison de penser que l'impact du programme ne serait pas reproduit avec un échantillon différent d'enfants francophones vivant en milieu minoritaire.

11.2.4. Limites des résultats et prochains travaux

L'utilisation d'un devis quasi expérimental à groupes non équivalents s'accompagne généralement de certaines réserves vis-à-vis des résultats découlant de la recherche. Plusieurs stratégies ont été utilisées pour assurer la validité des résultats. Une des plus grandes menaces à la validité interne est que les constats résultent d'un biais lié à la composition des groupes plutôt qu'à l'effet du programme. La possibilité de la présence d'une telle source de biais est diminuée en raison de l'attention particulière qui a été accordée au recrutement de participants pour les groupes témoins possédant un profil sociodémographique similaire à ceux du groupe programme (p. ex., le niveau socio-économique) et vivant dans le même arrondissement, assurant ainsi qu'ils ont accès aux mêmes ressources et services francophones que les participants du groupe

programme.⁹⁴ Ainsi, un premier contrôle sur la composition des groupes a été fait lors de l'inscription des familles au projet, notamment au niveau de l'emplacement du domicile des membres potentiels des groupes témoins. Un deuxième contrôle fut l'utilisation des mesures pré-intervention. C'est prochaine section.

Une deuxième menace à la validité interne de l'étude est que les évaluatrices, les éducatrices et les parents étaient conscients de qui était traité et qui ne l'était pas. Cette source de biais est inévitable quand on considère que le statut des garderies était connu dans la communauté (garderie offrant le programme, garderie n'offrant pas le programme) et, par association, le statut des enfants qui fréquentaient les garderies. Cette menace n'est pas aussi importante qu'on pourrait le croire de prime abord, puisque ce sont les garderies et non les enfants qui étaient recrutées et assignées aux groupes expérimentaux. Il est néanmoins difficile de répondre à la critique que les résultats sont dus à un biais lié à cette connaissance. Par contre, il est difficile d'imaginer que ces sources potentielles de biais auraient, individuellement ou en combinaison, produit le patron de résultats obtenus. Personne ne connaissait le niveau relatif de fidélité et de qualité du programme offert dans les garderies programme ou témoins. Personne ne connaissait les hypothèses de l'effet modérateur du dosage ou du profil linguistique. Par conséquent, il est peu probable qu'un biais quelconque (p. ex., dans la composition des groupes, de la part des évaluatrices) aurait donné des réponses congruentes à nos hypothèses de recherche.

Une troisième critique possible est la modeste taille de l'échantillon sur lequel reposent les constats ($N \approx 330$) alors que la robustesse des analyses statistiques atteint son maximum avec des échantillons très grands ($N > 1000$). Plusieurs stratégies ont été utilisées pour vérifier la robustesse des résultats notamment au niveau des analyses. Les analyses comparant les groupes sont les plus robustes à cet égard, reposant sur sept périodes de collecte de données auprès de plus de 300 participants (2 100 observations en tout). Les analyses plus fines comptent 100 participants au sein de chaque groupe (p. ex., l'analyse par dosage, qualité de la prestation du programme, profil linguistique). L'erreur d'échantillonnage associée à ces analyses complémentaires est alors comparativement plus importante. C'est pour cette raison que nous avons privilégié une interprétation des résultats reposant sur l'ensemble des résultats et non sur un résultat en particulier. Au niveau communautaire, il est important de noter que les résultats peuvent être généralisés uniquement aux communautés étudiées ou à des communautés semblables.⁹⁵

Enfin, les analyses d'impact ont porté uniquement sur des indicateurs de la préparation à la scolarisation et des indicateurs du vocabulaire français. Les indicateurs de la préparation à la scolarisation sont généraux de par leur nature, ce qui peut obscurcir des effets importants spécifiques à une dimension. Dans les rapports à venir, nous examinerons l'impact du programme testé sur des mesures plus ciblées des compétences cognitives et langagières, prédisant la réussite scolaire des enfants. Ces analyses serviront à élucider l'apport du programme testé pour les enfants et permettront de tracer l'impact du programme à court et à moyen terme.

⁹⁴ Le lecteur intéressé à en apprendre plus est invité à lire le Rapport sur les plans de travail et de méthodologie révisés remis à RHDCC le 30 mars 2007.

⁹⁵ Le facteur 'communauté' était considéré un facteur fixe lors des analyses d'impact en raison du nombre restreint de communautés.

11.3. CONCLUSION

La présente recherche comportait deux objectifs : mener une étude de la mise en œuvre d'une nouvelle intervention à deux volets, et en évaluer l'impact sur de jeunes enfants de minorité francophone et leurs parents. Les principaux constats de l'étude de la mise en œuvre sont que la prestation du programme se caractérise par un degré élevé d'intégrité tant au niveau des éléments structurels que de la fidélité et la qualité du contenu éducatif durant les deux années du programme. Ces constats sont pertinents pour les deux cohortes. Ces dimensions de la mise en œuvre du programme sont des éléments centraux de l'étude des impacts du programme testé. Un dernier élément central de l'étude d'impact concerne la vérification empirique établissant de façon quantitative que le programme testé se distingue des programmes existants offerts dans les garderies témoin de par sa fidélité et sa qualité. Aux fins des analyses d'impact, cette différenciation du programme des deux groupes de garderies se conçoit comme un appui empirique voulant que *le programme* en garderie soit considéré comme responsable de l'effet observé du programme.

L'étude d'impact a, quant à elle, révélée *des effets du programme positifs tant sur les résultantes concernant les enfants que celles des parents*. En ce qui concerne les résultantes touchant les enfants, les effets positifs se sont manifestés comme des gains développementaux supérieurs relativement aux groupes témoins en garderie et hors garderie. Généralement, les effets ont émergé tôt et sont demeurés constants au cours des deux années examinées dans le présent rapport. La nature de ces gains dépendait de l'exposition de l'enfant au français au début du projet. Les enfants ayant l'exposition la plus faible ont profité le plus du programme en matière de développement linguistique (p. ex., vocabulaire expressif), tandis que les enfants présentant l'exposition la plus forte ont connu des gains pour ce qui est de leur développement cognitif. Le volet du programme s'adressant aux parents a connu un succès partiel. Les parents ont bel et bien connu des changements positifs par rapport à la fréquence des activités de littératie, leurs connaissances et leur sentiment d'auto-efficacité. Toutefois, il est moins évident que les effets du programme Ateliers familles chez les parents influencent, à leur tour, le développement des enfants. Si le programme des Ateliers familles est responsable d'une partie de l'impact du programme sur les enfants, cet effet est limité au développement cognitif, du moins selon les variables considérées dans les analyses menées jusqu'à présent. À ce stade, le patron de résultats suggère que le programme en garderie est la source dominante de l'effet du programme sur les enfants.

Nous pouvons donc conclure que le programme testé a eu un impact modeste sur la préparation à la scolarisation des enfants de minorité francophone. La taille de l'effet observé équivaut à un gain de quelques mois de développement. Dans les prochains rapports, nous aborderons une nouvelle question de recherche: *Le nouveau programme préscolaire permet-il aux enfants francophones grandissant en milieu minoritaire d'être mieux outillés pour réussir dans les tâches essentielles à la réussite scolaire que sont la lecture et les mathématiques?* Cette deuxième question de recherche se penche sur la période de scolarisation englobant la première et la deuxième année scolaire, alors que les enfants sont âgés de 6 et 7 ans. Dans cette deuxième phase du projet Capacité d'apprentissage, il s'agira d'établir si une meilleure préparation à la scolarisation et une plus grande maîtrise de la langue française augmenteront les chances d'une réussite scolaire. Les réponses à cette question proviendront de l'étude des impacts du

programme sur des précurseurs plus immédiats de la réussite scolaire tels que la conscience phonologique, la connaissance alphabétique et le vocabulaire (Lonigan, 2008).

Références

- Abadie, A. (2005). Semiparametric Difference-in-Differences estimators. *Review of Economic Studies*, 72(1), 1-19.
- Adams, G. R., & Ryan, B. A. (2000). *A longitudinal analysis of family relationships and children's school achievement in one- and two-parent families*. Hull, QC : Human Resources Development Canada.
- Allen, M. (2004). Profiles of minority language schools in five Canadian provinces. *Education Quarterly Review*, 9(4), 17-63.
- Arnold, D. H., Zeljo, A., Doctoroff, G. L., & Ortiz, C. (2008). Parent involvement in preschool: Predictors and the relation of involvement to preliteracy development. *School Psychology Review*, 37(1), 74-90.
- Ball, J. (2010). *Enhancing learning of children from diverse language backgrounds: Mother tongue-based bilingual or multilingual education in early childhood and early primary school years*. Washington, D.C., DC: UNESCO. Disponible en ligne au : <http://www.ecdip.org/docs/pdf/UNESCO%20Mother-tongue%20based%20EY%202010.pdf>.
- Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51(6), 1173-1182.
- Becker, G. S. (1993). Nobel lecture: The economic way of looking at behavior. *The Journal of Political Economy*, 101(3), 385-409.
- Benoît, S. (s.d.). *Questionnaire utilisés par CIRCEM dans l'étude de l'impact de l'alphabétisation familiale sur les familles vivant en milieu minoritaire*. Vanier, ON : Coalition francophone pour l'alphabétisation et la formation de base en Ontario.
- Berger, L. M., Paxson, C., & Waldfogel, J. (2005). *Income and child development*. Princeton, NJ: Princeton University, Center for Research on Child Wellbeing.
- Bertrand, J. (2007). Programmes d'enseignement efficaces au préscolaire. Commentaires sur Kagan et Kauerz et sur Schweinhart. Dans R. E. Tremblay, R. G. Barr, & R. de V. Peters (éds.), *Encyclopédie sur le développement des jeunes enfants* [sur Internet] (pp.1-7). Montréal, QC : Centre d'excellence pour le développement des jeunes enfants. Disponible en ligne au : <http://www.enfant-encyclopedie.com/Pages/PDF/BertrandFRxp.pdf>
- Bertrand, M., Duflos, E., & Mullianathan, S. (2004). How much should we trust Difference-in-Differences estimates?. *The Quarterly Journal of Economics*, 119(1), 249-275.
- Bérubé, A., Legault, L., Janisse, L., Carson, R., Saucier, M.-A. et Lefebvre, L. (2014). *Capacité d'apprentissage dans les communautés francophones en situation minoritaire : Rapport de mise en œuvre du projet*. Ottawa (Ontario) : Société de recherche sociale appliquée.

- Bloom, B. S. (1976). *Human characteristics and school learning*. Columbus, OH : McGraw-Hill.
- Bornstein, L., & Bornstein, M. H. (2007). Pratiques parentales et développement social de l'enfant. Dans R. E. Tremblay, R. de V. Peters, M. Boivin, & R. G. Barr (éds.), *Encyclopédie sur le développement des jeunes enfants* [sur Internet] (pp. 1-2). Montréal, QC : Centre d'excellence pour le développement des jeunes enfants. Disponible en ligne au : <http://www.enfant-encyclopedie.com/Pages/PDF/BornsteinFRxp.pdf>.
- Bredenkamp, S. & Copple, C. (Eds.). (1999). *Developmentally appropriate practice in early childhood programs*. Washington, D.C., DC: National Association for the Education of Young Children.
- Breton, R. (1964). Institutional completeness of ethnic communities and the personal relations of immigrants. *The American Journal of Sociology*, 70(2), 193-205.
- Brigham, J. C. (1991). Attitudes and persuasion. *Social psychology* (2nd ed., pp. 132-224). New York, NY: Harper Collins.
- Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design*. Cambridge, MA: Harvard University Press.
- Bronfenbrenner, U., & Morris, P. A. (1998). The ecology of developmental processes. Dans W. Damon & R. M. Lerner (éds.), *Handbook of child psychology: Theoretical models of human development* (5th ed., Vol. 1, pp. 993-1028). New York, NY: Wiley.
- Brooks-Gunn, J., Berlin, L. J., & Fuligni, A. S. (2000). Early childhood intervention programs: What about the family?. Dans J. P. Shonkoff & S. J. Meisels (éds.), *Handbook of early childhood intervention* (2nd ed., pp. 549-588). New York, NY: Cambridge University Press.
- Burchinal, M. R., Roberts, J. E., Riggins, R., Jr., Zeisel, S. A., Neebe, E., & Bryant, D. (2000). Relating quality of center-based child care to early cognitive and language development longitudinally. *Child Development*, 71(2), 339-357.
- Bussière, P., Cartwright, F., Crocker, R., Ma, X., Oderkirk, J., & Zhang, X. (2001). *À la hauteur: la performance des jeunes du Canada en lecture, en mathématiques et en sciences*. Ottawa, ON: Développement des ressources humaines Canada, Statistique Canada, Conseil des ministres de l'Éducation (Canada).
- Carroll, C., Patterson, M., Wood, S., Booth, A., Rick, J., & Balain, S. (2007). A conceptual framework for implementation fidelity. *Implementation Science*, 2(40), 1-9. Disponible en ligne au : <http://www.implementationscience.com/content/pdf/1748-5908-2-40.pdf>.
- Charlebois, P., Brendgen, M., Vitaro, F., Normandeau, S., & Bourdreau, J.-F. (2004). Examining dosage effects on prevention outcomes: Results from a multi-modal longitudinal preventive intervention for young disruptive boys. *Journal of School Psychology*, 42(3), 201-220.
- Charte canadienne des droits et libertés*, s. 23, Part I de la *Loi constitutionnelle de 1982*, étant l'annexe B de la *Loi de 1982 sur le Canada* (U.K.), 1982, c.11.

