

**Université de Montréal**

**Modélisation des liens entre des variables  
associées à l'environnement scolaire et la performance  
à l'écrit en français des finissants du secondaire au Québec**

**Par  
Patrick Charles**

**Département d'administration et fondements de l'éducation  
Faculté des sciences de l'éducation**

**Thèse présentée à la Faculté des études supérieures  
en vue de l'obtention du grade de Philosophiae Doctor  
en mesure et évaluation en éducation**

**Mai 2012**

**Copyright, Patrick Charles, 2012**

Université de Montréal  
Faculté des études supérieures

Cette thèse intitulée :

*Modélisation des liens entre des variables associées à l'environnement scolaire  
et la performance à l'écrit en français des finissants du secondaire au Québec*

présentée par :

Patrick Charles

a été évaluée par un jury composé des personnes suivantes :

---

Président-rapporteur

---

Directeur de recherche

---

Membre du jury

---

Membre du jury

---

Examineur externe

---

Représentant du doyen

## SOMMAIRE

Le ministère de l'Éducation, du Loisir et du Sport (MELS) publie tous les ans des indicateurs (MELS, 2007) qui traitent de plusieurs aspects du système scolaire québécois. Quoique le MELS insiste sur l'obligation pour les écoles primaires et secondaires d'obtenir des résultats probants en termes d'« efficacité », les indicateurs utilisés pour apprécier la performance des écoles ne sont pas nécessairement conçus à cette fin.

Une étude du MELS rapporte que les élèves de 5<sup>e</sup> secondaire éprouvent de la difficulté à obtenir de bons résultats à deux critères de correction (syntaxe et ponctuation, orthographe), parmi les six de l'épreuve unique de français écrit (MELS, 2008). Ce fait nous amène à nous intéresser à l'étude de la modélisation des liens entre des facteurs associés à l'environnement scolaire et les résultats des élèves de la région métropolitaine de Montréal, en ce qui a trait à ces deux critères pour les cohortes des années 2006, 2007 et 2008. Nous procédons d'abord à des analyses descriptives des variables pour chacune des trois populations. Nous poursuivons l'analyse en effectuant plusieurs modélisations multiniveaux multivariées des deux critères en fonction de variables indépendantes, caractéristiques de l'élève et de l'école.

Les résultats de la présente recherche indiquent une relative stabilité dans la performance des élèves pour les trois années, avec une légère amélioration de la performance pour la cohorte de 2007, et qui se maintient pour la cohorte de 2008. Les élèves du secteur privé obtiennent de meilleurs résultats que ceux du public. Le résultat des filles est supérieur à celui des garçons et les élèves de langue maternelle française obtiennent de meilleurs résultats que ceux de langues maternelles différentes du français. Il importe cependant d'apporter quelques nuances dans l'interprétation de ces résultats.

En outre, la part de variance dans la performance des élèves attribuable à l'élève est de l'ordre de 75,0 % et de l'ordre de 25,0 % à l'école. On note un effet différentiel des facteurs d'élève et d'école selon le critère considéré. Les variables caractéristiques de l'élève comptent pour 13,9 % de la variance totale de la performance des élèves en syntaxe et ponctuation et pour 9,8 % en orthographe. Les variables caractéristiques de l'école comptent pour 3,7 % de la variance totale de la performance des élèves en syntaxe et ponctuation et pour 6,4 % en orthographe. Certains facteurs d'école, comme la taille, la mixité ne semblent pas présenter un lien significatif avec la performance des élèves.

**Mots clés :** Indicateurs, modélisation, analyse multiniveaux multivariée.

## SUMMARY

The *ministère de l'Éducation, du Loisir et du Sport* (MELS) of Quebec publishes every year educational indicators which describe many aspects of the school system. Although the MELS insists for elementary and high schools to have good results in terms of the effectiveness of their performance, the indicators that are used presently by the school's system are not designed to appreciate their performance.

A research by the MELS shows that secondary five high school students have some difficulty with two criteria (syntax and punctuation, spelling) out of the six for the French language examination (MELS, 2008). The actual research focuses on the modeling of links between school environment and student performances in the Montreal area over a period of three years. By using data from the MELS for years 2006, 2007 and 2008, this study investigates the relations between variables characteristics of student and school with the student performances. The analysis used descriptive statistics of the variables for each of the three years. We complete the analysis by doing some multilevel multivariate analysis.

The results of this study indicated that there is a relative stability in student performances over the three years with a slight improvement in 2007 which continue in 2008. Students from the private sector performed better than those of the public sector. The girls continuously have better results than the boys over the three years period. Students with French as their mother's tongue do have better results than those with mother's tongue different than French. Nevertheless, some nuances are necessary for better understanding the interpretation of the results.

Moreover, the part of variation in student performances is responsible for about 75,0 % to student level and for about 25,0 % to school level. We note some differential effects of student and school factors on the criteria considered. The student variables are responsible for about 13,9 % of the total variance in syntax and punctuation and for about 9,8 % in spelling. School variables are responsible for 3,7 % for the total variance in student performance for syntax and punctuation and for about 6,4 % for spelling. Some school factors, like the size of the school, the mix sex school do not seem to have any significant link with students' performances.

**Key words:** Indicators, statistics, multilevel multivariate analysis, modeling.

## TABLE DES MATIÈRES

<b>SOMMAIRE .....</b>	<b>III</b>
<b>SUMMARY .....</b>	<b>IV</b>
<b>LISTE DES TABLEAUX.....</b>	<b>XI</b>
<b>LISTE DES FIGURES .....</b>	<b>XIII</b>
<b>LISTE DES ABRÉVIATIONS, SIGLES ET ACRONYMNES.....</b>	<b>XIV</b>
<b>REMERCIEMENTS .....</b>	<b>XV</b>
<b>INTRODUCTION .....</b>	<b>1</b>
<b>CHAPITRE I.....</b>	<b>5</b>
<b>LA PROBLÉMATIQUE.....</b>	<b>5</b>
1.1 LE DÉVELOPPEMENT DE LA RECHERCHE SUR L'« EFFICACITÉ » DES ÉCOLES .....	5
1.2 DES ÉLÉMENTS CONTEXTUELS DU SYSTÈME ÉDUCATIF QUÉBÉCOIS .....	10
1.2.1 Le contexte historique.....	10
1.2.2 Le contexte législatif .....	11
1.2.3 Le contexte démographique .....	12
1.2.4 Le contexte financier.....	12
1.3 LE CONTEXTE DE L'EFFICACITÉ DE L'ÉCOLE .....	13
1.4 LA PERTINENCE DE LA RECHERCHE .....	15
1.5 LE PROBLÈME DE RECHERCHE.....	18
1.6 L'OBJECTIF DE LA RECHERCHE .....	22
<b>CHAPITRE II.....</b>	<b>24</b>
<b>LE CADRE THÉORIQUE .....</b>	<b>24</b>
2.1 LES INDICATEURS ÉDUCATIONNELS .....	24
2.1.1 Les indicateurs des « effets écoles » .....	26
2.2 LE RAPPORT COLEMAN (1966).....	27

2.2.1	Les éléments déterminants du Rapport Coleman.....	28
2.2.2	Des critiques du Rapport Coleman .....	31
2.3	LA RECENSION DES ÉCRITS SUR L'« EFFICACITÉ » DES ÉCOLES EN FILIATION AVEC LE RAPPORT COLEMAN.....	35
2.3.1	The Plowden Report (Peaker, 1967) .....	35
2.3.2	L'étude de Ainsworth et Batten (1974).....	37
2.3.3	L'étude de Lee et Bryk (1898).....	38
2.3.4	L'étude de Hanushek (1989).....	41
2.3.5	L'étude de Hedges, Laine et Greenwald (1994) .....	42
2.3.6	L'étude de Opdenakker et Van Damme (2000) .....	43
2.3.7	L'étude de Rumberger et Palardy (2005).....	44
2.3.8	L'étude de Konstantopoulos (2006) .....	47
2.3.9	L'étude de Petrucci (2010).....	49
2.3.10	L'étude de Blais (2003) .....	51
2.3.11	L'étude de McAndrew, Ledent et Ait-Said (2005) .....	54
2.3.12	L'étude de Ouedraogo (2006).....	56
2.4	SYNTHÈSE DES RECHERCHES PRÉSENTÉES.....	57
2.5	RAPPEL SUR LES OBJECTIFS DE LA RECHERCHE .....	58
	<b>CHAPITRE III.....</b>	<b>59</b>
	<b>LA MÉTHODOLOGIE.....</b>	<b>59</b>
3.1	LES SOURCES DE DONNÉES.....	59
3.1.1	Les données secondaires .....	60
3.2	LES POPULATIONS À L'ÉTUDE .....	61
3.3	L'INSTRUMENTATION .....	62
3.3.1	Les variables dépendantes.....	63
3.3.2	Les formes et les conditions de passation de l'épreuve.....	65
3.3.3	La grille d'évaluation de l'épreuve considérée dans cette étude.....	66
3.3.4	Les variables caractéristiques de l'élève.....	67
3.3.5	Les variables caractéristiques de l'école.....	69
3.3.6	Les variables caractéristiques de la classe.....	72
3.3.7	La liste des variables .....	73
3.4	LES MÉTHODES D'ANALYSE DES DONNÉES.....	74
3.4.1	Les modèles .....	75
3.4.2	La modélisation multiniveaux multivariée.....	76
3.4.3	Les précisions sur les étapes d'analyse.....	79

3.4.4	Les considérations éthiques .....	80
<b>CHAPITRE IV .....</b>		<b>82</b>
<b>LES ANALYSES DESCRIPTIVES DES DONNÉES .....</b>		<b>82</b>
4.1	LA DESCRIPTION DES DONNÉES DES TROIS COHORTES .....	83
4.2	LES RELATIONS ENTRE LES VARIABLES CARACTÉRISTIQUES DES ÉLÈVES ET LEUR PERFORMANCE .....	87
4.2.1	La relation entre l'âge et la performance des élèves .....	88
4.2.2	La relation entre le genre et la performance des élèves .....	90
4.2.3	La relation entre le genre, l'âge et la performance des élèves.....	91
4.2.4	La relation entre la langue maternelle et la performance des élèves.....	92
4.2.5	La relation entre l'âge, la langue maternelle et la performance des élèves.....	93
4.3	LES RELATIONS ENTRE LES VARIABLES CARACTÉRISTIQUES DE L'ÉCOLE ET LA PERFORMANCE DES ÉLÈVES.....	94
4.3.1	La performance des élèves selon le réseau .....	94
4.3.2	La performance des élèves selon l'âge et le réseau .....	95
4.3.3	La performance des élèves selon la mixité .....	96
4.3.4	La performance des élèves selon la taille .....	97
4.3.5	La performance des élèves selon la proportion des élèves de l'école de langue maternelle française.....	98
4.4	LA DESCRIPTION DES DONNÉES DES SOUS-POPULATIONS D'ÉLÈVES SANS LE NOMBRE ÉLEVÉ D'ERREURS .....	99
4.5	LES RELATIONS ENTRE LES VARIABLES CARACTÉRISTIQUES DES ÉLÈVES ET LA PERFORMANCE DES ÉLÈVES SANS LE NOMBRE ÉLEVÉ D'ERREURS.....	102
4.5.1	La relation entre l'âge et la performance des élèves sans le nombre élevé d'erreurs.....	102
4.5.2	La relation entre le genre et la performance des élèves sans le nombre élevé d'erreurs.....	103
4.5.3	La relation entre la langue maternelle et la performance des élèves sans le nombre élevé d'erreurs.....	104
4.6	LES RELATIONS ENTRE LES VARIABLES CARACTÉRISTIQUES DE L'ÉCOLE ET LA PERFORMANCE DES ÉLÈVES SANS LE NOMBRE ÉLEVÉ D'ERREURS.....	105

4.6.1	La relation entre le réseau d'appartenance de l'école et la performance des élèves sans le nombre élevé d'erreurs.....	105
4.6.2	La relation entre la mixité de l'école et la performance des élèves sans le nombre élevé d'erreurs.....	105
4.6.3	La relation entre la taille de l'école et la performance des élèves sans le nombre élevé d'erreurs.....	106
4.6.4	La relation entre la proportion des élèves de l'école de langue maternelle française et la performance des élèves sans le nombre élevé d'erreurs.....	107
4.7	L'ANALYSE DE LA PERFORMANCE DES ÉLÈVES EN DIFFICULTÉ SELON LES VARIABLES CARACTÉRISTIQUES DES ÉLÈVES.....	107
4.7.1	La relation entre l'âge et la performance des élèves en difficulté .....	108
4.7.2	La relation entre le genre et la performance des élèves en difficulté.....	108
4.7.3	La relation entre la langue maternelle et la performance des élèves en difficulté.....	109
4.8	L'ANALYSE DE LA PERFORMANCE DES ÉLÈVES EN DIFFICULTÉ SELON LES VARIABLES CARACTÉRISTIQUES DE L'ÉCOLE .....	110
4.8.1	La relation entre le réseau et la performance des élèves en difficulté.....	110
4.8.2	La relation entre la mixité et la performance des élèves en difficulté.....	111
4.8.3	La relation entre la proportion d'élèves de l'école de langue maternelle française et la performance des élèves en difficulté .....	111
4.9	UNE BRÈVE SYNTHÈSE DES RÉSULTATS DE L'ANALYSE DES CRIPTIVE.....	111
<b>CHAPITRE V .....</b>		<b>113</b>
<b>LES ANALYSES MULTINIVEAUX MULTIVARIÉES .....</b>		<b>113</b>
5.1	LES MODÈLES À UN NIVEAU À VIDE.....	113
5.1.1	Le modèle 1 à vide pour l'ensemble des trois cohortes .....	114
5.1.2	Le modèle 4 vide pour les sous-populations d'élèves sans le nombre élevé d'erreurs.....	116
5.2	LES MODÈLES À DEUX NIVEAUX À VIDE .....	118
5.2.1	Le modèle 2 à vide pour l'ensemble des trois cohortes .....	119
5.2.2	Le modèle 5 à vide pour les sous-populations d'élèves sans le nombre élevé d'erreurs.....	121

5.3	LES MODÈLES À TROIS NIVEAUX À VIDE .....	123
5.3.1	Le modèle 3 à trois niveaux à vide des trois cohortes .....	123
5.3.2	Le modèle 6 à trois niveaux à vide pour les sous-populations d'élèves sans le nombre élevé d'erreurs.....	127
5.4	LA MODÉLISATION DES VARIABLES CARACTÉRISTIQUES DES ÉLÈVES .....	130
5.4.1	La modélisation des variables caractéristiques des élèves pour les trois cohortes.....	130
5.4.2	La modélisation des variables caractéristiques de l'élève pour les sous-populations d'élèves sans le nombre élevé d'erreurs..	132
5.5	LA MODÉLISATION DES VARIABLES CARACTÉRISTIQUES DE L'ÉLÈVES ET DE L'ÉCOLES .....	133
5.5.1	La modélisation des variables caractéristiques des écoles pour l'ensemble des trois cohortes .....	134
5.5.2	La modélisation des variables caractéristiques des écoles pour les sous-populations d'élèves sans le nombre élevé d'erreurs..	137
	<b>CHAPITRE VI .....</b>	<b>140</b>
	<b>DISCUSSION.....</b>	<b>140</b>
6.1	L'EXAMEN DU RÉSULTAT DE L'ANALYSE DESCRIPTIVE DES VARIABLES CARACTÉRISTIQUES DES ÉLÈVES.....	141
6.1.1	La performance des élèves selon l'âge.....	141
6.1.2	La performance des élèves selon le genre .....	143
6.1.3	La performance des élèves selon langue maternelle.....	144
6.2	L'EXAMEN DU RÉSULTAT DE L'ANALYSE DESCRIPTIVE DES VARIABLES CARACTÉRISTIQUES DES ÉCOLES .....	146
6.2.1	La performance des élèves selon Le secteur .....	146
6.2.2	La performance des élèves selon la mixité .....	148
6.2.3	La performance des élèves selon la taille .....	148
6.2.4	La performance des élèves selon la proportion des élèves de langue maternelle française.....	149
6.3	L'EXAMEN DU RÉSULTAT DE L'ANALYSE MULTINIVEAUX.....	149
6.3.1	La modélisation multiniveaux multivariée à vide .....	150
6.3.2	La modélisation des variables caractéristiques des élèves.....	151
6.3.3	La modélisation des variables caractéristiques des écoles.....	152

<b>CHAPITRE VII .....</b>	<b>154</b>
<b>CONCLUSION.....</b>	<b>154</b>
7.1 UN BREF RAPPEL DES OBJECTIFS ET DE LA MÉTHODE DE LA RECHERCHE .....	154
7.2 LES LIMITES DE LA RECHERCHE.....	156
7.2.1 Le choix des variables traitées et le nombre de niveaux appliqué .....	156
7.2.2 Des relations dont l'interprétation reste à préciser .....	157
7.2.3 Le concept du déterminisme scolaire .....	158
7.2.4 Les résultats des élèves et les facteurs liés à la stabilité scolaire.....	159
7.2.5 L'importance des notes initiales de l'élève .....	160
7.3 UN SOMMAIRE DE L'ÉTUDE .....	161
<b>RÉFÉRENCES .....</b>	<b>165</b>

## LISTE DES TABLEAUX

Tableau 1 - Performance des élèves des trois populations .....	87
Tableau 2 - Performance des élèves de l'ensemble de la population selon l'âge .....	88
Tableau 3 - Performance des élèves selon le genre .....	91
Tableau 4 - Performance des élèves selon l'âge et le genre .....	92
Tableau 5 - Performance des élèves selon la langue maternelle .....	93
Tableau 6 - Performance des élèves selon l'âge et la langue maternelle .....	94
Tableau 7 - Performance des élèves selon le réseau .....	95
Tableau 8 - Performance des élèves selon l'âge et le réseau .....	95
Tableau 9 - Performance des élèves selon la mixité de l'école .....	96
Tableau 10 - Performance des élèves selon la taille de l'école .....	97
Tableau 11 - Performance des élèves selon la proportion des élèves de l'école de langue maternelle française .....	98
Tableau 12 - Performance des élèves sans le nombre élevé d'erreurs .....	102
Tableau 13 - Performance des élèves sans le nombre élevé d'erreurs selon l'âge .....	103
Tableau 14 - Performance des élèves sans le nombre élevé d'erreurs selon le genre .....	103
Tableau 15 - Performance des élèves sans le nombre élevé d'erreurs selon la langue maternelle .....	104
Tableau 16 - Performance des élèves sans le nombre élevé d'erreurs selon le réseau .....	105
Tableau 17 - Performance des élèves sans le nombre élevé d'erreurs selon la mixité .....	106
Tableau 18 - Performance des élèves sans le nombre élevé d'erreurs selon la taille .....	106

Tableau 19 - Performance des élèves sans le nombre élevé d'erreurs selon le pourcentage des élèves de l'école de langue maternelle française .....	107
Tableau 20 – Effectifs des élèves en difficulté selon l'âge .....	108
Tableau 21 - Performance des élèves en difficulté selon le genre.....	109
Tableau 22 - Performance des élèves en difficulté selon la langue maternelle .....	110
Tableau 23 - Performance des élèves en difficulté selon le réseau .....	110
Tableau 24 - Performance des élèves en difficulté selon la mixité de l'école .....	111
Tableau 25 - Estimation des paramètres et de leurs erreurs types pour l'ensemble des élèves des trois populations.....	126
Tableau 26 - Estimations des paramètres et des erreurs types des élèves sans le nombre élevé d'erreurs .....	129
Tableau 27 - Estimations des coefficients et de leurs erreurs types des variables caractéristiques de l'élève et de l'école des trois cohortes .....	136
Tableau 28 - Estimation des coefficients et de leurs erreurs types des variables caractéristiques des élèves et des écoles des sous-populations d'élèves sans le nombre élevé d'erreurs.....	139

## LISTE DES FIGURES

Figure 1 - Un modèle de base de l'environnement de l'école .....	7
Figure 2 - Distribution des résultats des élèves en syntaxe et ponctuation en 2006.....	83
Figure 3 - Distribution des résultats des élèves en orthographe en 2006 .....	84
Figure 4 - Distribution des résultats des élèves en syntaxe et ponctuation en 2007 .....	85
Figure 5 - Distribution des résultats des élèves en orthographe en 2007 .....	85
Figure 6 - Distribution des résultats des élèves en syntaxe et ponctuation en 2008.....	86
Figure 7 - Distribution des résultats des élèves en orthographe en 2008 .....	86
Figure 8 - Performance des élèves en syntaxe et ponctuation selon l'âge .....	89
Figure 9 - Performance des élèves en orthographe selon l'âge.....	90
Figure 10 - Résultat des élèves sans le nombre élevé d'erreurs en syntaxe et ponctuation en 2006 .....	99
Figure 11 - Résultat des élèves sans le nombre élevé d'erreurs en orthographe en 2006.....	100
Figure 12 - Résultat des élèves sans le nombre élevé d'erreurs en syntaxe et ponctuation en 2007 .....	101
Figure 13 - Résultat des élèves sans le nombre élevé d'erreurs en orthographe en 2007.....	101
Figure 14 - Modélisation à un niveau : les élèves ( $E_i$ ) et deux variables dépendantes, syntaxe et ponctuation ( $Y_1$ ), orthographe ( $Y_2$ ) .....	114
Figure 15 - Modélisation multivariée à deux niveaux (élève, école) et deux variables dépendantes : syntaxe et ponctuation ( $Y_1$ ), orthographe ( $Y_2$ ) .....	118
Figure 16 - Modélisation multivariée à trois niveaux (élève, école, année) et deux variables dépendantes : syntaxe et ponctuation ( $Y_1$ ), orthographe ( $Y_2$ ) .....	123

## LISTES DES ABRÉVIATIONS, SIGLES ET ACRONYMNES

AANB	Acte de l'Amérique du Nord britannique
ACSQ	Association des cadres scolaires du Québec
CIEP	Contexte, l'intrant, l'extrant et le processus
CRIFPE	Centre de recherche interuniversitaire sur la formation et la profession enseignante
CAI	Commission d'accès à l'information
CS	Commission scolaire
CSE	Conseil supérieur de l'éducation
Éi	Écart interquartile
Et	Écart-type
ET	Erreur type
IMSE	Indice du milieu socioéconomique
LABRIPROF	Laboratoire de recherche et d'intervention portant sur les politiques et les professions en éducation
LIP	Loi sur l'instruction publique
M	Moyenne
Md	Médiane
MELS	Ministère de l'Éducation, du Loisir et du Sport
MEQ	Ministère de l'Éducation du Québec
NELS	<i>National Education Longitudinal Study</i>
OCDE	Organisation de coopération et de développement économiques
PFÉQ	Programme de formation de l'école québécoise
PISA	<i>Programme for International Student Assessment</i>

## **REMERCIEMENTS**

Je veux exprimer d'abord ma profonde gratitude à monsieur Jean-Guy Blais, mon directeur de thèse, professeur titulaire au Département d'administration et fondements de l'éducation de l'Université de Montréal. Il m'a guidé tout au long de cette recherche. J'ai beaucoup apprécié sa compétence.

Mes remerciements s'adressent également au Département de la recherche et de la statistique du ministère de l'Éducation, du Loisir et du Sport (MELS) du Québec qui a mis à ma disposition les données qui ont fait l'objet d'analyse dans le cadre de la présente étude.

Je désire remercier particulièrement les membres de ma famille, mes enfants Larra Charles, Emmanuelle Charles, Valérie Charles, Didier Charles, Sébastien Charles et ma femme Florence Jules-François. Ils ont tous témoigné envers moi une grande compréhension dans la poursuite de ce projet.

Ma gratitude s'adresse également à tous ceux et à toutes celles qui de diverses façons m'ont encouragé à persévérer dans l'accomplissement de ce travail, à madame Huguette Daneau pour la mise en page du document et à mon ami Lionel Jean qui m'a offert sa contribution pour la relecture du texte.

## INTRODUCTION

La maîtrise de la langue d'enseignement est généralement considérée comme un facteur important de la réussite scolaire et de l'intégration sociale (MEQ, 1992). Il devient alors important pour les acteurs d'un système éducatif donné de mettre à la disposition des élèves tous les moyens qui favoriseront cette maîtrise. Dans cette perspective, l'école peut être considérée comme un lieu qui promeut l'égalisation des chances en imposant les mêmes règles à tous les élèves dans un environnement contrôlé où chacun a la possibilité de réussir, indépendamment de son origine sociale. L'école peut aussi être définie à partir de considérations sociopolitiques : elle devient un lieu de rencontre, une organisation d'un «marché» social où se rejoignent fournisseurs de services et demandeurs d'instruction (Rouet, 1993). De ce point de vue, l'école deviendra un lieu d'interactions entre de nombreux acteurs, aux prérogatives, aux pouvoirs et aux devoirs différents. Les auteurs Tyack et Cuban (2001, p. 86) décrivent le fonctionnement de l'école au quotidien comme suit : «Dans une école secondaire, les élèves sont regroupés à l'intérieur d'une classe souvent en fonction de leur âge. À chaque heure, ils changent d'une discipline à l'autre et d'un enseignant à l'autre.» En réfléchissant sur la portée de cette définition, on remarque bien qu'au concept d'école en tant qu'entité institutionnelle, se greffe celui de la salle de classe dans laquelle se retrouvent les enseignants et les élèves. Et pour ce qui est de l'apprentissage que ces enseignants font effectuer aux élèves, on s'est souvent posé la question suivante : « L'école remplit-elle adéquatement la mission qui lui est assignée? »

C'est ainsi qu'au début des années 1980, au nom de la reddition des comptes, on a vu apparaître une demande de production d'indicateurs pour

apprécier la performance des écoles. La définition de l'école en tant qu'institution devient ainsi inséparable du paradigme de l'école « efficace ». Ce courant de pensée cherche à déterminer et à isoler les facteurs permettant d'augmenter l'efficacité des écoles et d'améliorer les résultats scolaires des élèves. Cependant, bien qu'il s'appuie sur la conception de l'évaluation des apprentissages par les tests, le courant de recherche sur l'école efficace se situe dans la lignée des travaux du rapport Coleman (Coleman, Campbell, Hobson, McPartland, Mood et York, 1966) dont les analyses tendaient à montrer que l'école jouait un rôle négligeable dans l'amélioration des résultats scolaires et la réduction de l'inégalité des chances des élèves. Depuis la publication de ce rapport, de nombreuses études issues du courant de recherche sur l'école « efficace » ont tenté d'en vérifier les conclusions et de réexaminer la problématique de l'efficacité des écoles.

Mais les chercheurs n'ont pas toujours adopté les mêmes approches méthodologiques ni obtenu des résultats similaires pour des problématiques semblables. Par exemple, Pelgrum, Voogt et Plomp (1995) ont recensé plus de 170 indicateurs sous différentes appellations pour le seul aspect du curriculum. La question du choix des indicateurs pour représenter la relation entre les caractéristiques individuelles des élèves et des facteurs d'école et le rendement des élèves dans les matières scolaires, comme le français et les mathématiques, est l'une des préoccupations du mouvement de recherche sur les écoles efficaces.

Au Québec, les préoccupations relatives à la réussite scolaire apparaissent depuis plusieurs années dans les travaux de certains chercheurs (MEQ, 1992; Blais, 2003; McAndrew et collab., 2005; Ouedraogo, 2006; MELS, 2008). En utilisant des données du Ministère sur une période de trois ans, pour les cohortes d'élèves des années 2006, 2007 et 2008, relatives à l'épreuve uniforme de français écrit de la 5<sup>e</sup> année du secondaire, la présente

recherche tente de déterminer les variables caractéristiques de l'élève et de l'école qui seraient davantage liées aux résultats des élèves finissants du secondaires de la région de Montréal. Cette région urbaine a été retenue parce qu'on relève la présence de plusieurs groupes ethniques dans ses écoles. Une démarche de compréhension de l'apprentissage du français par les finissants du secondaire dans cette région présente alors des défis supplémentaires. Deux parmi les six critères d'évaluation de l'épreuve écrite, à la satisfaction desquels certains élèves éprouvent de la difficulté, ont été retenus pour la recherche : la syntaxe et la ponctuation, ainsi que l'orthographe.

Nous effectuons d'abord trois séries d'analyses pour examiner les relations entre les variables dépendantes, la syntaxe et la ponctuation, l'orthographe, et les variables indépendantes caractéristiques des élèves (âge, genre, langue maternelle) et des écoles (réseau, mixité, taille, pourcentage d'élèves de l'école de langue maternelle française) pour chacune des années : la population complète, une première sous-population d'élèves sans le nombre élevé d'erreurs et une deuxième sous-population d'élèves qui font 24 erreurs de syntaxe et de ponctuation et 35 erreurs d'orthographe. Nous poursuivons l'analyse avec plusieurs modélisations multiniveaux multivariées à deux niveaux (élève, école) et à trois niveaux (élève, école, année) pour dégager la variabilité des résultats obtenus par les élèves relativement aux deux critères de correction à l'épreuve unique de français écrit du MELS, selon les variables caractéristiques de l'élève et de l'école.

Le chapitre I traite de la problématique de recherche sur l'« efficacité » des écoles. Il décrit le développement de la recherche sur l'« efficacité » des écoles, le contexte, la pertinence, le problème général, l'objectif et la question spécifiques de recherche. Ce chapitre précise aussi les fondements de notre étude sur la modélisation des liens entre les variables caractéristiques de

l'élève et de l'école et les résultats des élèves en français écrit. Le chapitre 2 traite du cadre théorique de notre recherche. Il fait la recension des écrits. Nous examinons les éléments théoriques de la recherche sur la modélisation des liens entre des variables caractéristiques de l'environnement scolaire et la performance des élèves : la question des « indicateurs » éducationnels est traitée; sont aussi présentés les faits saillants de la recherche sur les « écoles efficaces ». Le chapitre 3 qui porte sur la méthodologie présente les sources des données, nos choix méthodologiques, les méthodes d'analyse des données, les modèles et les modélisations multiniveaux multivariées. Il désigne des variables individuelles caractéristiques de l'élève et des variables caractéristiques de l'école qui seront modélisées dans la recherche. Le chapitre 4 fournit la description des données, présente les analyses descriptives des variables individuelles caractéristiques des élèves et des écoles des trois populations pour les années 2006, 2007 et 2008. Le chapitre 5 expose les résultats de la modélisation multiniveaux multivariée. Le chapitre 6 discute des résultats de la recherche. Le chapitre 7 résume brièvement la recherche, il en souligne les limites et en dégage les conclusions; il propose enfin des perspectives pour de futures recherches.

## **CHAPITRE I**

### **LA PROBLÉMATIQUE**

Ce chapitre présente les éléments de la problématique de notre recherche sur l'« efficacité » des écoles : le développement de la recherche sur l'« efficacité » des écoles, certains des éléments contextuels du système éducatif québécois, le concept de l'« efficacité » de l'école, la pertinence de la recherche, le problème général de recherche, l'objectif et la question spécifiques de recherche.

#### **1.1 LE DÉVELOPPEMENT DE LA RECHERCHE SUR L'« EFFICACITÉ » DES ÉCOLES**

La mise en place des écoles primaires et secondaires qui assurent aujourd'hui à des jeunes l'éducation obligatoire a duré presque deux siècles. Avant les années 1800, l'instruction élémentaire des enfants était laissée à l'initiative des religieux (Rouet, 1993; Spring, 1996). C'est alors que l'État, en France et aux États-Unis, par exemple, a pris graduellement en charge la scolarisation des enfants. La diversité des points de vue des responsables politiques particulièrement aux États-Unis a cependant amené les établissements scolaires à se doter d'objectifs différents et à l'occasion contradictoires (Tyack et Cuban, 2001, p. 43). On demandait aux écoles de :

- socialiser les élèves à l'obéissance et, en même temps, de les initier à l'esprit critique;
- faire acquérir aux élèves les meilleures connaissances scolaires ainsi que les habiletés pratiques pour le marché du travail;
- favoriser entre les élèves à la fois l'esprit de coopération et de compétition;
- mettre en valeur les connaissances de base tout en encourageant la créativité et les habiletés complexes.

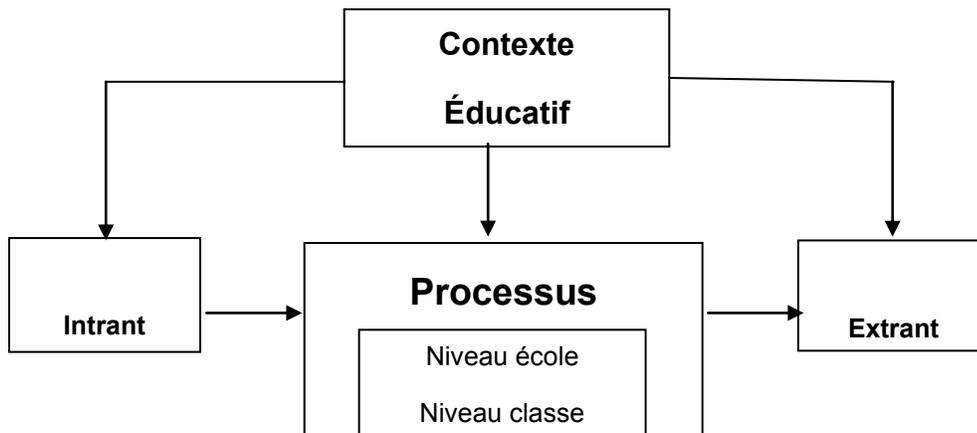
C'est dans ce contexte que se sont développées, depuis les années 1840, les écoles secondaires en Amérique du Nord. Compte tenu du passé esclavagiste des États-Unis, la scolarisation des élèves des différents groupes ethniques se faisait dans les écoles primaires et secondaires en fonction de leur appartenance ethnique jusqu'au milieu des années 1960. Constatant une différence dans la réussite scolaire des jeunes à la fin de leurs études secondaires selon leur appartenance ethnique, les autorités politiques ont commandé le rapport Coleman sur l'égalité des chances scolaires des minorités. Bien que cette étude ait eu pour but, entre autres, de vérifier dans quelle mesure le rendement scolaire des élèves était relié à leurs origines sociales et ethniques, l'influence des facteurs<sup>1</sup> d'école sur les résultats scolaires des élèves a aussi été examinée (Coleman, Campbell, Hobson, McPartland, Mood et York, 1966). D'après Reynolds, Teddlie, Creemers, Scheerens et Townsend (2000), c'est le rapport Coleman qui a amorcé le mouvement de recherche sur l'« efficacité » des écoles. Ce paradigme de recherche occupe aujourd'hui une place importante dans le

---

<sup>1</sup>Les facteurs d'école sont des variables caractéristiques de l'école, comme la taille, le réseau d'appartenance (public, privé), la mixité, l'indice du milieu socio-économique, etc. Ces facteurs seront expliqués plus en détail au chapitre 3 traitant du cadre théorique.

discours sur l'école et sur son rôle. Plusieurs études sur l'« efficacité » des écoles à la suite du rapport Coleman, dont celle de Peaker (1967) ainsi que celle de Ainsworth et Batten (1974) ont généralement été perçues comme une réponse à des conclusions de ce rapport, selon lesquelles l'école jouait un rôle négligeable dans l'amélioration des résultats scolaires des élèves et dans la réduction de l'inégalité des chances.

Le modèle de base du mouvement de recherche sur l'« efficacité » des écoles a été proposé par Stufflebeam (1971). Les auteurs Scheerens, Glass et Thomas (2007, p.18) reprenaient ce modèle dont ses principales composantes sont : le contexte, l'intrant, l'extrant et le processus (CIEP). La figure 1 expose l'objectif principal de ce courant de recherche qui vise à déterminer dans quelle mesure les variables caractéristiques du contexte éducatif, de l'intrant et du processus influencent les extrants, c'est-à-dire les résultats scolaires des élèves. Ce modèle met en évidence les relations entre une variation hypothétique, le plus souvent une amélioration des conditions de l'environnement, et les résultats scolaires des élèves.



**Figure 1 - Un modèle de base de l'environnement de l'école**

Il est utile de noter le caractère dynamique de ce modèle, en ce sens que le système éducatif est présenté comme une fonction de production qui

transforme les intrants en extrants éducationnels. Les intrants représentent les ressources matérielles et financières. Les processus représentent les structures organisationnelles et pédagogiques de l'institution scolaire. Les extrants peuvent être considérés, par exemple, comme les résultats scolaires des élèves à des examens officiels. Chacun de ces aspects du modèle peut faire l'objet d'une évaluation spécifique.

La plupart des chercheurs, dont Scheerens, Glass et Thomas (2007), reconnaissent toutefois l'existence de plusieurs courants à l'intérieur de ce mouvement de recherche. Chacun d'eux varie selon les intrants sur lesquels on met l'accent pour produire les extrants, tout en considérant le fait que les éléments du processus et du contexte restent dynamiques. Ces courants se rattachent chacun à un champ d'études particulier : l'économie, la sociologie, la pédagogie, etc. Ce sont :

- a) les études sur l'égalité des chances en éducation et le sens de l'école;
- b) les études économiques sur les fonctions de production éducatives;
- c) les études sur les programmes compensatoires en éducation;
- d) les études sur l'efficacité des écoles;
- e) les études sur l'efficacité de l'enseignement (salle de classe) et des approches pédagogiques.

Plusieurs études récentes sur l'« efficacité » des écoles intègrent ces différentes approches dans la manière de modéliser les données et dans le

choix des variables explicatives. À ce sujet, la modélisation multiniveaux a apporté une contribution importante. Cette modélisation permet un calcul adéquat des variances et des erreurs types à chacun des niveaux de l'analyse. On obtient généralement les mêmes estimés que dans la régression linéaire, mais les erreurs types sont corrigées pour les corrélations et l'hétérogénéité de la variance impliquées qui affectent les tests de signification statistiques.

Le système scolaire est alors considéré comme une organisation à plusieurs niveaux hiérarchiques, dont la prémisse principale veut que les responsables des niveaux supérieurs – dans le cas du Québec, le ministère de l'Éducation, du Loisir et du Sport (MELS) et les commissions scolaires (CS) – facilitent la tâche des acteurs des niveaux inférieurs, c'est-à-dire les écoles, les enseignants et les élèves. Ainsi, une intégration des éléments des différents courants de recherche sur l'« efficacité » des écoles devient possible par l'introduction dans l'analyse des variables et des concepts méthodologiques de chacun de ces courants de recherche.

Quoique cette intégration permette l'avancée des connaissances, dans la mesure où elle inclut dans l'analyse les variables caractéristiques de l'intrant, du processus, de l'extrant, des conditions du contexte ainsi que la reconnaissance de la structure hiérarchique du système éducatif, l'approche de la modélisation quantitative multiniveaux comporte un certain nombre de difficultés dont il faut tenir compte. Ainsi, dans son avis au ministère de l'Éducation du Québec (MEQ), l'Association des cadres scolaires du Québec (ACSQ, 2002, p. 9) formule cette observation : « Loin d'être une entreprise de production, l'école en est une d'éducation, et son efficacité ne se mesure pas uniquement en chiffres. » En outre, lorsque l'analyse se concentre uniquement sur l'école, elle ne tient évidemment pas compte de tous les éléments qui pourraient influencer le fonctionnement du système éducatif et les résultats scolaires des élèves.

## **1.2 DES ÉLÉMENTS CONTEXTUELS DU SYSTÈME ÉDUCATIF QUÉBÉCOIS**

Le développement du système éducatif, dans chaque pays, est influencé par son histoire et un contexte particulier (Dumay et Dupriez, 2009). Cette section explique sommairement le développement historique, démographique, financier et législatif du système éducatif québécois.

### **1.2.1 Le contexte historique**

Outre la marque de l'héritage de la France, le système scolaire québécois a été largement influencé par le contexte nord-américain. L'industrialisation des sociétés occidentales au début des années 1800 a été à la base du développement de l'école secondaire. Elle a nécessité la mise en place d'un cadre de formation à grande échelle pour les besoins de son fonctionnement. Au Canada, l'Acte de l'Amérique du Nord britannique (AANB) a donné des garanties constitutionnelles en matière d'éducation aux deux grandes confessions religieuses qui s'identifiaient aux deux communautés linguistiques présentes sur le territoire, les communautés française et anglaise. C'est ainsi qu'au Québec, pendant plus d'un siècle, il y a eu des commissions scolaires confessionnelles catholiques, fréquentées par la majorité francophone, et protestantes, fréquentées par les anglophones et la plupart des allophones.

Le développement des systèmes scolaires en Amérique du Nord s'est cependant réalisé selon les contextes bien particuliers de chaque pays. Contrairement à d'autres nations industrialisées, les États-Unis ont connu une augmentation rapide de la scolarisation des jeunes : en 1900, 10 % des enfants âgés de quatorze à dix-sept ans fréquentaient les écoles secondaires ; ils étaient 70 % en 1940 et 90 % en 1980 (Tyack et Cuban, 2001, p. 47). Au Québec, s'inspirant de la tradition et de l'héritage français (Rapport Parent, 1963), l'éducation primaire et secondaire a été assurée par

les congrégations religieuses jusqu'à la création du ministère de l'Éducation du Québec (MEQ) en 1964. C'est la mise en place du MEQ qui va accélérer le développement et la démocratisation du système éducatif québécois. Brassard (2005, p.1) estime que cette initiative témoigne de la volonté de l'État québécois d'assurer son autorité en matière d'éducation primaire et secondaire. Ce sont les législations successives du gouvernement qui ont conduit le système scolaires québécois vers sa forme actuelle.

