

Université de Montréal

**Utilisation d'une approche écologique pour l'analyse des résultats d'évaluations
standardisées : cas des performances en lecture aux tests PASEC2014-Cameroun**

Par

ALIOUM

Département d'administration et fondements de l'éducation

Faculté des sciences de l'éducation

**Mémoire présenté à la Faculté des études supérieures en vue de l'obtention du
grade de maître ès arts en mesure et évaluation en éducation**

Septembre 2020

© Alioum, 2020

Ce mémoire intitulé :

**Utilisation d'une approche écologique pour l'analyse des résultats
d'évaluations standardisées : cas des performances en lecture aux
tests PASEC2014-Cameroun**

Présenté par Alioum

a été évalué par un jury composé des personnes suivantes :

Sébastien Béland, président

Christophe Chénier, membre

Nathalie Loye, directrice

RÉSUMÉ

L'objectif de la présente étude est de soutenir l'utilisation d'une approche écologique pour l'analyse des données d'une évaluation standardisée. Nous implémentons cette approche en utilisant une Analyse de Classes Latentes (ACL) avec covariables sur les données du test de lecture du Programme d'Analyse des Systèmes Éducatifs de la CONFEMEN (PASEC) des élèves francophones de 6^e année du primaire au Cameroun (N=617) et montrons comment une telle approche peut susciter un regard nouveau sur les résultats de cette évaluation.

En effet, les programmes d'Évaluations standardisées à Grandes Échelles (ELE) visent à évaluer les apprentissages et les compétences des individus et fournissent des données d'aide à la décision dans beaucoup de pays (Hogan, 2017 ; Loye, 2011 ; Wagemaker, 2014). Dans le cadre de ces ELE, les performances sont estimées uniquement sur la base des réponses que fournissent les candidats et donc de leurs capacités cognitives (Zumbo et al, 2015). Pourtant, de plus en plus de recherches suggèrent de considérer la performance à un test comme un phénomène qui se réalise dans un réseau interconnecté de connaissances, de caractéristiques individuelles et de contextes particuliers (McNamara, 2007 ; McNamara et Roever, 2006 ; Mislevy, 2018 ; Zumbo et al, 2015). L'approche écologique, qui s'inscrit dans cette perspective, s'intéresse aux hypothèses contextuelles, sociales et culturelles lors de l'estimation des performances à un test (McNamara, 2007 ; McNamara et Roever, 2006 ; Zumbo et al, 2015).

Nos résultats mettent en évidence des écologies de performances en lecture qui varient en fonction des régions. Dans chacune de ces écologies, les profils de performances en lecture qui s'y dégagent dépendent des caractéristiques des élèves et des caractéristiques des milieux scolaires et extrascolaires dans lesquels ils évoluent. Nous mettons ainsi l'emphase sur le caractère situé des performances à un test permettant une lecture plus équitable des performances des différents candidats. Ce faisant, nous formulons des recommandations qui tiennent compte des réalités contextuelles des candidats contrairement aux recommandations uniformes souvent formulées à la suite des résultats aux ELE.

***Mots-clés :** Évaluation à Large Échelle, test, performance, approche écologique, Analyse des Classes Latentes.*

ABSTRACT

The purpose of this study is to support the use of an ecological approach for the analysis of standardized data assessment. This study is implemented by Latent Classes Analysis (ACL) with covariables on the test reading data of the « Programme d'Analyse des Systèmes Éducatifs de la CONFEMEN » (PASEC) french-speaking students of 6th grade primary school in Cameroon (N = 617) and shows how this approach can change the way we interpret the results of this study.

Indeed, Standardized Large Scale Assessments (ELE) programs aim to appraise the learning and skills of individuals and provide decision support data in many countries (Hogan, 2017 ; Loye, 2011; Wagemaker, 2014). In the context of these ELEs, performance is estimated solely based on the responses provided by the candidates and therefore of their cognitive abilities (Zumbo et al, 2015). Yet a growing body of research suggests viewing test performance as a phenomenon that occurs in an interconnected network of knowledge, individual characteristics, and particular contexts (McNamara, 2007 ; McNamara and Roever, 2006 ; Mislavy, 2018 ; Zumbo et al. , 2015). The ecological approach, which is part of this perspective, is interested in the contextual, social and cultural hypotheses when estimating the performance of a test (McNamara and Roever, 2006 ; McNamara, 2007 ; Zumbo et al, 2015).

Our results highlight reading performance ecologies that vary by region. In each of these ecologies, the reading performance profiles that emerge depend on the characteristics of the students and the characteristics of the school and out-of-school environments in which they operate. We therefore put the emphasis on the situated character of the performances on a test allowing a fairer reading of the performances of the different candidates. In doing so, we formulate recommendations that take into account the contextual realities of applicants, unlike the standard recommendations often made following ELE results.

Keywords : *International Large-Scale Assessment, test, performance, ecological approach, Latent Class Analysis.*

REMERCIEMENTS

Je tiens à exprimer mes sincères remerciements à ma directrice de recherche, Nathalie Loye, pour son encadrement et sa générosité intellectuelle. Elle a su, par ses encouragements et sa patience, m'initier à la recherche. Les mots me manquent pour exprimer toute ma gratitude...

Je tiens également à remercier les membres du jury de ce mémoire pour leurs rétroactions riches en apprentissages. A Sébastien Béland, tout d'abord comme président de ce jury, qui a toujours prêté une oreille attentive à mes différentes questions. Ensuite, à Sébastien Béland, comme enseignant, qui m'a beaucoup inspiré et orienté dans le cadre de ma recherche.

Mes sincères remerciements également à l'endroit des responsables du département d'administration et fondements de l'éducation. A M. Serge Larivée, vice-doyen aux études supérieures et à la recherche, Claudine Jomphe et Chantal Viau qui m'ont tenu la main durant toute ma formation.

Mes remerciements vont également à ma famille pour son soutien sans faille : mes deux mamans biologique et adoptive pour leurs appuis constants, ma tendre épouse pour sa patience et ses encouragements, Housnia et Souleyman pour les moments de joies qu'ils m'ont procurés malgré la distance, Bouba Moumini et Hamidou.

Je tiens également à remercier tous mes collègues et amis, Mohamadou Bello, Mohamed Younouss, Lamine Sibama, Ben Moustapha Diedhiou, Saïdou Séguèda, Walph Ferenzi, Tya Collins. Votre présence a éloigné de moi la solitude et m'a procuré du réconfort moral.

DÉDICACE

*À la mémoire de mon frère et tuteur Ousmanou Moumini,
... tes mots d'encouragement résonnent encore dans mes oreilles,
et que dire de tous ce que tu as entrepris pour que ce rêve se matérialise,
Homme franc, Homme humble,
tu laisseras à jamais un grand vide dans nos cœurs.*

TABLE DES MATIERES

RÉSUMÉ.....	i
ABSTRACT	ii
REMERCIEMENTS.....	iii
DÉDICACE.....	iv
TABLE DES MATIERES	v
LISTE DES TABLEAUX	viii
LISTE DES FIGURES.....	x
LISTE DES ACRONYMES ET ABRÉVIATIONS	xii
INTRODUCTION	1
CHAPITRE1 : PROBLÉMATIQUE.....	5
1.1 Contexte.....	5
1.2 Les évaluations à larges échelles (ELE).....	8
1.2.1 Impact des ELE	9
1.2.2 La mesure des performances dans les ELE.....	13
1.3 L'approche écologique	19
1.3.1 L'écologie du développement humain	20
1.3.2 L'écologie de la réponse à l'item	24
1.4 Question générale de recherche	27
CHAPITRE 2 : L'ANALYSE DES CLASSES LATENTES (ACL).....	29
2.1 Le modèle de l'ACL	31
2.1.1 Expression du modèle.....	31
2.1.2 Choix du nombre de classes latentes.....	32
2.2 Analyse de classes latentes avec covariables et invariance de mesure en ACL.....	34
2.2.1 Analyse de classes latentes avec covariables	34
2.2.2 Invariance de mesure en ACL.....	36
2.3 Exemples empiriques de l'utilisation de l'ACL	39
2.3.1 En sociologie.....	39
2.3.2 En médecine	41
2.3.3 En éducation	43
2.4 Questions spécifiques de recherche.....	47
CHAPITRE 3 : MÉTHODOLOGIE	48
3.1 Stratégie de recherche.....	48
3.1.1 Données PASEC2014.....	49
3.1.2 Données retenues pour nos analyses	52

3.2	Démarche d'analyse	60
3.2.1	Différentes étapes.....	60
3.2.2	Outils d'analyse	62
3.3	Considérations éthiques	62
CHAPITRE 4 : RÉSULTATS.....		64
4.1	Description de l'échantillon	64
4.1.1	Echantillon global.....	64
4.1.2	Caractéristiques personnelles des élèves dans les strates (Couche 2)	67
4.1.3	Environnement scolaire, classe, enseignants et directeurs dans les strates (Couche 3)	70
4.1.4	Écologie extrascolaire dans les strates (couche 4).....	74
4.1.5	Caractéristiques de la communauté et voisinage dans les strates (Couche 5).....	75
4.2	Taux de réussites aux différents items	76
4.2.1	Taux de réussites globaux	76
4.2.2	Taux de réussites dans les strates.....	78
4.3	Ecologie de performances en lecture	82
4.3.1	Le caractère situé des performances en lecture : analyse de l'invariance de mesure en fonction des strates	82
4.3.2	Profils de performances pour le test de lecture dans les strates.....	84
4.1.1	Éléments actifs dans l'écologie de réponse au test de lecture	91
CHAPITRE 5 : INTERPRETATION ET DISCUSSION.....		96
5.1	Le regard écologique sur les résultats du test de lecture PASEC2014	97
5.1.1	Des marqueurs propres à chaque strate.....	97
5.1.2	Des élèves et des milieux hétérogènes	99
5.1.3	Des écologies différenciées dans les strates.....	102
5.2	Apports de l'utilisation de l'approche écologique	107
5.2.1	D'une performance uniforme vers une performance contextualisée.....	107
5.2.2	Constats et recommandations au regard des écologies de réponses au test de lecture	114
5.3	Principales limites.....	118
CONCLUSION.....		120
BIBLIOGRAPHIE		122
ANNEXES.....		127
Annexe 1 : Variables utilisées pour le calcul du niveau socioéconomique (SES) des familles des élèves (PASEC, 2017b, p. 167).....		127
Annexe 2 : Variables utilisées pour le calcul de l'indice de ressources pédagogiques de la classe (PASEC, 2017b, p. 168).....		127

Annexe 3 : Variables utilisées pour le calcul de l'indice d'aménagement du territoire (PASEC, 2017b, p. 171).....	128
Annexe 4 : Variables utilisées pour le calcul de l'indice d'implication de la communauté (PASEC, 2017b, p. 173).....	128
Annexe 5 : Niveau de compétences, scores minimums et description de compétences au test de lecture en 6 ^e année primaire au Cameroun (PASEC, 2016, p. 44).....	129
Annexe 6 : Statistiques descriptives des données	130
Annexe 7 : Distribution des indicateurs de prévalences et probabilités conditionnelles par strates et par profil.....	131

LISTE DES TABLEAUX

Tableau 1 : Les cinq approches théoriques du développement (Papalia et al, 2010, p.12)	20
Tableau 2 : Les quatre différents modèles de variables latentes (Collins et Lanza, 2009, p.7).....	30
Tableau 3 : Proportion des réponses « oui » aux questions sur la délinquance adolescente (traduction libre de Collins et Lanza, 2009, p.11)	40
Tableau 4 : Classes latentes issues de l'ACL des données sur la délinquance adolescente (traduction libre de Collins et Lanza, 2009, p.12)	41
Tableau 5 : Qualité d'ajustement des modèles d'ACL sur les données d'état de santé (Lafortune, 2009, p. 102).....	42
Tableau 6 : Échantillon de l'évaluation PASEC2014 (PASEC, 2015)	49
Tableau 7 : Récapitulatif des sous-domaines évalués et leur poids en début de scolarité (PASEC, 2015).....	51
Tableau 8 : Récapitulatif des sous-domaines évalués et leur poids en fin de scolarité (PASEC, 2015)	51
Tableau 9 : Échantillon réalisé en 6e année primaire au Cameroun (PASEC, 2016)	52
Tableau 10 : Poids des processus évalués en lecture et leurs poids dans le test PASEC2014 (PASEC, 2017a).....	53
Tableau 11 : Répartition des blocs d'items des tests de lecture PASEC2014 dans les livrets (PASEC, 2017c, p. 32).....	53
Tableau 12 : Répartition des items de lecture PASEC2014 dans les livrets en fonction du processus évalué et des blocs (PASEC, 2017a, p. 25)	54
Tableau 13 : Échantillon d'élèves ayant effectué les 23 items du bloc A des tests de lecture PASEC2014 et retenu pour nos travaux.....	54
Tableau 14 : Classification PASEC de ses données contextuelles par niveau (PASEC, 2017c)	56
Tableau 15 : Sommaire des variables utilisées et leur classification dans le modèle écologique de Zumbo et ses collègues (2015).....	59
Tableau 16 : Répartition des 25% des élèves issus des milieux les plus défavorisés et des 25% des élèves issus des milieux les plus favorisés en fonction des strates.....	70
Tableau 17 : Indicateurs d'ajustement des modèles d'ACL pour les données globales.....	83
Tableau 18 : Indicateurs d'ajustement des modèles ACL à 4 classes latentes sur les données globales (modèle restreint et modèle non restreint).....	83
Tableau 19 : Indicateurs d'ajustement des modèles d'ACL pour la strate Grand Ouest.....	84
Tableau 20 : Indicateurs d'ajustement des modèles d'ACL pour la strate Grand Centre	84
Tableau 21 : Indicateur d'ajustement des modèles d'ACL pour la strate Grand Nord	84
Tableau 22 : Items communs aux profils de performances dans la strate Grand Ouest.....	86
Tableau 23 : Items communs aux profils de performances dans la strate Grand Centre.....	88
Tableau 24 : Items communs aux profils de performances dans la strate Grand Nord	90
Tableau 25 : Covariables participant à la formation des classes latentes dans la strate Grand Ouest ..	93
Tableau 26 : Covariables participant à la formation des classes latentes dans la strate Grand Centre	94
Tableau 27 : Covariables participant à la formation des classes latentes dans la strate Grand Nord...	95
Tableau 28 : Marqueurs des élèves de 6e année du primaire au Cameroun selon les strates.....	98
Tableau 29 : Portraits potentiels des élèves à fortes prédispositions et des élèves très vulnérables ..	100
Tableau 30 : Répartition des élèves à fortes prédispositions et des élèves très vulnérables dans les différentes strates.....	100
Tableau 31 : Caractéristiques des milieux scolaires et extrascolaires dans les différentes strates	102
Tableau 32 : Récapitulatif des résultats obtenus de l'ACL avec covariables.....	103
Tableau 33 : Écologies de performances dans les différentes strates.....	104

Tableau 34 : Pourcentages d'élèves des profils extrêmes à faire partie des profils de performances dans l'écologie à « conditions favorables et facteurs scolaire et extrascolaire déterminants » (Grand Ouest)	105
Tableau 35 : Pourcentages d'élèves des profils extrêmes à faire partie des profils de performances dans l'écologie à « conditions modérées et facteurs scolaire et extrascolaire déterminants » (Grand Centre).....	105
Tableau 36 : Pourcentages d'élèves des profils extrêmes à faire partie des profils de performances dans l'écologie à « conditions défavorables et facteur extrascolaire déterminant » (Grand Nord).....	105
Tableau 37 : Piste de réflexions du PASEC à l'issus des évaluations en lecture et en mathématiques des élèves de 6e année primaire (PASEC, 2015, p. 149-154).....	114

LISTE DES FIGURES

Figure 1 : Processus de mesure des attributs psychologiques (Bertrand et Blais, 2004, p.32).....	15
Figure 2 : Le modèle écologique du développement humain de Bronfenbrenner (Papalia et al, 2010, p.27).....	22
Figure 3 : Modèle écologique de la réponse à l’item (Zumbo et al, 2015, p.05).....	26
Figure 4 : Variable latente avec trois variables observées (Collins et Lanza, 2009, p.05).....	30
Figure 5 : Données collectées dans le cadre du PASEC2014 (PASEC, 2014, p.19).....	50
Figure 6 : Etapes de la recherche et résultats obtenus	62
Figure 7 : Répartition des élèves par sexe	64
Figure 8 : Distribution des âges des élèves de l’échantillon global.....	64
Figure 9 : Distribution des âges des élèves par sexe	64
Figure 10 : Pourcentages d’élèves en fonction du redoublement.....	65
Figure 11 : Pourcentage d’élèves en fonction de la fréquentation du préscolaire	65
Figure 12 : Répartition des élèves en fonction du statut de l’école	66
Figure 13 : Répartition des élèves en fonction de la zone de fréquentation	66
Figure 14 : Pourcentages d’enseignants et de directeurs par sexe.....	66
Figure 15 : Pourcentages d’enseignants et de directeurs en fonction du niveau d’études atteint.....	66
Figure 16 : Distribution des années d’expériences des enseignants	67
Figure 17 : Répartition des élèves par sexe dans les strates	68
Figure 18 : Distribution des âges des élèves dans les strates.....	68
Figure 19 : Pourcentages d’élèves en fonction du redoublement dans les strates	69
Figure 20 : Pourcentage d’élèves en fonction de la fréquentation du préscolaire dans les strates	69
Figure 21 : Distribution du statut socio-économique des élèves dans les strates	70
Figure 22 : Répartition des élèves en fonction de la zone de fréquentation dans les strates	71
Figure 23 : Répartition des élèves en fonction du statut de l’école dans les strates	71
Figure 24 : Distribution de l’indice de ressources pédagogiques des classes dans les strates.....	72
Figure 25 : Répartition des enseignants en fonction de leur sexe dans les strates.....	72
Figure 26 : Répartition des directeurs en fonction de leur sexe dans les strates.....	72
Figure 27 : Répartition des enseignants en fonction du niveau d’études atteint dans les strates.....	73
Figure 28 : Répartition des directeurs en fonction du niveau d’études atteint dans les strates.....	73
Figure 29 : Distribution des années d’expériences des enseignants dans les strates	74
Figure 30 : Répartition des élèves en fonction de la pratique de la langue d’enseignement dans les strates.....	74
Figure 31 : Distribution de l’indice d’aménagement de la localité de l’école en fonction des strates .	75
Figure 32 : Distribution de l’indice d’implication de la communauté en fonction des strates	76
Figure 33 : Pourcentages de bonnes réponses aux items pour l’ensemble de l’échantillon	77
Figure 34 : Pourcentage d’items selon les taux de réussite	78
Figure 35 : Pourcentages de bonnes réponses aux items pour la strate Grand Ouest.....	79
Figure 36 : Pourcentage d’items selon les taux de réussite dans la strate Grand Ouest	79
Figure 37 : Pourcentages de bonnes réponses aux items pour la strate Grand Centre	80
Figure 38 : Pourcentage d’items selon les taux de réussite dans la strate Grand Centre.....	80
Figure 39 : Pourcentages de bonnes réponses aux items pour la strate Grand Nord.....	81
Figure 40 : Pourcentage d’items selon les taux de réussite dans la strate Grand Nord	81
Figure 41 : Prévalences d’appartenance aux classes latentes dans la strate Grand Ouest	85
Figure 42 : Probabilités à fournir des réponses justes par item et par classe latente dans la strate Grand Ouest.....	86

Figure 43 : Prévalences d'appartenance aux classes latentes dans la strate Grand Centre.....	88
Figure 44 : Probabilités à fournir des réponses justes par item et par classe latente dans la strate Grand Centre	88
Figure 45 : Prévalences d'appartenance aux classes latentes dans la strate Grand Nord	90
Figure 46 : Probabilités à fournir des réponses justes par item et par classe latente dans la strate Grand Nord.....	90
Figure 47 : Performances des strates par rapport à la moyenne des scores nationaux en lecture – Fin de scolarité, Sous système francophone (PASEC, 2014, p. 53).....	108
Figure 48 : Pourcentage d'élèves de 6e année primaire par niveau de compétences en lecture, sous-système francophone (PASEC, 2014, p.52)	108
Figure 49 : L'écologie de performances au test de lecture dans la strate Grand Ouest.....	110
Figure 50 : L'écologie de performances au test de lecture dans la strate Grand Centre.....	111
Figure 51 : L'écologie de performances au test de lecture dans la strate Grand Nord.....	112

LISTE DES ACRONYMES ET ABRÉVIATIONS

ACL	Analyse de Classes Latentes
CP	Cours Préparatoire
CL2	Class two (deuxième année du primaire)
CL5	Class five (cinquième année du primaire)
CL6	Class six (sixième année du primaire)
CM1	Cours Moyen première année
CM2	Cours Moyen deuxième année
CONFEMEN	Conférence des ministres de l'Éducation des États et gouvernements de la Francophonie
DSSEF	Document de Stratégie du Secteur de l'Éducation et de la Formation
ELE	Évaluations à Larges Échelles
MINEDUB	Ministère de l'Éducation de Base
MINEPAT	Ministère de l'Économie, de la Planification et de l'Aménagement du Territoire
MLA	Monitoring Learning Achievement
OCDE	Organisation de Coopération et de Développement Économiques
PIRLS	Programme international de recherche en lecture scolaire
PISA	Programme International pour le Suivi des Acquis des élèves
SACMEQ	Consortium d'Afrique australe et orientale pour le pilotage de la qualité de l'éducation
TIMMS	Tendances dans les Études Internationales de Mathématiques et de Sciences

INTRODUCTION

Les évaluations standardisées d'apprentissages encore appelées Évaluations à Grandes Échelles (ELE) ont connu un grand essor durant ces trois dernières décennies (Loye, 2011; Rutkowski et al, 2013). Aujourd'hui, ces types d'évaluations sont les mieux connues et les plus répandues sur le globe (Hogan, 2017). Cette situation se justifie d'abord par la fièvre de l'évaluation qui s'est emparée des systèmes éducatifs dans les années 1990 (Damon, 2009 ; Gingras, 2008). Des années durant lesquelles les politiques de l'*accountability* ont su imposer aux acteurs du secteur public le fait de rendre compte dans le but de légitimer leurs actions (Felouzis et Hanhart, 2011). Durant ces périodes fastes du pilotage par les résultats, les ELE ont également profité de l'absence d'appareils statistiques fiables permettant d'évaluer les politiques publiques en matière d'éducation dans beaucoup de pays (Mons et Crahay, 2011). Ces deux facteurs ont donc permis aux ELE de s'imposer aujourd'hui comme un outil de bonne gouvernance (Mons et Crahay, 2011) et une boussole dans l'orientation de politiques éducatives (Loye, 2011; Wagemaker, 2014).

Parmi ces ELE, on peut notamment citer le Programme International pour le Suivi des Acquis des élèves (PISA), les Tendances dans les Études Internationales de Mathématiques et de Sciences (TIMSS), le Programme international de recherche en lecture scolaire (PIRLS), le Consortium d'Afrique australe et orientale pour le pilotage de la qualité de l'éducation (SACMEQ) et le Programme d'analyse des systèmes éducatifs de la CONFEMEN¹ (PASEC). Ces quelques programmes couvrent ensemble plus d'une centaine de pays (Loye, 2011). À travers des tests psychométriques et des questionnaires contextuels², la plupart de ces ELE se proposent de mesurer les performances des individus et par là, celles des systèmes éducatifs (Loye, 2011). Cependant, la nature complexe de la performance en éducation rend cette tâche de mesure assez ardue (Bertrand et Blais, 2004). La performance à un test serait en effet non seulement liée aux capacités cognitives des individus, mais également à certaines de leurs caractéristiques individuelles, à leur environnement de vie, à l'environnement dans lequel ils réalisent leurs apprentissages ainsi qu'aux conditions dans lesquelles ils effectuent le test (Mislevy, 2018 ; Pons, 2011; Zumbo et al, 2015). Toutefois, la prise en compte de cette complexité ne transparait pas dans les stratégies d'analyses des tests dans les ELE.

¹ Conférence des ministres de l'Éducation des États et gouvernements de la Francophonie.

² Caractéristiques de l'élève, de son milieu familial, de l'environnement de l'école et de la classe.

Les analyses des tests, dans le cadre des ELE, consistent à estimer les performances uniquement sur la base des réponses des individus aux différentes tâches d'un test, donc de leurs capacités cognitives, et ce n'est qu'à posteriori que les analyses s'intéressent aux effets des variables contextuelles sur ces performances. Une telle approche met plus d'emphasis sur la dimension cognitive de l'évaluations sacrifiant ainsi sa dimension sociale qui est tout aussi importante. Pourtant, des recherches en sciences de l'éducation et en sciences sociales proposent aujourd'hui de nouveaux outils conceptuels qui permettent d'appréhender la performance à un test dans toute sa complexité (Carley, 2001 ; Hox, 2002 ; Mislevy, 2018). Ces recherches qui présentent des stratégies de modélisation de la mesure très utiles, ouvrent de nouvelles perspectives pour les chercheurs en sciences de l'éducation (Bertrand et Blais, 2004 ; Von Davier et Sinharay, 2013). La perspective écologique que Zumbo et ses collègues (2015) proposent est une approche qui s'inscrit dans cette logique où la performance est considérée comme un phénomène qui se réalise dans un réseau interconnecté de connaissances, de caractéristiques individuelles et de contextes particuliers. Voir la performance à travers de telles lentilles nous plonge dans l'univers social de l'évaluation en nous invitant à lire une performance à un test à l'aune des caractéristiques des candidats, et des environnements scolaire et extrascolaire dans lesquels ils évoluent. Une telle perspective permettrait non seulement une interprétation moins surreprésentée des résultats, mais aiderait également à améliorer les prises de décisions politiques (Mislevy, 2018 ; Zumbo et al, 2015).

Le présent travail de recherche mobilise donc cette approche écologique et propose de revisiter les résultats du test de lecture du PASEC2014 au Cameroun des élèves francophones de sixième année du primaire, à l'aide de l'analyse des classes latentes (ACL) avec covariables. Notre choix de travailler sur le test de lecture PASEC2014 repose en grande partie sur l'enjeu de plus en plus crucial que jouent les compétences littéraires et notamment la lecture au sein de la société. En effet, en plus de servir de support d'apprentissage, la connaissance fonctionnelle de la lecture est nécessaire, voire indispensable pour trouver sa place dans la société (Giasson et Escoyez, 2013 ; Giasson et Vandecasteele, 2011). A ce titre, l'enseignement de la lecture reste une tâche importante pour les enseignants à la maternelle et au primaire (Giasson et Escoyez, 2013 ; Giasson et Vandecasteele, 2011). De même, la capacité des apprenants à pouvoir lire demeure une préoccupation majeure pour les décideurs du système éducatif (PASEC, 2015).

Ainsi, notre travail est structuré en cinq chapitres. Le premier chapitre est consacré à la problématique. Dans ce chapitre, nous déroulons de prime abord la genèse de notre projet de

recherche puis nous présentons les ELE ainsi que les approches qui servent à mesurer les performances dans le cadre de ces évaluations. Dans la deuxième partie du chapitre, nous portons un autre regard sur la notion de performance à un test. Nous aborderons donc dans cette partie les approches écologiques ; l'écologie du développement humain, pour ce qui est du processus d'acquisition des connaissances et l'écologie de la réponse à l'item, pour ce qui est du processus de restitution de cette connaissance notamment lors d'un test. Nous terminons le premier chapitre en énonçant notre question générale de recherche.

Au deuxième chapitre, nous présentons une alternative aux modèles de mesures utilisés pour estimer les performances aux ELE. La présentation de cette alternative, à savoir le modèle de l'Analyse des Classes Latentes (ACL), meuble la première partie de ce chapitre. Dans la deuxième partie du chapitre, nous regardons le modèle ACL avec covariables qui étend le modèle ACL et permet ainsi aux covariables de participer aux estimations au moment de mesurer. Dans cette deuxième partie, nous regardons également l'analyse de l'invariance de la mesure en ACL et nous montrons en quoi cette analyse peut servir d'argument pour tenir compte des variables contextuelles lors de la mesure des performances à un test. Nous clôturons ce chapitre en présentant les questions spécifiques de notre recherche à la suite de quelques exemples empiriques d'utilisation de l'ACL dans le domaine de la psychologie, de la santé et de l'éducation.

Le troisième chapitre de notre travail est réservé au cadre méthodologique. Dans une première partie, nous détaillons la stratégie retenue pour notre recherche. Nous nous attardons dans cette partie sur le contexte de l'évaluation PASEC2014, les données de cette évaluation et les données que nous retenons pour notre recherche. Ensuite, nous déroulons la démarche retenue pour nos analyses. Le but de cette deuxième section est de présenter les différentes étapes de nos analyses et les différents outils qui ont servi à réaliser ces analyses. Nous clôturons ce chapitre en relevant quelques considérations éthiques liées à notre recherche.

Le quatrième chapitre est consacré à la présentation des résultats de nos analyses. La première partie de cette présentation est dédiée aux statistiques descriptives des différentes variables contextuelles retenues pour nos analyses. Nous présentons ensuite les résultats aux différents items du test de lecture PASEC2014. Et, dans la dernière partie, nous déroulons l'écologie de performances en lecture qui résulte de nos analyses ACL avec covariables. Nous présentons ainsi les différents profils de performances qui résultent de nos analyses et les différentes covariables qui contribuent à la formation de ces profils.

Dans le cinquième chapitre, nous discutons de nos résultats en lien avec notre question générale recherche et nos deux questions spécifiques de recherche. Nous engageons ces discussions autour des différents profils de performance et des écologies de réponses au test de lecture PASEC2014 que nous avons obtenu. Nous soulignons ainsi, tous les éléments nouveaux que nous apporte le regard écologique pour la lecture des performances des élèves au test de lecture PASEC2014. Nous esquissons ensuite, les possibles pistes de décisions qui pourraient être prises à la lumière de ces résultats. Nous ne manquons pas également de souligner les différentes limites de nos travaux dans ce dernier chapitre. La partie conclusion clôt ensuite notre travail en le revisitant sommairement, et en abordant différentes pistes de recherches futures.

CHAPITRE 1 : PROBLÉMATIQUE

1.1 Contexte

Notre recherche, comme toute recherche d'ailleurs, a commencé autour d'une question intrigante qui tire ses racines dans notre pays natal, le Cameroun, où nous exerçons comme cadre d'études à la cellule de la planification du Ministère de l'Education de Base (MINEDUB). Informaticien de formation, nous avons la responsabilité de gérer la base de données de la carte scolaire ; une mine d'informations qui donne non seulement accès à une vue holistique sur le sous-secteur de l'éducation de base, mais qui ouvre également des portes à des collaborations avec des partenaires techniques et financiers du MINEDUB et des projets dans ce sous-secteur. C'est dans ce cadre que nous avons fait partie de l'équipe nationale du Programme d'Analyse des Systèmes Educatifs de la CONFEMEN (PASEC) Cameroun en tant qu'informaticien de l'équipe.

Le programme PASEC, comme beaucoup de programmes d'évaluations standardisées, s'emploie à évaluer les acquis scolaires des élèves afin de comparer les performances des systèmes éducatifs à l'aide des données recueillies avec des tests associés à des questionnaires d'enquête (Loye, 2011). Ces tests, pour le cas du PASEC, mesurent les compétences fondamentales en lecture et en mathématiques des élèves en début et en fin de scolarité primaire (PASEC, 2015). Le Cameroun prend part aux évaluations PASEC depuis 2004. Toutefois, la culture d'évaluation des acquis scolaires dans le pays date des années 1990 avec sa participation au Monitoring Learning Achievement (MLA I) ou Suivi Permanent des Acquis Scolaires en 1992. En effet, après la conférence de Jomtien³ en 1990, plusieurs pays en voie de développement⁴ ont décidé d'élaborer une batterie d'indicateurs permettant de mesurer la qualité de l'éducation. Parmi ces indicateurs figure en pole position, le niveau d'atteinte des apprentissages qui est souvent mesuré grâce aux données fournies par les scores aux tests nationaux et internationaux (UNESCO, 2005).

Entre 1992 et 2014, le Cameroun a participé ainsi à cinq évaluations de grande envergure. La première évaluation fut celle du MLA I en 1992 où participaient les élèves anglophones des grades 4 et 5 (4^e et 5^e année primaire). Ces élèves ont répondu à des questions sur la littérature, les nombres et la culture générale (Chinapah, 2003). Puis, le pays a participé

³ Conférence mondiale sur « l'éducation pour tous » tenue à Jomtien, du 5 au 9 mars 1990.

⁴ Ou pays du Sud. Ce sont des expressions consacrées pour désigner des pays dont le niveau de développement économique est inférieur à celui des pays du Nord ou pays développés.

au MLA II en 2003. Lors de cette évaluation, ce sont les élèves du grade 8 (2^e année secondaire) qui ont composé en mathématiques et sciences avec une approche sur la culture générale (Chinapah, 2003). Le Cameroun participera ensuite aux enquêtes PASEC2007⁵, SOFRECO et PASEC2014 respectivement en 2004, 2011 et 2014.