- Chartier, M. J., Dumaine, J., & Sabourin, E. (2011). Vivre en français à la petite enfance et apprendre à l'école française en milieu minoritaire, y a-t-il un lien? *Cahiers franco-canadiens de l'Ouest*, 23(1-2), 3-61.
- Chen, H. T. (Ed.). (2005). *Practice program evaluation: Assessing and improving planning, implementation, and effectiveness*. Thousand Oaks, CA: Sage Publications.
- Cleveland, G., Corter, C., Pelletier, J., Colley, S., Bertrand, J., Ontario Institute for Studies in Education/University of Toronto, & Jamieson, J. (2006). *A review of the state of the field of early childhood learning and development in child care, kindergarten and family support programs*. Toronto, ON: Atkinson Centre for Society and Child Development.
- Cohen, J. (1988). *Statistical power analysis for the behavioural sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioural sciences* (3rd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Commission nationale des parents francophones. (2005). *Le meilleur est...avenir! La petite enfance dans les communautés francophones en milieu minoritaire au Canada : une analyse transformative de la situation*. Ottawa, ON : Auteur. Disponible en ligne au : http://cnpf.ca/documents/SCAN_FRANCAIS_2005.pdf.
- Committee for Economic Development Research and Policy Committee. (1993). *Why child care matters: Preparing young children for a more productive America. A statement by the Research and Policy Committee of the Committee for Economic Development*. New York, NY: Committee for Economic Development.
- Conduct Problems Prevention Research Group. (1999). Initial impact of the Fast Track prevention trial for conduct problems: I. The high-risk sample. *Journal of Consulting and Clinical Psychology*, 67(5), 631-647.
- Connor, S., & Brink, S. (1999). *Understanding the Early Years: Community impacts on child development*. Hull, QC: Applied Research Branch, Strategic Policy, Human Resources Development Canada. Disponible en ligne au : <http://www.RHDCC.gc.ca/eng/cs/sp/sdc/pkrf/publications/nlscy/uey/1999-000092/w-99-6e.pdf>.
- Connors, L. J., & Epstein, J. L. (1995). Parent and school partnerships. Dans M. H. Bornstein (éd.), *Handbook of parenting. Volume 4: Applied and practical parenting* (pp. 437-458). Mahwah, NJ: Lawrence Erlbaum Associates.
- Conseil canadien sur l'apprentissage. (2007). *Rapport sur l'apprentissage au Canada : rapport sur l'état de l'apprentissage chez les jeunes enfants*. Ottawa, ON : Auteur.
- Conseil canadien sur l'apprentissage. (2008). *Literacy among Canadian students in minority language contexts*. Rapport non publié.
- Corbeil, J.-P., & Blaser, C. (2007). *Le portrait linguistique en évolution : Recensement de 2006*. Ottawa, ON: Statistique Canada. Disponible en ligne au : <http://www12.statcan.ca/census-recensement/2006/as-sa/97-555/pdf/97-555-XIE2006001-fra.pdf>.

- Corbeil, J.-P., Grenier, C., & Lafrenière, S. (2007). *Les minorités prennent la parole : résultats des l'Enquête sur la vitalité des minorités de langue officielle 2006*. Disponible en ligne au : <http://www.statcan.gc.ca/pub/91-548-x/91-548-x2007001-fra.pdf>.
- Corter, C., & Pelletier, J. (2005). Parents and community involvement in schools: Policy panacea or pandemic? Dans N. Bascia, A. Cumming, A. Datnow, K. Leithwood, & D. Livingstone (éds.), *International Handbook of Educational Policy* (pp. 295-327). Great Britain: Springer.
- Cummins, J. (1979). *Cognitive/academic language proficiency, linguistic interdependence, the optimum age question and some other matters. Working papers on bilingualism, No. 19*. Toronto, ON: The Ontario Institute for Studies in Education.
- Dane, A. V., & Schneider, B. H. (1998). Program integrity in primary and early secondary prevention: Are implementation effects out of control?. *Clinical Psychology Review*, 18(1), 23-45.
- Davis-Kean, P. E. (2005). The influence of parent education and family income on child achievement: The indirect role of parental expectations and the home environment. *Journal of Family Psychology*, 19(2), 294-304.
- Dearing, E., Berry, D., & Zaslow, M. (2006). Poverty during early childhood. Dans K. McCartney, & D. Phillips (éds.), *The Blackwell handbook of early childhood development* (pp. 399-423). Malden, MA: Blackwell Publishing.
- Desrosiers, H., & Ducharme, A. (2006). *Commencer l'école du bon pied : facteurs associés à l'acquisition du vocabulaire à la fin de la maternelle*, 4(1). Québec, QC : Institut de la statistique Québec. Disponible en ligne au : http://www.jesuisjeserai.stat.gouv.qc.ca/pdf/publications/feuille/fascicule_ecole_bon_pied.pdf.
- Dionne, A.-M. (2007). *Incidence des pratiques de littératie familiale sur les compétences en lecture de l'enfant dans les familles immigrantes*. Présenté au l'Actualité de la Recherche en Éducation et en Formation, Stasbourg, France.
- Doherty, G. (1997). *Zero to six: The basis for school readiness*. Hull, QC: Applied Research Branch of Strategic Policy, Human Resources Development Canada. Disponible en ligne au : <http://www.RHDCC.gc.ca/eng/cs/sp/sdc/pkrf/publications/1997-002557/SP-241-02-01E.pdf>.
- Donald, S. G., & Lang, K. (2007). Inference with Difference-in-Differences and other panel data. *The Review of Economics and Statistics*, 89(2), 221-233.
- Duncan, G. J., & Brooks-Gunn, J. (Eds). (1997). *Consequences of growing up poor*. New York, NY: Russell Sage Foundation.
- Dunn, L. M., Thériault-Whalen, C. M., & Dunn, L. M. (1993). *Échelle de Vocabulaire en Images Peabody adaptation française du Peabody Picture Vocabulary Test-Revised: Manuel pour les Formes A et B*. Toronto, ON : Pearson Canada Assessment Inc.
- Durlak, J. A., & DuPre, E. P. (2008). Implementation matters: A review on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology*, 41(3-4), 327-350.

- Education Quality and Accountability Office. (2009). *Ontario student achievement: EQAO's provincial report on the results of the 2008-2009 assessments of reading, writing and mathematics, primary division (grades 1-3) and junior division (grades 4-6), and the grade 9 assessment of mathematics. English-language students*. Toronto, ON. Disponible en ligne au : http://www.eqao.com/pdf_e/09/Cprp_PJ9e_0909_web.pdf.
- Engle, P. L., Black, M. M., Behrman, J. R., Cabral de Mello, M., Gertler, P., Kapiriri, L., ... International Child Development Steering Group. (2007). Child development in developing countries 3: Strategies to avoid the loss of developmental potential in more than 200 million children in the developing world. *The Lancet*, 369, 229-242. Récupérée de : http://www.who.int/child_adolescent_health/documents/pdfs/lancet_child_dev_series_paper3.pdf. Consulté le 23 mars 2010.
- Erikson, E. H. (1994). *Identity and the life cycle*. New York, NY: W. W. Norton & Company.
- Forgues, É., & Landry, R. (2006). *Définitions de la francophonie en situation minoritaire : analyse de différentes définitions statistiques et leurs conséquences*. Ottawa, ON : Institut Canadien de recherche sur les minorités linguistiques.
- Frazier, P. A., Tix, A. P., & Barron, K. E. (2004). Testing moderator and mediator effects in counseling psychology research. *Journal of Counseling Psychology*, 51(1), 115-134.
- Gardner, M. (1979). *Expressive One-Word Picture Vocabulary Test: Manual and forms*. Novato, CA: Academic Therapy Publications.
- Gauthier St-Onge, D. (s.d.). *Questionnaire utilisés dans l'Évaluation du programme alpha District 3, Edmundston, Nouveau-Brunswick*.
- Gilbert, A. (2003). *La petite enfance : porte d'entrée à l'école de langue française : une vision nationale. Rapport final de la recherche*. Ottawa, ON : Centre interdisciplinaire de recherche sur la citoyenneté et les minorités, Université d'Ottawa en partenariat avec la Fédération canadienne des enseignantes et des enseignants. Disponible en ligne au : <http://www.ctf-fce.ca/documents/Resources/Francaise/petiteenfance/Visionnationale.pdf>.
- Guimond, L. (2003). *Revue des travaux sur la vitalité des communautés francophones du Canada*. Ottawa, ON : Université d'Ottawa.
- Hargrave, A. C., & Sénéchal, M. (2000). A book reading intervention with preschool children who have limited vocabularies: The benefits of regular reading and dialogic reading. *Early Childhood Research Quarterly*, 15(1), 75-90.
- Harms, T., & Clifford, R. M. (1990). *Infant/Toddler Environment Rating Scale*. New York, NY: Teachers College Press.
- Harms, T., Clifford, R. M., & Cryer, D. (1998). *Early Childhood Environment Rating Scale – Revised Edition*. New York, NY: Teachers College Press.
- Harms, T., Clifford, R. M., & Cryer, D. (2010). *Échelle d'évaluation de l'environnement préscolaire – révisée*. Traduction de M. Baillargeon et H. Larouche. Québec, QC : Presses de l'Université du Québec.
- Haveman, R., & Wolfe, B. (1995). The determinants of children's attainments: A review of methods and results. *Journal of Economic Literature*, 33(4), 1829-1878.

- Hayes, A. F., & Cai, L. (2007). Using heteroskedasticity-consistent standard error estimators in OLS regression: An introduction and software implementation. *Behavior Research Methods*, 39(4), 709-722.
- Hertzman, C., & Kohen, D. (2003). Neighbourhoods matter for child development. *Transition*, 33, 3-5. Disponible en ligne au : http://www.vifamily.ca/media/node/339/attachments/Neighbourhoods_Matter_for_Child_development_Hertzman_Kohen.pdf.
- Hill, C. J., Bloom, H. S., Black, A. R., & Lipsey, M. W. (2008). Empirical benchmarks for interpreting effect sizes in research. *Child Development Perspectives*, 2(3), 172-177.
- Hindman, A. H., Skibbe, L. E., Miller, A., & Zimmerman, M. (2010). Ecological contexts and early learning: Contributions of child, family, and classroom factors during Head Start, to literacy and mathematics growth through first grade. *Early Childhood Research Quarterly*, 25(2), 235-250.
- Howard, G. S. (1980). Response-shift bias: A problem in evaluating interventions with pre/post self-reports. *Evaluation Review*, 4(1), 93-106.
- Howes, C., Burchinal, M., Pianta, R., Bryant, D., Early, D., Clifford, R., & Barbarin, O. (2008). Ready to learn? Children's pre-academic achievement in pre-kindergarten programs. *Early Childhood Research Quarterly*, 23(1), 27-50.
- Hox, J. (2002). *Multilevel analysis: Techniques and applications* (1st ed.). New York, NY: Routledge.
- Institute de la statistique Québec. (2003). *En 2002...j'avais 5 ans! Questionnaire informatisé rempli par l'intervieweuse (QIRI) – E6*. Québec, QC : Auteur.
- Izzo, C. V., Weissberg, R. P., Kaspro, W. J., & Fendrich, M. (1999). A longitudinal assessment of teacher perceptions of parent involvement in children's education and school performance. *American Journal of Community Psychology*, 27(6), 817-839.
- Jamieson, J. (2003). Analysis of covariance (ANCOVA) with difference scores. *International Journal of Psychophysiology*, 52(3), 277-283.
- Japel, C., Tremblay, R. E., & Côté, S. (2005). La qualité, ça compte! Résultats de l'Étude longitudinale du développement des enfants du Québec concernant la qualité des services de garde. *Choix !RPP*, 11(4), 1-46.
- Jordan, G. E., Snow, C. E., & Porche, M. V. (2000). Project EASE: The effect of a family literacy project on kindergarten students' early literacy skills. *Reading Research Quarterly*, 35(4), 524-546.
- Kane, T. J. (2004). *The impact of after-school programs: Interpreting the results of four recent evaluations*. New York, NY: William T. Grant Foundation. Disponible en ligne au : http://www.wtgrantfoundation.org/publications_and_reports/browse_reports/kane_working_paper.
- Klebanov, P. K., Brooks-Gunn, J., & Duncan, G. J. (1994). Does neighbourhood and family poverty affect mother's parenting, mental health, and social support?. *Journal of Marriage and Family*, 56(2), 441-455.