### **1.2.2 Le contexte législatif**

Les difficultés économiques des années 1990 ont créé des pressions sur la demande de production d'indicateurs de plusieurs systèmes sociaux, notamment le système éducatif. Selon Blais, Laurier et Pelletier (2001, p. 132), les gouvernements ont pris durant ces années le virage de l'imputabilité, notamment sous la forme de reddition de comptes, puis celle de l'obligation de résultats. Dans ce même esprit, Brassard (2005, p. 1) souligne que les différentes réformes, notamment la loi 124 du ministère de l'Éducation du Québec (MEQ), constituent la dernière étape d'un long processus de recherche d'un modèle de répartition du pouvoir au sein du système scolaire québécois qui contribuerait à la fois à l'efficacité de l'établissement scolaire et à la satisfaction des différents acteurs associés de près ou de loin à l'activité éducative.

La loi 124 du MEQ, promulguée en décembre 2002, impose l'obligation pour l'ensemble des établissements scolaires publics de produire un plan de réussite. On peut lire dans le rapport annuel de gestion du MEQ (MEQ, 2002, p. 79 à 82) que l'école doit « accroître l'efficacité du système d'éducation en mettant l'accent sur *les résultats*, sur l'imputabilité et sur la transparence de la gestion publique ».

### **1.2.3 Le contexte démographique**

Le ministère de l'Éducation, du Loisir et du Sport (MELS) détient la responsabilité d'adapter l'offre de services à la population en âge de fréquentation scolaire obligatoire, à savoir les jeunes âgés de 6 à 16 ans. Depuis les années 1960, l'effet conjugué du vieillissement de la population et de la diminution du taux de natalité au Québec a provoqué une baisse du nombre de jeunes âgés de 6 à 24 ans, malgré une légère augmentation des ressortissants de l'immigration qui se retrouvent surtout dans la région métropolitaine de Montréal. Le nombre de jeunes de cette tranche d'âge est passé de 2 030 000 pour une population totale de 5 259 000 en 1961, soit 38,0 %, à 1 900 000 pour une population de 7 600 000, en 2001, soit 25,0 %. Paradoxalement, malgré la diminution de l'effectif scolaire, les édifices sont devenus plus grands, l'organisation s'est complexifiée et l'offre de formation plus diversifiée (CSE, 2007). Cette nouvelle situation a eu un impact important sur la gestion et le financement du système éducatif québécois.

### **1.2.4 Le contexte financier**

De la création du ministère de l'Éducation du Québec (MEQ) en 1964 à nos jours, le financement de l'éducation par l'État s'est accru de manière considérable. Le ministère de l'Éducation, du Loisir et du Sport (MELS, 2007, p. 18) estime que les dépenses du gouvernement pour l'éducation en 2006-2007 étaient de 12,8 milliards de dollars, soit 1,6 milliard de dollars de plus qu'en 2002-2003. Ces dépenses constituent le deuxième poste budgétaire du gouvernement du Québec.

C'est dans ce contexte historique, démographique, législatif et financier que se situe la présente étude. Du point de vue historique, le milieu scolaire québécois est influencé par les idéologies éducatives dominantes en France et aux États-Unis. Durant ces dernières années, fortement régulé par les législations gouvernementales, le système scolaire québécois a été marqué

par une baisse importante du nombre de jeunes en âge de fréquentation scolaire obligatoire et une augmentation de la diversité culturelle et ethnique dans les grandes villes, particulièrement Montréal.

### **1.3 LE CONCEPT DE L'« EFFICACITÉ » DE L'ÉCOLE**

La définition du concept d'« efficacité » d'une école ne fait pas consensus au sein de la communauté éducative. Cette situation est due en partie au fait que les chercheurs ne s'entendent pas sur les déterminants qui pourraient expliquer l'« efficacité » d'un établissement d'enseignement. Selon Cheong Cheng (1996), l'« efficacité » d'une école peut être appréciée, d'une part, selon le degré de satisfaction des personnes qui y sont associées, à savoir les parents, les élèves, le personnel de direction, les enseignants et, d'autre part, par la mesure des opérations internes à l'école, c'est-à-dire la bonne gestion, la capacité d'une organisation à utiliser de façon efficace ses ressources, etc.

Au contraire de Coleman et ses collaborateurs (1966) qui concluaient que les facteurs d'école entraînaient peu d'influence sur les résultats scolaires des élèves, plusieurs auteurs (Hedges, Greenwald et Laine, 1994; Rumberger et Palardy, 2005; Konstantopoulos, 2006; Bressoux, 2007; Dumay et Dupriez, 2009; Petrucci, 2010) ont montré que certaines variables associées à l'établissement scolaire peuvent présenter un lien significatif avec le rendement des élèves, rendement exprimé le plus souvent par leurs résultats scolaires à des tests standardisés dans les matières de base, comme la langue d'enseignement, les sciences et les mathématiques.

D'autre part, certains auteurs – dont Nutall (1991), Cheong Cheng (1996) ainsi que l'Association des cadres scolaire du Québec (ACSQ, 2002, p. 1) –, en quête d'une meilleure compréhension de la qualité d'une institution

scolaire, suggèrent de ne pas dissocier le concept d'« efficacité » d'une école de sa mission, plus générale, ni de son projet éducatif, plus spécifique. Au Québec, le ministère de l'Éducation, du Loisir et du Sport (MELS, 2003) dans son *Programme de formation de l'école québécoise* (PFÉQ, 2003, p. 5) résume ainsi la mission de l'école québécoise : « Instruire, socialiser et qualifier ». Les auteurs Carufel et Moreau (2002, p. 5) proposent la définition suivante du projet éducatif :

« C'est une démarche dynamique par laquelle une école, compte tenu des besoins de ses élèves et en concertation avec ses divers agents, s'applique à assurer un environnement éducatif favorable à l'élève, détermine ses orientations propres et se donne un plan d'action en lien avec sa mission. Le projet éducatif s'appuie sur la mission de l'école qui est « d'instruire, de socialiser et de qualifier » les élèves tout en les rendant aptes à entreprendre et à réussir un parcours scolaire. Il émerge d'un long processus cohérent de réflexion et de discussion dans lequel les acteurs d'un établissement scolaire expriment clairement les idées directrices qui les animent et se bâtissent progressivement une identité collective qui les distingue des autres écoles.»

Il faut ajouter que dans la mesure où les projets éducatifs peuvent être différents d'une école à l'autre, l'« efficacité » d'une école devient relative. C'est d'ailleurs dans ce sens que Nutall (1991) parle de l'équité ou de l'efficacité différentielle dans l'évaluation des établissements scolaires. Par exemple, plusieurs écoles publiques et privées de la région métropolitaine de Montréal recrutent leurs élèves en fonction de leur projet éducatif. Conséquemment, ces écoles se différencient tant par rapport aux critères de sélection que selon leurs objectifs d'apprentissage : des programmes d'enrichissement dans les matières de base, des écoles sports-études, des

écoles d'art (musique, danse), des écoles confessionnelles privées, des écoles selon le genre de l'élève (mixte, filles, garçons), des écoles d'éducation internationale, des écoles d'immersion qui favorisent l'apprentissage d'une langue seconde, etc.

La problématique de la présente recherche s'articule sur la thématique de l'« efficacité » des écoles. Elle porte sur la modélisation des liens entre des variables associées à l'environnement scolaire, c'est-à-dire au contexte éducatif, à l'intrant, au processus et au rendement des élèves à des épreuves officielles dont les résultats interpellent les acteurs du système éducatif québécois. La modélisation consistera, dans ce contexte, d'une part, à déterminer des variables susceptibles d'affecter les résultats scolaires des élèves et, d'autre part, à apprécier leur degré d'influence les unes par rapport aux autres, tout en mettant en relief, lorsqu'il s'avérera pertinent de le faire, leurs interrelations.

#### **1.4 LA PERTINENCE DE LA RECHERCHE**

Depuis 1986, le ministère de l'Éducation du Québec (MEQ) publie tous les ans les *Indicateurs de l'éducation* qui traitent de l'ensemble du système scolaire québécois. Cette publication, outre l'information qu'elle met à la disposition des usagers du système éducatif (les parents, les administrateurs d'école, certaines entreprises, etc.), poursuit aussi un objectif de reddition de comptes (MELS, 2007, p. 9). Il s'agit alors pour le ministère de l'Éducation, du Loisir et du Sport (MELS), au nom de la transparence, d'informer la population de l'état du système éducatif et de son évolution.

Les publications du MELS (MEQ, 1992; MELS, 2008) ont relevé, de l'école primaire à l'université, l'existence d'un problème sur le plan de la qualité dans l'usage de la langue : nombreuses fautes d'orthographe, connaissance

insuffisante de la grammaire, pauvreté du vocabulaire, etc. Déjà en 1992, le Ministère (MEQ, 1992, p. 3) reconnaissait la présence de difficultés quant à l'apprentissage du français dans les écoles primaires et secondaires. À cette époque, il formulait l'hypothèse que ces difficultés découlaient en partie de l'omniprésence de l'image à la télévision et dans les jeux vidéo ainsi que de la diversité ethnique croissante dans les classes dans la région métropolitaine de Montréal. En même temps, il admettait que l'école ne saurait prendre prétexte de cette situation sociale et culturelle pour éviter de se poser la question de l'efficacité de ses propres pratiques en matière d'enseignement de la langue, particulièrement du français écrit.

Depuis l'entrée en vigueur de la Réforme de l'éducation en l'an 2000, le ministère de l'Éducation, du Loisir et du Sport (MELS, 2008) se préoccupe des retombées de celle-ci sur l'apprentissage de la langue d'enseignement. À cet égard, il a constitué un comité d'experts indépendants sur l'apprentissage du français dans le but d'analyser le rendement scolaire des élèves aux ordres d'enseignement primaire et secondaire, en vue de mieux soutenir le développement de la compétence à écrire, particulièrement en ce qui a trait à l'orthographe, à la syntaxe et à la ponctuation.

Ce comité reconnaît déjà que, pour être compétent en écriture, l'élève doit avoir acquis des connaissances en ce qui a trait à la grammaire et au lexique et qu'il faut donc relier les termes «connaissances» et «compétences»; il souligne en même temps que le développement de la compétence à écrire est impossible sans la maîtrise du code linguistique. Ces observations ont été confirmées dans une étude longitudinale récente (MELS, 2008) portant sur les résultats obtenus par les élèves de 2002 à 2006 à l'épreuve unique d'écriture de cinquième secondaire du MELS. Cette même étude rappelle que certains élèves éprouvent de la difficulté dans la maîtrise de deux des six éléments liés aux critères de la grille de correction, la syntaxe et la ponctuation ainsi que l'orthographe.

Relativement à cette constante préoccupation de vouloir améliorer la qualité de l'apprentissage du français, le *Programme de formation de l'école québécoise, Domaine des langues* (MELS, 2003, p. 1) signale que la langue et la culture sont deux dimensions essentielles des cours de français. La maîtrise de la langue française par tous les élèves est un facteur primordial de la réussite scolaire et représente un défi majeur pour l'école québécoise.

Pour cette raison, le MELS a mis en œuvre deux plans d'action au cours des dernières années. D'une part, des mesures pour l'amélioration du français écrit et, d'autre part, des mesures sur la réussite et la persévérance scolaire. Il recommandait que la maîtrise du code linguistique et l'accès aux œuvres littéraires soient davantage pris en compte dans l'enseignement de la langue écrite.

De plus, la place déterminante de la langue d'enseignement dans le curriculum incitera les responsables scolaires à choisir la réussite aux épreuves ministérielles relatives à la langue d'enseignement au secondaire, c'est-à-dire le français, comme cible parmi les principaux indicateurs nationaux des plans de réussite (ACSQ, 2002, p. 2). L'importance de cette épreuve s'explique également par le fait que la réussite du cours de français de 5<sup>e</sup> secondaire est nécessaire à l'obtention du diplôme de fin d'études secondaires.

La pertinence de la présente recherche sur la modélisation des liens entre un certain nombre de variables caractéristiques de l'élève et de l'école et les résultats des élèves finissants du secondaire à l'épreuve unique de français écrit du MELS sur une période de trois ans (2006, 2007 et 2008) se justifie donc dans le contexte actuel. Notre étude propose, comme on pourra le constater, d'ajouter quelques éléments d'explication au débat actuel sur l'apprentissage de la langue d'enseignement.

## 1.5 LE PROBLÈME DE RECHERCHE

Dans les recherches sur l'« efficacité » des écoles, la question de l'importance des variables caractéristiques de l'élève, de la salle de classe et de l'école en rapport avec le rendement scolaire des élèves est essentielle. Il n'a pas toujours été facile cependant de déterminer non seulement les variables les plus pertinentes associées à l'« efficacité » d'une école, mais aussi de départager leurs effets en fonction des différents niveaux d'appartenance (élève, classe, école), selon le rendement scolaire des élèves.

Dans une analyse conceptuelle, Felouzis et Perroton (2007) ont souligné que la notion même d'« efficacité » d'une école recouvre plusieurs réalités pour lesquelles il reste à construire l'articulation au sein d'une véritable théorie. Goldstein et Woodhouse (2000), en Grande-Bretagne, ont quant à eux reproché aux chercheurs du mouvement de recherche sur l'« efficacité » des écoles de considérer les établissements scolaires comme des entités indépendantes et de négliger ainsi leur rapport d'interdépendance. En d'autres termes, ils sont d'avis que dans leur évaluation certains chercheurs ne tiennent pas compte de la spécialisation des écoles en fonction de leur clientèle spécifique ni du projet éducatif de l'établissement scolaire.

Outre ce problème du rapport d'interdépendance des écoles, les auteurs Blais, Laurier et Pelletier (2001) dans une réflexion sur l'utilisation des indicateurs comme outils de régulation en éducation ont souligné le fait suivant :

« L'attention à l'égard de la qualité générale de l'établissement scolaire s'est aujourd'hui largement transposée sur les indicateurs d'extrant. Ainsi, aux États-Unis, on favorise maintenant une appréciation de la qualité du système scolaire à l'aide des résultats à des tests cognitifs standardisés; on

apprécie plus la qualité en termes de résultats qu'en termes d'intrant, de processus ou de contexte. En fait, nous sommes passés de l'obligation de moyens à l'obligation de résultats.»

Les premières recherches sur l'« efficacité » des écoles (Coleman et collab., 1966; Peaker, 1967) ont conclu à l'existence de peu de relations entre les facteurs caractéristiques de l'école et le rendement scolaire des élèves. Hanushek (1989), quant à lui, n'a trouvé aucune relation significative entre les ressources de l'école et le rendement scolaire des élèves. En revanche, Hedges, Greenwald et Laine (1994), en utilisant un sous-ensemble des données provenant de l'étude de Hanushek (1989), ont trouvé qu'il y avait au contraire un impact significatif des ressources de l'école sur le rendement scolaire des élèves. Dans une étude de comparaison internationale des systèmes scolaires, Wößmann (2003) a pour sa part montré que la différence dans les résultats scolaires des élèves ne peut pas être attribuée à la différence des ressources investies dans les écoles, mais plutôt à la structure organisationnelle propre à chaque institution scolaire.

Si les différences entre les établissements scolaires dans leur capacité à favoriser la réussite des élèves font partie des résultats établis, l'explication de ces différences reste à construire à l'intérieur d'une théorie formelle (Goldstein et Woodhouse, 2000). La controverse quant à la manière de déterminer l'« efficacité » d'une école persiste au sein de la communauté scientifique à propos des éléments suivants : la définition du concept de l'« efficacité » de l'école (Felouzis et Perroton, 2007), les méthodologies utilisées (Goldstein, 1991; Thrupp, 1999; Goldstein et Woodhouse, 2000; Stringfield, 2002; Blais, 2003; Stevens, 2005), l'impact réel des ressources de l'école sur le rendement scolaire des élèves (Hanushek, 1989; Hedges, Greenwald et Laine, 1994; Wößmann, 2003), ainsi que les modalités de publication et d'utilisation des résultats de recherche (Goldstein, Huiqi, Rath et Hill, 2000; Meyer, 2002; Blais, 2003).

Notre questionnement, dans le cadre de la présente recherche, se veut surtout à résonance méthodologique. Il porte sur la difficulté du choix des variables dépendantes et des variables indépendantes, sur la collecte et la disponibilité des données, ainsi que sur la manière de modéliser les liens entre les variables contextuelles et le rendement scolaire des élèves finissants du secondaire qui ont passé des épreuves standardisées du Ministère en français écrit, dans le contexte de la formation générale des jeunes.

Quelles seraient les variables plus sensibles à la différence des caractéristiques de l'élève et de l'école susceptibles d'être associées à l'« efficacité » d'une école secondaire? Comment départager les effets de chacune de ces variables en fonction de son niveau d'appartenance (élève, école) au rendement scolaire des élèves? Comment évoluent les résultats scolaires des élèves dans le temps? De ces interrogations générales découlent d'autres problèmes spécifiques de nature disciplinaire et méthodologique, problèmes dont nous nous préoccupons également.

Dans le cadre de la présente recherche, c'est la relation entre des variables caractéristiques des élèves et des écoles et les résultats scolaires des élèves finissants du secondaire en français écrit qui retient notre attention. À cet égard, le rapport d'un comité d'experts sur l'apprentissage de l'écriture (MELS, 2008, p. 33) a émis un certain nombre de recommandations pour l'amélioration de l'apprentissage de la langue d'enseignement. Les plus pertinentes pour l'orientation de notre recherche sont les suivantes :

- a) entreprendre une réflexion sur l'épreuve unique de 5e secondaire pour examiner la possibilité d'y offrir d'autres choix que le texte argumentatif et étudier le poids relatif du lexique dans les critères d'évaluation de cette épreuve;

- b) encourager les recherches pour mieux adapter l'enseignement de l'écriture et permettre son évaluation;
- c) mettre en place des conditions qui facilitent l'accompagnement dans l'apprentissage de l'écriture des élèves allophones, des élèves à besoins particuliers, telle la réduction du nombre d'élèves par classe.

Notre recherche se situe dans la foulée des recommandations d'une étude longitudinale de la Direction de l'évaluation du MELS, avec la collaboration du Laboratoire de recherche et d'intervention traitant des politiques et des professions en éducation (LABRIPROF) de l'Université de Montréal, relativement à la performance des élèves à l'épreuve unique d'écriture de 5<sup>e</sup> secondaire de 2002 à 2006 (MELS, 2008, p. 6). Selon les résultats de cette étude, les élèves semblent s'en tirer généralement assez bien en ce qui concerne l'organisation et la cohérence du texte. Toutefois, une partie d'entre eux éprouvent de la difficulté dans l'utilisation du lexique, de la grammaire et de l'orthographe.

À notre connaissance, aucune recherche n'a modélisé les résultats des élèves de 5<sup>e</sup> secondaire à des critères de correction de l'épreuve unique de français écrit du MELS, en fonction des variables caractéristiques de l'élève et de l'école. Pour les besoins de notre étude, nous avons donc examiné les résultats des élèves finissants du secondaire selon deux critères spécifiques de correction, la syntaxe et la ponctuation ainsi que l'orthographe. Signalons par ailleurs que d'un point de vue méthodologique, les résultats des élèves selon ces critères de correction est plus robuste que la note finale au bulletin, qui représente une moyenne pondérée des résultats en lecture, en écriture et à l'oral.

## 1.6 L'OBJECTIF DE LA RECHERCHE

L'objectif de la présente recherche, à portée méthodologique, est de modéliser les liens entre un certain nombre de variables indépendantes caractéristiques de l'élève et de l'école et les résultats des élèves finissants du secondaire, suivant deux critères spécifiques de correction (syntaxe et ponctuation, orthographe) de l'épreuve unique de français écrit du Ministère, sur une période de trois ans. En d'autres termes, nous comptons analyser l'ampleur et la différence de variabilité dans les résultats des élèves relativement aux deux critères spécifiques de correction, sous l'angle des facteurs caractéristiques de l'élève et de l'école, et ce, pour les années 2006, 2007 et 2008.

Cet objectif nous amène à formuler la question suivante : Comment modéliser les liens entre un certain nombre de variables caractéristiques des élèves et des écoles et les résultats des élèves finissants du secondaire selon deux critères spécifiques de correction à l'épreuve unique de français écrit du MELS, afin de déterminer l'ampleur de la variabilité des résultats des élèves?

Cette question nous conduira à :

- a) choisir, à partir de la littérature, les variables pertinentes caractéristiques de l'élève et de l'école parmi les nombreuses variables susceptibles d'influencer les résultats des élèves de 5<sup>e</sup> secondaire à l'épreuve unique de français écrit du MELS;
- b) analyser les différentes méthodes utilisées dans les recherches sur l'« efficacité » des écoles de manière à mieux orienter notre étude;
- c) modéliser les relations entre les variables pertinentes pour lesquelles nous disposons de données et les résultats des élèves selon les deux critères spécifiques de correction (syntaxe et

ponctuation, orthographe) de l'épreuve unique de français écrit du MELS.

Dans le chapitre suivant, qui présente le cadre théorique de notre étude, nous examinerons les données conceptuelles du courant de recherche sur l'« efficacité » des écoles.

## **CHAPITRE II**

### **LE CADRE THÉORIQUE**

Ce chapitre traite des éléments théoriques de la recherche : les indicateurs éducationnels, le rapport Coleman (Coleman, Campbell, Hobson, McPartland, Mood et York, 1966) et ses éléments déterminants, la recension des écrits sur l'« efficacité » de l'école en filiation avec le rapport Coleman. Il rappelle aussi les objectifs spécifiques de la présente recherche.

#### **2.1 LES INDICATEURS ÉDUCATIONNELS**

Un indicateur éducationnel se définit comme la description d'une manifestation observable de l'atteinte d'un niveau de compétence (MEQ, 2003, p. 7), dans laquelle peuvent apparaître les dimensions d'approximation, de description, de relation, de diagnostic, de variation dans le temps, de comparaison et d'évaluation d'un système éducatif. Il peut être représenté par des indices statistiques qui permettent d'exprimer des jugements de valeur à propos des aspects importants d'un système éducatif (Scheerens, Glass et Thomas, 2007). Ces formules de définition d'un indicateur éducationnel comprennent les composantes suivantes :

- a) l'idée que nous avons affaire à des caractéristiques mesurables d'un système éducatif;
- b) la délimitation des aspects importants d'un système éducatif peut fournir seulement de façon superficielle (Nutall, 1989) le profil des conditions actuelles du système, mais non de manière approfondie;
- c) l'exigence que les indicateurs précisent des éléments de la qualité de l'école implique qu'ils sont des statistiques présentant des points de comparaison à partir desquels on peut émettre des jugements de valeur.

L'utilisation d'indicateurs éducationnels pour apprécier la performance des établissements scolaires n'est pas toujours simple. Il peut s'avérer peu fécond d'avoir recours à un seul type d'indicateurs pour étudier les relations entre les résultats des élèves à des examens standardisés et les caractéristiques d'un système éducatif. Les auteurs Blais, Laurier et Pelletier (2001, p. 135) font remarquer que la principale difficulté observée lors de la mise au point d'un ensemble d'indicateurs ne réside pas tant dans le manque de sources d'informations que dans la nature imprécise de la définition de l'objet conceptuel du recueil de l'information. Ils ajoutent cette mise en garde :

« Un indicateur donné n'a que peu de sens s'il est traité isolément, hors contexte, et s'il n'est pas mis dans une perspective comparative, en regard des autres indicateurs du contexte. Il trouve sa pertinence dans une structure comparative longitudinale, car c'est l'accumulation successive de données dans le temps qui permet de contraster les différences et les évolutions. »

Un indicateur éducationnel devrait donc être interprété en fonction du contexte éducatif pour lequel il a été conçu. Laveault (2009, p. 178) apporte des précisions quant au suivi et à l'interprétation des tendances liées au rendement scolaire des élèves à l'aide d'indicateurs éducationnels. D'après cet auteur, ils présentent de nombreux défis qui se résument en trois points :

1. un choix d'indicateurs pertinents pour l'apprentissage;
2. le développement d'indicateurs fondés sur des statistiques appropriées;
3. la répartition des résultats parmi les différentes composantes d'une matière ou d'une habileté.

Les statistiques d'un système éducatif ne sont pas toujours des indicateurs utiles (De Landsheere, 1979; Bottani et Tuijnman, 1994; Fitz-Gibbon, 1996; Camilli et Firestone, 1999). Le choix des indicateurs éducationnels doit être déterminé non seulement par le contexte de la recherche, mais aussi en fonction de la disponibilité des données.

### **2.1.1 Les indicateurs des «effets écoles»**

Plusieurs auteurs (Teddlie, Reynolds et Sammons, 2000; Scheerens, Glass et Thomas, 2007) ont recensé une liste d'indicateurs qui représentent les «effets écoles» : ce sont les indicateurs du processus, les indicateurs de résultats, les indicateurs de réussite à long terme, les indicateurs d'impact et les indicateurs de pertinence. Chacun d'eux se rapporte à une composante du modèle de base du mouvement de recherche sur l'école efficace. Mais ceux qui nous intéressent représentent l'influence de la scolarisation sur la performance moyenne des élèves, lorsqu'ils sont ajustés aux caractéristiques familiales et/ou à l'habileté initiale des élèves. Selon Willms (1992), les

« effets écoles » peuvent être définis comme étant la mesure de l'effet unique de chaque école sur le rendement scolaire d'un élève au terme de sa scolarisation, dans le cas où celui-ci aurait fréquenté plusieurs établissements scolaires. Konstantopoulos (2006) utilise, quant à lui, le concept « effets écoles » pour indiquer la relation entre les ressources de l'école et le rendement scolaire des élèves tout en contrôlant leurs caractéristiques individuelles. Il estime que les facteurs d'école et le rendement scolaire entretiennent un lien strictement corrélational, non causal, dans le sens où on ne peut pas prédire avec certitude le rendement des élèves à partir des caractéristiques de l'école. En résumé, l'« effet école », tel qu'il est présenté par les auteurs Willms (1992), Teddlie, Reynolds et Sammons (2000) ainsi que Konstantopoulos (2006), semble être approprié pour représenter les relations entre des variables caractéristiques de l'école et les résultats des élèves à des épreuves standardisées.

L'utilisation que nous ferons de certains indicateurs éducationnels tiendra compte de l'objectif spécifique de la présente recherche et du contexte de sa réalisation : la modélisation des liens entre des variables caractéristiques de l'élève et de l'école et les résultats obtenus par les élèves finissants du secondaire de la région métropolitaine de Montréal à l'épreuve unique de français écrit du ministère de l'Éducation, du Loisir et du Sport (MELS).

## **2.2 LE RAPPORT COLEMAN (1966)**

Cette section examine l'étude qui est à l'origine du développement de la recherche sur l'« efficacité » des écoles. Cette étude a été effectuée par Coleman et ses collaborateurs (1966). Depuis sa publication, elle demeure l'une parmi les études qui ont provoqué énormément de commentaires au sein de la communauté scientifique.

### 2.2.1 Les éléments déterminants du rapport Coleman

La première étude sur les politiques éducatives, *The Equality of Educational Opportunity* (Coleman et collab., 1966), mieux connue sous le nom de « rapport Coleman », fut réalisée dans les années 1960. Son objectif consistait à fournir des informations précises sur l'état du système éducatif aux États-Unis à un moment donné de son histoire. De façon générale, cette recherche concluait que les facteurs d'école exerçaient peu d'influence sur les résultats scolaires des élèves. Plus spécifiquement, le rapport Coleman formulait les constats qui suivent.

- a) La scolarisation des élèves de la majorité blanche et de ceux de la minorité noire se fait dans les établissements scolaires en fonction de leur appartenance ethnique.
- b) La réussite scolaire varie principalement en fonction de l'appartenance sociale et ethnique des élèves. Les élèves des minorités ethniques autres que les minorités asiatiques réussissent moins bien que ceux de la majorité blanche.
- c) Les ressources matérielles et humaines investies dans les établissements scolaires fréquentés par les Blancs sont sensiblement équivalentes à celles investies dans les établissements scolaires fréquentés par les Noirs.
- d) Les enfants de la minorité noire des milieux défavorisés réussissent mieux lorsqu'ils sont intégrés dans les écoles qui sont fréquentées par les élèves de la classe moyenne de diverses origines ethniques. Ainsi, la ségrégation raciale serait un facteur explicatif des écarts de réussite entre les différentes minorités ethniques et le groupe majoritaire.

- e) De manière générale, la disparité entre la réussite scolaire des élèves de la minorité noire et de ceux de la majorité blanche se maintient tout au long des douze années de scolarité qui s'étalent du début du primaire à la fin du secondaire.

À l'aide de modèles mathématiques de régression linéaire multiple et de plus d'une centaine de variables, cette étude utilise cinq groupes de tests pour évaluer la réussite scolaire des élèves : les tests de performance verbale et non verbale, de compréhension à la lecture, de mathématique, de sciences de la nature, des arts et des humanités. Les scores obtenus à ces tests ont été transformés de sorte que la moyenne générale des élèves pour chaque matière est égale à 50 et l'écart-type à 10.

Au total, plus de 645 000 élèves venant de 4 000 écoles publiques ont passé des tests standardisés et ont rempli des questionnaires. La plupart des enseignants, les directeurs d'établissements, les directeurs généraux des districts scolaires ont eux aussi répondu aux questionnaires qui leur étaient destinés.

L'échantillon représentatif de l'ensemble de la population des écoles primaires et secondaires comprenait cinq des douze niveaux d'études de la maternelle à la fin des études secondaires. Il faut noter que 30 % des écoles qui avaient été sélectionnées n'ont pas participé à cette recherche. Les élèves de troisième et de sixième années du primaire ainsi que ceux de neuvième et de douzième années du secondaire ont répondu aux questionnaires. Seulement la moitié des élèves de première année des écoles primaires l'ont fait.

Cette recherche divisait le territoire américain en onze grandes régions. Le taux de participation des écoles variait selon les régions. Il fut de 72 % dans

les régions métropolitaines nord et ouest, de 65 % dans les régions métropolitaines sud et sud-ouest, de 82 % dans les régions non métropolitaines nord et ouest et de 61 % dans les régions non métropolitaines sud et sud-ouest.

Coleman et ses collaborateurs (1966) ont trouvé que la performance des élèves dépendait plus de l'influence de facteurs associés au statut social des parents que de facteurs d'école. Ils ont rapporté qu'entre 5 % et 35 % de la variance dans la performance des élèves étaient attribuables à des facteurs d'école (le nombre moyen d'élèves par classe, l'auditorium, la cantine, le gymnase, l'infirmerie, les livres gratuits, les livres récents, un libraire, les repas gratuits, etc.) de ces différents groupes ethniques : les Amérindiens, les Asiatiques, les Blancs, les Latinos et les Noirs. Les auteurs ont alors conclu que cette variabilité dans le score des élèves des groupes minoritaires était vraisemblablement attribuable à la petite taille de l'échantillon tiré des groupes minoritaires autres que les Noirs.

Le rapport Coleman a également fait état des désavantages qui pourraient être associés au fait que les groupes minoritaires utiliseraient dans le cadre familial une langue maternelle différente de la langue d'enseignement. Les auteurs cherchaient alors à vérifier si les élèves qui parlaient à la maison une langue différente de la langue d'enseignement étaient désavantagés lorsqu'ils commençaient l'école en première année du primaire. Et si c'était le cas, comment ce désavantage évoluait dans le temps, durant leur cheminement scolaire de la première année du primaire jusqu'à la fin du secondaire.

Selon ce rapport, il appert que les élèves des groupes minoritaires qui n'ont pas fréquenté une garderie avant leur entrée à l'école primaire et qui utilisaient chez eux une langue différente de la langue d'enseignement subissaient quelques désavantages mineurs comparativement à ceux du

groupe majoritaire dont la langue maternelle était l'anglais, la langue d'enseignement. Quant aux élèves de la douzième année du secondaire qui utilisaient à la maison une langue différente de la langue d'enseignement, il est apparu, dans leur cas, que les désavantages mineurs qu'on avait observés chez les élèves en première année du primaire diminuaient considérablement et même, pour plusieurs d'entre eux, finissaient par disparaître.

Mentionnons par ailleurs qu'il est difficile de pondérer l'usage des langues parlées à la maison par les élèves lorsque ceux-ci en emploient plusieurs. Par exemple, à l'époque du rapport Coleman, on présumait que les élèves des groupes minoritaires, d'origines ethniques différentes, utilisaient à la maison leur langue maternelle selon une fréquence indéterminée par rapport à l'anglais. À ce sujet, Coleman et collab. (1966) faisaient remarquer que les élèves qui utilisaient à la maison une langue maternelle autre que l'anglais étaient souvent victimes à l'école de situations non reliées à l'activité d'apprentissage proprement dite.

### **2.2.2 Des critiques du rapport Coleman**

Le rapport Coleman a suscité de nombreuses controverses en ce qui a trait aux concepts de réussite scolaire, de l'égalité des chances scolaires, aux choix méthodologiques et à ses conclusions. L'étude de Cain et Watts (1970) rapportée dans Cherkaoui (1978) remet en question la méthodologie utilisée par les auteurs du rapport Coleman. Selon cette étude, la méthodologie conditionne les conclusions du rapport sur l'égalité des chances scolaires. L'ensemble de l'entreprise est mis en doute à cause de l'absence de spécifications d'un modèle théorique. Cain et Watts (1970) proposent qu'un cadre théorique soit défini au préalable, qui ordonnerait le très grand nombre de variables indépendantes introduites dans le modèle de régression linéaire,

sans quoi l'interprétation des résultats de l'analyse de régression serait hasardeuse.

D'autre part, l'ordre selon lequel les variables ont été introduites dans le modèle touche l'un des points les plus sensibles de la démarche méthodologique adoptée dans l'analyse des données. Les auteurs du rapport Coleman (1966) introduisent dans l'équation d'abord les variables liées au statut social de la famille et en dernier lieu les variables scolaires. Selon eux, cette façon de procéder se justifie par le fait que les premières variables caractéristiques de la famille sont indépendantes de toute influence des caractéristiques de l'école. En tenant constant le statut social, ils estiment pouvoir mesurer l'influence des variables scolaires sur la réussite en les introduisant dans le modèle de régression linéaire et d'analyse de la variance. Mais Murnane (1975) a montré que la décision des auteurs du rapport Coleman d'introduire les variables caractéristiques de la famille dans l'équation de la régression linéaire avant celles de l'école a biaisé l'analyse et l'a empêchée de montrer l'importance de l'effet des établissements scolaires sur le rendement des élèves. Hanushek et Kain (1972) ont de plus fait valoir que, sur la base d'une analyse secondaire des données du rapport Coleman, lorsque l'on introduit d'abord les variables scolaires, celles-ci expliquent une part de variance dans la réussite des élèves beaucoup plus importante que les variables liées à l'origine sociale.

Les auteurs Bowles et Levin (1968), pour leur part, ont signalé deux sources d'erreur importantes dans le rapport Coleman. La première a trait au très fort taux de non-réponses qui biaise en partie la structure générale des données. On compte en effet pour les écoles secondaires plus de 40 % de non-réponses qui ne seraient pas le fait du hasard. Un nombre élevé d'écoles urbaines ont refusé de participer à l'enquête, de sorte que les écoles des banlieues se sont trouvées surreprésentées. La deuxième concerne les dépenses par école. Les auteurs ont ainsi observé qu'on a mesuré les

dépenses, non par école, mais par groupes d'écoles. Ce qui a pu minimiser la dispersion des dépenses par école.

D'autres observations portent sur les limites des résultats du rapport Coleman qui sont principalement dues à la démarche formelle utilisée. Celle-ci est orientée vers l'explication des relations entre les « entrées » et les « sorties ». Une telle démarche laisse échapper une catégorie de phénomènes dont l'influence sur la réussite scolaire des élèves est tout aussi importante que celle des entrées. On a reproché au rapport Coleman de ne pas modéliser les variables qui pourraient mesurer les vrais effets de l'école sur le rendement scolaire des élèves. Les auteurs du rapport ont, il est vrai, intégré quelques variables relatives à l'attitude de l'élève (le contrôle de l'environnement, l'intérêt pour la lecture, la perception de soi) dans le modèle d'analyse de régression linéaire multiple. Mais quelques analyses ont montré que ces variables avaient peu de lien avec les résultats scolaires des élèves. Coleman et collab. (1966) considèrent les caractéristiques suivantes de la classe : la compétence des enseignants, la disponibilité du matériel pédagogique et la qualité du curriculum, comme étant les vrais facteurs d'école. Ce sont ces indicateurs que certains auteurs – dont Bressoux, Leroy-Audouin et Coustère (1998) ainsi que Dumay et Dupriez (2009) – ont appelé par la suite les « effets classes » et qui ont été assimilés à des « effets écoles » dans le rapport Coleman.

Avec le choix de l'élève comme unité d'analyse, les auteurs du rapport Coleman ont déterminé que la réussite est une fonction croissante de la taille de l'établissement scolaire. Toutefois, cette relation disparaît lorsqu'on contrôle les équipements qu'offre l'école. Par la suite, les chercheurs qui choisissaient l'école comme unité d'analyse ont obtenu des résultats différents de ceux du rapport Coleman : la variance du rendement scolaire des élèves était davantage liée aux facteurs d'école qu'à ceux de la maison (Mayeske, Wisler, Beaton, Weinfeld, Cohen, Okada, Proshek et Tabler, 1969;

Brookover, Beady, Flood, Schweitzer et Wisenbaker, 1979; Teddlie et Stringfield, 1993).

Une autre critique adressée au rapport Coleman vient du fait que les auteurs ont modélisé les résultats des élèves à un examen standardisé dans la langue d'enseignement à l'oral. Un peu plus tard, les chercheurs Madaus, Kellaghan, Rakow et King (1979) ont recommandé de ne pas utiliser les résultats des élèves à des examens oraux comme variable dépendante : les tests standardisés qui portent sur la langue d'enseignement sont moins sensibles à la différence des caractéristiques de l'école que les tests de contenu spécifique, comme les sciences et les mathématiques, sur le rendement des élèves. Ces auteurs considéraient que les parents plus scolarisés peuvent influencer davantage la qualité de l'expression orale de leurs enfants.

Le rapport Coleman avait utilisé plusieurs modèles mathématiques de régression linéaire multiple. Les logiciels statistiques dédiés comme *The Hierarchical Linear Model* (HLM) des auteurs Raudenbush, Bryk et Congdon (2010) et MLwiN (Rasbash, Steele, Browne et Goldstein, 2009), qui permettent l'utilisation de modèles de régressions linéaires hiérarchiques et non linéaires pour l'analyse des données éducatives, n'étaient pas encore disponibles. Selon Plewis (1997, p. 35), en considérant la variabilité de chaque niveau du système éducatif (élève, classe, école), les modèles statistiques (hiérarchiques, multiniveaux) qui tiennent compte de la structure imbriquée des données apportent de meilleures précisions dans l'estimation des coefficients des variables explicatives que ne peut le faire la régression linéaire multiple. Rappelons que le rapport Coleman a quand même été produit dans les années 1960 : comme toute étude scientifique, elle est limitée à l'état des connaissances de l'époque et donc fortement liée au contexte de sa réalisation.

## **2.3 LA RECENSION DES ÉCRITS SUR L'« EFFICACITÉ » DES ÉCOLES EN FILIATION AVEC LE RAPPORT COLEMAN**

Les premières recherches sur l'« efficacité » des écoles examinaient la relation entre les ressources de l'école et le rendement scolaire des élèves à des tests standardisés. Les résultats de ces études ont révélé l'importance des caractéristiques familiales dans l'explication de la variance du rendement scolaire des élèves. Le rapport Coleman, initiateur de ce mouvement de recherche, a suscité un nombre important d'études sur l'« efficacité » des établissements scolaires aux États-Unis et ailleurs dans le monde.

Cette section présente quelques-unes des études qui ont suivi le rapport Coleman (1966). Elles ont été perçues comme des réponses aux résultats obtenus par les auteurs du rapport. Nous en faisons état ici parce qu'elles ont contribué au développement du mouvement de recherche sur l'école « efficace » et aussi parce qu'elles ont utilisé plusieurs variables et des méthodes d'analyse qui ont influencé des études subséquentes. Parmi les études réalisées hors Québec, nous avons retenu celles de Peaker (1967), Ainsworth et Batten (1974) en Angleterre, celles de Lee et Bryk (1989), Hanushek (1989), Hedges, Laine et Greenwald (1994) aux États-Unis, Opdenakker et Van Damme (2000) en Belgique, Rumberger et Palardy (2005), Konstantopoulos (2006), aux États-Unis et Petrucci (2010) en Suisse, pour leur pertinence à notre problématique de recherche. En ce qui concerne les études réalisées au Québec qui s'inscrivent dans le courant de recherche de l'école « efficace », nous retenons les suivantes : Blais (2003), McAndrew, Ledent et Ait-Said (2005) ainsi que Ouedraogo (2006). Les sections suivantes présentent un examen critique de chacune de ces études.

### **2.3.1 The Plowden Report (Peaker, 1967)**

The *Plowden Report* (Peaker, 1967), intitulé *Children and Their Primary School*, a suivi le rapport Coleman (Coleman et collab., 1966). Cette étude fut

réalisée en Angleterre et au pays de Galles. Elle a utilisé un échantillon de 2 348 élèves venant de 132 écoles primaires et un nombre considérable de variables, plus d'une centaine. Une première série d'analyses ont été effectuées sur une longue liste de variables, à la suite desquelles on est arrivé à en retenir 104 pour les écoles et 73 pour les élèves. Le choix définitif de ces variables découlait davantage toutefois de la taille de l'échantillon que des caractéristiques des écoles et des élèves. Ce traitement s'explique en partie par le fait que la procédure d'échantillonnage stratifiée appliquée déterminait un ensemble spécifique d'écoles et d'élèves; une procédure d'échantillonnage aléatoire aurait peut-être amené l'auteur à choisir des variables différentes de celles liées à une stratégie stratifiée qui aurait imposé un certain nombre d'écoles par région.