Pour ces trois évaluations, les matières ciblées sont identiques : il s'agit de la langue d'enseignement et des mathématiques. De même, ces enquêtes tiennent compte de la spécificité bilingue du système éducatif camerounais avec la participation des élèves anglophones, mais également des élèves francophones. Toutefois, l'enquête SOFRECO a évalué les élèves du CM1/CL5 et les enquêtes PASEC2007 et PASEC2014 quant à elles ont ciblé les élèves du CP/CL2 et ceux des CM2/CL6. Les différents rapports de ces évaluations pointent une insuffisance des acquis des élèves et surtout des disparités importantes au niveau des performances entre les différents groupes d'élèves (MINEPAT, 2013). A la suite de ces publications, le gouvernement camerounais à travers le MINEDUB a pris certaines initiatives dans le sens d'améliorer la qualité des apprentissages des élèves.

C'est dans cette perspective que Le Programme d'Amélioration de la qualité de l'Éducation de Base (PAQUEB) et le Programme d'Amélioration de l'Équité et de la Qualité de l'Éducation (PAEQUE) vont voir le jour (MINEPAT, 2013). À travers ces programmes, les responsables du secteur éducatif prennent des actions parmi lesquelles : la contractualisation des maîtres de parents⁶, le recrutement d'enseignants dans des zones défavorisées, le renforcement du réseau scolaire à travers la construction de nouvelles salles de classe ou la distribution gratuite des manuels essentiels aux élèves du primaire (MINEPAT, 2013). Toutefois, malgré ces efforts, le Document de Stratégie du Secteur de l'Éducation et de la Formation (DSSEF) élaboré en 2013 déplore des résultats mitigés (MINEPAT, 2013). Ce constat est appuyé par la dernière évaluation PASEC2014 qui souligne les mêmes déficits et les mêmes disparités que ceux relevés dans les précédentes évaluations (PASEC, 2015).

Les groupes de travail ministériels ou interministériels, auxquels nous participions, posaient régulièrement des réflexions autour de ces faibles résultats des élèves et des disparités qui en découlaient. Le rapport PASEC pointe du doigt les politiques de redoublement en vigueur au Cameroun, l'absence de ressources pédagogiques suffisantes dans les classes et les

⁵ PASEC2007 en raison de l'année de parution du rapport de l'enquête en 2007.

⁶ Au Cameroun, il existe une catégorie d'enseignants du préscolaire ou du primaire appelée « Maîtres des parents ». Ce sont des enseignants recrutés et payé par des associations des parents d'élèves afin de pallier le déficit en enseignants dans certaines zones.

disparités entre les écoles en zones urbaines et celles en zones rurales comme possibles raisons aux contre-performances des élèves (PASEC, 2015). En dehors de ces raisons, les comités mettaient également en avant des arguments liés à la validité des tests et les difficultés à implémenter les recommandations issues des enquêtes sur le terrain. Sur le plan personnel, les disparités entre les élèves des différentes régions du pays étaient ce qui nous préoccupait le plus. Ces disparités opposent des élèves des régions de l'Ouest du pays - les plus performants selon les résultats - à ceux des régions du Centre – élèves moyens – et ceux des régions du Nord⁷ du pays – élèves en grande difficulté-. L'une des pistes qui nous semblait pertinente à explorer était le niveau de familiarité des élèves avec la langue d'enseignement.

Dans la partie septentrionale du pays, il est bien connu que la pratique des langues locales est très ancrée dans les habitudes des populations contrairement aux langues d'enseignement que sont le français et l'anglais (Tadadjeu et Mba, 1997). A l'opposé, dans les régions du Centre et de l'Ouest plus cosmopolites, les langues officielles font office de langue de communication et souvent, au détriment des langues locales (Tadadjeu et Mba, 1997). La pratique de la langue par des communautés est souvent perçue comme un facteur pouvant permettre ou non la maîtrise de cette langue. Nous avons alors pour hypothèse que cet aspect aurait un impact sur la maîtrise des langues d'enseignement par les élèves et par ricochet, sur leur maîtrise des autres disciplines scolaires. La principale question que nous nous posions était celle de savoir :

« *Quel rôle joue le niveau de familiarité des élèves avec la langue de passation du test dans les résultats PASEC2014 au Cameroun ?* »

Cependant, tout au long de notre formation en mesure et évaluation en éducation et à travers une revue de littérature importante, nous avons constaté que les enjeux actuels autour des ELE allaient au-delà de cette question. Celui que nous avons choisi d'aborder dans le cadre de notre travail tourne autour de la recherche du meilleur modèle de mesure pour appréhender les performances aux évaluations des ELE. Toutefois, cette recherche légitime est souvent taxée d'utopique ; l'utilisation de chaque modèle de mesure comportant ses avantages et ses risques (Bertrand et Blais, 2004). Cependant, nous commencerons dans la section suivante à débayer le chemin vers ce but à travers une présentation des ELE. Nous allons revenir brièvement sur

⁷ Les termes régions du Nord, du Centre et de l'Ouest font références aux strates de l'enquête PASEC2014. Ainsi, par régions de l'Ouest, il faut comprendre les régions du Nord-Ouest, de l'Ouest, du Littoral et du Sud-Ouest. Les régions du Centres regroupent elles, les régions du Centre, de l'Est et du Sud. Les régions du Nord renvoient aux régions de l'Adamaoua, de l'Extrême-Nord et du Nord.

la genèse de ces types d'évaluation puis nous présenterons les rôles de ces évaluations et les approches qui servent à mesurer les performances dans le cadre de ces évaluations. Nous clôturons cette partie en soulignant quelques limites de nature à remettre en cause les comparaisons entre groupes d'individus à l'aide des résultats des ELE.

1.2 Les évaluations à larges échelles (ELE)

Évaluer des individus à grande échelle à l'aide des tests standardisés pour ensuite les classer et les comparer ne date pas d'hier. Cette pratique remonterait vers l'an 1115 avant J.-C en chine (Bertrand et Blais, 2004). A cette époque, la dynastie Chan procédait à la sélection des bureaucrates à travers un concours constitué de tests standardisés (Bertrand et Blais, 2004). Au fil du temps, d'autres gouvernants de par le monde vont également essayer de comparer ou de classer leur population (Bertrand et Blais, 2004). Les évaluations standardisées des apprentissages des élèves, tel que nous les connaissons aujourd'hui, datent des années 1960 (Rutkowski et al, 2013). L'Association Internationale pour l'Évaluation des Acquis Scolaires (IEA) avait alors, à travers la First International Mathematics Study (FIMS), lancé la toute première évaluation internationale en mathématiques (Rutkowski et al, 2013). C'est cependant dans les années 1990, à la faveur de l'instauration des politiques de pilotage par les résultats dans plusieurs pays, que ces programmes vont se multiplier et gagner en prestige (Felouzis et Hanhart, 2011 ; Loye, 2011 ; Mons et Crahay, 2011).

La mise en œuvre de ces évaluations consiste à collecter des données sur le niveau d'apprentissage des individus à l'aide de tests et d'autres informations telles que les caractéristiques personnelles et socio-économiques des participants ou les conditions de passation. (American Psychological Association, 2003). Elles se déroulent généralement à l'échelle nationale ou internationale et les participants se voient soumis à des conditions de passation identiques (Loye, 2011 ; Rutkowski et al, 2013). Les étapes les plus récurrentes durant ces évaluations sont entre autres : la définition des objectifs de l'évaluation, la confection des instruments de collecte de données, le choix des participants, les prétests, l'analyse des données et la dissémination des résultats (OCDE, 2016 ; PASEC, 2015). Ces opérations peuvent s'étaler sur trois à cinq ans (Loye, 2011) et impliquer des équipes pluridisciplinaires composées de pédagogues, de psychopédagogues, de statisticiens, d'informaticiens, de psychométriciens et beaucoup d'autres spécialistes (OCDE, 2016 ; PASEC, 2015).

Tous ces programmes d'évaluation partagent des objectifs communs qui sont : (i) de fournir des informations permettant d'améliorer les pratiques d'enseignement et le processus

d'apprentissage ; (ii) de permettre aux différents pays participants de prendre des décisions éclairées en matière de politiques éducatives et (iii) de comparer les différents systèmes éducatifs des pays du monde en matière d'organisation, de curricula, de ressources et de pratiques favorisant la réussite des élèves (OCDE, 2016 ; PASEC, 2015). A ces objectifs communs peuvent s'ajouter d'autres objectifs spécifiques en lien avec les populations cibles. Le PISA par exemple vérifie si les systèmes éducatifs préparent les élèves à faire face aux défis futurs de leur vie d'adultes (Loye, 2011). Toutefois, certaines voix s'élèvent pour dire que ces différents objectifs sont loin d'être réalistes. Les raisons souvent évoquées portent sur la capacité de ces évaluations à fournir des mesures valides ou de qualité ; ce d'autant plus que chaque pays ou région affiche des spécificités contextuelles et scolaires propres difficilement comparables à celles des autres pays ou régions (Heyneman et Lee, 2014 ; Pons, 2011). Malgré ces doutes, les traces des effets des ELE sur les politiques éducatives nationales sont assez visibles. Les analystes et les chercheurs s'accordent à regrouper ces effets en lien avec leur acceptation au sein des pays, leurs impacts sur les discours politiques, les programmes scolaires, les méthodes d'enseignement, la recherche et le partenariat en éducation avec certaines organisations de développement.

1.2.1 Impact des ELE

Consensus grandissant sur la nécessité de mener des ELE

Au fil du temps, de nombreux pays ont rejoint ceux qui participaient déjà aux ELE marquant ainsi leur intérêt pour ces évaluations. L'évaluation PIRLS qui a regroupé 37 pays participants en 2001 a connu ensuite la participation de 58 pays en 2011 (Loye, 2011 ; Wagemaker, 2014). L'évaluation PISA a réuni 63 pays participants à son édition de 2009 contre 41 pays en 2003 (Loye, 2011 ; Wagemaker, 2014). Le nombre de pays participants à l'évaluation TIMSS a quant à lui presque doublé entre 1999 et 2011 passant de 40 pays participants à 79 pays (Loye, 2011 ; Wagemaker, 2014). De même, l'évaluation PASEC, qui a connu trois pays à savoir le Congo, Djibouti et le Mali lors de sa première édition en 1994, a connu la participation de dix pays d'Afrique subsaharienne durant son édition de 2014 (CONFEMEN, 2018). Cette évolution en termes de nombre de pays participants semble souligner un consensus au sein des pays à utiliser les résultats des différentes ELE.

Changement de discours politique dans certains pays

Plusieurs résultats aux ELE ont poussé les pouvoirs publics de certains pays à inscrire à l'ordre du jour de leurs discussions des questions sur l'éducation. C'est le cas des résultats des toutes premières évaluations TIMSS, qui ont créé un « scandale » au sein de l'élite politique en Afrique du Sud et poussé les autorités à débattre sur ces résultats (Howie, 2011 dans Wagemaker, 2014). Ces mêmes résultats ont suscité des inquiétudes en Israël et donné lieu à plusieurs titres à la une des journaux, dont le plus célèbre est celui du journal Education Week intitulé : « Down in Rankings : Israel Seeks Changes in Education » (Goldstein, 2004 dans Wagemaker, 2014). Aux États-Unis, Alan Greenspan, alors président de la Réserve fédérale américaine, avait utilisé des résultats de l'évaluation TIMSS en 2004 dans son allocution devant la Commission de l'éducation et du personnel de la Chambre des représentants pour attirer l'attention sur ces résultats (Wagemaker, 2014). En Allemagne, ce sont les résultats à l'évaluation PISA-2000 qui ont provoqué un grand choc dans le pays. La faute à la note du pays qui était inférieure à la moyenne internationale (Gruber, 2006 dans Wagemaker, 2014). Le parlement allemand a dû organiser une session spéciale « PISA » pour débattre sur ces résultats (Gruber, 2006 dans Wagemaker, 2014). À l'Île Maurice comme aux Seychelles, les autorités ont dû mettre sur la table des questions autour du redoublement et du recours à des tuteurs privés à la suite des résultats aux évaluations SACMEQ (Murimba, 2005 dans Wagemaker, 2014). De même, au Sénégal, le ministre de l'Éducation et les membres de son cabinet ont beaucoup échangé sur la question du redoublement après la présentation des résultats PASEC (Bernard et Michaelowa 2006 dans Wagemaker, 2014). Ces différentes réactions ont probablement servi de catalyseurs aux prises de décisions en matière de politiques éducatives dans ces pays.

Modifications des programmes scolaires

Les résultats aux ELE ont eu pour effet la modification des curricula dans plusieurs pays afin de pallier les déficits des résultats des élèves. L'évaluation TIMSS-1995 a servi de catalyseur à la révision des programmes d'études dans les pays comme l'Islande, le Koweït, la Nouvelle-Zélande, la Norvège, la Roumanie ou l'Afrique du Sud (Wagemaker, 2014). En Australie, des programmes compensatoires ont vu le jour pour répondre aux mauvais résultats des élèves à la suite des résultats de l'évaluation PIRLS (Wagemaker, 2014). En Russie, les résultats TIMSS et PIRLS ont conduit à l'élaboration de nouvelles normes éducatives pour le cycle primaire en 2011 (Kovaleva, 2011 dans Wagemaker, 2014). Au Chili, les autorités n'ont pas hésité à se pencher sur le déséquilibre qui existait entre les programmes en mathématiques

et en sciences pour les réformer à la suite des résultats de l'évaluation TIMSS 1999 (Cariola et al., 2011 dans Wagemaker, 2014). Le Japon a procédé également à la révision de ses programmes d'enseignement suite aux conclusions de l'évaluation PISA 2009 (Nakajima, 2010 dans Wagemaker, 2014).

Changements des pratiques pédagogiques

Beaucoup de pays ont réajusté les pratiques pédagogiques courantes dans leurs écoles à la suite des différents résultats aux ELE. En Islande, les résultats des évaluations TIMSS ont conduit à une augmentation des heures d'enseignement en mathématiques et en sciences au niveau primaire (Wagemaker, 2014). En Slovénie, ce sont les résultats aux évaluations TIMSS de 2003 qui ont conduit à la mise en place d'un programme de formation et de développement professionnel des enseignants (Klemencic, 2010 dans Wagemaker, 2014). Même son de cloche aux Philippines où les pouvoirs publics ont procédé au renforcement des capacités de 100 000 enseignants. Au Japon, les résultats PISA-2009 ont servi de catalyseur à une augmentation du nombre de journées d'instruction annuelles pour les écoles et du temps consacré à l'enseignement (Nakajima, 2010 dans Wagemaker, 2014). En Afrique du Sud, les résultats de l'évaluation PIRLS-2006 ont boosté la mise en place d'une campagne d'apprentissage et d'une stratégie de lecture nationale (Howie, 2011 dans Wagemaker, 2014). Ces résultats ont servi également de base à un cadre de politique nationale pour la formation et le développement des enseignants (Howie, 2011 dans Wagemaker, 2014).

Impacts sur la recherche

Les recherches sur les ELE s'intéressent aux influences des caractéristiques individuelles, de l'environnement éducatif, des caractéristiques socioculturelles et économiques sur la qualité des apprentissages et participent au suivi des objectifs de développement du millénaire (Von Davier et al, 2013 ; Wagemaker, 2014). De même, la multiplication de ces programmes d'enquêtes à large échelle a eu également pour effet, l'accroissement de la recherche fondée sur les données produites (Loye, 2011). Ce sont généralement des analyses secondaires qui revisitent les différents outils et méthodologies utilisés dans le cadre de ces évaluations. A travers ces regards rétrospectifs, les chercheurs proposent de nouvelles perspectives théoriques et pratiques en vue d'améliorer la qualité et la validité des données et des mesures produites (Wagemaker, 2014). Cependant, il arrive souvent que les résultats de certaines recherches réfutent l'utilité ou la pertinence de ces enquêtes (Wagemaker, 2014). Les pays s'approprient ces recherches en favorisant le développement de ressources humaines capables de mener ces évaluations ou d'utiliser les données de ces enquêtes. Les résultats des

évaluations TIMSS et PIRLS 2012 ont conduit par exemple à la création de structures nationales pour les évaluations d'apprentissages en Hongrie et en Macédoine (Wagemaker, 2014). C'est également le cas des pays comme l'Afrique du Sud et le Malawi qui ont développé sur le plan local, des compétences en matière de recherche et d'évaluation des apprentissages (Wagemaker, 2014). L'académie russe de l'éducation de son côté, utilise les bases de données TIMSS et PIRLS lors de ses séminaires conçus pour former des coordonnateurs et spécialistes régionaux dans l'évaluation et l'analyse des données des apprentissages (Wagemaker, 2014).

Renforcement du partenariat en éducation

Les stratégies en éducation des grandes organisations de développement et des bailleurs de fonds telles que l'Agence américaine pour le développement international (USAID), le Département britannique pour le développement international (DIFID), la Banque Mondiale (BM), L'Organisation des Nations unies pour l'éducation, la science et la culture (UNESCO), le Partenariat Mondial pour l'Education (PME) et bien d'autres reposent de plus en plus sur les résultats issus des ELE. Ces organisations qui plaident en faveur de la participation d'un plus grand nombre de pays à ces évaluations font la promotion de l'utilisation des résultats qui en découlent dans les négociations en vue des aides qu'ils fournissent aux pays en voie de développement (Wagemaker, 2014). Les demandes de financements que ces pays adressent à ces partenaires techniques intègrent le plus souvent, leurs résultats aux différentes évaluations (Wagemaker, 2014).

Les différents impacts des ELE que nous venons de présenter doivent néanmoins être nuancés. En effet, les prises de décisions politiques tiennent compte de plusieurs autres facteurs et aléas et ne sauraient être imputables aux seuls résultats issus des ELE (Wagemaker, 2013). Cependant, il ne demeure pas moins que les ELE constituent, au regard de ce qui vient d'être présenté, des enjeux importants pour les pays et leur système éducatif. La tâche qui consiste à fournir aux décideurs des éléments d'aide à la décision, basés sur l'évaluation des apprentissages, est cependant ardue. La mesure des performances des individus dans le cadre des ELE repose principalement sur des attributs psychologiques observables indirectement, complexes, et cette mesure peut être instable dans le temps et l'espace (Bertrand et Blais, 2004). Cette tâche devient dès lors délicate et sujette, le plus souvent, à des critiques. Dans la section suivante, nous allons donc porter notre attention sur ce qui est considéré comme la mesure des performances aux ELE et souligner quelques limites sur cette approche.

1.2.2 La mesure des performances dans les ELE

Mesure des attributs psychologiques

Grâce aux succès de la recherche quantitative durant des siècles, la mesure est devenue pour beaucoup, l'essence même de la science (Bertrand et Blais, 2004). Dans la perspective pythagoricienne par exemple, tout est mesurable et tout doit être mesuré (Bertrand et Blais, 2004). Les différentes définitions de la mesure vont cependant connaître deux tendances : une tendance qui soutient que la quantité est essentiellement empirique et une autre qui soutient que les nombres ont une existence au-delà du monde réel (Bertrand et Blais, 2004). Dans le premier groupe, on retrouve Aristote et Euclide pour qui, mesurer c'est déterminer le rapport entre deux quantités (Bertrand et Blais, 2004). Cette définition est d'ailleurs à l'origine de la définition dite « classique » de la mesure qui considère que mesurer consiste à estimer ou découvrir le rapport entre la grandeur d'un attribut quantitatif et une unité de cette même grandeur (Bertrand et Blais, 2004). Cette première posture est toutefois plus adaptée à la mesure des attributs physiques tel que la taille, le poids, la température, et bien d'autres.

La deuxième tendance se développera au courant du 20^e siècle et connaîtra plus de succès avec les travaux de Campbell et notamment ceux de Stevens (Bertrand et Blais, 2004). Pour Campbell d'abord, la mesure consiste à représenter par des nombres des propriétés d'objets ou de phénomènes et des relations entre ces objets. Stevens reprendra la définition de Campbell et posera les jalons d'une définition plus souple de la mesure. Pour ce dernier, mesurer consiste à assigner des nombres à des objets ou à des phénomènes selon des règles (Stevens, 1951 dans Bertrand et Blais, 2004). Avec cette nouvelle perspective, le plus important est de mettre en place des règles pour définir ce que représentent les nombres et la façon de les attribuer (Bertrand et Blais, 2004). Les différentes définitions de la mesure en sciences sociales et en éducation vont beaucoup s'inspirer de celle de Stevens. Des auteurs dans ces domaines précisent que l'assignation de ces nombres peut toucher des propriétés des individus tels que l'anxiété, l'attitude ou l'habileté dans une discipline à l'instar des mathématiques et bien d'autres (Torgerson, 1958 dans Bertrand et Blais, 2004). Cette assignation doit également transformer les relations entre les propriétés des individus (X est plus habile en langue que Y) sous forme de relations entre les nombres à l'instar de $74\% > 55\%$ (Gruijter et van der Kamp, 1984 dans Bertrand et Blais, 2004). Cependant, le processus de mesure des attributs psychologiques tel que l'habileté en mathématique ou en langue d'un individu ne se limite pas uniquement à cette activité d'assignation de nombre. Bertrand et Blais (2004) identifient 6 étapes indispensables au processus de mesure des attributs psychologiques.

La première étape commence donc par le choix du ou des construits qui seront mesurés comme le présente la figure 1. Un construit représente la caractéristique que le test est censé mesurer (American Educational Research Association et al, 2003). Il peut correspondre par exemple, à l'habileté à résoudre un certain type d'exercice en mathématiques, ou à lire un texte (Bertrand et Blais, 2004). Un construit est théorique et est généralement le résultat d'une concertation entre professionnels de la mesure (American Educational Research Association et al, 2003 ; Bertrand et Blais, 2004). Ceux-ci déterminent et fixent ce qu'ils souhaitent mesurer (Bertrand et Blais, 2004). Cependant, rien ne garantit que ce construit existe ou qu'il existera à la fin du processus de mesure (Bertrand et Blais, 2004). L'étape 2 de spécifications est une étape qui consiste à cibler des domaines qui pourraient matérialiser ces construits à mesurer (Bertrand et Blais, 2004). Dans les ELE, il s'agit de mettre en place les cadres de références qui soutiendront les futurs tests (Loye, 2011). Ces cadres de référence cibleront non seulement les domaines et compétences qui seront évalués, mais peuvent également indiquer comment les tests seront administrés et à quel moment (Loye, 2011).

L'étape numéro 2 concerne la rédaction des items en cohérence avec le cadre de référence rédigé (Bertrand et Blais, 2004). Cette étape aboutit généralement à la mise en place d'une banque d'items (Bertrand et Blais, 2004). Les professionnels choisissent et administrent un échantillon de ces items aux individus ciblés pour participer à ces tests (Bertrand et Blais, 2004). A ce stade, les ELE sont souvent confrontées au défi d'évaluer des domaines de sujets larges à travers des tests administrés en 1 heure ou 2 heures (Von Davier et Sinharay, 2013). Pour y faire face, elles utilisent plusieurs blocs de tests liés entre eux par des items d'ancrages et chaque groupe de participants n'est évalué que sur une partie du contenu du cadre (Von Davier et Sinharay, 2013). Généralement, avant une administration grandeur nature des tests, les ELE procèdent à des prétests dans le but de s'assurer que les instruments et les procédures de passation fonctionnent bien (Loye, 2011). L'étape 4 symbolise l'étape de consolidation des différentes réponses des participants (Bertrand et Blais, 2004). Les trois dernières étapes consistent à procéder à un ensemble d'analyses qui vont aboutir à la construction d'échelle ou à l'assignation de scores sensés refléter le construit qui est mesuré (Bertrand et Blais, 2004).

Ces trois dernières étapes de la partie du processus de mesure, qui se font à l'aide de modèles de mesures, sont des étapes sur lesquelles nous nous attarderons le plus dans le cadre de ce travail ; l'instrumentation et la passation des tests étant des domaines assez vastes qui font également l'objet de plusieurs recherches. La section suivante présente donc les modèles de mesures utilisées dans le cadre des ELE.

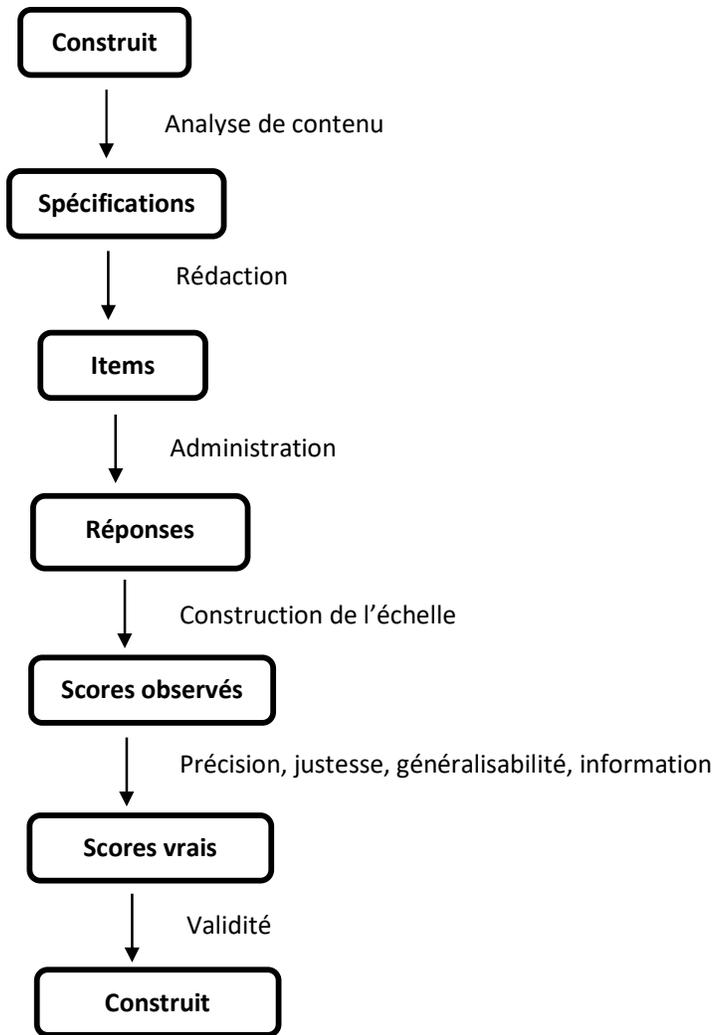


Figure 1 : Processus de mesure des attributs psychologiques (Bertrand et Blais, 2004, p.32)

Modèles de mesure utilisés dans les ELE

Dans le cadre des ELE, les participants fournissent, en remplissant un ou plusieurs questionnaires, une variété d'informations en plus des réponses aux questions du test (Loye, 2011 ; Von Davier et Sinharay, 2013). Ce questionnaire permet de recueillir des informations dites covariables sur l'environnement scolaire et extrascolaire des participants, leurs activités académiques et non académiques, leurs attitudes, leurs impressions ainsi que des variables démographiques les concernant (Loye, 2011 ; Von Davier et Sinharay, 2013). Le traitement de cet ensemble de données se fait en deux temps à l'aide de modèles de mesures distincts (Von Davier et Sinharay, 2013). Dans un premier temps, un modèle estime les habiletés ou performances des participants, puis un autre modèle intègre les covariables aux résultats pour

décélérer des éventuels liens entre ces covariables et les performances aux tests (Von Davier et Sinharay, 2013).

Le dictionnaire Legendre (2005) définit un modèle comme une représentation simplifiée d'un phénomène pour mieux l'étudier. C'est le cas des modèles d'échelles, des modèles physiques, des modèles de jouets ou des modèles de mesure (Magnani et Bertolotti, 2017). En psychométrie, un modèle de mesure a pour objectif d'encadrer la construction d'un instrument de mesure d'un attribut psychologique, tel que l'habileté en lecture ou l'attitude envers l'école, en assignant à un individu une valeur qui reflète la quantité de cet attribut qu'il possède (Bertrand et Blais, 2004). La mesure des attributs psychologiques se fait généralement à l'aide de trois modèles de mesure : (i) le modèle de la théorie classique, (ii) le modèle de la théorie de la généralisabilité et (iii) le modèle de la théorie des réponses à l'item (TRI) (Bertrand et Blais, 2004). Ces trois modèles de mesures sont cependant essentiellement axés sur les processus mentaux qui se rapportent aux fonctions de connaissance, à la mémoire, au raisonnement, à la prise de décision et autres aspects liés à la cognition (Von Davier et Sinharay, 2013). Ainsi, ils ne prennent pas en considération des facteurs externes dans la mesure d'un construit (Von Davier et Sinharay, 2013).

Dans la philosophie cognitiviste, les connaissances antérieures d'un individu détermineraient ce que cet individu peut apprendre et seuls son système cognitif et une organisation propre lui permettraient de restituer ces connaissances (Wambach, 2003). Cependant, l'idée d'un individu intériorisant des connaissances dans lesquelles il puiserait pour résoudre des problèmes s'efface progressivement pour laisser place à celle d'un individu qui interagirait avec son milieu et qui mobiliserait des ressources personnelles et externes pour résoudre ces problèmes (Mondada et Pekarek 2000). Dans cette perspective, les processus mentaux sont des processus sociaux situés, liés ainsi à des valeurs culturelles ou institutionnelles et des possibles interactions entre les individus et leur environnement (Mondada et Pekarek 2000). Cette dimension sociale et située des performances, pose souvent des défis importants aux programmes d'enquêtes standardisés et soulève des critiques au sein de la communauté scientifique (Sireci, 2011). La section suivante présente donc certaines des limites qui pourraient justifier l'idée de redéfinir la performance dans les ELE et par là, justifier l'utilisation de modèles de mesures qui intégreraient, au moment de la mesure de la performance, des données individuelles et contextuelles.

Quelques limites liées à la mesure des performances dans les ELE

Les ELE se déroulent dans des pays et des régions où résident des populations de culture, de langue et de situations économiques différentes (Sireci, 2011). De même, chacun de ces pays offre des cadres d'apprentissages scolaires à sa population en fonction de sa propre politique en matière d'éducation (Sireci, 2011). Dans la littérature, plusieurs écrits interrogent la capacité de ces évaluations à fournir des tests et des résultats qui tiennent compte de toute cette diversité (Zumbo, 2007). Une partie de cette littérature se développe autour des problématiques liées aux questions d'équité, à l'analyse du fonctionnement différentiel des items et à l'invariance de la mesure (Zumbo, 2007). Dans le meilleur des mondes, des individus de niveau d'habileté similaires devraient avoir les mêmes chances de réussir à un test indépendamment de leurs caractéristiques individuelles et de leur milieu de vie (American Educational Research Association et al, 2003 ; Sireci, 2011 ; Zumbo, 2007). En psychométrie, lorsque cette hypothèse est vérifiée dans les résultats d'un test, on parle alors de l'invariance de la mesure, d'absence de biais ou d'absence de fonctionnement différentiel des items (American Educational Research Association et al, 2003 ; Sireci, 2011 ; Zumbo, 2007). Dans une telle situation, il est donc équitable de comparer les résultats obtenus entre les différents groupes de participants (American Educational Research Association et al, 2003 ; Sireci, 2011 ; Zumbo, 2007). Cet idéal est cependant difficilement atteint à cause de la diversité qui entoure les tests.

Des recherches montrent par exemple que des versions traduites d'un même test, en raison des spécificités linguistiques propres aux pays ou à des régions, fonctionnent différemment (Ercikan,1998). C'est le cas des tests PIRLS et TIMSS en langue et en sciences qui favoriseraient des élèves canadiens anglophones comparativement à leurs camarades francophones (Ercikan,1998). C'était également le cas pour les tests PISA de 2003 entre ces mêmes groupes d'élèves (Oliveri et al, 2012). Ces différentes recherches soutiennent que des versions différentes d'un même test pourraient évaluer différents construits à cause des biais liés à la traduction et compromettre ainsi la comparabilité des résultats entre pays (Ercikan,1998 ; Oliveri et al, 2012). Dans le même sillage, des études montrent que les différences en termes de contenu de programmes scolaires au sein des pays pouvaient entraîner un fonctionnement différentiel sur des items (Ercikan, 2002 ; Ercikan et Koh, 2005). C'est cela que Ercikan (2002) et Ercikan et Koh (2002) ont pu montrer à travers les résultats aux tests de mathématiques et de sciences de TIMSS de 1995. Leurs études avaient comparé les résultats des élèves canadiens à ceux de leurs camarades en France et ceux des élèves anglais à ceux des élèves américains (Ercikan, 2002 ; Ercikan et Koh, 2005). Les conclusions de ces études

soulignent une sérieuse limite à comparer les performances des élèves de ces différents pays (Ercikan, 2002 ; Ercikan et Koh, 2005).