- Knighton, T., Brochu, P., & Gluszynski, T. (2010). *Measuring up: Canadian results of the OECD PISA study. The performance of Canada's youth in reading, mathematics, and science. 2009 first results for Canadians aged 15*. Disponible en ligne au : <http://www.cmec.ca/Publications/Lists/Publications/Attachments/254/PISA2009-can-report.pdf>.
- Krull, J. L., & MacKinnon, D. P. (1999). Multilevel mediation modeling in group-based intervention studies. *Evaluation Review*, 23(4), 418-444.
- Lafrance, F. (1993). Les conditions sociolinguistiques de l'enseignement du français minoritaire et leurs conséquences sur la pédagogie du français langue maternelle en Ontario. *Revue du Nouvel-Ontario*, 15, 11-34.
- Lamb, T., & Tschillard, R. (2003). *An underutilization design in applied research: The retrospective pré-test*. Présenté au Annual meeting of the American Sociological Association.
- Lamb, T. A., & Tschillard, R. (2005). Evaluating learning in professional development workshops: Using the retrospective pre-test. *Journal of Research in Professional Learning*. Récupérée au : http://www.mdecgateway.org/olms/data/resource/6794/Evaluating%20Learning%20in%20OPD%20Workshops_OST-PD.pdf. Consulté février 2008.
- Landry, R. (1994). Diagnostic sur la vitalité de la communauté acadienne du Nouveau-Brunswick. *Égalité*, 36, 11-39.
- Landry, R. (2010). *Petite enfance et autonomie culturelle : Là ou le nombre le justifie... V*. Rapport de recherche réalisé pour la Commission nationale des parents francophones. Moncton, NB : Institut canadien de recherche sur les minorités linguistiques.
- Landry, R., & Allard, R. (1997). L'exogamie et le maintien de deux langues et de deux cultures : le rôle de la francité familioscolaire. *Revue des sciences de l'éducation*, 23(3), 561-592.
- Landry, R., & Allard, R. (2000). Langue de scolarisation et développement bilingue : le cas des acadiens et francophones de la Nouvelle-Écosse, Canada. *DiversCité Langues*, V. Récupérée au : <http://www.telug.quebec.ca/diverscite/entree.htm>. Consulté le 23 avril, 2009.
- Landry, R., Allard, R., & Deveau, K. (2007). Bilingual schooling of the Canadian francophone minority: A cultural autonomy model. *International Journal of the Sociology of Language*, 185, 133-162.
- Landry, R., Allard, R., & Deveau, K. (2009). Self-determination and bilingualism. *Theory and Research in Education*, 7(2), 203-213.
- Laperrière, A. (1997). Les critères de scientificité des méthodes qualitatives. Dans J. Poupart, J.-P. Deslauriers, L. H. Groulx, A. Laperrière, R. Mayer & A. P. Pires (éds.). *La recherche qualitative : enjeux épistémologiques et méthodologiques* (pp. 365-389). Boucherville, QC: Gaëtan Morin.
- Legault, L., Mák, G., Verstraete, M. et Bérubé, A. (2014). *Capacité d'apprentissage dans les communautés francophones en situation minoritaire : Rapport de référence*. Ottawa (Ontario) : Société de recherche sociale appliquée.

- Lemelin, J.-P., & Boivin, M. (2007). *Mieux réussir dès la première année : l'importance de la préparation à l'école*. Québec, QC : Institut de la statistique Québec. Disponible en ligne au : <http://www.jesuisjeserai.stat.gouv.qc.ca/pdf/publications/feuille/Fasc2Vol4.pdf>.
- LeTouzé, S. (2005). *Pour mon enfant d'abord : étude de l'impact de l'alphabétisation familiale sur les familles vivant en milieu minoritaire. Étape 1 : 2004-2005*. Ottawa, ON : Centre interdisciplinaire de recherche sur la citoyenneté et les minorités.
- LeTouzé, S. (2006). *Pour mon enfant d'abord : étude de l'impact de l'alphabétisation familiale sur les familles vivant en milieu minoritaire. Étape 2 : Cohortes 2 et 3 : 2005-2006*. Ottawa, ON : Centre interdisciplinaire de recherche sur la citoyenneté et les minorités.
- LeTouzé, S. (2007). *Pour mon enfant d'abord : étude de l'impact de l'alphabétisation familiale sur les familles vivant en milieu minoritaire. Étape 3 : Cohortes 4 et 5. 2006-2007*. Ottawa, ON : Centre interdisciplinaire de recherche sur la citoyenneté et les minorités.
- Lipsey, M. W., & Wilson, D. B. (1993). The efficacy of psychological, educational, and behavioral treatment. *American Psychologist*, 48(12), 1181-1209.
- Little, R. J. A., & Rubin, D. B. (1987). *Statistical analysis with missing data*. New York, NY: John Wiley & Sons.
- Lonigan, C. J. (2008). *Effective strategies for promoting young children's literacy skills*. Papier présenté au 40th Banff International Conferences on Behavioural Science. Effective Early Learning Programs: Research, Policy and Practice, Banff, AB.
- Lonigan, C. J., Schatschneider, C., Westberg, L., & The National Early Literacy Panel. (2008). Identification of children's skills and abilities linked to later outcomes in reading, writing, and spelling. Dans *Developing early literacy: Report of the National Early Literacy Panel* (pp. 55-101). Jessup, MD: National Institute for Literacy.
- Maltais, C. (2007). Relation entre la littératie familiale en milieu francophone et l'incidence d'un programme de maternelle quatre ans à temps plein sur le développement du langage et de la lecture des enfants. *Éducation francophone en milieu minoritaire*, 2(1), 6-18. Disponible en ligne au : <http://www.ustboniface.mb.ca/cusb/reefmm/notrerevue/v2n1Maltais.pdf>.
- Martel, A. (2001). *Droits, écoles et communautés en milieu minoritaire : 1986 - 2002. Analyse pour un aménagement du français par l'éducation*. Ottawa, ON : Commissariat aux langues officielles.
- Meisels, S. J., & Atkins-Burnett, S. (2006). Evaluating early childhood assessments: A differential analysis. Dans K. McCartney & D. Phillips (éds.), *The Blackwell handbook of early childhood development* (pp. 533-549). Malden, MA: Blackwell Publishing.
- Moore, T. (2005). *Towards improved support for families of young children: The role of communities*. Présenté au 9th Australian Institute of Family Studies Conference. Disponible en ligne au : <http://aifs.gov.au/institute/afrc9/moore1.pdf>.
- Moore, D., & Tananis, C. A. (2009). Measuring change in a short-term educational program using a retrospective pretest design. *American Journal of Evaluation*, 30(2), 189-202.

- Moulton, B. R. (1990). An illustration of a pitfall in estimating the effects of aggregate variables on micro units. *The Review of Economics and Statistics*, 72(2), 334-338.
- National Association for the Education of Young Children. (1984). *Accreditation criteria and procedures: Position statement of the National Academy of Early Childhood Programs*. Washington, D.C., DC: National Association for the Education of Young Children.
- National Institute of Child Health and Human Development Early Child Care Research Network, & Duncan, G. J. (2003). Modeling the impacts of child care quality on children's preschool cognitive development. *Child Development*, 74(5), 1454-1475.
- Patton, M. Q. (1990). *Qualitative evaluation and research methods* (2nd ed.). Newbury Park, CA: Sage Publications.
- Patton, M. Q. (2008). *Utilization-focused evaluation* (4th ed.). Thousand Oaks, CA: Sage Publications, Inc.
- Pearson, B. Z. (2007). Social factors in childhood bilingualism in the United States. *Applied Psycholinguistics*, 28(3), 399-410.
- Pearson, B. Z., Fernandez, S., Lewedag, V., & Oller, D. K. (1997). The relation of input factors in lexical learning by bilingual infants. *Applied Psycholinguistics*, 18(1), 41-58.
- Pelletier, J., & Corter, C. (2005). Design, implementation, and outcomes of a school readiness program for diverse families. *The School Community Journal*, 17(3), 89-116.
- Phillips, L. M., Hayden, R., & Norris, S. P. (2006). *Family literacy matters: A longitudinal parent-child interaction study*. Calgary, AB: Temeron Books.
- Raikes, H. H., Summers, J. A., & Roggman, L. A. (2005). Father involvement in Early Head Start programs. *Fathering: A Journal of Theory, Research, and Practice about Men as Fathers*, 3(1), 29-58.
- Reese, E., Sparks, A., & Leyva, D. (2010). A review of parent interventions for preschool children's language and emergent literacy. *Journal of Early Childhood Literacy*, 10(1), 97-117.
- Ressources humaines et développement des compétences Canada (2006). *Projet pilote de garde d'enfants : principes directeurs du projet pilote, 2006*. Disponible en ligne au : http://www.RHDCC.gc.ca/fra/sm/comm/ds/garde_enfants/2006/principes_directeurs.shtml.
- Rockwell, S. K., & Kohn, H. (1989). Post-then-pre evaluation. *Journal of Extension*, 27(2) [en ligne]. Récupérée au : <http://www.joe.org/joe/1989summer/a5.php>.
- Rogers, K. N., Fernandez, M., Thurber, L., & Smitley, A. (2004). Exploring differential attrition rates among system of care evaluation participants. *Journal of Community Psychology*, 32(2), 167-176.
- Sanders, M. R., & Morawska, A. (2006). Peut-on améliorer les résultats des enfants en modifiant les connaissances des parents, leurs attentes dysfonctionnelles et la régulation de leurs émotions? Dans R. E. Tremblay, R. de V. Peters, M. Boivin & R. G. Barr (éds.), *Encyclopédie sur le développement des jeunes enfants* [sur Internet] (pp. 1-13). Montréal,

- QC : Centre d'excellence pour le développement des jeunes enfants. Disponible en ligne au : <http://www.enfant-encyclopedie.com/Pages/PDF/Sanders-MorawskaFRxp.pdf>.
- Sénéchal, M., & LeFevre, J.-A. (2002). Parental involvement in the development of children's reading skills: A five-year longitudinal study. *Child Development*, 73(2), 445-460.
- Sénéchal, M., & Young, L. (2008). The effect of family literacy interventions on children's acquisition of reading from kindergarten to grade 3: A meta-analytic review. *Review of Educational Research*, 78(4), 880-907.
- Ska, B. (1995). *Épreuve de compréhension de CARROW-WOOLFOLK. Épreuve de dénomination EO-WPVT-R. Données normatives*. Montréal, QC : Groupe coopératif en orthophonie pour la région des Laurentides.
- Snow, C. E., Burns, M. S., & Griffin, P. (éds.). (1998). *Preventing reading difficulties in young children*. Washington, D.C., DC: National Academic Press.
- Statistique Canada. (2006). *Enquête sur la vitalité des minorités de langue officielle (EVMLO) : questionnaire*. Ottawa, ON : Auteur. Disponible en ligne au : http://www.statcan.gc.ca/imdb-bmdi/instrument/5099_Q1_V1-fra.pdf.
- Statistique Canada, & Ressources humaines et Développement social Canada. (2005). *Enquête longitudinale nationale sur les enfants et les jeunes. Aperçu d'enquête pour la collecte de données de 2002-2003. Cycle 5*. Ottawa, ON : Auteur.
- Statistique Canada, & Ressources humaines et Développement social Canada. (2006). *Enquête longitudinale nationale sur les enfants et les jeunes : Matériel d'enquête pour la collecte de données 2004-2005. Cycle 6. Livre 1 : Parent, enfant et jeune*. Ottawa, ON : Auteur.
- Strauss, A. L., & Corbin, J. (1998). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (2nd ed.). Thousand Oaks, CA: Sage Publications.
- Tabachnick, B. G., & Fidell, L. S. (2006). *Using multivariate statistics* (5th ed.). Boston, MA: Allyn & Bacon.
- Tamis-LeMonda, C. S., & Rodriguez, E. T. (2008). Rôle des parents pour favoriser l'apprentissage et l'acquisition du langage chez les jeunes enfants. Dans R. E. Tremblay, R. G. Barr, R. de V. Peters & M. Boivin (éds.), *Encyclopédie sur le développement des jeunes enfants* [sur Internet] (pp. 1-11). Montréal, QC : Centre d'excellence pour le développement des jeunes enfants. Disponible en ligne au : <http://www.ccl-cca.ca/pdfs/ECLKC/encyclopedia/TamisLemondaRodriguezFRxpCSAJELangage.pdf>.
- Taylor, P. J., Russ-Eft, D. F., & Taylor, H. (2009). Gilding the outcome by tarnishing the past: Inflationary biases in retrospective pretests. *American Journal of Evaluation*, 30(1), 31-43.
- Thomas, E. M. (2006). *La disposition à apprendre à l'école pour les jeunes de cinq ans au Canada*. Ottawa, ON : Statistique Canada. Disponible en ligne au : <http://www.statcan.gc.ca/pub/89-599-m/89-599-m2006004-fra.pdf>.
- Thompson, G., Legault, L., Lalonde, P. et Bérubé, A. (2014). *Capacité d'apprentissage dans les communautés francophones en situation minoritaire : Rapport des résultats de la première cohorte*. Ottawa (Ontario) : Société de recherche sociale appliquée.

- Tremblay, R. E., Nagin, D. S., Séguin, P. D., Boivin, M., Pérusse, D., & Japel, C. (2004). Physical aggression during early childhood: Trajectories and predictors. *Pediatrics*, *114*(1), e43-e50.
- van Steensel, R., McElvany, N., Kurvers, J., & Herppich, S. (2011). How effective are family literacy programs?: Results of a meta-analysis. *Review of Educational Research*, *81*(1), 69-96.
- Vihman, M. M., Lum, J. A. G., Thierry, G., Nakai, S., & Keren-Portnoy, T. (2006). The onset of word form recognition in one language and in two. Dans P. McCardle & E. Hoff (éds.), *Childhood bilingualism: Research on infancy through school age* (pp. 30–44). Clevedon, UK: Multilingual Matters.
- Vygotsky, L. S. (1978). *Mind in society: Development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Wasik, B. A., Bond, M. A., & Hindman, A. (2006). The effects of a language and literacy intervention on Head Start children and teachers. *Journal of Educational Psychology*, *98*(1), 63-74.
- White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica*, *48*(4), 817-838.
- Williams, R. L. (2000). A note on robust variance estimation for cluster-correlated data. *Biometrics*, *56*(2), 645-646.
- Willms, J. D. (2007). *Child Care Pilot Project: Final report on direct assessment measures and parent and community questionnaires*. Fredericton, NB: KSI Research International Inc.
- Willms, J. D. (2009). *Évaluation Petite Enfance : Appréciation Directe, version allongée*. Fredericton, NB : KSI Research International Inc.
- Woodbridge, J. M. (2002). *Econometric analysis of cross section and panel data*. Cambridge, MA: MIT Press.