Dans le but d'obtenir des résultats plus robustes, l'auteur suggérait de regrouper les variables en trois classes représentant :

- 1) les attitudes parentales;
- 2) les conditions à la maison;
- 3) la scolarité des parents.

L'étude a montré l'importance primordiale du milieu familial sur les résultats scolaires des élèves en fin de scolarité primaire. Elle concluait, comme le rapport Coleman, que la différence de statut socio-économique des parents expliquerait davantage la variation quant au rendement scolaire des élèves : les facteurs socio-économiques, en fait, expliquaient 58 % de la variance des résultats scolaires.

Dans son étude, Peaker (1967) a utilisé une méthodologie différente de celle du rapport Coleman : le nombre de sujets de l'échantillon était nettement plus

réduit que celui du rapport Coleman. L'auteur a aussi effectué des analyses de régression linéaire multiple par étapes. Dans ce processus, des équations de régression successive sont produites par l'ajout d'une variable à la fois. C'est ainsi qu'ont été sélectionnées les variables indépendantes qui présentent une forte corrélation avec la variable dépendante. Le regroupement des variables en trois catégories selon le statut socio-économique des parents apportait à l'époque une amélioration dans la compréhension des relations entre les facteurs relevant du milieu familial et les résultats des élèves.

### **2.3.2 L'étude de Ainsworth et Batten (1974)**

Un peu plus tard, Ainsworth et Batten (1974) ont fait le suivi d'un échantillon d'élèves du *Rapport Plowden*. Ils ont retenu 53 établissements scolaires et 1 544 élèves parmi l'ensemble des données initiales. Ils ont attentivement consulté une partie des données de ce rapport dans le but d'étudier les relations entre les résultats scolaires des élèves et certains aspects de l'environnement des établissements de l'ordre d'enseignement secondaire de la ville de Manchester en Angleterre. Des analyses factorielles ont été effectuées sur des matrices de corrélations comprenant jusqu'à 113 variables. Les auteurs ont trouvé, à deux moments différents de la scolarisation des élèves, une corrélation très élevée entre leur rendement scolaire et leur statut social. En résumé, cette enquête met en évidence les aspects qui suivent.

1. Le poids des facteurs tenant du milieu familial augmente à mesure que les études se prolongent. L'impact de ces facteurs est associé non seulement au niveau atteint par les élèves en fin de scolarité primaire, mais aussi à des caractéristiques du milieu scolaire dans l'enseignement secondaire.

2. L'influence des facteurs qui relèvent du statut socio-économique et des attitudes parentales peut être compensée ou cumulée avec le temps. S'ils sont défavorables, ces facteurs peuvent contribuer à aggraver le sentiment d'infériorité de l'élève et à le garder dans les classes de second degré.
3. La compréhension de la relation entre le type d'établissement scolaire fréquenté par les élèves et la compétence des enseignants en rapport avec le rendement scolaire des élèves peut aider à moduler les interventions éducatives non seulement à l'égard des enfants, mais aussi auprès des parents.

Cette étude longitudinale, réalisée presque une décennie plus tard que le rapport Coleman, a appliqué des outils méthodologiques plus avancés que ceux utilisés dans ce dernier, comme des logiciels informatiques plus sophistiqués permettant de calculer les coefficients des variables avec de meilleures précisions. Ces conclusions confirment celles obtenues par l'étude de Peaker (1967) en regard de la relation entre les facteurs relevant du statut socio-économique des parents et la performance des élèves. Les auteurs ont cependant analysé de nouveaux phénomènes, comme le lien entre les résultats des élèves et leur estime de soi, d'une part, et la compétence des parents, d'autre part.

### **2.3.3 L'étude de Lee et Bryk (1989)**

Pour leur part, Lee et Bryk (1989), aux États-Unis, ont eu recours à la modélisation multiniveaux pour explorer la relation entre les secteurs éducatifs (catholique, public) et le rendement scolaire des élèves. Les données de l'étude sont constituées d'un échantillon d'élèves venant de 160 écoles secondaires et d'étudiants de la première année de quelques universités. Ces établissements d'enseignement ont été choisis de manière

aléatoire. Les données ont été analysées afin de déterminer l'influence de l'organisation scolaire du secteur d'appartenance de l'école (catholique ou publique) dans le rendement scolaire des élèves. La recherche conclut à l'existence d'une faible corrélation entre l'origine sociale et le rendement scolaire des élèves dans les écoles catholiques.

Les résultats de cette étude, rapportés également dans Bryk et Raudenbush (1988, p. 183), confirment ceux d'autres recherches (Coleman et Hoffer, 1987; Coleman et collab. 1982) : la performance des élèves dans les cours de mathématiques de base est plus élevée dans les écoles catholiques que dans les écoles publiques; la différence de performance entre les élèves des groupes minoritaires et ceux du groupe majoritaire est plus faible dans les écoles catholiques; d'autre part, la relation entre certaines variables caractéristiques des différentes classes sociales et la performance des élèves est moins stable dans les établissements scolaires catholiques. Les chercheurs ont mené des analyses supplémentaires sur d'autres variables qui expliquent davantage la performance des élèves en termes de différence dans le choix du curriculum, des normes scolaires et de l'environnement disciplinaire.

Ces résultats fournissent un support empirique à l'idée que l'organisation scolaire et les normes environnementales des écoles secondaires produisent un impact significatif dans la distribution sociale du rendement scolaire des élèves. La différence entre les deux secteurs, catholique et public, apparaît clairement. Il y a peu de choix dans le curriculum des écoles catholiques : celles-ci adoptent une approche proactive dans la désignation des matières qui doivent être enseignées. Les écoles du secteur catholique sont relativement de petite taille par rapport à celles du secteur public. Elles sont sécuritaires et offrent un climat d'ordre qui facilite l'apprentissage des élèves. L'étude a appliqué la modélisation multiniveaux, ce qui a permis d'analyser

l'effet différentiel des facteurs d'école selon le secteur sur la performance des élèves.

Ces trois premières recherches, *The Plowden Report* (Peaker, 1967), celle de Ainsworth et Batten (1974) et celle de Lee et Bryk (1989), ont utilisé les notes moyennes des élèves dans plusieurs disciplines pour analyser les relations entre les éléments du contexte et la performance des élèves. Les deux premières n'ont cependant pas appliqué la modélisation multiniveaux. Ces trois études ont, tout compte fait, permis de dégager l'importance des éléments suivants dans l'analyse et la modélisation d'une situation éducative :

- l'identification des variables propres à chaque entité (élève, classe, école);
- le statut socio-économique des parents;
- l'appartenance ethnoculturelle des élèves;
- l'utilisation de nouvelles méthodes d'analyse;
- l'accessibilité à des données secondaires;
- la taille de l'école;
- le secteur de l'école (public, privé, catholique);
- la méthode d'échantillonnage.

Il s'agit pour la plupart de concepts importants pour notre recherche. Certains, comme la taille, le secteur, les données secondaires, le choix des variables, seront analysés plus en détail dans le chapitre portant sur la méthodologie. Les études qui suivent ont appliqué des techniques et des outils méthodologiques plus récents qui vont influencer davantage l'instrumentation de notre recherche.

### **2.3.4 L'étude de Hanushek (1989)**

L'étude de Hanushek (1989) cherche à évaluer l'« efficacité » des écoles par la mesure du poids relatif des variables associées à l'environnement scolaire sur le rendement des élèves. L'auteur définit l'« efficacité » d'une école comme étant le maximum d'extrants possibles (rendement de l'élève) qui peuvent être obtenus à partir d'une combinaison donnée d'intrants. C'est une définition qui se rapproche de celle de l'efficience, que Hanushek considère comme étant la capacité d'une organisation à utiliser de façon optimale ses ressources. Autrement dit, c'est le rapport entre ce qui est réalisé et les moyens mis en œuvre pour y parvenir. Plus ce rapport est élevé, plus l'établissement scolaire en question est efficient.

Dans son étude, l'auteur a retenu les variables suivantes : le rapport du nombre d'élèves au nombre d'enseignants, la scolarité du personnel enseignant, l'expérience du personnel enseignant, le salaire du personnel enseignant, les dépenses par élève, les intrants administratifs et les ressources physiques de l'établissement scolaire. Le rendement scolaire est exprimé par les résultats des élèves à des tests standardisés en mathématiques, en sciences, en lecture et en écriture.

À l'aide de la fonction de production éducative, une modélisation déterminant le maximum d'extrants (rendement de l'élève) qui peuvent être obtenus d'une combinaison donnée d'intrants, Hanushek (1989) a analysé 38 articles scientifiques dans lesquels il y avait 187 équations de régression linéaire. Il a conclu qu'on ne pouvait pas affirmer de manière convaincante l'existence d'une corrélation significative entre les variables associées à l'environnement de l'école et le rendement des élèves. Il n'a donc pas trouvé de corrélation significative entre la réduction du nombre d'élèves par classe par enseignant, la formation des enseignants, le salaire des enseignants, les dépenses par élève, les dépenses administratives, la qualité de l'environnement physique

de l'établissement scolaire et les résultats des élèves à des tests standardisés. Il a cependant attribué la corrélation positive observée entre l'expérience du personnel enseignant et le rendement scolaire des élèves au fait que les enseignants plus expérimentés accaparent les élèves les plus doués au moment de l'affectation des tâches aux enseignants. Il reconnaît implicitement que les aptitudes individuelles des élèves sont un facteur de prédiction important de leur rendement scolaire futur.

Cette étude utilise plusieurs variables relatives à la salle de classe, au personnel enseignant, au personnel administratif, aux ressources physiques et financières de l'établissement scolaire. Nous n'avons pas l'intention d'utiliser des variables de ce type dans notre recherche. Toutefois, les résultats de l'étude restent importants pour celle-ci dans la mesure où il importe de comprendre, en lien avec l'étude de Hedges, Laine et Greenwald (1994) présentée ci-dessous, que les choix méthodologiques appliqués dans la sélection de l'échantillon de sujets et des outils statistiques pour le traitement des données peuvent être déterminants dans les conclusions d'une étude.

### **2.3.5 L'étude de Hedges, Laine et Greenwald (1994)**

Dans la foulée des critiques provoquées par l'étude de Hanushek (1989), les auteurs Hedges, Laine et Greenwald (1994) ont analysé les mêmes variables avec un sous-ensemble des données de l'étude précédente. En faisant le choix d'une méthodologie différente, les auteurs ont appliqué une analyse statistique qui compte le nombre de résultats significatifs pour chacune des variables, retenant parmi les études consultées uniquement celles qui avaient révélé que la disponibilité des ressources de l'école, principalement le financement, était reliée systématiquement et positivement au rendement scolaire des élèves dans tout l'univers éducationnel. Ils ont conclu que cette

corrélation était suffisamment significative pour être considérée comme un indicateur éducationnel important.

Contrairement à la conclusion de l'étude de Hanushek(1989), les auteurs Hedges, Laine et Greenwald (1994) ont soutenu que les variables caractéristiques de la classe relatives à la compétence du personnel enseignant – par exemple leur scolarité et leur expérience – présentent une forte corrélation avec le rendement scolaire de l'élève.

L'intérêt de cette étude comme celle de Hanushek (1989), relativement à notre recherche sur la modélisation des liens entre des variables caractéristiques de l'élève et de l'école et les résultats scolaires des élèves, est de montrer à quel point les résultats et les conclusions d'une recherche sont tributaires des choix méthodologiques. Il importera pour nous de faire montre de prudence dans le choix des variables dépendantes, des variables indépendantes et des méthodes d'analyses de données pour éviter des biais qui risqueraient de compromettre nos résultats et par le fait même les conclusions de notre étude.

### **2.3.6 L'étude de Opdenakker et Van Damme (2000)**

Dans une perspective d'analyse méthodologique, les auteurs Opdenakker et Van Damme (2000), en Belgique, ont exploré l'impact du choix du nombre de niveaux d'un modèle linéaire hiérarchique sur l'estimation des coefficients des variables du contexte. Ils ont utilisé des modèles à deux, trois et quatre niveaux pour la modélisation d'un échantillon de données composé de 2 680 élèves, de 150 classes, de 81 enseignants et de 46 écoles secondaires.

Les auteurs ont montré que le choix du modèle ainsi que le nombre de niveaux peuvent influencer de manière significative les résultats d'une étude. En effet, le fait d'ignorer un ou plusieurs niveaux dans une modélisation entraînera des conséquences sur l'estimation des coefficients des variables

des autres niveaux. À ce propos, les auteurs Raudenbush et Bryk (2002) ont cependant prouvé que les modèles hiérarchiques qui déterminent les effets d'établissement dans le rendement scolaire des élèves sont mieux représentés par les modèles à deux niveaux (élève, école).

L'étude de Opdenakker et de Van Damme, quoique limitée à cause du peu de sujets de l'échantillon, nous renseigne sur l'importance de bien déterminer le nombre de niveaux dans la modélisation d'une structure éducative. Mais la difficulté provenant de la conceptualisation des variables dépendantes et indépendantes, des indicateurs éducationnels (Blais et collab., 2001) ainsi que de l'accessibilité à des données fiables sur les variables joue souvent dans le choix du nombre de niveaux dans la modélisation d'un système éducatif.

Quoique notre propre échantillon de données comprenne un nombre élevé de sujets, les résultats de l'étude de Opdenakker et de Van Damme sont importants pour la nôtre, car nous avons l'intention d'utiliser des modèles hiérarchiques à deux niveaux (élève, école) et à trois niveaux (élève, école, année). Ces résultats nous invitent à la prudence dans l'interprétation des coefficients des équations de la modélisation multiniveaux multivariée, pour éviter de surestimer les coefficients des variables relatives à un niveau donné, comme l'élève ou l'école, à cause justement de l'absence dans l'analyse des variables relatives à la salle de classe.

### **2.3.7 L'étude de Rumberger et Palardy (2005)**

Les auteurs Rumberger et Palardy (2005) de leur côté ont étudié la relation entre plusieurs indicateurs relatifs à la performance des écoles, dont le score aux tests, le taux de décrochage scolaire et le taux de transfert. Les données de leur recherche ont été recueillies de l'étude *The National Education Longitudinal Study* (NELS) entreprise en 1988 aux États-Unis. Elles provenaient des résultats des élèves à des tests standardisés en

mathématiques, en anglais, en sciences, en histoire et en sciences humaines. Des données complémentaires émanaient d'un questionnaire soumis à 14 199 élèves choisis aléatoirement de l'échantillon initial de 25 000 élèves de huitième année, ainsi qu'à des parents, à des enseignants et à des directeurs d'école pour les années 1988, 1990 et 1992. D'autres éléments obtenus de l'ensemble initial de données ont été ensuite recueillis pour le suivi des élèves en 1990, 1992, 1994 et en 2000.

Afin de s'assurer de la robustesse des résultats, les auteurs n'ont pas retenu dans leur analyse les élèves dont les données manquaient à certaines variables pour la période couvrant les années 1988 et 1990, et ceux qui fréquentaient une école secondaire dont l'effectif comprenait moins de six élèves. Ils ont examiné un nombre important de variables habituellement associées au contexte éducatif, à l'intrant, au processus et à l'extrant, selon le modèle de base du mouvement de recherche sur l'efficacité des écoles (Scheerens, Glass et Thomas, 2007, p. 18).

Les variables retenues pour l'élève étaient le score moyen de chaque élève à quatre tests standardisés en mathématiques, en lecture, en sciences, en histoire, ainsi que la durée de la passation des examens. Les variables sur le plan de l'école étaient les suivantes : la persévérance scolaire, c'est-à-dire la capacité d'une école à retenir ses élèves jusqu'à la fin des études secondaires; le taux d'abandon scolaire, c'est-à-dire le pourcentage des élèves qui quittent l'école avant l'obtention de leur diplôme; le taux de transfert des élèves, à savoir le nombre d'élèves qui changent d'école ou qui sont orientés vers des programmes non sanctionnés par un diplôme; le taux d'attrition exprimé par l'addition du taux d'abandon au taux de transfert des élèves; la proportion d'élèves issus des minorités ethniques; la proportion d'élèves de famille monoparentale; le secteur (catholique, privé, public); la taille de l'école (petite, grande); la localisation (urbaine, rurale). Les variables indépendantes pour la classe étaient les suivantes : le rapport élèves-

enseignant; le salaire moyen du personnel enseignant; l'expérience du personnel enseignant; la scolarité du personnel enseignant; le temps moyen accordé aux devoirs; la proportion d'élèves victimes d'intimidation, etc.

Les auteurs ont utilisé trois types de modèles linéaires hiérarchiques : un modèle à trois niveaux pour la modélisation de la courbe de croissance du rendement scolaire des élèves sur une période de quatre ans, un modèle de régression logistique multinomiale à deux niveaux pour estimer le taux d'abandon et de transfert des élèves, enfin un modèle de régression logistique binomiale à deux niveaux pour estimer le taux d'attrition.

Les résultats de cette étude ont montré que les écoles qui sont efficaces pour favoriser la performance scolaire des élèves ne le sont pas nécessairement dans la réduction du taux d'abandon des élèves. Les auteurs suggèrent, en plus des résultats scolaires des élèves à des tests standardisés, d'utiliser d'autres indicateurs, comme la persévérance scolaire, pour mieux apprécier l'« efficacité » des écoles.

Cette étude longitudinale s'avère pertinente pour notre recherche parce qu'elle utilise plusieurs modèles de régression à deux niveaux ainsi qu'un nombre important de variables indépendantes caractéristiques de l'élève et de l'école susceptibles de nous intéresser, comme le secteur (public, privé), la taille de l'établissement scolaire, la proportion d'élèves de l'école issus de l'immigration qu'on pourrait assimiler à la proportion d'élèves ayant une langue maternelle différente de la langue d'enseignement. Cette étude a en outre permis de déterminer la courbe de croissance du rendement scolaire des élèves selon des données longitudinales sur une période de quatre ans. Nous ne disposons pas cependant de ce type de données pour notre recherche. Et celles que nous sommes en mesure d'utiliser ne nous permettent pas de procéder à cette forme d'analyse.

### **2.3.8 L'étude de Konstantopoulos (2006)**

Dans une étude portant sur la tendance des facteurs d'école et des caractéristiques du personnel enseignant en rapport avec la performance des élèves, Konstantopoulos (2006) a appliqué, d'une part, un modèle hiérarchique à deux niveaux pour analyser le poids relatif des facteurs associés à l'établissement scolaire dans l'explication de la variance du rendement des élèves et, d'autre part, un modèle hiérarchique à trois niveaux, quand des variables caractéristiques des enseignants étaient disponibles, pour expliquer le poids relatif des facteurs associés au personnel enseignant sur la performance des élèves.

Les données de sa recherche proviennent de trois importantes études menées sur une période de trente ans aux États-Unis. Chacune d'elles a retenu un échantillon aléatoire d'élèves de la huitième et de la douzième année du secondaire. Les résultats des élèves à des tests standardisés ont été recueillis, sur deux intervalles de dix ans, dans des matières comme la lecture, les sciences et les mathématiques. Dans l'ensemble des États-Unis, une première étude portant sur un échantillon de 15 800 élèves des écoles publiques et privées de la douzième année du secondaire a été réalisée en 1972; une deuxième étude sur le suivi des élèves a été effectuée en 1982 avec un échantillon de 26 216 élèves et la troisième, conduite en 1992, au moyen d'un échantillon de 12 921 élèves.

L'auteur a utilisé plusieurs variables pour chacun des trois niveaux de la structure éducative de sa modélisation (élève, enseignant, école). Il a retenu pour le niveau de l'élève les éléments suivants : le genre, l'appartenance à une minorité ethnique, le statut socio-économique; au niveau de l'école : la localisation de l'établissement scolaire (banlieusarde, urbaine, rurale), le secteur (public, privé), le pourcentage d'élèves qui suivaient des cours préparatoires à l'université, la durée de l'année scolaire, la proportion

d'élèves issus d'une minorité ethnique et le taux d'abandon scolaire; au niveau de l'enseignant : le nombre d'élèves par enseignant, l'assiduité et le rendement scolaire des élèves.

Relativement au rendement scolaire des élèves, les résultats de cette recherche ont révélé que la différence de variabilité entre les écoles apparaît de plus en plus importante avec le temps. Autrement dit, la performance des élèves à l'intérieur de l'école devient de plus en plus homogène. D'autre part, les facteurs d'école expliquent systématiquement plus de 50 % de la variance du rendement scolaire moyen des élèves. On a noté un effet considérable des facteurs caractéristiques des enseignants sur le rendement scolaire des élèves en sciences et en mathématiques.

Les données sur le personnel enseignant n'étaient pas disponibles pour deux des trois années. Seule l'étude sur le suivi des écoles secondaires pour l'année 1992 reliait un échantillon d'élèves non seulement à l'école, mais aussi à la classe et à l'enseignant, ce qui a permis d'appliquer une modélisation hiérarchique à trois niveaux. La difficulté d'accès à des données sur le personnel enseignant pour les deux autres années a cependant posé des limites à l'analyse, ainsi qu'à l'interprétation des résultats de la recherche sur la tendance de la variabilité des facteurs associés au personnel enseignant en regard du rendement scolaire des élèves.

Cette étude s'avère importante pour notre recherche, car elle fait appel à plusieurs variables relatives à l'élève et à l'école ainsi qu'à la modélisation multiniveaux dans l'analyse des données. Elle prend en compte l'effet des variables relatives à l'école sur le rendement scolaire de l'élève par la modélisation hiérarchique à deux niveaux, ce que nous avons aussi l'intention de faire. Mais nous nous empressons de préciser que nous procéderons par la modélisation multiniveaux multivariée : cette approche paraît tout à fait appropriée à la problématique de notre recherche.

### **2.3.9 L'étude de Petrucci (2010)**

Récemment, Petrucci (2010) en Suisse a réalisé une étude sur les relations entre les facteurs caractéristiques des élèves et de la classe et le rendement scolaire des élèves en compréhension de l'écrit. L'objectif de l'étude était de vérifier dans quelle mesure les résultats obtenus par les élèves pour des épreuves externes standardisées, à deux moments clés de leur scolarité obligatoire, dépendent des caractéristiques individuelles des élèves et aussi des caractéristiques de la classe dans laquelle ils sont scolarisés. L'auteur a donc voulu savoir s'il s'agissait des mêmes caractéristiques tout au long de la scolarité obligatoire. En mettant en œuvre des modèles linéaires hiérarchiques à deux niveaux sur la base des scores obtenus à partir d'une épreuve externe standardisée durant l'année scolaire 2006 en compréhension de l'écrit, il a tenté de trouver des éléments de réponse à ces questions.

L'étude a porté sur une population d'élèves dans le canton de Genève, en Suisse. Petrucci (2010) a utilisé trois regroupements : 238 classes de la 6<sup>e</sup> année primaire, qui marque la fin des études primaires, totalisant 4 173 élèves; 106 classes de la 9<sup>e</sup> année du regroupement A, des élèves forts, totalisant 2 340 élèves et 69 classes du regroupement B, des élèves moyens totalisant 960 élèves. Les regroupements A et B du cycle d'orientation marquent la fin de la scolarité obligatoire en Suisse.

Pour chacun des élèves, l'auteur disposait d'un score en compréhension de l'écrit, la variable dépendante. Il a analysé quatre variables indépendantes du premier niveau (élève) : l'origine sociale ou la catégorie socioprofessionnelle des parents, la langue maternelle et la nationalité, l'âge scolaire, l'option étudiée et le niveau des enseignements. Il a également analysé des variables indépendantes du deuxième niveau (classe) : le pourcentage de filles, de non-francophones et de non-Suisse, la composition sociale des classes

ainsi que la prédominance d'une option et l'âge moyen des élèves. De plus, en tenant compte du nombre réduit d'élèves dans les classes du regroupement B, des élèves aux habiletés ordinaires, et dans le but de caractériser de façon plus significative des différences de composition, l'auteur a réparti le niveau « classe » en trois sous-groupes, selon les quartiles : 25 % d'élèves forts, 50 % d'élèves moyens et 25 % d'élèves faibles.

En effectuant une modélisation hiérarchique à deux niveaux (élève, classe) du score en compréhension de l'écrit, Petrucci a commencé son analyse par un modèle à vide et a inclus, par la suite, certaines variables caractéristiques du niveau de l'élève et du niveau de la classe. Les résultats ont révélé que les différences entre les élèves d'une même classe sont beaucoup plus marquées que les différences entre les classes. Ainsi, les variations du score obtenu à l'épreuve sont davantage imputables à des différences individuelles qu'à des différences reliées aux variables retenues pour caractériser la classe. Le niveau classe contribue dans une mesure de 10,2 % au primaire, de 10,0 % au regroupement A et de 17,8 % au regroupement B. Le niveau élève joue pour sa part dans une mesure de 89,8 % au niveau primaire, 90 % au regroupement A et 82,2 % au regroupement B.

De façon générale, l'auteur a observé un effet positif de la langue maternelle sur les performances en compréhension de l'écrit dans l'enseignement primaire et aux deux regroupements du cycle d'orientation de 9<sup>e</sup> année, avec une certaine nuance selon les profils choisis par les élèves. La catégorie socioprofessionnelle à laquelle appartiennent les parents de l'élève a également une influence positive sur le score obtenu par celui-ci à l'évaluation externe dans les deux ordres d'enseignement, en faveur des élèves de la catégorie favorisée. Par ailleurs, en 6<sup>e</sup> année du primaire comme au cycle d'orientation en 9<sup>e</sup> année, le fait d'être en retard dans la scolarité est toujours associé négativement au score obtenu par les élèves. En revanche,

le fait d'être en avance n'a qu'un effet positif sur les performances mesurées en 6<sup>e</sup> année du primaire.

Cette étude apparaît particulièrement pertinente pour notre recherche, puisqu'elle a été réalisée dans un contexte assez semblable à la situation montréalaise, soit le canton de Genève en Suisse, où plusieurs communautés linguistiques cohabitent au sein d'une population urbanisée. Elle porte en outre sur la compréhension de l'écrit en français. Notons enfin qu'elle utilise aussi des modélisations multiniveaux pour analyser le poids relatif des variables associées aux deux niveaux (élève et classe) sur les résultats des élèves dans une épreuve externe, en compréhension de l'écrit en français.

#### **2.3.10 L'étude de Blais (2003)**

Blais (2003) a étudié les différences entre les écoles secondaires de l'ensemble du Québec quant aux résultats de leurs élèves à l'épreuve unique d'écriture pour le français, langue d'enseignement. Son étude a été entreprise à la suite de la publication du bulletin des écoles secondaires de l'Institut économique de Montréal, un classement des écoles secondaires qui assimile la qualité de l'école à sa position dans une mise en rangs (Marceau et Cowley, 2001). Or Goldstein (1991) en Angleterre et Blais (2003) au Québec avaient auparavant soutenu qu'il est hasardeux de faire des inférences sur la qualité des écoles en se référant à une simple mise en rangs établie suivant les résultats moyens des élèves d'une école à des examens standardisés dans deux ou trois disciplines.

En utilisant des données sur une période de huit ans (de 1994 à 2001) provenant du MELS, approximativement 62000 élèves de 5<sup>e</sup> secondaire par année âgés de 15 à 18 ans du secteur francophone, Blais (2003) a d'abord analysé la distribution des résultats des élèves de certaines sous-populations (l'âge, le genre, le réseau) et de certaines écoles. Il a conclu que le seul

examen d'une statistique de tendance centrale, comme les moyennes pour la population ou les sous-populations, même si ces moyennes reflètent des tendances générales, ne permet pas d'obtenir un portrait très nuancé des résultats pour chaque école.

Il a aussi analysé la relation entre le nombre d'élèves présents à l'examen, qui est une indication de l'effectif total de l'établissement scolaire, et la moyenne de l'école. Il a relevé que les écoles où beaucoup d'élèves passent l'examen présentent des performances moyennes plus qu'acceptables, que ce soit au public ou au privé. D'autres écoles à vocation particulière de plus petite taille présentent des moyennes plus faibles. L'auteur recommande la prudence dans l'interprétation des relations entre la taille de l'école et les résultats des élèves à l'épreuve unique de français écrit. Compte tenu de leur projet éducatif et de la clientèle desservie, certaines écoles de petite taille présentent une moins bonne performance pour ce qui est des résultats moyens des élèves en français et en mathématiques.

Blais a ensuite analysé la relation entre, d'une part, deux indices socio-économiques : l'indice de défavorisation du ministère de l'Éducation, du Loisir et des Sports (MELS) et le revenu familial moyen par école dans les établissements scolaires publics pour lesquels ces indices sont disponibles et, d'autre part, la moyenne des notes à l'épreuve de l'école. Il a observé plusieurs disparités régionales quant à cette relation.

Pour la région de Montréal, les écoles où les moyennes des élèves sont hautes sont surtout des écoles privées et publiques qui sélectionnent les élèves au début du secondaire au moyen de tests d'admission. À propos de la relation entre l'âge des élèves et les résultats à l'épreuve du MEQ, l'auteur a fait plusieurs constatations. D'abord, les élèves de 15 et 16 ans obtiennent des notes moyennes plus élevées que ceux qui sont plus âgés. De plus, le

poids des élèves de 17 et 18 ans, les élèves en retard, est toujours beaucoup plus important pour le réseau public que pour le réseau privé.

En outre, en utilisant une modélisation à deux niveaux pour les données de l'année 2001, Blais (2003) a intégré les variables « genre » (fille, garçon) et « âge » pour le premier niveau, celui de l'élève, et les variables « réseau » (public, privé) et « nombre d'élèves présents à l'épreuve » pour le deuxième niveau, celui de l'école. Il a cependant écarté la variable « note de l'école par élève » comme variable explicative, parce qu'il est difficile de faire de cette mesure un standard pouvant être appliqué à une école ou à une autre. C'est un constat que confirment les auteurs Schmidt, Houang et McKnight (2004) qui soulignent le problème de l'équivalence des scores provenant des tests standardisés élaborés par des établissements scolaires différents, dans leurs études portant sur la « valeur ajoutée » de l'école au rendement scolaire des élèves. D'après les auteurs Blais (2003) ainsi que Schmidt, Houang et McKnight (2004), la note fournie par les écoles introduit des distorsions qui rendent difficile l'interprétation de certaines estimations.

Les résultats des analyses multiniveaux indiquent que le niveau de l'école contribue dans une mesure de 17 % à l'explication de la variance totale observée, alors que le niveau de l'élève y contribue, lui, à 83 %. De plus, pour la mesure de 17 % d'explication attribuée aux différences entre les écoles, la variable « âge de l'élève » est nettement plus dominante que les autres variables introduites dans la modélisation.

À la lumière des données analysées, Blais note qu'il existe une différence appréciable entre les garçons et les filles, qui se maintient depuis 1994 et qui devrait être étudiée plus attentivement. Les garçons présentent systématiquement des résultats moyens et médians moins élevés que ceux des filles. De plus, la relation entre le revenu familial moyen et la moyenne

des élèves à l'épreuve de français écrit du MEQ ne peut être étudiée sans la prise en compte des disparités régionales.

En général, l'auteur a conclu que, pour la majorité des élèves et la majorité des écoles, ce ne sont pas les caractéristiques des écoles qui sont déterminantes pour les élèves qui les fréquentent. Ce sont plutôt les caractéristiques des élèves qui sont déterminantes pour les écoles qu'ils fréquentent. D'où l'importance prise par la sélection des élèves à l'entrée à l'ordre d'enseignement secondaire.

L'étude de Blais (2003) est importante pour notre recherche, parce que nous avons l'intention d'examiner plusieurs des variables qu'il a traitées. De plus, cette étude effectue une analyse descriptive des données ainsi que plusieurs modélisations hiérarchiques à deux niveaux (élève, école) en fonction des notes moyennes de l'élève en français écrit, obtenues du Ministère pour l'ensemble du Québec. En faisant le lien avec notre problématique de recherche, nous pensons cependant qu'il y a lieu de pousser davantage l'analyse des relations entre les caractéristiques des élèves et leurs résultats en français écrit en modélisant, parmi les composantes de l'examen qui forment la note moyenne, celles avec lesquelles les élèves éprouvent certaines difficultés. Dans le cadre de notre recherche, parmi les six critères retenus pour la correction de l'épreuve unique de français écrit du Ministère, ce sont les résultats à deux critères spécifiques de correction, la syntaxe et la ponctuation ainsi que l'orthographe, qui seront analysés.

### **2.3.11 L'étude de McAndrew, Ledent et Ait-Said (2005)**

Dans une étude portant sur la réussite scolaire des jeunes des communautés noires dans la région métropolitaine de Montréal, les auteurs McAndrew, Ledent et Ait-Said (2005) ont présenté un bilan statistique qui précise le cheminement et les résultats scolaires de ces élèves du secondaire aux épreuves ministérielles en fonction du secteur (français, anglais), du genre

(fille, garçon), de la langue maternelle, de l'âge, du lieu de naissance, du niveau d'entrée dans le système scolaire québécois et du rang socio-économique occupé. Pour l'ensemble des trois cohortes (1995, 1996 et 1997), les auteurs ont utilisé les données concernant 6 617 élèves, dont 86,9 %, soit 5 747 élèves, fréquentent le secteur français et les autres 870 élèves, le secteur anglais. Compte tenu des caractéristiques de chaque secteur, les auteurs avaient décidé de les traiter indépendamment et de présenter les résultats séparément.

Les résultats indiquent que c'est en français écrit, langue d'enseignement, que les élèves des communautés noires éprouvent davantage de difficultés. Leur pourcentage de réussite dans cette matière s'établit à 79,8 %, comparativement à 85,1 % pour les élèves issus de l'immigration, et à 89,6 % pour l'ensemble de la population. Toutefois, la langue maternelle de ces élèves ne semble pas jouer automatiquement le rôle auquel on s'attendrait. Les élèves des communautés noires originaires d'Afrique anglophone obtiennent en effet des résultats supérieurs, ou au moins équivalents, à ceux des élèves originaires des Antilles et d'Afrique francophone.

D'après les auteurs, le fait d'arriver à un âge « normal » au secondaire, c'est-à-dire à 12 ans, s'avère un facteur de prédiction important pour la réussite scolaire des élèves des communautés noires. Ils notent de plus que les élèves des communautés noires nés au Québec continuent de réussir de manière nettement moins favorable que les autres élèves. Ce constat devrait mener vers une exploration plus large de l'origine des difficultés qu'ils connaissent, particulièrement en matière de réussite dans la langue d'enseignement.

Cette étude s'avère importante pour notre recherche parce que, d'une part, elle utilise des données qui proviennent du ministère de l'Éducation, du Loisir et du Sport (MELS) – la même source de données que nous comptons

utiliser – et, d'autre part, parce qu'elle analyse les caractéristiques des élèves de la région métropolitaine de Montréal, comme l'âge, le genre, l'origine ethnique, la langue maternelle, des variables qui seront également retenues dans notre recherche.

### **2.3.12 L'étude de Ouedraogo (2006)**

Dans le prolongement de la recherche de Blais (2003), Ouedraogo (2006) a étudié la possibilité d'hybridation des données du recensement de Statistique Canada et des fichiers de résultats scolaires du ministère de l'Éducation, du Loisir et du Sport (MELS), afin de modéliser de nouvelles variables caractéristiques de l'environnement scolaire de l'élève. Son étude ainsi que celle de Blais (2003) nous seront d'une grande utilité dans la mesure où elles ont été réalisées au Québec avec sensiblement les mêmes variables caractéristiques des élèves et des écoles que nous avons l'intention de traiter.

Outre les variables relatives à l'élève (âge, genre), à l'école (réseau, taille) et au statut socio-économique des parents (revenu moyen par ménage, sous-scolarisation des mères, monoparentalité féminine), Ouedraogo a examiné la relation que certaines variables spécifiques à la condition résidentielle (l'accès plus difficile au logement, la mobilité et l'exiguïté résidentielle) pouvaient avoir avec les résultats scolaires des élèves. Il a constaté l'importance du poids relatif des variables caractéristiques de l'école par rapport à celles de l'élève dans la région de Montréal. Plus précisément, les variables associées au niveau de l'école comptent pour 22 % à 26 % de la variance totale du rendement scolaire des élèves, alors que Blais (2003) montrait que l'«école» accusait un poids relatif de 17 % de la variance totale quant à la performance des élèves pour l'ensemble du Québec. De plus, les résultats ont montré que, parmi les variables qui concourent à l'explication de la variance, l'exiguïté résidentielle indique le poids le plus important (7,8 %)

devant l'âge (7,7 %), la taille de l'école (4,2 %) et le genre (3,1 %). La variable « sous-scolarisation des femmes », bien que s'étant révélée significative, affiche un poids marginal (0,03 %). Il importe de souligner que nous n'avons pas l'intention d'utiliser, dans notre recherche, des variables relatives au statut socio-économique des parents ni à la condition résidentielle des élèves. Nous reconnaissons toutefois que la rétention d'un nombre plus élevé de variables des deux niveaux hiérarchiques (élève, école) dans la modélisation pourrait amener à une meilleure répartition des coefficients des variables, et conséquemment à une meilleure compréhension de l'influence des facteurs d'élèves et d'école sur la performance des élèves.

## **2.4 UNE SYNTHÈSE DES RECHERCHES PRÉSENTÉES**

Cette section fournit une brève synthèse des éléments déterminants des études retenues qui nous aideront à établir le cadre théorique de notre recherche. Les premières études du mouvement de recherche sur l'école efficace (Coleman et collab. 1966; Peaker, 1967; Ainsworth et Batten, 1974) utilisaient des échantillons élevés d'élèves, plus d'une centaine de variables, des modèles mathématiques de régression linéaire multiple et des matrices de corrélation. Elles ont montré, comme le rapport Coleman, que les facteurs relevant du statut socioéconomique des parents de l'élève étaient plus importants dans le rendement scolaire de l'élève que les facteurs d'école. Par la suite, les résultats des études de ce mouvement de recherche sont devenus contradictoires. Hanushek (1989) a montré que les ressources investies dans l'école n'avaient pas d'influence significative sur la performance des élèves, alors que Hedges, Laine et Greenwald (1994) ont soutenu le contraire. Les outils méthodologiques utilisés dans ces recherches semblaient influencer leurs conclusions. Les techniques de la modélisation

multiniveaux devenaient de plus en plus répandues. Elles constituent depuis une avancée importante et apportent de meilleures précisions dans le calcul des coefficients des variables. Opdenakker et Van Damme (2000) ont montré que le choix du nombre de niveaux dans une modélisation multiniveaux influence la mesure des coefficients de variables. Toutefois, en utilisant cet outil d'analyse statistique, les études de Blais (2003), d'Ouedraogo (2006) et de Petrucci (2010) ont modélisé les liens entre les facteurs caractéristiques de l'élève (l'âge, le genre et la langue maternelle, les conditions résidentielles), de la classe (composition sociale des élèves, les niveaux d'habiletés) et de l'école (le réseau, la taille, la mixité, la proportion des élèves de l'école de langue maternelle française) et la performance scolaire des élèves dans différentes matières. Les résultats de ces études situent le poids relatif des facteurs d'élèves entre 75 % et 90 % de l'explication de la performance des élèves.

## **2.5 RAPPEL SUR LES OBJECTIFS DE LA RECHERCHE**

Il est utile de rappeler à cette étape-ci que l'objectif de la présente recherche est de modéliser les liens entre un certain nombre de variables caractéristiques de l'élève et de l'école et les résultats des élèves finissants du secondaire de la région métropolitaine de Montréal, suivant deux critères spécifiques de correction (syntaxe et ponctuation, orthographe) de l'épreuve unique de français écrit du Ministère sur une période de trois ans. En d'autres termes, nous comptons analyser l'ampleur et la différence de variabilité dans les résultats des élèves relativement aux deux critères spécifiques de correction, sous l'angle des facteurs caractéristiques de l'élève et de l'école, et ce, pour les années 2006, 2007 et 2008.

## **CHAPITRE III**

### **MÉTHODOLOGIE**

Ce chapitre présente les éléments méthodologiques de la recherche : la provenance des données, les populations à l'étude (celles des années 2006, 2007 et 2008), les méthodes d'analyse de données, les modèles, les étapes de la modélisation multiniveaux multivariée et les considérations éthiques.

#### **3.1 LES SOURCES DE DONNÉES**

Les analyses qui seront effectuées dans cette recherche reposent sur les résultats de trois cohortes d'élèves soumis aux épreuves uniques de français écrit du ministère de l'Éducation, du Loisir et du Sport (MELS) durant les années 2006, 2007 et 2008. Les données qui nous intéressent ont été recueillies et sont gérées par le Ministère, en ce sens qu'il en est le gardien et le seul à pouvoir y autoriser l'accès. Les épreuves d'où proviennent les données sont imposées aux élèves finissants du secondaire en formation générale des jeunes inscrits dans les écoles publiques et privées du Québec. Les résultats obtenus contribuent à la prise de décision pour la délivrance du diplôme d'études secondaires. Ces données n'ont donc pas été recueillies

spécifiquement pour les fins de la présente recherche. Il a fallu effectuer les démarches nécessaires auprès des autorités compétentes pour les acquérir.

### **3.1.1 Les données secondaires**

L'analyse de données secondaires concerne l'utilisation de données recueillies par des recherches antérieures dans le but de poursuivre des objectifs de recherche différents. Selon les auteurs Hinds, Vogel et Clarke-Steffen (1997), cette procédure utilise le plus souvent des ensembles de données quantitatives. À la suite de la recension des écrits présentée plus haut dans la présente recherche, il est ressorti qu'une même base de données peut être analysée sous des angles différents dans plusieurs études.