Les différences culturelles et les pratiques pédagogiques dans les pays s'avèrent être également des entraves aux comparaisons internationales des résultats aux ELE (Huang et al, 2016). Huang et ses collègues (2016) ont pu montrer que ces facteurs imposaient des limites à comparer les résultats aux tests PISA de 2006 entre des pays comme le Canada, les Etats Unis, Hong Kong et la chine. A cause des différences culturelles et des pratiques pédagogiques propres à ces pays, les résultats des tests PISA 2006 semblaient mesurer différents construits en fonction de ces pays (Huang et al, 2016). Asil et Brown (2016) ont également mis en évidence le fait que le niveau de développement économique des pays pouvait remettre en cause les comparaisons des résultats aux tests PISA de 2009 entre des pays de l'OCDE. Cet élément semblait jouer un rôle dans le fonctionnement différentiel de certains items de ce test (Asil et Brown, 2016).

Ces quelques exemples, parmi tant d'autres, soulignent la difficulté pour les résultats aux tests des ELE de tenir compte de la diversité des contextes dans lesquels évoluent les différents participants aux tests (Sireci, 2011). Pourtant, ces mêmes résultats comme nous l'avons exposé dans les sections précédentes, s'imposent de plus en plus dans les pays comme des outils indispensables d'aide à la décision en matière de politiques éducatives (Loye, 2011 ; Wagemaker, 2014). De même, les impacts possibles de ces résultats sur les réformes curriculaires et pédagogiques, sur la recherche et sur le renforcement des partenariats en éducation dans ces pays sont considérables (Wagemaker, 2014). De telles implications poussent aujourd'hui les chercheurs à envisager d'autres avenues pour réduire davantage, les problèmes liés à l'estimation des performances dans le cadre de ces évaluations. Les avancées de la recherche en sciences sociales et en sciences de l'éducation facilitent aujourd'hui ce travail, car elles fournissent des outils théoriques et pratiques précieux pour appréhender une performance à un test dans toute sa complexité (Mislevy, 2012).

Pour notre travail, nous avons fait le choix de regarder du côté des « écologistes ». L'approche écologique est une approche qui permet non seulement de comprendre le processus d'apprentissage chez un individu, mais également d'expliquer le processus par lequel il restitue ses acquis (Zumbo et al, 2015). Dans ce sens, elle permet de conceptualiser la performance à un test comme étant le produit à la fois des capacités cognitives des individus, de leurs caractéristiques individuelles, et des caractéristiques sociales et contextuelles de l'environnement dans lequel ils évoluent (Zumbo et al, 2015). Ce point de vue met ainsi en

exergue les dimensions sociale et située des performances à un test (Zumbo et al, 2015). Nous présenterons donc dans la partie suivante cette approche. Mais avant de souligner les éclairages que celle-ci fournit pour comprendre les mécanismes qui entrent en jeu dans la réalisation d'une performance à un test, nous allons revenir brièvement sur l'écologie du développement humain du psychologue américain Bronfenbrenner. L'approche de Bronfenbrenner est en effet une lentille utile pour comprendre différents changements qui interviennent chez un individu à l'instar du développement de sa cognition (Zumbo et al, 2015). Cette approche est également la principale source d'inspiration des « écologistes » de la mesure (Zumbo et al, 2015).

1.3 L'approche écologique

Chaque être humain se développe et s'épanouit au sein d'une communauté disposant d'une certaine organisation et des structures d'activités identifiables et récurrentes appelées pratiques (Mislevy, 2018). Ces pratiques couplées aux caractéristiques innées de chaque individu et aux caractéristiques environnementales le façonnent dans sa manière de faire, d'agir et de penser (Mislevy, 2018 ; Papalia et al, 2010). Ces changements quantitatifs et qualitatifs qui s'opèrent de façon dynamique, cohérente et organisée durant toute la vie de l'individu prennent le nom de « développement humain » (Papalia et al, 2010). Ces métamorphoses peuvent être de trois ordres : (i) physiques en lien avec le corps, le cerveau, les capacités sensorielles et/ou motrices ; (ii) affectifs et sociaux liés aux émotions, à la personnalité et aux relations avec les autres et (iii) cognitifs relatifs aux habiletés mentales, à la perception, à l'apprentissage, à la mémoire, au langage, au raisonnement ou encore à la créativité (Papalia et al, 2010). Plusieurs approches se sont intéressées au développement humain pour mieux comprendre et expliquer ces différents changements qui interviennent (Papalia et al, 2010).

Les différentes théories issues de ces approches se regroupent en cinq grands courants à savoir : (i) l'approche psychanalytique ; (ii) l'approche behavioriste et néo-behavioriste ; (iii) l'approche cognitiviste ; (iv) l'approche humaniste et (v) l'approche écologique (Papalia et al, 2010). Aucun de ces courants n'explique entièrement le développement humain compte tenu de sa complexité (Papalia et al, 2010). Toutefois, chaque courant jette, en fonction de ses spécificités, un regard particulier sur certains pans de ce développement (Papalia et al, 2010). Le tableau 1 réunit chaque approche, ses principales théories et les postulats de base de ces théories. Nous développons brièvement, dans la section suivante, les postulats des modèles écologiques qui nous intéressent particulièrement.

Tableau 1 : Les cinq approches théoriques du développement (Papalia et al, 2010, p.12)

Approche	Principale théorie	Postulat de base
Psychanalytique	Théorie psychosexuelle (Freud)	Le comportement est contrôlé par des pulsions inconscientes puissantes.
	Théorie psychosociale (Erikson)	La personnalité est influencée par la société et se développe en passant par différentes crises ou alternatives critiques.
Béavioriste et néo-béavioriste	Béaviorisme ou théorie traditionnelle de l'apprentissage (Pavlov, Skinner, Watson)	Les individus réagissent aux stimuli de l'environnement ; le comportement est contrôlé par l'environnement.
	Théorie de l'apprentissage social ou sociocognitive (Bandura)	Les enfants apprennent dans un contexte social donné, en observant et en imitant des modèles, et contribuent activement à leur apprentissage.
Cognitiviste	Théorie cognitiviste (Piaget)	Des changements qualitatifs de la pensée se produisent dans l'enfance et l'adolescence. L'enfant initie activement son propre développement qui va dans le sens d'une adaptation croissante.
	Théorie socioculturelle (Vygotsky)	Les interactions sociales sont au centre du développement cognitif de l'enfant.
	Théorie du traitement de l'information	Les êtres humains sont en quelque sorte des processeurs de symboles.
Humaniste	L'actualisation de soi (Rogers)	L'être humain tend vers le développement de son plein potentiel en étant authentique et congruent.
	La pyramide des besoins (Maslow)	Les besoins de base doivent être satisfaits avant qu'une personne n'atteigne l'actualisation de soi.
Écologique	Modèles écologiques (Bronfenbrenner)	L'individu se développe en fonction de six systèmes d'influence qui sont en interaction constante.

1.3.1 L'écologie du développement humain

L'approche écologique est une approche qui s'inspire de l'écologie en biologie. L'écologie stipule en effet qu'un organisme se développe à l'intérieur d'un système qui peut favoriser ou empêcher sa croissance (Papalia et al, 2010). Pour les « écologistes » en psychologie, cette réalité est applicable à l'être humain ; faisant ainsi de son développement physique et mental, le produit d'un ensemble de facteurs (géographiques, culturels, familiaux, politiques, économique, etc.) et des liens qui peuvent exister d'un côté entre l'individu et ces facteurs, mais également des liens possibles entre ces différents facteurs (Malo, 2000 ; Papalia et al, 2010). La théorie écologique de Bronfenbrenner (1979) s'inspire des questions sur le développement de la personnalité de Vygotsky et de Lewin. Cette théorie accorde une grande importance à l'influence du contexte sur le développement de l'individu et reste attentive aux effets des changements de milieu sur celui-ci (Absil et al, 2012). Elle se fonde ainsi sur l'idée centrale selon laquelle l'environnement, au sens large, a une emprise sur le développement de

l'être humain (Absil et al, 2012 ; Bronfenbrenner, 1994 ; Hayes et al, 2017 ; Malo, 2000). La théorie de Bronfenbrenner (1979) suggère de considérer une approche systémique dans laquelle quatre types de milieux (microsystèmes), interagiraient entre eux pour concourir au développement de l'être humain (Absil et al, 2012).

Le modèle écologique du développement humain

La représentation conceptuelle de la théorie de Bronfenbrenner se fait généralement à l'aide un modèle à plusieurs couches comme le montre la figure 2 (Papalia et al, 2010). La première couche de ce modèle est l'ontosystème qui représente l'individu avec toutes ses caractéristiques personnelles (Hayes et al, 2017 ; Malo, 2000 ; Papalia et al, 2010). Ces dernières englobent ce qui est inné ou acquis dans le temps chez l'individu. C'est notamment son bagage génétique, ses habiletés intellectuelles ou ses valeurs sociales (Hayes et al, 2017 ; Malo, 2000 ; Papalia et al, 2010). Ensuite vient le microsystème qui représente l'environnement immédiat dans lequel se trouve l'individu et où il a une participation active et directe (Hayes et al, 2017 ; Malo, 2000 ; Papalia et al, 2010). Cet environnement contient des personnes, des lieux physiques ou des objets qui s'y trouvent ou encore des activités qui peuvent s'y dérouler ou des rôles qu'il peut y tenir (Hayes et al, 2017 ; Malo, 2000 ; Papalia et al, 2010).

L'exosystème est la troisième couche de ce modèle. Contrairement au microsystème, l'exosystème désigne l'environnement contextuel dans lequel l'individu n'est pas directement impliqué, mais qui a néanmoins un impact dans sa vie (Hayes et al, 2017 ; Malo, 2000 ; Papalia et al, 2010). En plus de référer aux lieux physiques, aux personnes et objets qu'ils contiennent, et aux activités et rôles qui s'y déroulent, l'exosystème compte également toutes les décisions qui sont prises dans cet environnement (Malo, 2000). La quatrième couche, le macrosystème, inclut tout ce qui est culture, croyance, idéologie, valeurs politique ou économique véhiculées dans une société (Hayes et al, 2017 ; Malo, 2000 ; Papalia et al, 2010).

Le mésosystème, la couche 5, est la couche qui renferme les différents liens et interactions qui peuvent exister entre deux ou plusieurs microsystèmes (Hayes et al, 2017 ; Malo, 2000 ; Papalia et al, 2010). Le chronosystème quant à lui, représente les influences liées au temps à l'exemple de l'effet que pourrait avoir l'expérience sur le comportement futur d'un individu ou l'effet de l'âge sur ses capacités cognitives (Malo, 2000).

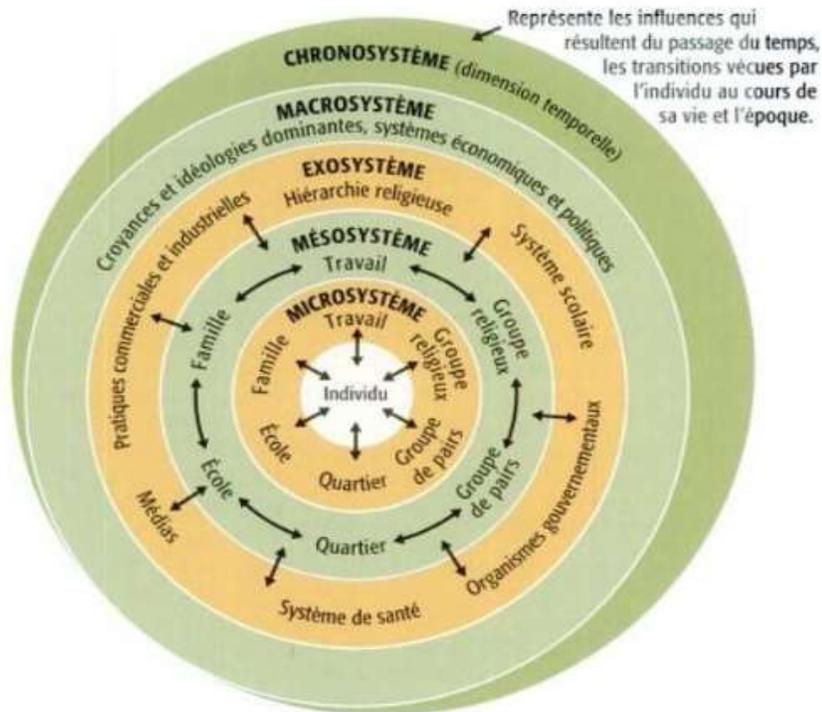


Figure 2 : Le modèle écologique du développement humain de Bronfenbrenner (Papalia et al, 2010, p.27)

Usages du modèle écologique du développement humain

Le modèle écologique du développement humain trouve son application principalement en psychologie. Il permet de comprendre certains comportements ou certains troubles qui pourraient survenir dans le développement des individus notamment chez les enfants (Absil et al, 2012). Des résultats dans ce domaine nous apprennent par exemple qu'un enfant ayant à sa naissance des prédispositions désavantageuses et qui grandit dans un environnement riche en opportunités pourrait se développer de façon harmonieuse (Absil et al, 2012 ; Malo, 2000 ; Papalia et al, 2010). De même, un enfant né avec un potentiel élevé peut développer certains troubles s'il grandit dans un environnement qui comporte un niveau élevé de risques (Absil et al, 2012 ; Malo, 2000 ; Papalia et al, 2010). De telles recherches suggèrent que le développement physique ou cognitif chez un individu ne serait pas un acte isolé, mais plutôt, le résultat d'un ensemble de facteurs en constante interactions (Absil et al, 2012 ; Malo, 2000 ; Papalia et al, 2010).

Le modèle de Bronfenbrenner trouve également des applications dans le domaine de la santé. Des chercheurs s'en servent pour notamment dresser des diagnostics plus fiables en tenant compte d'un maximum de facteurs et des possibles interactions entre ces facteurs (Absil et al, 2012). Ils s'en servent également pour élaborer des stratégies plus efficaces de lutte contre des maladies (Absil et al, 2012). A l'aide de ce modèle, les chercheurs en médecine arrivent à mieux cerner les acteurs, leurs milieux et les interactions qui favoriseraient ou réduiraient la propagation d'une maladie (Absil et al, 2012).

En psychométrie, dans le domaine de la mesure, l'idée d'une approche écologique fait également son chemin. Dans ses travaux par exemple, Fox insiste sur la nécessité de tenir compte des conditions de passation des tests pour estimer les performances des participants (2003). Ses travaux mettent en évidence l'existence d'une possible interaction évaluateurs-participants-tâches lors d'un test et celle-ci participerait aux choix des réponses des différents participants (Fox, 2003). McNamara (1997, 2007) et McNamara et Roever (2006) suggèrent également de regarder au-delà des réponses aux questions d'un test pour estimer les performances. La prise en compte du contexte social des participants au moment de cette estimation fournirait un aperçu plus juste de leurs performances (McNamara, 1997 ; McNamara, 2007 ; McNamara et Roever, 2006). Les travaux de Fox (2003), McNamara (1997, 2007) et McNamara et Roever (2006) rassemblent ainsi des preuves qui montrent que les processus cognitifs qui permettent à un candidat de choisir une réponse à un item ne sont pas isolés.

Transition vers l'écologie de la réponse à l'item

En mesure, il ne suffit pas seulement de rassembler des preuves comme le font Fox (2003), McNamara (1997, 2007) et McNamara et Roever (2006). Il faudrait également mettre en place un modèle écologique qui montre clairement comment des facteurs contextuels pourraient affecter les réponses à des items en médiant les processus cognitifs normalement supposés générer ces réponses (Zumbo et al, 2015). Et c'est justement à ce travail que vont s'atteler Zumbo et ses collègues (2015). Ces derniers adhèrent aussi au principe de l'écologie de Bronfenbrenner qui stipule que les êtres humains, dès leur naissance, évoluent et s'imprègnent d'une culture ; cette culture et l'environnement naturel où ils évoluent affectent leur psychologie et leur cognition dans un système d'interactions. Pour Zumbo et ses collègues (2015), l'analogie avec l'écologie de la réponse à l'item est que les candidats à un test apportent également un bagage social, culturel et historique pour passer ce test. Ce bagage, en plus des processus cognitifs participerait donc au choix d'une réponse à un item (Zumbo et al, 2015).

Avec une telle posture, Zumbo et ses collègues (2015) invitent les praticiens de la mesure à se détacher d'une structure binaire dans laquelle chaque variable contextuelle aurait de façon isolée un effet sur les performances à un test vers une forme contextualisée des performances.

1.3.2 L'écologie de la réponse à l'item

Le modèle écologique de Zumbo et ses collègues (2015) repose sur un ensemble de cinq couches comme le montre la figure 3. Ces différentes couches concourent simultanément au choix par un individu d'une réponse à un item contribuant ainsi à l'estimation de sa performance (Zumbo et al, 2015). Ces différentes couches représentent : (i) le format et le contenu du test, (ii) les caractéristiques individuelles, (iii) l'environnement scolaire, (iv) l'environnement extra-scolaire et, (v) les caractéristiques de la communauté (Zumbo et al, 2015).

Le format et le contenu d'un test

Cette couche renferme les différentes caractéristiques d'un test en termes de format et de contenu. Ces deux aspects d'un test jouent en effet un rôle important dans le choix des réponses des candidats (Zumbo et al, 2015). Des individus peu familiers avec le format des questions d'un test, ou les termes et expressions contenues dans ces questions, auraient de faibles probabilités à fournir de bonnes réponses (McNamara, 2007 ; McNamara et Roever, 2006). Pourtant, ces individus ne seraient pas pour autant moins habiles que ceux qui fourniraient de bonnes réponses à ces questions (McNamara, 2007 ; McNamara et Roever, 2006). Des versions traduites d'un même test pourraient également mesurer différents construits (Zumbo et al, 2015). Les biais liés à la traduction sont assez récurrents et pourraient compromettre l'idée de comparer des individus qui passent différentes versions d'un même test lorsque cet aspect n'est pas pris en compte (Ercikan, 1998 ; Ercikan, 2002 ; Ercikan et Koh, 2005).

Les caractéristiques individuelles

La deuxième couche du modèle écologique de Zumbo et ses collègues (2015) repose sur les caractéristiques individuelles des participants aux tests. Pendant longtemps, les caractéristiques individuelles ont en effet focalisé l'attention des chercheurs sur les disparités au niveau des résultats des tests (Zumbo et al, 2015). Plusieurs travaux de recherches indexent le sexe comme un élément justifiant les performances des individus (Martinková et al, 2017 ; Ryan et Chiu, 2001 ; Taylor et Lee, 2012). Cependant, certains auteurs réfutent de plus en plus cette idée d'une performance à un test qui serait liée au sexe (Zumbo et al, 2015). Ces derniers se tournent vers d'autres facteurs tel que le statut socio-économique (Willms, 2003), la maîtrise

de la langue d'enseignement (Mc Andrew et al, 2008), la personnalité, l'état de santé, le niveau de concentration ou encore le parcours scolaire (Papalia, et al, 2010). Ces différents éléments justifieraient les performances des individus aux tests (Zumbo et al, 2015).

L'environnement scolaire

L'environnement scolaire du modèle écologique de Zumbo et ses collègues (2015) est constitué d'éléments liés à la salle de classe, à l'enseignement, et à l'école. Cependant, dans les approches multi niveaux, une distinction est souvent faite entre les facteurs niveau classe et les facteurs niveau école (Luke, 2004). Les facteurs liés aux salles de classe seraient nichés dans des écoles et donc agissant à un niveau inférieur (Luke, 2004). Beaucoup d'études qui s'intéressent aux performances scolaires révèlent des effets significatifs qu'auraient les ressources d'une école et d'une salle de classe sur les performances des élèves (Willms, 2003). Ces ressources peuvent être pédagogiques, matérielles ou humaines (Willms, 2003). De même, dans une recension récente d'écrits, Poulin et ses collègues (2015) indiquent que beaucoup d'études lient la performance scolaire au climat qui prévaut dans les établissements scolaires. Des élèves victimes de violences de la part de leurs pairs auraient de faibles performances scolaires (Poulin et al, 2015).

L'environnement extrascolaire

L'environnement extrascolaire fait référence à la famille de l'individu, aux gens et aux choses qu'il côtoie quotidiennement en dehors du contexte scolaire (Zumbo et al, 2015). Dans son étude, Koné (2007) montre que le cadre familial a un impact sur les performances aux tests de langues pour des élèves allophones arabes et créoles vivant à Montréal. Elle souligne également que la pratique de la langue d'enseignement dans un contexte familial et le soutien scolaire des proches ont des effets positifs sur les résultats aux tests de langues des élèves (Koné, 2007). A l'inverse, ceux des élèves n'ayant pas ce soutien ou qui ne pratiquent la langue d'enseignement que dans un contexte d'apprentissages auraient de faibles performances (Koné, 2007).

Les caractéristiques de la communauté

La dernière couche du modèle écologique de réponse à l'item que proposent Zumbo et ses collègues (2015) est constituée des caractéristiques de la communauté. La communauté peut faire référence au voisinage, à un groupe social, ou de façon plus large à des groupes d'individus vivant dans un même pays (Zumbo et al, 2015). Cette couche renferme également toute les règles qui régissent le fonctionnement de cette communauté (Zumbo et al, 2015). Willms

(2003), soulignait justement que le statut socioéconomique moyen d'une collectivité aurait un effet sur les performances scolaires des individus de cette collectivité. Asil et Brown (2016) montrent également que le niveau de développement économique des pays aurait un impact sur les performances des élèves aux tests de lecture de PISA 2009.

Le modèle écologique de Zumbo et ses collègues (2015) essaye donc de tenir compte du maximum de facteurs qui concourent à générer la réponse à un item. Ils préconisent également la prise en compte de tous ces facteurs au moment d'estimer la performance à un test (Zumbo et al, 2015). De ce fait, ils se détachent de la posture qui consiste à estimer les performances uniquement sur la base des réponses des participants et à les expliquer ensuite à l'aide des facteurs contextuels. Pour notre recherche, nous nous alignons sur la posture de Zumbo et ses collègues (2015). C'est cette posture qui guidera nos différentes hypothèses et qui nous aidera à répondre à notre question de recherche.

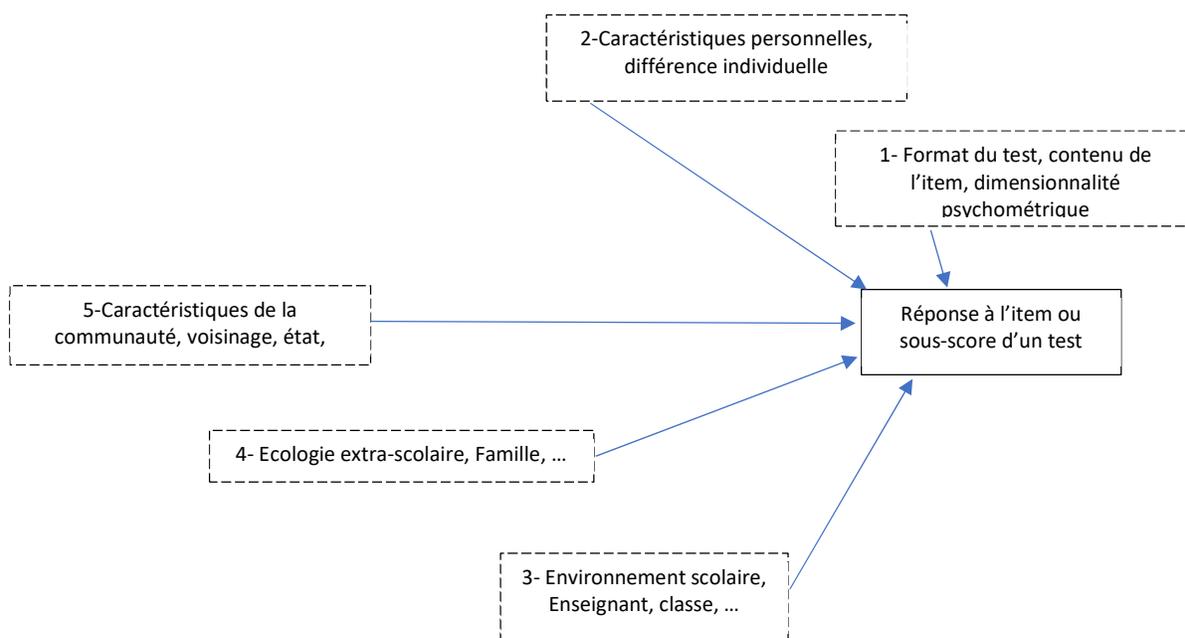


Figure 3 : Modèle écologique de la réponse à l'item (Zumbo et al, 2015, p.5)

1.4 Question générale de recherche

A travers les sections précédentes, nous avons montré que les ELE occupent une place de choix dans les instruments de décision pour les politiques éducatives au sein de beaucoup de pays (Loye, 2011 ; Wagemaker, 2014). Dans la plupart des cas, les ELE utilisent des modèles de mesure qui ne tiennent compte que des réponses des participants pour mesurer leurs performances aux tests (Bertrand et Blais, 2004 ; Von Davier et Sinharay, 2013). Cette approche présente cependant des limites qui pourraient remettre en cause la comparabilité des résultats entre groupes d'individus (Asil et Brown, 2016 ; Ercikan, 1998 ; Ercikan, 2002 ; Ercikan et Koh, 2005 ; Huang et al, 2016 ; Oliveri et al, 2012 ; Sireci, 2011 ; Zumbo, 2007). Toutefois, nous avons également montré que des avancées en recherche dans les domaines des sciences sociales et de l'éducation nous permettent aujourd'hui d'envisager d'autres perspectives. Les résultats de ces recherches suggèrent de considérer une performance comme un phénomène qui se réalise dans un réseau interconnecté de connaissances, de caractéristiques individuelles et de contextes particuliers (Mislevy, 2018 ; Zumbo et al, 2015). La perspective écologique de Zumbo et ses collègues (2015), qui est fort intéressante, s'inscrit dans cette logique.

De même, comme nous l'avons souligné à l'entame de ce travail, au Cameroun, les résultats issus des différents tests des ELE sont de plus en plus utilisés pour le pilotage du système éducatif (MINEPAT, 2013). C'est le cas de l'évaluation en littérature, en nombres et en culture générale du MLA I durant l'année 1992, en mathématiques, en sciences et en culture générale pour le MLA II de 2003 et en langue d'enseignement et en mathématiques pour les évaluations PASEC-2007, SOFRECO-2011 et PASEC2014 respectivement durant les années 2004, 2011 et 2014 (Chinapah, 2003 ; PASEC, 2007 ; PASEC, 2015 ; SOFRECO, 2011). Les recommandations issues de ces évaluations ont conduit les responsables du système éducatif camerounais entre autres à recruter plus d'enseignants au primaire, notamment dans des zones défavorisées, à renforcer le réseau scolaire par des constructions de salles de classe et à distribuer gratuitement des manuels essentiels aux élèves du primaire (MINEPAT, 2013). De tels impacts pour un système éducatif poussent nécessairement la communauté scientifique à s'interroger sur la qualité des instruments de collecte, le choix des modèles de mesure pour estimer les performances des élèves, la validité des résultats, voir sur la nécessité même à utiliser de telles évaluations (Pons, 2011 ; Wagemaker, 2014).

Le choix que nous avons fait pour notre travail est celui de nous attarder sur le processus de mesure des performances aux ELE, notamment à celui du test de lecture PASEC2014 puis, de rechercher et de proposer une approche qui pourrait être plus pertinente pour un tel

processus. L'hypothèse que nous posons est qu'un processus de mesure qui intègre à la fois des réponses aux items et des données contextuelles pour estimer les performances à un test pourrait fournir une lecture fort intéressante des résultats et suggérer de nouvelles pistes pour les prises de décisions. L'objectif principal de notre recherche sera donc de répondre à la question générale suivante :

Comment la mesure des performances qui intègre une approche écologique influence-t-elle les résultats du test en lecture du PASEC2014 ?

Répondre à cette question nécessitera, bien évidemment, d'opérationnaliser la mesure dans une approche écologique. A cette préoccupation, nous retenons la proposition de Zumbo et ses collègues (2015) qui nous disent de nous tourner du côté du modèle d'analyse des classes latentes avec covariables (ACL). Comme nous le verrons dans la section suivante, le modèle ACL avec covariables, qui est une combinaison du modèle ACL et du modèle de régression logistique multinomiale, permet d'implémenter cette approche écologique (Collins et Lanza, 2009 ; Zumbo et al, 2015).

CHAPITRE 2 : L'ANALYSE DES CLASSES LATENTES (ACL)

Les psychologues anglais et américain Charles Edward Spearman et Louis Leon Thurstone ont introduit, au début du 20^e siècle, la théorie des variables latentes dans le but de faire face à un problème inhérent à la psychologie qui était celui d'appréhender l'inobservable à l'instar des attitudes, des pensées, des opinions et bien d'autres (McCutcheon, 1987 ; Lazarsfeld et Henry, 1968). Pour de tels concepts, cette théorie postule l'existence de variables non observables directement (variables latentes), mais dont les valeurs peuvent être estimées à partir de variables manifestes (Collins et Lanza, 2009 ; Lazarsfeld et Henry, 1968). Par exemple, nous ne pouvons pas mesurer directement l'engagement religieux chez un individu (McCutcheon, 1987). Cependant, nous pouvons supposer qu'un individu est très engagé sur le plan religieux lorsque celui-ci fréquente de manière assidue les lieux de culte, s'adonne constamment à la prière ou affirme que ses croyances religieuses ont de l'importance (McCutcheon, 1987). La régularité de fréquentation des lieux de cultes, la fréquence de prières et le degré d'appréciation de sa croyance, qui sont manifestes, servent alors à mesurer l'engagement religieux qui n'est pas directement observable et donc latent (McCutcheon, 1987). La figure 4 illustre une telle situation où trois variables manifestes (X_1 , X_2 et X_3) permettent d'estimer une variable latente. Des erreurs de mesure (e_1 , e_2 et e_3) sont également associées aux variables manifestes lors de cette estimation.

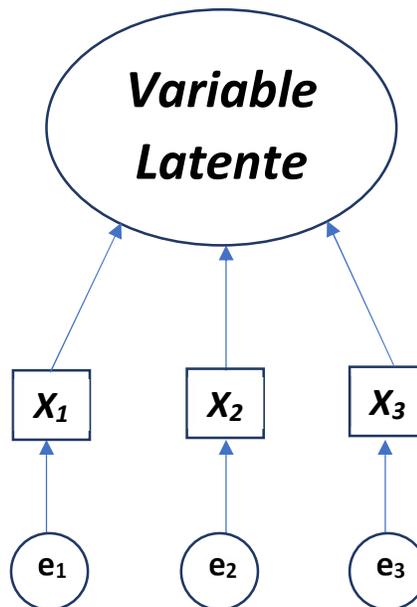


Figure 4 : Variable latente avec trois variables observées (Collins et Lanza, 2009, p.05)

Les premiers modèles de mesure ont servi à analyser uniquement des variables manifestes continues (Collins et Lanza, 2009 ; McCutcheon, 1987). Ce n'est qu'en 1950 que Lazarsfeld posera les bases théoriques d'un modèle à variables latentes qui utilise des variables manifestes catégorielles (Collins et Lanza, 2009 ; McCutcheon, 1987). Aujourd'hui, quatre types de modèles de variables latentes se dégagent des différents travaux d'auteurs : (i) l'analyse factorielle ; (ii) l'analyse de profils latents ; (iii) l'analyse de traits latents ou théorie des réponses à l'item et (iv) l'analyse de classes latentes (Collins et Lanza, 2009 ; McCutcheon, 1987). Le tableau 2 présente ces modèles de mesure en fonction de la nature des variables observées, mais également de la nature de la variable latente qui est estimée.

Tableau 2 : Les quatre différents modèles de variables latentes (Collins et Lanza, 2009, p.7)

	Variable latente continue	Variable latente catégorielle
Variable manifeste continue	Analyse factorielle	Analyse de profils latents
Variable manifeste catégorielle	Analyse de traits latents, ou théorie des réponses à l'item	Analyse de classes latentes*

*En gras, l'analyse des classes latentes qui nous intéresse

Chacun de ces modèles sert à résumer les caractéristiques partagées des multiples variables manifestes dans une ou plusieurs variables latentes (Collins et Lanza, 2009 ; McCutcheon, 1987). Ils permettent ainsi aux chercheurs d'explorer des structures sous-jacentes parmi un ensemble de variables observées dans une perspective exploratoire ou de tester des hypothèses sur ces structures dans une perspective confirmatoire (McCutcheon, 1987). Les variables latentes qui en résultent sont de deux sortes : (i) les variables latentes continues qui supposent l'existence de différences quantitatives entre les groupes d'individu ou d'objets sur le construit mesuré et (ii) les variables latentes catégorielles qui supposent l'existence, en plus de ces différences quantitatives, des différences qualitatives entre les groupes d'individus ou objets sur le construit mesuré (Collins et Lanza, 2009).