Annexe A: Calendrier et outils d'évaluation

Le tableau présenté ci-dessous fournit le calendrier des évaluations des enfants. De plus, on y retrouve, pour chacune des vagues de collecte de données, le construit à l'étude (*en italique*) et l'outil utilisé pour évaluer son développement auprès des enfants.

Tableau A.1 : Calendrier des évaluations des enfants

Niveau de base	+ 4 mois (1 ^{er} post-test)	+8 mois (2 ^e post-test)	+ 12 mois (3 ^e post-test)	+16 mois (4 ^e post-test)	+20 mois (5 ^e post-test)	+24 mois (6 ^e post-test)
<i>Préparation à la scolarisation</i>						
ÉPE-AD ⁹⁶ : Conscience de soi (A)						
ÉPE-AD : Habiletés cognitives (B)						
ÉPE-AD : Langage et Communication (C)						
ÉPE-AD : Habiletés physique et motrice (D)						
<i>Vocabulaire expressif</i>						
Sous-échelle de l'ÉPE-AD						
					EOWPVT-R	
<i>Vocabulaire réceptif</i>						
Sous-échelle de l'ÉPE-AD						
						ÉVIP-R

¹⁰⁵ Évaluation de la petite enfance — Appréciation directe (Willms, 2007)

Annexe B : Contenu des trois versions de l'ÉPE-AD

Cette section concerne le contenu de trois versions de l'ÉPE-AD (le EYE-DA en anglais) utilisées au cours des deux premières années de mise en œuvre du projet Capacité d'apprentissage. De nouvelles versions de l'ÉPE-AD ont été créées en réponse à deux enjeux notés à l'automne 2008 et suite aux décisions prises lors d'une rencontre avec le Comité consultatif de RHDCC qui a eu lieu le 17 décembre 2008. Le premier enjeu fait référence au risque élevé d'un effet plafond soulevé par les projections de scores pour la 5e évaluation (février 2009) à partir des résultats aux 3e et 4e évaluations. Les services du concepteur de l'outil ont donc été retenus par RHDCC pour créer de nouveaux items plus difficiles permettant ainsi de suivre la trajectoire développementale des enfants quant aux domaines A, B, et C. Le deuxième enjeu concerne la nécessité d'ajouter une mesure d'habiletés linguistiques afin de faire une discrimination plus fines des enfants. Bien que l'instrument ÉPE-AD mesure bien la préparation à scolarisation des enfants, il n'était pas suffisamment « sensible » pour capter les dimensions développementales chez les enfants et il ne mesure qu'en surface leurs habiletés de communication.

Pour les fins de la discussion qui suit, nous nommons « ÉPE-AD initiale » l'outil utilisé en février 2009 dans le cadre du projet Capacité d'apprentissage et nous nommons « ÉPE-AD étendue » l'outil administré au cours de l'hiver 2009 (comprenant les questions plus difficiles développées par Willms et excluant les questions plus faciles de l'ÉPE-AD initiale). Enfin, nous nommons « ÉPE-AD modifiée » la version de l'outil qui a été retravaillée par la SRSA et qui a été utilisée lors des évaluations menées de juin à octobre 2009.

Domaines mesurés par l'ÉPE-AD initiale

L'ÉPE-AD conçue par Doug Willms comprend quatre domaines et un cinquième conçu spécifiquement pour le projet Capacité d'apprentissage :

- Domaine A = Conscience de soi;
- Domaine B = Habiletés cognitives;
- Domaine C = Langue et communication;
- Domaine D = Habiletés physique et motrice; et
- Domaine E = Conscience et engagement envers la culture francophone.

Domaines mesurés par l'ÉPE-AD modifiée

L'ÉPE-AD modifiée comprend sept dimensions, notamment:

- Vocabulaire expressif;
- Conscience de soi;
- Conscience phonologique;
- Numératie;

- Mémoire pour les informations personnelles;
- Connaissances alphabétiques; et
- Lecture orale.

Ces échelles permettent de poursuivre la trajectoire développementale des enfants pour les domaines A, B et C. La version de l'ÉPE-AD modifiée comprend la totalité des items du domaine A (moins la question A18). Pour le domaine B, la totalité des items a été conservée, soit dans les sous-échelles de Conscience phonologique, de Numératie, ou de Connaissances alphabétiques. Enfin, le domaine C est en partie conservé dans la sous-échelle de Vocabulaire expressif, avec laquelle il présente une corrélation de 0,87, confirmant ainsi qu'il sera possible de suivre la trajectoire développementale de ce domaine.

Les tableaux suivants offrent une comparaison des items retrouvés dans les trois versions de l'ÉPE-AD. La version modifiée de l'ÉPE-AD a été conçue en français seulement; les items présentés dans les tableaux sont des verbatims issus de cette évaluation. Le lecteur est prié de noter que les items non retenus dans la version de l'EPE-AD modifiée ne se retrouvent pas dans le tableau. Ces items incluent ceux éliminés par Willms dans la version de l'EPE-AD étendue.

Tableau B.1 : Sous-échelle de Vocabulaire expressif

Items	ÉPE-AD initiale	ÉPE-AD étendue	ÉPE-AD modifiée
Vocabulaire expressif			
Comment s'appelle cet objet? <i>un seau/une chaudière, des boutons, des chandelles, une horloge/un réveille-matin</i>	A11	x	V1
Pointez chacune des images, une à la fois. Peux-tu me nommer quatre couleurs? Peux-tu me nommer quatre fruits? Peux-tu me nommer quatre animaux? Peux-tu me nommer quatre vêtements?	A12	x	V2
Comment se nomme cette partie du corps? <i>le menton, le coude, le poignet, l'épaule</i>	A14	x	V3
Qu'est-ce que cette personne fait comme métier? <i>le policier, l'enseignant(e), le fermier/jardinier, le médecin</i>	A15	x	V4
Qu'est-ce qu'on utilise pour écrire? pour s'asseoir? pour se promener? pour balayer?	C8	x	V5
Nomme le plus d'animaux possible.	C9 (nomme huit animaux)	x	V6
Quel temps fait-il dans chacune des images? <i>Il vente, il pleut, il neige, il fait soleil</i>		A21	V7
Comment s'appelle ceci? <i>thermomètre, tasse à mesurer, règle, balance</i>		A22	V8

Note : L'échelle de Vocabulaire expressif (six items) est corrélée à 0,87 aux items du domaine C de l'échelle ÉPE-AD étendue (C7 à C14). Nous avons conservé uniquement les items qui demandent à l'enfant de nommer un objet, de façon à être cohérent avec la méthodologie des tests de vocabulaire conventionnels, tel l'EOWPVT.

Tableau B.2 : Sous-échelle de Conscience de soi

Items	ÉPE-AD initiale	ÉPE-AD étendue	ÉPE-AD modifiée
Conscience de soi			
Que devrais-tu faire quand tu : es fatigué, as faim, as froid, as soif?	A7	x	CS1
La souris est petite; l'éléphant est ____ (gros ou grand). Cet oiseau est dehors : cet oiseau est ____ (en dedans ou à l'intérieur). Ce pot à biscuits est plein; ce pot à biscuits est ____ (vide). Cette échelle est courte; cette échelle est ____ (longue).	A8	x	CS2
Placer des objets pour démontrer sa compréhension de : premier, dernier, devant, derrière .	A10	x	CS3
Nommer le moment de la journée (le matin, l'après-midi, la soirée, la nuit) associé à deux de ces situations communes : des étoiles dans le ciel, prendre le petit déjeuner, retourner à la maison après l'école	A13	x	CS4
Quel jour est-ce aujourd'hui? Peux-tu me dire quel jour on sera demain? Hier, quel jour était-ce? Peux-tu me nommer un jour de la fin de semaine?		A17	CS5
Avant de traverser la rue, que dois-tu faire? Que dois-tu faire si un inconnu te demande d'aller avec lui? Dans la voiture, que portes-tu toujours pour être en sécurité? Si tu entends le détecteur de fumée, que dois-tu faire?		A20	CS6

Tableau B.3 : Sous-échelle de Conscience phonologique

Items	ÉPE-AD initiale	ÉPE-AD étendue	ÉPE-AD modifiée
Conscience phonologique			
<i>Écoute-moi prononcer le mot jouet : jouet commence par le son /j/.</i> Par quel son commence le mot piano? Par quel son commence le mot tomate? Par quel son commence le mot boîte? Par quel son commence le mot rouge?		C16	CP1
Dis-moi si les mots riment, s'ils finissent par le même son. pomme/homme; boîte/lune; chien/chat; pain/main	B11	x	CP2
Trouve deux mots qui commencent par /b/ comme dans « ballon ».	B13	x	CP3

Dis-moi si les mots suivants commencent par le même son : mère/lait, balle/beau, plat/clou, peau/pas	B14	x	CP4
Écoute pendant que je prononce le mot canif. Le mot canif se termine par le son /f/. Quel est le son qui termine le mot cheval? Quel est le son qui termine le mot album? Quel est le son qui termine le mot autobus? Quel est le son qui termine le mot neuf (9)?		C19	CP5
Voici un bateau (une pomme, une table, une fenêtre et du lait). Le mot bateau commence par un son /b/. Écoute /b/, bateau. Quelle image commence par le son t? Quelle image commence par le son f? Quelle image commence par le son l? Quelle image commence par le son p?		B18	CP6

Note : Pour l’item C16, nous demandons seulement de dire par quel son commence le mot. Nous ne demandons pas d’identifier la lettre. Le but étant de mesurer la conscience phonologique, nous tentons de garder cette mesure pure. Les mots « piano » et « tomate » ont été ajoutés pour arriver à un score maximum de 4, tout comme avec les autres versions du test.

Tableau B.4 : Sous-échelle de Numératie

Items	ÉPE–AD initiale	ÉPE–AD étendue	ÉPE–AD modifiée
Numératie			
Dire le nombre de parties : Un chat a combien de queues? Un chien a combien de pattes? Un oiseau a combien d’ailes? Ta main a combien de doigts?	A9	x	N1
Compte toutes les étoiles et dis-moi combien il y en a.	B9	x	N2
Regarde chaque chiffre et dis-moi lequel est <u>le plus</u> grand. 6 ou 8, 12 ou 10, 9 ou 7, 11 ou 12.	B10	x	N3
Mets ensemble des jetons pour faire un groupe de... 5, 7, 8, 9.	B12	x	N4
Montre-moi le chiffre... 13, 20, 45, 112.		B19	N5

Note : Pour l’item B19, l’ordre des chiffres a été changé sur l’image afin que les chiffres ne soient pas présentés dans le même ordre que demandés.

Tableau B.5 : Mémoire pour les informations personnelles

Items	ÉPE–AD initiale	ÉPE–AD étendue	ÉPE–AD modifiée
Mémoire pour les informations personnelles			
Peux-tu me dire quel âge tu as et quels sont le jour et le mois de ton anniversaire (de ta fête)?	A16	x	MP1
Savoir où l’on habite : Quel est le nom de la rue où tu habites?		A19	MP2

Quel est le nom de la ville où nous habitons? Quel est le nom de notre province? Quel est le pays où nous habitons?			
---	--	--	--

Tableau B.6 : Sous-échelle de connaissances alphabétiques

Items	ÉPE-AD initiale	ÉPE-AD étendue	ÉPE-AD modifiée
Connaissances alphabétiques			
Dire le son de la lettre majuscule	B16	B17	Livret
Dire la lettre majuscule	B15		
Dire la lettre minuscule		B20	Livret

Note : Les lettres sont présentées en ordre de difficulté croissante pour les francophones canadiens. Nous avons présenté un total de 26 lettres, incluant celles avec des accents « é, è, ê ». Chaque réponse correcte obtient un score de 1. L'étendue des scores est donc, pour B17 et B20, de 0 à 26.

Tableau B.7 : Lecture orale

Items	ÉPE-AD initiale	ÉPE-AD étendue	ÉPE-AD modifiée
Lecture orale			
Lire huit mots perçus de façon globale à fréquence élevée. Dis-moi quel mot est écrit ici.	B17	B16 Chien Oui Non Balle Chat Le La Nez	L1 Le La Oui Non Balle Maman Nez Papa

Note : Nous avons changé l'ordre des mots de façon à ce que leur fréquence d'utilisation en français soit décroissante. Nous avons également changé deux mots pour les remplacer par d'autres mots plus communs en français. Les mots anglais « dog » et « cat » font souvent partie des livres d'alphabet anglophones alors que leurs équivalents en français présentent une orthographe plus complexe (c.-à-d. les lettres « ch » (pour chien et chat) constituent un symbole de plusieurs lettres pour représenter un seul son). Nous avons substitué ces mots par les mots « maman » et « papa », qui sont plus long, mais plus fréquents dans les livres francophones pour enfants.