Par exemple, les données provenant de *The National Education Longitudinal Study* (NELS, 1982) ont été utilisées dans l'étude de Rumberger et Palardy (2005) ainsi que dans celle de Konstantopoulos (2006), pour atteindre des objectifs de recherche différents. La première visait l'étude de la relation entre les facteurs associés à l'environnement scolaire et la performance des élèves, alors que la seconde examinait la tendance des facteurs caractéristiques des établissements scolaires sur le rendement scolaire des élèves.

Dans le même ordre d'idées, Hanushek (1989), d'une part, et Greenwald, Hedges et Laine (1996), d'autre part, ont analysé les mêmes données secondaires avec deux méthodologies distinctes. Ils ont abouti à des résultats différents : le premier affirmait que les facteurs d'école et du personnel enseignant n'exerçaient aucune influence sur le rendement scolaire des élèves, et les autres chercheurs, par la suite, ont soutenu le contraire.

Blais (2003), McAndrew et collab. (2005) et Ouedraogo (2006) ont utilisé les fichiers de données du ministère de l'Éducation, du Loisir et du Sport (MELS) normalement recueillies pour les fins de régulation du système éducatif. Ouedraogo (2006) a en outre utilisé les fichiers du MELS et ceux de Statistique Canada pour étudier la possibilité d'hybridation des fichiers de données provenant de plusieurs sources, en vue de prédire la relation entre certaines variables relatives au statut socio-économique des élèves et leur performance scolaire.

Dans le cadre de la présente recherche, nous allons nous aussi utiliser des données provenant du MELS pour modéliser les liens entre certaines variables caractéristiques des élèves et des écoles et les résultats des élèves à deux critères spécifiques de correction de l'épreuve unique du français écrit de 5<sup>e</sup> secondaire du MELS.

### **3.2 LES POPULATIONS À L'ÉTUDE**

La présente recherche porte sur la totalité des écoles et des élèves finissants du secondaire du secteur francophone de la région de Montréal qui ont passé l'épreuve unique de français écrit du ministère de l'Éducation, du Loisir et du Sport (MELS) durant les années 2006, 2007 et 2008. Ce territoire comprend trois commissions scolaires francophones et plusieurs établissements scolaires privés de langue française.

Les fichiers de données originaux reçus du MELS, en format Excel, contiennent les résultats des élèves selon les deux critères de correction de l'épreuve de français, la syntaxe et la ponctuation ainsi que l'orthographe. Ces fichiers présentent aussi les données sur les variables caractéristiques de l'élève (l'âge, le genre, la langue maternelle) et de l'école (l'effectif, le nombre d'élèves de 5<sup>e</sup> secondaire présents à l'examen, le secteur, l'indice du

milieu socio-économique des écoles publiques (IMSE) pour les trois populations à l'étude, soit celles des années 2006, 2007 et 2008. Nous ne retenons pas la variable IMSE pour l'analyse, parce que les mesures concernant cette variable sont disponibles seulement pour les écoles publiques.

Après «nettoyage» des fichiers originaux, nous avons considéré pour l'analyse les élèves qui sont âgés entre 15 ans et 18 ans et les écoles qui présentaient au moins 20 élèves à l'épreuve unique de français écrit du Ministère. Ainsi, la population de la première cohorte, celle de l'année 2006, contient 10208 élèves venant de 75 écoles secondaires, dont 44 écoles publiques et 31 écoles privées. La population de la deuxième cohorte, celle de l'année 2007, contient 11 044 élèves venant de 77 écoles secondaires, dont 45 écoles publiques et 32 écoles privées, et la troisième, celle de l'année 2008, contient 10 231 élèves venant de 75 écoles, dont 41 écoles publiques et 34 écoles privées. L'effectif de chaque établissement scolaire varie considérablement, indépendamment de la population considérée (2006, 2007 et 2008), de 105 à 2 200 élèves.

La section suivante sur l'instrumentation de la recherche justifie nos choix des deux variables dépendantes et d'un certain nombre de variables indépendantes caractéristiques des élèves et de l'école qui seraient susceptibles d'influencer les résultats des élèves finissants du secondaire à l'épreuve unique de français écrit du ministère de l'Éducation, du Loisir et du Sport (MELS).

### **3.3 L'INSTRUMENTATION**

La plupart des études du mouvement de recherche sur l'« efficacité » de l'école qui ont été réalisées depuis la publication du rapport Coleman

(Coleman et collab., 1966) modélisent la relation entre certaines variables associées à l'environnement de l'école et le rendement scolaire des élèves. Plusieurs chercheurs prétendent avoir étudié toutes les variables qu'il était possible de conceptualiser.

### **3.3.1 Les variables dépendantes**

Le choix des variables dépendantes dans les recherches sur l'« efficacité » des écoles se fait non seulement en fonction des objectifs spécifiques de ces recherches, mais aussi selon la disponibilité des données. Des chercheurs ont montré que les résultats des élèves à des tests standardisés en mathématiques sont plus sensibles à la différence des caractéristiques de l'école, au contraire de la langue d'enseignement, où l'apprentissage peut être largement influencé par les facteurs externes à l'école, comme le statut socio-économique de la famille (Madaus et collab, 1979; Mandeville et Anderson, 1987).

Rappelons que notre étude prendra en compte deux critères spécifiques de correction de l'épreuve de français relatifs au code linguistique comme variables dépendantes : la syntaxe et la ponctuation ainsi que l'orthographe. La décision de modéliser les deux critères de correction comme variables dépendantes est justifiée par les constats suivants :

1. les résultats des élèves selon ces deux critères spécifiques de correction sont faibles (MELS, 2008; p. 6) parmi l'ensemble des six critères de correction de l'épreuve de production écrite qui regroupent l'organisation du texte et la maîtrise du code linguistique;
2. à notre connaissance, aucune étude n'a modélisé conjointement ces deux critères spécifiques de correction de l'épreuve unique de français écrit du MELS par une analyse multiniveaux multivariée, comme nous avons l'intention de le faire;

3. la correction centralisée permet d'obtenir une certaine uniformité dans la correction de l'épreuve;
4. les résultats des élèves selon ces critères sont constitués de données brutes, plus robustes pour l'analyse que la note finale en pourcentage qui représente une moyenne pondérée des six critères de correction de l'épreuve.

Les auteurs Bock, Wolfe et Fisher (1996) ainsi que Meyer (2002) utilisent à des fins de comparaison et d'évaluation des établissements d'enseignement le score moyen des élèves aux tests pour calculer la « valeur ajoutée » de l'école sur le rendement scolaire des élèves relevant d'un même district aux États-Unis. D'autres chercheurs – dont Plewis (1997), Goldstein, Huiqi, Rath et Hill (2002), en Angleterre; Blais (2003), au Québec, Felouzis (2004) et Bressoux, (2007) en France – ont cependant montré que cette mesure, si elle est la seule information utilisée pour apprécier l'apport de l'établissement scolaire sur le rendement des élèves, non seulement ne traduit pas la réalité de l'école, mais introduit des distorsions qui rendent difficile l'estimation des résultats à cause des risques de biais d'agrégation.

Le biais d'agrégation indique qu'une corrélation au niveau macro ne peut pas être utilisée pour tirer des conclusions sur le niveau micro. En d'autres termes, on ne peut pas se fonder uniquement sur des données agrégées, comme la note moyenne des élèves de l'école, pour tirer des conclusions sur des comportements individuels. On court alors le risque d'affirmer des faussetés. Ce n'est pas qu'il soit inadmissible de raisonner au niveau macro, mais il faut dans ce cas limiter l'interprétation à ce niveau, sans chercher à inférer directement ou indirectement en termes de comportement des individus du niveau micro (Bressoux, 2007, p. 274, 275).

La présente étude modélisera des données brutes, soit le nombre d'erreurs des élèves selon les deux critères spécifiques de correction, la syntaxe et la ponctuation ainsi que l'orthographe. Pour situer le contexte dans lequel on fait passer l'épreuve, les sous-sections qui suivent apportent des justifications supplémentaires quant aux formes, aux conditions de passation de l'épreuve ainsi que la grille d'évaluation de celle-ci.

### **3.3.2 Les formes et les conditions de passation de l'épreuve**

Bien des années avant la Réforme de l'éducation commencée en l'an 2000, le système scolaire québécois s'est doté d'une épreuve unique en français écrit à caractère certificatif. Elle vise à évaluer les connaissances et les compétences acquises par les élèves finissants du secondaire en fonction des objectifs d'apprentissage définis dans le *Programme de formation de l'école québécoise* (MELS, 2003). Dans le but de répondre aux exigences de l'épreuve ministérielle, du moins jusqu'en 2009, puisque des modifications y ont été apportées depuis, l'élève devait rédiger un texte argumentatif dans lequel il défendait une position sur un sujet à enjeu critique dans le but de convaincre un destinataire donné du bien-fondé de son point de vue. Pour être en mesure d'écrire un texte de qualité et de soutenir la position adoptée à l'aide d'arguments pertinents, il disposait d'environ une semaine avant l'épreuve de façon à bien se documenter sur le sujet abordé. À cette fin, un dossier de départ, constitué de textes courants et littéraires dans lesquels était abordée la thématique retenue pour l'épreuve, avait été remis à l'élève qui était invité à compléter la recherche par ses propres moyens. Au moment de l'épreuve, on précisait la situation d'écriture dont l'élève devait tenir compte. La production attendue était un texte d'environ 500 mots rédigé en un maximum de trois heures. L'élève pouvait avoir à sa disposition une feuille de notes de lecture, qui ferait l'objet d'un contrôle, ainsi que quelques ouvrages de référence sur le code linguistique.

C'est ainsi que, chaque année, on proposait aux candidats un sujet nouveau pour la production écrite, d'où notre choix d'analyser les trois cohortes, celles des années 2006, 2007 et 2008, séparément. Rappelons à ce propos que le problème de l'équivalence des scores provenant de tests standardisés différents a été soulevé dans plusieurs études, dont celle de Blais (2003). Les tests standardisés, provenant d'établissements scolaires différents, dans un même domaine d'apprentissage, ne sont pas toujours équivalents, que ce soit en termes d'habiletés ou de contenu (Schmidt, Houang et McKnight, 2004).

### **3.3.3 La grille d'évaluation de l'épreuve considérée dans cette étude**

La grille d'évaluation de la compétence à écrire un texte argumentatif, qui s'intitule « Épreuve unique d'écriture du français, langue d'enseignement » (MEQ, 2003), précise que les trois critères du code linguistique représentent 50 % de la pondération de la note de passage, l'utilisation du vocabulaire approprié compte pour 5 %, la syntaxe et la ponctuation pour 20 % et l'orthographe pour 25 %.

Élaborée pour l'évaluation de la compétence à écrire, cette grille prend appui sur le profil du scripteur de textes courants défini dans le *Programme d'études, le français enseignement secondaire* et décrit les différents niveaux que doivent atteindre les finissants du secondaire. Il faut souligner que la grille ne décrit pas tous les aspects du programme, mais bien les principaux traits caractéristiques nécessaires pour juger de la compétence à écrire un texte dans un contexte d'évaluation sommative des élèves de fin d'études secondaires au Québec. La grille a changé en 2009. Mais les ajustements apportés n'ont aucune incidence sur les exigences qui sont maintenues.

Le Ministère a cependant émis depuis 2005 une disposition selon laquelle les textes comportant 35 erreurs d'orthographe et plus sont «examinés attentivement». Ce nombre d'erreurs est présenté comme un point de repère

pour l'évaluation formelle d'un texte de 500 mots, rédigé dans un temps limité et avec des ressources restreintes. L'évaluation du critère *Orthographe* devrait faire appel, comme celle des autres critères, au jugement professionnel de l'enseignant. Elle ne devrait pas se réduire au simple comptage des erreurs, mais prendre en compte leur nature, leur récurrence, la complexité des phrases, la longueur du texte (MELS, 2011). On demande donc aux correcteurs d'arrêter de compter le nombre d'erreurs des élèves à partir de 24 erreurs pour la syntaxe et ponctuation et de 35 erreurs pour l'orthographe.

Cette disposition a eu pour effet de produire une fréquence anormalement élevée de ces valeurs. Devant une telle situation, nous envisageons d'appliquer trois séries d'analyses à chaque variable dépendante pour chacune des cohortes de l'étude, celles des années 2006, 2007 et 2008 : l'ensemble de la population de chaque cohorte qui réunit tous les élèves de l'année à l'étude; une première sous-population qui comprend les élèves dont les résultats se situent entre zéro et vingt-trois erreurs de syntaxe et de ponctuation et entre zéro et trente-quatre erreurs d'orthographe; une deuxième sous-population qui contient des élèves en difficulté, ceux qui font 24 erreurs et plus en syntaxe et ponctuation et 35 erreurs et plus en orthographe. En procédant de cette façon, nous comptons dégager les caractéristiques des élèves de chacune des sous-populations à l'étude. Les sections suivantes présentent les justifications du choix des variables indépendantes caractéristiques de l'élève et de l'école.

### **3.3.4 Les variables caractéristiques de l'élève**

**L'âge de l'élève** : l'âge normal des élèves de 5<sup>e</sup> secondaire au 30 septembre de l'année où ils ont à passer l'examen de français écrit du Ministère est de 16 ans. Cependant, certains d'entre eux, mais peu nombreux, sont âgés de 14, 15, 17, 18 ans et plus. Plusieurs chercheurs – dont Goldstein, Huiqi, Rath

et Hill (2000), en Angleterre, Blais (2003), au Québec, Felouzis (2004), en France, et Petrucci (2010), en Suisse – ont établi que l'âge de l'élève est un facteur significatif du résultat scolaire obtenu. Plus spécifiquement, Blais (2003) a montré que les élèves qui arrivent à l'âge normal, c'est-à-dire à 16 ans, en cinquième secondaire au Québec réussissent mieux en français que ceux qui sont plus âgés. Également, selon McAndrew, Ledent et Ait-Said (2005), arriver à l'âge normal au secondaire, c'est-à-dire à 12 ans, s'avère un facteur de prédiction important de la réussite scolaire pour les élèves des communautés noires de Montréal. Il nous semble donc approprié d'examiner plus attentivement la relation entre l'âge et les résultats obtenus par tous les élèves de la région de Montréal selon les deux critères spécifiques de correction de l'épreuve unique de français écrit du MELS.

**Le genre de l'élève** : une étude de Blais (2003) pour le Québec a montré que les filles obtiennent de meilleurs résultats que les garçons à l'épreuve unique de français écrit du MELS en 5<sup>e</sup> secondaire. Selon le rapport du *Program for International Student Assessment* (PISA, 2006, p. 11) sur les résultats obtenus par les élèves en compréhension de l'écrit, les élèves québécois se classent au 4<sup>e</sup> rang des 57 pays participants. Tout en tenant compte de la bonne performance des élèves du Québec en compréhension de l'écrit et du faible écart entre les garçons et les filles en comparaison des autres provinces canadiennes, il y a lieu de vérifier plus en profondeur la relation entre le genre (garçon, fille) et les résultats des élèves de la région de Montréal aux deux critères spécifiques retenus pour notre recherche.

**La langue maternelle** : selon le ministère de l'Éducation, du Loisir et du Sport (MELS, 2006), la langue maternelle est la première langue apprise à la maison dans l'enfance et encore comprise par le recensé au moment du recensement fédéral. Au Québec, la majorité des immigrants, dont la langue maternelle n'est pas le français, habitent la région métropolitaine de Montréal. Selon McAndrew, Ledent et Ait-Said (2005, p. 33), c'est en français écrit,

langue d'enseignement, que les élèves des communautés noires de la région métropolitaine de Montréal éprouvent le plus de difficulté.

Les études de Fertig et Wright (2004, p. 6), qui analysent les données du *Program for International Student Assessment* (PISA), malgré l'absence des données provenant du Canada et du Québec pour l'année 2004 – notre contexte de recherche –, ont montré que les élèves qui utilisent à la maison une langue différente de celle de l'enseignement obtiennent des résultats scolaires inférieurs à ceux dont la langue maternelle est celle de l'enseignement. La présente recherche envisage donc de modéliser la langue maternelle des élèves comme variable indépendante en fonction des deux critères spécifiques de correction.

### **3.3.5 Les variables caractéristiques de l'école**

La diversité des caractéristiques des établissements d'enseignement dans la région métropolitaine de Montréal nous amène à analyser plusieurs variables associées à l'environnement de l'école : le réseau scolaire (public, privé), la proportion dans l'école des élèves de langue maternelle française, la mixité de l'école (mixte, fille, garçon), la taille de l'école.

**Le réseau scolaire (public, privé) :** au Québec, le secteur privé est particulièrement important dans les régions urbaines, comme Montréal, où se concentrent la majorité des écoles de ce secteur (Blais, 2003). Rumberger et Palardy (2005), aux États-Unis, ont conclu à l'existence d'une relation différenciée entre l'appartenance de l'école à un réseau public ou privé (catholique) et le rendement des élèves. La présente recherche s'intéresse à la différence dans les résultats obtenus par les élèves finissants du secondaire pour chacun des deux critères de correction, la syntaxe et la ponctuation ainsi que l'orthographe à l'épreuve unique de français écrit du MELS en fonction du réseau d'appartenance de l'école (public ou privé).

**La proportion des élèves de l'école de langue maternelle française :** la proportion d'élèves issus de l'immigration dans une école peut influencer l'apprentissage de la langue d'enseignement. Plusieurs études concluent, d'une part, à l'existence des « effets écoles » sur le rendement scolaire des élèves en fonction de leur origine ethnique (Coleman et collab., 1966; Nuttal et collab., 1989; Meyer, 2002; Stevens, 2005) et, d'autre part, à une forte corrélation entre la proportion élevée d'élèves immigrants dans une école et leur rendement scolaire (Coleman et collab. 1966; Rumberger et Palardy, 2005; Konstantopoulos, 2006).

Le Québec reçoit de plus en plus d'immigrants en provenance de pays de cultures, de langues et de religions différentes. Certaines écoles de la région métropolitaine de Montréal accueillent un pourcentage d'élèves immigrants qui dépasse les 50 % de la clientèle. Selon McAndrew, Ledent et Ait-Said (2005), 77 % des jeunes Noirs du secteur francophone issus de l'immigration au Québec se retrouvent dans la région de Montréal. Le pourcentage des élèves de l'école de langue maternelle française est calculé selon l'identification de la langue maternelle déclarée par chaque élève de l'école : c'est le rapport entre le nombre d'élèves de l'école de langue maternelle française et l'effectif total de l'école.

**La mixité de l'école (mixte, fille, garçon) :** la mixité de l'établissement scolaire est déterminée par le genre des élèves de chaque école, une information qui est consignée dans les fichiers de données reçus du MELS. On peut savoir si l'inscription à l'école est réservée aux filles, aux garçons et si celle-ci est composée des élèves des deux sexes. Si l'établissement scolaire accueille des élèves des deux sexes, il est catégorisé comme étant mixte. Autrement, il sera qualifié du genre des élèves qui le fréquentent. Les données reçues pour les trois cohortes (2006, 2007 et 2008) nous indiquent qu'il n'y avait aucune école réservée aux garçons dans la région

métropolitaine de Montréal. On comptait cependant une dizaine d'écoles réservées aux filles.

Plusieurs études révèlent l'existence d'un effet différentiel des écoles mixtes sur le rendement scolaire des élèves en faveur des filles (Nuttal et collab., 1989; Goldstein, Huiqi, Rath et Hill, 2000; Meyer, 2002; Stevens, 2005; Konstantopoulos, 2006). La présente recherche s'intéresse à la variabilité des résultats obtenus par les élèves de 5<sup>e</sup> secondaire pour chacun des deux critères de correction (la syntaxe et la ponctuation, l'orthographe) à l'épreuve unique de français écrit du MELS en fonction de la mixité des élèves de l'école (filles, mixte), étant donné qu'il n'y a pas d'école réservée aux garçons.

**La taille de l'école** : à Montréal, certaines écoles secondaires présentent un effectif qui dépasse les 2 000 élèves, alors que d'autres sont de plus petite taille, avec un effectif de moins de 500 élèves. Les résultats d'études récentes (Konstantopoulos, 2006) confirment l'existence d'une corrélation significative entre la taille de l'école et les résultats des élèves : les écoles de plus petite taille seraient plus performantes. En ce qui concerne le Québec, Blais (2003) a cependant montré qu'il faut user de prudence dans l'interprétation des relations entre la taille de l'école et les résultats des élèves. Compte tenu de leur projet éducatif et de la clientèle desservie, certaines écoles de petite taille présentent une moins bonne performance en termes de résultat moyen des élèves en français et en mathématiques.

Il faut signaler que, quelle que soit sa taille, l'établissement scolaire doit présenter au moins 20 élèves à l'épreuve unique de français écrit du MELS pour être retenu dans les deux types d'analyse : descriptive et multiniveaux. Cette exigence constitue l'une des conditions d'application de la modélisation multiniveaux pour des raisons de robustesse des résultats de la modélisation.

D'autres variables intéressantes relatives à l'établissement scolaire pourraient évidemment être utilisées comme variables indépendantes, mais elles ne sont pas accessibles. Cet empêchement représente l'une des limites de notre recherche.

### **3.3.6 Les variables caractéristiques de la classe**

Les variables relatives à la salle de classe ou du personnel enseignant ne seront pas analysées dans la présente recherche, principalement à cause des difficultés d'accès à des données relatives à ces variables. À titre d'information, citons-en quelques-unes : la scolarité du personnel enseignant, l'expérience du personnel enseignant, le nombre d'élèves par enseignant et le genre du personnel enseignant. Toutefois, pour une meilleure appréciation de notre étude, il convient de présenter ici ce que retient la littérature à propos de la relation entre ces variables et le rendement scolaire des élèves.

Les études de Darling-Hammond, Holtzman, Gatlin et Heilig (2005) ont conclu à l'influence positive du niveau élevé de scolarité du personnel enseignant sur le rendement scolaire des élèves. Les résultats des études de Zvoch et Stevens (2004) rapportés dans Stevens (2005) estimaient au contraire que la scolarité du personnel enseignant ne constitue pas un facteur significatif dans le rendement scolaire des élèves. Darling-Hammond et ses collaborateurs (2005) ont aussi soutenu que l'expérience du personnel enseignant demeure un facteur important pour la réussite scolaire des élèves, et Hanushek (1989) avait lié un tel résultat à la capacité des enseignants expérimentés à choisir les élèves les plus forts.

Le rapport nombre d'élèves par enseignant dans la salle de classe influence les résultats scolaires des élèves. Les auteurs Rumberger et Palardy (2005) ainsi que Konstantopoulos (2006) ont établi l'influence positive du nombre réduit d'élèves par enseignant sur le rendement des élèves. Les recherches de Nye, Hedges et Konstantopoulos (2000) avaient aussi indiqué en outre

que cette influence positive du nombre réduit d'élèves par classe par enseignant est constante pour ce qui est de la performance des élèves dans tous les types d'établissements d'enseignement, qu'ils soient privés ou publics.

Petrucci (2010), pour sa part, n'a pas considéré les caractéristiques du personnel enseignant. Il s'est limité à l'étude de l'effet de composition des différents groupes d'élèves dans la classe. Il a montré que le regroupement d'élèves en fonction de leur catégorie sociale et de leurs habiletés causait un impact sur leur réussite scolaire.

En résumé, les conclusions des études sur les relations entre les variables caractéristiques de la classe ou du personnel enseignant et le rendement scolaire des élèves divergent. D'où la nécessité d'approfondir cet aspect de la problématique en tenant compte de chaque contexte scolaire particulier.

### **3.3.7 La liste des variables**

Cette section présente la liste des variables dépendantes et des variables indépendantes ainsi que leur niveau d'appartenance qui seront analysés dans la présente recherche. La justification du choix de ces variables a été présentée dans la section précédente.

- Variables dépendantes

Deux critères de correction de l'épreuve unique de français écrit du MELS

- Syntaxe et ponctuation
- Orthographe

- Variables indépendantes au niveau de l'élève

- Âge de l'élève au 30 septembre de l'année scolaire (15, 16, 17 et 18 ans)
- Genre de l'élève (fille, garçon)

- Langue maternelle (français, autres)
- Variables indépendantes au niveau de l'école
  - Réseau d'appartenance (public, privé)
  - Mixité de l'école (mixte, fille)
  - Taille de l'école
  - Proportion des élèves de l'école de langue maternelle française

### **3.4 LES MÉTHODES D'ANALYSE DES DONNÉES**

Avant d'entreprendre une analyse, il est nécessaire de vérifier les données afin de s'assurer que les mesures des variables à l'étude, sur l'ensemble des nombres réels, sont recueillies autant que possible sans erreurs, qu'elles répondent aux attentes et qu'elles ne contiennent pas de valeurs aberrantes.

Les valeurs aberrantes sont définies comme étant la mesure des variables qui sont situées à plus de trois écarts-types de la valeur moyenne de la distribution. Généralement, quatre raisons expliquent leur présence qui peut fausser radicalement le résultat d'une analyse; leur traitement statistique requiert alors des actions différentes (Osborne, 2002) :

1. si les valeurs aberrantes proviennent de la collecte des données, il faut vérifier attentivement les mesures des variables et noter celles qui sont en dehors des valeurs attendues;
2. selon le logiciel statistique utilisé, on peut ne pas spécifier les codes recommandés pour le calcul des valeurs manquantes; si c'est le cas, il faut s'assurer que les valeurs manquantes ne sont pas traitées de la même manière que les données réelles;

3. on peut aussi recueillir un mauvais échantillon de données; il faut dans ce cas éliminer de l'échantillon les sujets qui ne font pas partie de la population;
4. pour les distributions atypiques, on peut choisir soit d'opérer une transformation des données pour ramener les valeurs extrêmes dans des proportions attendues, ou de les analyser séparément sans les transformer; c'est également la position adoptée par Howell (2002, p. 32).

Compte tenu des particularités des données qui proviennent des résultats des élèves selon les deux critères spécifiques de correction, à savoir une fréquence élevée de la valeur 24 en syntaxe et ponctuation et de la valeur 35 en orthographe, nous avons choisi de les traiter séparément. Nous présenterons dans les sections suivantes de notre étude des précisions sur les caractéristiques des modèles, la modélisation multiniveaux multivariée et les étapes de la modélisation.

### **3.4.1 Les modèles**

Les modèles sont utilisés dans les recherches en sciences humaines pour faciliter notre compréhension des phénomènes sociaux. Bradley et Schaefer (1998) définissent un modèle comme un cadre d'analyse formel ou informel qui cherche à synthétiser la complexité à partir des caractéristiques critiques du monde réel pour la compréhension des relations institutionnelles et techniques sous-jacentes, afin de faciliter l'explication du phénomène et de faire des prédictions. Modéliser, c'est construire une représentation générale et simplifiée du réel; c'est ébaucher un plan, un schéma, à partir duquel on pourra essayer différentes fictions particulières, ou simulations, de la chose que l'on veut représenter (Van der Maren, 1995). Dormont (1999, p. 10) pour sa part apporte une précision supplémentaire sur les exigences de la

démarche de modélisation : «modéliser exige une certaine parcimonie», c'est-à-dire que le chercheur doit utiliser ce qui est strictement nécessaire et suffisant pour expliquer le phénomène. Il faut donc veiller à formuler une représentation acceptable de la relation étudiée.

Dans le cadre de la présente recherche, en complémentarité avec les analyses descriptives, nous ferons plusieurs modélisations multiniveaux multivariées pour mieux comprendre, d'une part, la variabilité dans les résultats obtenus par les élèves en fonction des variables de chacun des niveaux (élève, école, année) et, d'autre part, les relations entre les variables caractéristiques de l'élève et de l'école et les résultats des élèves par rapport aux deux critères spécifiques de correction (la syntaxe et la ponctuation, l'orthographe) à l'épreuve unique de français écrit du MELS.

### **3.4.2 La modélisation multiniveaux multivariée**

La modélisation multiniveaux multivariée vise à l'ajustement de modèles qui ont plus d'une variable dépendante. Snijders et Bosker (1999, p. 201) précisent que la spécification des modèles est l'un des aspects les plus difficiles de l'inférence statistique, car elle repose sur deux leviers principaux qu'il convient de manipuler simultanément : des considérations liées au sujet de la recherche et des considérations statistiques. Le but de la spécification est de parvenir à un modèle qui décrit correctement les données observées sans complications inutiles, d'où la notion de parcimonie de Dormont (1999). En considérant les difficultés inhérentes à la modélisation multiniveaux, Petrucci (2010) rappelle que les étapes de la spécification des modèles multiniveaux est une tâche très complexe pour laquelle il n'existe toutefois aucune règle claire et précise.

La plupart des données éducatives présentent une structure hiérarchique dans laquelle les élèves sont regroupés à l'intérieur d'une classe et les classes à l'intérieur d'une école. Ce type de données est particulièrement

adapté aux modèles multiniveaux. L'équation de leur modélisation à coefficients aléatoires peut contenir plusieurs termes d'erreurs, un pour chaque niveau. Selon certains auteurs (Hill et Rowe, 1996; Plewis, 1997), l'utilisation de la modélisation multiniveaux se justifie par le fait qu'elle permet d'obtenir de meilleures précisions dans l'analyse des données, particulièrement celles qui proviennent d'un système éducatif.

D'autres auteurs, dont Cohen et Baldi (1998), ont cependant montré qu'il est préférable d'utiliser *The Clustering-robust linear regression* (CRLR), quand toutes les conditions à l'application des modèles multiniveaux ne sont pas satisfaites, particulièrement les conditions d'indépendance et de normalité de l'échantillon. Mais plusieurs raisons justifient l'utilisation de la modélisation multiniveaux pour l'analyse des données éducatives. Les modèles multiniveaux permettent d'introduire en même temps dans l'analyse les variables individuelles et les variables agrégées (Bressoux, Leroy-Audoin et Coustère, 1998). Ces modèles rendent en effet caduque la question du choix de l'unité d'analyse, entre l'élève et l'école, qui avait eu cours à l'époque du rapport Coleman (1966). Ils permettent d'éviter le dilemme agrégation versus désagrégation. On peut en effet intégrer directement et simultanément les différentes variables indépendantes, chacune au niveau qui est le sien dans la structure hiérarchisée des données éducatives.

Même s'il est possible d'analyser chacun des deux critères de correction séparément, les auteurs Snijders et Bosker (1999, p. 201) recommandent de les analyser conjointement pour plusieurs raisons. Premièrement, on peut conclure sur l'importance de la corrélation entre les deux variables dépendantes, notamment en indiquant dans quelle mesure cette corrélation dépend des variables caractéristiques du niveau de l'élève ou du niveau de l'école. Deuxièmement, les tests des effets spécifiques sur chaque variable dépendante sont plus puissants dans les analyses multivariées. Cela est évident sous la forme d'une plus petite erreur type dans l'estimation des

coefficients. Troisièmement, vérifier si l'effet d'une variable indépendante  $x_i$  sur une variable dépendante  $Y_1$  est plus important que sur une autre variable dépendante  $Y_2$  quand les données, partielles ou totales, sur les deux variables dépendantes ont été observées sur les mêmes individus, comme c'est le cas de nos données, est possible seulement avec les analyses multivariées. Et enfin, si on veut faire un test unique de l'effet conjoint d'une variable indépendante sur plusieurs variables dépendantes, alors la modélisation multivariée devient nécessaire. De surcroît, ce type de test permet d'éviter l'effet du hasard qui est inhérent à l'utilisation d'un test spécifique pour chaque variable dépendante. Néanmoins, l'analyse d'une situation éducative par la modélisation multiniveaux multivariée est plus compliquée que par la régression linéaire multiple.

Dans la présente recherche, nous appliquerons plusieurs modélisations hiérarchiques multivariées. L'équation 1 représente un modèle à trois niveaux et deux variables dépendantes sans les variables indépendantes et sans interactions entre celles-ci.

$$Y_{hijkl} = \gamma_{01} + \gamma_{02} + v_{1l} + v_{2l} + \mu_{1kl} + \mu_{2kl} + \lambda_{1jkl} + \lambda_{2jkl}, \quad h=1,2 \quad (\text{Éq. 1})$$

- $\gamma_{01}$  = interception de la première variable dépendante
- $\gamma_{02}$  = interception de la deuxième variable dépendante
- $v_{1l}$  = résidu première variable pour année l
- $v_{2l}$  = résidu deuxième variable pour année l
- $\mu_{1kl}$  =résidu première variable de l'école k pour l'année l
- $\mu_{2kl}$  =résidu deuxième variable de l'école k pour l'année l
- $\lambda_{1jkl}$  =résidu première variable de l'élève j pour l'école k et l'année l
- $\lambda_{2jkl}$  =résidu deuxième variable de l'élève j pour l'école k et l'année l.

Les variables dépendantes sont représentées par  $Y_{ijkl}$ , où l'indice  $j$  représente les élèves,  $k$  les écoles et  $l$  les années. Ainsi la variable  $Y_{1jkl}$  représente la première variable dépendante, la syntaxe et la ponctuation, de l'élève  $j$  dans l'école  $k$  pour l'année  $l$ . La variable  $Y_{2jkl}$  représente la deuxième variable dépendante, l'orthographe, de l'élève  $j$  dans l'école  $k$  pour l'année  $l$ .

Dans la présente recherche, nous utilisons plusieurs modèles multiniveaux multivariés avec deux variables dépendantes : la syntaxe et la ponctuation, l'orthographe. Chaque variable dépendante aura sa propre équation dans laquelle nous pourrions ajouter plusieurs variables indépendantes. Dans l'analyse exploratoire des variables indépendantes, nous avons retenu pour le niveau élève : l'âge, le genre et la langue maternelle; pour le niveau école : le réseau, la proportion des élèves de l'école de langue maternelle française, la taille et la mixité. Ces variables ont été choisies tant au niveau de l'élève que de l'école parce qu'elles possèdent une corrélation significative avec les résultats des élèves selon les deux critères de correction à l'épreuve unique de français écrit du MELS au niveau de signification de 5 %.

### **3.4.3 Les précisions sur les étapes d'analyse**

La poursuite de ces objectifs nous conduit à préciser les étapes d'analyse de la présente recherche. Nous analysons les données en deux étapes. La première consiste en trois séries d'analyses descriptives avec le logiciel SPSS (2010). La deuxième consiste en plusieurs modélisations multiniveaux multivariées avec le logiciel MLwiN (Rasbash, Steele, Browne and Goldstein, 2009).

1. Trois séries d'analyses descriptives des variables dépendantes (syntaxe et ponctuation, orthographe) et de leur lien avec les variables indépendantes caractéristiques de l'élève et de l'école pour chacune des trois cohortes, celles des années 2006, 2007 et 2008. La première série d'analyses se fera sur l'ensemble des données de

chaque cohorte. La deuxième série d'analyses se fera sur une première sous-population d'élèves sans les mesures de 24 erreurs en syntaxe et ponctuation et de 35 erreurs en orthographe. La troisième série d'analyses se fera sur une deuxième sous-population, celles des élèves en difficulté, c'est-à-dire ceux qui font 24 erreurs en syntaxe et ponctuation et 35 erreurs en orthographe.

2. Pour l'ensemble des élèves de chacune des cohortes et des sous-populations d'élèves sans les valeurs de 24 erreurs de syntaxe et ponctuation et de 35 erreurs en orthographe, nous ferons plusieurs modélisations multiniveaux multivariées à vide en ajoutant au fur et à mesure un niveau (élève, école, année) à la structure des données. Nous passerons ensuite aux modélisations multiniveaux multivariées qui intègrent dans le modèle retenu les variables caractéristiques du niveau de l'élève et du niveau de l'école.

#### **3.4.4 Les considérations éthiques**

Le comité d'éthique de la faculté des études supérieures de l'Université de Montréal nous a d'abord délivré un certificat d'éthique, lequel nous permet de poursuivre la recherche selon les objectifs précis définis dans le protocole de recherche. À la suite de notre demande auprès de la Commission d'accès à l'information (CAI), la Direction de la statistique, de la recherche et de l'information du ministère de l'Éducation, du Loisir et du Sport (MELS) nous a fourni les données suivantes dont il demeure le dépositaire légal : les informations sur le nombre d'erreurs relevées dans les copies des élèves par les correcteurs du MELS à propos des deux critères spécifiques de correction (la syntaxe et la ponctuation, l'orthographe) à l'épreuve unique de français écrit du MELS ainsi que les données relatives à d'autres variables contextuelles, caractéristiques de l'élève (l'âge, le genre, la langue maternelle) et de l'école (le secteur, l'effectif, le nombre d'élèves qui se

présentent à l'examen, l'indice du milieu socioéconomique (IMSE) des écoles publiques). L'autorisation est assortie de certaines exigences d'utilisation qu'il importe de signaler : les données doivent être utilisées uniquement pour les fins de la présente recherche; l'interdiction formelle de les croiser avec d'autres banques de données existantes ou futures; les données seront conservées une année après la fin de l'étude, puis elles seront détruites.

C'est maintenant le moment de présenter les analyses de données et les résultats de la recherche.

## CHAPITRE IV

### LES ANALYSES DESCRIPTIVES DES DONNÉES

Les nombreuses procédures statistiques utilisées dans l'analyse des données, comme les tests d'hypothèses, la distribution des résidus dans les analyses multiniveaux, etc., reposent sur l'hypothèse de la normalité, c'est-à-dire que la distribution de fréquence des mesures des variables et de leurs résidus est présumée conforme à la loi normale. Cette conformité, il est vrai, n'est pas toujours assurée au moment de la collecte des données brutes. Il a été toutefois établi de manière générale qu'une violation mineure de la condition de la normalité des données a peu ou n'entraîne pas d'effets significatifs sur les conclusions des études (Cohen, 1969; p. 266, 267).

Nous effectuerons dans ce chapitre, d'abord la description des données, les trois séries d'analyses statistiques descriptives pour chacune des trois cohortes à l'étude, celles de 2006, de 2007 et de 2008. La première série d'analyses se fera sur l'ensemble des élèves de chaque cohorte. La deuxième série portera sur les résultats de chaque sous-population d'élèves sans les valeurs de 24 erreurs en syntaxe et ponctuation et de 35 erreurs en orthographe. La troisième série d'analyses sera réalisée sur la sous-population d'élèves en difficulté, c'est-à-dire sur les résultats de ceux qui font 24 erreurs en syntaxe et ponctuation et 35 erreurs en orthographe. Nous

adopterons cette approche de trois séries d'analyses pour chaque cohorte, parce que nous estimons que ceux qui font beaucoup de fautes font partie d'une sous-population différente nécessitant par le fait même une analyse séparée.

#### 4.1 LA DESCRIPTION DES DONNÉES DES TROIS COHORTES

Cette section présente l'analyse descriptive des trois cohortes, celles des années 2006, 2007 et 2008. Pour l'année 2006, après vérification et nettoyage des données, nous avons retenu 10208 élèves venant de 75 écoles, dont 31 établissements d'enseignement privé et 44 écoles publiques, qui ont présenté respectivement 3555 et 6653 élèves à l'épreuve unique de français écrit du ministère de l'Éducation, du Loisir et du Sport (MELS). Les figures 2 et 3 montrent les distributions de fréquence pour les deux variables à l'étude : la syntaxe et la ponctuation ainsi que l'orthographe.

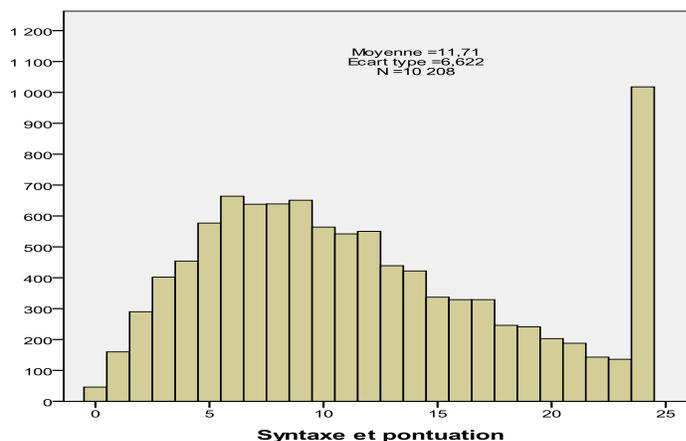
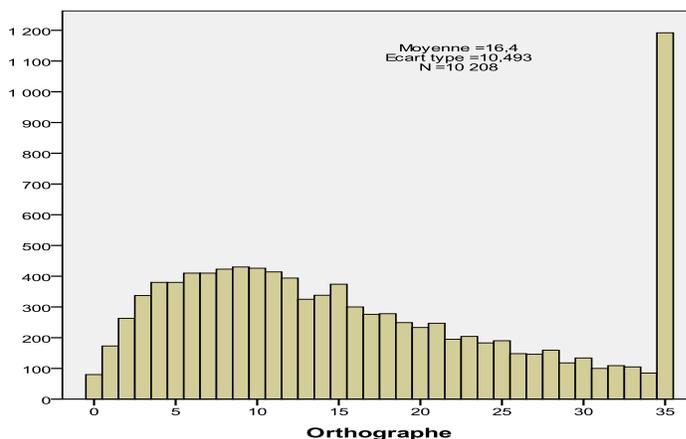


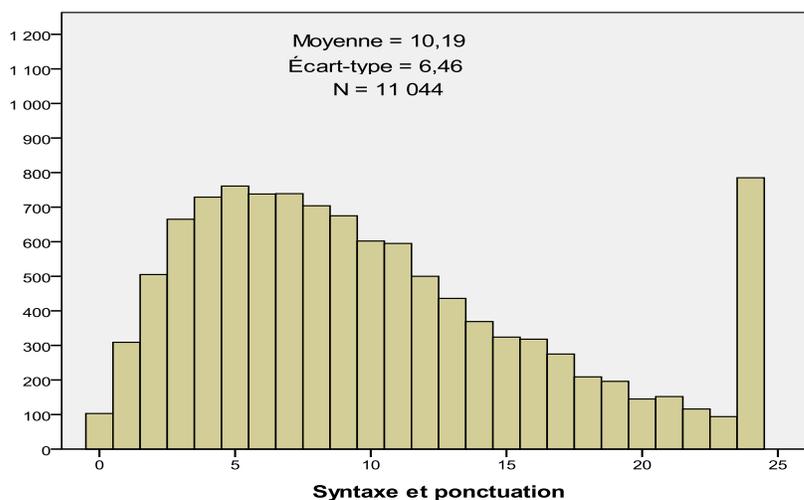
Figure 2 - Distribution des résultats des élèves en syntaxe et ponctuation en 2006



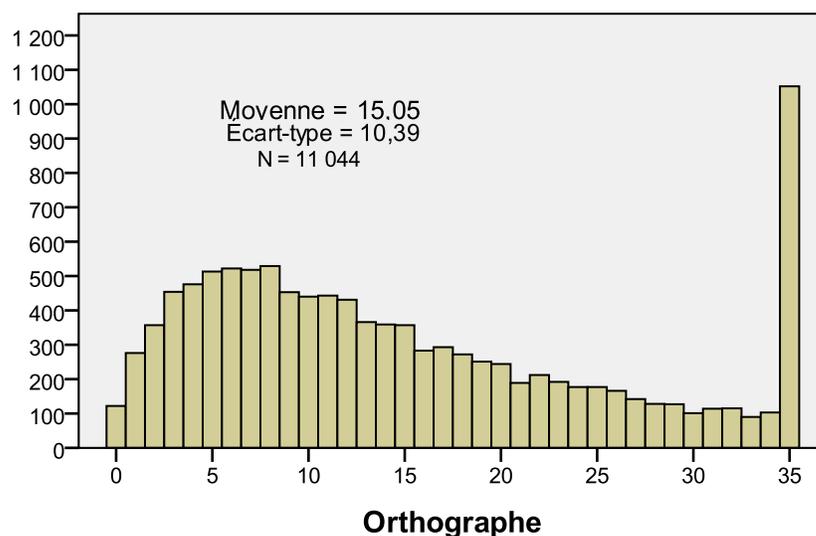
**Figure 3 - Distribution des résultats des élèves en orthographe en 2006**

Ces figures indiquent une distribution de données avec une fréquence très élevée d'une valeur à l'extrême droite. La fréquence élevée de cette mesure représente celle à partir de laquelle les correcteurs du MELS arrêtent de compter les erreurs des élèves, soit 24 pour la syntaxe et la ponctuation et 35 pour l'orthographe.