L'ACL est en effet très utile et très robuste pour analyser des données dont la structure de base est un schéma de réponses catégorielles (Huang et Bandeen-Roche, 2004) à l'instar des réponses aux tests standardisés (dichotomiques ou polytomiques). En plus, comme nous le verrons dans les sections suivantes, l'ACL avec covariables étend le modèle de base de l'ACL et admet l'introduction, dans un même modèle, de covariables (Collins et Lanza, 2009). Cela a pour conséquence de permettre aux covariables d'avoir des effets directs sur la formation des variables latentes (Collins et Lanza, 2009). Nous présenterons donc tour à tour le modèle de

l'ACL, puis celui de l'ACL avec covariables. Nous détaillerons également l'analyse de l'invariance de la mesure en ACL et montrerons en quoi elle est utile pour soutenir l'idée d'une mesure de performances à un test qui tient compte des variables contextuelles. Nous allons ensuite clôturer le chapitre en présentant les questions spécifiques de notre recherche après avoir donné quelques exemples empiriques d'utilisation de l'ACL dans le domaine de la psychologie, de la santé et de l'éducation.

2.1 Le modèle de l'ACL

2.1.1 Expression du modèle

L'ACL se base sur une approche « orientée-personne » contrairement aux autres modèles de variables latentes qui sont axés sur des variables manifestes (Collins et Lanza, 2009). Dans une approche axée sur les variables telle que l'analyse factorielle, l'accent est mis sur l'identification des relations entre les variables, et ces relations sont supposées s'appliquer à tous les individus (Collins et Lanza, 2009). L'approche orientée-personne en revanche se focalise sur les individus et ce, sur la base de leurs schémas de réponses aux variables observées qui seraient pertinentes pour le problème considéré (Collins et Lanza, 2009). Le modèle identifie ensuite des sous-types d'individus relativement homogènes à l'aide de ces schémas de réponses (Collins et Lanza, 2009).

Un schéma de réponses encore appelé « patron de réponses » est une combinaison de réponses possibles sur l'ensemble des variables d'intérêts (Collins et Lanza, 2009 ; McCutcheon, 1987). Par exemple, un patron de réponses sur trois variables d'intérêts dont les modalités de réponses sont dichotomiques {**oui** ou **non**} peut se décliner en « **oui** » pour la première variable, « **oui** » pour la deuxième variable et « **non** » pour la troisième variable ; soit (**oui**, **oui**, **non**). D'autres individus peuvent répondre par (**oui**, **oui**, **oui**) ou (**oui**, **non**, **oui**) ou encore (**non**, **non**, **oui**) aux différentes variables.

Supposons que j indique la variable observée avec $j = 1, \dots, J$ et que chaque variable observée j soit associée à un nombre de catégories de réponse noté r_j avec $r_j = 1, \dots, R_j$. Dans ce cas, l'ensemble des patrons de réponse possibles forme une table de contingence comportant $W = \prod_{j=1}^J R_j$ cellules. Dans notre exemple, $W = \prod_{j=1}^3 R_j = 2 \times 2 \times 2 = 2^3 = 8$ patrons de réponse possible. Chaque case du tableau correspond donc à un vecteur noté $y = (r_1, \dots, r_j)$. L'ensemble des patrons de réponse fait l'objet d'une matrice Y ayant W lignes et J colonnes, chaque ligne de cette matrice correspond à un patron de réponse et est associée à une probabilité $P(Y = y)$, la somme des probabilités pour l'ensemble des patrons de réponse est égale à 1.

L'ACL consiste à dégager une variable latente catégorielle notée L constituée de classes latentes notées c avec $c = 1, \dots, C$. Pour un individu donné, chaque classe latente c est caractérisée par la probabilité d'appartenir à cette classe latente, ce paramètre appelé prévalence est noté γ_c . Les classes sont mutuellement exclusives et exhaustives, ce qui se traduit par $\sum_{c=1}^C \gamma_c = 1$. Chaque individu est ensuite classé dans la classe latente ayant la prévalence la plus élevée.

Chaque classe latente est également caractérisée par les probabilités de choisir chacune des modalités de réponse de chaque item, notées ρ_{j,r_j} (Collins et Lanza, 2009). C'est ce paramètre qui est au cœur de la mesure associée à la variable latente. Ainsi, le paramètre $\rho_{j,r_j|c}$ correspond à la probabilité conditionnelle qu'un individu fasse le choix de la modalité de réponse r_j pour la variable j sachant qu'il appartient à la classe c . Étant donné que chaque individu ne choisit qu'une seule modalité de réponse à une question donnée, on a $\sum_{r_j=1}^{R_j} \rho_{j,r_j|c} = 1$.

Soit y_j la valeur de la réponse donnée par un individu à la question j . L'indicateur $I(y_j = r_j)$ prend la valeur 1 lorsque $y_j = r_j$ et la valeur 0 sinon. L'équation du modèle correspond à la probabilité d'observer un vecteur particulier de réponses y . Cette probabilité est alors fonction des probabilités d'appartenance de l'individu aux classes latentes et des probabilités conditionnelles d'observer chaque réponse en fonction de l'appartenance à une classe latente (voir équation 1).

$$P(Y = y) = \sum_{c=1}^C \gamma_c \prod_{j=1}^J \prod_{r_j=1}^{R_j} \rho_{j,r_j|c}^{I(y_j=r_j)} \quad (\text{Équation 1})$$

2.1.2 Choix du nombre de classes latentes

Le choix du nombre de classes latentes à identifier ou à retenir dans le cadre d'une ACL est une étape très importante (McCutcheon, 1987 ; Collins et Lanza, 2009). Elle permet de retenir le nombre de profils qui décrit au mieux les données observées (Collins et Lanza, 2009). Ce choix est fait selon que le chercheur est dans une posture confirmatoire ou exploratoire (McCutcheon, 1987). Dans le cadre d'une démarche confirmatoire, il est question d'étudier la correspondance entre les caractéristiques hypothétiques de la variable latente et celles découvertes dans la variable latente identifiée empiriquement (McCutcheon, 1987). Le modèle de classe latente est alors restreint ; le chercheur impose des limites a priori sur les probabilités de classes latentes et/ou sur les probabilités conditionnelles de réponses aux items selon les spécificités des hypothèses à vérifier (McCutcheon, 1987).

Les restrictions sur les probabilités de classes latentes consistent à fixer a priori un nombre donné de classes latentes ou à astreindre deux ou plusieurs classes latentes à avoir la même proportion d'individus (McCutcheon, 1987). Un chercheur peut opérer ces choix lorsqu'il existe dans la littérature des preuves solides qui laissent entrevoir qu'un concept mesuré par un certain nombre d'indicateurs aboutit nécessairement à un certain nombre de profils (McCutcheon, 1987). Il peut également tester l'hypothèse selon laquelle une classe latente donnée représente une proportion spécifique de la population (McCutcheon, 1987).

Les restrictions de probabilités conditionnelles de réponses aux items se font lorsque la théorie qui guide le chercheur requiert une valeur spécifique pour une modalité de réponse à une variable donnée (McCutcheon, 1987). La valeur de la probabilité pour cette réponse particulière est alors fixée. Cela arrive lorsque la littérature soutient qu'un type particulier de répondants est prédisposé à choisir une certaine modalité de réponse à une question donnée (McCutcheon, 1987). Le chercheur observera alors la nature des contributions de ces restrictions aux autres paramètres de l'ACL. Cependant, McCutcheon (1987) affirme que dans la pratique, les restrictions sur ces paramètres de l'ACL ont une application quelque peu limitée, car elles peuvent conduire à des conclusions scientifiques erronées. Et en général, les chercheurs associent les modèles non restreints aux modèles restreints pour fournir des analyses plus riches (McCutcheon, 1987).

Dans le cadre d'une démarche exploratoire, qui est la plus fréquemment utilisée, aucune tentative n'est faite pour tester une hypothèse concernant les caractéristiques des profils (McCutcheon, 1987). Le modèle de classe latente n'est pas restreint, car le chercheur n'impose aucune contrainte a priori sur l'un ou l'autre type de paramètres du modèle (McCutcheon, 1987). Celui-ci se base alors sur certains indices d'ajustement du modèle pour opérer des choix sur le nombre de classes latentes à retenir.

Collins et Lanza (2009) conseillent de regarder et de comparer les critères d'information bayésien (BIC), d'information bayésien ajusté (ABIC), d'information d'Akaike (AIC) et d'information d'Akaike (AIC) ainsi que le rapport de vraisemblance (G^2) pour les différents modèles. Le meilleur modèle est alors celui qui minimise au mieux les différents indicateurs d'ajustement (Hagenaars et McCutcheon, 2002 ; Collins et Lanza, 2009).

Cependant, et le plus souvent d'ailleurs, il est difficile d'avoir un modèle où tous ces paramètres sont inférieurs comparativement à d'autres modèles (McCutcheon, 1987 ; Collins et Lanza, 2009). Dans une telle situation, Agresti et ses collègues (1990) préconisent le choix

du modèle qui possède le plus faible G^2 . Cependant, cette option ne marche que lorsqu'on a une taille d'échantillon N assez grande soit $\frac{N}{W} > 5$; étant le nombre de cellules de notre tableau de contingence (Collins et Lanza, 2009). Collins et Lanza (2009) suggèrent quant à eux de faire le choix du modèle le plus parcimonieux en fonction du BIC et de l'AIC. Dans la pratique, ce choix est fait sur des modèles dont les nombres de classes latentes varient entre 2 et 7 (Patrick et al, 2016). Les résultats de recherches antérieures et la cohérence des modèles avec la théorie sont également des paramètres importants qui permettent de faire ce choix (Collins et Lanza, 2009).

2.2 Analyse de classes latentes avec covariables et invariance de mesure en ACL

2.2.1 Analyse de classes latentes avec covariables

L'ajout des covariables à un modèle ACL permet de décrire la formation des classes latentes et de les caractériser (Collins et Lanza, 2009).

Pour une covariable X , et les mêmes notations que pour l'équation 1, le modèle ACL avec covariable correspond à l'équation 2.

$$P(Y = y|X = x) = \sum_{c=1}^C \gamma_c(x) \prod_{j=1}^J \prod_{r_j=1}^{R_j} \rho_{j,r_j|c}^{I(y_j=r_j)} \quad (\text{Équation 2})$$

dans laquelle $\gamma_c(x)$ est un modèle logistique multinomial correspondant à l'équation 3 pour une seule covariable (l'indice $c' = 1, \dots, C - 1$ correspond à l'utilisation de la classe C comme référence). L'équation 3 peut être généralisée afin de tenir compte de plusieurs covariables.

$$\gamma_c(x) = P(L = c|X = x) = \frac{e^{\beta_{0c} + \beta_{1c}x}}{1 + \sum_{c'=1}^{C-1} e^{\beta_{0c'} + \beta_{1c'}x}} \quad (\text{Équation 3})$$

Comme dans toutes les régressions, l'ACL avec covariables prend en compte des variables catégorielles et continues comme covariables (Collins et Lanza, 2009). Ce modèle traite toutes les covariables comme des variables numériques; cela implique souvent une codification des variables prédictives catégorielles en « variables factices » binaires (0/1) (Collins et Lanza, 2009). La variable sexe par exemple peut être recodifiée dans une variable factice qui prend la valeur 0 pour les femmes et 1 pour les hommes ou inversement. Les variables continues ne nécessitent pas de recodification. Cependant, Collins et Lanza (2009) proposent de standardiser ces variables pour faciliter l'interprétation des résultats.

Lubke et Muthén (2007) proposent de tester le modèle en incluant une covariable à la fois, puis, d'inclure dans le modèle final, uniquement les covariables qui ont eu un effet significatif sur la formation des classes. Ils suggèrent ensuite d'étudier les autres variables a posteriori pour caractériser les différents profils (Lubke et Muthén, 2007).

L'ACL avec covariables nécessite également la désignation d'une variable latente comme catégorie de référence (Collins et Lanza, 2009). Le choix de cette catégorie de référence est arbitraire et n'affecte pas les résultats de manière substantielle. Cependant, ce choix doit être judicieux afin de faciliter l'interprétation des résultats (Collins et Lanza, 2009).

Malgré l'intérêt croissant sur l'utilisation des modèles ACL ces dernières années, très peu de recherches se sont penchées sur la question de l'inclusion des covariables dans le modèle ACL (Nylund-Gibson et Masyn, 2016 ; Wurpts et Geiser, 2014). Les quelques débats qui existent autour de ce sujet semblent également quelque peu contradictoires (Nylund-Gibson et Masyn, 2016). Clark et Muthén (2009) par exemple, soutiennent qu'en raison de la présence de possibles relations entre les covariables et les indicateurs permettant d'étudier un phénomène, l'introduction des covariables dans le modèle ACL pourrait conduire à la formation de profils latents difficilement interprétables même par les experts du domaine. De tels résultats pourraient également donner lieu à des classes latentes, taxées de « fausses classes latentes », qui ne correspondraient pas à la réalité en ce qui concerne le phénomène étudié (Clark et Muthén, 2009). Pour d'autres auteurs en revanche, c'est pourtant là que résiderait tout l'intérêt d'ajouter des covariables dans un modèle ACL (Wurpts et Geiser, 2014). L'introduction des covariables dans le modèle ACL serait même bénéfique à plusieurs égards (Wurpts et Geiser, 2014).

Pour Muthén (2004), un modèle ACL avec covariables aiderait plutôt à déterminer correctement le nombre de classes latentes. En effet, dans la réalité également, il existe des situations dans lesquelles les covariables peuvent influencer directement le processus de mesure (Muthén, 2004 ; Reboussin et al, 2008). L'exemple que Reboussin et ses collègues (2008) donnent pour illustrer cette situation est celle de l'acceptation des problèmes liés à l'alcoolisme (sociaux, juridiques et autres) par les alcooliques en fonction des sexes. Les hommes auraient tendance à plus accepter ces problèmes que les femmes malgré le fait que les deux peuvent être considérés comme des buveurs à problèmes (Reboussin et al, 2008). Dans une étude qui étudierait donc le degré d'approbation de ces problèmes, un modèle ACL avec covariables (sexe pour le cas présent) fournirait une classification plus correcte des profils compte tenu de l'effet de la covariable (sexe) sur les réponses des individus (Reboussin et al, 2008, Vermunt et

Magidson, 2005). Pour de telles situations, le modèle ACL avec covariables estimerait des indicateurs d'ajustement de meilleure qualité que le modèle ACL sans covariables (Reboussin et al, 2008, Vermunt et Magidson, 2005). Le modèle ACL avec covariables présenterait également un atout dans les cas où la taille de l'échantillon serait petite ou lorsque le chercheur ne disposerait pas d'assez d'indicateurs pour étudier un phénomène.

L'inclusion de covariables dans un modèle ACL aiderait en effet à mieux estimer le nombre de classes latentes et à fournir une meilleure classification des individus lorsque la taille de l'échantillon est petite ; soit comprise entre $N = 70$ à $N = 200$ (Wurpts et Geiser, 2014). Le modèle ACL avec covariables peut également s'avérer très utile lorsque le nombre d'indicateurs disponibles pour étudier un phénomène est réduit ; 4 à 5 indicateurs (Wurpts et Geiser, 2014). Les covariables compenseraient ainsi ces limites en fournissant des informations supplémentaires qui peuvent être utilisées dans le processus d'estimation des classes latentes (Wurpts et Geiser, 2014). Un modèle ACL avec covariables permettrait donc d'obtenir de meilleurs résultats qu'un modèle ACL sans covariables (Wurpts et Geiser, 2014).

A l'instar de Zumbo et ses collègues (2015), nous constatons donc que le modèle ACL avec covariables présente des caractéristiques qui vont dans le sens des idéaux que soutient l'approche écologique. Comme nous l'avons déroulé dans le chapitre précédent, dans une approche écologique, estimer les performances à un test invite à considérer à la fois les connaissances des individus, leurs caractéristiques individuelles et les contextes dans lesquels ils évoluent (Mislevy, 2018 ; Zumbo et al, 2015). Ces individus vivent toutefois en groupes et partagent certaines caractéristiques contextuelles communes. Cette réalité invite les chercheurs à s'assurer au préalable que les modèles de mesures qu'ils utilisent peuvent s'appliquer à tous les groupes d'individus de leur échantillon sans risque de biais (Collins et Lanza, 2009). Dans la section suivante, nous évoquons donc l'invariance de mesure en ACL qui traite de ces questions.

2.2.2 Invariance de mesure en ACL

Dans une opération de mesure d'attributs psychologiques où la population cible peut être subdivisée en sous-populations plus ou moins hétérogènes, comme c'est le cas pour les ELE, la question d'invariance de la mesure revêt un aspect fondamental (Engelhard, 1992 ; Engelhard, 2008). L'objectif des praticiens de la mesure en sciences sociales est en effet d'éliminer l'effet de l'échantillon sur la mesure (Engelhard, 1992 ; Engelhard, 2008). Une telle mesure pourrait être stable et s'appliquer à toute la population cible en faisant abstraction des caractéristiques personnelles des individus (Engelhard, 1992 ; Engelhard, 2008). Par exemple,

les résultats des tests pour un groupe d'individus plus ou moins hétérogène doivent tenir pour l'ensemble de ces individus indépendamment du degré de décomposition (filles/garçons, riches/pauvres...) qui peut être appliqué aux données (Engelhard, 1992 ; Engelhard, 2008). Dans le cadre de l'ACL également, lorsque deux ou plusieurs groupes doivent être comparés sur une variable latente, il est important de s'assurer que la variable latente a ou non les mêmes caractéristiques de mesure pour chacun des groupes (Collins et Lanza, 2009).

L'existence d'une invariance de mesure dans l'ACL au sens strict, signifierait que le nombre de classes latentes et toutes les probabilités de réponses aux items sont identiques, quel que soit le groupe d'individus qui proviendrait de l'échantillon (Collins et Lanza, 2009). Une telle situation implique également que l'interprétation des classes latentes est identique d'un groupe à l'autre et que les individus d'une classe latente particulière, mais issus de groupes différents ont des probabilités identiques de fournir un patron de réponse donné (Collins et Lanza, 2009). Dans la pratique, il existe deux façons de vérifier si l'invariance de mesure tient lors d'une ACL.

La première option consiste à mener des ACL séparées sur des sous-ensembles de données distincts qui représentent les différents groupes d'individus pour lesquels nous souhaitons faire des vérifications d'invariance de mesure (Collins et Lanza, 2009). Pour chaque groupe, le choix du modèle le plus parcimonieux se fait sur la base des indices d'ajustement tel que présenté dans les sections précédentes. Une telle démarche ne s'applique cependant que lorsque les tailles des échantillons sont assez grandes et qu'elles ne varient pas considérablement entre les différents groupes (Collins et Lanza, 2009). Si cette condition n'est pas respectée, il est plus judicieux de conduire une seule ACL sur l'ensemble des données et de considérer que les résultats s'appliquent à l'ensemble des données (Collins et Lanza, 2009).

S'il ne se pose cependant aucun problème de taille d'échantillon et que les résultats suggèrent des nombres de classes latentes différents entre les groupes, l'hypothèse de l'invariance de mesure est rejetée et les analyses ACL devraient être conduites de façons séparées pour chaque groupe afin d'assurer une interprétation plus équitable des résultats (Collins et Lanza, 2009). Dans le cas contraire, si les résultats laissent supposer des nombres de classes latentes identiques pour tous les groupes, alors l'hypothèse de l'invariance de mesure entre les groupes tient et une seule ACL devrait être faite pour l'ensemble des données (Collins et Lanza, 2009).

La deuxième option permettant de vérifier l'invariance de mesure en ACL consiste à comparer les indices d'ajustement de deux modèles d'ACL sur les données globales (Collins et Lanza, 2009). Les deux modèles doivent avoir le même nombre de classes latentes préalablement choisi en sélectionnant le modèle qui s'ajuste au mieux aux données globales (Collins et Lanza, 2009). Ce choix s'opère selon la procédure que nous avons présentée dans les sections précédentes. Dans le premier modèle, non restreint, les probabilités à fournir des patrons de réponses au sein des classes latentes peuvent varier d'un groupe à l'autre tandis qu'elles sont contraintes d'être égales entre les groupes dans le deuxième modèle (Collins et Lanza, 2009). Une comparaison des indicateurs d'ajustement des deux modèles au moyen d'un test de χ^2 permet alors de prendre une décision. Si, comparativement au premier modèle, le deuxième modèle fournit un ajustement significativement plus faible aux données, alors l'hypothèse de l'invariance de mesure entre les différents groupes est compromise ; la conduite des ACL séparées sur les données des différents groupes est alors plus équitable (Collins et Lanza, 2009).

Si par contre le deuxième modèle s'ajuste mieux aux données que le premier modèle, alors les restrictions des paramètres dans le deuxième modèle sont plausibles et l'hypothèse de l'invariance de mesure est alors retenue (Collins et Lanza, 2009). Une seule ACL sur les données globale est dans ce cas suffisante et les résultats peuvent s'appliquer à l'ensemble des individus (Collins et Lanza, 2009).

Dans la pratique, bien que les procédures statistiques soient utiles, rien ne remplace les antécédents du chercheur dans le domaine étudié et sa connaissance des données (Collins et Lanza, 2009). Le chercheur doit en effet, et le plus souvent, faire preuve de jugement pour décider s'il doit ou non conduire une vérification de l'invariance de mesure sur des groupes issus de son échantillon (Collins et Lanza, 2009). Ainsi, ses connaissances du domaine et des données peuvent lui suggérer de mener une seule ACL sur l'ensemble des données même si les résultats de l'analyse de l'invariance de mesure lui suggèrent le contraire (Collins et Lanza, 2009). Il ne doit cependant pas omettre de le souligner au moment de rapporter les résultats de ses analyses ACL (Collins et Lanza, 2009).

Nous venons donc de présenter de façon théorique le modèle de l'ACL qui est une approche orientée personne intéressante pour traiter des données dont la structure de base est un schéma de réponses catégorielles et classer les individus en fonction des réponses qu'ils fournissent (Collins et Lanza, 2009 ; Huang et Bandeen-Roche, 2004). Nous avons ensuite introduit le modèle de l'ACL avec covariables qui étend le modèle de base de l'ACL et admet

l'introduction, dans un même modèle, de variables d'intérêts et de covariables (Collins et Lanza, 2009). Nous avons également parlé de l'analyse de l'invariance de la mesure en ACL qui est indispensable lorsque nous traitons des données sur des groupes d'individus hétérogènes (Collins et Lanza, 2009). Dans une telle situation, nous devons donc nous assurer que tous nos résultats tiennent pour l'ensemble des individus (Collins et Lanza, 2009). Dans la section suivante, nous présentons quelques exemples de recherches empiriques qui utilisent l'ACL pour analyser des données. Les principaux objectifs de ces présentations sont entre autres de souligner l'avantage à utiliser l'ACL pour la classification des individus, de regarder la procédure de sélection du meilleur modèle ACL à appliquer aux données et d'évoquer la possibilité d'utiliser les modèles ACL aux données d'une ELE.

2.3 Exemples empiriques de l'utilisation de l'ACL

Dans cette section, nous présentons trois études où l'ACL est appliquée aux données. Nous déroulons tout d'abord la recherche de Udry (2003) qui relève du domaine de la sociologie. Nous regardons ensuite l'étude de Lafortune (2009) en médecine et terminons par les travaux de Abarda (2013) en éducation. Pour chacune des recherches, nous insistons uniquement sur un aspect particulier de l'ACL. Ainsi, dans la première présentation, nous soulignons l'utilité d'utiliser l'ACL dans une recherche pour classer des individus. Pour la seconde présentation, nous portons un regard sur la procédure de sélection du meilleur modèle ACL pour l'analyse des données. Dans la dernière présentation, nous évoquons la possibilité d'implémenter une approche écologique à l'aide de l'ACL pour les données d'une ELE.

2.3.1 En sociologie

L'exemple en sociologie que nous avons choisi de présenter porte sur les données de la National Longitudinal Study of Adolescent to Adult Health (Add Health) en Caroline, aux États-Unis. Add Health mène des recherches longitudinales sur la santé et les comportements à risque chez les jeunes et les influences des facteurs sociaux sur ces comportements (Udry, 2003). Celle qui nous intéresse est l'étude sur la délinquance adolescente de Udry (2003) présentée par Collins et Lanza (2009). Udry (2003) a analysé les données d'adolescents (N = 2087) qui ont participé à l'enquête Add Health de 1994-1995. Ces adolescents étaient alors âgés en moyenne de 16,4 ans et faisaient la 4^e et 5^e année du secondaire (Collins et Lanza, 2009). Les participants devaient répondre par « oui » ou par « non » à une série de questions parmi lesquelles : (i) a déjà menti aux parents ? (ii) a déjà été bruyant/tapageur/indiscipliné en public ? (iii) a déjà endommagé un bien ? a déjà volé quelque chose dans un magasin ? (iv) a déjà volé

quelque chose d'une valeur <50 \$? a déjà participé à un combat de groupe ? Le tableau 3 présente le pourcentage de « oui » obtenu pour chacune des questions.

Tableau 3 : Proportion des réponses « oui » aux questions sur la délinquance adolescente (traduction libre de Collins et Lanza, 2009, p.11)

Item du questionnaire	Proportion de réponses « oui »
A déjà menti aux parents	0,57
A déjà été bruyant/tapageur/indiscipliné en public	0,49
A déjà endommagé un bien	0,17
A déjà volé quelque chose dans un magasin	0,24
A déjà volé quelque chose d'une valeur <50 \$	0,20
A déjà participé à une bagarre de groupe	0,19

Les résultats du tableau 3 permettent de souligner que certains comportements délinquants sont plus normatifs que d'autres à l'exemple du fait de mentir à ses parents qui est assez courant (Collins et Lanza, 2009). Les résultats du tableau 3 mettent également en exergue le fait que la participation à une bagarre de groupe est moins courante chez les personnes interrogées (Collins et Lanza, 2009). Cependant, une ACL sur les données permet d'approfondir les investigations notamment en recherchant s'il existe des sous-groupes distincts d'adolescents qui adopteraient des schémas particuliers de comportement délinquant (Collins et Lanza, 2009). Si c'est le cas, l'ACL permet également de connaître la répartition des adolescents dans ces éventuels sous-groupes (Collins et Lanza, 2009). Ainsi, comme le montre le tableau 4, les résultats d'une ACL sur ces données a permis de dégager 4 classes latentes : les « non/légers délinquants », les « antagonistes verbaux », les « voleurs à l'étalage » et les « délinquants généraux » (Collins et Lanza, 2009).

Les non/légers délinquants, plus nombreux (49% de l'échantillon), regroupent les adolescents qui ont des probabilités presque nulles de pratiquer les différents comportements déviants (Collins et Lanza, 2009). Cependant, ils sont quelque peu enclins à mentir aux parents et à être bruyants/tapageurs/indisciplinés en public (Collins et Lanza, 2009). Les antagonistes verbaux (26% de l'échantillon) ont de leur côté des probabilités assez élevées de mentir aux parents et d'être bruyants/tapageurs/indisciplinés en public tandis qu'ils ont des probabilités faibles de pratiquer les autres comportements déviants (Collins et Lanza, 2009). Les voleurs à l'étalage (18% de l'échantillon) regroupent les adolescents qui, en plus d'avoir des probabilités élevées de mentir aux parents et d'être bruyants/tapageurs/indisciplinés, présentent également des probabilités élevées de voler quelque chose dans un magasin et de voler quelque chose d'une valeur <50 \$ (Collins et Lanza, 2009). Les délinquants généraux, certes très marginaux

(6% de l'échantillon), ont de fortes probabilités de pratiquer tous les différents comportements déviants (Collins et Lanza, 2009).

Tableau 4 : Classes latentes issues de l'ACL des données sur la délinquance adolescente (traduction libre de Collins et Lanza, 2009, p.12)

	Classe Latente			
	Non/légers délinquants	Antagonistes verbaux	Voleurs à l'étalage	Délinquants généraux
Probabilité d'appartenance à la classe latente	0,49	0,26	0,18	0,06
Probabilité conditionnelle de réponse « oui »				
A déjà menti aux parents	0,33	0,81*	0,78	0,89
A déjà été bruyant/tapageur/indiscipliné en public	0,20	0,82	0,62	1,00
A déjà endommagé un bien	0,01	0,25	0,25	0,89
A déjà volé quelque chose dans un magasin	0,03	0,02	0,92	0,88
A déjà volé quelque chose d'une valeur <50 \$	0,00	0,03	0,73	0,88
A déjà participé à un combat de groupe	0,04	0,31	0,24	0,64

*Probabilités conditionnelles >0,5 en gras pour faciliter l'interprétation

Ainsi, dans cette première présentation, nous soulignons l'utilité à utiliser l'ACL dans une recherche pour classer des individus lorsque nos variables d'intérêts sont dichotomiques ou polytomiques. Dans une telle situation, le modèle ACL scrute en profondeur les données et met en exergue des sous-groupes d'individus qui auraient des probabilités similaires de donner un schéma particulier de réponses. Ainsi, au lieu d'avoir des relations entre variables, nous avons plutôt des relations entre les individus ; chaque patron de réponse représentant un individu en particulier. L'exemple suivant que nous présentons est de celui Lafortune (2009) qui utilise l'ACL pour ses besoins de recherche en médecine.

2.3.2 En médecine

Notre choix en médecine porte donc sur le travail de Lafortune (2009) qui étudie les rapports entre les besoins hétérogènes des personnes âgées vivant à Montréal et leur utilisation conséquente des services médicaux et sociaux. Cependant, nous nous intéressons uniquement au volet d'identification des profils d'état de santé de son travail. Son échantillon (N=1164) provient d'une revue des dossiers de personnes âgées de 65 ans et plus qui recevaient des services à domicile des centres locaux de services communautaires (CLSC) Bordeaux-Cartierville et Côtes-des-Neiges (Lafortune, 2009). Pour établir les profils d'état de santé de ces personnes âgées, Lafortune (2009) a effectué une ACL sur 17 indicateurs d'état de santé. Ces indicateurs sont regroupés en six grands groupes : (i) problèmes de santé chroniques, (ii) état cognitif, (iii) dépression, (iv) limitations fonctionnelles, (v) incapacités et (vi) limitations sensorielles (Lafortune, 2009).

Pour le choix du meilleur modèle pour classer les individus de son échantillon, Lafortune (2009) a comparé 4 modèles de 2 à 5 classes latentes ; au-delà de 5 classes latentes, il ne constate aucune amélioration significative sur la qualité de la classification. La comparaison des 4 modèles montre que le modèle à 4 classes est le modèle le plus parcimonieux (Lafortune, 2009). Les indicateurs AIC, BIC et ABIC ont les valeurs les plus petites et la qualité de la classification y est élevée (entropie : 0,805). Le tableau 5 présente les indicateurs d'ajustement des différents modèles.

Tableau 5 : Qualité d'ajustement des modèles d'ACL sur les données d'état de santé (Lafortune, 2009, p. 102)

	Modèles			
	2 classes	3 classes	4 classes*	5 classes
Comparaison des modèles séquentiels	2 vs 1 classes	3 vs 2 classes	4 vs 3 classes	5 vs 4 classes
LMR-LRT				
Log vraisemblance (c+1 classes)	13265,31	13595,67	12239,68	12093,26
-2log vraisemblance	1468,19	2711,98	292,85	107,55
Valeur p	0,000	0,000	0,0006	0,307
Critères d'information				
AIC	25148,43	24483,26	24192,52	24296,97
BIC	25365,96	24493,48	24207,69	24848,37
BIC ajusté	25229,38	24487,13	24198,16	24502,15
Entropie	0,814	0,791	0,805	0,765

LMR-LRT= Test du ratio de vraisemblance de Low-Mendell-Rubin ; AIC : Critère d'information d'Akaike; BIC : Critère d'information Bayésien.

* En gras, les valeurs des indicateurs qui soutiennent le choix du modèle à 4 classes (valeurs les plus petites sauf entropie qui doit être élevée) ; aucune amélioration significative n'est observée à plus de 4 classes latentes (p value > 0,05).

Dans le modèle à 4 classes latentes, les 4 profils qui résultent de l'analyse ACL se distinguent sur le plan qualitatif par les dimensions physique et cognitive, et sur le plan quantitatif par la gravité des incapacités que les individus manifestent (Lafortune, 2009). Ces différents profils sont : (i) limitations physiques et cognitives, (ii) limitations cognitives, (iii) limitations physiques et (iv) relativement sains. Le profil « limitations physiques et cognitives » qui représente 23 % de l'échantillon regroupe des individus qui ont des probabilités élevées de présenter des troubles cognitifs, des maladies chroniques, des limitations sensorielles et fonctionnelles, et d'avoir souffert d'un accident cérébro-vasculaire (Lafortune, 2009). Le deuxième profil, les « limitations cognitives » regroupe des personnes caractérisées principalement par des limitations cognitives et peu de limitations physiques (Lafortune, 2009). Ces individus, qui ont un poids de 11,4 % de l'échantillon, ont peu de chances d'avoir des maladies chroniques et des limitations fonctionnelles (Lafortune, 2009).

Les individus qui font partie du troisième profil, « limitations physiques » ont des probabilités très élevées de souffrir de maladies chroniques, mais ne présentent aucun trouble cognitif (Lafortune, 2009). Ils représentent 35,6 % de l'échantillon et ont également de fortes chances de déclarer des symptômes dépressifs, des limitations fonctionnelles et des incapacités pour la mobilité (Lafortune, 2009). Le dernier profil, « relativement sains », regroupe 29,9 % de l'échantillon et rassemble des individus qui ont peu de maladies chroniques et de faibles probabilités d'avoir des incapacités, des limitations fonctionnelles et des troubles cognitifs (Lafortune, 2009).