Annexe C : Sommaire des thèmes abordés dans les Ateliers familles

N° 1 – Apprendre en français, c’est amusant!

- intro au projet Capacité d’apprentissage;
- nos enfants à 18 ans

N° 2 – Je suis le premier éducateur de mon enfant

- ce que vous faites comme premier éducateur de l’enfant
- l’étayage
- les styles de parentage
- les intelligences multiples (Gardner)

N° 3 – L’éveil à l’écrit

- ce que l’enfant apprend avant de savoir lire et écrire
- choisir un bon livre
- la lecture partagée

N° 4 – L’apprentissage : stimulation des sens de l’enfant

- la stimulation des sens de l’enfant
- le développement global
- comment stimuler l’apprentissage de l’enfant à la maison

N° 5 – Mon enfant en quête d’autonomie

- l’autodiscipline pour développer l’estime de soi
- le rôle des émotions dans l’apprentissage et la réussite dans la vie
- les stratégies pour amener l’enfant vers l’autodiscipline, comme nommer les émotions

N° 6 – Le développement langagier, culturel, et identitaire de mon enfant

- l’importance des chansons, contes et comptines
- le bilinguisme additif
- les variétés de français
- comment stimuler le langage à la maison

N° 7 – La communication

- ce qui facilite la communication
- la communication verbale et non verbale
- le rôle de la position physique et de l'attitude dans la communication

N° 8 – Le bien-être de notre famille

- le changement dans les familles
- les besoins de l'enfant (c.-à-d., selon Maslow)
- les valeurs à transmettre à nos enfants

N° 9 – Notre vie au sein de la communauté francophone

- les activités, ressources et services dans la communauté francophone
- savoirs en Francophonie
- les différences entre l'école française et l'immersion

N° 10 – Les célébrations

- retour sur les apprentissages des Ateliers familles
- les traditions
- les célébrations sont une partie intégrante de la famille

Annexe D : Procédure pour l'administration de l'ÉPE–AD (Mesure pré-intervention)

Les évaluatrices qui ont administré l'ÉPE–AD (le EYE-DA en anglais) aux enfants ont été recrutées à compter de la fin de l'été 2007. Elles ont reçu de la SRSA une formation théorique et pratique d'une durée d'environ six heures aux mois d'août, septembre et, pour la communauté d'Orléans, octobre. Cette formation visait, en plus de la présentation du protocole d'administration du test, la présentation du projet Capacité d'apprentissage ainsi que des procédures relatives à la confidentialité. Les évaluatrices se sont engagées par contrat à respecter le protocole d'administration et de confidentialité. Les étapes complètes du protocole pour l'administration du test sont :

1. Les évaluatrices appellent les parents pour prendre les rendez-vous des évaluations à la maison ou pour les prévenir du moment de l'évaluation en garderie. Ces appels ont aussi pour but de confirmer l'âge en mois de l'enfant ainsi que les réponses aux questions concernant les langues parlées avec les parents et les amis pour le domaine E (*Conscience et engagement envers la culture francophone*).
2. Les évaluatrices et la coordonnatrice communautaire prennent contact avec les garderies participantes pour définir l'horaire et aménager un endroit dans la classe favorable à une bonne évaluation.
3. L'évaluatrice applique la méthode « médicale », c'est-à-dire qu'elle attend de vérifier que l'enfant soit le bon avant de remplir les renseignements d'identification sur le questionnaire papier.
4. L'évaluatrice s'adresse à l'enfant dans sa langue maternelle tout d'abord, puis applique le protocole pour déterminer la langue du test.
5. L'évaluatrice suit les indications de notation fournies par le concepteur de l'outil, ce qui veut dire qu'elle arrondit les points à l'entier inférieur dans le but de réaliser une évaluation prudente.
6. L'évaluatrice encourage l'enfant, mais n'offre pas d'indices sur les réponses à donner, sauf si le protocole l'indique.
7. Si l'enfant est fatigué en cours de test, l'évaluatrice s'arrête et peut recommencer plus tard au début du domaine où elle s'était arrêtée.
8. À la fin de l'évaluation, l'évaluatrice remet à l'enfant un autocollant pour le remercier de sa participation.
9. Dans les cas où l'enfant ne voudrait vraiment pas participer, l'évaluatrice doit tenter d'évaluer l'enfant au moins une autre fois (un autre jour).

Le point 4 est définitivement un point clé pour l'évaluation puisque le respect ou le non-respect de cette règle peut influencer les résultats de manière considérable. Le protocole

d'administration pour déterminer la langue du test, suggéré par M. Willms (présenté le 8 décembre 2006 et révisé en juillet 2007), est présenté au tableau D.1.

Tableau D.1 : Arbre décisionnel pour déterminer la langue d'administration du test

<i>Étapes</i>	<i>Protocole</i>
1. Administration du domaine E	Administrer les six questions du domaine E directement <u>à l'enfant, et au parent pour les questions E4 à E6</u> , au début du test.
2. Arbre de décision pour déterminer la langue du test (domaine E)	Si le score est supérieur à 6 pour le domaine E, administrer le reste de l'ÉPE-AD en français. Si le score est inférieur ou égal à 6 pour le domaine E, évaluer le domaine C en français et en anglais.
3. Arbre de décision pour déterminer la langue du test (domaine C)	Si le score en français pour ce domaine est supérieur à 14, le reste de l'évaluation se déroule en français. Si le score en français pour ce domaine est inférieur ou égal à 14 et le score en anglais est inférieur ou égal à 14, le reste du test se déroule également en français. Si le score pour le domaine C est inférieur ou égal à 14 en français, mais le score est supérieur à 14 en anglais, le reste de l'évaluation se fait en anglais.

Également, il y a deux points de départ pour l'administration du test, dépendant de l'âge de l'enfant. Si l'enfant est âgé de moins de quatre ans, l'évaluatrice débute avec le premier item pour le domaine. Autrement, elle débute plus loin dans le test, avec l'option de retourner au point de départ si l'enfant présente des difficultés. Le Rapport de référence n'a pas fait mention de cette règle puisque tous les enfants étaient âgés de moins de quatre ans lors de l'évaluation de base.

Annexe E : Validation de l'imputation de l'ÉPE-AD

La stratégie d'imputation adoptée dans le présent rapport se base sur la performance antérieure et ultérieure de l'enfant sur la même échelle. Il est toutefois utile de quantifier la justesse de l'algorithme que nous avons utilisé pour une situation particulière. Dans ce qui suit, nous présentons les résultats d'une analyse dont l'objectif est de valider la stratégie d'imputation utilisée pour pallier aux valeurs manquantes de l'ÉPE-AD.

Pour ce faire, nous avons pris les données de la deuxième cohorte comme étude de cas, retenant uniquement les enfants avec des données complètes. Dans un premier temps, nous avons supprimé les valeurs observées de la deuxième évaluation. Ensuite, ces valeurs manquantes artificielles ont été estimées par l'algorithme d'imputation. L'évaluation de l'algorithme s'est fait en comparant les valeurs observées réelles aux valeurs imputées. Si l'imputation est satisfaisante, les scores originaux et les scores imputés seront distribués de façons semblables (c.-à-d. des moyennes et des écarts-type semblables) et ils seront fortement corrélés.

Les résultats de cette analyse sont rapportés dans le tableau E.1. Pour les échelles de Communication, Cognition, et Conscience de soi, l'algorithme a tendance à surestimer la performance réelle des enfants. Toutefois, la différence entre la moyenne des scores originaux et la moyenne des scores imputés n'est pas significative. Pour ces résultantes des enfants, l'association entre les deux distributions est très forte, avec des corrélations variant de 0,70 à 0,88. En somme, la performance de l'algorithme pour ces trois échelles confirme nos attentes de la justesse de la méthode d'imputation utilisée.

Exceptionnellement, l'algorithme a tendance à sous-estimer significativement la performance des enfants au domaine physique, soit d'environ 0,30 écart-type. La corrélation entre les scores originaux et les scores imputés est toutefois relativement forte pour ce domaine, avec des corrélations observées de 0,50. En somme, la justesse de l'algorithme d'imputation pour cette échelle est moindre que prévu.

Précisons que peu importe la magnitude du biais observée, les conséquences pour les analyses d'impact ne sont pas sérieuses. Les valeurs manquantes se répartissent également parmi les groupes expérimentaux, sauf au temps quatre (voir le tableau 7.2), indiquant ainsi que la validité interne des analyses examinant le domaine Physique n'est pas en jeu. Nous avons jugé préférable de garder la composition de l'échantillon d'une analyse à l'autre aussi constant que possible. L'imputation des valeurs manquantes est alors effectuée pour toutes les échelles de l'ÉPE-AD et ce, en dépit du biais potentiel pour le domaine Physique.

Tableau E.1 : Moyennes et écarts types des scores originaux et imputés ainsi que leur inter-corrélation

Domaine ÉPÉ-AD	Données originales		Données imputées		Taille de la différence (<i>d</i> Cohen)	Intervalle de confiance sur <i>r</i> (95 %)		
	M	É.-T.	M	É.-T.		<i>r</i>	Borne inférieure	Borne supérieure
Cognition	24,36	± 10,23	26,50	± 11,61	0,20	0,70	0,57	0,79
Communication	26,85	± 11,18	28,40	± 11,62	0,14	0,77	0,67	0,84
Conscience de soi	33,95	± 11,44	34,23	± 12,74	0,02	0,88	0,82	0,92
Physique	14,17	± 7,01	11,39	± 9,20	- 0,34**	0,50	0,32	0,64

Note : M = moyenne ; É.-T. = écart-type ; N=84 ; Seuils de signification : *** $p < 1$ % ; ** $p < 5$ % ; * $p < 10$ %.

Annexe F : Résultat de l'analyse des variables confusionnelles

Cette annexe rapporte les résultats d'une analyse préliminaire visant l'identification des covariables pertinentes pour les analyses d'impact. Les résultats portent sur l'échantillon combinant les deux cohortes et excluant les retraits de la première année (N = 342). À moins d'indication contraire, l'appartenance des enfants aux groupes expérimentaux est celle du niveau de base pour les tests d'association impliquant cette variable.

Les variables qui rencontrent les deux critères d'inclusion sont retenues comme covariables lors de l'analyse d'impact. Ces deux critères sont : i) une association significative avec au moins une variable dépendante et ii) une association significative avec l'appartenance aux groupes expérimentaux. Nous rapportons dans les tableaux suivants les tests de signification de ces deux critères pour chaque covariable potentielle.

La présentation des résultats est organisée de la façon suivante. Les variables examinées sont regroupées dans des tableaux de la façon suivante : les variables démographiques et de composition familiale, les variables socioéconomiques, les variables linguistiques, les variables d'engagement à la francophonie, les variables de style parental, les caractéristiques de l'environnement en garderie et finalement une série de facteurs méthodologiques. Enfin, un dernier tableau contient les résultats d'une analyse vérifiant la stabilité de l'effet de certaines covariables dans le temps. Dans ce qui suit, les variables retenues comme covariables sont en caractères gras.

Tableau F.1 : Pertinence des variables de démographie et de composition familiale

Covariable	Corrélations significatives ($p < .10$) avec les variables dépendantes au niveau de base									Tests d'association avec l'appartenance aux groupes expérimentaux			
	Com	Con	Cog	Phy	Voc. expressif	Voc. réceptif	Fréq. litt.	Langue litt.	Paramètre de l'échantillon % ou M (É.-T.)	Test d'association avec les groupes	Groupe programme % ou M (É.-T.)	Groupe témoin en garderie % ou M (É.-T.)	Groupe témoin hors garderie % ou M (É.-T.)
Sexe (+Femme) ^a	0,11	0,13	0,19	0,17	0,12	0,13	-	-	50,9 %	$\chi^2(2) = 1,37$	46,1 %	53,4 %	52,3 %
Âge de l'enfant (en mois) ^b	0,30	0,35	0,44	0,47	0,29	0,25	-	-	38,42 (3,66)	F(2, 339) = 0,24	38,35 (3,78)	38,59 (3,52)	38,28 (3,75)
Jumeaux (+Jumeaux)	-	-	-	-	-	-	-	-	2,3 %	SO	3,9 %	1,5 %	1,9 %
Fratrie plus âgée	-	-0,13	-	-	-0,13	-	-0,25	-0,11	52,3 %	$\chi^2(2) = 2,74$	48,0 %	57,9 %	49,5 %
Fratrie plus jeune	-	-	0,10	-	-	-	0,20	-	34,5 %	$\chi^2(2) = 22,19^{***}$	24,5 %	27,8 %	52,3 %
Foyer monoparental (+monoparental) ^a	-	-	-	-	-	-	-	-	8,5 %	$\chi^2(2) = 5,51^*$	6,9 %	12,8 %	4,7 %
Taille du ménage	-	-0,12	-	-	-0,10	-	-0,12	-,13	4,04 (0,93)	F(2, 339) = 4,97^{***}	3,87 (0,83)	3,99 (1,01)	4,26 (0,88)

Note : Les variables retenues comme covariables dans les analyses d'impact sont en caractères gras. ^aToutes les variables catégorielles sont binaires (codes 0, 1) et les statistiques descriptives rapportés dans le tableau sont celles de la catégorie « 1 » indiquée entre parenthèses et précédée par le symbole (+). Com= Communication, Con = Conscience de soi, Cog = Cognition, Phy = Physique, Voc. = Vocabulaire, Fréq litt. =Fréquences des activités de littératie, Langue litt. = Langue des activités de littératie, M = moyenne ; É.-T. = écart-type ; N=342 ; Seuils de signification : *** $p < 1$ % ; ** $p < 5$ % ; * $p < 10$ %.