Pour 2007, on a retenu 11044 élèves. Ces élèves viennent de 77 écoles : 7093 d'entre eux sont inscrits dans 45 écoles publiques et 3951 dans 32 écoles privées. Les figures 4 et 5 illustrent les distributions de fréquence respectivement pour la syntaxe et la ponctuation et l'orthographe. Ces distributions de fréquence des résultats des élèves pour les deux variables à l'étude sont assez semblables à celles de 2006; une majorité des élèves se répartissent selon une distribution montrant une fréquence très élevée au bout de chaque distribution.



**Figure 4 - Distribution des résultats des élèves en syntaxe et ponctuation en 2007**



**Figure 5 - Distribution des résultats des élèves en orthographe en 2007**

Pour 2008, nous avons retenu 10231 élèves venant de 75 écoles secondaires : 6485 d'entre eux sont inscrits au secteur public et 3746 au secteur privé. Les figures 6 et 7 illustrent les distributions de fréquence pour les variables à l'étude. On constate une grande similitude dans la distribution

des résultats des élèves pour les trois cohortes à l'étude, soient celles des années 2006, 2007 et 2008.

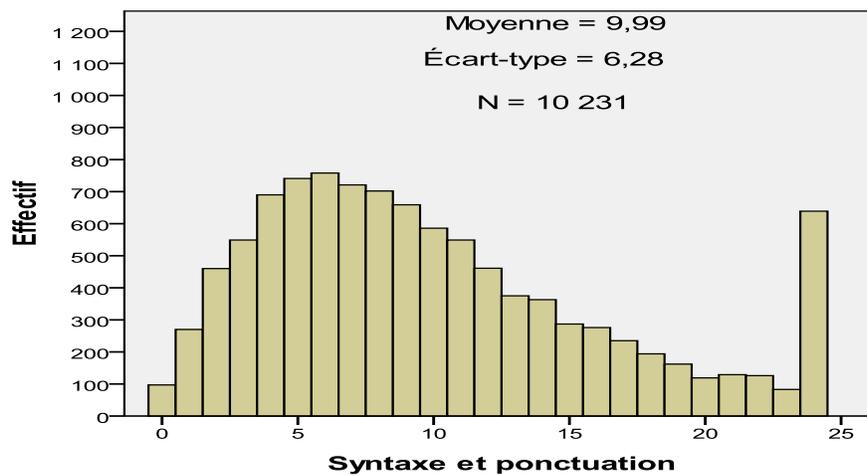


Figure 6 - Distribution des résultats des élèves en syntaxe et ponctuation en 2008

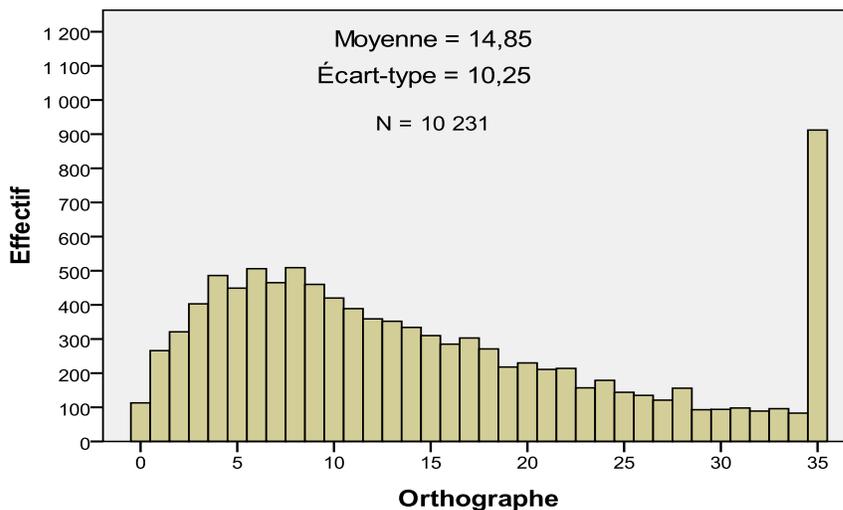


Figure 7 - Distribution des résultats des élèves en orthographe en 2008

Le tableau 1 illustre les résultats moyens (M) avec leurs écarts-types (Ét) et les résultats médians (Md) des élèves pour chacune des cohortes. En 2007, on observe une augmentation de l'effectif scolaire de près de 8,0 % par rapport à 2006 et de près de 7,8 % par rapport à 2008. À notre connaissance, aucun événement particulier ne justifie cette augmentation de l'effectif scolaire en 2007.

**Tableau 1 - Performance des élèves des trois populations**

Année	N	Syntaxe et ponctuation			Orthographe		
		Md	M	Ét	Md	M	Ét
2006	10208	11	11,70	6,60	14	16,40	10,50
2007	11044	9	10,20	6,50	12	15,05	10,40
2008	10231	9	10,00	6,30	12	14,90	10,30

Les résultats indiquent toutefois une légère amélioration de la performance des élèves en 2007, laquelle se traduit par une diminution du nombre moyen d'erreurs pour les deux variables et se maintient en 2008. En effet, la médiane de la distribution de la performance des élèves en syntaxe et ponctuation en 2007 et 2008 est de 9 erreurs alors qu'elle était de 11 erreurs en 2006. En orthographe, elle est de 12 erreurs en 2007 et 2008 alors qu'elle était de 14 erreurs en 2006. Quoique moins robuste que la médiane, la tendance du résultat moyen des cohortes d'élèves évolue dans le même sens que celle de la médiane.

#### **4.2 LES RELATIONS ENTRE LES VARIABLES CARACTÉRISTIQUES DES ÉLÈVES ET LEUR PERFORMANCE**

Cette section présente l'analyse de la relation entre les variables caractéristiques des élèves, l'âge, le genre, la langue maternelle et leur performance aux deux critères de correction. On a analysé également le

croisement des variables *âge* et *genre* ainsi que *âge* et *langue maternelle* avec la performance des élèves.

#### 4.2.1 La relation entre l'âge et la performance des élèves

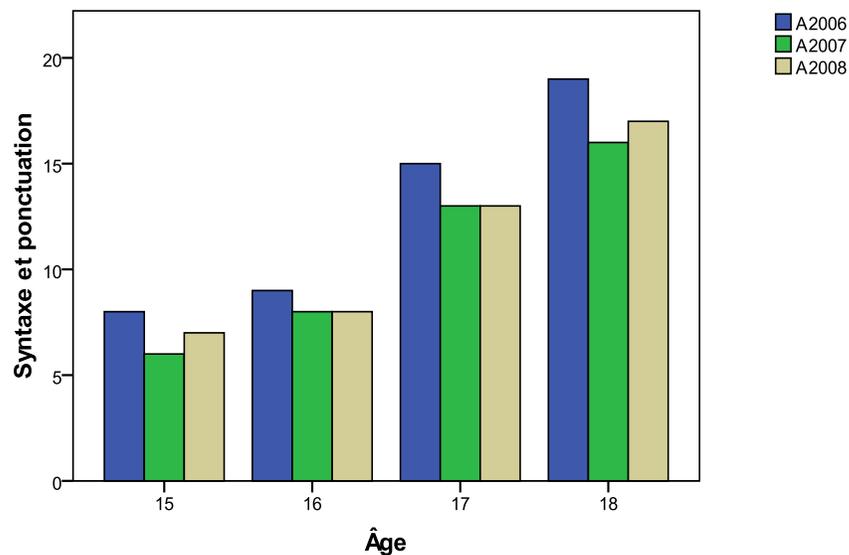
Parmi les variables individuelles caractéristiques des élèves qui ont été signalées dans la recension des écrits, la variable *âge* est l'une des plus importantes (Blais, 2003; Felouzis, 2004; Petrucci, 2010). Le tableau 2 présente les statistiques relatives à la distribution de la performance des élèves des trois cohortes en fonction de l'âge. On y présente les résultats médians (Md) avec leurs écarts interquartiles (Éi).

Tableau 2 - Performance des élèves de l'ensemble de la population selon l'âge

Âge	Syntaxe et ponctuation						Orthographe					
	2006		2007		2008		2006		2007		2008	
	Md	Éi	Md	Éi	Md	Éi	Md	Éi	Md	Éi	Md	Éi
15	8	7	6	6	7	7	11	12	8	10	8	10
16	9	8	8	7	8	7	13	14	11	13	11	13
17	15	12	13	10	13	10	21	20	19	18	19	18
18	19	11	16	11	17	13	25	20	20	20	23,5	22

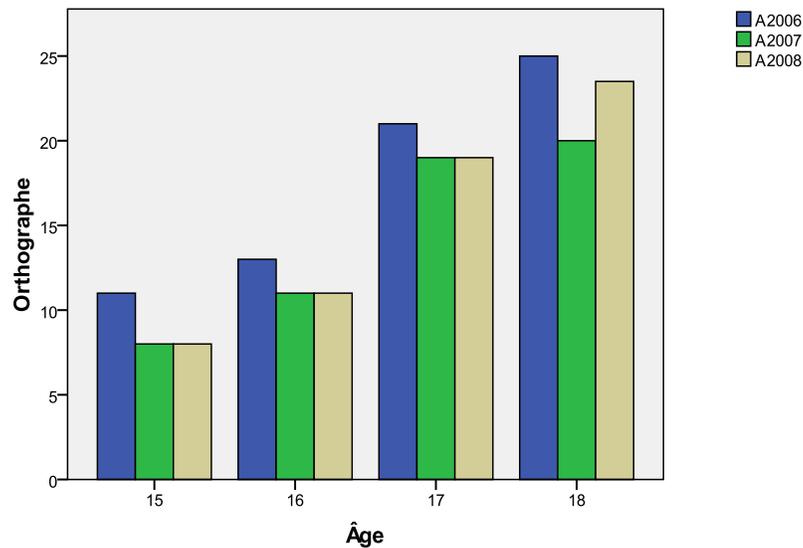
La tendance générale est que le nombre d'erreurs médian des élèves pour les deux variables croît avec l'âge. La différence de résultat la plus importante se trouve entre les élèves de 16 ans, ceux qui arrivent à l'âge normal en 5<sup>e</sup> secondaire, et ceux de 17 ans, qui accusent une année de retard. En syntaxe et ponctuation, la distance médiane entre ces deux groupes d'âge est de 6 erreurs en 2006 et de 5 erreurs en 2007 et en 2008. En orthographe, cette distance médiane est de 8 erreurs entre ces deux groupes d'âge pour chacune des trois cohortes. Notons que la différence de résultat est plus marquée entre les élèves de 16 ans et de 17 ans qu'entre ceux de 15 ans et de 16 ans.

Les figures 8 et 9 illustrent la différence dans la distribution de la performance des élèves, respectivement en syntaxe et ponctuation et en orthographe, en fonction de l'âge pour les trois cohortes. Comme on l'avait remarqué pour l'ensemble des résultats de chaque cohorte au tableau 1, on observe une amélioration du résultat des élèves en 2007 et 2008 par rapport à 2006, qui se traduit par un plus petit nombre d'erreurs pour chaque groupe d'âge considéré dans l'étude.



**Figure 8 - Performance des élèves en syntaxe et ponctuation selon l'âge**

Notons qu'en 2007, les élèves de 15 ans et de 18 ans obtiennent de meilleurs résultats en syntaxe et ponctuation que ceux de leur groupe d'âge en 2006 et en 2008.



**Figure 9 - Performance des élèves en orthographe selon l'âge**

Blais (2003) a montré que les élèves québécois qui arrivent à l'âge normal en 5<sup>e</sup> secondaire, c'est-à-dire à 16 ans, réussissent mieux en français que ceux qui sont plus âgés. Les résultats obtenus par l'analyse de nos données pour les trois années confirmeraient ceux qu'à exposés Blais (2003). Nous rappelons que McAndrew, Ledent et Ait-Said (2005) avaient indiqué qu'arriver à l'âge normal au secondaire s'avère un facteur de prédiction important de la réussite scolaire pour les élèves des communautés noires de Montréal. Il serait opportun d'examiner plus attentivement la relation entre l'âge et les résultats obtenus par les élèves de la région de Montréal selon les deux critères de correction de l'épreuve unique de français écrit du MELS, durant les années 2006, 2007 et 2008, par la modélisation multiniveaux multivariée.

#### **4.2.2 La relation entre le genre et la performance des élèves**

Le tableau 3 présente les résultats médians (Md) avec leurs écarts interquartiles (Éi) de la performance des élèves en syntaxe et ponctuation et

en orthographe, selon le genre. En syntaxe et ponctuation, on n'observe pas de différence importante dans la médiane de la distribution des résultats de deux années sur les trois. La distance médiane dans la distribution des résultats des élèves est tout au plus d'une erreur en faveur des filles en 2006 et en 2008.

**Tableau 3 - Performance des élèves selon le genre**

	Syntaxe et ponctuation						Orthographe					
	2006		2007		2008		2006		2007		2008	
Genre	Md	Éi	Md	Éi	Md	Éi	Md	Éi	Md	Éi	Md	Éi
Fille	10	10	9	10	8	8	12	14	11	13	11	13
Garçon	11	9	9	9	9	9	17	18	15	17	15	15

En orthographe, les filles obtiennent toutefois de meilleurs résultats que les garçons. La différence des résultats médians entre les filles et les garçons est de 5 erreurs en 2006 et de 4 erreurs en 2007 et 2008 en faveur des filles. On sait que selon plusieurs études réalisées au Québec et ailleurs, les filles réussissent mieux que les garçons dans les épreuves de français écrit (Blais, 2003; PISA, 2006, p. 11). D'après nos données, il y aurait lieu d'examiner plus en profondeur la relation entre le genre (garçon, fille) et les résultats des élèves de la région métropolitaine de Montréal par rapport aux deux critères spécifiques retenus pour notre recherche.

Nous allons maintenant croiser la variable *genre* de l'élève avec celle de *l'âge*.

#### **4.2.3 La relation entre le genre, l'âge et la performance des élèves**

Le tableau 4 indique les médianes (Md) de la distribution de la performance des élèves en fonction de l'âge et du genre pour les trois années, celles de 2006, de 2007 et de 2008. En syntaxe et ponctuation, les élèves des deux sexes réussissent de manière équivalente en 2007, mais en 2006 et 2008, les garçons de 18 ans réussissent mieux que les filles du même âge. En

orthographe, les filles réussissent mieux que les garçons, quel que soit le groupe d'âge considéré à chacune des trois cohortes.

**Tableau 4 - Performance des élèves selon l'âge et le genre**

Age	Syntaxe et ponctuation						Orthographe					
	2006		2007		2008		2006		2007		2008	
	Fille Md	Garçon Md	Fille Md	Garçon Md	Fille Md	Garçon Md	Fille Md	Garçon Md	Fille Md	Garçon Md	Fille Md	Garçon Md
15	7	8	6	6	7	6	10	12	7	9	8	9
16	9	10	8	8	8	8	11	15	10	13	10	13
17	15	15	13	13	13	13	19	23	18	22	17	22
18	20	18	16	16	18	17	24	27	23	25	20	24

La maîtrise de l'orthographe semble être problématique pour les garçons, surtout ceux qui sont âgés de 18 ans. Peu nombreuse, il est vrai dans l'effectif, cette catégorie d'âge représente autour de 5,50 % en 2006, 5,40 % en 2007 et 4,30 % en 2008. Ces pourcentages représentent quand même un nombre important d'élèves dans chaque cohorte. Il convient d'approfondir la réflexion pour déterminer les raisons qui font que certains élèves éprouvent de la difficulté en orthographe. La variable *genre* de l'élève sera retenue dans l'analyse par la modélisation multiniveaux.

#### **4.2.4 La relation entre la langue maternelle et la performance de l'élève**

Le tableau 5 présente les médianes (Md) avec leurs écarts-interquartiles (Éi) de la distribution de la performance des élèves en fonction de la langue maternelle. Les élèves dont la langue maternelle est le français réussissent mieux que ceux d'autres langues maternelles en syntaxe et ponctuation et en orthographe. Cependant, l'ampleur de la différence varie selon chaque cohorte à l'étude. En syntaxe et ponctuation, la distance médiane est de 3,5 erreurs en 2007 et de 4 erreurs en 2006 et 2008, en faveur des élèves de

langue maternelle française. En orthographe, la distance médiane dans la distribution des résultats des élèves entre les deux catégories linguistiques est de 2 erreurs en 2006, et d'une erreur en 2007 et 2008, en faveur des élèves de langue maternelle française.

**Tableau 5 - Performance des élèves selon la langue maternelle**

Langue maternelle	Syntaxe et ponctuation						Orthographe					
	2006		2007		2008		2006		2007		2008	
	Md	Éi	Md	Éi	Md	Éi	Md	Éi	Md	Éi	Md	Éi
<b>Français</b>	9	8	7,5	8	7	8	13	16	12	16	12	15
<b>Autres</b>	13	11	11	10	11	10	15	16	13	16	13	15

Au Québec, la majorité des immigrants de langues maternelles différentes du français habitent la région métropolitaine de Montréal. Plusieurs recherches indiquent que les élèves de langues maternelles différentes de la langue d'enseignement éprouvent des difficultés dans la réussite de leurs études (Fertig et Wright, 2004, p. 6; McAndrew, Ledent et Ait-Said, 2005, p. 33).

Nous allons croiser la variable la *langue maternelle* de l'élève avec *l'âge* pour approfondir notre compréhension du phénomène.

#### **4.2.5 La relation entre l'âge, la langue maternelle et la performance des élèves**

Le tableau 6 présente la performance des élèves en fonction de l'âge et de la langue maternelle. En 2006, les élèves dont la langue maternelle est le français réussissent mieux que ceux d'autres langues maternelles en syntaxe et ponctuation et en orthographe, quel que soit le groupe d'âge considéré. En 2007 et 2008, les élèves de langue maternelle française réussissent mieux que ceux de langues maternelles différentes du français en syntaxe et ponctuation. En orthographe toutefois, les élèves de langues maternelles différentes du français qui sont âgés de 16, 17 et 18 ans réussissent de manière équivalente à ceux de langue maternelle française ou mieux que ces derniers.

**Tableau 6 - Performance des élèves selon l'âge et la langue maternelle**

Âge	Syntaxe et ponctuation						Orthographe					
	2006		2007		2008		2006		2007		2008	
	Fran.	Autre	Fran.	Autre	Fran.	Autre	Fran.	Autre	Fran.	Autre	Fran.	Autre
	Md	Md	Md	Md	Md	Md	Md	Md	Md	Md	Md	Md
15	7	9	5	8	6	9	10	12	7	10	8	10
16	8	12	7	10	8	10	12	13	11	11	11	11
17	14	15	12	14	12	14	19	26	24	17	22	17
18	17	19,5	15	17	16	18	25	29	29	22	26	21

Les élèves de langues maternelles différentes du français éprouvent par contre plus de difficulté dans la construction de phrases où les exigences de maîtrise de la syntaxe et de la ponctuation sont plus complexes. La réussite différentielle des groupes linguistiques en fonction de l'âge nous incite à retenir la langue maternelle pour la suite de l'analyse par la modélisation multiniveaux multivariée.

### **4.3 LES RELATIONS ENTRE LES VARIABLES CARACTÉRISTIQUES DE L'ÉCOLE ET LA PERFORMANCE DES ÉLÈVES**

Pour chaque école de la région métropolitaine de Montréal, nous disposons de données sur les variables suivantes : le réseau, le nombre d'élèves présents à l'épreuve et la taille. Nous avons déterminé d'autres variables : la proportion des élèves de l'école de langue maternelle française et la mixité de l'école. Les sections qui suivent présenteront l'analyse de ces variables contextuelles selon les deux critères de correction à l'étude.

#### **4.3.1 La performance des élèves selon le réseau**

Le tableau 7 indique les médianes de la distribution des résultats des élèves en fonction du réseau de l'école pour les trois cohortes à l'étude. En syntaxe et ponctuation, la distance médiane entre les deux secteurs est de 4 erreurs en 2006 et de 3 erreurs en 2007 et 2008, en faveur du secteur privé. En orthographe, elle est de 7 erreurs en 2006 et 2007, et de 6,50 erreurs en

2008, également en faveur du secteur privé. Les élèves des écoles privées obtiennent de meilleurs résultats que ceux du réseau public durant les trois années.

**Tableau 7 - Performance des élèves selon le réseau**

Réseau	Syntaxe et ponctuation						Orthographe					
	2006		2007		2008		2006		2007		2008	
	Md	Éi	Md	Éi	Md	Éi	Md	Éi	Md	Éi	Md	Éi
<b>Privé</b>	8	7	7	7	7	6	10	12	8	10	8,50	10
<b>Public</b>	12	11	10	10	10	10	17	18	15	17	15	17

Considérant l'importance du secteur privé dans la région métropolitaine de Montréal, les résultats de nos analyses sur cette variable ainsi que ceux obtenus par d'autres études sur les différences des résultats des élèves selon le secteur, privé ou public (Blais, 2003; Rumberger et Palardy, 2005), nous allons poursuivre l'analyse en croisant la variable réseau avec l'âge de l'élève.

#### 4.3.2 La performance des élèves selon le réseau et l'âge

Le tableau 8 présente les résultats des élèves pour les deux critères de correction en fonction de l'âge et du réseau. En 2006 et 2008, les élèves du réseau privé obtiennent de meilleurs résultats que ceux du réseau public quant au plus petit nombre d'erreurs pour chacun des groupes d'âge considérés par rapport aux deux variables, la syntaxe et la ponctuation ainsi que l'orthographe.

**Tableau 8 - Performance des élèves selon l'âge et le réseau**

Âge	Syntaxe et ponctuation						Orthographe					
	2006		2007		2008		2006		2007		2008	
	Privé Md	Public Md	Privé Md	Public Md	Privé Md	Public Md	Privé Md	Public Md	Privé Md	Public Md	Privé Md	Public Md
<b>15</b>	7	8,50	6	7	6	8	9	12	7	11	8	11
<b>16</b>	8	11	6	9	6	9	10	15	8	14	8	14
<b>17</b>	13	15	11,50	13	10	14	18	22	15	20	14	20
<b>18</b>	17	19	16,50	16	15	18	23	26	27	24	16	24

En 2007, les élèves du secteur public qui sont âgés de 18 ans obtiennent cependant de meilleurs résultats que ceux du même âge qui fréquentent le secteur privé, tant en syntaxe et ponctuation qu'en orthographe. Ces résultats sont peut-être dus au fait qu'une plus large proportion d'élèves de 18 ans étaient inscrits dans les écoles publiques durant cette année; le nombre élevé d'élèves de ce groupe d'âge stabilise la distribution de leur performance, par opposition à un plus petit nombre d'élèves inscrits au secteur privé dont les résultats seraient plus variables. Si la différence de résultats entre les élèves des deux secteurs de ce groupe d'âge est de 0,5 erreurs en syntaxe et ponctuation, elle est cependant de 3 erreurs en orthographe en faveur des élèves du secteur public.

#### 4.3.3 La performance des élèves selon la mixité

Le tableau 9 indique les médianes (Md) de la distribution de la performance des élèves avec leurs écarts interquartiles (Éi) en fonction de la mixité des écoles. Il n'existe pas d'écoles réservées aux garçons dans la région métropolitaine de Montréal. Les filles inscrites aux écoles qui leur sont réservées obtiennent de meilleurs résultats que les élèves, filles et garçons, des écoles mixtes. En syntaxe et ponctuation, la distance médiane est d'une erreur pour chacune des années. En orthographe, la distance médiane est de 3 erreurs en 2006, de 2 erreurs en 2007 et de 3 erreurs en 2008, en faveur des écoles réservées aux filles.

Tableau 9 - Performance des élèves selon la mixité de l'école

Mixité	Syntaxe et ponctuation						Orthographe					
	2006		2007		2008		2006		2007		2008	
	Md	Éi	Md	Éi	Md	Éi	Md	Éi	Md	Éi	Md	Éi
Fille	10	10	8	8	8	7	12	13	11	12	10	11
Mixte	11	10	9	9	9	9	15	17	13	16	13	15

Le pourcentage très faible de filles qui fréquentent des écoles réservées aux filles, moins de 10 % de la population, nous invite à la prudence quant à

l'interprétation de ce résultat. Comme nous l'avons déjà noté, plusieurs études indiquent l'existence d'un effet différentiel des écoles mixtes sur le rendement scolaire des élèves, en faveur des filles (Nuttal et collab., 1989; Goldstein, Huiqi, Rath et Hill, 2000; Meyer, 2002; Stevens, 2005; Konstantopoulos, 2006). La mixité de l'école, telle qu'elle a été préalablement définie dans notre recherche, sera modélisée dans les analyses multiniveaux multivariées.

#### 4.3.4 La performance des élèves selon la taille

Le tableau 10 présente les médianes de la distribution des résultats des élèves en fonction de la taille de l'école. En vue de faciliter la présentation de ces résultats, la variable taille de l'école est divisée en trois catégories. Les écoles de petite taille contiennent moins de 600 élèves. Les écoles de taille moyenne comptent entre 600 et 1 200 élèves et les écoles de grande taille sont fréquentées par plus de 1 200 élèves. Les résultats des élèves des écoles de grande taille semblent montrer un léger avantage comparativement à ceux des élèves des écoles de petite taille et de taille moyenne, tant en syntaxe et ponctuation qu'en orthographe.

**Tableau 10 - Performance des élèves selon la taille de l'école**

	Syntaxe et ponctuation						Orthographe					
	2006		2007		2008		2006		2007		2008	
Taille	Md	Éi	Md	Éi	Md	Éi	Md	Éi	Md	Éi	Md	Éi
Petite	11	9	10	9	9	8	14	16	14	16	12	14
Moyenne	11	10	9	9	9	9	15	17	13	16	13	16
Grande	10	10	8	8	8	8	14	16	12	15	12	15

Contrairement à nos résultats, une étude récente aux États-Unis confirme l'existence d'une corrélation positive entre la taille de l'école et les résultats des élèves (Konstantopoulos, 2006) : les écoles de plus petite taille seraient plus performantes.

Il convient de rappeler que pour être considéré dans la présente recherche, l'établissement scolaire doit présenter au moins 20 élèves à l'épreuve unique de français écrit du MELS. Cette exigence est l'une des conditions d'application de la modélisation multiniveaux. Nous retenons donc la variable taille de l'école pour la suite de l'analyse par la modélisation multiniveaux multivariée.

#### 4.3.5 La performance des élèves selon la proportion des élèves de l'école de langue maternelle française

Le tableau 11 indique les médianes de la distribution des résultats en fonction de la proportion des élèves de l'école de langue maternelle française. Pour faciliter la présentation des données, nous avons classé les écoles en trois catégories : celles qui ont moins de 25 % de leur effectif d'élèves de langue maternelle française, celles qui ont entre 25 % et 75 % de cet effectif et celles qui ont plus de 75 % de cet effectif. Nos données indiquent que les élèves des écoles de cette dernière catégorie présentent de meilleurs résultats que ceux des deux autres catégories.

**Tableau 11 - Performance des élèves selon la proportion des élèves de l'école de langue maternelle française**

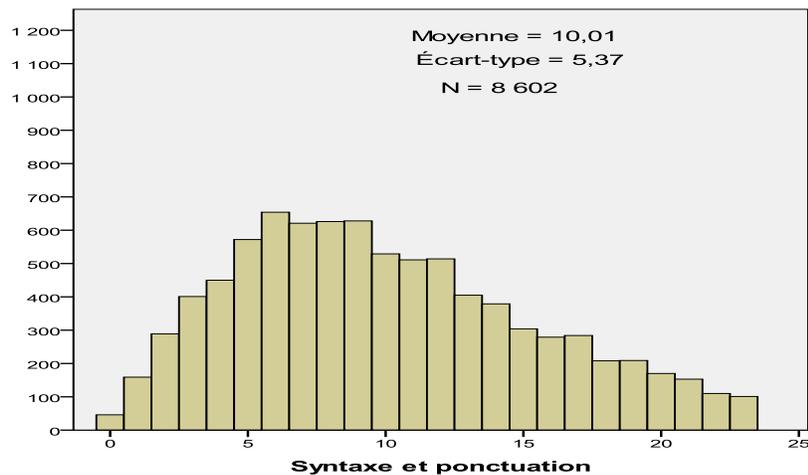
Année	Syntaxe et ponctuation			Orthographe		
	25 %	25 %-75 %	+ 75 %	25 %	25 %-75 %	+ 75 %
2006	13	11	9	15	16	11
2007	12	10	7	13	14	10
2008	11	9	7	14	13	10

La proportion des élèves de l'école de langue maternelle française présente un lien avec la variable individuelle la *langue maternelle* caractéristique de l'élève. Des études récentes ont conclu à une forte corrélation entre la proportion élevée d'élèves immigrants dans une école et leur rendement scolaire (Coleman et collab. 1966; Rumberger et Palardy, 2005;

Konstantopoulos, 2006). Considérant que 77 % des jeunes Noirs du secteur francophone issus de l'immigration au Québec se retrouvent dans la région de Montréal (McAndrew, Ledent et Ait-Said, 2005), nous avons retenu cette variable pour la suite de l'analyse par la modélisation multiniveaux multivariée.

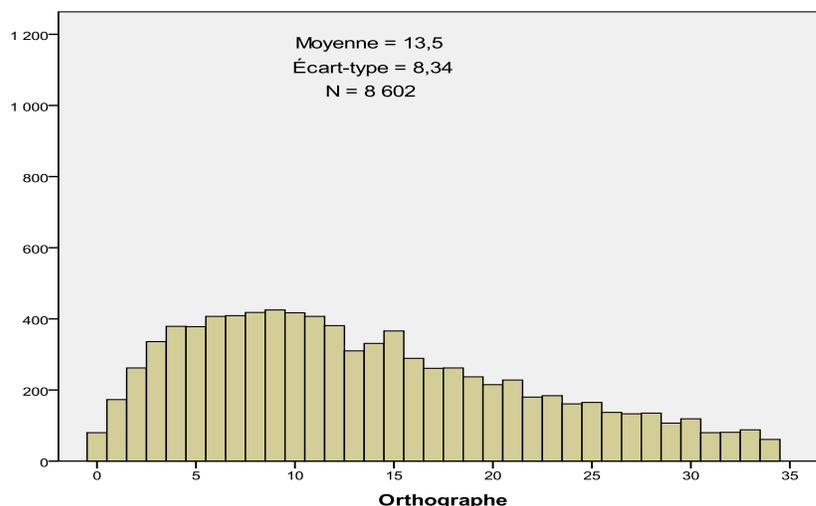
#### 4.4 LA DESCRIPTION DES DONNÉES DES SOUS-POPULATIONS D'ÉLÈVES SANS LE NOMBRE ÉLEVÉ D'ERREURS

Les figures 10 et 11 illustrent la distribution de fréquence des résultats des élèves sans les valeurs de 24 erreurs en syntaxe et ponctuation et de 35 erreurs en orthographe en 2006. Durant cette année, cette sous-population comprenait 8602 élèves. En syntaxe et ponctuation, les résultats des élèves présentent une distribution normale. Le nombre d'erreurs moyen des élèves est de 10,01 avec un écart-type de 5,37.



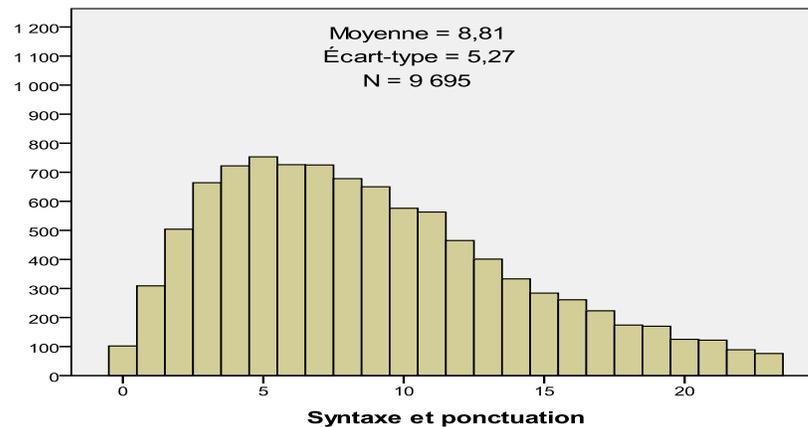
**Figure 10 - Résultat des élèves sans le nombre élevé d'erreurs en syntaxe et ponctuation en 2006**

En orthographe, les résultats présentent une distribution normale avec une légère asymétrie positive. Les élèves de cette sous-population font en moyenne 13,50 erreurs avec un écart-type de 8,34.

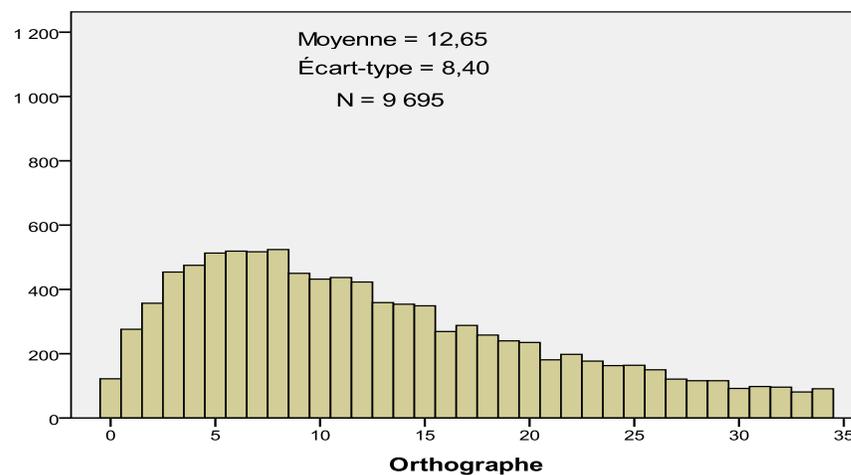


**Figure 11 - Résultat des élèves sans le nombre élevé d'erreurs en orthographe en 2006**

Pour l'année 2007, les figures 12 et 13 illustrent la distribution de fréquence des résultats des élèves de cette sous-population. Celle-ci contient 9695 élèves. Ces résultats montrent une distribution normale avec une légère asymétrie positive. En syntaxe et ponctuation, les élèves font en moyenne 8,81 erreurs avec un écart-type de 5,27. En orthographe, ils font en moyenne 12,65 erreurs avec un écart-type de 8,40.



**Figure 12 - Résultat des élèves sans le nombre élevé d'erreurs en syntaxe et ponctuation en 2007**



**Figure 13 - Résultat des élèves sans le nombre élevé d'erreurs en orthographe en 2007**

Le tableau 12 présente les médianes (Md) et les moyennes (M) avec leurs écarts-types de la distribution des résultats des sous-populations d'élèves sans le nombre élevé d'erreurs pour chacune des trois cohortes, celles des années 2006, 2007 et 2008. La meilleure performance pour une sous-population d'élèves a été réalisée en 2008. Durant cette année, en syntaxe et ponctuation, les élèves font en moyenne 8,7 erreurs avec un écart-type de

5,2. En orthographe, ils font en moyenne 12,6 erreurs avec un écart-type de 8,3.

**Tableau 12 - Performance des élèves sans le nombre élevé d'erreurs**

Année	N	Syntaxe et ponctuation			Orthographe		
		Md	M	Ét	Md	M	Ét
<b>2006</b>	8602	9	10,0	5,4	12	13,5	8,3
<b>2007</b>	9695	8,0	8,8	5,3	11,0	12,7	8,4
<b>2008</b>	9100	8,0	8,7	5,2	11,0	12,6	8,3

La tendance dans la médiane de la distribution des résultats des élèves sans le nombre élevé d'erreurs pour les trois années de ces sous-populations évolue dans le même sens que pour les élèves de l'ensemble de la population. On note une légère amélioration à partir de 2007, qui se maintient en 2008 pour les deux variables à l'étude.

#### **4.5 LES RELATIONS ENTRE LES VARIABLES CARACTÉRISTIQUES DES ÉLÈVES ET LA PERFORMANCE DES ÉLÈVES SANS LE NOMBRE ÉLEVÉ D'ERREURS**

Cette section présente l'analyse descriptive de la relation entre les variables caractéristiques des élèves : l'âge, le genre et la langue maternelle, et les résultats des élèves sans le nombre élevé d'erreurs en syntaxe et ponctuation et en orthographe.

##### **4.5.1 La relation entre l'âge et la performance des élèves sans le nombre élevé d'erreurs**

Le tableau 13 indique la performance des élèves selon l'âge, pour chacune des trois sous-populations. Comme pour l'ensemble de la population, les élèves les plus jeunes réussissent mieux que les plus âgés. La différence médiane la plus importante apparaît entre les élèves de 16 et de 17 ans. Elle

est de 3 erreurs en 2006, de 4 erreurs en 2007 et 2008 en syntaxe et ponctuation, et en orthographe, elle est de 5 erreurs en 2006 et de 6 erreurs en 2007 et 2008, en faveur des élèves de 16 ans. Encore ici, la différence de résultats est plus marquée entre les élèves de 16 et de 17 ans qu'entre ceux de 15 et de 16 ans. Le facteur d'âge est relié plus fortement à l'orthographe qu'à la syntaxe et la ponctuation.

**Tableau 13 - Performance des élèves sans le nombre élevé d'erreurs selon l'âge**

Âge	Syntaxe et ponctuation						Orthographe					
	2006		2007		2008		2006		2007		2008	
	Md	Éi	Md	Éi	Md	Éi	Md	Éi	Md	Éi	Md	Éi
15	7	6	6	7	7	7	10	10	7	9	8	9
16	9	7	7	7	7	7	11	12	10	12	10	11
17	12	7	11	7	11	8	16	13	16	13	16	13
18	14	9	13	8	14	8	18	13	19	14	17	14

#### 4.5.2 La relation entre le genre et la performance des élèves sans le nombre élevé d'erreurs

Le tableau 14 présente les résultats médians (Md) avec leurs écarts interquartiles (Éi) de la distribution de la performance des élèves sans les valeurs de 24 erreurs en syntaxe et ponctuation et de 35 erreurs en orthographe, en fonction du genre. Selon les données analysées, en syntaxe et ponctuation on n'observe pas de différence dans la médiane de la distribution des résultats des deux groupes. Les filles font en moyenne moins d'erreurs d'orthographe que les garçons. On observe une distance médiane de 3 erreurs en 2006, de 2 erreurs en 2007 et de 3 erreurs en 2008, en faveur des filles.

**Tableau 14 - Performance des élèves sans le nombre élevé d'erreurs selon le genre**

Genre	Syntaxe et ponctuation						Orthographe					
	2006		2007		2008		2006		2007		2008	
	Md	Éi	Md	Éi	Md	Éi	Md	Éi	Md	Éi	Md	Éi
Fille	9	8	8	7	8	7	11	11	10	11	10	12
Garçon	9	8	8	7	8	7	14	13	12	13	13	13

Les garçons semblent éprouver plus de difficulté que les filles dans la maîtrise de l'orthographe.

#### 4.5.3 La relation entre la langue maternelle et la performance des élèves sans le nombre élevé d'erreurs

Le tableau 15 présente les médianes de la distribution de la performance des élèves en fonction de la langue maternelle. Les élèves de langue maternelle française réussissent mieux que ceux de langues maternelles différentes du français en syntaxe et ponctuation de même qu'en orthographe, pour chacune des trois sous-populations. L'ampleur de la différence varie cependant selon chacune des sous-populations.