L'originalité des travaux de Lafortune (2009) tient du fait d'avoir modélisé l'hétérogénéité de l'état de santé de ces personnes âgées avant d'étudier les conséquences sur l'accès, l'intensité et les coûts des services sociaux et médicaux pour ce type de population. Son travail est ainsi le premier à décrire les différents rapports entre l'état de santé et les coûts en fin de vie pour ces personnes âgées fragiles (Lafortune, 2009).

Pour cette seconde présentation, nous avons porté un regard sur la procédure de sélection du meilleur modèle ACL pour l'analyse des données. Ce choix, comme nous venons de le voir repose sur un ensemble d'indicateur d'ajustement de modèles ; et c'est le modèle le plus parcimonieux qui est retenu pour les analyses. Généralement, le modèle le plus parcimonieux est celui qui offre une meilleure interprétabilité des classes latentes (capacité à caractériser chaque classe latente de façon distincte). C'est le cas des classes latentes de Lafortune (2009) où chaque profil a ses propres caractéristiques qui le distingue des autres profils : (i) limitations physiques et cognitives, (ii) limitations cognitives, (iii) limitations physiques et (iv) relativement sains. Le troisième exemple que nous présentons est un travail de Abarda (2013) qui utilise des ACL sur des données d'une ELE.

2.3.3 En éducation

L'étude que nous avons choisi de présenter en éducation est une ACL sur les données d'un test de langue d'une ELE qui s'est déroulée dans 178 écoles au Cameroun en 2011 (Abarda 2013). Les participants à cette évaluation étaient des élèves du cours moyen 1 (5e année du primaire) au nombre de 2553 à raison de 608 élèves anglophones et 1945 élèves francophones (Abarda 2013). Trois domaines étaient évalués dans ce test de langue : (i) l'écrit, (ii) la lecture et (iii) les outils de la langue notamment la grammaire, la conjugaison et l'orthographe (Abarda 2013). Sur les 39 items de ce test de langue, 31% portaient sur l'écrit, 28% sur la lecture et 41% les outils de la langue. Une partie de ces items (14) sont des Questions à Choix multiples (QCM) tandis que 19 items sont des questions à réponses courtes et 6 items, des questions à réponses

ouvertes. Les analyses ACL de Abarda (2013) sur les réponses des élèves ont permis d'identifier 5 classes latentes.

La première classe latente est composée de « lecteurs confirmés ». Elle regroupe 23% de la totalité des élèves ayant passé le test. Ces élèves ont la particularité d'avoir de grandes chances de répondre correctement à tous les items (Abarda 2013). Ils présentent ainsi, au regard du test qu'ils ont effectué, des capacités de lire et de comprendre, de tirer des relations entre plusieurs mots, d'interpréter le sens d'une partie du texte et de déduire plusieurs informations à partir d'un texte (Abarda 2013).

La deuxième classe latente qui s'est dégagée des analyses ACL de Abarda (2013) regroupe des « non lecteurs ». Ces types d'élèves ont des probabilités élevées d'échec sur tous les items (Abarda 2013). Ils semblent avoir beaucoup de difficulté à lire des textes simples, à reconnaître des informations ou à comprendre la signification d'un extrait d'un texte et donc ont moins de chances à trouver des réponses justes aux items, même les plus faciles du test (Abarda 2013). Ils représentent 43% des élèves de l'échantillon (Abarda 2013).

La troisième classe latente regroupe des « lecteurs partiels » qui représentent 14% des élèves de l'ensemble de l'échantillon (Abarda 2013). Ces élèves ont des probabilités élevées à trouver de bonnes réponses uniquement au niveau de quatre items qui semblent être les plus faciles du test et qui portent sur le repérage d'une information explicite (Abarda 2013). Ils éprouvent cependant beaucoup de difficultés face aux autres items du test (Abarda 2013).

La quatrième classe latente est celle des « lecteurs créatifs » (Abarda 2013). Ce type d'élèves qui présente des probabilités élevées à trouver de bonnes réponses aux questions à réponses courtes et à réponses ouvertes a cependant des difficultés à trouver les bonnes réponses aux QCM (Abarda 2013). Les élèves dans cette situation représentent cependant une faible proportion (5%) des élèves de l'échantillon (Abarda 2013).

La cinquième classe latente identifiée par Abarda (2013) est celle des « lecteurs ayant besoin d'aide ». Cette classe latente regroupe les 16% des élèves de l'échantillon qui sont capables de répondre aux questions qui demandent de repérer une information simple (Abarda 2013). Ces élèves ont cependant des probabilités inférieures à 0,5 de réussir les autres items du test (Abarda 2013).

Abarda (2013) a également conduit une ACL distincte sur certaines variables liées à la situation socioéconomique des élèves ainsi que sur le niveau d'instruction des mères de ces

élèves. Cette analyse lui a permis de classer les élèves en cinq profils socio-économiques : (i) les élèves issus des ménages pauvres, mais alphabétisés (22% de l'effectif), (ii) les élèves issus de familles vulnérables et alphabétisées (8% de l'effectif), (iii) les élèves issus de familles vulnérables et analphabètes (20% de l'effectif), (iv) les élèves issus des ménages moyens et alphabétisés (30% de l'effectif) et (v) les élèves issus de familles riches et alphabétisées (20% de l'effectif).

Le travail de Abarda (2013) a permis de souligner des profils de performances particuliers à l'exemple des « lecteurs créatifs » qui réussissent à produire des réponses élaborées, mais qui ont des difficultés face à des QCM qui semblent relativement plus faciles. De même, le croisement entre les profils de performances et les profils socio-économiques qu'il a obtenu a permis d'identifier des groupes d'élèves qui performant, malgré certains désavantages (Abarda 2013). C'est le cas des élèves faisant partie des « ménages pauvres, mais alphabétisés » qui réussissent aux items du test de lecture (Abarda 2013). De tels détails ne sont pas visibles dans les rapports des résultats aux tests des ELE, notamment aux tests PASEC 2006 que Abarda (2013) compare à ses résultats.

Le travail de Abarda (2013), souligne la possibilité mais également les bénéfices à utiliser une ACL pour analyser des données d'une ELE. Les profils de performances fort intéressants et les cas atypiques que ses travaux ont pu mettre en exergue montrent bien la prévalence à utiliser les modèles ACL pour les données d'une ELE au lieu des modèles classiques couramment utilisés dans ces programmes. Cependant, l'approche d'analyse de Abarda (2013) pour estimer les performances aux tests reste celle proche de celle utilisée dans les ELE. Abarda (2013) utilise en effet, dans un premier modèle ACL, uniquement les réponses aux items des tests pour estimer les profils de performances. C'est dans un second modèle ACL qu'il estime les profils socio-économiques à base des données socio-économiques des participants (Abarda, 2013). Cette façon de faire est cependant contraire à la philosophie d'une approche écologique où l'estimation des performances à un test devrait se faire à l'aide d'un modèle unique qui tiendrait compte à la fois des réponses aux items, mais également des éventuelles données contextuelles comme le ferait un modèle ACL avec covariables. Nous n'avons malheureusement trouvé dans la littérature aucune étude qui utilise le modèle de l'ACL avec covariables sur les données d'une ELE ; la seule étude que nous avons recensée et qui utilise l'ACL sur les données d'une ELE étant celle de Abarda (2013) avec les limites que nous venons de souligner.

Les trois études empiriques que nous venons de présenter nous ont donc permis tout d'abord de souligner l'utilité à utiliser l'ACL dans une recherche pour classer des individus. Ensuite, nous avons montré que la procédure de sélection du meilleur modèle ACL obéit à des règles strictes et que le résultat de la sélection permet généralement une meilleure classification des individus concernés en fonction des réponses qu'ils fournissent aux différentes variables d'intérêts. Pour finir, nous avons évoqué la possibilité d'implémenter une approche écologique à l'aide de l'ACL sur des données d'une ELE. Notre choix pour ce dernier exemple présente certes des limites pour une implémentation d'une approche écologique. Cependant, cet exemple souligne les bénéfices qui peuvent résulter de l'utilisation d'une ACL sur les données d'une ELE. Dans la section suivante, nous présenterons donc les différentes questions spécifiques de notre recherche.

2.4 Questions spécifiques de recherche

Tel que nous l'avons présenté dans le chapitre précédent, la question générale de recherche à laquelle nous souhaitons répondre est :

Comment la mesure des performances qui intègre une approche écologique influence-t-elle les résultats du test en lecture du PASEC2014 ?

Dans le présent chapitre, nous venons de voir le modèle ACL, qui est préconisé par Zumbo et ses collègues (2015) pour implémenter l'approche écologique sur les données d'une ELE. De plus, l'extension de ce modèle, l'ACL avec covariables, permet d'introduire dans un même modèle de mesure, les différentes réponses aux items d'un test ainsi qu'une liste de covariables. Nous avons également présenté des trois exemples de recherches empiriques dont le premier avait pour but de montrer que l'ACL est intéressante pour traiter des données catégorielles et classifier des individus. La seconde présentation avait pour but de présenter les mécanismes de sélection du modèle ACL et la troisième, de souligner les avantages à utiliser une ACL sur les données d'une ELE. Cette dernière présentation, celle de Abarda (2013) est fort intéressante pour nos travaux et aurait pu nous servir de référence pour nos analyses. Cependant, Abarda (2013) analyse uniquement les réponses aux tests pour estimer ses profils de performances. Une telle approche est contraire aux idéaux d'une approche écologique où l'estimation des performances à un test devrait se faire à l'aide d'un modèle unique qui tiendrait compte à la fois des réponses aux items, mais également des éventuelles données contextuelles.

Pour nos travaux, nous allons donc étendre les travaux de Abarda (2013) dans une perspective écologique en analysant les résultats du test de lecture PASEC2014-Cameroun à l'aide d'un modèle ACL avec covariables. Ce modèle comme nous l'avons présenté, permet la participation des covariables à l'estimation des profils qui pourraient se dégager des résultats. Ainsi, nous déclinons notre question générale de recherche en deux questions spécifiques qui sont :

- (i) *Comment interpréter les résultats du test en lecture PASEC2014-Cameroun lorsqu'on les analyse à partir de l'approche écologique de Zumbo et ses collègues (2015) en utilisant des analyses des classes latentes avec covariables ?*
- (ii) *Comment la compréhension des résultats des tests de lecture PASEC2014-Cameroun est-elle influencée par l'usage de l'approche écologique de Zumbo et ses collègues (2015) en comparaison avec les résultats PASEC ?*

CHAPITRE 3 : MÉTHODOLOGIE

Ce chapitre, qui présente le cadre méthodologique de notre recherche, est le lieu pour nous d'explicitier toute la démarche qui nous permettra de répondre aux différentes questions spécifiques que nous venons de présenter. Nous allons donc présenter brièvement dans un premier temps, la stratégie retenue pour notre recherche. Dans cette première section, nous nous attarderons également sur le contexte de l'évaluation PASEC2014, les données de cette évaluation et les données que nous retenons pour notre recherche. Ensuite, nous déroulons la démarche retenue pour nos analyses. Le but de cette deuxième section est de présenter les différentes étapes de nos analyses et les différents outils qui ont servi à réaliser ces analyses. Nous clôturons ce chapitre en relevant quelques considérations éthiques liées à notre recherche.

3.1 Stratégie de recherche

Pour nos travaux, nous avons fait le choix de réexaminer de façon rétrospective les données de l'évaluation PASEC2014. Notre choix de travailler sur les données de cette évaluation repose sur plusieurs raisons. Tout d'abord parce que notre pays, le Cameroun, a pris part à l'évaluation PASEC2014. Ensuite, comme nous l'avons présenté dans la partie contexte de notre recherche, ce sont les résultats de cette évaluation qui ont suscité en nous les premières réflexions pour notre recherche. Une autre raison qui nous a poussé à choisir ces données est notre relative maîtrise du contexte des pays ayant pris part à l'évaluation PASEC2014 de façon générale et du contexte camerounais en particulier. La dernière raison, et pas la moindre, qui justifie notre choix se base sur l'accessibilité des données PASEC2014 et des manuels techniques destinés à l'exploitation des données de cette évaluation. En effet, Le PASEC met facilement à la disposition des chercheurs, des étudiants et de toute personne qui le sollicite, les différentes bases de données et documents techniques relatifs aux évaluations qu'il mène. Tout requérant qui remplit un formulaire sur le site internet du PASEC à l'adresse <http://www.pasec.confemen.org/donnees>, peut avoir accès à ces données.

Le fait pour nous d'utiliser des données existantes ne présente cependant pas que des avantages, car nous sommes dans une situation où les causes et les effets du phénomène que nous souhaitons étudier, se sont déjà produits (Simon, et Goes, 2013). Une telle situation met le chercheur dans l'impossibilité de modifier les variables indépendantes ou d'ajouter celles qu'il jugerait pertinentes pour sa recherche (Simon, et Goes, 2013). Cette situation met également tout chercheur face à un manque de contrôle au niveau de la sélection aléatoire de l'échantillon (Simon, et Goes, 2013). Dans un tel contexte, Lord (1973) suggère de justifier le

choix de toute variable issue de la base de données et d'expliquer clairement la démarche d'analyse qui conduira à tout nouveau résultat. C'est ce à quoi s'attèlent les sections suivantes.

3.1.1 Données PASEC2014

Échantillon PASEC2014

L'évaluation PASEC2014 s'est déroulée entre 2012 et 2014 dans dix pays d'Afrique subsaharienne que sont : le Bénin, le Burkina Faso, le Burundi, le Cameroun, le Congo, la Côte d'Ivoire, le Niger, le Sénégal, le Tchad et le Togo (PASEC, 2015). La langue d'enseignement dans ces pays est le français ; exception pour le Cameroun qui dispose de deux langues d'enseignement⁸ et du Burundi où les élèves en début du cycle primaire (2e année), étudient en langue locale qui est le Kirundi (PASEC, 2015). Le PASEC a ainsi confectionné les différents instruments de cette évaluation en français et traduit en anglais pour les élèves anglophones au Cameroun et en Kirundi pour les élèves de 2^e années du primaire au Burundi (PASEC, 2015). Comme le montre le tableau 6, l'évaluation PASEC2014 a concerné près de 40 000 élèves en début (2e année) et en fin de scolarité primaire (6e année) fréquentant plus de 1 800 écoles⁹ (PASEC, 2015).

Tableau 6 : Échantillon de l'évaluation PASEC2014 (PASEC, 2015)

Pays	Participants en début de scolarité		Participants en fin de scolarité	
	Ecoles	Elèves	Ecoles	Elèves
Bénin	80	732	165	3 033
Burkina Faso	100	969	182	3 416
Burundi	90	855	180	3 461
Cameroun	134	1 071	266	3 817
Congo	90	871	164	2 673
Côte d'Ivoire	88	840	169	2 972
Niger	86	813	176	3 196
Sénégal	83	807	160	2 905
Tchad	88	813	157	2 484
Togo	95	911	189	3 256
Total	934	8 682	1 808	31 213

Données collectées

Les élèves qui ont participé à cette évaluation PASEC2014 ont répondu aux questions des tests en langues et en mathématiques (PASEC, 2015). Ils ont également renseigné des

⁸ Le Cameroun a pour langues officielles l'anglais et le français.

⁹ Les **934** écoles, où se trouvent les élèves de 2^e années du primaire, font partie du total des **1 808** écoles.

questionnaires portant sur leurs caractéristiques individuels, leurs comportements, leurs attitudes, leurs stratégies d'études et leur milieu familial (PASEC, 2015). Les directeurs d'écoles et les enseignants ont, quant à eux, fourni des informations sur l'environnement de l'école et de la classe (PASEC, 2015). Le PASEC, en plus de ces données, a entrepris également de collecter des informations sur les pays participants notamment sur leur système éducatif, et leur organisation administrative (PASEC, 2015). La figure 5 présente le récapitulatif des différentes données de l'évaluation PASEC2014.

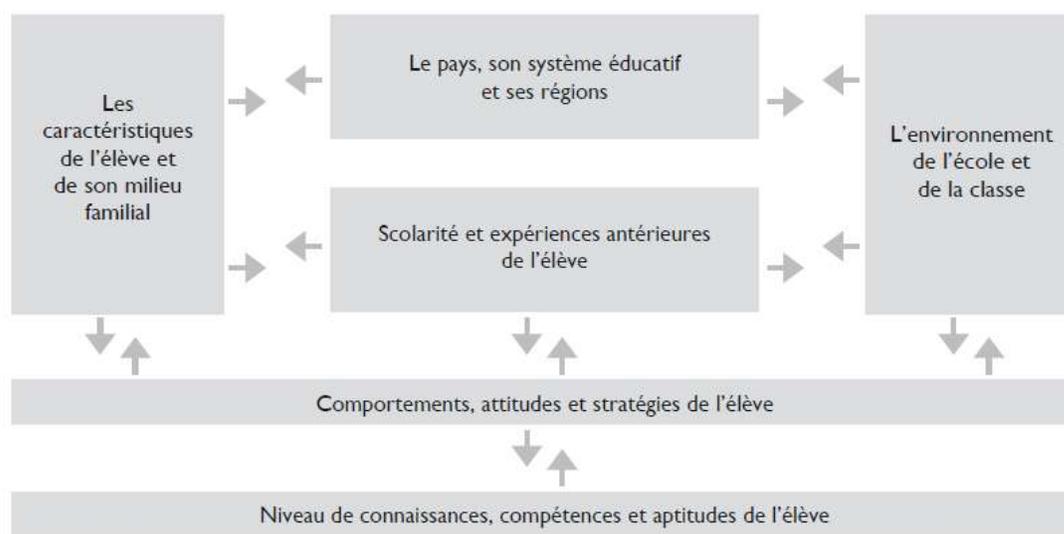


Figure 5 : Données collectées dans le cadre du PASEC2014 (PASEC, 2015, p.19)

Caractéristiques des tests cognitifs

En début de scolarité, les tests PASEC2014 mesurent les compétences acquises au cours des premiers apprentissages des tous petits (PASEC, 2015). En langue, ces tests ont porté sur la compréhension de l'oral, la familiarisation avec l'écrit et lecture-décodage et la compréhension de l'écrit (PASEC, 2015). En mathématiques les élèves ont répondu à des questions sur l'arithmétique et la géométrie (PASEC, 2015). Des administrateurs de tests, recrutés et formés pour l'occasion, ont administré individuellement les tests aux élèves pour une durée de 30 minutes par élève et par discipline (PASEC, 2015). Le tableau 7 présente un récapitulatif des sous-domaines évalués en début de scolarité ainsi que leurs poids respectifs dans ces tests.

Tableau 7 : Récapitulatif des sous-domaines évalués et leur poids en début de scolarité (PASEC, 2015)

	Sous-domaines évalués	Poids dans le test
Langue	Compréhension de l'oral	32%
	Familiarisation avec l'écrit et lecture-décodage	30%
	Compréhension de l'écrit	38%
Mathématiques	Arithmétique	65%
	Géométrie	35%

En fin de scolarité, les tests PASEC2014 de type « papier-crayon », mesurent des compétences qui doivent d'une part permettre aux élèves de comprendre, d'apprendre et de s'adapter à des situations quotidiennes courantes et d'autre part, leur permettre de poursuivre une scolarité post-primaire dans de meilleures conditions (PASEC, 2015). Ces élèves ont donc effectué des exercices sur le décodage de mots et de phrases isolés et sur la compréhension de textes pour ce qui est des tests de lecture (PASEC, 2015). Pour les tests de mathématiques, ils ont répondu à des questions sur l'arithmétique, la mesure et la géométrie et espace (PASEC, 2015). Pour chacune des disciplines, les élèves avaient 2 heures pour répondre à une série de questions à choix multiples (QCM). Le tableau 8 présente un récapitulatif des sous-domaines évalués en fin de scolarité ainsi que leurs poids respectifs dans ces tests.

Tableau 8 : Récapitulatif des sous-domaines évalués et leur poids en fin de scolarité (PASEC, 2015)

	Sous-domaines évalués	Poids dans le test
Lecture	Décodage de mots et de phrases isolés	26%
	Compréhension de textes	74%
Mathématiques	Arithmétique	47%
	Mesure	36%
	Géométrie et espace	17%

La conduite des analyses ACL, que nous avons choisi de mener sur les données PASEC2014, et l'interprétation des résultats de ces analyses est un travail minutieux et chronophage (Collins et Lanza, 2009 ; McCutcheon, 1987). Compte tenu de cet aspect et du temps alloué à la rédaction d'un mémoire de maîtrise, nous avons procédé à des choix sur les données pour nos analyses sans pour autant compromettre les résultats de ces analyses. Ainsi, nos analyses portent uniquement sur les données des tests de lecture des élèves francophones de la 6^e année du primaire au Cameroun. Un tel choix n'exclut pas des analyses ultérieures similaires à celles que nous avons mené sur l'ensemble des données PASEC2014 ou sur une autre partie de ces données. Ce choix va également dans le sens des travaux du PASEC, qui a mené des analyses sur l'ensemble de ses données incluant tous les pays, mais également mené

des analyses sur les données de chaque pays. La section suivante présente donc les données retenues pour nos analyses.

3.1.2 Données retenues pour nos analyses

Échantillon 6^e année francophone au Cameroun

Dans le but d'avoir suffisamment de données pour permettre des comparaisons, le PASEC a procédé au regroupement des écoles des 10 régions du Cameroun par strates relativement homogènes sur la base des caractéristiques socio-économiques et culturelles desdites régions (PASEC, 2015). Ainsi, dans le plan d'échantillonnage PASEC2014, les écoles francophones sont regroupés en trois strates : (i) la strate du Grand Nord qui contient des écoles des régions de l'Adamaoua, de l'Extrême-Nord et du Nord, (ii) la strate du Grand Centre composée d'écoles des régions du Centre, de l'Est et du Sud et (iii) la strate du Grand Ouest qui regroupe des écoles des régions du Littoral, du Nord-Ouest, de l'Ouest et du Sud-Ouest (PASEC, 2016). Dans les faits, 167 écoles francophones sont représentées à cette évaluation pour un total de 2 186 élèves francophones en 6^e année primaire (PASEC, 2016). Le tableau 9 présente la répartition des écoles et des élèves par strates.

Tableau 9 : Échantillon réalisé en 6^e année primaire au Cameroun (PASEC, 2016)

Strates	Ecoles	Elèves
Grand Ouest	50	851
Grand Centre	62	690
Grand Nord	55	645
Total	167	2 186

Les tests de lectures en fin de scolarité (6^e année primaire)

Comme nous l'avons mentionné plus haut, les tests de lecture en 6^e année PASEC2014 portent sur deux sous-domaines que sont : (i) le décodage de mots et de phrases isolés et (ii) la compréhension de textes. Pour le décodage de mots et de phrases isolés, les élèves devaient déchiffrer le sens de mots et de phrases isolés et procéder à des reconnaissances graphophonologiques de ces mots (PASEC, 2017a). En compréhension de textes, ils se sont attelés à extraire des informations explicites, à réaliser des inférences simples et à interpréter et combiner des informations (PASEC, 2017a). Le tableau 10 regroupe ces processus cognitifs par sous-domaines et présente leurs poids dans le test.

Tableau 10 : Poids des processus évalués en lecture et leurs poids dans le test PASEC2014 (PASEC, 2017a)

Processus cognitifs évalués	Poids des processus dans le test
Décodage de mots et de phrases isolés	
Déchiffrage du sens de mots et de phrases isolés	Non Défini
Reconnaissance graphophonologique de mots	Non Défini
Compréhension de textes	
Extraire des informations explicites	52%
Réaliser des inférences simples	21%
Interpréter et combiner des informations	27%

Organisation des items dans les tests PASEC2014

Le PASEC utilise la méthode des cahiers tournants pour ses différents tests (PASEC, 2017c). La méthode des cahiers tournants, couramment utilisée par les différentes ELE, permet de scinder un test en plusieurs blocs d'items et de les regrouper en des sous-ensembles de tests contenus dans des livrets (PASEC, 2017c). L'appellation « cahiers tournants » tient également du fait que l'ordre de passation (première et deuxième partie du test) des blocs d'items est organisé de façon à ce que chaque bloc d'items occupe tous les rangs d'un livret à l'autre (PASEC, 2017c). Ainsi, les tests de lecture PASEC2014 pour les élèves de 6^e année du primaire sont constitués de quatre livrets (A, B, C et D) disposant chacun de deux blocs d'items (PASEC, 2017c). De même, la distribution des livrets aux élèves est aléatoire et chaque élève ne répond qu'aux items d'un seul livret (PASEC, 2017c). Le tableau 11 présente la répartition des blocs d'items par livret pour les tests de lecture PASEC2014.

Tableau 11 : Répartition des blocs d'items des tests de lecture PASEC2014 dans les livrets (PASEC, 2017c, p. 32)

Livrets	Première partie	Deuxième partie
Livret A	Bloc 1	Bloc 2
Livret B	Bloc 2	Bloc 3
Livret C	Bloc 3	Bloc 4
Livret D	Bloc 4	Bloc 1

*Même couleur pour le Blocs de même nature pour en faciliter le repérage.

Les différents blocs contiennent des nombres identiques d'items qui sont censés être équivalents sur leur niveau de difficulté (PASEC, 2017a, 2017b, 2017c). Ces items sont répartis dans les blocs de façon à suivre une même progression graduelle de leur niveau de difficulté (PASEC, 2017a, 2017b, 2017c). Les blocs couvrent également des sous-domaines identiques et des processus cognitifs comparables (PASEC, 2017a, 2017b, 2017c). Dans ce sens,

l'ensemble des tests de lecture PASEC2014 est composé de 92 items à raison de 23 items par blocs et 46 items par livret (PASEC, 2017a, 2017b, 2017c). Le tableau 12 présente la répartition des items de lecture dans les différents livrets en fonction des processus cognitifs et des blocs.

Tableau 12 : Répartition des items de lecture PASEC2014 dans les livrets en fonction du processus évalué et des blocs (PASEC, 2017a, p. 25)

Processus cognitifs	Bloc 1	Bloc 2	Bloc 3	Bloc 4	Total
Décoder et reconnaître une information	6	5	6	6	23
Extraire des informations explicites	8	10	9	9	36
Réaliser des inférences simples	4	3	4	4	15
Interpréter et combiner des informations	5	5	4	4	18
Total	23	23	23	23	92

Blocs d'items du test de lecture et échantillon final retenus pour nos analyses

Le but de l'utilisation des cahiers tournants est de permettre aux ELE de tester un maximum d'items auprès des élèves sans pour autant allonger la durée prévue de passation de ces tests (PASEC, 2017a, 2017b, 2017c). Les items d'ancrage, qui sont les blocs d'items que partagent les livrets, rendent possible l'établissement d'un lien entre tous les livrets (PASEC, 2017a, 2017b, 2017c). Et au moment d'estimer les performances des élèves, le modèle de réponse à l'item (IRT) permet de placer tous les items du test sur une seule et même échelle et de prédire le score des élèves aux items qui n'étaient pas dans leur livret (PASEC, 2017a, 2017b, 2017c). Cependant, nous n'avons pas trouvé dans la littérature, la documentation qui souligne la possibilité d'établir de tel liens entre des items contenus dans des livrets différents à l'aide des modèles ACL. De plus, nous n'avons pas besoin de faire les analyses sur la totalité des livrets pour répondre à notre question de recherche. Pour cela, nos analyses ne porteront que sur les 23 items du bloc 1 des tests de lecture PASEC2014. Ce choix, comme le montre le tableau 13, reconfigure notre échantillon à un total de 565 élèves francophones de 6^e primaire.

Tableau 13 : Échantillon d'élèves ayant effectué les 23 items du bloc A des tests de lecture PASEC2014 et retenu pour nos travaux

Strates	Ecoles	Elèves
Grand Ouest	50	218
Grand Centre	61	178
Grand Nord	54	169
Total	165	565

Nous notons également que le contenu de la quasi-totalité des items des tests PASEC2014 sont malheureusement confidentiels (PASEC, 2015). Pour le test de lecture en fin

de scolarité par exemple, seuls 8 items sur les 92 items que compte le test sont libérés (PASEC, 2015). Ces 8 items servent en fait d'items d'illustration dans le rapport PASEC2014 et il n'est pas possible de faire le lien entre ces items et leur place dans un quelconque livret encore moins dans un bloc d'items (PASEC, 2015). Dans la base de données PASEC2014, seuls les numéros d'items sont présents ainsi que les différentes réponses des élèves. Le PASEC met cependant à la disposition des chercheurs un manuel d'utilisation de la base de données dans lequel sont indiqués les réponses justes. Cette situation nous impose donc d'analyser le test de lecture PASEC2014 comme un tout sans nous permettre d'analyser en détails les processus cognitifs sous-jacents aux items.

Les variables contextuelles retenues pour nos analyses

Les questionnaires contextuels soumis aux élèves, aux enseignants, aux directeurs d'écoles et aux responsables des ministères en charge de l'éducation ont permis au PASEC de recueillir une quantité importante d'informations. Les questionnaires destinés aux élèves en fin de scolarité primaire captent des informations sur les caractéristiques familiales des élèves, leur parcours scolaire, leurs ressentis à propos de leur bien-être à l'école, leur appréciation du travail qu'ils fournissent à l'école, leur goût pour la lecture et les mathématiques, leur santé et sur les ressources éducatives et les occasions d'apprentissage auxquels ils ont accès à domicile (PASEC, 2017a, 2017b, 2017c). Les questionnaires pour les enseignants de 6^e année primaire collectent des informations sur ces enseignants et sur la classe dans laquelle ils donnent cours à l'instar des infrastructures et des ressources éducatives (PASEC, 2017a, 2017b, 2017c). Les directeurs des écoles ont de leur côté répondu à des questions sur des thématiques en lien avec l'environnement scolaire et leur fonction (PASEC, 2017a, 2017b, 2017c). En plus de toutes ces données, les différents pays ayant pris part à l'évaluation PASEC2014 ont fourni des informations relatives aux stratégies éducatives et à leur mise en œuvre (PASEC, 2017a, 2017b, 2017c). La classification PASEC regroupe l'ensemble de ces données contextuelles en trois différents niveaux : (i) le niveau « élève et son milieu familial » ; (ii) le niveau « école, classe et communauté locale » et (iii) le niveau « pays, son système éducatif et ses zones éducatives ». Le tableau 14 présente cette classification.

Tableau 14 : Classification PASEC de ses données contextuelles par niveau (PASEC, 2017c)

Niveau	Regroupement	Données contextuelles
Niveau 1	L'élève et son milieu familial	Sexe
		Âge
		Indice socioéconomique
		Redoublement
		Préscolaire
Niveau 2	L'environnement de la classe	Indice de ressources pédagogiques de la classe
		Sexe de l'enseignant
		Niveau universitaire de l'enseignant
		Ancienneté de l'enseignant
Niveau 3	L'environnement de l'école	Statut de l'école
		Zone de l'école
		Sexe du directeur
		Niveau universitaire du directeur

Dans le plan d'analyse des données PASEC2014, les indices représentent des variables calculées à partir d'un groupe de variables. C'est le cas entre autres de l'indice socioéconomique de la famille de l'élève, et l'indice de ressources pédagogiques de la classe que nous présenterons plus bas. De même, le PASEC en concertation avec les pays participants a procédé à des choix de variables contextuelles d'intérêt en fonction de pays. Dans le cadre de nos analyses, nous avons donc utilisé des variables contextuelles d'intérêt choisi par le PASEC pour les analyses des données du Cameroun et des variables qui semblent être pertinentes pour représenter certains niveaux du modèle écologique de Zumbo et ses collègues (2015) ; le modèle qui nous sert de référence dans nos analyses. Ce choix répond à la fois à une volonté de rendre possible la comparaison entre nos résultats et ceux du PASEC2014 pour le Cameroun, mais aussi à une volonté de rester conforme au plan d'analyse que nous avons choisi.

➤ ***Niveau 1 : L'élève et son milieu familial***

Au niveau de l'élève et de son milieu familial, les variables contextuelles PASEC2014-Cameroun sont le sexe de l'élève, son âge, le fait qu'il ait redoublé ou pas (redoublement), le fait qu'il ait fait le préscolaire ou pas (préscolaire) et l'indice socio-économique de sa famille. La réduction des inégalités d'accès et d'apprentissage entre les filles et les garçons, le redoublement, l'accès au préscolaire sont est en effet des problématiques qui préoccupent les responsables en charge de l'éducation au Cameroun (PASEC, 2015). L'indice socioéconomique de la famille de l'élève renferme des informations sur la disponibilité de biens

matériels dans les ménages et les caractéristiques de l'habitation (PASEC, 2015). La disponibilité de biens matériels renvoie notamment au nombre de livres à la maison ainsi qu'à la possession de biens d'équipements à l'instar du téléviseur, de l'ordinateur, de la radio, et des biens similaires (PASEC, 2015). La disponibilité de biens matériels renvoie également à la possession de biens durables et de moyens de transport comme la machine à coudre, le fer à repasser les engins motorisés et autres (PASEC, 2015). Les caractéristiques de l'habitation font référence aux matériaux utilisés pour la construction de l'habitat, la présence de latrines avec ou sans eau courante, la présence de l'électricité à la maison et la présence d'un puits ou d'un robinet d'eau courante à la maison (PASEC, 2015). L'indice socioéconomique ainsi calculé repose sur une échelle dont la moyenne est de 50 et l'écart-type de 10 (PASEC, 2015). Une valeur élevée de l'indice correspond à une condition de vie favorable tandis qu'une valeur faible est associée à un ménage défavorisé (PASEC, 2015). L'annexe 1 présente les différentes variables ayant servi au calcul de l'indice socioéconomique de la famille.