^b Exceptionnellement, nous retenons cette variable comme covariable en dépit du fait qu'elle n'est pas associée à l'appartenance aux conditions de traitement. Cette décision est justifiée par sa très forte association avec les variables dépendantes.

Tableau F.2 : Pertinence des variables socio-économiques

Covariable	Corrélations significatives ($p < .10$) avec les variables dépendantes au niveau de base									Tests d'association avec l'appartenance aux groupes expérimentaux			
	Com	Con	Cog	Phy	Voc. expressif	Voc. réceptif	Fréq. litt.	Langue litt.	Paramètres globaux % ou M (É.-T.)	Test d'association avec les groupes	Groupe programme % ou M (É.-T.)	Groupe témoin en garderie % ou M (É.-T.)	Groupe témoin hors garderie % ou M (É.-T.)
Âge à la première naissance (+Jeune mère < 24ans) ^a	-0,17	-	-	-0,09	-0,13	-0,14	-	-0,12	16,1 %	$\chi^2(2) = 0,77$	14,7 %	15,2 %	18,7 %
Revenu (+60 000 \$ et plus) ^a	0,14	0,13	0,16	0,10	0,13	0,11	-	-	76,6 %	$\chi^2(2) = 1,53$	76,5 %	79,7 %	72,9 %
Éducation mère ^b (+Au moins un diplôme collégial)	0,24	0,25	0,18	-	0,26	0,20	-	-	79,8 %	$\chi^2(2) = 0,27$	81,4 %	79,7 %	78,5 %
Éducation père ^b (+Au moins un diplôme collégial) ^a	-	-	-	-	-	-	-	-	67,0 %	$\chi^2(2) = 2,77$	70,6 %	69,2 %	60,7 %
Statut d'immigrant	-	-	-	-	-	-	-	-0,15	4,7 %	$\chi^2(2) = 0,55$	5,9 %	4,5 %	3,7 %
Capital social	-	-	0,11	-	-	-	-	-	15,43 (3,91)	$F(2, 338) = 0,26$	15,58 (3,91)	15,50 (4,02)	15,21 (3,78)
Soutien social	-	-	-	-0,11	-	-	0,10	-	19,46 (1,69)	$F(2, 182) = 1,89^c$	19,14 (2,23)	19,59 (1,02)	19,59 (1,73)

Note : Les variables retenues comme covariables pour les analyses d'impact sont en caractères gras. ^a Toutes variables catégorielles sont binaires (codes 0, 1) et les pourcentages rapportés sont pour la catégorie « 1 » indiquée entre parenthèses par le symbole (+). Com= Communication, Con = Conscience de soi, Cog = Cognition, Phy = Physique, Voc. = Vocabulaire, Fréq litt. = Fréquences des activités de littératie, Langue litt. = Langue des activités de littératie, M = moyenne ; É.-T. = écart-type ; N=342 ; Seuils de signification : *** $p < 1$ % ; ** $p < 5$ % ; * $p < 10$ %.

^b Exceptionnellement, nous retenons cette variable comme covariable, seulement pour les analyses par profile linguistique. ^c Test F robuste à l'hétérogénéité des variances de Welch.

Tableau F.3 : Pertinence des variables linguistiques

Covariable	Corrélations significatives ($p < .10$) avec les variables dépendantes au niveau de base								Tests d'association avec l'appartenance aux groupes expérimentaux				
	Com	Con	Cog	Phy	Voc. expressif	Voc. réceptif	Fréq. litt.	Langue litt.	Paramètres globaux % ou M (É.-T.)	Test d'association avec les groupes	Groupe programme % ou M (É.-T.)	Groupe témoin en garderie % ou M (É.-T.)	Groupe témoin hors garderie % ou M (É.-T.)
Fréq. litt. (Niveau de base)	-	0,10	0,10	-	-	0,10	SO	-0,17	15,12 (2,94)	F(2, 339) = 4,63**	15,37 (2,56)	14,53 (3,00)	15,62 (3,11)
Langue litt. (Niveau de base)	0,45	0,38	-	-	0,36	0,34	-0,17	SO	19,96 (5,97)	F(2, 339) = 3,14**	18,95 (6,47)	19,89 (5,74)	21,00 (5,63)
Langue parlée à l'enfant (mère)	0,45	0,37	-	-	0,37	0,31	-0,15	0,81	4,06 (1,36)	F(2, 336) = 4,57**	3,83 (1,45)	3,99 (1,39)	4,38 (1,17)
Langue parlée à l'enfant (père)	0,37	0,29	-	-	0,30	0,24	-0,09	0,70	3,69 (1,65)	F(2, 324) = 5,23***	3,36 (1,75)	3,62 (1,66)	4,09 (1,46)
Continuum de français parlé par l'enfant	0,52	0,41	-	-	0,42	0,37	-0,10	0,84	3,79 (1,47)	F(2, 339) = 5,02***	3,46 (1,58)	3,79 (1,45)	4,10 (1,34)
Langue de garde de jour (0-12)	0,18	0,11	-	-	-	0,11	-	0,32	2,25 (0,87)	F(2, 337) = 1,21	2,21 (0,86)	2,19 (0,89)	2,36 (0,86)
Langue de garde de jour (13-24)	0,12	-	-	-	-	-	-	0,22	2,25 (0,86)	F(2, 337) = 1,17	2,17 (0,84)	2,23 (0,89)	2,35 (0,84)
Langue de garde de jour (25-36)	-	-	-0,12	-	-	-	-	0,17	2,33 (0,85)	F(2, 337) = 0,27	2,29 (0,85)	2,33 (0,88)	2,37 (0,82)
Type de foyer selon la PLOP (+Endog-français)	0,30	0,26	-	-	0,26	0,21	-0,13	0,69	50,1 %	$\chi^2(2) = 3,81$	46,1 %	47,0 %	57,9 %
Type de foyer selon la langue parlée à l'enfant (+Endog-français) ^a	0,30	0,26	-	-	0,26	0,21	-0,13	0,60	52,2 %	$\chi^2(2) = 4,57$	45,1 %	51,5 %	59,8 %

Note : Les variables retenues comme covariables pour les analyses d'impact sont en caractères gras. ^aToutes variables catégorielles sont binaires (codes 0, 1) et les pourcentages rapportés sont pour la catégorie « 1 » indiquée entre parenthèses et précédée par le symbole (+). Com= Communication, Con = Conscience de soi, Cog = Cognition, Phy = Physique, Voc. = Vocabulaire, Fréq litt. = Fréquences des activités de littératie, Langue litt. = Langue des activités de littératie, M = moyenne ; É. T. = écart-type ; N=342 ; Seuils de signification : *** $p < 1$ % ; ** $p < 5$ % ; * $p < 10$ %.

Tableau F.4 : Pertinence des variables sociolinguistiques

Covariable	Corrélations significatives ($p < .10$) avec les variables dépendantes au niveau de base									Tests d'association avec l'appartenance aux groupes expérimentaux			
	Com	Con	Cog	Phy	Voc. expressif	Voc. réceptif	Fréq. litt.	Langue litt.	Paramètre(s) global(aux) (% ou M (É.-T.))	Test d'association avec les groupes	Groupe programme (% ou M (É.-T.))	Groupe témoin en garderie (% ou M (É.-T.))	Groupe témoin hors garderie (% ou M (É.-T.))
Vitalité	0,26	0,21	-	-	0,20	0,22	-0,17	0,53	15,61 (5,34)	F(2, 288) = 5,01***	15,18 (5,05)	14,88 (5,32)	17,16 (5,41)
Engagement envers la culture francophone	-	-	-	-	-	-	-	0,24	8,16 (3,04)	F(2, 308) = 0,06	8,22 (3,00)	8,08 (2,99)	8,20 (3,16)
Sentiment d'appartenance	0,41	0,35	-	-	0,32	0,35	-0,17	0,62	2,42 (0,75)	F(2, 337) = 2,79*	2,30 (0,82)	2,41 (0,75)	2,55 (0,66)

Note : Les variables retenues comme covariables pour les analyses d'impact sont en caractères gras. Com= Communication, Con = Conscience de soi, Cog = Cognition, Phy = Physique, Voc. = Vocabulaire, Fréq litt. =Fréquences des activités de littératie, Langue litt. = Langue des activités de littératie, M = moyenne ; É-T = écart-type ; N=342 ; Seuils de signification : *** $p < 1$ %; ** $p < 5$ %; * $p < 10$ %.

Tableau F.5 : Pertinence de variables de style parental

Covariable	Corrélations significatives ($p < .10$) avec les variables dépendantes au niveau de base									Tests d'association avec l'appartenance aux groupes expérimentaux			
	Com	Con	Cog	Phy	Voc. expressif	Voc. réceptif	Fréq. litt.	Langue litt.	Paramètres globaux % ou M (É.-T.)	Test d'association avec les groupes	Groupe programme % ou M (É.-T.)	Groupe témoin en garderie % ou M (É.-T.)	Groupe témoin hors garderie % ou M (É.-T.)
Fonctionnement familial	0,16	0,15	-	-	0,14	0,14	-	-	29,75 (3,16)	$F(2, 204) = 1,78^a$	29,17 (4,12)	30,02 (2,57)	29,98 (2,69)
Dépression (mère)	0,15	0,15	-0,10	-0,12	-0,12	-0,12	-	-0,13	10,41 (3,22)	$F(2, 331) = 2,23$	10,71 (3,42)	9,95 (2,70)	10,72 (3,57)
Style parental autoritaire	0,15	0,15	0,15	0,12	0,14	0,15	0,12	-	12,97 (2,10)	$F(2, 336) = 9,39^{***}$	12,25 (2,27)	13,37 (2,05)	13,16 (1,80)
Style parental positif	-	-	-	-	-	-	0,30	-	23,09 (1,90)	$F(2, 336) = 2,35^*$	22,95 (1,98)	22,94 (2,00)	23,42 (1,65)
Empowerment	-	-	-	-	-	-	0,18	-	16,64 (2,31)	$F(2, 289) = 1,41$	16,35 (2,36)	16,63 (2,23)	16,93 (2,36)

Note : Les variables retenues comme covariables pour les analyses d'impact sont en caractères gras. Com= Communication, Con = Conscience de soi, Cog = Cognition, Phy = Physique, Voc. = Vocabulaire, Fréq litt. =Fréquences des activités de littératie, Langue litt. = Langue des activités de littératie, M = moyenne ; É-T = écart-type ; N=342 ; Seuils de signification : *** $p < 1$ % ; ** $p < 5$ % ; * $p < 10$ %.

^a Test F robuste à l'hétérogénéité des variances de Welch.

Tableau F.6 : Pertinence des caractéristiques de l'environnement en garderie

Covariable	Corrélations significatives ($p < .10$) avec les variables dépendantes au niveau de base									Tests d'association avec l'appartenance aux groupes expérimentaux			
	Com	Con	Cog	Phy	Voc. expressif	Voc. réceptif	Fréq. litt.	Langue litt.	Paramètres globaux % ou M (É.-T.)	Test d'association avec les groupes	Groupe programme (% ou M (É.-T.))	Groupe témoin en garderie % ou M (É.-T.)	Groupe témoin hors garderie % ou M (É.-T.)
Dosage – hrs/semaine (niveau de base)	0,21	0,12	-	-	0,10	0,12	-	0,29	30,77 (10,93)	F(1, 233) = 0,00	30,81 (10,31)	30,73 (11,42)	SO
Dosage – hrs/semaine (à 4 mois)	0,26	0,13	-	-	0,12	0,19	-0,15	0,41	27,51 (9,15)	F(1, 230) = 0,48	27,03 (8,63)	27,87 (9,55)	SO
Dosage – hrs/semaine (à 8 mois et 12 mois)	0,18	-	-	-0,12	-	0,14	-	0,26	27,49 (9,43)	F(1, 224) = 0,00	27,48 (9,27)	27,49 (9,58)	SO
Éducatrices :													
Années d'expérience	-	-	0,12	-	-	-	-	-	6,15 (5,40)	F(1, 38) = 0,12	5,69 (3,25)	6,35 (6,14)	SO
Formation spéciale	0,21	0,12	-	-	0,12	0,24	-	0,20	80,5 %	$\chi^2(1) = 0,09$	83,3 %	79,3 %	SO
Éducation (+Au moins un diplôme collégial) ^a	-	-	-	-0,13	-	-0,14	-	-0,12	51,2 %	$\chi^2(1) = 0,01$	50,0 %	51,7 %	SO
Langue maternelle (+Français) ^a	-	-	-	-	-	-	-	-	90,2 %	$\chi^2(1) = 0,92$	83,3 %	93,1 %	SO

Note : Les variables retenues comme covariables pour les analyses d'impact sont en caractères gras. ^aToutes variables catégorielles sont binaires (codes 0, 1) et les pourcentages rapportés sont pour la catégorie « 1 » indiquée entre parenthèses et précédée par le symbole (+). Com= Communication, Con = Conscience de soi, Cog = Cognition, Phy = Physique, Voc. = Vocabulaire, Fréq litt. =Fréquences des activités de littératie, Langue litt. = Langue des activités de littératie, M = moyenne ; É.-T. = écart-type ; N=342 ; Seuils de signification : *** $p < 1$ % ; ** $p < 5$ % ; * $p < 10$ %.