**Tableau 15 - Performance des élèves sans le nombre élevé d'erreurs selon la langue maternelle**

Langue maternelle	Syntaxe et ponctuation						Orthographe					
	2006		2007		2008		2006		2007		2008	
	Md	Éi	Md	Éi	Md	Éi	Md	Éi	Md	Éi	Md	Éi
<b>Français</b>	8	7	7	7	7	6	12	13	10	12	11	12
<b>Autres</b>	11	8	10	8	10,5	8	13	12	12	13	11	11

En syntaxe et ponctuation, la distance médiane est de 3 erreurs en 2006 et 2007 et de 3,5 erreurs en 2008, en faveur des élèves de langue maternelle française. En orthographe, la distance médiane entre les deux groupes est de 2 erreurs en 2007 et d'une erreur en 2006, en faveur des élèves de langue maternelle française. Comme nous l'avons déjà indiqué pour l'ensemble des élèves de chaque cohorte, cette variable sera retenue pour la suite de l'analyse par la modélisation multiniveaux multivariée.

#### 4.6 LES RELATIONS ENTRE LES VARIABLES CARACTÉRISTIQUES DE L'ÉCOLE ET LA PERFORMANCE DES ÉLÈVES SANS LE NOMBRE ÉLEVÉ D'ERREURS

Dans les sections qui suivent, nous analyserons les variables contextuelles caractéristiques des écoles en fonction du résultat des élèves sans le nombre élevé d'erreurs, selon les deux critères de correction à l'étude.

##### 4.6.1 La relation entre le réseau d'appartenance de l'école et la performance des élèves sans le nombre élevé d'erreurs

Le tableau 16 indique la performance des élèves sans le nombre élevé d'erreurs en fonction du réseau d'appartenance de l'école. Les élèves du secteur privé réussissent mieux que ceux du secteur public, quelle que soit l'année considérée pour chacun des deux critères de correction à l'étude. La distance médiane entre les résultats des élèves des deux secteurs est cependant plus importante en orthographe qu'en syntaxe et ponctuation : elle est de 2 erreurs en 2006 et de 3 erreurs en 2007 et 2008 en syntaxe et ponctuation; en orthographe, elle est de 5 erreurs pour chacune des sous-populations. Encore ici, ce sont les résultats des élèves en orthographe qui sont problématiques.

Tableau 16 - Performance des élèves sans le nombre élevé d'erreurs selon le réseau

Réseau	Syntaxe et ponctuation						Orthographe					
	2006		2007		2008		2006		2007		2008	
	Md	Éi	Md	Éi	Md	Éi	Md	Éi	Md	Éi	Md	Éi
Privé	8	7	6	6	6	6	9	10	8	10	8	10
Public	10	8	9	7	9	7	14	14	13	14	13	12

##### 4.6.2 La relation entre la mixité de l'école et la performance des élèves sans le nombre élevé d'erreurs

Le tableau 17 illustre les résultats des élèves sans le nombre élevé d'erreurs, selon la mixité de l'établissement scolaire. En 2006, 2007 et 2008, pour la syntaxe et la ponctuation, on n'observe pas de différence dans la performance des élèves selon la mixité de l'école. En orthographe, les filles

inscrites aux écoles qui leur sont réservées obtiennent cependant de meilleurs résultats que les élèves, filles et garçons, des écoles mixtes. Dans cette sous-population, c'est la variable orthographe qui semble être problématique pour les élèves qui fréquentent les écoles mixtes, soit 91 % de l'effectif total d'élèves.

**Tableau 17 - Performance des élèves sans le nombre élevé d'erreurs selon la mixité**

	Syntaxe et ponctuation						Orthographe					
	2006		2007		2008		2006		2007		2008	
	Md	Éi	Md	Éi	Md	Éi	Md	Éi	Md	Éi	Md	Éi
<b>Mixité</b>												
<b>Fille</b>	9	8	8	7	8	6	10	10	10	11	9	11
<b>Mixte</b>	9	8	8	7	8	7	12	12	11	12	11	12

#### 4.6.3 La relation entre la taille de l'école et la performance des élèves sans le nombre élevé d'erreurs

Le tableau 18 présente les résultats des élèves sans les valeurs de 24 erreurs en syntaxe et ponctuation et de 35 erreurs en orthographe, en tenant compte de la taille de l'école. En syntaxe et ponctuation, les résultats indiquent qu'en 2007 et 2008, la performance des élèves des écoles de taille moyenne et de grande taille est supérieure à celle des élèves des écoles de petite taille. Les résultats des élèves en syntaxe et ponctuation semblent être plus problématiques qu'en orthographe si l'on tient compte de la taille de l'école. Pour cette raison, ainsi que pour d'autres qui ont été évoquées dans l'analyse des résultats des élèves de l'ensemble des trois cohortes, la variable taille de l'école sera traitée par la modélisation multiniveaux multivariée.

**Tableau 18 - Performance des élèves sans le nombre élevé d'erreurs selon la taille**

	Syntaxe et ponctuation						Orthographe					
	2006		2007		2008		2006		2007		2008	
	Md	Éi	Md	Éi	Md	Éi	Md	Éi	Md	Éi	Md	Éi
<b>Taille</b>												
<b>Petite</b>	9	7	9	7	9	7	12	12	12	12	11	12
<b>Moyenne</b>	10	8	8	7	8	7	12	12	11	13	11	13
<b>Grande</b>	9	7	8	8	8	7	12	12	10	11	11	12

#### 4.6.4 La relation entre la proportion des élèves de l'école de langue maternelle française et la performance des élèves sans le nombre élevé d'erreurs

Le tableau 19 indique la performance des élèves sans le nombre élevé d'erreurs en fonction de la proportion des élèves de l'école de langue maternelle française. Les élèves des écoles dont l'effectif de langue maternelle française est de plus de 75 % obtiennent de meilleurs résultats que ceux des écoles fréquentées par un pourcentage d'élèves de langue maternelle française moins élevé. Cette variable est d'autant plus importante qu'elle est liée à la variable individuelle *langue maternelle* caractéristique de l'élève. Il y aura lieu de l'examiner davantage dans la suite de l'analyse par la modélisation multiniveaux multivariée.

Tableau 19 - Performance des élèves sans le nombre élevé d'erreurs selon le pourcentage des élèves de l'école de langue maternelle française

% langue maternelle	Syntaxe et ponctuation						Orthographe					
	2006		2007		2008		2006		2007		2008	
	Md	Éi	Md	Éi	Md	Éi	Md	Éi	Md	Éi	Md	Éi
Moins de 25	11	9	11	8	10	7	13	13	12	11	12	12
25 et 75	10	8	9	8	8	7	13	12	12	13	11	13
Plus de 75	8	7	6	6	7	7	10	11	9	10	9	12

#### 4.7 L'ANALYSE DE LA PERFORMANCE DES ÉLÈVES EN DIFFICULTÉ SELON LES VARIABLES CARACTÉRISTIQUES DE L'ÉLÈVE

Cette section présente l'analyse des variables individuelles caractéristiques des élèves en difficulté, c'est-à-dire ceux qui font 24 erreurs de syntaxe et ponctuation et 35 erreurs d'orthographe, selon les deux critères de correction, la syntaxe et la ponctuation et l'orthographe. Ces analyses seront effectuées sous l'angle de la représentativité du nombre d'élèves dans chaque catégorie pour chaque sous-population à l'étude.

#### 4.7.1 La relation entre l'âge et la performance des élèves en difficulté

Le tableau 20 présente le nombre d'élèves en difficulté dans chaque catégorie d'âge. En 2006, on compte 10 élèves de 15 ans qui font 24 erreurs de syntaxe et ponctuation et 19 élèves de ce groupe d'âge qui font 35 erreurs d'orthographe.

Tableau 20 – Effectifs des élèves en difficulté selon l'âge

Âge	Syntaxe et ponctuation			Orthographe		
	2006 Effectif	2007 Effectif	2008 Effectif	2006 Effectif	2007 Effectif	2008 Effectif
15	10	13	9	19	20	8
16	487	359	294	637	537	499
17	333	282	214	363	341	285
18	192	132	122	175	159	120

En 2008, on compte 294 élèves de 16 ans qui font 24 erreurs de syntaxe et ponctuation et 499 élèves de ce groupe d'âge qui font 35 erreurs d'orthographe. Dans cette sous-population, la variable orthographe semble être plus problématique, quel que soit le groupe d'âge considéré.

La proportion des élèves de 16 ans diminue considérablement dans ces trois sous-populations d'élèves en difficulté par rapport à la population d'élèves de chaque cohorte. En syntaxe et ponctuation, elle passe d'environ 75 % dans l'ensemble des trois cohortes à environ 45 % dans ces sous-populations d'élèves en difficulté. Par contre, la proportion des élèves de 17 ans et de 18 ans augmente. De 20 % qu'elle était dans l'ensemble de la population d'élèves, elle passe à plus de 50 % dans ces trois sous-populations d'élèves en difficulté.

#### 4.7.2 La relation entre le genre et la performance des élèves en difficulté

Le tableau 21 présente l'effectif des élèves qui éprouvent des difficultés, en fonction du genre pour les trois sous-populations. En général, les garçons sont plus nombreux que les filles dans ces trois sous-populations, à

l'exception de 2006, où on compte 509 garçons et 513 filles pour la variable *syntaxe et ponctuation*. Par exemple, en 2007, on compte 432 garçons et 354 filles qui font 24 erreurs de syntaxe et ponctuation et plus, et 655 garçons et 402 filles qui font 35 erreurs d'orthographe et plus. En 2008, on compte 261 filles qui font 24 erreurs de syntaxe et ponctuation, et 399 filles qui font 35 erreurs d'orthographe; on dénombre 378 garçons qui font 24 erreurs de syntaxe et ponctuation et 573 garçons qui font 35 erreurs d'orthographe. On relève en effet une augmentation de la proportion de garçons dans cette sous-population par rapport à celle de la population des trois cohortes.

**Tableau 21 - Performance des élèves en difficulté selon le genre**

	Syntaxe et ponctuation			Orthographe		
	2006	2007	2008	2006	2007	2008
Genre	Effectif	Effectif	Effectif	Effectif	Effectif	Effectif
Fille	513	354	261	471	402	339
Garçon	509	432	378	723	655	573

#### **4.7.3 La relation entre la langue maternelle et la performance des élèves en difficulté**

Le tableau 22 présente l'effectif de chaque sous-population d'élèves en difficulté, selon la langue maternelle. En syntaxe et ponctuation, les élèves de langue maternelle française sont moins nombreux que ceux de langues maternelles différentes du français en 2006 et 2008. Ils sont cependant légèrement plus nombreux en 2007. En orthographe, les élèves de langue maternelle française sont plus nombreux que ceux de langues maternelles différentes du français dans chacune des trois sous-populations d'élèves en difficulté. Ce résultat sera approfondi dans l'autre étape d'analyse par la modélisation multiniveaux multivariée.

Tableau 22 - Performance des élèves en difficulté selon la langue maternelle

Langue maternelle	Syntaxe et ponctuation			Orthographe		
	2006 Effectif	2007 Effectif	2008 Effectif	2006 Effectif	2007 Effectif	2008 Effectif
Français	438	396	273	660	651	541
Autres	584	390	366	534	406	371

#### 4.8 L'ANALYSE DE LA PERFORMANCE DES ÉLÈVES EN DIFFICULTÉ SELON LES VARIABLES CARACTÉRISTIQUES DE L'ÉCOLE

Cette section présente les analyses sur les relations entre les variables caractéristiques de l'école et les élèves en difficulté. Ces analyses seront effectuées sous l'angle de la représentativité du nombre d'élèves dans chaque catégorie de variable pour les sous-populations à l'étude.

##### 4.8.1 La relation entre le réseau et la performance des élèves en difficulté

Le tableau 23 présente l'effectif de la sous-population des élèves en difficulté selon le réseau. Quelle que soit l'année considérée, les élèves en difficulté sont toujours plus nombreux dans les écoles publiques que dans les écoles privées. En 2006, on comptait 144 élèves du secteur privé et 878 élèves du secteur public qui faisaient 24 erreurs en syntaxe et ponctuation. Pour cette même année, on dénombrait 138 élèves du secteur privé et 1056 élèves du secteur public qui faisaient 35 erreurs d'orthographe.

Tableau 23 - Performance des élèves en difficulté selon le réseau

Réseau	Syntaxe et ponctuation			Orthographe		
	2006 Effectif	2007 Effectif	2008 Effectif	2006 Effectif	2007 Effectif	2008 Effectif
Privé	144	102	75	138	130	90
Public	878	684	564	1056	927	822

#### 4.8.2 La relation entre la mixité et la performance des élèves en difficulté

Le tableau 24 présente l'effectif de chaque sous-population d'écoles, filles et mixtes, comptant des élèves en difficulté. Le nombre de filles qui fréquentent des écoles réservées aux filles est nettement moins élevé que celui d'élèves, filles et garçons, qui fréquentent des écoles mixtes. Il faut tenir compte du nombre assez bas de filles qui fréquentent des écoles qui leur sont réservées. La variable mixité de l'école sera quand même analysée par la modélisation multiniveaux multivariée.

Tableau 24 - Performance des élèves en difficulté selon la mixité de l'école

Mixité	Syntaxe et ponctuation			Orthographe		
	2006 Effectif	2007 Effectif	2008 Effectif	2006 Effectif	2007 Effectif	2008 Effectif
Fille	74	45	20	75	56	30
Mixte	948	741	619	1119	1001	882

#### 4.8.3 La relation entre la proportion d'élèves de l'école de langue maternelle française et la performance des élèves en difficulté

Il n'y a pas de tendance significative entre la proportion d'élèves de l'école de langue maternelle française et le nombre d'élèves en difficulté. Compte tenu du lien entre cette variable et la variable individuelle caractéristique de l'élève, la *langue maternelle*, celle-ci a été retenue pour la série d'analyses multiniveaux multivariée.

### 4.9 UNE BRÈVE SYNTHÈSE DES RÉSULTATS DE L'ANALYSE DESCRIPTIVE

De façon générale, on observe donc les mêmes tendances dans la performance des élèves en fonctions de leurs caractéristiques individuelles pour les trois cohortes à l'étude, celles des années 2006, 2007 et 2008 et des

trois sous-populations sans le nombre élevé d'erreurs. Ces tendances sont plus marquées chez la population des élèves en difficulté, ceux qui font 24 erreurs et plus de syntaxe et ponctuation et ceux qui font 35 erreurs et plus d'orthographe.

Les élèves les plus jeunes réussissent mieux que les plus âgés. La différence de performance la plus importante est observée entre les élèves qui arrivent à l'âge normal en 5<sup>e</sup> secondaire, c'est-à-dire à 16 ans, et ceux qui sont âgés de 17 ans. Les filles réussissent mieux que les garçons et les élèves de langue maternelle française réussissent mieux que ceux de langues maternelles différentes du français en syntaxe et ponctuation.

Au niveau des variables caractéristiques de l'école, les élèves du réseau privé obtiennent de meilleurs résultats que ceux du réseau public. Les élèves qui fréquentent les écoles qui ont une proportion élevée d'élèves de langue maternelle française obtiennent de meilleurs résultats que les autres. Les élèves qui fréquentent les écoles de grande taille semblent avoir de meilleure performance que ceux qui fréquentent des écoles de petite taille.

## **CHAPITRE V**

### **LES ANALYSES MULTINIVEAUX MULTIVARIÉES**

La démarche d'analyse multiniveaux multivariée qu'adopte la présente recherche vise à faire ressortir la relation entre les deux critères de correction (la syntaxe et la ponctuation et l'orthographe) et les variables des trois niveaux (élève, école, année) qui forment la structure hiérarchique des données relatives aux élèves des écoles francophones de la région de Montréal durant une période de trois ans. L'élève (l'âge, le genre, la langue maternelle) y représente le premier niveau, l'école (le réseau, le pourcentage d'élèves de langue maternelle française, la mixité, la taille), le deuxième niveau et l'année (2006, 2007, 2008), le troisième niveau.

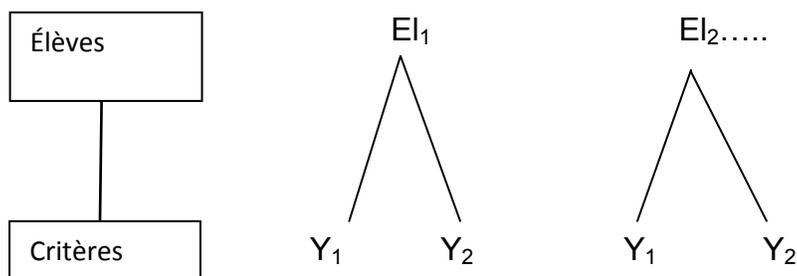
#### **5.1 LES MODÈLES À UN NIVEAU À VIDE**

Avant d'entamer l'analyse, nous avons calculé la corrélation entre les deux variables dépendantes (la syntaxe et la ponctuation, l'orthographe) pour l'ensemble des résultats obtenus par les élèves de chacune des trois cohortes sur lesquelles serait faite la régression multiniveaux multivariée. La valeur du coefficient de corrélation entre les deux variables dépendantes est de 0,60 en 2006, de 0,62 en 2007 et de 0,62 en 2008 au niveau de signification de 5 %. On note donc une corrélation positive significative à peu

près équivalente entre ces deux variables dépendantes pour chacune des trois années.

### 5.1.1 Le modèle 1 à vide pour l'ensemble des trois cohortes

L'analyse débute avec le modèle 1 à vide pour le premier niveau, l'élève, avec les deux variables dépendantes, la syntaxe et la ponctuation et l'orthographe. Ce modèle ne contient pas de variables indépendantes. La figure 14 illustre les éléments formant la structure de la modélisation multiniveaux multivariée de ce modèle avec l'élève et les deux critères de correction, où les symboles  $El_j$  ( $j = 1, \dots, n$ ) représentent les élèves, la variable dépendante  $Y_1$  représente la syntaxe et la ponctuation et la variable dépendante  $Y_2$  représente l'orthographe.



**Figure 14. Modélisation à un niveau : les élèves ( $El_j$ ) et deux variables dépendantes, syntaxe et ponctuation ( $Y_1$ ), orthographe ( $Y_2$ )**

Le tableau 25 de la page 126 indique pour les deux variables dépendantes les valeurs moyennes, les variances, les covariances, leurs erreurs types ainsi que la déviance de la modélisation.

Le nombre d'erreurs moyen des élèves en syntaxe et ponctuation est de 10,80 avec une erreur type (ET) de 0,04. Cette information est représentée sous la forme suivante 10,80 (0,04). En orthographe, le nombre d'erreurs moyen avec son erreur type est de 15,70 (0,06). Ces valeurs moyennes sont légèrement différentes de celles de chacune des trois cohortes à l'étude.

La variance de la performance des élèves en syntaxe et ponctuation est de 43,70 (0,34). Elle est plus petite que la variance de la performance des élèves en orthographe qui est de 110,70 (0,87). Ces résultats indiquent que la performance des élèves a tendance à être beaucoup moins constante en orthographe qu'en syntaxe et ponctuation. Ce premier modèle 1 à vide avec les coefficients aléatoires des variables ainsi que leurs erreurs types est représenté par les équations 2, 3, 4 et 5 ci-dessous.

$$(\text{Syntaxe et ponctuation})_{1j} = \beta_{0j} \cdot \text{Cons} \cdot \text{SyntaxePonct}_j + e \quad (\text{Éq. 2})$$

$$\beta_{0j} = 10,80 (0,04) + u_{0j} \quad (\text{Éq. 3})$$

$$\sigma_{0j}^2 = 43,70 (0,34)$$

$$(\text{Orthographe})_{2j} = \beta_{1j} \cdot \text{Cons} \cdot \text{Orthographe}_j + e \quad (\text{Éq. 4})$$

$$\beta_{1j} = 15,70 (0,06) + u_{1j} \quad (\text{Éq. 5})$$

$$\sigma_{1j}^2 = 110,70 (0,87)$$

$$\text{Cov} = 42,80 (0,45)$$

La covariance entre les deux variables est positive et elle vaut 42,80 (0,45). Elle signifie qu'en moyenne, les élèves finissants du secondaire qui font beaucoup d'erreurs de syntaxe et de ponctuation à l'épreuve unique de français écrit du MELS ont tendance à commettre beaucoup d'erreurs d'orthographe. Ceux qui font peu d'erreurs d'orthographe ont tendance à commettre également peu d'erreurs de syntaxe et de ponctuation, comme

l'indiquait la corrélation d'environ 0,62 entre les deux variables dépendantes pour chacune des trois cohortes. Pour vérifier la signification de cette covariance, nous avons effectué un test du rapport de vraisemblance. La valeur de la probabilité associée à cette covariance est très faible ( $p=0,000$ ). Nous pouvons ainsi conclure que la covariance entre les deux variables est significative.

### **5.1.2 Le modèle 4 à vide pour les sous-populations d'élèves sans le nombre élevé d'erreurs**

Le tableau 26 de la page 129 présente les valeurs moyennes, les variances, les covariances ainsi que les erreurs types et la déviance qui résultent de la modélisation des variables des sous-populations sans les valeurs de 24 erreurs en syntaxe et ponctuation et de 35 erreurs en orthographe. La valeur du coefficient de corrélation entre ces deux variables dépendantes est de 0,47 au niveau de signification de 5 %. Quoique ce coefficient de corrélation soit plus petit que celui de l'ensemble des élèves, il demeure positif et significatif.

Le nombre d'erreurs moyen des élèves des sous-populations est de 9,14 (0,03) en syntaxe et ponctuation et de 12,90 (0,05) en orthographe. Ces valeurs moyennes sont légèrement différentes de celles obtenues à chacune des sous-populations dans l'analyse descriptive. Elles sont calculées en tenant compte des élèves des trois sous-populations.

La variance de la performance des élèves en syntaxe et ponctuation est de 28,06 (0,26). Elle est plus petite que celle de la performance des élèves en orthographe, qui est de 70,05 (0,60). Comme pour l'ensemble des élèves, la performance des élèves des sous-populations a davantage tendance à être beaucoup plus variable en orthographe qu'en syntaxe et ponctuation. Ce modèle avec les coefficients aléatoires avec leurs erreurs types est représenté par les équations 6, 7, 8 et 9 ci-dessous.

$$(\text{Syntaxe et ponctuation})_{1j} = \beta_{0j} \cdot \text{Cons} + \text{SyntaxePonct}_j + \epsilon \quad (\text{Éq. 6})$$

$$\beta_{0j} = 9,14 (0,03) + u_{0j} \quad (\text{Éq. 7})$$

$$\sigma_{0j}^2 = 28,06 (0,26)$$

$$(\text{Orthographe})_{2j} = \beta_{1j} \cdot \text{Cons} + \text{Orthographe}_j + \epsilon \quad (\text{Éq. 8})$$

$$\beta_{1j} = 12,90 (0,05) + u_{1j} \quad (\text{Éq. 9})$$

$$\sigma_{1j}^2 = 70,05 (0,60)$$

$$\text{cov} = 20,88 (0,45)$$

La covariance entre les deux variables dépendantes de ces sous-populations est positive. Elle est de 20,88 (0,45). Ce qui signifie qu'en moyenne, les élèves qui font beaucoup d'erreurs de syntaxe et de ponctuation ont tendance à commettre beaucoup d'erreurs d'orthographe. À l'inverse, ceux qui font peu d'erreurs d'orthographe ont aussi tendance à commettre également peu d'erreurs de syntaxe et de ponctuation, comme l'indiquait la corrélation positive de 0,47 entre les résultats des élèves pour les deux variables dépendantes. Nous avons vérifié la signification de cette covariance par un test du rapport de vraisemblance. La valeur de la probabilité est extrêmement faible. Nous pouvons par conséquent conclure que cette covariance est significative.

Les modèles 1 et 4 que nous venons d'ajuster pour l'ensemble des élèves des trois cohortes et des trois sous-populations supposent l'indépendance

entre les élèves. Il faut rappeler que les groupes sont constitués d'élèves qui viennent d'une même école. Nous voulons vérifier dans quelle mesure il existe une différence dans la variabilité de la performance des élèves à l'intérieur des écoles et entre ces écoles. Pour ce faire, nous serons amené à inclure le deuxième niveau, l'école, dans notre modèle.

## 5.2 LES MODÈLES À DEUX NIVEAUX À VIDE

La figure 15 ci-dessous illustre les éléments de la structure de la modélisation multiniveaux multivariée à coefficients aléatoires à deux niveaux (élève, école), sans les variables indépendantes. La variable dépendante  $Y_1$  représente la syntaxe et la ponctuation et la variable dépendante  $Y_2$  représente l'orthographe; les  $Ei$ , où ( $i = 1 \dots \dots \dots I$ ) représentent les élèves et les  $Ec_j$ , ( $j = 1 \dots \dots \dots J$ ), les écoles.

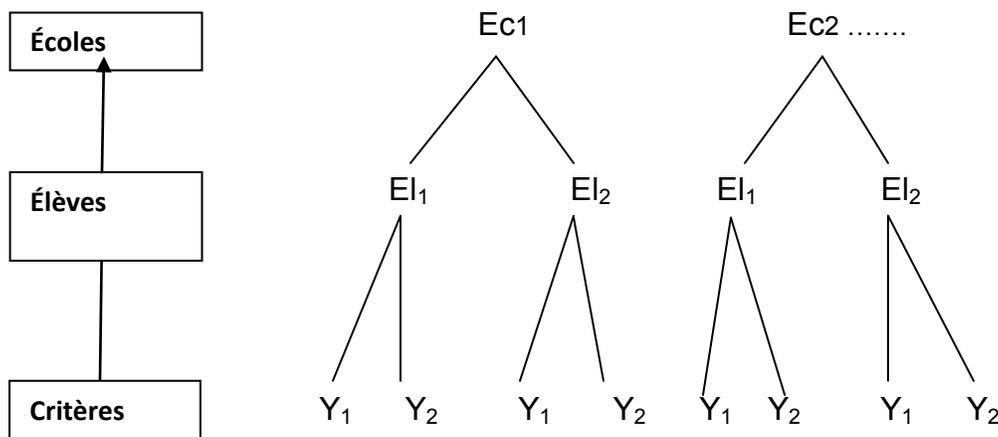


Figure 15. Modélisation multivariée à deux niveaux (élève, école) et deux variables dépendantes : syntaxe et ponctuation ( $Y_1$ ), orthographe ( $Y_2$ )

### 5.2.1 Le modèle 2 à vide pour l'ensemble des trois cohortes

Le tableau 25 de la page 126 présente les résultats de la modélisation du modèle 2 à vide à deux niveaux. La moyenne du nombre d'erreurs des élèves est de 11,16 (0,22) en syntaxe et ponctuation et de 16,07 (0,35) en orthographe. Nous avons cependant départagé la variance et la covariance de chacune des variables dépendantes entre une composante au niveau de l'élève et une composante au niveau de l'école.

Les résultats de cette modélisation indiquent que 75,72 % de la variance totale de la performance des élèves en syntaxe et ponctuation et 74,93 % de la variance totale de la performance des élèves en orthographe se retrouvent au niveau de l'élève, le premier niveau. Et 24,28 % de la variance totale de la performance des élèves en syntaxe et ponctuation de même que 25,07 % de la variance totale de la performance des élèves en orthographe se retrouvent au niveau de l'école, le deuxième niveau. Ce modèle 2 est représenté par les équations 10, 11, 12 et 13 ci-dessous :

$$(\text{Syntaxe et ponctuation})_{1jk} = \beta_{0jk} \cdot \text{Cons} \cdot \text{SyntaxePonct}_{jk} + e \quad (\text{Éq. 10})$$

$$\beta_{0jk} = 11,16 (0,22) + v_{0k} + u_{0jk} \quad (\text{Éq. 11})$$

$$\sigma_{v0k}^2 = 10,52 (1,01)$$

$$\sigma_{v1k}^2 = 27,79 (2,70)$$

$$\text{cov} = 14,70 (1,50)$$

$$(\text{Orthographe})_{2jk} = \beta_{1jk} \cdot \text{Cons} \cdot \text{Orthographe}_{jk} + e \quad (\text{Éq. 12})$$

$$\beta_{1jk} = 16,07 (0,35) + v_{1k} + u_{1jk} \quad (\text{Éq. 13})$$

$$\sigma u_{0jk}^2 = 32,80 (0,30)$$

$$\sigma u_{1jk}^2 = 82,98 (0,65)$$

$$\text{cov} = 26,94 (0,30)$$

Les résultats de cette modélisation indiquent que 35,30 % de la covariance entre les deux variables se situent au niveau de l'école. La covariance entre les deux variables dépendantes est positive aux deux niveaux. Elle est de 26,94 (0,30) au niveau de l'élève et de 14,70 (1,50) au niveau de l'école.

Il importe de préciser le résultat de cette modélisation. D'abord au niveau de l'établissement scolaire, on parle d'une moyenne d'école, c'est-à-dire d'une valeur agrégée; dans les écoles où les élèves font en moyenne beaucoup d'erreurs de syntaxe et de ponctuation, ils ont aussi tendance à commettre en moyenne beaucoup d'erreurs d'orthographe. Sur le plan individuel, on parle d'une mesure pour chaque élève; à l'intérieur d'une école, l'élève qui fait beaucoup d'erreurs de syntaxe et de ponctuation, répétons-le, a également tendance à commettre beaucoup d'erreurs d'orthographe.

Nous avons calculé la corrélation entre les deux variables dépendantes associées à chacun des deux niveaux : elle est de 0,86 au niveau de l'école et de 0,52 au niveau de l'élève. Ce résultat indique une plus forte corrélation entre le nombre moyen d'erreurs de syntaxe et ponctuation et d'orthographe au niveau de l'école que celle qui existe entre les élèves à l'intérieur d'une école. En d'autres termes, il y a plus d'homogénéité entre les écoles qu'il y en a entre les élèves à l'intérieur de l'école.

### 5.2.2 Le modèle 5 à vide des sous-populations d'élèves sans le nombre élevé d'erreurs

Le tableau 26 de la page 129 présente les moyennes relatives à la distribution de la performance des élèves appartenant aux sous-populations d'élèves sans le nombre élevé d'erreurs. Le nombre d'erreurs moyen est de 9,74 (0,15) en syntaxe et ponctuation et de 13,71 (0,24) en orthographe. Il s'agit de moyennes très semblables à celles que nous avons obtenues à la modélisation précédente de ces sous-populations par le modèle 4 à un niveau. Les résultats de cette modélisation indiquent que 82,37 % de la variance totale de la performance des élèves en syntaxe et ponctuation et 82,29 % de la variance totale de la performance des élèves en orthographe sont attribuables à l'élève. Et 17,63 % de la variance totale de la performance des élèves en syntaxe et ponctuation de même que 17,71 % de la variance totale de la performance des élèves en orthographe sont attribuables à l'école. Ce modèle 5 est représenté ci-dessous par les équations 14, 15, 16 et 17 :

$$(\text{Syntaxe et ponctuation})_{1jk} = \beta_{0jk} \cdot \text{Cons} \cdot \text{SyntaxePonct}_{jk} + e \quad (\text{Éq. 14})$$

$$\beta_{0jk} = 9,74 (0,15) + v_{0k} + u_{0jk} \quad (\text{Éq. 15})$$

$$\sigma_{v_{0k}}^2 = 4,95 (0,49)$$

$$\sigma_{v_{1k}}^2 = 12,52 (1,24)$$

$$\text{cov} = 5,84 (0,68)$$

$$(\text{Orthographe})_{2jk} = \beta_{1jk} \cdot \text{Cons} \cdot \text{Orthographe}_{jk} + e \quad (\text{Éq. 16})$$

$$\beta_{1jk} = 13,71 (0,24) + v_{1k} + u_{1jk} \quad (\text{Éq. 17})$$

$$\sigma_{u_{0jk}}^2 = 23,12 (0,20)$$

$$\sigma_{u_{1jk}}^2 = 58,16 (0,50)$$

$$\text{cov} = 14,53 (0,24)$$

La covariance entre les deux variables dépendantes est positive aux deux niveaux. Elle est de 14,53 (0,24) au niveau de l'élève et de 5,84 (0,68) au niveau de l'école. Comme pour l'ensemble des élèves de la population, ce résultat indique que, de façon générale, dans les écoles où les élèves font en moyenne beaucoup d'erreurs de syntaxe et de ponctuation, ils ont tendance à commettre également beaucoup d'erreurs d'orthographe. De plus, à l'intérieur d'une école, l'élève qui fait beaucoup d'erreurs de syntaxe et de ponctuation a aussi tendance à faire beaucoup d'erreurs d'orthographe.

Le fait de retirer les données sur les élèves en difficulté, c'est-à-dire ceux qui font 24 erreurs en syntaxe et ponctuation et 35 erreurs en orthographe, neutralise dans une certaine mesure l'effet des facteurs d'école sur la performance des élèves. La variance totale de la performance des élèves passe de l'ordre de grandeur de 25 % pour l'ensemble de la population à 17 % dans cette sous-population au niveau de l'école. La variance totale de la performance des élèves attribuable à l'élève passe de près de 75 % pour l'ensemble de la population à près de 82 % dans cette sous-population. Les facteurs d'école semblent avoir une plus grande influence sur la performance des élèves en difficulté, ceux qui font 24 erreurs de syntaxe et ponctuation et 35 erreurs d'orthographe. Par ailleurs, les coefficients des variables de la modélisation de ces sous-populations d'élèves sans le nombre élevé d'erreurs sont estimés avec une meilleure précision que ceux de l'ensemble de la population, pour ce qui est d'une plus petite erreur type associée à chaque estimation.

### 5.3 LE MODÈLE À TROIS NIVEAUX À VIDE

Nous allons maintenant intégrer dans le modèle un troisième niveau, les années des cohortes (2006, 2007 et 2008) pour lesquelles nous disposons de données. La figure 16 présente les éléments de la structure de la modélisation multiniveaux multivariée à trois niveaux dans laquelle la variable dépendante  $Y_1$  représente la syntaxe et la ponctuation, la variable dépendante  $Y_2$  représente l'orthographe; les  $Ei_i$  représentent les élèves, les  $Ec_j$  (représentent) les écoles et les symboles  $A_k$  (représentent) les années.

#### 5.3.1 Le modèle 3 à trois niveaux à vide des trois cohortes

Les résultats de la modélisation à vide par le modèle 3 des trois cohortes d'élèves sont présentés au tableau 25 de la page 126. La plus grande part de variabilité dans la performance des élèves se retrouve au niveau de l'élève et de l'école.

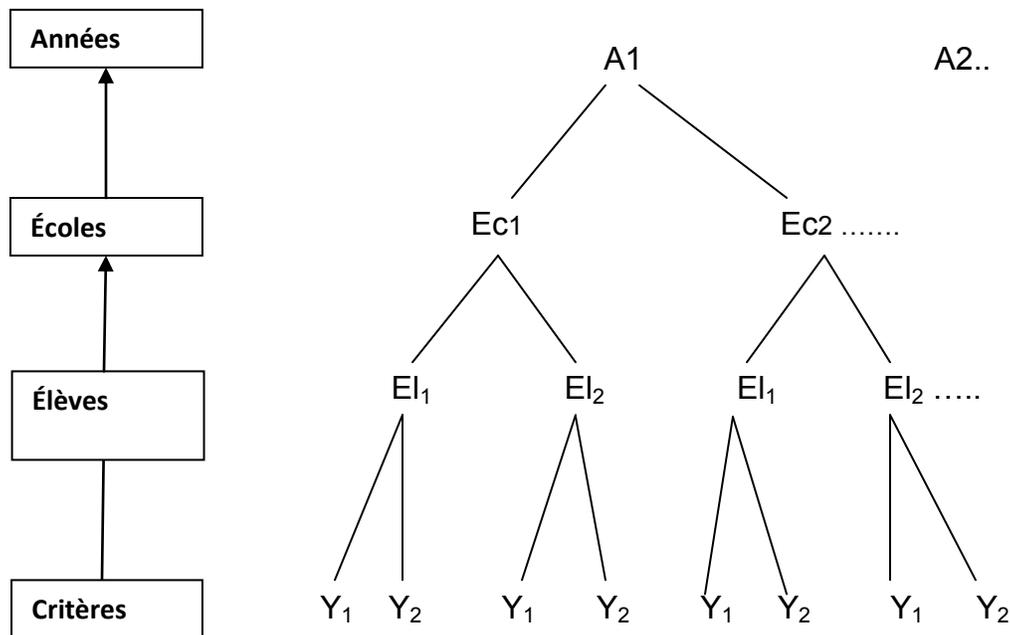


Figure 16. Modélisation multivariée à trois niveaux (élève, école, année) et deux variables dépendantes : syntaxe et ponctuation ( $Y_1$ ), orthographe ( $Y_2$ )

En syntaxe et ponctuation, 76,04 % de la variance totale de la performance des élèves est attribuable à l'élève, 23,70 % à l'école et 0,29 % à l'année. En orthographe, 74,90 % de la variance totale de la performance des élèves est attribuable à l'élève, 25,10 % à l'école et 0,00 % à l'année. Cette analyse montre dans quelle mesure les relations entre les performances des élèves en syntaxe et ponctuation et en orthographe et les facteurs caractéristiques des élèves et des écoles sont plus marquées que celles entre les performances des élèves selon les années. Les résultats de cette modélisation sont représentés par les équations 18, 19, 20 et 21 ci-dessous.

$$(\text{Syntaxe et ponctuation})_{1jkl} = \beta_{0jkl} \cdot \text{Cons} \cdot \text{SyntaxePonct}_{jkl} + e \quad (\text{Éq. 18})$$

$$\beta_{0jkl} = 11,16 (0,30) + f_{0l} + v_{0kl} + u_{0jkl} \quad (\text{Éq. 19})$$

$$(\text{Orthographe})_{2jkl} = \beta_{1jkl} \cdot \text{Cons} \cdot \text{Orthographe}_{jkl} + e \quad (\text{Éq. 20})$$

$$\beta_{1jkl} = 16,07 (0,35) + f_{1l} + v_{1kl} + u_{1jkl} \quad (\text{Éq. 21})$$

$$\sigma f_{0l}^2 = 00,13 (0,13)$$

$$\sigma f_{1l}^2 = 00,00 (0,00)$$

$$\text{cov} = 00,00 (0,00)$$

$$\sigma v_{0kl}^2 = 10,21 (0,98)$$

$$\sigma v_{1kl}^2 = 27,80 (2,67)$$

$$\text{cov} = 14,50 (1,50)$$

$$\sigma_{u_{0jkl}}^2 = 32,80 (0,30)$$

$$\sigma_{u_{1jkl}}^2 = 82,97 (0,65)$$

$$\text{cov} = 26,94 (0,33)$$

Nous n'avons pas constaté de variabilité significative au niveau de l'année. La même observation s'applique d'ailleurs à notre première série d'analyses descriptives des données provenant des trois populations, où les corrélations entre les deux variables dépendantes à l'étude ainsi que les moyennes de l'ensemble des élèves de chacune des cohortes et des sous-populations étaient assez semblables.

Ces résultats indiquent, toutes choses étant égales par ailleurs, que les performances des élèves des différentes cohortes, celles des années 2006, 2007 et 2008, sont sensiblement les mêmes sans égard aux variables qui seront analysées. Par conséquent, nous n'allons pas retenir le niveau année dans les modélisations subséquentes, celui-ci n'apportant aucune information supplémentaire dans la compréhension de la variabilité de la performance des élèves.



### 5.3.2 Le modèle 6 à trois niveaux à vide pour les sous-populations d'élèves sans le nombre élevé d'erreurs

Procédons maintenant à l'intégration du troisième niveau, celui de l'année, dans la modélisation multiniveaux multivariée des sous-populations d'élèves sans le nombre élevé d'erreurs. Les résultats de cette modélisation par le modèle 6 sont présentés au tableau 26 de la page 129. Comme pour l'ensemble de la population examinée, la plus grande part de la variabilité dans la performance des élèves de ces sous-populations se retrouve aux niveaux de l'élève et de l'école. En syntaxe et ponctuation, 82,63 % de la variance totale de la performance des élèves est attribuable à l'élève, 16,98 % à l'école et 0,39 % à l'année. En orthographe, 82,27 % de la variance totale de la performance des élèves est attribuable à l'élève, 17,73 % à l'école et 0,00 % à l'année.

Notre analyse montre également pour ces sous-populations, à l'instar de la modélisation des données de l'ensemble des trois cohortes, dans quelle mesure les relations entre la performance des élèves en syntaxe et ponctuation et en orthographe et les facteurs caractéristiques des élèves et des écoles, sont plus marquées que celles entre la performance des élèves et les années. Les résultats de cette modélisation sont représentés par les équations 22, 23, 24 et 25 ci-dessous.

$$(\text{Syntaxe et ponctuation})_{1jkl} = \beta_{0jkl} \cdot \text{Cons} \cdot \text{SyntaxePonct}_{jkl} + e \quad (\text{Éq. 22})$$

$$\beta_{0jkl} = 9,74 (0,24 + f_{0l} + v_{0kl} + u_{0jkl}) \quad (\text{Éq. 23})$$

$$(\text{Orthographe})_{2jkl} = \beta_{1jkl} \cdot \text{Cons} \cdot \text{Orthographe}_{jkl} + e \quad (\text{Éq. 24})$$

$$\beta_{1jkl} = 13,71 (0,24) + f_{1l} + v_{1kl} + u_{1jkl} \quad (\text{Éq. 25})$$

$$\sigma f_{0l}^2 = 00,11 (0,12)$$

$$\sigma f_{1l}^2 = 00,00 (0,00)$$

$$\text{cov} = 00,00 (0,00)$$

$$\sigma v_{0kl}^2 = 4,75 (0,48)$$

$$\sigma v_{1kl}^2 = 12,53 (1,25)$$

$$\text{cov} = 5,47 (0,60)$$

$$\sigma u_{0jkl}^2 = 23,12 (0,20)$$

$$\sigma u_{1jkl}^2 = 58,16 (0,50)$$

$$\text{cov} = 14,53 (0,24)$$

La modélisation des variables par le modèle à vide à trois niveaux de ces sous-populations ne nous a pas indiqué de variabilité significative au niveau de l'année. Nous n'allons donc pas retenir le niveau année dans les modélisations subséquentes.