➤ *Niveau 2 : L'environnement de la classe*

Le niveau 2 de la classification des données contextuelles PASEC regroupe l'indice de ressources pédagogiques de la classe, le sexe de l'enseignant, le fait qu'il ait un niveau universitaire ou pas et son ancienneté (en année). L'indice de ressources pédagogiques de la classe regroupe des informations sur le niveau d'équipement de la classe (voir annexe 2). C'est notamment la disponibilité des manuels pour les élèves, les documents et matériels pédagogiques pour les enseignants et le mobilier de classe (PASEC, 2015). Au niveau des manuels, le PASEC a collecté des informations sur le nombre de manuels de mathématiques et de lecture disponibles par élève (PASEC, 2015). Pour les documents et le matériel pédagogique, le programme a recensé le nombre de guides pédagogiques et de programmes de lecture et de mathématiques pour les enseignants (PASEC, 2015). Le matériel pédagogique recensé par le PASEC est constitué du tableau, des craies, du dictionnaire, de l'horloge, des cartes du monde, de l'Afrique et du pays, et du matériel de mesure tel que l'équerre, le compas et la règle (PASEC, 2015). Pour le mobilier de classe, le PASEC a réuni des informations sur la disponibilité d'un bureau et d'une chaise pour le maître, de l'armoire et des étagères de rangement pour les livres, d'un coin de lecture et des tables-bancs en nombre suffisant (PASEC, 2015). Ces données sont synthétisées sur une échelle ayant une moyenne de 50 et un écart-type de 10 dans le but de construire l'indice d'équipement de la classe (PASEC, 2015).

➤ *Niveau 3 : L'environnement de l'école*

Au niveau de l'environnement de l'école, les données sont constituées du statut de l'école, de la zone où est située l'école, de l'indice d'aménagement de la localité de l'école, du sexe du directeur et du fait qu'il ait un niveau universitaire ou pas. Au Cameroun, une école peut avoir un statut public, privé ou communautaire¹⁰ (PASEC, 2015). Cependant, les écoles communautaires accueillent moins de 2% d'élèves du pays et n'ont pas été prises en compte dans le plan d'échantillonnage du PASEC (PASEC, 2015). Les écoles sont également situées soit en zone urbaine, soit en zone rurale (PASEC, 2015). Le PASEC a calculé l'indice d'aménagement de la localité de l'école sur la base de la présence dans la localité de certains biens et services que sont la route goudronnée, l'électricité, le collège, le lycée, l'hôpital, le centre de soin ou de santé, le poste de gendarmerie ou de police, la banque, la caisse d'épargne, le bureau de poste, le centre culturel ou la bibliothèque (PASEC, 2015). Une grande valeur de cet indice signale une zone bien aménagée et une faible valeur de l'indice signale une zone faiblement aménagée. L'annexe 3 regroupe les différentes variables utilisées pour le calcul de l'indice d'aménagement de la localité de l'école.

Comme nous l'avons mentionné plus haut, aux données contextuelles PASEC2014 contenues dans le plan d'analyse des résultats du Cameroun, nous avons sélectionné d'autres variables contextuelles pour représenter toutes les couches du modèle écologique de Zumbo et ses collègues (2015). Pour rappel, ce modèle est constitué de 5 couches : (i) le format et le contenu du test, (ii) les caractéristiques individuelles, (iii) l'environnement scolaire, (iv) l'environnement extrascolaire et (v) les caractéristiques de la communauté. Nous avons ainsi retenu la variable « langue » qui représente le fait qu'un élève parle ou pas la Langue d'enseignement (français) à la maison. Cette variable peut être logée dans la cinquième couche du modèle écologique, l'environnement extrascolaire. Cette couche renferme en effet comme nous l'avons présenté dans les sections précédentes, les personnes et les choses que l'individu côtoie quotidiennement en dehors du contexte scolaire et les différentes interactions entre cet individu et ces personnes ou ces choses.

Pour la pratique de la langue d'enseignement, le Cameroun, comme la plupart des pays d'Afrique subsaharienne d'ailleurs, est marqué par un contexte national plurilingue assez riche (Alidou, 2006 ; Kody, 2005). En raison de cette diversité et des usages quotidiens de ces langues

¹⁰ Les écoles communautaires au Cameroun sont des écoles gérées par des associations communautaires.

par les différentes communautés, la pratique de la langue d’enseignement reste donc une dimension importante à considérer pour cerner le contexte d’apprentissage des élèves dans ces pays (PASEC, 2015).

Nous avons également intégré dans nos analyses la variable indice d’implication de la communauté. Cette variable se classe dans la dernière couche du modèle écologique de Zumbo et ses collègues (2015) à savoir les caractéristiques de la communauté. L’indice d’implication de la communauté synthétise 22 autres variables contextuelles (voir annexe 4 pour le détail complet). Dans cette synthèse figure la présence ou non au sein de l’école d’une association de parents d’élèves et d’enseignants (APEE), d’une association des mères éducatrices (AME), d’une coopérative scolaire, d’un conseil d’école ou encore d’un comité de gestion (PASEC, 2017b). A ces variables s’ajoutent également des variables qui soulignent différentes collaborations entre l’école et la collectivité locale¹¹, la fréquence de ces collaborations au cours d’une année scolaire, et divers appuis que la collectivité apporte aux écoles (PASEC, 2017b).

Le regroupement des variables contextuelles selon le modèle écologique de Zumbo et ses collègues (2015) nous a également amené à redistribuer les différentes variables dans les différentes couches du modèle. La couche 3 du modèle écologique regroupe par exemple toutes les variables de l’environnement scolaire et donc des variables niveau école et niveau classe du regroupement PASEC. Le tableau 15 présente un sommaire des variables utilisées dans nos travaux et leur classification dans le modèle écologique de Zumbo et ses collègues (2015). Ce tableau contient également une description de ces variables, leur format et leurs modalités.

Tableau 15 : Sommaire des variables utilisées et leur classification dans le modèle écologique de Zumbo et ses collègues (2015)

Couche	Variable	Libellé	Format	Modalités
Couche 1 : format et contenu du test	f1-f23	Items f1 à f3	Binaire	1 : Bonne réponse 2 : Mauvaise réponse
	Sexe	Sexe de l’élève	Binaire	0 : Fille 1 : Garçon
Couche 2 : caractéristiques individuelles	Age		Continue	Numérique
	SES	Indice socioéconomique	Continue	Numérique
	Redoublement		Binaire	Numérique
Couche 3 : environnement scolaire	Préscolaire	Genre de l’élève	Binaire	1 : Fille 2 : Garçon
	Sexe_ens	Sexe de l’enseignant	Binaire	0 : Femme 1 : Homme
	Niveau_univ_ens	L’enseignant a un niveau universitaire	Binaire	0 : Non 1 : Oui

¹¹ APEE, conseil d’école, comité de gestion ou autre association locale de la société civile

Couche	Variable	Libellé	Format	Modalités
Couche 4 : environnement extrascolaire Couche 5 : caractéristiques de la communauté	Anc_ens	Ancienneté de l'enseignant (en années)	Continue	Numérique
	Sexe_dir	Sexe du directeur	Binaire	0 : Femme 1 : Homme
	Niveau_univ_dir	Le directeur a un niveau universitaire	Binaire	0 : Non 1 : Oui
	Zone	Zone où est située l'école	Nominale	0 : Rurale 1 : Urbaine
	Indice_resso_peda_mt	Indice de ressources pédagogiques de la classe	Continue	Numérique
	Indice_resso_peda_ecol	Indice de ressources pédagogiques de l'école	Continue	Numérique
	Langue	Langue française parlée à la maison	Binaire	1 : Toujours ou parfois 2 : Jamais
	lindice_impli_communau	Indice d'implication de la communauté	Continue	Numérique
Indice_amenag_terri	Indice d'aménagement du territoire	Continue	Numérique	

3.2 Démarche d'analyse

3.2.1 Différentes étapes

Comme annoncé dans les sections précédentes, les deux questions spécifiques de recherche auxquelles nous avons décidé de répondre sont :

- (i) *Comment interpréter les résultats du test en lecture PASEC2014-Cameroun lorsqu'on les analyse à partir de l'approche écologique de Zumbo et ses collègues (2015) en utilisant des analyses des classes latentes avec covariables ?*
- (ii) *Comment la compréhension des résultats des tests de lecture PASEC-Cameroun2014 est-elle influencée par l'usage de l'approche écologique de Zumbo et ses collègues (2015) en comparaison avec les résultats PASEC ?*

Les données qui nous ont servi pour nos analyses et que nous venons de présenter sont constituées des résultats du livret A au test de lectures PASEC2014 de 565 élèves francophones de 6^e année primaire au Cameroun et des données contextuelles pour trois strates. Les trois strates sont entre autres la strate du Grand Ouest, du Grand Centre et du Grand Nord. Pour répondre à nos deux questions de recherche, nous avons donc procédé, dans un premier temps, à des analyses descriptives de ces données à la fois sur l'échantillon global, mais également par strate. Les analyses descriptives sur l'échantillon global donnent une idée générale sur nos données tandis que les analyses par strate font ressortir des variabilités plus ou moins prononcées sur certaines variables entre les strates. Compte tenu du nombre différent d'élèves

par strate dans l'échantillon, nous avons tenu à rapporter toute différence statistiquement significative et à mentionner celles qui ne le sont pas. Les deux tests utilisés pour analyser ces différences sont : l'analyse de la variance (ANOVA) pour les moyennes et le test de Khi deux pour les proportions. Nous avons tenu également à rapporter la taille des effets pour ces tests.

Pour les ANOVA, nous rapportons la taille de l'effet r qui est la racine carrée du rapport de la somme des carrés inter-groupes par la somme totale des carrés. Cet indice varie entre 0 et 1 et les balises suivantes ont été élaborées par Cohen (1988) pour guider son interprétation : (i) Autour de 0,1 : effet de petite taille ; autour de 0,3 : effet de taille moyenne et (iii) autour de 0,5 et plus : effet de grande taille¹². Pour les tests de khi2 nous rapportons le V de Cramer dont la valeur se situe également entre 0 et 1. De même que pour la taille d'effet pour l'ANOVA, le V de Cramer s'interprète de la façon suivante : (i) Autour de 0,1 : effet de petite taille ; autour de 0,3 : effet de taille moyenne et (iii) autour de 0,5 et plus : effet de grande taille¹³.

La deuxième partie de nos analyses, qui en est la principale, porte sur l'analyse des classes latentes (ACL). D'une part, nous avons mené des analyses de l'invariance de mesure en fonction des strates. Cette partie des analyses nous a permis de déterminer qu'il était plus judicieux de mener des ACL séparées par strates que de mener une ACL sur les données globales. D'autre part, nous avons conduit des analyses ACL avec covariables sur nos données. Cette analyse nous a permis de dégager les profils de performances par strate et l'écologie de réponse au test de lecture PASEC2014 par strate.

Nous portons ensuite un regard écologique sur les résultats du test de lecture PASEC2014. Ce regard est le lieu pour nous de mettre en exergue des nouveaux éclaircissements et les différents apports d'une analyse écologique sur les données PASEC2014. Pour finir, nous présentons quelques constats et esquissons quelques recommandations à la lumière de nos différents résultats. La figure 6 récapitule les différentes étapes de la recherche et les résultats obtenus.

¹² <http://spss.espaceweb.usherbrooke.ca/pages/statistiques-inferentielles/analyse-de-variance.php>

¹³ <http://spss.espaceweb.usherbrooke.ca/pages/statistiques-inferentielles/test-de-chi-2.php>

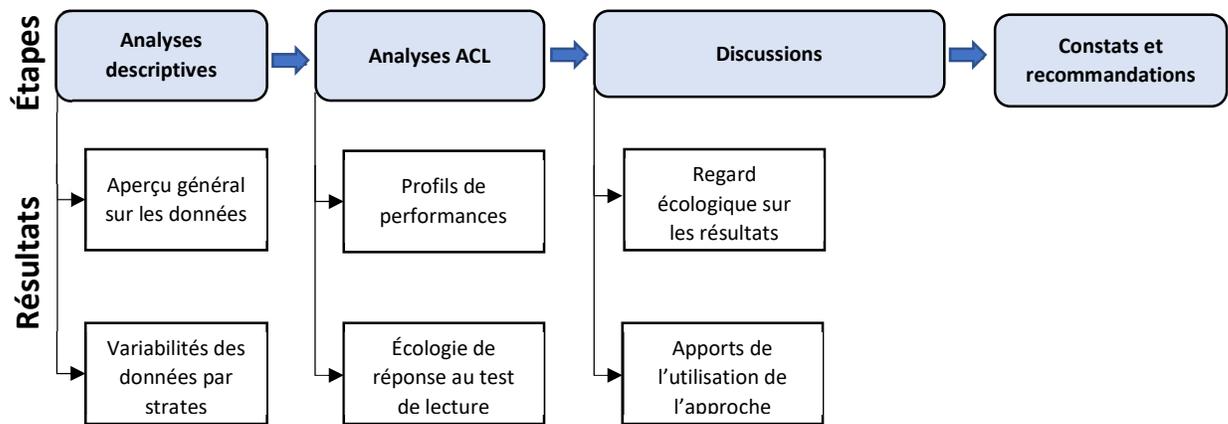


Figure 6 : Etapes de la recherche et résultats obtenus

3.2.2 Outils d'analyse

Les principaux logiciels que nous avons utilisé dans le cadre de notre recherche sont entre autres : le Statistical Package for the Social Sciences (SPSS), le Statistical Analysis System (SAS) et le tableur Excel. L'Université de Montréal met gratuitement à la disposition de ses étudiants les logiciels SPSS et SAS. De nos jours, beaucoup de chercheurs en sciences sociales utilisent ces logiciels qui ont l'avantage d'être facile d'utilisation et de permettre également une interprétation rapide et facile des résultats d'analyses à l'aide de la gamme de graphiques dont ils disposent.

3.3 Considérations éthiques

La mise à disposition des données PASEC2014 aux chercheurs et étudiants est soumise à certaines conditions qui peuvent être lues sur le site internet du PASEC à l'adresse <http://www.pasec.confemen.org/donnees>. Sur cette adresse internet, les termes suivants apparaissent dans la rubrique « conditions d'utilisation » :

- Informer le PASEC et, le cas échéant, fournir les résultats (publication, document de travail, article scientifique, etc.) issus de l'exploitation des données obtenues ;
- Citer clairement le PASEC dans ses références de ses travaux ;
- Si ses travaux de recherche sont disponibles sur internet, le PASEC se propose, après approbation de l'utilisateur, de diffuser le lien sur le site internet de la CONFEMEN et du PASEC ;
- L'utilisation faite des données n'engage que leur utilisateur et en aucun cas le PASEC ne pourra en être tenu responsable.

Les données PASEC2014 sont des données anonymes dépouillées de toute indication qui pourrait permettre de reconnaître une école participante ou un élève. Des codes numériques permettent de représenter des écoles et des élèves. Nous avons également obtenu auprès des responsables de l'Université de Montréal un certificat d'approbation éthique pour notre recherche. Le comité d'éthique de la recherche en éducation et en psychologie après analyse de notre devis de recherche nous a en effet délivré le certificat No CEREP-20-006-D.

CHAPITRE 4 : RÉSULTATS

4.1 Description de l'échantillon

4.1.1 Echantillon global

Cette partie présente brièvement les caractéristiques de notre échantillon global¹⁴ en décrivant les élèves, leurs lieux d'accueil et le personnel qui les encadre. Comme nous l'avons mentionné dans la section précédente, notre échantillon se compose de 565 élèves francophones de sixième année du primaire qui proviennent de 165 écoles différentes (Tableau 13). Parmi les 565 élèves, 54,2% sont des garçons et 45,8% sont des filles (Figure 7). Les élèves sont âgés de 9 à 18 ans et ont une moyenne d'âge de 12,3 ans. Cette moyenne d'âge est de 12,4 ans pour les garçons et de 12,2 ans pour les filles. Les figures 8 et 9 montrent la distribution des âges dans l'échantillon global, et par sexe.

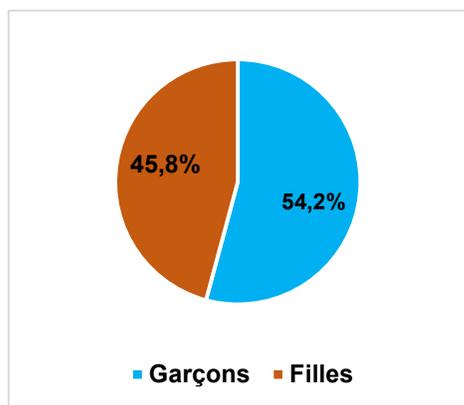


Figure 7 : Répartition des élèves par sexe

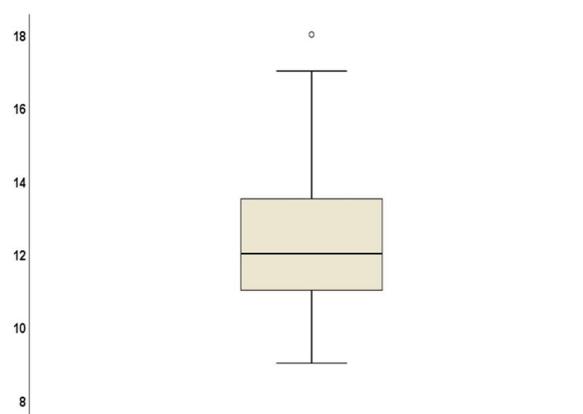


Figure 8 : Distribution des âges des élèves de l'échantillon global

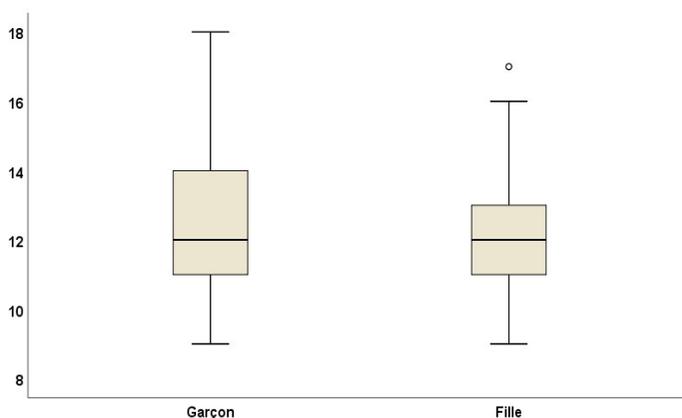


Figure 9 : Distribution des âges des élèves par sexe

¹⁴ L'annexe 6 contient les statistiques descriptives pour toutes nos données.

L'examen du parcours scolaire des élèves révèle qu'un peu plus de la moitié de ceux-ci (56%) a déjà redoublé au moins une fois (Figure 10). Ce pourcentage est plus élevé chez les garçons (63%) que chez les filles (48%). Moins de la moitié de ces élèves (41,9%) est passé par le préscolaire (Figure 11). Cette proportion est plus importante (78,6%) dans les écoles situées en zones urbaines comparativement à celles situées en zones rurales où seul un élève sur quatre a eu à faire le préscolaire. De même, la proportion des filles ayant fait le préscolaire (46,8%) est légèrement plus importante que celle des garçons (37,6%).

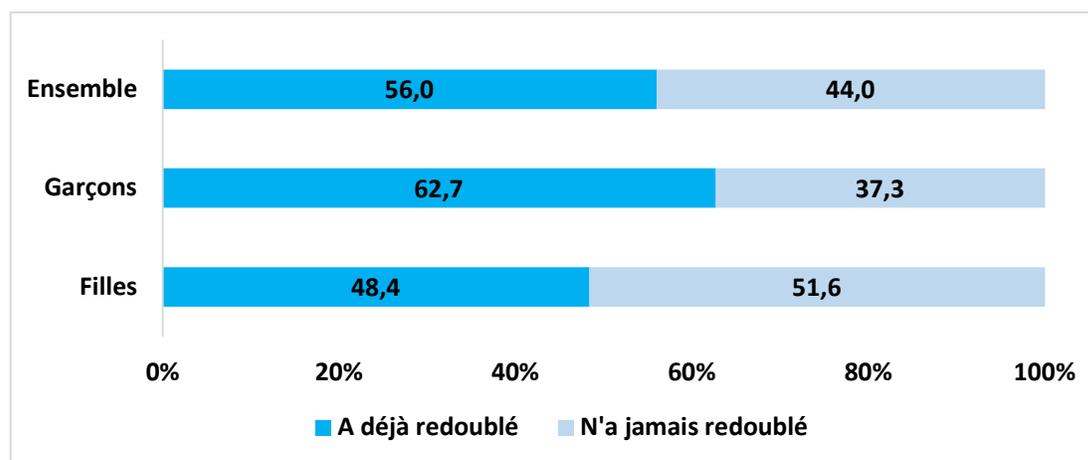


Figure 10 : Pourcentages d'élèves en fonction du redoublement

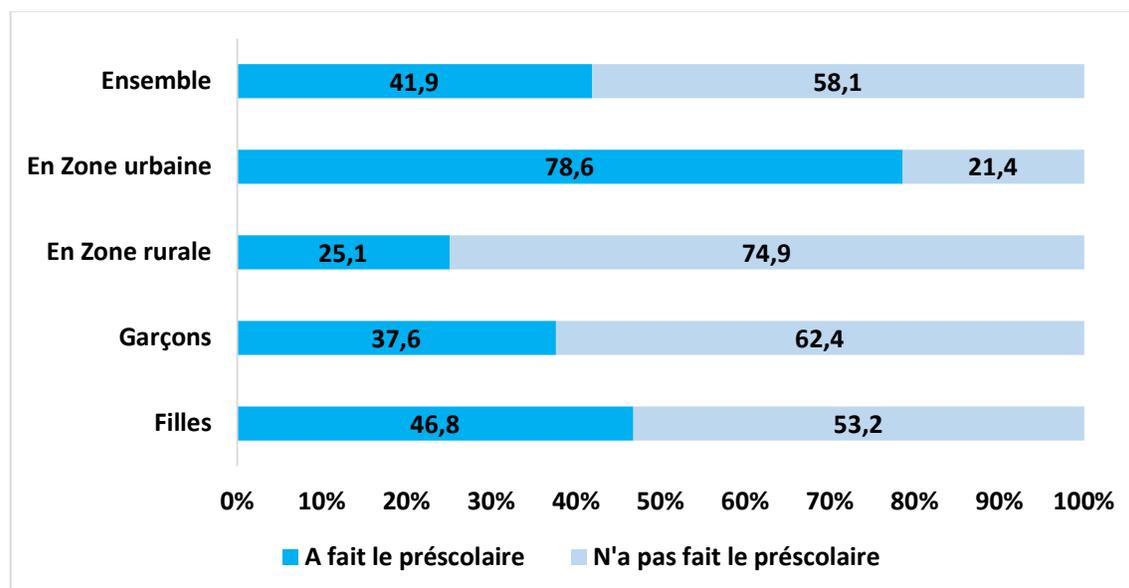


Figure 11 : Pourcentage d'élèves en fonction de la fréquentation du préscolaire

Au niveau des structures d'accueil, les écoles publiques encadrent 70,3% d'élèves de l'échantillon contre 29,7% pour les écoles privées (Figure 12). Une forte proportion (68%) de ces élèves fréquente une école en zone rurale contre 32% en zone urbaine (Figure 13).

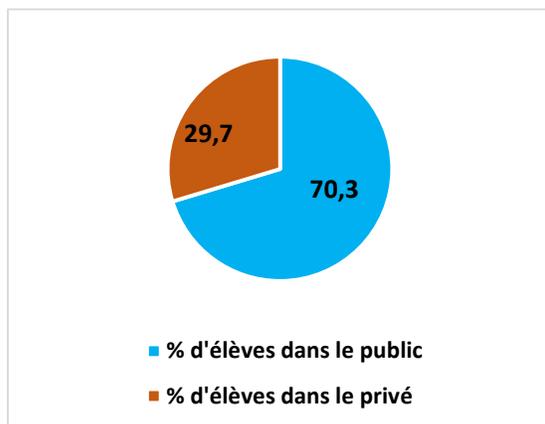


Figure 12 : Répartition des élèves en fonction du statut de l'école

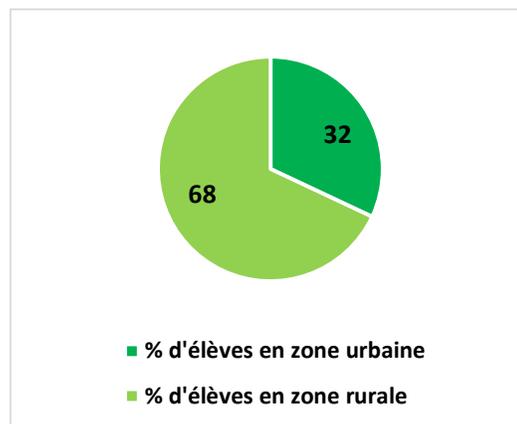


Figure 13 : Répartition des élèves en fonction de la zone de fréquentation

Au niveau de l'encadrement, 82,4% des enseignants sont des hommes et 17,6% des femmes (Figure 14). Des proportions similaires se reflètent pour les directeurs d'écoles (Figure 15) dont 83% sont des hommes et 17% des femmes. De faibles proportions d'enseignants (17%) et de directeurs (20%) ont un niveau universitaire. Le nombre d'années d'expérience des enseignants, affiche une moyenne de 11 années. La figure 16 présente la distribution des années d'expérience des enseignants au sein de l'échantillon.

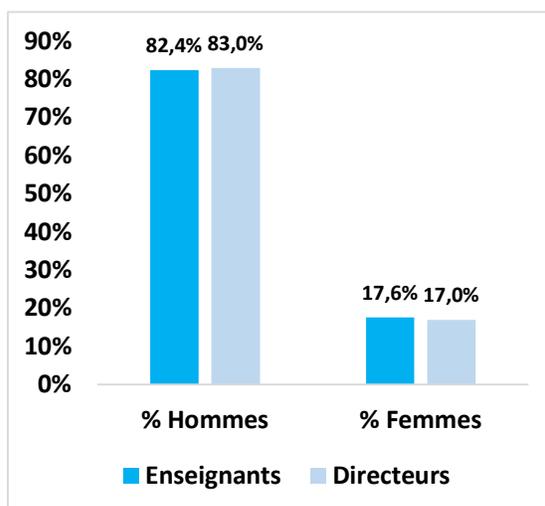


Figure 14 : Pourcentages d'enseignants et de directeurs par sexe

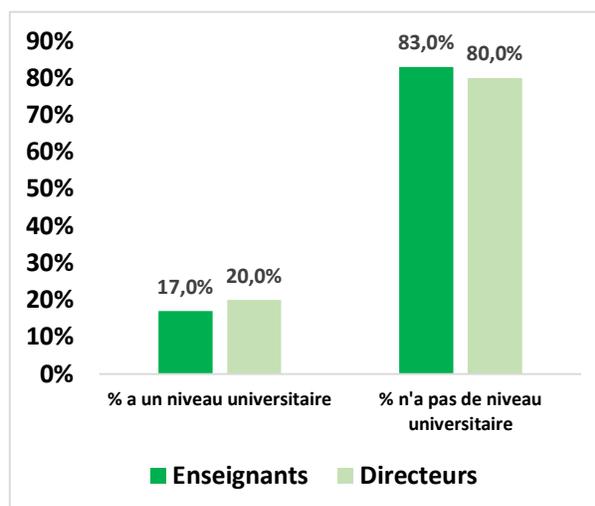


Figure 15 : Pourcentages d'enseignants et de directeurs en fonction du niveau d'études atteint

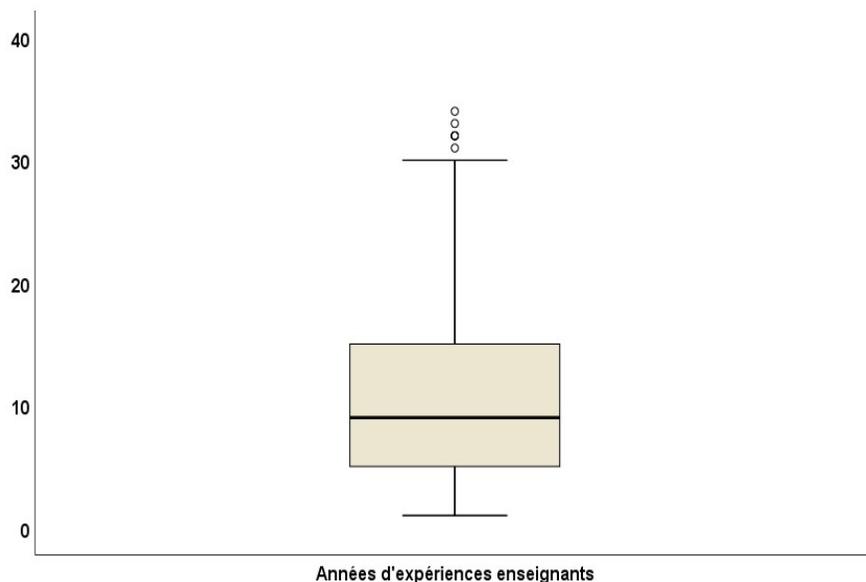


Figure 16 : Distribution des années d'expérience des enseignants

Comme nous l'avons présenté dans les sections précédentes, l'approche écologique de Zumbo et ses collègues (2015) soutient l'idée d'une performance située. Nous avons également souligné que le plan d'échantillonnage PASEC2014 pour les écoles francophones du Cameroun est constituée de trois strates qui regroupent chacune des écoles des régions qui partagent certaines caractéristiques sur le plan économique, linguistique et culturel. Ces trois strates sont les strates Grand Ouest, Grand Centre et Grand Nord. Après la présentation de l'échantillon global, nous présentons donc dans la partie suivante, les statistiques descriptives dans les strates pour les données contextuelles. Cette présentation est faite en regroupant les différentes variables par niveau du modèle écologique. Notons que la couche 1 qui inclut le contenu du test, les items de lecture dans notre cas, est présentée à la suite de cette section.

4.1.2 Caractéristiques personnelles des élèves dans les strates (Couche 2)

Comme nous l'avons présenté dans le cadre méthodologique de notre travail, notre échantillon de 565 élèves se subdivise comme suit : 218 élèves dans la strate Grand Ouest, 178 élèves dans la strate Grand Centre et 169 élèves dans la strate Grand Nord (Tableau 13). La strate Grand Nord est celle qui contient la plus grande proportion de garçons (62,7%) suivie de la strate Grand Centre (53,4% de garçon) et de la strate Grand Ouest qui compte 48,2% de garçon (Figure 17). Ces différences sont statistiquement significatives ($p=0,02$), mais de faible effet (V de Cramer=0,1).

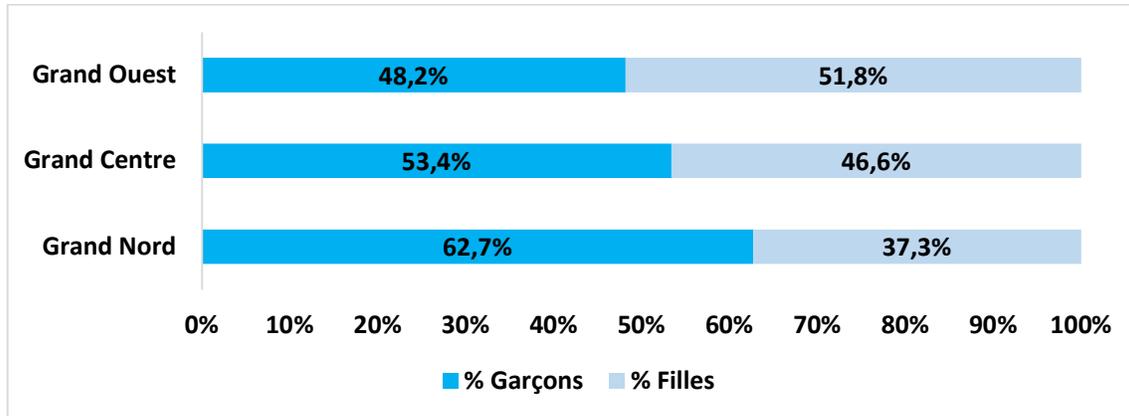


Figure 17 : Répartition des élèves par sexe dans les strates

Les moyennes d'âge des élèves diffèrent aussi en fonction des strates ($p < 0,05$ et $r = 0,3$). Les élèves de la strate Grand Ouest semblent être moins âgés (11,8 ans en moyenne) que ceux des strates Grand Centre (12,3 ans en moyenne) et Grand Nord (12,9 ans en moyenne). La figure 18 présente la distribution des âges des élèves dans les différentes strates.