Tableau F.7 : Pertinence des facteurs méthodologiques

Covariable	Corrélations significatives ($p < .10$) avec les variables dépendantes au niveau de base									Tests d'association avec l'appartenance aux groupes expérimentaux			
	Com	Con	Cog	Phy	Voc. expressif	Voc. réceptif	Fréq. litt.	Langue litt.	Paramètres globaux % ou M (É.-T.)	Test d'association avec les groupes	Groupe programme (% ou M (É.-T.))	Groupe témoin en garderie (% ou M (É.-T.))	Groupe témoin hors garderie (% ou M (É.-T.))
Inscription à l'école (2^e année) ^b	SO	SO	SO	SO	SO	SO	SO	SO	SO	$\chi^2(4) = 24,90^{***}$	SO	SO	SO
À temps plein (+ Temps plein) ^a	0,27	0,21	0,28	SO	0,18	SO	-	0,14	36,1 %	$\chi^2(2) = 3,63$	38,9 %	31,0 %	42,5 %
À temps partiel (+Temps partiel) ^a	-0,37	-0,27	-0,15	SO	-0,28	SO	-	-	29,0 %	$\chi^2(2) = 23,23^{***}$	32,2 %	38,0 %	9,2 %
Langue d'administration (+Français) ^a :													
Sondage au niveau de base	0,25	0,26	-	-	0,24	0,19	-0,11	0,47	92,1 %	SO	91,2 %	87,2 %	99,1 %
à 4 mois	0,36	0,31	0,16	-	0,27	0,29	-	0,54	82,7 %	$\chi^2(2) = 1,83$	80,8 %	80,8 %	86,8 %
à 8 mois	0,36	0,23	-	-	0,21	0,25	-	0,57	87,6 %	$\chi^2(2) = 0,60$	85,4 %	88,5 %	88,4 %
à 12 mois	0,37	0,30	-	-	0,30	0,28	-0,12	0,54	86,6 %	$\chi^2(2) = 1,38$	83,9 %	86,0 %	89,4 %
ÉPE-AD au niveau de base	0,44	0,20	-	-	0,20	0,13	-	0,59	87,2 %	$\chi^2(2) = 2,32$	83,0 %	88,5 %	89,6 %
à 4 mois	0,57	0,30	0,24	0,16	0,26	0,27	-0,13	0,62	86,5 %	$\chi^2(2) = 1,61$	86,1 %	89,2 %	83,6 %
à 8 mois	0,65	0,32	0,23	-	0,30	0,33	-	0,59	88,2 %	$\chi^2(2) = 2,66$	87,5 %	91,5 %	84,8 %
à 12 mois	0,63	0,30	0,22	-	0,26	0,44	-0,14	0,49	92,3 %	$\chi^2(2) = 7,84^{**}$	93,7 %	96,1 %	86,7 %
Changement de groupe (+changement) ^a	-0,11	-0,14	-	-	-0,10	-	-	-	4,1 %	SO	5,9 %	3,8 %	2,8 %
Cohorte ^a (+ 1 ^{ère}) ^c	-	-	-	-	-	-	-	-	71,3 %	$\chi^2(2) = 0,08$	70,6 %	72,2 %	71,0 %
Communauté	SO	SO	SO	SO	SO	SO	SO	SO	SO	$\chi^2(6) = 13,30^{**}$	SO	SO	SO
Orléans (+Orléans) ^a	0,26	0,19	-	-	0,19	0,21	0,10	0,21	31,0 %	$\chi^2(2) = 0,94$	29,4 %	29,3 %	34,6 %
Cornwall (+Cornwall) ^a	-0,37	-0,30	-0,10	-0,12	-0,27	-0,37	0,10	-0,39	33,6 %	$\chi^2(2) = 1,65$	38,2 %	33,1 %	29,9 %
Durham (+Durham) ^a	-	-	0,12	0,30	-	-	-	-0,22	10,8 %	$\chi^2(2) = 8,17^{**}$	14,7 %	13,5 %	3,7 %
Edmunston (+Edmunston) ^a	0,13	0,13	-	-	0,11	0,13	-0,23	0,36	24,6 %	$\chi^2(2) = 5,66^*$	17,6 %	24,1 %	31,8 %

Note : Les variables retenues comme covariables pour les analyses d'impact sont en caractères gras. ^aToutes variables catégorielles sont binaires (codes 0, 1) et les pourcentages rapportés sont pour la catégorie « 1 » indiquée entre parenthèses et précédée par le symbole (+). Com= Communication, Con = Conscience de soi, Cog = Cognition, Phy = Physique, Voc. = Vocabulaire, Fréq litt. = Fréquences des activités de littératie, Langue litt. = Langue des activités de littératie, M = moyenne ; É.-T. = écart-type ; N=342 ; Seuils de signification : *** $p < 1$ % ; ** $p < 5$ % ; * $p < 10$ % . ^b Enfants tels que regroupés à la cinquième période d'évaluation. ^c Exceptionnellement, nous retenons la cohorte comme covariable dans les analyses d'impact.

Tableau F.8 : Tests robustes de l'instabilité des effets de variables confusionnelles dans le temps

Covariable	Variables dépendantes							
	Com	Con	Cog	Phy	Voc. expressif	Voc. réceptif	Fréq. Litt.	Langue litt.
Communauté - Wald F(9,14)	2,88**	0,57	0,56	1,16	0,37	3,28**	2,26*	2,03
Cohorte - Wald F(3, 20)	2,61*	1,25	0,70	15,79***	0,62	3,22**	0,30	8,87***
Fratrie plus jeune - Wald F(3, 20)	0,52	0,18	3,36**	4,24**	0,16	0,77	0,06	3,56**
Taille du ménage - Wald F(3, 20)	1,45	1,98	1,04	0,55	1,22	0,93	2,31	0,64
Fréq. litt. (Niveau de base) - Wald F(3, 20)	0,86	2,01	3,12**	0,26	0,69	0,42	40,92***	0,67
Langue litt. (Niveau de base) - Wald F(3, 20)	5,25***	2,41*	0,56	1,01	1,42	0,30	1,24	15,96***
Langue parlée à l'enfant (mère) - Wald F(3, 20)	8,32***	6,62***	1,45	1,34	1,94	0,89	4,39**	2,72*
Langue parlée à l'enfant (père) - Wald F(3, 20)	6,73***	1,19	0,60	0,37	0,32	2,73*	4,81**	5,84***
Continuum de français parlé par l'enfant - Wald F(3, 20)	1,76	0,29	0,93	0,42	0,45	1,02	4,73**	7,74***
Âge de l'enfant (en mois) - Wald F(3, 20)	2,94*	3,01*	0,51	0,61	0,88	1,74	2,29	2,29
Vitalité - Wald F(3, 20)	0,05	2,95*	1,10	0,78	0,49	0,37	0,60	2,57*
Style autoritaire - Wald F(3, 20)	1,36	1,46	3,55**	0,18	0,31	2,85*	1,64	1,50

Note : Les tests rapportés dans le tableau découlent d'une spécification dont : des dichotomies de « temps », l'effet principal des covariables et l'interaction des covariables avec les dichotomies de temps. Les termes d'erreurs sont regroupés par garderie et les données portent sur les quatre premières évaluations. Com= Communication, Con = Conscience de soi, Cog = Cognition, Phy = Physique, Voc. = Vocabulaire, Fréq litt. =Fréquences des activités de littératie, Langue litt. = Langue des activités de littératie Seuils de signification : *** $p < 1\%$; ** $p < 5\%$; * $p < 10\%$.

Annexe G : Résultats de l'analyse des échelles ÉVIP-R et EOWPVT-F

Tableau G.1 : Impact du programme sur le vocabulaire français : Échelles normalisées de Vocabulaire réceptif (ÉVIP-R) et Vocabulaire expressif (EOWPVT-F)

Type de différence	Vocabulaire réceptif (ÉVIP-R)				Vocabulaire expressif (EOWPVT-F)			
	Non-ajusté (N = 333)		Ajusté (N = 328)		Non-ajusté (N = 328)		Ajusté (N = 321)	
	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.
G1 vs G2								
Score brute	-3,64	2,43	-3,74**	1,33	-3,09**	1,48	-2,66	1,64
Score standardisé	-0,20	0,13	-0,20**	0,07	-0,23**	0,11	-0,19	0,12
G1 vs G3								
Score brute	-2,10	3,63	-3,59**	1,64	-1,42	2,91	-3,37	2,48
Score standardisé	-0,11	0,20	-0,20**	0,09	-0,10	0,21	-0,25	0,18

*Note : Les valeurs négatives des estimations DD pour toutes les comparaisons entre les groupes représentent un effet traitement positif (c.-à-d. un avantage pour le groupe programme). Les scores standardisés sont des données brutes transformées en score Z ayant une moyenne de zéro et un écart-type de 1. L'affectation aux groupes à l'étude au moment de l'administration des échelles est utilisée pour les comparaisons rapportées dans ce tableau. L'ÉVIP-R a été administré à 20 mois et l'EOWPVT-F à 24 mois. Les erreurs types robuste de Huber-White sont rapportées avec groupement par garderie. Le seuil de signification est indiqué à *** $p < 1\%$, ** $p < 5\%$; * $p < 10\%$. Diff = différence, E.T. = erreur type.*

Annexe H : Résultats de l'analyse des échelles de l'ÉPE-AD (deuxième année)

Tableau H.1 : Impact du programme non-ajustés sur la préparation à la scolarisation lors de la deuxième année – Scores standardisés

Type de différence	Indicateur de la préparation à la scolarisation							
	Communication		Conscience de soi		Cognition		Vocabulaire expressif	
	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.
G1 vs G2								
Niveau de base	0,44	0,31	0,35	0,25	0,26**	0,12	0,42	0,26
12 mois (DD)	-0,34***	0,11	-0,22*	0,11	-0,16	0,14	-0,26**	0,11
16 mois (DD)	-0,23**	0,09	-0,15*	0,07	-0,15	0,19	-0,12	0,12
20 mois (DD)	-0,32	0,20	-0,35***	0,09	-0,18*	0,09	-0,38***	0,11
24 mois (DD)	-0,38***	0,12	-0,36***	0,10	-0,28*	0,14	-0,40***	0,12
G1 vs G3								
Niveau de base	0,24	0,32	0,21	0,28	0,07	0,10	0,30	0,26
12 mois (DD)	-0,38**	0,17	-0,07	0,07	-0,07	0,13	-0,08	0,13
16 mois (DD)	-0,21	0,16	0,20*	0,11	0,12	0,22	0,20	0,19
20 mois (DD)	-0,31	0,20	-0,32*	0,18	-0,26*	0,15	-0,38	0,26
24 mois (DD)	-0,25*	0,13	-0,49***	0,16	-0,25	0,25	-0,50**	0,21
Groupe x temps Wald F	3,02**		5,69***		2,99**		2,28*	

Note : Contrairement aux figures, les valeurs négatives des estimés DD présentées dans ce tableau représentent un effet traitement positif (c.-à-d. un avantage pour le groupe programme). Degrés de liberté du test Wald F sont 8 et 15. Les erreurs types robuste de Huber-White sont rapportées avec groupement par garderie. Le seuil de signification est indiqué à *** $p < 1\%$, ** $p < 5\%$; * $p < 10\%$. Diff = différence, E.T. = erreur type.