À cette autre étape de la modélisation, nous allons intégrer les variables individuelles caractéristiques de l'élève dans une première modélisation à deux niveaux (élève, école), pour tenter d'expliquer les relations entre la performance des élèves selon les deux critères de correction, la syntaxe et la ponctuation et l'orthographe, et les variables indépendantes caractéristiques des élèves. Nous intégrerons ensuite les variables caractéristiques de l'école dans une dernière étape pour compléter notre analyse.



## 5.4 LA MODÉLISATION DES VARIABLES CARACTÉRISTIQUES DES ÉLÈVES

Cette nouvelle étape de nos analyses teste le modèle à deux niveaux (élève, école) avec les trois variables catégorielles indépendantes caractéristiques de l'élève : l'âge, le genre et la langue maternelle. Le coefficient des variables dichotomiques s'interprètera ainsi : une valeur positive indique que c'est la catégorie de la variable codée par le chiffre 1 qui fait plus d'erreurs. Une valeur négative indique que c'est la catégorie de la variable codée par 1 qui fait moins d'erreurs. Une valeur proche de zéro indique que les deux catégories font sensiblement le même nombre d'erreurs. Dans la modélisation qui suit, le genre de l'élève est codé ainsi : le chiffre 1 représente les garçons et 0 les filles. Pour la langue maternelle, la donnée des élèves de langues maternelles différentes du français est codée par le chiffre 0, et la donnée des élèves de langue maternelle française est codée par le chiffre 1. La variable âge comprend les élèves qui sont âgés de 15, 16, 17 et 18 ans ; plus le coefficient de cette variable est élevé, plus ce sont les élèves plus âgés qui font plus de fautes.

### 5.4.1 La modélisation des variables caractéristiques de l'élève pour les trois cohortes

Les équations 26, 27, 28 et 29 ci-dessous ainsi que le tableau 27 de la page 136 présentent les résultats de la modélisation par le modèle 7 des variables individuelles caractéristiques de l'élève. Celles-ci expliquent 13,85 % de la variance totale de la performance des élèves en syntaxe et ponctuation et 9,81 % de la variance totale de la performance des élèves en orthographe.

$$(\text{Syntaxe et ponctuation})_{ijk} = \beta_{0jk} + 1,78(0,05)\text{Âge} + 0,34(0,07)\text{Genre} - 2,08(0,07)\text{Français}$$

**(Éq. 26)**

$$\beta_{0jk} = -16,80(0,90) + v_{ok} + u_{0jk}$$

**(Éq.27)**

$$\sigma_{v_{0k}}^2 = 6,63 (0,65)$$

$$\sigma_{v_{1k}}^2 = 20,65 (2,00)$$

$$\text{cov} = 9,99 (1,05)$$

$$(\text{Orthographe})_{2jk} = \beta_{1jk} + 2,29 (0,09)\text{Âge} + 3,04 (0,11)\text{Genre} + 0,71 (0,11)\text{Français} \quad (\text{Éq. 28})$$

$$\beta_{1jk} = -23,04 (1,51) + v_{1k} + u_{1jk} \quad (\text{Éq. 29})$$

$$\sigma_{u_{0jk}}^2 = 30,69 (0,24)$$

$$\sigma_{u_{1jk}}^2 = 79,25 (0,60)$$

$$\text{cov} = 25,25 (0,31)$$

En syntaxe et ponctuation, l'analyse de la variabilité de la performance des élèves nous indique que l'effet le plus important est relié à la langue maternelle de l'élève, avec un coefficient de 2,08 (0,07) en faveur des élèves de langue maternelle française, suivi de l'âge avec un coefficient de 1,78 (0,05) en faveur des élèves les plus jeunes, et du genre avec un coefficient de 0,34 (0,07) en faveur des filles.

En orthographe, les résultats de cette modélisation indiquent que le genre affiche le poids relatif le plus important avec un coefficient de 3,04 (0,11) en faveur des filles, suivi de l'âge avec un coefficient de 2,30 (0,09) en faveur des élèves les plus jeunes, et de la langue maternelle avec un coefficient de 0,71 (0,11) en faveur des élèves de langues maternelles différentes du français.

#### 5.4.2 La modélisation de variables caractéristiques de l'élève pour les sous-populations d'élèves sans le nombre élevé d'erreurs

Pour les sous-populations d'élèves sans les valeurs de 24 erreurs en syntaxe et ponctuation et de 35 erreurs en orthographe, l'analyse multiniveaux multivariée applique le modèle à deux niveaux (élève, école) avec les trois variables indépendantes caractéristiques de l'élève : l'âge, le genre et la langue maternelle. Les équations 30, 31, 32 et 33 présentent les résultats de cette modélisation avec les coefficients des variables individuelles caractéristiques de l'élève. Le tableau 28 de la page 139 indique que les variables indépendantes caractéristiques de l'élève expliquent 11,18 % de la variance totale de la performance des élèves en syntaxe et ponctuation, et 6,44 % de la variance totale de leur performance en orthographe.

$$(\text{Syntaxe et ponctuation})_{1jk} = \beta_{0jk} + 1,46(0,06)\hat{\text{Age}} - 0,05(0,06) \text{ Genre} - 2,10(0,07) \text{ Français} \quad (\text{Éq. 30})$$

$$\beta_{0jk} = -12,69 (0,97) + v_{0k} + u_{0jk} \quad (\text{Éq.31})$$

$$\sigma_{v_{0k}}^2 = 3,27 (0,33)$$

$$\sigma_{v_{1k}}^2 = 10,00 (1,00)$$

$$\text{cov} = 4,28 (0,50)$$

$$(\text{Orthographe})_{2jk} = \beta_{1jk} + 1,92(0,09)\hat{\text{Age}} + 2,38(0,09) \text{ Genre} + 0,46(0,10) \text{ Français} \quad (\text{Éq. 32})$$

$$\beta_{1jk} = -16,23 (1,60) + v_{1k} + u_{1jk} \quad (\text{Éq. 33})$$

$$\sigma_{u_{0jk}}^2 = 21,66$$

$$\sigma_{u_{1jk}}^2 = 56,13 (0,48)$$

$$\text{cov} = 13,96 (0,29)$$

La variabilité de la performance des élèves en syntaxe et ponctuation indique que l'effet le plus important est relié à la langue maternelle de l'élève, avec un coefficient de 2,10 (0,07) en faveur des élèves de langue maternelle française, suivi de l'âge avec un coefficient de 1,46 (0,06) en faveur des élèves les plus jeunes et du genre avec un coefficient de 0,05 (0,06), plutôt négligeable, en faveur des filles.

En orthographe, contrairement à la syntaxe et à la ponctuation, cette modélisation indique que le genre obtient le poids relatif le plus important avec un coefficient de 2,38 (0,09) en faveur des filles, suivi de l'âge avec un coefficient de 1,92 (0,09) en faveur des élèves les plus jeunes, et de la langue maternelle avec un coefficient de 0,46 (0,11) en faveur des élèves de langues maternelles autres que le français. Les coefficients des variables de cette sous-population sont estimés avec une meilleure précision pour ce qui est d'une plus petite erreur type comparativement à l'ensemble de la population.

## 5.5 LA MODÉLISATION DES VARIABLES CARACTÉRISTIQUES DE L'ÉLÈVE ET DE L'ÉCOLE

Le modèle à deux niveaux (élève, école) avec les variables caractéristiques des élèves, sans les variables d'école, servira de référence pour l'analyse du modèle qui intègre les facteurs d'école : le réseau, la mixité, la taille et le pourcentage des élèves de l'école dont la langue maternelle est le français. Dans ce modèle, la variable *réseau* sera codée ainsi : le chiffre 0 représente le secteur privé et le chiffre 1 le secteur public. Pour la mixité, le chiffre 0 représente les écoles réservées aux filles et le chiffre 1 les écoles mixtes. La

taille de l'école et le pourcentage des élèves seront modélisés avec leur valeur brute.

### 5.5.1 La modélisation des variables caractéristiques des écoles pour l'ensemble des trois cohortes

Le tableau 27 de la page 136 ainsi que les équations 34, 35, 36 et 37 représentent le modèle 8 avec les coefficients aléatoires des variables caractéristiques de l'élève et de l'école pour l'ensemble des trois cohortes.

$$\begin{aligned} (\text{Syntaxe et ponctuation})_{1jk} = & \beta_{0jk} + 1,77 (0,05) \text{ Âge} + 0,34 (0,07) \text{ Genre} - 2,05 (0,07) \\ & \text{Français} + 2,20 (0,37) \text{ Réseau} + 0,00 (0,00) \text{ Taille} + 0,02 (0,50) \\ & \text{Mixité} - 1,46 (0,07) \% \text{LangueMaternelle} \end{aligned} \quad (\text{Éq. 34})$$

$$\beta_{0jk} = -15,83 (1,50) + v_{0k} + u_{0jk} \quad (\text{Éq. 35})$$

$$\sigma_{v_{0k}}^2 = 5,27 (0,52)$$

$$\sigma_{v_{1k}}^2 = 14,23 (1,40)$$

$$\text{cov} = 7,24 (0,78)$$

$$\begin{aligned} (\text{Orthographe})_{2jk} = & \beta_{1jk} + 2,30 (0,09) \text{ Âge} + 3,03 (0,11) \text{ Genre} + 0,71 (0,11) \text{ Français} + \\ & 5,63 (0,60) \text{ Réseau} + 0,00 (0,00) \text{ Taille} - 0,12 (0,81) \text{ Mixité} \\ & + 0,43 (1,14) \% \text{LangueMaternelle} \end{aligned} \quad (\text{Éq. 36})$$

$$\beta_{1jk} = -23,94 (1,74) + v_{1k} + u_{1jk} \quad (\text{Éq. 37})$$

$$\sigma_{u_{0jk}}^2 = 30,69 (0,24)$$

$$\sigma_{u_{1jk}}^2 = 79,26 (0,63)$$

$$\text{cov} = 25,55 (0,31)$$

Les résultats de cette modélisation pour l'ensemble des élèves des trois cohortes indiquent que les variables caractéristiques de l'école expliquent 3,65 % de la variance totale de la performance des élèves en syntaxe et ponctuation, et 6,42 % de la variance totale de la performance des élèves en orthographe. Dans cette modélisation, la réseau obtient le poids relatif le plus important avec un coefficient de 5,63 (0,60) en orthographe et 2,20 (0,37) en syntaxe et ponctuation en faveur des écoles privées, suivi du pourcentage des élèves de l'école de langue maternelle française avec un coefficient de 1,46 (0,70) en syntaxe et ponctuation, et 0,43 (1,14) en orthographe en faveur des écoles qui comptent un haut pourcentage d'élèves de langue maternelle française. La taille et la mixité avec des coefficients respectivement de 0,00 (0,00) et 0,02 (0,50) en syntaxe et ponctuation et 0,00 (0,00) et 0,12 (0,81) en orthographe présentent un poids plutôt négligeable dans la performance des élèves.

**Tableau 27 - Estimations des coefficients et de leurs erreurs types des variables caractéristiques de l'élève et de l'école des trois cohortes**

Coefficients	Modèle 2		Modèle 7		Modèle 8	
	Syntaxe	Orthographe	Syntaxe	Orthographe	Syntaxe	Orthographe
Élève	Estimation (ET)	Estimation (ET)	Estimation (ET)	Estimation (ET)	Estimation (ET)	Estimation (ET)
Moyenne	11,16 (0,22)	16,07 (0,35)	-16,81 (0,90)	-23,04 (1,50)	-15,85 (1,07)	-23,94 (1,74)
Genre			-0,34 (0,07)	-3,04 (0,11)	-0,34 (0,07)	-3,03 (0,11)
Langue m.			-2,08 (0,07)	0,71 (0,11)	-2,05 (0,07)	0,71 (0,11)
Âge			1,78 (0,05)	2,29 (0,09)	1,78 (0,05)	2,30 (0,09)
<b>École</b>						
Réseau					2,20 (0,37)	5,63 (0,60)
Mixité					0,02 (0,50)	-0,12 (0,81)
Taille					-0,00 (0,00)	-0,00 (0,00)
%Langue m.					-1,46 (0,70)	0,43 (1,14)
<b>Aléatoires</b>						
Var. année						
Covar. année						
Var. école	10,52 (1,01)	27,79 (2,70)	6,63 (0,65)	20,65 (2,00)	5,27 (0,52)	14,23 (1,40)
Covar. école	14,678(1,53)		9,99 (1,05)		7,24 (0,78)	
Var. élève	32,80 (0,30)	82,98 (0,70)	30,69 (0,24)	79,25 (0,60)	30,69 (0,24)	79,26 (0,63)
Covar. élève	26,94 (0,30)		25,251 (0,31)		25,55 (0,31)	
<b>2 log (vraisemblance)</b>	429187,73		425354,52		425252,17	
<b>Déviance</b>			3833,2		102.35	
	<b>Pourcentage de variance expliquée du modèle 1 à vide</b>		<b>Pourcentage de variance expliquée du modèle 2 à vide</b>		<b>Pourcentage de variance expliquée du modèle 3 à vide</b>	
Niveau année						
Niveau école	24,28 %	25,09 %			<b>3,65 %</b>	<b>6,42 %</b>
Niveau élève	75,72 %	74,91 %	<b>13,85 %</b>	<b>9,81 %</b>		
<b>VTE</b>	<b>100 %</b>	<b>100 %</b>				

### 5.5.2 La modélisation des variables caractéristiques des écoles pour les sous-populations d'élèves sans le nombre élevé d'erreurs

Cette section présente la modélisation des variables explicatives caractéristiques de l'élève et de l'école pour les sous-populations d'élèves sans les valeurs de 24 erreurs en syntaxe et ponctuation et de 35 erreurs en orthographe. Le modèle 9 déjà traité avec les variables caractéristiques des élèves servira de référence pour l'analyse de ce modèle 10 qui intègre les variables caractéristiques de l'école : le réseau, la mixité, la taille de l'école et le pourcentage des élèves de l'école ayant le français comme langue maternelle. Le tableau 28 à la page 139 ainsi que les équations 38, 39, 40 et 41 présentent les résultats de cette modélisation avec les coefficients des variables caractéristiques de l'élève et de l'école.

$$\begin{aligned} (\text{Syntaxe et ponctuation})_{1jk} = & \beta_{0jk} + 1,46 (0,06) \hat{\text{Âge}} + 0,05 (0,06) \text{Genre} - 2,07 (0,07) \text{Français} \\ & + 1,29 (0,27) \text{Réseau} - 0,16 (0,37) \text{Mixité} + 0,00 (0,00) \text{Taille} \\ & - 1,31 (0,50) \% \text{LangueMaternelle} \end{aligned} \quad (\text{Éq. 38})$$

$$\beta_{0jk} = -11,65 (1,07) + v_{0k} + u_{0jk} \quad (\text{Éq. 39})$$

$$\sigma_{v_{0k}}^2 = 2,74 (0,28)$$

$$\sigma_{v_{1k}}^2 = 6,96 (0,72)$$

$$\text{cov} = 3,17 (0,39)$$

$$\begin{aligned} (\text{Orthographe})_{2jk} = & \beta_{1jk} + 1,90 (0,09) \hat{\text{Âge}} - 2,37 (0,09) \text{Genre} + 0,46 (0,10) \text{Français} \\ & + 3,88 \text{Réseau} - 0,00 (0,00) \text{Taille} - 0,28 (0,58) \text{Mixité} \\ & + 0,56 (0,82) \% \text{LangueMaternelle} \end{aligned} \quad (\text{Éq. 40})$$

$$\beta_{1jk} = -16,71 (1,71) + v_{1k} + u_{1jk} \quad (\text{Éq. 41})$$

$$\sigma u_{0jk}^2 = 21,65 (0,19)$$

$$\sigma u_{1jk}^2 = 56,14 (0,48)$$

$$\text{cov} = 13,96 (0,23)$$

Les résultats de cette modélisation indiquent que le réseau présente le poids relatif le plus important, avec un coefficient de 3,88 (0,43) en orthographe et 1,29 (0,27) en syntaxe et ponctuation en faveur du secteur privé, suivi de la proportion des élèves de l'école de langue maternelle française avec 1,31 (0,50) en syntaxe et ponctuation, et 0,56 (0,82) en orthographe, ce qui est très peu significatif. Ces variables caractéristiques de l'école expliquent 2,46 % de la variance totale de la performance des élèves en syntaxe et ponctuation, et 5,43 % de la variance totale de la performance des élèves en orthographe.

La contribution respective des variables caractéristiques de l'élève, le premier niveau, et de l'école, le deuxième niveau, diffère selon la variable dépendante considérée. Les variables caractéristiques des élèves, avec un poids relatif de 13,85 %, toutes choses étant égales par ailleurs, contribuent davantage à l'explication de la performance des élèves en syntaxe et ponctuation qu'en orthographe, critère pour lequel elles obtiennent un poids relatif de 9,81 %. En revanche, les variables caractéristiques des écoles, avec un poids relatif de 6,42 %, contribuent davantage à l'explication de la performance des élèves en orthographe qu'en syntaxe et ponctuation, critère pour lequel elles obtiennent un poids relatif de 3,65 %. Ce résultat nous indique que les facteurs d'école contribuent davantage à l'apprentissage des élèves en orthographe qu'en syntaxe et ponctuation.

**Tableau 28 - Estimation des coefficients et de leurs erreurs types des variables caractéristiques des élèves et des écoles des sous-populations d'élèves sans le nombre élevé d'erreurs**

Coefficients	Modèle 5		Modèle 9		Modèle 10	
	Syntaxe	Orthographe	Syntaxe	Orthographe	Syntaxe	Orthographe
<b>Élève</b>	<b>Estimation (ET)</b>	<b>Estimation (ET)</b>	<b>Estimation (ET)</b>	<b>Estimation (ET)</b>	<b>Estimation (ET)</b>	<b>Estimation (ET)</b>
Moyenne	9,74 (0,15)	13,71 (0,24)	-12,69 (0,97)	-16,23 (1,60)	-11,65 (1,07)	-16,71 (1,71)
Genre			- 0,05 (0,06)	- 2,38 (0,09)	- 0,05 (0,06)	- 2,37 (0,09)
Langue M.			- 2,10 (0,07)	+ 0,46 (0,10)	- 2,07 (0,07)	+0,46 (0,10)
Âge			+1,46 (0,06)	+ 1,92 (0,09)	+ 1,46 (0,06)	+1,90 (0,09)
<b>École</b>						
Réseau					+ 1,29 (0,27)	+3,88 (0,43)
Mixité					- 0,16 (0,37)	- 0,28 (0,58)
Taille					- 0,00 (0,00)	- 0,00 (0,00)
%Langue m.					- 1,31 (0,50)	+0,56 (0,82)
<b>Aléatoires</b>						
Var. année						
Covar. année						
Var. école	4,95(0,50)	12,52(1,24)	3,27(0,33)	10,00(1,00)	2,74(0,28)	6,96
Covar. école	5,84 (0,68)		4,28 (0,50)		3,17 (0,39)	
Var. élève	23,12(0,20)	58,16(0,50)	21,66(0,18)	56,13(0,48)	21,65(0,19)	56,14(0,48)
Covar. Élève	14,53 (0,20)		13,96 (0,29)		13,96 (0,23)	
<b>2 log (vraisemblance)</b>	349451,18		346461,38		346457,56	
<b>Déviance</b>			2989,80		3,82	
	<b>Pourcentage de variance expliquée</b>		<b>Pourcentage de variance expliquée</b>		<b>Pourcentage de variance expliquée</b>	
Niveau année						
Niveau école	17,63 %	17,71 %			<b>2,46 %</b>	<b>5,43 %</b>
Niveau élève	82,37 %	82,29 %	<b>11,18 %</b>	<b>6,44 %</b>		
<b>VTE</b>	<b>100 %</b>	<b>100 %</b>				

## **CHAPITRE VI**

### **DISCUSSION**

Les analyses qui précèdent nous ont permis de vérifier que dans la région métropolitaine de Montréal, les résultats des élèves finissants du secondaire à l'épreuve unique de français écrit du ministère de l'Éducation, du Loisir et du Sport (MELS), selon les deux critères spécifiques de correction, la syntaxe et la ponctuation ainsi que l'orthographe, s'expliquent principalement par un certain nombre de variables individuelles caractéristiques des élèves, à savoir l'âge (15, 16, 17 et 18 ans), le genre (fille, garçon), la langue maternelle (français, autres). Ces variables, rappelons-le, ont souvent été évoquées dans plusieurs rapports de recherche réalisés au Québec (MEQ, 1992; Blais, 2003; McAndrew, Ledent et Ait-Said, 2005; Ouedraogo, 2006; Rapport du MELS, 2008) et ailleurs (Petrucci, 2010) pour expliquer des différences dans l'acquisition et la maîtrise du français écrit. Nous avons aussi relevé, avec les modélisations multiniveaux multivariées, qu'une très faible partie de ces différences est imputable à certaines variables caractéristiques de l'école, comme le réseau (public, privé), la proportion d'élèves de langue maternelle française, la taille et la mixité (fille, mixte). Les sections suivantes présentent la discussion des résultats des analyses descriptives et multiniveaux.

## **6.1 L'EXAMEN DU RÉSULTAT DE L'ANALYSE DESCRIPTIVE DES VARIABLES CARACTÉRISTIQUES DES ÉLÈVES**

Cette section présente une discussion du résultat de l'analyse descriptive de la performance des élèves selon les variables caractéristiques des élèves (l'âge, le genre, la langue maternelle).

### **6.1.1 La performance des élèves selon l'âge**

Les résultats de la présente recherche indiquent donc que l'âge est un facteur important dans la performance des élèves. Les élèves qui arrivent à l'âge normal en 5<sup>e</sup> secondaire, c'est-à-dire à 16 ans, représentent autour de 75 % de l'effectif total de chacune des trois cohortes. Ils réussissent mieux que ceux de 17 ans et de 18 ans qui représentent à peine 20 % de l'effectif. La différence de résultats est plus marquée entre les élèves de 16 ans et de 17 ans qu'entre ceux de 15 ans et de 16 ans. Donc, si le fait d'être en avance ne pose pas de problème, celui d'être en retard peut s'avérer très hasardeux. Rappelons que les données reçues du MELS comptent également des élèves qui sont âgés de 14 ans et même quelques-uns de 19 ans. Leurs résultats n'ont pas été analysés dans cette étude. Il est probable que les élèves âgés de 14 ans réussiraient mieux que ceux qui sont plus âgés et que les élèves de 19 ans obtiendraient de moins bonnes notes que ceux qui sont plus jeunes, selon la logique de la distribution des résultats que nous avons obtenus en fonction de l'âge.

On est porté à s'interroger sur la présence de plusieurs élèves de 14 ans et de 19 ans en 5<sup>e</sup> secondaire. Il semble que les élèves âgés de 14 ans montreraient des qualités exceptionnelles dans leur apprentissage dès l'école primaire. De ce fait, les responsables de certains établissements scolaires surtout privés, psychologues, conseillers pédagogiques et directeurs, favoriseraient leur passage dans des classes supérieures qui ne correspondent pas à leur âge selon le système scolaire québécois. À l'opposé, les élèves de 19 ans affichent un retard de trois ans par rapport à

l'âge normal des élèves en 5<sup>e</sup> secondaire. Leur apprentissage connaîtrait de réelles complications.

Les résultats obtenus sur les relations entre la performance des élèves et l'âge sont, nous l'avons reconnu, en accord avec les conclusions d'autres études qui ont été réalisées sur des notes moyennes reçues par des élèves à des épreuves standardisées menées au Québec (Blais, 2003; McAndrew et collab., 2005, Ouedraogo, 2006), ainsi qu'avec celle de Petrucci (2010) en Suisse. Blais (2003) avait montré pour l'ensemble du Québec, incluant les élèves de la région de Montréal qui représentaient à l'époque à peu près 20 % de l'effectif de sa population, que les élèves arrivant à l'âge normal en 5<sup>e</sup> secondaire réussissent mieux que ceux qui sont plus âgés. McAndrew et collab. (2005) pour leur part, en analysant les données relatives aux élèves de la région métropolitaine de Montréal, avaient indiqué que le fait de commencer les études secondaires à l'âge normal, c'est-à-dire à 12 ans, constitue un important facteur de prédiction de réussite des études secondaires pour les élèves issus des communautés noires. De son côté, Petrucci (2010) a montré que si le fait d'être en avance dans son cheminement scolaire n'entraîne aucune incidence particulière sur la performance des élèves dans une épreuve de compréhension à l'écrit à la fin de la scolarité obligatoire en Suisse, celui d'être en retard entraîne par contre un impact négatif sur la performance des élèves. Précisons que les élèves de 17 ans et de 18 ans représentent respectivement à peu près 35 % et 18 % dans la sous-population de ceux qui éprouvent beaucoup de difficulté en syntaxe et ponctuation, et autour de 33 % et 15 % en orthographe. Ces proportions se situent respectivement autour de 15 % en syntaxe et ponctuation et de 5 % en orthographe pour l'ensemble de la population.

Il se dégage une constante déterminante dans les résultats des trois cohortes d'élèves, celles des années 2006, 2007 et 2008 : l'écart le plus important entre les différents groupes d'âge apparaît entre les élèves de 16 ans et de

17 ans. Le fait pour un élève d'accuser un retard d'une année dans son cheminement scolaire, indépendamment du moment où il a pris ce retard durant son parcours, au primaire ou au secondaire, peut causer un handicap majeur pour sa réussite à l'examen de français écrit du MELS en 5<sup>e</sup> secondaire, par rapport aux deux critères de correction précités. Le retard en tant que tel ne constitue pas un problème d'apprentissage, mais les difficultés d'apprentissage qu'éprouvent ces jeunes de 17 et de 18 ans sont attribuables probablement à toutes sortes de circonstances défavorables qui les ont occasionnées.

### **6.1.2 La performance des élèves selon le genre**

En ce qui a trait au genre de l'élève, les résultats indiquent que les filles font en moyenne moins d'erreurs de syntaxe et de ponctuation et moins d'erreurs d'orthographe que les garçons durant les trois années à l'étude, soit 2006, 2007 et 2008. Ces résultats confirment également ceux obtenus par Blais (2003) qui avait relevé, pour l'ensemble du Québec, une différence appréciable entre la performance des garçons et celle des filles qui s'est maintenue durant sept ans, de 1994 à 2001, en faveur des filles.

L'explication des résultats supérieurs constants des filles comparativement à ceux des garçons n'a pas été approfondie dans les études menées jusqu'ici. Il semble que certaines jeunes filles auraient fréquenté dès leur entrée au primaire des écoles réservées aux filles qui maintiennent une tradition d'excellence dans l'enseignement de la langue française. Les filles des écoles secondaires seraient aussi plus responsables dans leurs études que les garçons qui s'adonneraient davantage aux loisirs, et particulièrement aux jeux vidéo (MEQ, 1992). On sait que la qualité de la langue parlée et écrite de ces jeux en français, qui ont été traduits de l'anglais, laisse souvent à désirer. Elle exercerait une influence négative sur l'apprentissage de la syntaxe, de la ponctuation et de l'orthographe par les garçons. L'explication

des jeux vidéo en est une parmi tant d'autres. La représentation des rôles sociaux joue probablement davantage chez les élèves.

Si dans l'ensemble de la population les filles représentent approximativement 54 % et les garçons 46 % des élèves, par contre dans la sous-population de ceux qui font 24 erreurs de syntaxe et ponctuation et 35 erreurs d'orthographe, les garçons sont généralement plus nombreux que les filles, surtout en orthographe. Il faudrait cependant apporter certaines nuances dans l'explication de la performance des élèves selon le genre. Nous avons croisé la variable genre de l'élève avec l'âge. Pour les trois cohortes étudiées, les filles réussissent mieux que les garçons en orthographe pour chaque groupe d'âge considéré. En revanche, en syntaxe et ponctuation, pour l'année 2006, les garçons de 17 ans et de 18 ans réussissent de manière équivalente aux filles de ces mêmes groupes d'âge ou mieux que celles-ci; pour l'année 2007, la performance des élèves en fonction du genre est équivalente pour les différents groupes d'âge et en 2008, les garçons réussissent mieux que les filles ou de manière équivalente. Comme on peut le constater, c'est la maîtrise de l'orthographe qui s'avère plus difficile pour les garçons que pour les filles.

### **6.1.3 La performance des élèves selon la langue maternelle**

En ce qui concerne la langue maternelle des élèves, rappelons qu'il existe plus d'une quinzaine de langues maternelles dans la région métropolitaine de Montréal. De façon générale, les résultats indiquent que les élèves de langue maternelle française font en moyenne moins d'erreurs en syntaxe et ponctuation et en orthographe que les élèves de langues maternelles différentes du français. Dans la mesure où l'on peut considérer ces élèves comme des ressortissants de l'immigration, les résultats de la présente recherche s'accordent avec ceux plusieurs études (Coleman et collab., 1966; McAndrew et collab., 2005; Petrucci, 2010). Ceux-ci ont montré que la

performance des élèves dont la langue maternelle est celle de l'enseignement est supérieure à la performance de ceux qui sont issus de l'immigration et dont la langue maternelle est différente de la langue d'enseignement. Les élèves de langue maternelle française représentent 60 % et les autres 40 % de l'effectif de l'ensemble de la population. Dans la sous-population d'élèves en difficulté, les élèves de langue maternelle française représentent autour de 45 % de l'effectif en syntaxe et ponctuation et 57 % de l'effectif en orthographe; les élèves de langues maternelles différentes du français représentent à peu près 55 % de l'effectif en syntaxe et ponctuation et 43 % en orthographe.

Lorsque nous avons croisé la variable la *langue maternelle* des élèves à celle de l'âge, les résultats ont indiqué qu'en 2007 et 2008, les élèves de langue maternelle française ont obtenu de meilleurs résultats en syntaxe et ponctuation que ceux de langues maternelles différentes du français pour chaque groupe d'âge considéré dans la recherche. En revanche, en orthographe, les élèves de 16, 17 et 18 ans de langues maternelles différentes du français réussissent mieux que ceux de langue maternelle française ou de manière équivalente. Ce résultat nous porte à supposer que les élèves de langues maternelles différentes du français mémorisent plus facilement l'orthographe des mots qu'ils ne peuvent construire des phrases en respectant les règles propres à la syntaxe et à la ponctuation en français. Les caractéristiques structurelles de leurs langues maternelles peuvent être très différentes de celles de la langue française. D'où la difficulté pour eux à composer des phrases qui respectent rigoureusement les règles syntaxiques et celles de la ponctuation du français, et à accéder au niveau de maîtrise des élèves qui s'expriment naturellement en français. Il est probable que la relative facilité des allophones à acquérir l'orthographe tient aussi au fait qu'il s'agit principalement de connaissances déclaratives, même si l'on tient

compte du fait que l'orthographe lexicale et l'orthographe grammaticale sont fondues dans le même critère.

En outre, il aurait été utile de déterminer depuis combien de temps les élèves de langues maternelles autres que le français fréquentent l'école québécoise. Cette information nous aurait permis d'apprécier leur rythme d'apprentissage et d'appropriation de la langue française. Et de tenter de vérifier aussi l'efficacité de la pratique d'enseignement du français dans les écoles de la région métropolitaine de Montréal.

## **6.2 L'EXAMEN DU RÉSULTAT DE L'ANALYSE DESCRIPTIVE DES VARIABLES CARACTÉRISTIQUES DES ÉCOLES**

Cette section fait la discussion du résultat de l'analyse descriptive des variables caractéristiques des écoles (le secteur, la mixité, la taille, la proportion des élèves de langue maternelle française) selon la performance des élèves.

### **6.2.1 La performance des élèves selon le secteur de l'école**

Les résultats indiquent qu'il existe une différence majeure dans la performance des élèves selon le secteur d'appartenance de l'école, public ou privé. Les élèves qui fréquentent les établissements d'enseignement privés de la région de Montréal obtiennent de meilleurs résultats aux deux critères de correction que ceux des écoles publiques. Les élèves de 18 ans du secteur public ont cependant obtenu en 2007 de meilleurs résultats que ceux du même âge du secteur privé, tant en syntaxe et ponctuation qu'en orthographe. Ainsi, le secteur ne serait pas dans tous les cas un facteur déterminant dans la performance des élèves plus âgés.

La meilleure performance des élèves du secteur privé peut s'expliquer par la sélection qu'opèrent la plupart des écoles de ce secteur dès l'entrée au

secondaire. Celles-ci ne sont pas obligées de scolariser les élèves en difficulté, c'est-à-dire de les garder jusqu'à la fin de leurs études secondaires. À ce sujet, Rumberger et Palardy (2005) considèrent le fait d'amener les élèves en difficulté jusqu'à la fin de leurs études secondaires comme un critère d'efficacité de l'école. Même si la performance des élèves du secteur privé dépasse celle du secteur public, il est intéressant de noter que le secteur public scolarise beaucoup plus d'élèves en difficulté et les amène en 5<sup>e</sup> secondaire à passer l'épreuve unique de français écrit du MELS. Blais (2003) faisait en effet remarquer que le poids des élèves de 17 ans et de 18 ans, les élèves en retard, est toujours beaucoup plus important dans le réseau public que dans le réseau privé. Pour l'ensemble des élèves des trois cohortes, soit celles des années 2006, 2007 et 2008, les résultats de notre recherche indiquent que le pourcentage des élèves en retard, garçons et filles, correspond respectivement à 33,50 % et 27,59 % du réseau public. Dans le réseau privé, les élèves en retard représentent 8,54 % de garçons et 6,79 % de filles. Ces données expliquent en partie les performances moyennes et médianes plus faibles des élèves du secteur public, compte tenu des difficultés d'apprentissage qu'éprouvent les élèves plus âgés qui représentent un poids relatif plus important au secteur public comparativement au secteur privé. Mentionnons qu'on ne dispose pas de données précises sur la sélection des élèves. Il ne nous a pas été possible de comparer les écoles en fonction de la sélection des élèves à l'entrée de l'ordre d'enseignement secondaire. Une hypothèse serait qu'il n'y a pas de différence entre les résultats des élèves du secteur public et ceux du secteur privé lorsqu'on compare les écoles qui sélectionnent les élèves, par exemple les écoles du secteur public qui offrent le programme d'éducation internationale et celles du secteur privé.

### **6.2.2 La performance des élèves selon la mixité de l'école**

La mixité de l'école est une variable dérivée, c'est-à-dire qu'elle a été calculée à partir des informations reçues du ministère de l'Éducation, du Loisir et du Sport (MELS). Elle est reliée à la variable individuelle genre de l'élève. Comme nous l'avons noté, elle n'avait pas été analysée dans les études réalisées antérieurement au Québec. Les résultats de notre recherche indiquent que les filles qui fréquentent des écoles réservées aux filles, quoique peu nombreuses, avec moins de 10 % de l'effectif total de chacune des cohortes, font en moyenne moins d'erreurs de syntaxe et de ponctuation et moins d'erreurs d'orthographe que les élèves qui fréquentent les écoles mixtes.

Nous sommes cependant porté à user de prudence quant à l'interprétation de ce dernier résultat, à cause du très petit nombre de filles qui fréquentent des écoles réservées aux filles dans la région de Montréal par rapport à l'ensemble des élèves de chaque cohorte. Dans l'interprétation de la relation entre la performance des élèves et le genre, nous avons par ailleurs fourni quelques éléments de réponses qui expliqueraient pourquoi les filles obtiendraient de meilleurs résultats que les garçons. Une interprétation qui vaut encore dans l'explication de la relation entre la mixité de l'école et la performance des élèves.

### **6.2.3 La performance des élèves selon la taille**

D'après nos résultats, la taille des écoles ne serait pas un facteur déterminant dans la performance des élèves. On observe toutefois que les élèves qui fréquentent les écoles de grande taille, celles comptant un effectif de plus de 1200 élèves, obtiennent de meilleurs résultats que ceux qui étudient dans des écoles de petite taille ou de taille moyenne, c'est-à-dire des écoles qui accueillent entre 105 et 1199 élèves. Blais (2003) avait déjà souligné qu'il fallait faire montre de prudence dans l'analyse de la relation entre la taille

d'une école et les résultats de ses élèves, à cause de la vocation particulière de certaines écoles de petite taille qui pourraient présenter des résultats plus faibles que la moyenne générale. Ces écoles, pour la plupart, reçoivent des élèves de 4<sup>e</sup> et de 5<sup>e</sup> secondaire qui viennent terminer leur diplôme d'études secondaires, comme c'est le cas des écoles destinées à la formation générale des jeunes adultes. Elles reçoivent des élèves plus âgés que ceux qui arrivent à l'âge normal en 5<sup>e</sup> secondaire. Les élèves de ces écoles éprouvent beaucoup de difficulté dans leur apprentissage en général, et particulièrement dans les cours de français.

#### **6.2.4 La performance des élèves selon la proportion des élèves de langue maternelle française**

La proportion des élèves de l'école de langue maternelle française est aussi une variable dérivée. On peut constater un léger avantage quant à la performance des élèves inscrits dans les écoles dont la proportion d'élèves de langue maternelle française est élevée, c'est-à-dire les écoles dont plus de 75 % de l'effectif est de langue maternelle française, comparativement aux écoles dont la proportion d'élèves de langue maternelle française est inférieure à ce pourcentage. Les explications relatives à la relation entre la langue maternelle et la performance de l'élève sont également valides pour ce qui concerne la relation entre la proportion des élèves de l'école de langue maternelle française et la performance des élèves.

### **6.3 L'EXAMEN DU RÉSULTAT DES ANALYSES MULTINIVEAUX**

Nous avons poursuivi notre examen des relations entre les variables caractéristiques des élèves et des écoles et la performance des élèves en effectuant des analyses multiniveaux multivariées en vue de mesurer l'ampleur et la différence de variabilité dans les résultats obtenus par les

élèves selon les deux critères spécifiques de correction, la syntaxe et la ponctuation ainsi que l'orthographe, pour les années 2006, 2007 et 2008.

### **6.3.1 La modélisation multiniveaux multivariée à vide**

Nos données indiquent pour l'ensemble des trois cohortes considérées que la part de variance totale attribuable à l'élève est de l'ordre de 75 %, et cela selon les deux critères spécifiques de correction, la syntaxe et la ponctuation ainsi que l'orthographe. Celle qui est attribuable à l'école est de l'ordre de 25 %. Pour la sous-population d'élèves sans les valeurs de 24 erreurs en syntaxe et ponctuation et de 35 erreurs en orthographe, la part de variance totale attribuable à l'élève est de l'ordre de 83 %, et de l'ordre de 17 % à l'école.

Un premier constat est que l'absence du nombre élevé d'erreurs des élèves dans la sous-population neutralise l'influence de l'école dans la performance des élèves. En effet, la part de variance totale de l'école dans la performance des élèves passe de 25 % pour l'ensemble de la population à 17 % pour la sous-population d'élèves sans le nombre élevé d'erreurs. Le rôle de l'école serait donc plus important dans l'apprentissage du français des élèves en difficulté.

Les résultats de la présente recherche, il importe encore de répéter, concordent avec ceux des études précédentes menées au Québec. L'étude de Blais (2003) avait montré, à partir d'une population qui présentait les caractéristiques d'une distribution normale, plus proche de la distribution de la sous-population sans les valeurs élevées des erreurs des élèves, que la variance totale était attribuable à près de 17 % au niveau de l'école et à près de 83 % au niveau de l'élève. Pour sa part, Ouedraogo (2006), de façon plus spécifique, avait trouvé qu'entre 22 % et 26 % de la variance totale de la performance des élèves de 5<sup>e</sup> secondaire à l'épreuve unique de français écrit du MELS dans la région de Montréal se situent au niveau de l'école. Des

résultats qui se rapprochent de ceux de l'ensemble des trois cohortes que nous avons étudiées. Les résultats de notre recherche concordent également avec ceux d'autres études menées ailleurs qu'au Québec. Petrucci (2010) avait trouvé que le niveau école contribue dans une mesure de 10,20 % au primaire, de 10,00 % au regroupement A et de 17,8 % au regroupement B. Également en Suisse, le rôle de l'école est aussi plus important chez les élèves aux habiletés ordinaires que chez ceux qui sont plus forts. Le niveau élève contribue pour sa part dans une mesure de 89,80 % au niveau primaire, de 90 % au regroupement A et de 82,20 % au regroupement B. Konstantopoulos (2006), pour sa part, a trouvé que la part de variance totale attribuable à l'école était autour de 10% et autour de 90% attribuable à l'élève.

### **6.3.2 La modélisation des variables caractéristiques des élèves**

En ce qui se rapporte aux variables individuelles caractéristiques des élèves, nos résultats concordent, avec certaines nuances, avec ceux des auteurs qui ont modélisé celles-ci en fonction d'une note moyenne et sur des populations différentes. Blais (2003) a montré que le facteur d'âge était plus important que les autres sur la performance des élèves dans une épreuve unique du MELS. Ouedraogo (2006) a lui aussi trouvé que la variable âge apportait plus d'informations que le genre dans la performance d'un échantillon d'élèves de la région métropolitaine de Montréal. Pour l'ensemble des trois cohortes, celles des années 2006, 2007 et 2008, l'analyse de la variabilité de la performance des élèves en syntaxe et ponctuation révèle que l'effet le plus important est relié à la langue maternelle de l'élève, avec un coefficient de 2,08 (0,07) en faveur des élèves de langue maternelle française, suivi de l'âge avec un coefficient de 1,78 (0,05) en faveur des élèves les plus jeunes, et du genre avec un coefficient de 0,34 (0,07) en faveur des filles. Ces résultats confirment ceux que nous avons présentés dans la première partie de nos analyses descriptives.