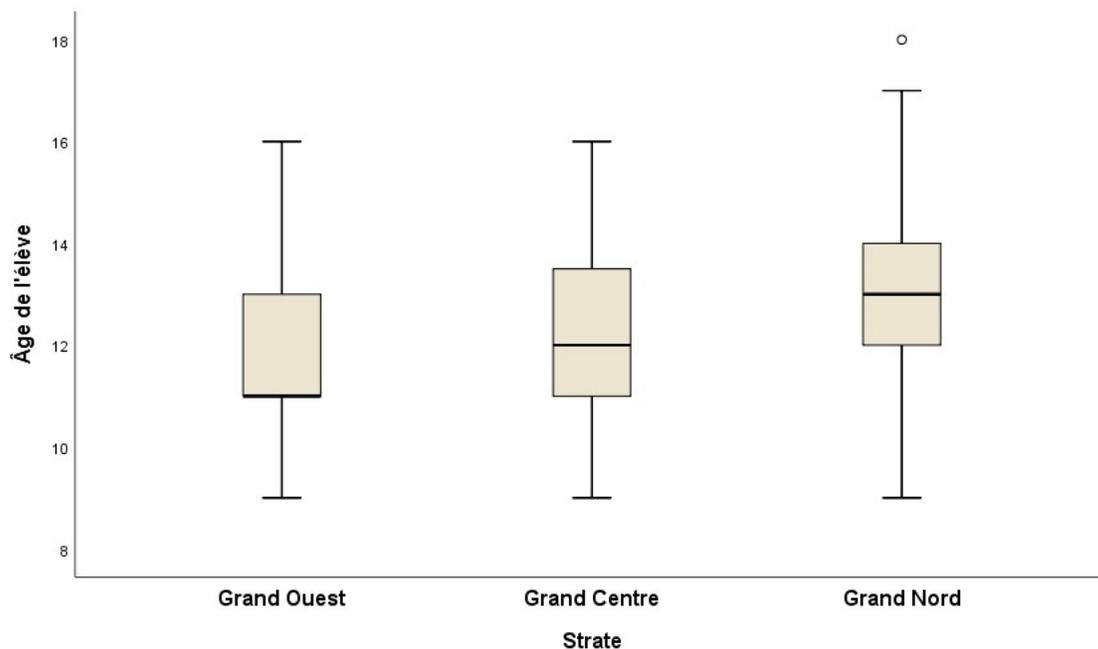


Figure 18 : Distribution des âges des élèves dans les strates

Les strates Grand Centre et Grand Nord comptent plus de redoublants comparativement à la strate Grand Ouest ($p < 0,01$ et V de Cramer = 0,2). Les élèves des strates Grand Centre et Grand Nord qui ont déjà redoublé au moins une fois représentent respectivement 63,5% et 61,1% des effectifs dans ces strates. À l'opposé, dans la strate du Grand Ouest, c'est 46,1% de l'effectif qui est dans cette situation (Figure 19)

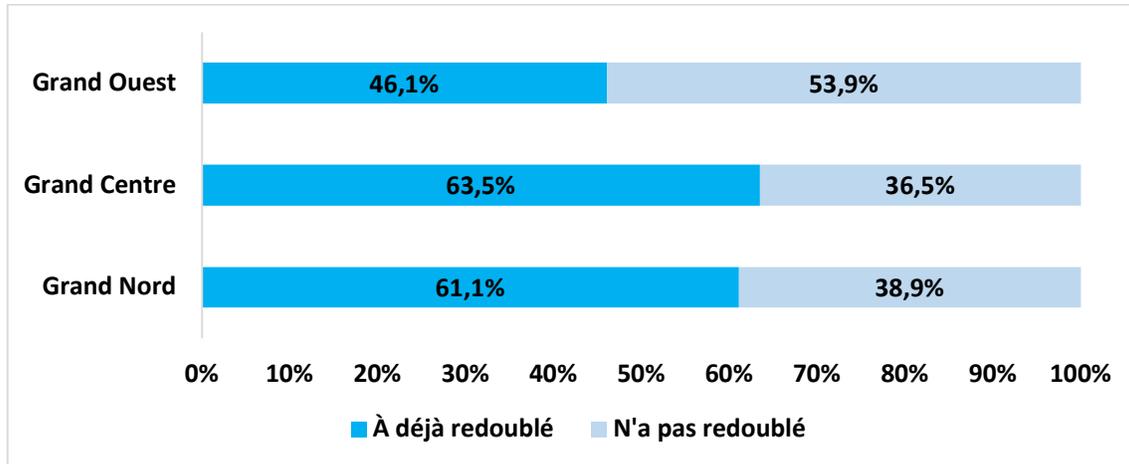


Figure 19 : Pourcentages d'élèves en fonction du redoublement dans les strates

La fréquentation du préscolaire présente une variabilité considérable (Figure 20) entre les élèves des différentes strates ($p < 0,01$ et V de Cramer=0,5). Dans la strate Grand Ouest, 60,6% d'élèves sont passés par le préscolaire tandis que seuls 7,5% d'élèves de la strate Grand Nord ont fait le préscolaire. Dans la strate Grand Centre, c'est la moitié (51,5%) des élèves qui est passée par le préscolaire.

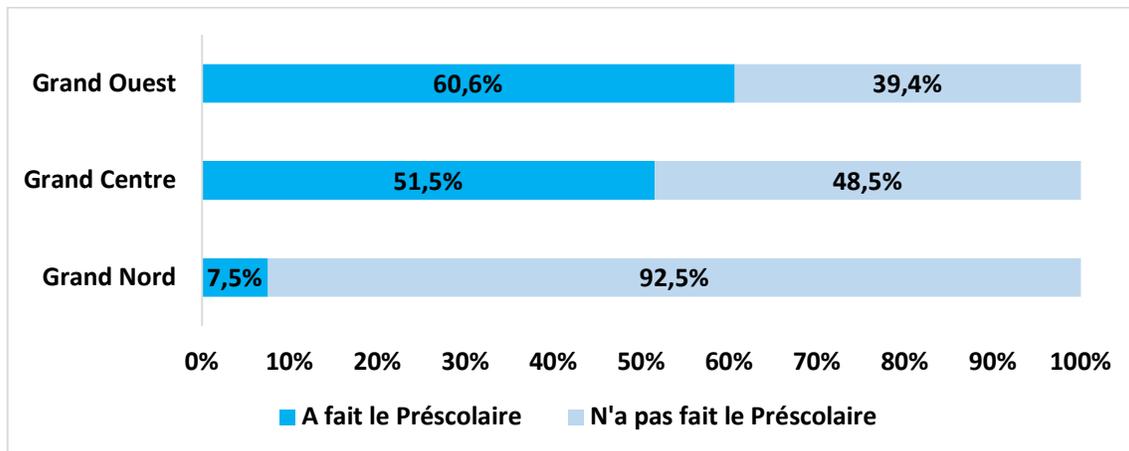


Figure 20 : Pourcentage d'élèves en fonction de la fréquentation du préscolaire dans les strates

De même, le niveau socioéconomique moyen des élèves diffère significativement en fonction des strates ($p < 0,01$ et $r = 0,3$). Il est plus élevé au Grand Ouest et se situe à 56,6 tandis qu'il est de 50,8 au Grand Centre et 44,4 au Grand Nord. La figure 21 présente la distribution du statut socio-économique des élèves dans les différentes strates.

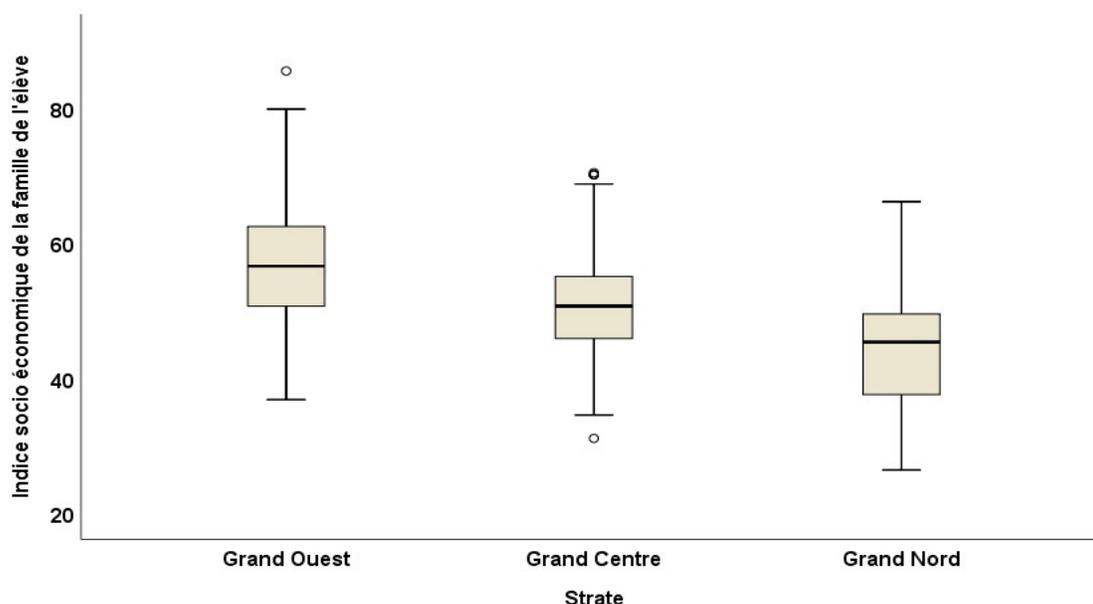


Figure 21 : Distribution du statut socio-économique des élèves dans les strates

Rapporté au plan national, plus de la moitié des élèves (58,6%) issus des milieux les plus défavorisés proviennent de la strate Grand Nord. Cependant, les deux tiers des élèves (70%) issus des milieux les plus favorisés sont dans la strate Grand Ouest. Le tableau 16 présente la répartition des 25% des élèves issus des milieux les plus défavorisés et des 25% des élèves issus des milieux les plus favorisés en fonction des strates.

Tableau 16 : Répartition des 25% des élèves issus des milieux les plus défavorisés et des 25% des élèves issus des milieux les plus favorisés en fonction des strates

	Strate			Total
	Grand Ouest	Grand Centre	Grand Nord	
25% des élèves issus des milieux les plus défavorisés	11,3%	30,1%	58,6%	100,0%
25% des élèves issus des milieux les plus favorisés	70,4%	21,1%	8,5%	100,0%

4.1.3 Environnement scolaire, classe, enseignants et directeurs dans les strates (Couche 3)

Plus de la moitié (57,3%) des élèves de la strate du Grand Ouest est scolarisée dans des écoles situées en zone urbaine alors qu'ils ne sont que 24,7% et 7,1% respectivement dans les

strates du Grand Centre et du Grand Nord (Figure 22). Cette disparité est assez marquée ($p < 0,01$ et V de Cramer=0,5).

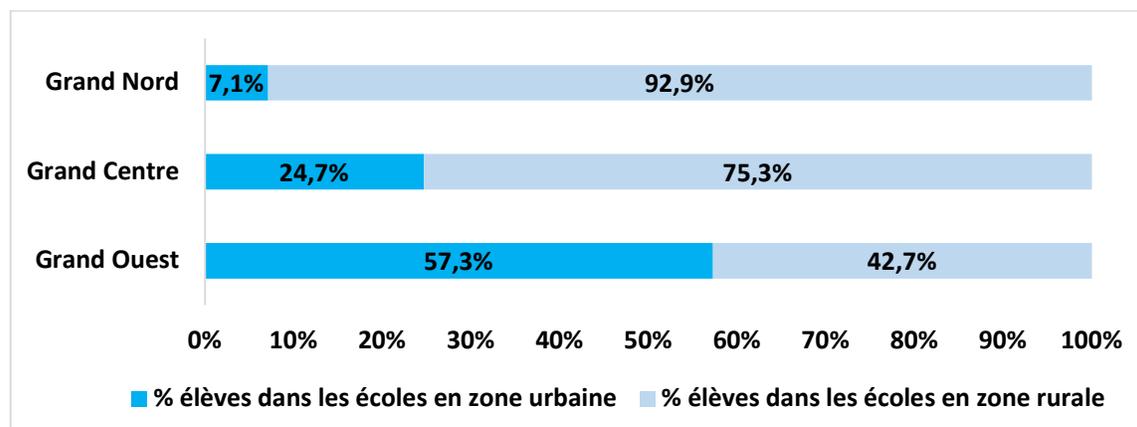


Figure 22 : Répartition des élèves en fonction de la zone de fréquentation dans les strates

Les écoles qui accueillent les élèves sont majoritairement des écoles publiques dans toutes les strates (Figure 23), malgré quelques légères différences ($p < 0,01$ et V de Cramer=0,2). Ainsi, ce sont 89,9% d'élèves qui sont scolarisés dans une école publique dans la strate Grand Nord, 65,7%, d'élèves dans la strate Grand Centre et 58,7% d'élèves dans la strate du Grand Ouest.

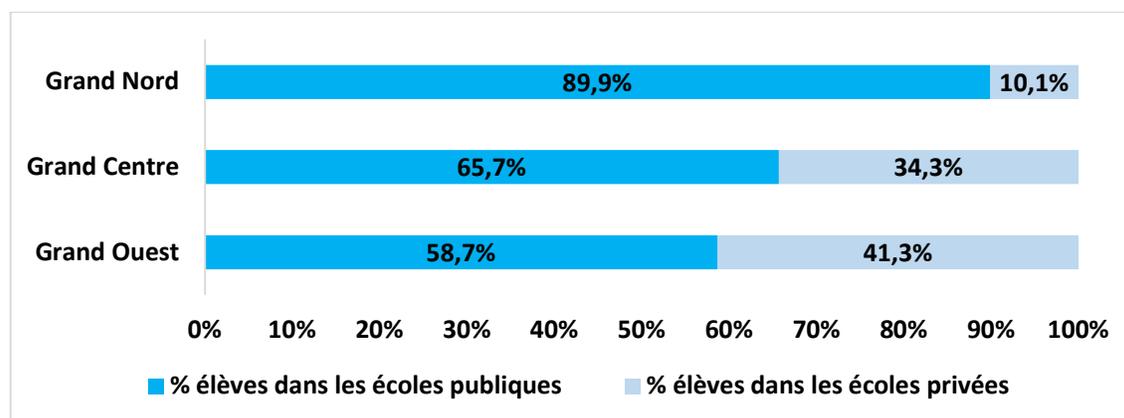


Figure 23 : Répartition des élèves en fonction du statut de l'école dans les strates

Les salles de classe qui accueillent les élèves dans les différentes strates sont diversement équipées en matière de ressources pédagogiques ($p < 0,01$ et $r = 0,6$). L'indice moyen de ressources pédagogiques de la classe dans la strate Grand Ouest (50,1) est largement au-dessus de ceux de la strate Grand Centre (45,6) et de la strate du Grand Nord (38,3). La figure 24 présente la distribution de l'indice de ressources pédagogiques des classes dans les différentes strates.

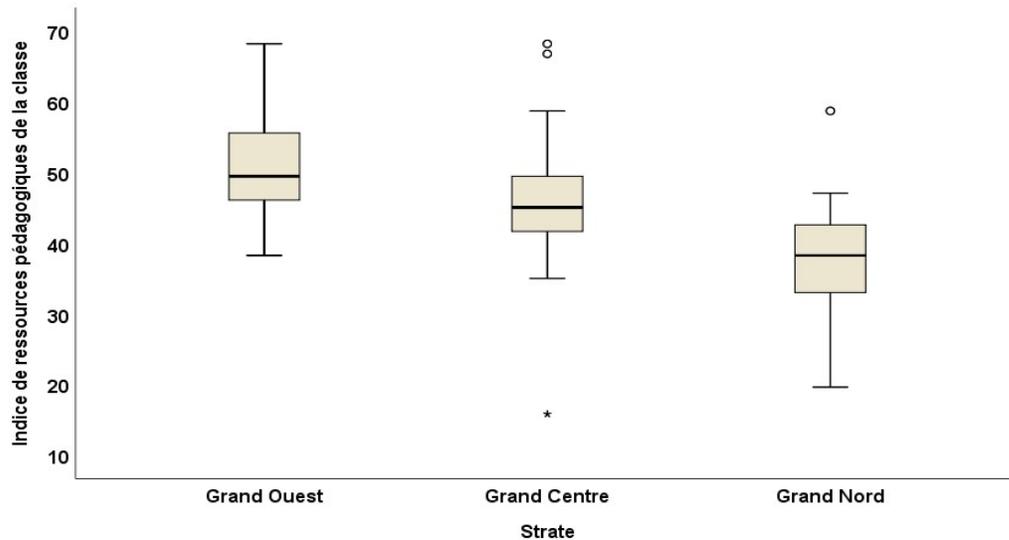


Figure 24 : Distribution de l'indice de ressources pédagogiques des classes dans les strates

Les enseignants qui encadrent les élèves dans les salles de classe sont en majorité des hommes (Figure 25) malgré quelques différences entre les strates ($p < 0,01$ et V de Cramer = 0,3). Ils représentent ainsi 68,0% de l'effectif dans la strate Grand Ouest et 78,7% dans la strate Grand Centre. Dans la strate du Grand Nord, la totalité de l'effectif (100,0%) est masculine. Tout comme les enseignants, les responsables des établissements scolaires sont à majorité des hommes (Figure 26) avec des variations légères par strates ($p < 0,01$ et V de Cramer = 0,2). Ainsi, les hommes qui officient comme directeurs dans les établissements représentent 72,0% de l'effectif dans la strate Grand Ouest, 82,0% dans la strate Grand Centre et 94,4% dans la strate Grand Nord.

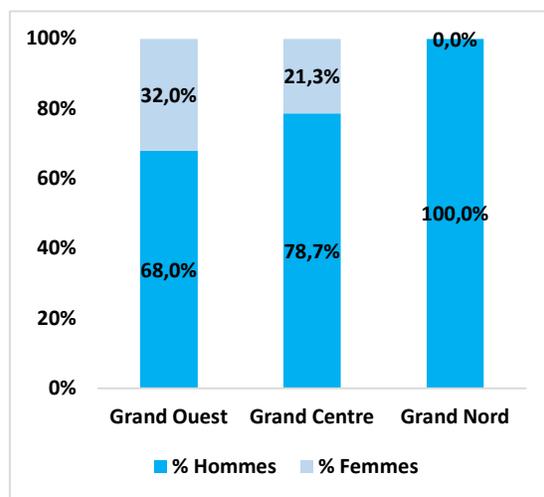


Figure 25 : Répartition des enseignants en fonction de leur sexe dans les strates

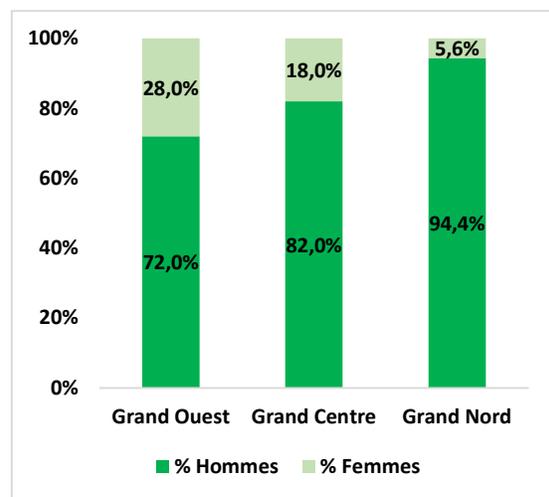


Figure 26 : Répartition des directeurs en fonction de leur sexe dans les strates

La strate du Grand Centre concentre le plus grand taux d’enseignants ayant un niveau universitaire comparativement aux deux autres strates ($p < 0,01$ et V de Cramer=0,3). La proportion d’enseignants qui va au-delà des études secondaires est en effet de 29,5% dans la strate Grand Centre contre 16,0% dans la strate Grand Ouest et seulement 3,7% dans la strate Grand Nord (Figure 27). C’est également la strate du Grand Centre qui possède la plus grande proportion de directeurs ayant atteint un niveau universitaire (Figure 28) comparativement aux deux autres strates ($p < 0,01$ et V de Cramer=0,3). Ils sont ainsi 34,4% à avoir franchir le cap des études secondaires dans la strate Grand Centre contre 18,0% dans la strate Grand Ouest et seulement 5,6% dans la strate Grand Nord.

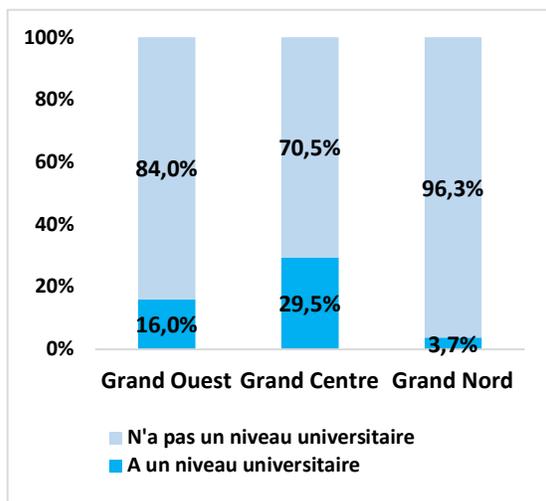


Figure 27 : Répartition des enseignants en fonction du niveau d’études atteint dans les strates

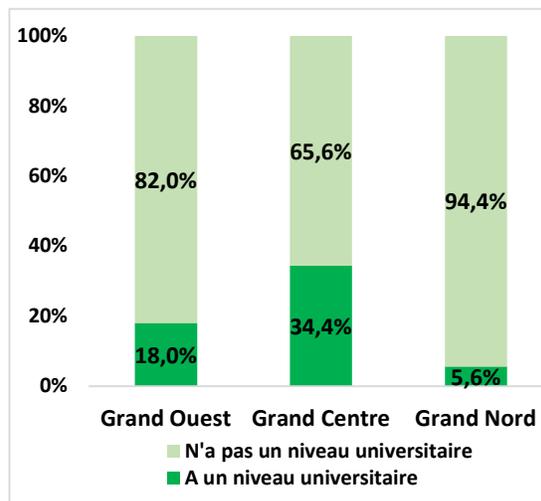


Figure 28 : Répartition des directeurs en fonction du niveau d’études atteint dans les strates

Sur le plan de l’expérience, les enseignants de la strate Grand Ouest cumulent en moyenne un nombre d’années d’expériences légèrement supérieur à ceux des enseignants des deux autres strates ($p < 0,01$ et $r = 0,2$). La moyenne d’années d’exercices des enseignants est de 13,3 années dans la strate Grand Ouest, de 11,5 années dans la strate du Grand Centre et de 8,3 années dans la strate Grand Nord. La figure 29 présente la distribution des années d’expériences des enseignants dans les différentes strates.

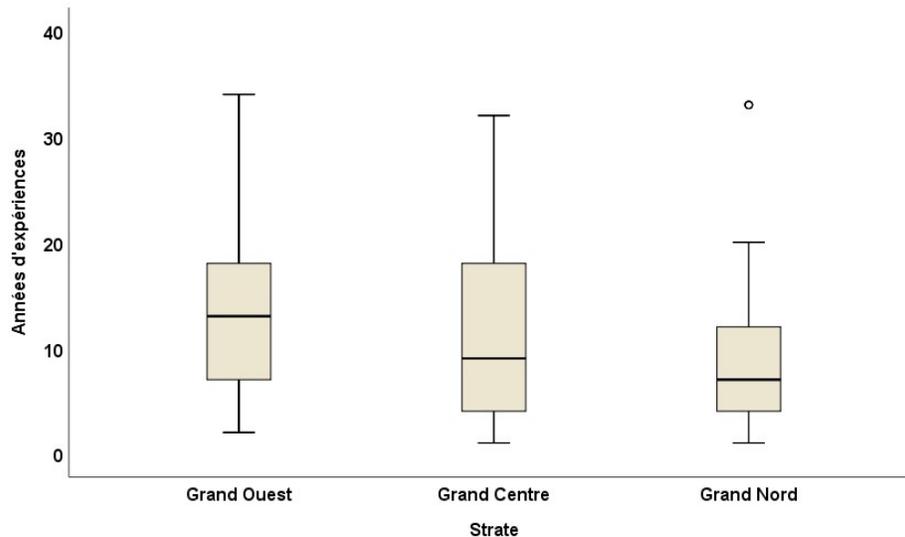


Figure 29 : Distribution des années d’expériences des enseignants dans les strates

4.1.4 Écologie extrascolaire dans les strates (couche 4)

La seule variable en lien avec l’écologie extrascolaire qui nous intéresse dans le cadre de notre travail est la pratique du français à la maison. La pratique de cette langue d’enseignement hors du milieu scolaire varie légèrement en fonction des différentes strates ($p < 0,01$ et V de Cramer = 0,2). Dans la strate Grand Ouest, c’est la quasi-totalité (94,3%) des élèves qui a affirmé « parler toujours ou souvent » le français à la maison. Dans la strate Grand Centre, 82,6% d’élèves ont répondu par « oui » à cette question. Dans la strate Grand Nord, ils représentent 70,1% de l’effectif (Figure 30).

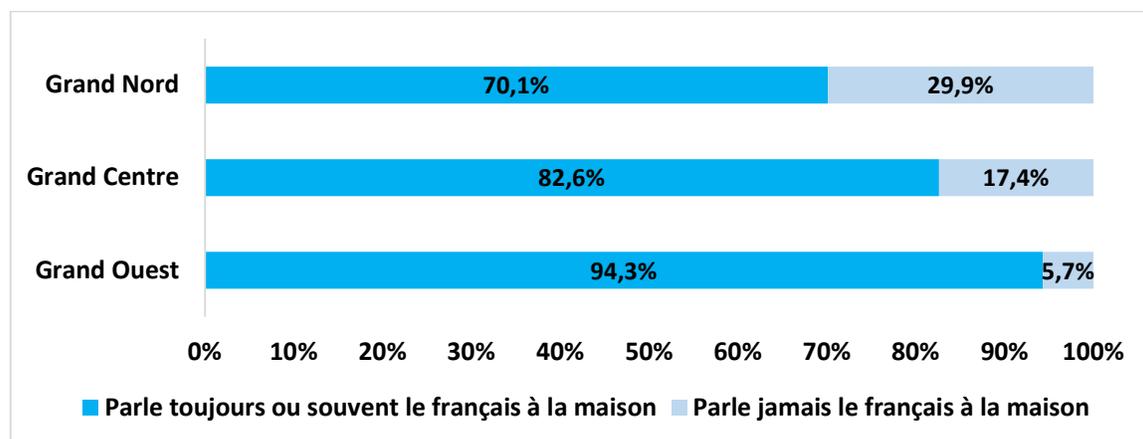


Figure 30 : Répartition des élèves en fonction de la pratique de la langue d’enseignement dans les strates

4.1.5 Caractéristiques de la communauté et voisinage dans les strates (Couche 5)

Les zones où sont situées les écoles de la strate du Grand Ouest affichent en moyenne un indice d'aménagement de la localité de l'école supérieur à ceux des deux autres strates ($p < 0,01$ et $r = 0,4$). Pour rappel, l'indice d'aménagement de la localité de l'école tient compte la présence dans la localité de certains biens et services que sont la route goudronnée, l'électricité, le collège, le lycée, l'hôpital, le centre de soin ou de santé, le poste de gendarmerie ou de police, la banque, la caisse d'épargne, le bureau de poste, le centre culturel ou la bibliothèque (PASEC, 2015). Un indice élevé pour une localité signifierait que cette localité est assez dotée en ces biens et services. En moyenne, cet indice est de 52,3 dans la strate Grand Ouest largement supérieur à la moyenne nationale qui est de 47,4. Par contre, dans les strates du Grand Centre et du Grand Nord, les indices moyens d'aménagement sont respectivement de 46,3 et 44,1. La figure 31 présente la distribution de l'indice d'aménagement de la localité de l'école en fonction des strates.

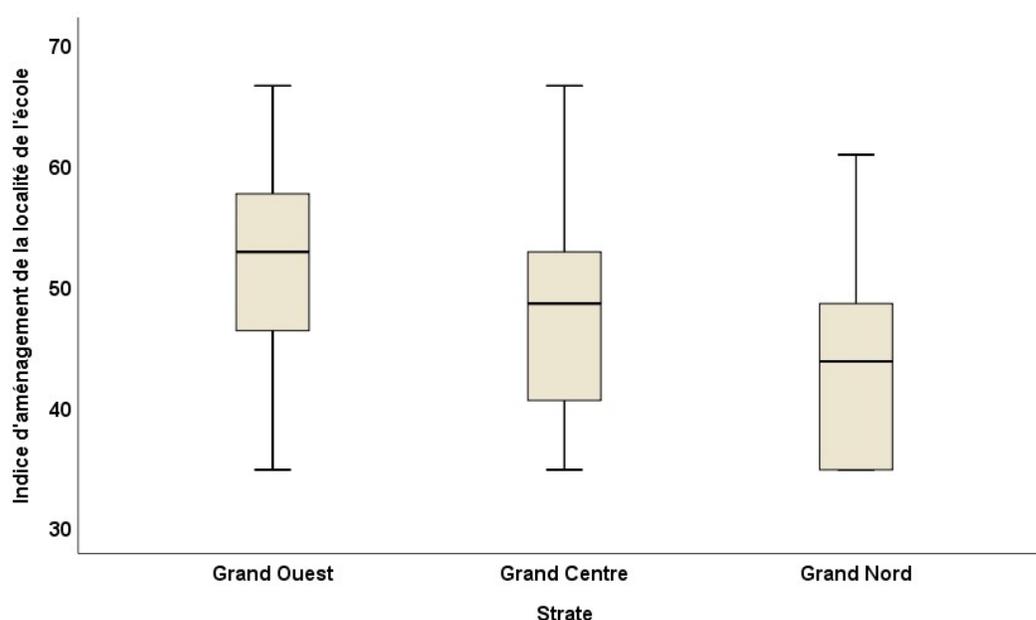


Figure 31 : Distribution de l'indice d'aménagement de la localité de l'école en fonction des strates

Aucune différence statistiquement significative ne se dégage cependant au niveau de l'indice d'implication de la communauté dans les écoles pour les différentes strates. Cet indice est en moyenne de 47,8 pour les écoles de la strate du Grand Ouest, de 44,5 pour celles de la strate du Grand Centre et 47,8 pour celles de la strate du Grand Nord. Comme nous l'avons mentionné dans les sections précédentes, l'indice d'implication de la communauté synthétise 22 autres variables contextuelles en lien avec la présence ou non au sein de l'école de diverses

associations de la collectivité locale¹⁵, les différentes collaborations qu’elles entretiennent avec le milieu scolaire et les divers appuis qu’elles fournissent à ces écoles (PASEC, 2017b). La figure 32 présente la distribution d’implication de la communauté en fonction des strates.

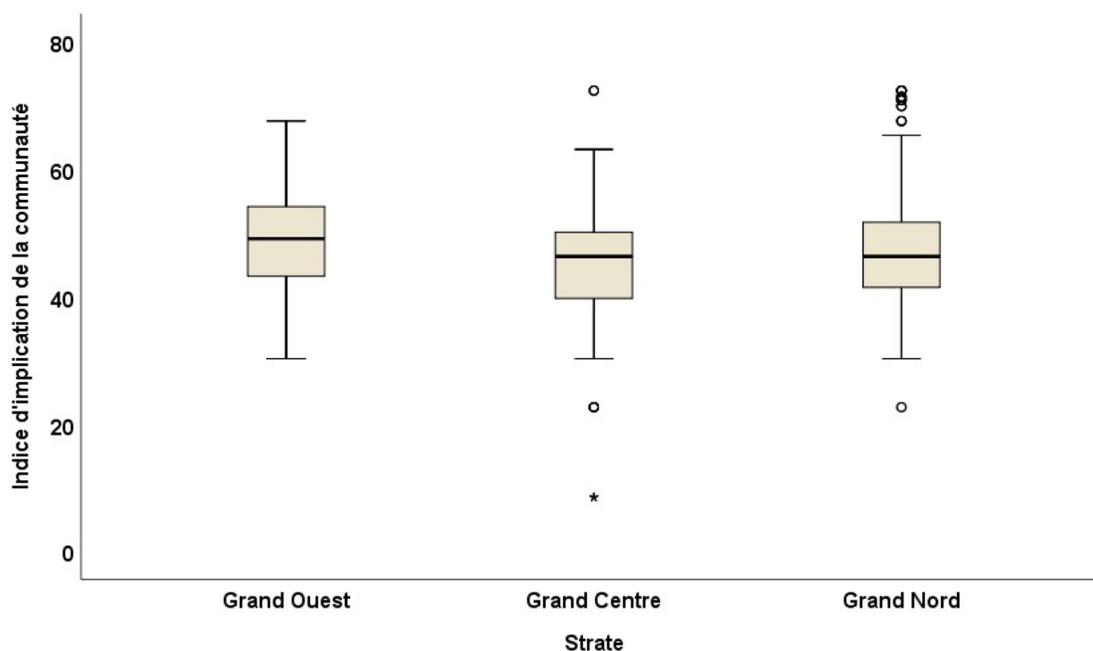


Figure 32 : Distribution de l’indice d’implication de la communauté en fonction des strates

Nous venons de présenter des statistiques descriptives des variables contextuelles retenues pour nos analyses. Ces premiers résultats mettent en exergue des disparités contextuelles plus ou moins marquées entre les différentes strates et ce, au niveau de la plupart de nos variables d’intérêts. La section suivante présente, de façon descriptive également, les résultats aux différents items en lecture. Dans cette partie, nous présenterons également les résultats de façon globale puis en fonction des strates.