Tableau H.2 : Impact ajusté du programme sur la préparation à la scolarisation lors de la deuxième année – Scores standardisés

Type de différence	Indicateur de la préparation à la scolarisation							
	Communication		Conscience de soi		Cognition		Vocabulaire expressif	
	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.
G1 vs G2								
Niveau de base	0,30***	0,08	0,20**	0,08	0,09	0,09	0,28***	0,09
12 mois (DD)	-0,41***	0,11	-0,19*	0,10	-0,09	0,11	-0,29**	0,11
16 mois (DD)	-0,17	0,11	-0,18**	0,08	0,01	0,14	-0,11	0,11
20 mois (DD)	-0,20*	0,11	-0,31***	0,08	-0,01	0,11	-0,35***	0,08
24 mois (DD)	-0,27**	0,13	-0,31***	0,08	-0,13	0,10	-0,37***	0,11
G1 vs G3								
Niveau de base	-0,01	0,05	-0,04	0,10	-0,06	0,09	0,01	0,08
12 mois (DD)	-0,38***	0,09	-0,02	0,12	-0,09	0,08	-0,07	0,11
16 mois (DD)	-0,23	0,15	0,16	0,11	0,22	0,14	0,17	0,17
20 mois (DD)	-0,18	0,15	-0,27	0,16	-0,06	0,09	-0,35	0,20
24 mois (DD)	-0,07	0,15	-0,38**	0,15	-0,02	0,06	-0,40**	0,16
Groupe x temps Wald F	3,18**		14,30***		2,65**		5,19***	

*Note : Contrairement aux figures, les valeurs négatives des estimés DD présentés dans ce tableau représentent un effet traitement positif (c.-à-d. un avantage pour le groupe programme). Degrés de liberté du test Wald F sont 8 et 15. Les erreurs types robuste de Huber-White sont rapportées avec groupement par garderie. Le seuil de signification est indiqué à *** $p < 1\%$, ** $p < 5\%$; * $p < 10\%$. Diff = différence, E.T. = erreur type.*

Annexe I : Résultats de l'analyse des impacts en fonction du profil linguistique

Tableau I.1 : Tests de l'hypothèse que le profil linguistique du foyer agit comme modérateur de l'effet programme lors de la deuxième année

Test de modération	Indicateur de la préparation à la scolarisation							
	Communication		Conscience de soi		Cognition		Vocabulaire expressif	
	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.
G1 vs G2								
Temps 4 (DDD)	0,22	0,15	0,00	0,19	-0,55***	0,19	0,12	0,24
Temps5 (DDD)	0,52***	0,18	0,17	0,20	-0,30	0,22	0,31	0,23
Temps 6 (DDD)	0,26	0,22	0,03	0,21	-0,31	0,32	0,06	0,22
Temps 7 (DDD)	-0,20	0,25	-0,15	0,21	-0,25	0,29	-0,14	0,24
G1 vs G3								
Temps 4 (DDD)	0,43**	0,19	-0,40**	0,16	-0,67***	0,21	-0,09	0,28
Temps5 (DDD)	0,56*	0,31	-0,21	0,18	-0,56	0,35	0,17	0,35
Temps 6 (DDD)	0,64**	0,30	0,43	0,38	0,13	0,35	0,65	0,55
Temps 7 (DDD)	0,00	0,36	0,26	0,43	-0,17	0,28	0,46	0,60

Note : Les estimés DDD représentent la différence dans l'impact du programme pour les enfants issus de foyers à forte et à faible exposition. La catégorie de référence est « faible exposition ». *L'interprétation des effets DDD se fait de façon suivante : (a) les estimés positifs indiquent que l'effet du programme testé diminue avec une grande exposition au français (b) les estimés négatifs indiquent que l'effet du programme testé est accentué avec une plus forte exposition au français.* Les erreurs types robustes de Huber-White sont rapportées avec groupement d'erreurs par garderie. Le seuil de signification est rapporté à *** $p < 1\%$, ** $p < 5\%$; * $p < 10\%$. Diff = différence, E.T. = erreur type.

Tableau I.2 : Impact du programme testé selon le profil linguistique du foyer – (2e année) – Scores ajustés

		Indicateur de la préparation à la scolarisation							
		<i>Foyer à faible exposition au français</i>							
Type de différence	Communication		Conscience de soi		Cognition		Vocabulaire expressif		
	Diff.	ET	Diff.	ET	Diff.	ET	Diff.	ET	
G1 vs G2									
	Niveau de base	0,40***	0,10	0,27**	0,11	0,11	0,14	0,34***	0,10
	Temps 4 (DD)	-0,50***	0,11	-0,21	0,14	0,15	0,15	-0,35**	0,13
	Temps5 (DD)	-0,39***	0,11	-0,24*	0,13	0,12	0,19	-0,31**	0,13
	Temps 6 (DD)	-0,29*	0,15	-0,34**	0,15	0,13	0,20	-0,39***	0,13
	Temps 7 (DD)	-0,14	0,19	-0,25	0,16	-0,04	0,22	-0,30*	0,15
G1 vs G3									
	Niveau de base	-0,23**	0,10	-0,29**	0,12	-0,36**	0,16	-0,10	0,18
	Temps 4 (DD)	-0,59***	0,09	0,25*	0,12	0,28	0,18	-0,00	0,12
	Temps5 (DD)	-0,48**	0,20	0,37**	0,14	0,59**	0,25	0,06	0,15
	Temps 6 (DD)	-0,46*	0,26	-0,50*	0,28	-0,11	0,23	-0,70*	0,36
	Temps 7 (DD)	0,00	0,34	-0,48	0,30	0,11	0,18	-0,59	0,40
		<i>Foyer à forte exposition au français</i>							
Type de différence	Communication		Conscience de soi		Cognition		Vocabulaire expressif		
	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.	
G1 vs G2									
	Niveau de base	0,19	0,11	0,15	0,12	0,15	0,11	0,24	0,15
	Temps 4 (DD)	-0,30**	0,14	-0,21	0,14	-0,40***	0,12	-0,23	0,19
	Temps5 (DD)	0,09	0,15	-0,07	0,12	-0,17	0,16	0,00	0,17
	Temps 6 (DD)	-0,07	0,16	-0,31**	0,12	-0,19	0,17	-0,33**	0,14
	Temps 7 (DD)	-0,37**	0,16	-0,40***	0,11	-0,29**	0,13	-0,44**	0,17
G1 vs G3									
	Niveau de base	0,09	0,08	0,05	0,14	0,12	0,13	0,07	0,17
	Temps 4 (DD)	-0,19	0,13	-0,15	0,18	-0,40***	0,10	-0,09	0,22
	Temps5 (DD)	0,03	0,18	0,16	0,16	0,03	0,20	0,23	0,29
	Temps 6 (DD)	0,12	0,12	-0,07	0,18	0,02	0,16	-0,04	0,27
	Temps 7 (DD)	-0,05	0,09	-0,22	0,19	-0,06	0,13	-0,13	0,25

*Note : Contrairement aux figures, les valeurs négatives des effets DD rapportés dans ce tableau comparant les groupes DD représentent un effet traitement positif (c.-à-d. un avantage pour le groupe programme). Le Type de foyer est construit à partir de la langue habituellement parlée par l'enfant à sa mère croisée avec la langue habituellement parlée par l'enfant à son père) : foyer à forte exposition au français, autre. Les erreurs types robuste de Huber-White sont rapportées avec groupement des erreurs par garderie. Le seuil de signification est indiqué à *** $p < 1\%$, ** $p < 5\%$; * $p < 10\%$. Diff = différence, E.T. = erreur type.*

Tableau I.3 : Impact du programme testé sur le vocabulaire selon le profil linguistique : Échelles normalisées de vocabulaire réceptif (ÉVIP-R) et expressif (EOWPVT)

Indicateur de la préparation à la scolarisation									
<i>Foyer à faible exposition au français</i>									
Type de différence	Vocabulaire réceptif (ÉVIP-R)				Vocabulaire expressif (EOWPVT)				
	Modèle initial		Modèle ajusté		Modèle initial		Modèle ajusté		
	Diff.	ET	Diff.	ET	Diff.	ET	Diff.	ET	
G1 vs G2									
Score brute	-5,96***	1,58	-6,40***	1,39	-2,44	1,62	-2,64*	1,42	
Score standardisé	-0,33***	0,09	-0,36***	0,08	-0,18	0,12	-0,20*	0,11	
G1 vs G3									
Score brute	-2,82**	1,31	-3,41**	1,53	0,17	2,01	-0,42	1,44	
Score standardisé	-0,16**	0,07	-0,19***	0,09	0,01	0,15	-0,03	0,11	
<i>Foyer à forte exposition au français</i>									
Type de différence	Vocabulaire réceptif (ÉVIP-R)				Vocabulaire expressif (EOWPVT)				
	Modèle initial		Modèle ajusté		Modèle initial		Modèle ajusté		
	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.	
G1 vs G2									
Score brute	-0,58	2,98	-0,79	2,23	-2,52	1,49	-2,33	1,94	
Score standardisé	-0,03	0,17	-0,04	0,13	-0,19	0,11	-0,17	0,14	
G1 vs G3									
Score brute	-1,95	5,92	-3,10	3,80	-5,89	4,98	-6,37	3,89	
Score standardisé	-0,11	0,33	-0,17	0,21	-0,44	0,37	-0,47	0,29	

*Note : Contrairement aux figures, les valeurs négatives des effets présentés dans ce tableau représentent un effet traitement positif (c.-à-d. un avantage pour le groupe programme). N = 321 pour le Vocabulaire expressif (EOWPVT) et N = 328 pour le Vocabulaire réceptif (ÉVIP-R). Les score standardisés sont des données brutes transformés en score Z ayant une moyenne de zéro et un écart-type de 1. L'affectation aux groupes à l'étude lors de l'administration des échelles est utilisée pour les comparaisons rapportées dans ce tableau. L'ÉVIP-R a été administré à 20 mois et l'EOWPVT à 24 mois. Les erreurs types robuste de Huber-White sont rapportées avec groupement par garderie. Le seuil de signification est indiqué à *** $p < 1\%$, ** $p < 5\%$; * $p < 10\%$. Diff = différence, E.T. = erreur type.*

Annexe J : Résultats de l'analyse de la fréquence et de la langue des activités de littératie

Tableau J.1: Impact du programme testé sur la fréquence des activités de littératie ainsi que la langue utilisée lors de ces activités (non-ajusté)

Type de différence	Comportement des parents à la maison							
	Fréquence des activités				Langue utilisée			
	Score brut		Score standardisé		Score Brut		Score standardisé	
	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.
G1 vs G2								
Niveau de base	-.918*	.488	-.312*	.166	.933	2.151	.156	.360
4 mois (DD)	-.005	.673	.001	.228	-.987	.616	-.166*	.093
8 mois (DD)	-.007	.438	.001	.148	-.580	.713	-.089	.099
12 mois (DD)	-.374	.633	-.121	.214	-.775	.976	-.123	.125
16 mois (DD)	.106	.750	.047	.248	-.724	1.677	-.131	.271
20 mois (DD)	.290	.704	.108	.232	-1.230	1.398	-.198	.249
24 mois (DD)	.289	.759	.096	.259	-1.313	1.325	-.214	.230
G1 vs G3								
Niveau de base	0.244	.674	.083	.229	2.051	2.244	.344	.376
4 mois (DD)	-1.262	.895	-.426	.303	-1.662***	.577	-.274***	.092
8 mois (DD)	-.915	.624	-.308	.212	-2.000**	.782	-.334**	.125
12 mois (DD)	-1.290*	.743	-.433*	.252	-1.322	.859	-.191	.116
16 mois (DD)	-.771	1.038	-.255	.346	-3.381*	1.741	-.506*	.282
20 mois (DD)	-1.001	.924	-.329	.308	-3.079**	1.331	-.487*	.242
24 mois (DD)	-.884	.961	-.302	.327	-2.893**	1.344	-.472*	.235
Groupe x temps Wald F	3.04**		2.92**		3.407**		4.78***	

Note : Degrés de liberté pour le test F de Wald sont 12 et 11. Les erreurs types robuste de Huber-White sont rapportées avec groupement des erreurs par garderie. Le seuil de signification est indiqué à *** $p < 1\%$, ** $p < 5\%$; * $p < 10\%$. Diff = différence, E.T. = erreur type.

Tableau J.2: Impact ajusté du programme testé sur la Fréquence des activités de littératie ainsi que la Langue des activités de littératie

Type de différence	Comportement des parents à la maison							
	Fréquence des activités				Langue utilisée			
	Score brut		Score standardisé		Score brut		Score standardisé	
	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.	Diff.	E.T.
G1 vs G2								
Niveau de base	-0.009	.016	-0.003	.005	-0.016	.028	-0.002	.005
4 mois (DD)	-0.301	.298	-0.102	.101	-0.895**	.401	-0.161**	.072
8 mois (DD)	-0.569**	.267	-0.191**	.090	-0.562	.363	-0.092	.067
12 mois (DD)	-0.909**	.353	-0.304**	.118	-0.563	.401	-0.118	.084
16 mois (DD)	-0.492	.373	-0.161	.122	-1.361	.933	-0.093	.123
20 mois (DD)	-0.383	.513	-0.124	.167	-1.162**	.477	-0.185**	.069
24 mois (DD)	-0.441	.342	-0.151	.117	-0.234	.429	-0.035	.064
G1 vs G3								
Niveau de base	0.006	.035	0.002	.012	0.027	.043	0.007	.007
4 mois (DD)	-1.041***	.249	-0.351***	.084	-1.309**	.511	-0.235**	.092
8 mois (DD)	-0.683***	.197	-0.229***	.066	-1.38***	.418	-0.249***	.074
12 mois (DD)	-1.093***	.254	-0.366***	.085	-0.49	.517	-0.103	.108
16 mois (DD)	-0.63	.397	-0.206	.130	-2.83***	.628	-0.273***	.086
20 mois (DD)	-0.848*	.488	-0.276*	.159	-1.021	.673	-0.207**	.089
24 mois (DD)	-0.723*	.364	-0.248*	.125	-0.396	.628	0.016	.101
Groupe x temps Wald F	4.382**		4.383**		7.429***		4.084**	

Note : Degrés de liberté pour le test F de Wald sont 12 et 11. Les erreurs types robuste de Huber-White sont rapportées avec groupement des erreurs par garderie. Le seuil de signification est indiqué à *** $p < 1\%$; ** $p < 5\%$; * $p < 10\%$. Diff = différence, E.T. = erreur type.

SOCIÉTÉ DE RECHERCHE SOCIALE APPLIQUÉE

BUREAU D'OTTAWA

55, rue Murray, bureau 400
Ottawa (Ontario) K1N 5M3
Tél. : 613.237.4311
Télééc. : 613.237.5045

BUREAU DE TORONTO

481, avenue University, bureau 705
Toronto (Ontario) M5G 2E9
Tél. : 416.593.0445
Télééc. : 647.725.6293

BUREAU DE VANCOUVER

128, rue Pender Ouest, bureau 301
Vancouver (Colombie-Britannique) V6B 1R8
Tél. : 604.601.4070
Télééc. : 604.601.4080