De plus, quant à la quantité d'informations apportée par les variables caractéristiques des élèves, signalons que l'ordre d'importance de ces variables sur la performance des élèves en orthographe ne suit pas celle que l'on peut observer en syntaxe et ponctuation. Les résultats indiquent que le genre semble valoir le poids relatif le plus important avec un coefficient de 3,04 (0,10) en faveur des filles, suivi de l'âge avec un coefficient de 2,30 (0,08) en faveur des élèves les plus jeunes, et de la langue maternelle avec un coefficient de 0,71 (0,11) en faveur des élèves de langues maternelles autres que le français.

Les variables individuelles caractéristiques des élèves (âge, genre, langue maternelle), qui montrent un poids relatif de 13,90 %, contribuent davantage dans l'explication des résultats des élèves en syntaxe et ponctuation qu'en orthographe, critère pour lequel ces variables indépendantes indiquent un poids relatif de 9,80 %.

Ces résultats concordent eux aussi avec ceux que nous avons exposés précédemment. Les élèves de langues maternelles différentes du français auraient obtenu des résultats légèrement supérieurs en orthographe à ceux attribués aux élèves de langue maternelle française. Rappelons que des considérations à propos de la langue maternelle des élèves ont été fournies à la présentation de la première série d'analyses descriptives sur la relation entre cette variable et la performance des élèves.

### **6.3.3 La modélisation des variables caractéristiques des écoles**

Au sujet des variables caractéristiques de l'école, les résultats indiquent que la variable réseau représente le poids relatif le plus important avec un coefficient de 5,63 (0,60) en orthographe et 2,20 (0,37) en syntaxe et ponctuation en faveur des écoles privées, suivi du pourcentage des élèves de l'école de langue maternelle française avec un coefficient de 1,46 (0,70) en syntaxe et ponctuation et 0,43 (1,14) en orthographe, en faveur des écoles

qui comptent un nombre élevé d'élèves de langue maternelle française. Les variables mixité et taille de l'école ne semblent pas présenter des relations significatives avec la performance des élèves. La valeur de leur coefficient dans la modélisation est plutôt négligeable. Par contre, Ouedraogo (2006) rapportait que la variable taille de l'école comptait pour une mesure de 4,2 %.

Les variables caractéristiques des écoles (réseau, taille, mixité, proportion d'élèves de l'école de langue maternelle française), avec un poids relatif de 6,40 % en orthographe, jouent davantage dans l'explication des résultats des élèves qu'en syntaxe et ponctuation, critère pour lequel on obtient un pourcentage plutôt faible de 3,65 %.

Si l'on prend en considération les sous-populations sans les valeurs de 24 erreurs en syntaxe et ponctuation et de 35 erreurs en orthographe, les variables caractéristiques du niveau école (réseau, mixité, taille, proportion d'élèves de langue maternelle française) entrent en jeu à hauteur de 2,50 % des 17,60 % en syntaxe et ponctuation, et en orthographe à hauteur de 5,40 % des 17,70 % dans la variabilité de la performance des élèves. Pour les sous-populations d'élèves sans le nombre d'erreurs élevé, l'effet des facteurs d'école, nous l'avons déjà noté, diminue par comparaison à celui de l'ensemble de la population qui compte les élèves de tous les niveaux de compétences.

D'après ces résultats, les écoles semblent maintenir entre elles une plus grande homogénéité que les élèves à l'intérieur d'une école dans la région métropolitaine de Montréal. Cependant, qu'il s'agisse de l'ensemble des trois cohortes ou des trois sous-populations sans les valeurs de 24 erreurs en syntaxe et ponctuation et de 35 erreurs en orthographe, certains facteurs caractéristiques de l'école comme la taille et la mixité ne semblent pas entraîner un effet significatif sur la performance des élèves pour les deux critères de correction retenus.

## **CHAPITRE VII**

### **CONCLUSION**

#### **7.1 UN RAPPEL DES OBJECTIFS ET DE LA MÉTHODE DE LA RECHERCHE**

L'étude que nous avons effectuée s'inscrit dans le courant de recherche sur l'école « efficace » (*The School Effectiveness Research*). Elle avait pour but de déterminer les relations entre, d'une part, les variables indépendantes caractéristiques des élèves, comme l'âge (15, 16, 17 et 18 ans), le genre (fille, garçon) et la langue maternelle (français, autres); de l'école, comme le secteur (public ou privé), la mixité (mixte, fille), la proportion des élèves de langue maternelle française, la taille et, d'autre part, les deux variables dépendantes (la syntaxe et la ponctuation, l'orthographe) afin de vérifier dans quelle mesure les résultats obtenus par les élèves finissants du secondaire de la région de Montréal, selon ces deux critères spécifiques de correction à l'épreuve unique de français écrit du MELS, dépendent de certaines caractéristiques individuelles des élèves, mais aussi de certaines caractéristiques de l'établissement dans lequel ils sont scolarisés, et cela, sur une période de trois ans, soit 2006, 2007 et 2008.

La poursuite de ces objectifs nous conduit à fixer les étapes d'analyse des données de la présente recherche. L'étude de ces données s'est déroulée en deux étapes.

D'abord, dans la première étape, nous avons effectué trois séries d'analyses descriptives des variables dépendantes (syntaxe et ponctuation, orthographe) et de leur lien avec les variables indépendantes caractéristiques de l'élève et de l'école pour chacune des trois cohortes, celles des années 2006, 2007 et 2008. La première série d'analyses a été menée sur l'ensemble des données de chaque cohorte. La deuxième série d'analyses a été réalisée sur une première sous-population d'élèves sans les mesures de 24 erreurs en syntaxe et ponctuation et de 35 erreurs en orthographe. La troisième série d'analyses a été exécutée sur une deuxième sous-population, celle des élèves en difficulté, c'est-à-dire ceux qui font 24 erreurs en syntaxe et ponctuation et 35 erreurs en orthographe.

Dans une deuxième étape, nous avons effectué des analyses multiniveaux multivariées. Pour les élèves de chacune des trois cohortes et des trois sous-populations sans les valeurs de 24 erreurs de syntaxe et ponctuation et de 35 erreurs en orthographe, nous avons appliqué plusieurs modélisations multiniveaux multivariées à vide en ajoutant au fur et à mesure un niveau (élève, école, année) à la structure des données. Nous avons fait ensuite des modélisations multiniveaux multivariées qui intègrent dans le modèle à coefficients aléatoires retenu à deux niveaux les variables caractéristiques de l'élève et ensuite de l'école.

La non-disponibilité des données relatives au personnel enseignant ainsi que la difficulté de conceptualisation de plusieurs variables caractéristiques de l'environnement scolaire ont cependant imposé des limites à notre démarche d'analyse.

## **7.2 LES LIMITES DE LA RECHERCHE**

Nous sommes conscient du fait qu'il existe un grand nombre de variables que l'on pourrait étudier aux niveaux de l'élève et de l'école. Mais nous nous sommes intéressé à celles qui nous paraissent les plus pertinentes à notre problématique relativement à l'élève (l'âge, le genre, la langue maternelle) et à l'école (le secteur, la taille, le nombre d'élèves présents à l'épreuve) et pour lesquelles nous avons obtenu des données du Ministère. Nous avons aussi retenu des variables relatives à l'école (la mixité, la proportion des élèves de l'école de langue maternelle française) pour lesquelles nous avons déterminé les mesures selon les données disponibles. Nous reconnaissons que certaines variables relatives à l'environnement scolaire, de prime abord importantes, mais difficiles à conceptualiser et à mesurer, comme le climat de l'école, le leadership de la direction, l'implication du personnel, l'adhésion du personnel au projet éducatif de l'école ainsi que les variables relatives au personnel enseignant et/ou à la salle de classe n'ont toutefois pas été analysées dans notre recherche.

### **7.2.1 Le choix des variables traitées et le nombre de niveaux appliqué**

L'une des limites de cette étude réside dans l'absence des données relatives au personnel enseignant. Cet inconvénient nous a contraints à ne pas modéliser les variables caractéristiques du niveau classe. Nous reconnaissons que le fait d'ignorer un niveau intermédiaire dans une modélisation, comme la salle de classe, peut amener à surestimer l'importance de l'effet des variables des autres niveaux, comme les facteurs d'école ou les variables caractéristiques des élèves sur le rendement scolaire des élèves (Opdenakker et Van Damme, 2000).

Nous n'avons pas modélisé non plus les pratiques pédagogiques des enseignants de français de 5<sup>e</sup> secondaire. Dans les modèles quantitatifs, il est plus aisé de décrire un enseignant suivant des traits facilement

observables, des données tangibles comme l'âge, la scolarité, la formation, l'expérience, le degré d'exigence, etc., que de définir des pratiques pédagogiques plus informelles et moins nettes, comme les interactions entre les enseignants et les élèves (Petrucci, 2010). D'autre part, les auteurs De Ketele (2009) ainsi que Duru-bellat et Leroy-Audoïn (1990) signalent que les pratiques pédagogiques sont trop souvent définies sur la base de données secondaires, collectées à d'autres fins, qui sont le fruit de déclarations des enseignants eux-mêmes et qui n'ont pas été observées directement. Notre étude s'est donc limitée à la modélisation d'un nombre bien défini de variables quantitatives pour les deux niveaux (élève, école).

### **7.2.2 Des relations dont l'interprétation reste à préciser**

D'autre part, comme cela a été le cas dans les recherches de Blais (2003), d'Ouedraodo (2006) et de Petrucci (2010), l'analyse du pouvoir explicatif des modèles multiniveaux multivariés mis en œuvre dans la présente recherche a révélé qu'avec les informations dont nous disposons, soit les données sur les trois variables individuelles caractéristiques de l'élève (l'âge, le genre, la langue maternelle) et les quatre variables caractéristiques de l'école (le secteur, la mixité, la taille, la proportion des élèves de l'école de langue maternelle française), la plus grande partie du phénomène, celle qui a trait aux relations entre les variables du contexte et la performance des élèves par rapport aux deux critères de correction à l'épreuve unique du MELS, reste tout compte fait inexplicée.

Pour l'ensemble des trois cohortes retenues, celles des années 2006, 2007 et 2008, les trois variables caractéristiques de l'élève qui ont été modélisées, c'est-à-dire l'âge, le genre et la langue maternelle, expliquent à peine 10,48 % de la variance totale de la performance des élèves en syntaxe et ponctuation, et 7,35 % de la variance totale des résultats obtenus par les élèves en orthographe. Les variables caractéristiques des écoles, c'est-à-dire

le réseau, la mixité, la taille, la proportion d'élèves de l'école de langue maternelle française, expliquent très peu de la variance totale dans la performance des élèves, à peine 0,88 % pour la syntaxe et la ponctuation et 1,61 % pour l'orthographe. En plus des trois variables caractéristiques des élèves, nous avons modélisé quatre variables relatives aux écoles, mais nous n'avons retenu aucune variable relative au personnel enseignant et/ou à la salle de classe. La part de variance expliquée dans la présente recherche est du même ordre de grandeur que celle obtenue dans la plupart des études (Blais, 2003; Konstantopoulos, 2005; Petrucci, 2010) réalisées au Québec et ailleurs.

### **7.2.3 Le concept du déterminisme scolaire**

Nous reconnaissons aussi que la relation entre les résultats des élèves par rapport aux deux critères de correction retenus, la syntaxe et la ponctuation ainsi que l'orthographe et les variables individuelles caractéristiques des élèves de même que les variables caractéristiques des écoles considérées dans cette recherche, n'est pas déterministe. C'est que nous n'avons pas fait la démonstration d'une relation de cause à effet, ce qui d'ailleurs n'était pas l'objectif de la présente recherche. Comme le précise Blais (2003), il est difficile d'associer dans une relation de cause à effet l'école secondaire fréquentée au moment de la passation de l'épreuve unique et le résultat obtenu en français écrit qui est en fait l'aboutissement d'un apprentissage étalé sur douze ans marqué par des changements de tous ordres, comme la mobilité des élèves et du personnel enseignant. Les résultats obtenus par les élèves d'une école peuvent fluctuer d'une année à l'autre sans nécessairement indiquer en même temps une différence significative quant aux compétences collectives des élèves à l'écrit. Rappelons à ce sujet que notre étude a analysé des données sur trois cohortes d'élèves qui sont identifiées à trois populations différentes. Nous avons étudié ces trois

populations d'élèves, qui se ressemblent certes, mais dont les sujets n'ont pas passé le même test. Ce qui nous a amené à les traiter séparément.

Les chercheurs dont les travaux traitent de l'« efficacité » des écoles se préoccupent de la mobilité des élèves et du personnel enseignant d'une institution scolaire à une autre. Ainsi, certains auteurs dont Meyer (2002) et Keeves, Hungi et Afrassa (2005) constatent que les élèves qui changent d'école sont, le plus souvent, les plus faibles. Rumberger et Palardy (2005), de leur côté, remettent en question la décision de certaines écoles de transférer presque systématiquement les élèves faibles dans d'autres établissements scolaires en ne conservant que les plus forts. Selon eux, ce geste aurait pour conséquence de hausser le résultat moyen des élèves d'une école et de biaiser du même coup l'évaluation des établissements scolaires. Ils considèrent que la capacité d'une école à mener ses élèves à la diplomation doit être considérée comme un indicateur de son efficacité, tout en tenant compte du déménagement volontaire des familles. À cet égard, les écoles du secteur public de la région métropolitaine de Montréal qui amènent quatre fois plus d'élèves en retard à l'épreuve unique de français écrit de 5<sup>e</sup> secondaire du MELS seraient considérées plus efficaces que celles du secteur privé.

#### **7.2.4 Les résultats des élèves et les facteurs liés à la stabilité scolaire**

La question de la rémanence des effets d'établissements est liée à celle de la mobilité des élèves et du personnel enseignant. Elle n'a pas été prise en considération dans notre étude. Il aurait sans doute été utile de vérifier si les élèves finissants du secondaire qui se sont présentés à l'épreuve unique de français écrit du MELS ont effectivement fréquenté le même établissement scolaire durant tout le cycle secondaire, à savoir cinq années. À ce propos, nous n'avons pas pu obtenir non plus d'informations ni sur le secteur

d'appartenance (privé, public) de l'école, ni sur la mixité de l'école (réservée aux filles, mixte) que les élèves ont fréquentée au primaire.

Toujours au sujet de la mobilité des élèves et de la rémanence des effets d'établissement, les auteurs Bressoux, Leroy-Audouin et Coustère (1998) sont d'avis que les acquis mesurés à un temps  $t$  ne sont pas uniquement fonction des conditions de scolarité de ce temps. Ces acquis peuvent être également le fruit de conditions antérieures dont les effets peuvent être plus ou moins durables et plus ou moins marqués. Wiggan (2007) de son côté fait remarquer l'importance de la stabilité du personnel enseignant des écoles qui reçoivent les élèves faibles dans les quartiers défavorisés. Il juge que « le caractère transitoire des enseignants dans les écoles des quartiers pauvres des régions urbaines où les élèves auraient besoin d'un personnel enseignant compétent et stable affecte de manière négative leur rendement scolaire ».

### **7.2.5 L'importance des notes initiales de l'élève**

Plusieurs auteurs, dont Goldstein, Huiqi, Rath et Hill (2000), Meyer (2002), Felouzis (2004) et Bressoux (2007), considèrent explicitement le résultat initial de l'élève à l'entrée du cycle secondaire comme étant un facteur de prédiction important de ses résultats futurs. La majorité des écoles privées et certaines écoles publiques à vocation particulière de la région métropolitaine de Montréal recrutent leurs élèves selon leurs résultats scolaires à la demande d'inscription, souvent en complémentarité avec leur résultat à des séances d'examen de sélection dans les matières de base comme les mathématiques, le français et les connaissances générales. Tout en admettant l'importance du caractère prédictif du résultat initial de l'élève dans son rendement scolaire futur (McAndrew et collab., 2005), nous déplorons qu'il ne nous ait pas été possible de recueillir des données sur celui-ci. Le Comité de travail pour mieux soutenir le développement de la compétence à

écrire (MELS, 2008) reconnaît que l'information sur le profil de l'élève à l'entrée du secondaire faciliterait la planification de son apprentissage. Ce Comité recommande d'ailleurs aux commissions scolaires d'établir des mécanismes permettant le transfert de l'information sur le développement des compétences en écriture de l'élève au moment de son passage de l'ordre primaire à l'ordre secondaire, afin de faciliter une prise en charge différentielle des élèves dès le début du secondaire. L'absence de ces renseignements pose une autre limite à la présente recherche, parce que nous ne sommes pas ainsi en mesure d'affirmer que les effets des variables caractéristiques de l'école introduites dans la modélisation correspondent à l'apport réel de celle-ci dans les résultats des élèves.

### **7.3 UN SOMMAIRE DE L'ÉTUDE**

Tout en ayant fait état de ces réelles limites de notre recherche, nous estimons que celle-ci apporte une utile contribution à la clarification de cette complexe problématique de l'apprentissage du français écrit au Québec. L'étude a en effet mis en relief les importants points suivants :

- les élèves de langue maternelle française obtiennent de meilleurs résultats en syntaxe et ponctuation que ceux d'autres langues maternelles;
- les filles réussissent mieux que les garçons tant en syntaxe et ponctuation qu'en orthographe;
- les élèves qui arrivent à l'âge normal en 5<sup>e</sup> secondaire réussissent mieux que ceux qui accusent du retard dans leur apprentissage;

- la distance médiane la plus importante dans la performance des élèves pour une différence d'une année est entre les élèves de 16 ans et de 17 ans, tant en syntaxe et ponctuation qu'en orthographe;
- les élèves du secteur privé obtiennent en général de meilleurs résultats que ceux du secteur public en syntaxe et ponctuation et en orthographe;
- selon les années, certains élèves de 17 ans et de 18 ans du secteur public obtiennent de meilleurs résultats que ceux de leur groupe d'âge du secteur privé, particulièrement en orthographe;
- le secteur public présente quatre fois plus d'élèves en retard à l'épreuve unique du MELS que le secteur privé;
- l'effectif de la population en 2007 est environ de 7,8 % plus important qu'en 2008 et de 8 % en 2006;
- la variabilité dans la performance des élèves entre les années n'est pas significative;
- pour l'ensemble de la population des trois cohortes, la part de variance totale attribuable à l'élève est de l'ordre de 75 % et de l'ordre de 25 % à l'école;
- pour les sous-populations d'élèves sans le nombre élevé d'erreurs, la part de variance totale attribuable à l'élève est de l'ordre de 83 % et de l'ordre de 17 % à l'école;
- les écoles sont plus homogènes que le sont les élèves à l'intérieur d'une école;

- le rôle de l'école serait plus important dans l'apprentissage du français des élèves en difficulté;
- les variables caractéristiques des élèves contribuent davantage dans l'explication de la performance des élèves en syntaxe et ponctuation qu'en orthographe;
- les variables caractéristiques des écoles contribuent davantage dans l'explication de la performance des élèves en orthographe qu'en syntaxe et ponctuation;
- En syntaxe et ponctuation, l'effet le plus important est relié à la langue maternelle de l'élève, avec un coefficient de 2,08 (0,07) en faveur des élèves de langue maternelle française, suivi de l'âge avec un coefficient de 1,78 (0,05) en faveur des élèves les plus jeunes, et du genre avec un coefficient de 0,34 (0,07) en faveur des filles;
- En orthographe, le genre affiche le poids relatif le plus important avec un coefficient de 3,04 (0,11) en faveur des filles, suivi de l'âge avec un coefficient de 2,30 (0,09) en faveur des élèves les plus jeunes, et de la langue maternelle avec un coefficient de 0,71 (0,11) en faveur des élèves de langues maternelles différentes du français;
- la mixité et la taille de l'école sont des facteurs peu significatifs dans la performance des élèves.

Au demeurant, d'autres explications sur la performance des élèves finissants du secondaire par rapport aux deux critères de correction à l'épreuve unique de français écrit du Ministère doivent être recherchées non seulement aux deux niveaux de la structure éducative analysée (élève, école), mais aussi au

niveau de la classe, en considérant les pratiques pédagogiques des enseignants. En vue d'aider les jeunes à améliorer leur apprentissage du français, nos propositions rejoignent celles du Comité d'experts (MELS, 2008) qui recommande que le Ministère s'assure que le régime pédagogique du secondaire garantit un minimum de deux cents heures d'enseignement du français par année, durant les trois premières années du secondaire; de mettre en place des conditions qui facilitent l'accompagnement dans l'apprentissage de l'écriture des élèves allophones, des élèves à besoins particuliers, telle la réduction du nombre d'élèves par classe.

Les informations relatives à la classe et au personnel enseignant constituent un niveau intermédiaire qu'il serait important d'explorer dans la recherche d'explications sur la performance des élèves en français écrit. Même s'il est vrai que nous ne disposons pas de beaucoup d'informations sur les écoles, il y a lieu de poursuivre et d'approfondir les travaux en vue d'expliquer pourquoi les élèves de langue maternelle française obtiennent de meilleurs résultats en syntaxe et ponctuation que ceux de langues maternelles différentes de la langue d'enseignement, pourquoi la maîtrise de l'orthographe demeure pour eux plus difficile que celle de la syntaxe et de la ponctuation, pourquoi les filles obtiennent systématiquement de meilleurs résultats que les garçons, pourquoi les élèves plus jeunes obtiennent de meilleurs résultats que les plus âgés, et surtout pourquoi la différence de performance la plus importante est entre les élèves de 16 et de 17 ans. Il faut espérer que d'autres études continueront d'enrichir le sujet et ouvriront des voies inédites aux recherches sur la place du français dans la formation des jeunes du Québec.

## RÉFÉRENCES

- Ainsworth, M. E. et Batten, E. J. (1974). The Effects of Environmental Factors on Secondary Educational Attainment in Manchester : A Plowden Follow up. Dans C. Teddlie et D. Reynolds (dir.), *The International Handbook of School Effectiveness Research*. London, Grande Bretagne : Routledge Falmer Press.
- Association de cadres scolaires du Québec (2002). *Avis de l'Association des cadres scolaires du Québec au ministre de l'Éducation sur le projet de loi 124* (Dossier 02-05). Québec, Québec.
- Blais, J.-G. (2003). *Études des différences entre les écoles secondaires du Québec quant aux résultats de leurs élèves à certaines épreuves du ministère de l'Éducation de la fin du secondaire*. Rapport de Recherche. CRIFPE-LABRIPROF. Montréal, Québec. Faculté des sciences de l'éducation, Université de Montréal.
- Blais, J.-G., Laurier, M. et Pelletier, G. (2001). Les indicateurs comme outil de régulation en éducation : Entre culture de chercheur et culture de décideur. Dans G. Pelletier (dir.), *Autonomie et évaluation en éducation : entre projet et évaluation* (pp. 131-147). Montréal, Québec : Éditions de l'AFIDES.
- Bock, R. D. et Wolfe, R. (1996). *A review and analysis of the Tennessee Value-Added Assessment System. Part I: Audit and review of the Tennessee Value-Added Assessment System (TVAAS) : Rapport final*. Nashville, TN: Comptroller of the Treasury.

- Bottani, R. et Tuijnman, A. (1994). The Design of Indicator Systems. Dans A. C Tuijnman et T. N. Postelethwaite, (dir.), *Monitoring The Standards of Education* (pp. 44-47). Oxford :Pergamon Press.
- Bowles, S. et Levin, H. M. (1968). The Determinants of scholastic Achievement : An Appraisal of some recent Evidence. *The journal of human resources*, 3(1), 3-24.
- Bradley, W. J. et Schaefer, K. C. (1998). *The Uses and Misuses of Data and Models: The Mathematization of the Human Sciences*. California, USA : SAGE Publications.
- Brassard, A. (2005). *Regards multiples sur l'évolution de la centralisation-décentralisation dans le système d'enseignement fondamental du Québec depuis 1959*. Leçon inaugurale du Département d'Administration et fondements de l'éducation. Montréal, Québec : Université de Montréal.
- Bressoux, P., Leroy-Audouin, C. et Coustère, P. (1998). Les extensions des modèles multiniveaux et leur application pour l'évaluation en éducation. *Mesure et évaluation en éducation*, 21(1), 35-59.
- Bressoux, P. (2008). *Modélisation statistique appliquée aux sciences sociales : Méthodes en sciences humaines*. Belgique : Editions De boeck Université.
- Brookover, W. B., Beady, C., Flood, P., Schweitzer, J. et Wisenbaker, J. (1979). *School Social Systems and Student Achievement : School Can Make a Difference*. New-York, New York :Praeger.

- Cain, G. C. et Watts, H. W. (1970). Problems in making policy Interferences from the Coleman Report. Dans M. Cherkaoui, *Sur l'égalité des chances : à propos du «Rapport Coleman»*. Revue française de sociologie, 19(2), 237-260.
- Carufel, J. de et Moreau, A. (2002). *Avis de l'Association des cadres scolaires du Québec au Ministre de l'éducation sur le projet de loi 124* (Dossier 02-05, ACSQ). Québec, Québec.
- Camilli, G. et Firestone, W. A. (1999). Values and State Ratings : An Examination of the State-By-State Education Indicators in Quality Counts. *Educational Measurement :Issues And Practice*, 18(4), 17-26.
- Cheong Cheng, Y. (1996). *School Effectiveness and School Based Management: A mechanism for Development*. London, Grande Bretagne : The Falmer Press.
- Cherkaoui, M. (1978). Sur l'égalité des chances : à propos du «Rapport Coleman. *Revue française de sociologie*, 19(2), 237-260.
- Cohen, J. (1969). *Statistical Power Analysis for the behavioral sciences*. New York, New York : Academic Press.
- Cohen, J. et Baldi, S. (1998). *An Evaluation of the Relative Merits of HLM vs. Robust Linear Regression in Estimating Models with Multi-Level Data*, Mimeo, Washington, DC, American Institutes for Research, February.

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D. et York, R. L. (1966). *Equality of Educational Opportunity*. Washington DC: NCES, US Government Printing Office.

Conseil supérieur de l'éducation. (2007). *Les projets pédagogiques particuliers au secondaire: Diversifier en toute équité*. Québec, Québec : Gouvernement du Québec.

Darling-Hammond, L., Holtzman, D. J., Gatlin, S. J. et Heilig, J. V. (2005, April). *Does teacher preparation matter? Evidence about teacher certification, teach for America, and teacher effectiveness*. Communication présentée au congrès the American Educational Research Association. Montréal, Québec, Canada.

De Landsheere, G. (1979). *Dictionnaire de l'évaluation et de la recherche en éducation*. Paris : Les Presses universitaires de France.

Dumay, X. et Dupriez, V. (2005). Effet établissement : quelles relations entre composition sociale des établissements et processus internes? *Mesure et évaluation en éducation*, 28(2), 67-92.

Dumay, X. et Dupriez, V. (2009). *L'efficacité dans l'enseignement : Promesses et zones d'ombre?* Bruxelles : Éditions De Boeck Université.

Dormont, B. (1999). *Introduction à l'Économétrie*, Montchrestien, Paris, France.

- Duru-bellat, M. et Leroy-Audoin, C. (1990). Les pratiques pédagogiques au CP. Structure et incidence sur les acquisitions des élèves. *Revue française de pédagogie*. Dans F. Petrucci, *Compétences en lecture à Genève: influence des caractéristiques des élèves et des classes à différents moments de la scolarité obligatoire. Mesure et évaluation en éducation*, 33(1), 79-106.
- Felouzis, G. (2004). Les indicateurs de performances des lycées, une analyse critique. *Éducation et formations*, 70, 83-95.
- Felouzis, G. et Perroton, J. (2007). Repenser les effets d'établissement : marchés scolaires et mobilisation. *Revue française de pédagogie*, 159.
- Fertig, M. et Wright, R. E. (2004). *School Quality, Educational Attainment and Aggregation Biases*. Germany: Institut for the Study of Labor. 994.
- Fitz-Gibbon, C. T. (1996). *Monitoring Education : Indicators, quality and effectiveness*. London: Cassel.
- Goldstein, H. (1991). Better Ways to Compare Schools ? *Journal of Educational Statistics*, 16(2), 89-91.
- Goldstein, H. et Woodhouse, G. (2000). School Effectiveness Research and Educational Policy. *Oxford Review of Education*, 26(3), 353-363.
- Goldstein, H., Huiqi, P., Rath, T. et Hill, N. (2000). The use of value added information in judging school performance. *Correspondence by Harvey Goldstein. Institute of Education*, London, UK.

- Greenwald, R., Hedges, L. V. et Laine, R. D. (1996). The Effects of School Resources on Student Achievement. *Review of Educational Research*, 66(3), 361-396.
- Hanushek, E. A. (1994). The economics of schooling: Production and efficiency in public school. *Journal of Economic literature*, 24(3), 1141-1177.
- Hanushek, E. A. (1989). The Impact of Differential Expenditure on School Performance. *Educational Researcher*, 18(4), 45-51.
- Hanushek, E. A. et Kain, J. F. (1972). On the Value of Equality of educational Opportunity as a Guide to public Policy. Dans M. Cherkaoui, *Sur l'égalité des chances : à propos du «Rapport Coleman»*. Revue française de sociologie, 19(2), 237-260.
- Hedges, L. V., Greenwald, R. et Laine, R. (1994). Does money matter? A meta-analysis of studies of the effects of differential school inputs in student outcomes. Dans *Assessing The Effect of School Resources on Student Performance: An Update*. University of Rochester.NY : Center for Economic Research.
- Hill, P. W. et Rowe, K. J. (1996). Multilevel Modeling in School Effectiveness Research. *School Effectiveness and School Improvement*, 7(1), 1-34.
- Hinds, P. S., Vogel, R. J. et Clarke-Steffen, L. (1997). The possibilities and pitfalls of doing a secondary analysis of a qualitative data set. *Qualitative Health Research*, 7(3), 408-424.

- Howel, C. D. (2002). *Méthodes statistiques en sciences humaines*. Bruxelles : Éditions De Boeck Université.
- Jencks, C., Smith, M., Acland, H., Bane, M. J., Cohen, D., Gintis, H. et collab. (1972). *Inequality: A Reassessment of the Effect of Family and Schooling in America*. New York: Basic Books.
- Keeves, J. P., Hungi, N. et Afrassa, T. (2005). Measuring Value Added Effects Across Schools : Should Schools Be Compared In Performance? *Studies in Educational Evaluation*, 31(2-3), 247-266.
- Konstantopoulos, S. (2006). Trends of School Effects on Student Achievement : Evidence from NLS :72, HSB :82, and NELS :92. *Teachers College Record*, 108(12), 2550-2581.
- Laveault, D. (2009). L'amélioration de l'efficacité du système éducatif : sur quels indicateurs s'appuyer? Dans X. Dumay, et V. Dupriez. (dir.), *L'efficacité dans l'enseignement : Promesses et zones d'ombre?* (pp. 177-194). Bruxelles : Éditions De Boeck Université.
- Lee, V. et Bryk, A. S. (1989). A multilevel model of the social distribution of high school achievement. *Sociology of Education*, 62(3), 172-192.
- Legendre, R. (2005). *Dictionnaire actuel de l'éducation*. Montréal, Québec : Larousse.
- Madaus, G. F., Kellaghan, T., Rakow, E. A. et King, D. J. (1979). *The sensitivity of measures of school effectiveness*. Harvard Educational Review, 49(2), 207-220.

- Mandeville, G. K. et Anderson, L. W. (1987). The stability of school effectiveness indices across grade levels and subject areas. *Journal of Educational Measurement*, 24(3), 203-216.
- Marceau, R. et Cowley, P. (2001). *Études sur les Politiques Éducationnelles: Bulletin des écoles secondaires du Québec*. Québec : Institut économique de Montréal et Institut Fraser.
- Mayeske, G. W., Wisler, C. E., Beaton, A. E., Weinfeld, F. D., Cohen, W. M., Okada, T., Proshok J. M. et Tabler, K. A. (1969). *A study of our nation's schools*. Washington DC: US Government Printing Office.
- McAndrew, M., Ledent, J. et Ait-Said, R. (2005). *La réussite scolaire des jeunes noirs au secondaire*. Rapport de recherche. Immigration et Métropoles. Montréal, Québec.
- Meyer, R. H. (2002). *Value-Added Indicators: Do They Make an Important Difference? Evidence From the Milwaukee Public Schools*. Wisconsin Center For Educational Research. Communication présentée à Annual Meeting of the American Educational Research Association, New Orleans, Louisiane, USA.
- Ministère de l'Éducation du Québec (1992). *Pour une langue belle : Priorités dans l'enseignement du français, langue maternelle et langue d'adoption*. Québec, Québec : Gouvernement du Québec.

Ministère de l'Éducation du Québec, Direction de l'évaluation. (2003). *Grille d'évaluation de la compétence à écrire un texte argumentatif : Épreuve unique d'écriture du Français, langue d'enseignement, 5<sup>e</sup> année du secondaire*. Québec, Québec : Gouvernement du Québec.

Ministère de l'Éducation du Québec, Direction de l'évaluation (2004). *Politique de l'évaluation des apprentissages*. Québec, Québec : Gouvernement du Québec.

Ministère de l'Éducation du Québec (2004). *Le Programme de formation de l'école québécoise : Domaine des langues*. Québec, Québec : Gouvernement du Québec.

Ministère de l'Éducation, du Loisir et du Sport, Direction de la recherche des statistiques et des indicateurs (2005). *La carte des unités de peuplement de 2003 : Les principales données socio-économiques et démographiques du recensement 2001 selon les territoires des commissions scolaires*. Québec, Québec : Gouvernement du Québec.

Ministère de l'Éducation, du Loisir et du Sport (2007). *Indicateurs de l'Éducation*. Québec, Québec : Gouvernement du Québec.

Ministère de l'Éducation, du Loisir et du Sport, Direction de la formation générale des jeunes (2008). *Mieux soutenir le développement de la compétence à écrire : Rapport du Comité d'experts sur l'apprentissage de l'écriture*. Québec, Québec : Gouvernement du Québec.

- Ministère de l'Éducation du Québec, Direction de l'évaluation (2011). *Document d'information: Épreuve unique d'écriture du Français, langue d'enseignement, 5<sup>e</sup> année du secondaire*. Québec, Québec :Gouvernement du Québec.
- Murnane, R. J. (1975). The Impact of School Resources on the Learning of Inner City Children. Dans C. Teddlie et D. Reynolds (dir.), *The International Handbook of School Effectiveness Research* (55-133). London: Routledge Falmer Press.
- Nuttal, D. (1991). An Instrument to be honed: Tables do not reflect school's true performance. Dans *A Brief History of Educational Value Added: How Did We Get To Where We Are? School Effectiveness and School Improvement*, 10(2), 233-256.
- Nye, B. A., Hedges, L. V. et Konstantopoulos, S. (2000). The Effects of small classes on academic achievement : The results of the Tennessee class size experiment. *American Educational Research Journal*, 37(1), 123-151.
- Organisation de coopération et de développement économiques - OCDE (2006). *Programme international pour le suivi des acquis des élèves (PISA)*. Paris, France : Éditions OCDE.
- Opendakker, M.-C. et Van Damme, J. (2000). The Importance of Identifying Levels in Multilevel Analysis: An Illustration of the Effects of Ignoring the Top or Intermediate Levels in School Effectiveness Research. *School Effectiveness and School Improvement: An International Journal of Research, Policy and Practice*, 11(1), 103-130.

- Osborne, W. J. (2002). Notes on the uses of Data transformations. *Practical Assessment, Research and Evaluation*. Raleigh, North Carolina, North Carolina State University, 8(6).
- Ouedraogo, S. D. M. (2006). *L'hybridation des données du recensement et des fichiers des résultats scolaires, une base pour accroître le potentiel explicatif des indicateurs de performance en éducation*. (Thèse de doctorat inédite). Université de Montréal, Montréal, Québec.
- Peaker, G. F. (1967). *The regression analysis on a National Survey*. Central Advisory Council for Education (England). *Children and their primary schools. (Plowden Report)*. London. HSMO, vol. 2.
- Pelgrum, W., J., Voogt, J. et Plomp, T. (1995). Curriculum indicators in international comparative research, Measuring the Quality of Schools. Dans C. Teddlie et D. Reynolds (dir.), *The International Handbook of School Effectiveness Research* (55-133). London: Routledge Falmer Press.
- Petrucci, F. (2010). Compétences en lecture à Genève : influence des caractéristiques des élèves et des classes à différents moments de la scolarité obligatoire. *Mesure et évaluation en éducation*, 33(1), 79-106.
- Plewis, I. (1997). *Statistics in Education*. Great Britain: Arnold Publications.
- Raudenbush, S. W. et Wilms, J. D. (1995). The Estimation of school effects. *Journal of Educational and Behavioral statistics*, 20(4), 307-335.

- Raudenbush, S. W. et Bryk A. S. (2002). *Hierarchical Linear Models: Applications and Data Analysis Methods*. Second Edition. Newbury Park, California :SAGE Publications.
- Rapport Parent (1963). *Résumé du Rapport de la commission royale d'enquête sur l'enseignement dans la province de Québec*. Québec, Québec : Gouvernement du Québec.
- Reynolds, D., Teddlie, C., Creemers, B., Scheerens, J. et Townsend, T. (2000). An Introduction to School Effectiveness Research. Dans C. Teddlie et D. Reynolds (dir.), *The International Handbook of School Effectiveness Research* (3-25). London: Routledge Falmer Press.
- Rouet, G. (1993). *L'invention de l'école*. France : Les Presses universitaires de Nancy.
- Rumberger, R. W. et Palardy, G. J. (2005). Test Scores, Dropout Rates, and Transfer Rates as Alternative Indicators of High School Performance. *American Educational Research Journal*, 42(1), 3-42.
- Sammons, P., Mortimore, P. et Thomas, S. (1996). Do schools perform consistence across outcomes and areas? Dans C. Teddlie et D. Reynolds (dir.), *The International Handbook of School Effectiveness Research* (pp. 55-133). London: Routledge Falmer Press.
- Scheerens, J., Glass, C. et Thomas, S. M. (2007). *Educational Evaluation, assessment, and Monitoring: A Systemic Approach*. London : Taylor and Francis Publications.

- Schmidt, W. H., Houang, R. T. et McKnight. C. C. (2005). Value-Added Research : Right Idea but Wrong Solutions? Dans R. Lissitz (dir.), *Value Added Models in Education: Theory and Applications* (pp. 145-165). Minnesota: JAMPRESS. Meaple Grove.
- Smith, M. (1972). Equality of educational Opportunity: the basic Findings reconsidered. Dans M. Cherkaoui, *Sur l'égalité des chances : à propos du «Rapport Coleman»*. Revue française de sociologie (pp 237-260), 19(2).
- Snijders, T. A. B. et Bosker, R. J. (1999). *An Introduction to basic and advanced multilevel modeling*. London : SAGE Publications.
- Spring, J. (1996). *The American School*. New York : McGraw-Hill Inc.
- SPSS (2010). Base 17.0. SPSS Inc. Chicago, Illinois. USA.
- Stevens, J. (2005). The Study of School Effectiveness as a Problem in Research Design. Dans R. Lissitz (dir.), *Value Added Models in Education: Theory and Applications* (pp. 166-1208). Minnesota: JAMPRESS. Meaple Grove.
- Stringfield, S. (2002). Science Making a Difference : Let's be Realistic! *School Effectiveness and School Improvement*, 13(1), 15-29.
- Stufflebeam, D. L. (1971). The relevance of the CIPP evaluation model for educational accountability. *Journal of research and Development in Education*, 5, 19-25.

- Teddlie, C., Reynolds, D. et Sammons, P. (2000). The Methodology and Scientific Properties of School Effectiveness Research. Dans C. Teddlie et D. Reynolds (dir.), *The International Handbook of School Effectiveness Research* (55-133). London: Routledge Falmer Press.
- Teddlie C. et Stringfield, S. (1993). Schools Do Make a Difference: Lessons Learned from a 10-Year Study of School Effects. Dans C. Teddlie et D. Reynolds (dir.), *The International Handbook of School Effectiveness Research*. London : Routledge Falmer Press.
- Thrupp, M. (1999). *Schools Making a Difference: Let's be Realistic!* Buckingham, Grande-Bretagne : Open University Press.
- Tyack, D. et Cuban, L. (2001). *Tinkering Toward Utopia: A Century of Public School Reform*. Cambridge, Massachusetts : Harvard University Press.
- Van der Maren, J.-M. (1995). *Les méthodes de recherche pour l'éducation*. Montréal, Québec : Les Presses de l'Université de Montréal et De Boeck.
- Wiggan, G. (2007). Race, School Achievement, and Educational Inequality: Toward a Student-Based Inquiry Perspective. *Review of Educational Research*, 77(3), 310-331.
- Willms, J. D. (1992). Monitoring School Performance: A guide for Educators. Dans C. Teddlie et D. Reynolds (dir.), *The International Handbook of School Effectiveness Research*. London: Routledge Falmer Press.

Wößmann, L. (2003). Schooling Resources, Educational Institutions and Student Performance: the international Evidence. *Oxford Bulletin of Economics and Statistics*, 65(2), 117–170.

Zvoch, K. et Stevens, J. J. (2004). Longitudinal Effects of School Context and Practice on Mathematics Achievement. Dans R. Lissitz (dir.), *Value Added Models in Education: Theory and Applications*, Minnesota: JAMPRESS. Meaple Grove.