4.2 Taux de réussites aux différents items

4.2.1 Taux de réussites globaux

Pour rappel, le test de lecture que nous analysons comporte 23 items libellés de f1 à f23. De même, comme nous l’avons souligné dans le cadre méthodologique de notre travail, nous n’avons pas accès aux contenus de ces items. Le contenu des items PASEC est en effet confidentiel ; seuls les choix de réponses des élèves sont accessibles dans la base de données. Comme le montre la figure 33, les items f5, f10, f11, f18 et f20 affichent des taux de réussite de moins de 30%. Ces 5 items représentent 21,7% du nombre total des items (Figure 34). Les

¹⁵ APEE, conseil d’école, comité de gestion ou autre association locale de la société civile

8 items f4, f6, f7, f9, f12, f15, f19 et f21 ont des taux réussites qui varient entre 30% et 60% (Figure 33). Ces 8 items ont un poids de 34,8% dans l'ensemble des 23 items du test (Figure 34). Près de la moitié (43,5%) de ces items, soit 10 items, présentent néanmoins des taux de réussite supérieurs ou égaux à 60% (Figure 34). Ce sont notamment les items f1, f2, f3, f8, f13, f14, f16, f17, f22, f23 (Figure 33).

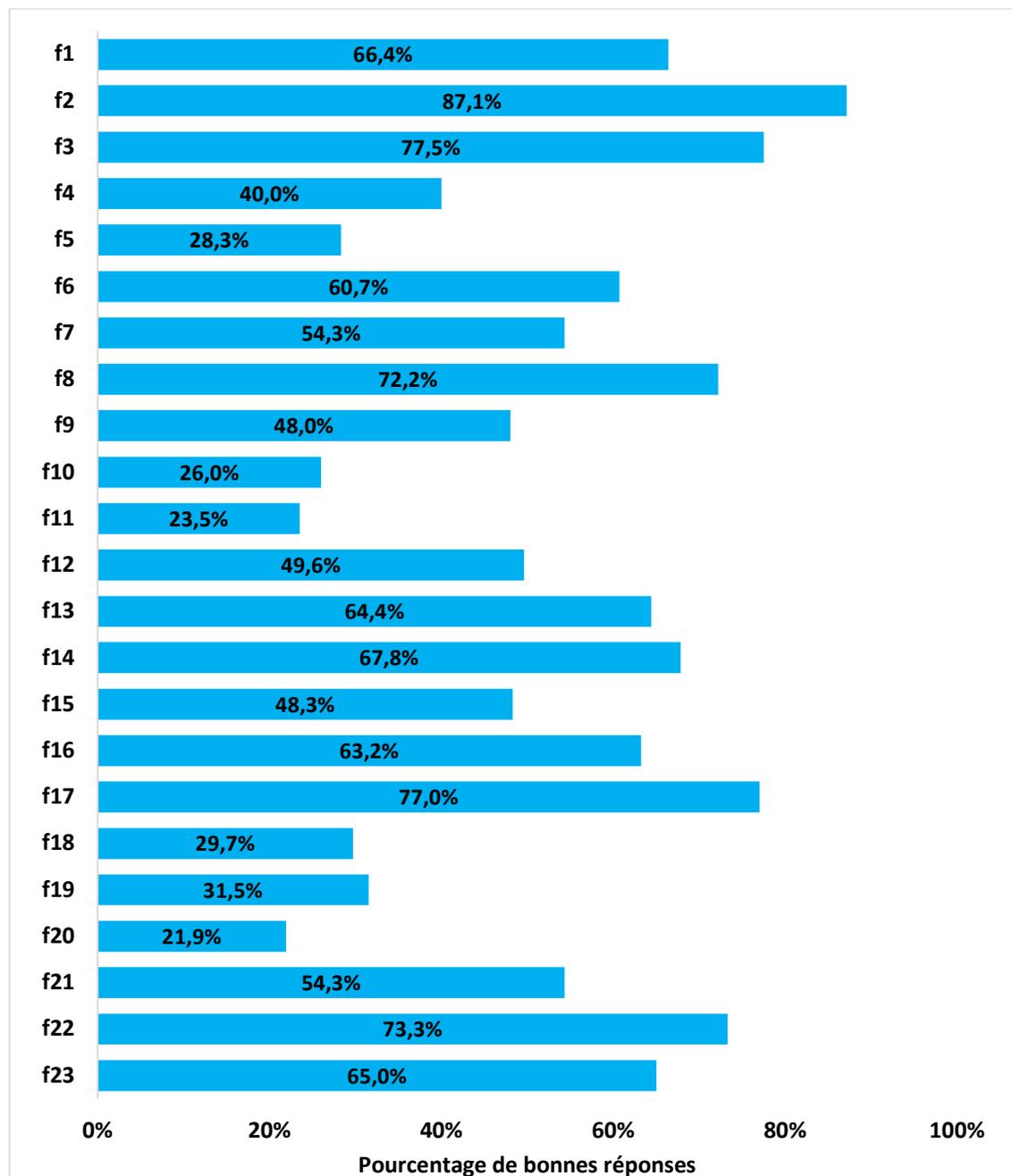


Figure 33 : Pourcentages de bonnes réponses aux items pour l'ensemble de l'échantillon

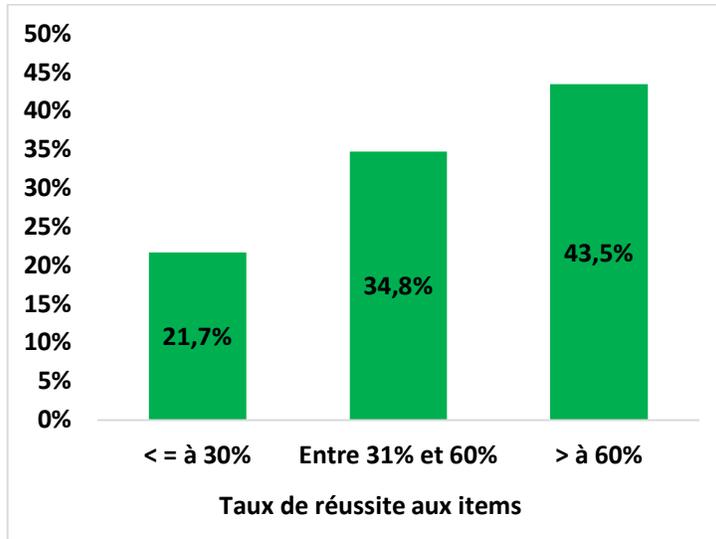


Figure 34 : Pourcentage d'items selon les taux de réussite

4.2.2 Taux de réussites dans les strates

Un examen des taux de réussite par strate fait apparaître des disparités entre celles-ci. Comme le montrent les figures 35 et 36, les valeurs des indicateurs sont assez élevées dans la strate Grand Ouest où plus de la moitié des items (60,9% soit 14 items) affichent un taux de réussite de plus de 60%. Dans cette strate, seuls 2 items (8,7%) ont un taux de réussite de 30% ou moins tandis que 7 items (30,4%) présentent des taux de réussite compris entre 30% et 60%. Dans le Grand Centre, les résultats semblent moyens (Figures 37 et 38). En effet, dans cette strate, 13% d'items (3 items) sont réussis à 30% ou moins, 34,8% d'items (8 items) sont réussis entre 30% et 60% et 52,2% d'items (12 items) sont réussis à plus de 60%. Dans la strate Grand Nord, les résultats semblent quelque peu mitigés (Figures 39 et 40). Dans cette strate, un nombre assez élevé d'items (8 items soit 34,8%) est réussi à 30% ou moins. Plus de la moitié des items (52,2% soit 12 items) a un taux de réussite compris entre 30% et 60% et seuls 3 items (13%) affichent des taux de réussite supérieurs à 60%.

Malgré ces disparités, quelques items se retrouvent dans les mêmes intervalles de réussite de part et d'autre des trois strates. Ainsi, les items f11 et f20 affichent des taux de réussite inférieurs à 30% dans les trois strates. De même l'intervalle 30%-60% de taux de réussite contient les items f9 et f 15 pour les trois strates. Les trois items f2, f17 et f22 ont également des taux de réussite supérieurs à 60% dans les trois strates.

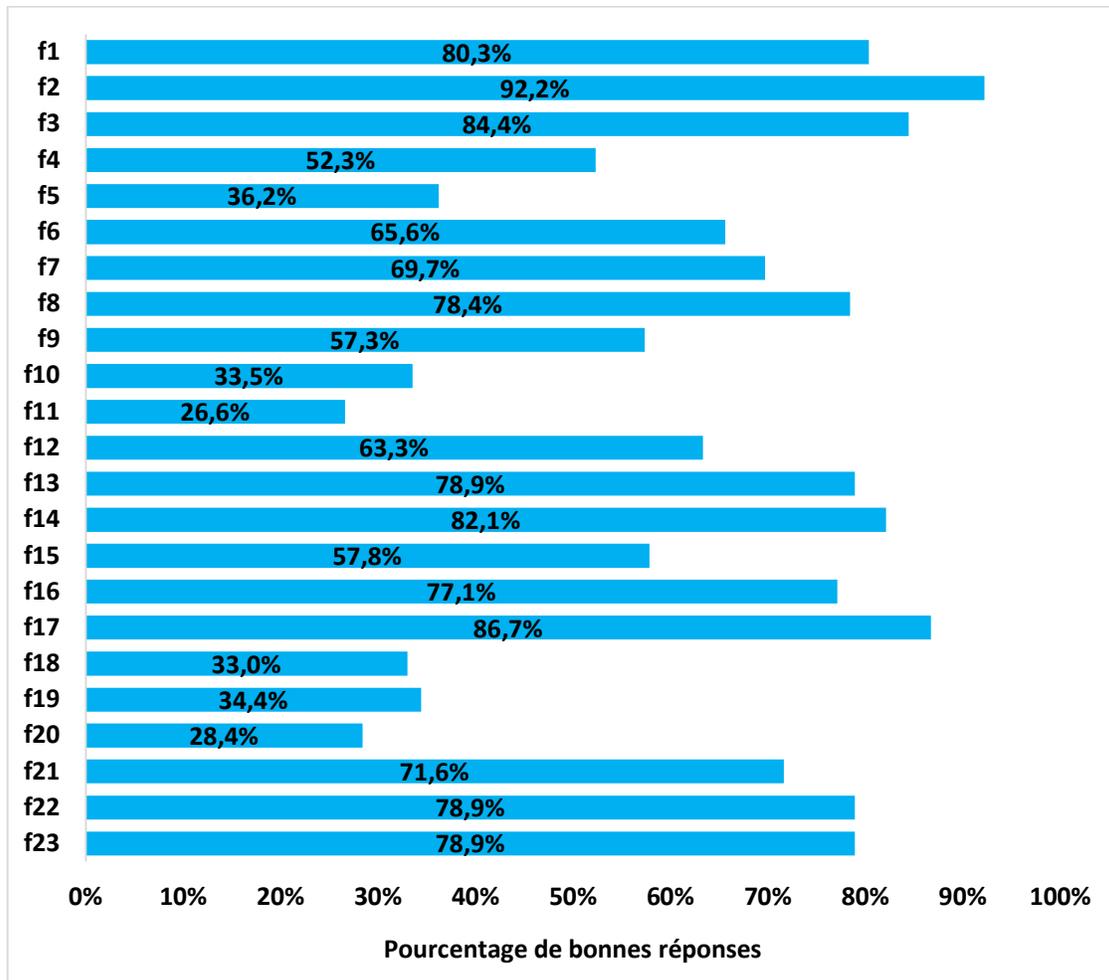


Figure 35 : Pourcentages de bonnes réponses aux items pour la strate Grand Ouest

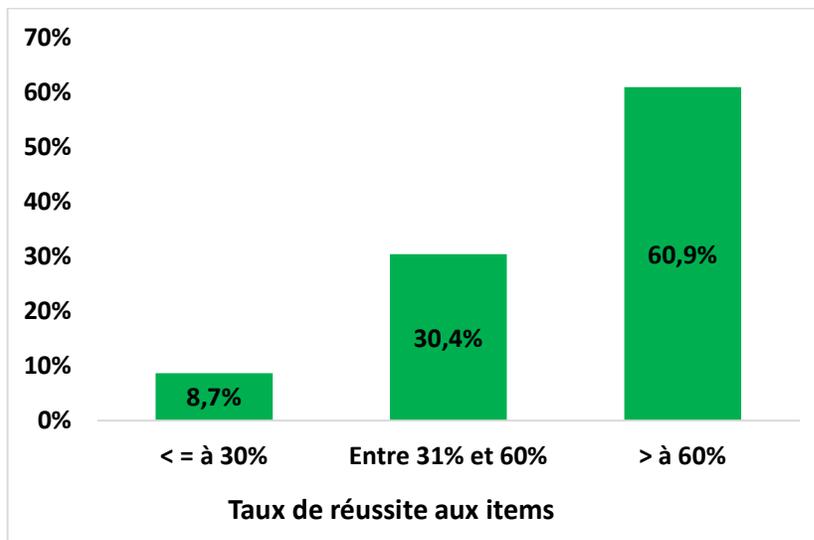


Figure 36 : Pourcentage d'items selon les taux de réussite dans la strate Grand Ouest

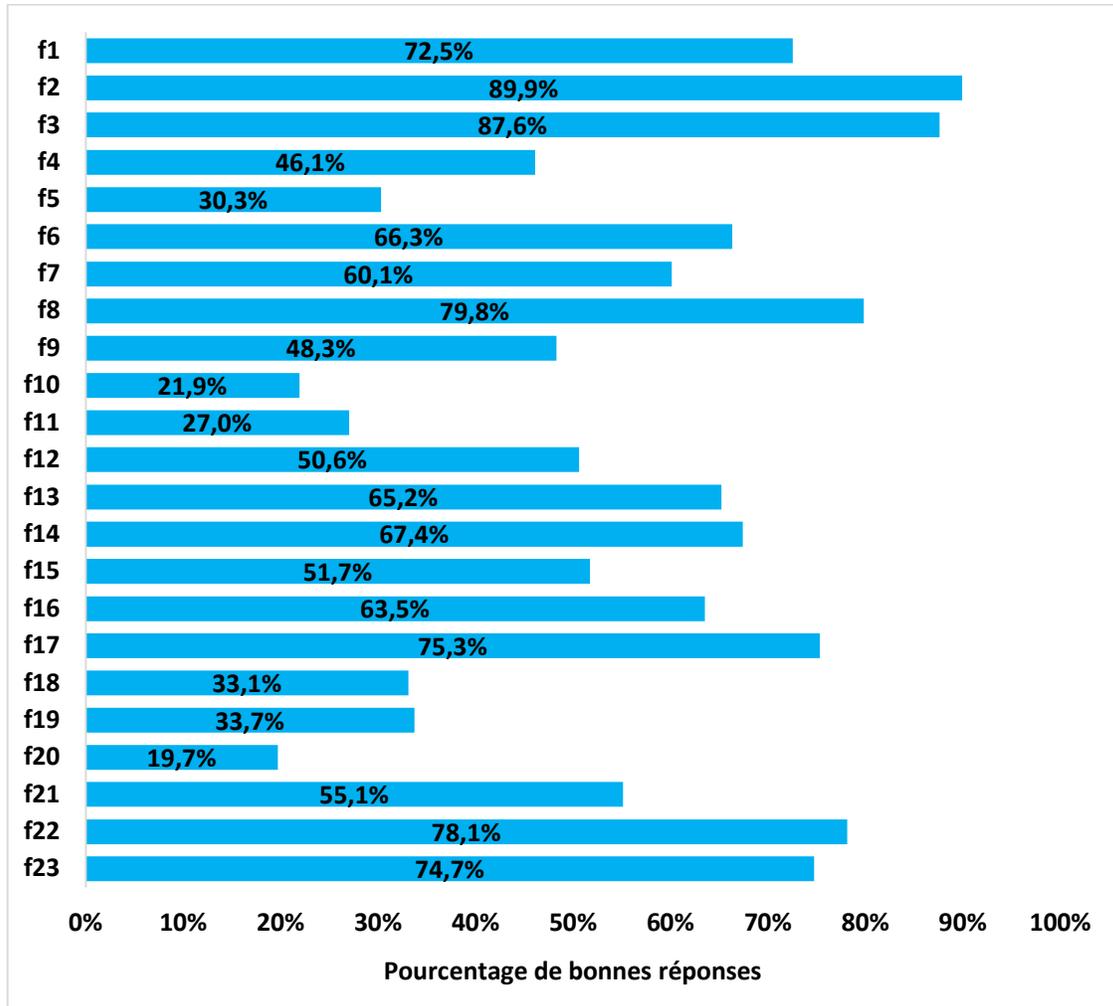


Figure 37 : Pourcentages de bonnes réponses aux items pour la strate Grand Centre

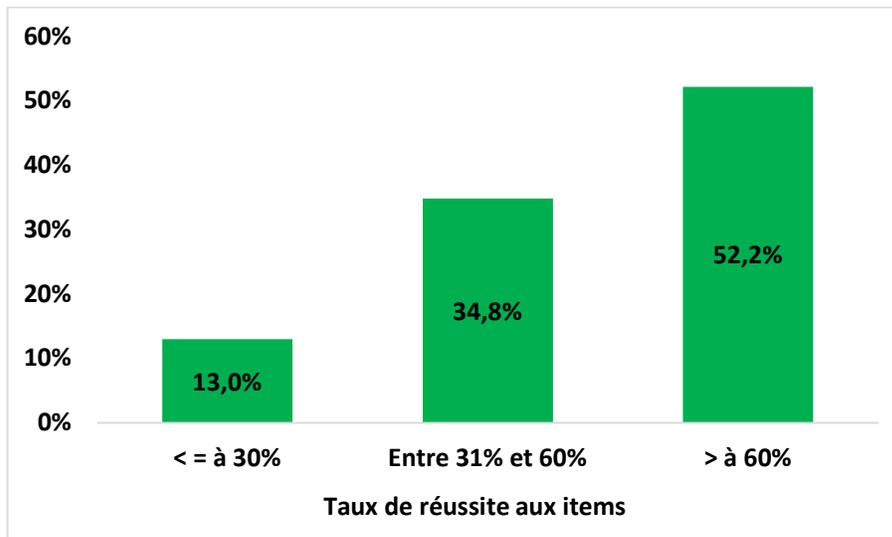


Figure 38 : Pourcentage d'items selon les taux de réussite dans la strate Grand Centre

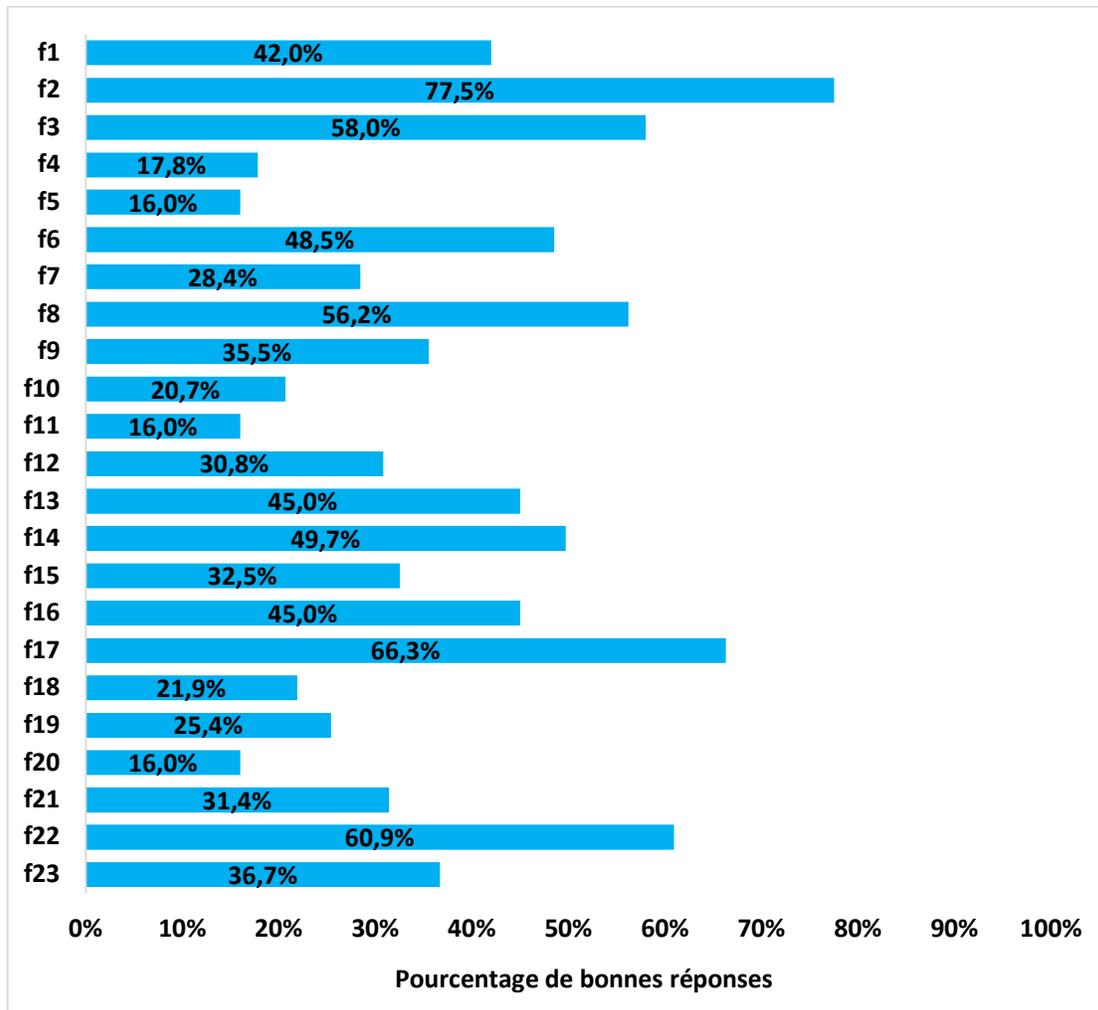


Figure 39 : Pourcentages de bonnes réponses aux items pour la strate Grand Nord

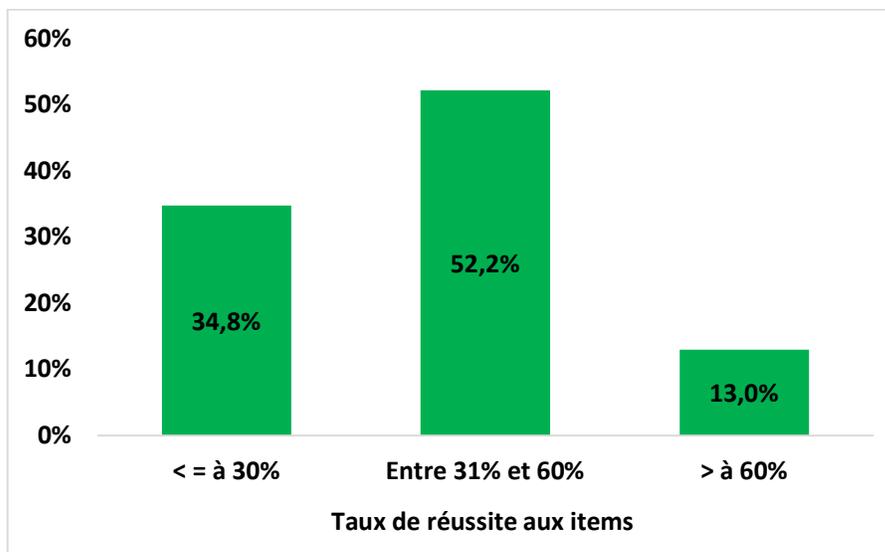


Figure 40 : Pourcentage d'items selon les taux de réussite dans la strate Grand Nord

4.3 Ecologie de performances en lecture

4.3.1 Le caractère situé des performances en lecture : analyse de l'invariance de mesure en fonction des strates

Comme nous venons de le présenter, le PASEC a regroupé les écoles de son échantillon en trois strates qui sont la strate Grand Ouest, Grand Centre et Grand Nord. De même, comme nous venons de le voir, les caractéristiques personnelles des élèves, l'environnement scolaire, l'environnement extrascolaire et les caractéristiques de la communauté dans ces strates diffèrent sur plusieurs points. Lorsque le chercheur mène des analyses ACL dans une situation où un échantillon est composé de sous-groupes plus ou moins hétérogènes, comme c'est le cas pour nos strates, le chercheur doit vérifier s'il doit conduire les analyses ACL sur l'ensemble des données ou le faire sur chacun des sous-groupes en les considérant comme des sous-ensembles de données distinctes (Collins et Lanza, 2009). Cette vérification s'opère à travers une analyse de l'invariance de la mesure ACL sur les données globales (Collins et Lanza, 2009).

Pour rappel, l'analyse de l'invariance de la mesure, telle que décrite par Collins et Lanza, (2009) et présentée dans la partie consacrée au cadre de référence de notre travail, comporte trois phases :

- Le choix du modèle le plus parcimonieux pour les données globales en comparant les différents indicateurs d'ajustements de plusieurs modèles : le modèle retenu s'appelle alors le modèle restreint, car il ne tient pas compte d'une éventuelle variance de la mesure en lien avec les sous-groupes visés ;
- L'estimation des indicateurs d'ajustements pour le modèle non restreint pour les données globales : ce modèle doit avoir le même nombre de classes latentes que le modèle restreint ;
- La comparaison des indicateurs d'ajustement des deux modèles (restreint et non restreint) au moyen d'un test de Khi2 : si le test est non significatif, les analyses ACL peuvent se faire sur les données globales et l'inférence faite sur les données s'appliquent indépendamment des sous-groupes. Si en revanche le test est significatif, les analyses ACL doivent se faire sur chacun des sous-groupes de façon distincte.

Les résultats, comme le montre le tableau 16, suggèrent de retenir un modèle à 4 classes latentes pour le modèle restreint. C'est en effet ce modèle à 4 classes latentes qui minimise au mieux les critères d'information bayésien (BIC), d'information bayésien ajusté (ABIC),

d'information d'Akaike (AIC) et d'information d'Akaike (AIC) et le rapport de vraisemblance (G^2). L'estimation de notre modèle non restreint s'est donc également faite sur la base d'un modèle à 4 classes latentes.

Tableau 17 : Indicateurs d'ajustement des modèles d'ACL pour les données globales

Nbre de classes	LL	DF	G^2	AIC	BIC	CAIC	ABIC	ENTROPIE
2	-7079,87	8388560	7 132,21	7 226,21	7 430,04	7 477,04	7 280,84	0,88
3	-6901,01	8388536	6 774,49	6 916,49	7 224,41	7 295,41	6 999,02	0,83
4*	-6842,21	8388512	6 656,89	6 846,89	7 258,89	7 353,89	6 957,31	0,83
5	-6812,56	8388488	6 597,58	6 835,58	7 351,67	7 470,67	6 973,90	0,85
6	-6751,59	8388464	6 475,65	6 761,65	7 381,81	7 524,81	6 927,86	0,84
7	-6698,74	8388440	6 369,96	6 703,96	7 428,21	7 595,21	6 898,06	0,82

* En gras les indicateurs du modèle retenu

Tous les modèles du tableau 17 sont des modèles restreints qui assument donc que les résultats s'appliquent à tous les sous-groupes d'individus dans l'échantillon. Cependant, comme nous l'avons mentionné plus haut, les premiers résultats que nous venons de présenter à la section 4.1 suggèrent que les groupes d'élèves des trois strates seraient assez hétérogènes et évolueraient dans des contextes assez différents. Ceci étant, nous avons estimé les paramètres d'ajustement d'un modèle à 4 classes latentes non restreint pour comparer ces indicateurs d'ajustements à ceux du modèle à quatre classes latentes du tableau 16. Le tableau 18 contient les indicateurs d'ajustements des deux modèles restreint et non restreint que nous avons comparés.

Tableau 18 : Indicateurs d'ajustement des modèles ACL à 4 classes latentes sur les données globales (modèle restreint et modèle non restreint)

Modèle à 4 classes	LL	DF	G^2	AIC	BIC	CAIC	ABIC	ENTROPIE
Modèle restreint	-6842,21	8388512	6 656,89	6 846,89	7 258,89	7 353,89	6 957,31	0,83
Modèle non restreint	-6769,41	25165722	7 695,93	7 897,93	8 335,95	8 436,95	8 015,33	0,84

Le test de Khi2 que nous avons effectué sur ces indicateurs d'ajustements est significatif ($p < 0,01$) ce qui suggère de rejeter l'hypothèse de l'invariance de la mesure indépendamment des strates. Nous avons donc considéré les résultats des élèves des différentes strates comme des sous-ensembles de données distinctes et mené des analyses ACL avec covariables séparées. Dans les sections suivantes, nous présenterons donc les résultats de ces analyses par strates. Nous déroulons d'abord les différents profils de performances par strates qui résultent des analyses puis, nous affichons les différentes écologies de réponses au test de lecture qui se dégagent.

4.3.2 Profils de performances pour le test de lecture dans les strates.

Les analyses ACL conduites sur les trois strates séparément suggèrent de retenir 3 classes latentes pour la strate Grand Ouest et 2 classes latentes pour chacune des strates Grand Centre et Grand Nord. Les tableaux 19, 20 et 21 contiennent les indicateurs d'ajustement des différents modèles analysés pour retenir les modèles les plus parcimonieux pour les trois strates. Les résultats détaillés de la distribution des indicateurs de prévalences et des probabilités conditionnelles pour chacune des strates sont contenus dans l'annexe 7.

Tableau 19 : Indicateurs d'ajustement des modèles d'ACL pour la strate Grand Ouest

Nbre de classes	LL	DF	G ²	AIC	BIC	CAIC	ABIC	ENTROPIE
2	-2497,39	8388560	2682,64	2776,64	2935,71	2982,71	2786,77	0,90
3*	-2406,67	8388536	2501,19	2643,19	2883,49	2954,49	2658,50	0,88
4	-2374,02	8388512	2435,90	2625,90	2947,43	3042,43	2646,38	0,87
5	-2344,28	8388488	2376,42	2614,42	3017,18	3136,18	2640,08	0,88
6	-2315,84	8388464	2319,54	2605,54	3089,52	3232,52	2636,37	0,91
7	-2293,75	8388440	2275,36	2609,36	3174,58	3341,58	2645,37	0,93

* En gras les indicateurs du modèle retenu

Tableau 20 : Indicateurs d'ajustement des modèles d'ACL pour la strate Grand Centre

Nbre de classes	LL	DF	G ²	AIC	BIC	CAIC	ABIC	ENTROPIE
2*	-2207,02	8388560	2577,65	2671,65	2821,19	2868,19	2672,35	0,89
3	-2166,97	8388536	2497,55	2639,55	2865,45	2936,45	2640,60	0,85
4	-2131,78	8388512	2427,15	2617,15	2919,42	3014,42	2618,57	0,89
5	-2104,45	8388488	2372,50	2610,50	2989,13	3108,13	2612,27	0,89
6	-2079,87	8388464	2323,35	2609,35	3064,34	3207,34	2611,48	0,90
7	-2054,70	8388440	2273,01	2607,01	3138,36	3305,36	2609,49	0,92

* En gras les indicateurs du modèle retenu

Tableau 21 : Indicateur d'ajustement des modèles d'ACL pour la strate Grand Nord

Nbre de classes	LL	DF	G ²	AIC	BIC	CAIC	ABIC	ENTROPIE
2*	-2168,52	8388560	2642,68	2736,68	2883,78	2930,78	2734,97	0,85
3	-2123,63	8388536	2552,91	2694,91	2917,13	2988,13	2692,32	0,89
4	-2086,88	8388512	2479,41	2669,41	2966,75	3061,75	2665,95	0,90
5	-2055,05	8388488	2415,74	2653,74	3026,19	3145,19	2649,40	0,92
6	-2021,94	8388464	2349,52	2635,52	3083,10	3226,10	2630,32	0,91
7	-2000,33	8388440	2306,31	2640,31	3163,00	3330,00	2634,23	0,92

* En gras les indicateurs du modèle retenu

Profils de performances dans la strate Grand Ouest

Nous présentons ici les trois classes latentes qui se dégagent des résultats d'analyse dans la strate Grand Ouest. La première classe latente de la strate du Grand Ouest regroupe 46% d'élèves de cette strate (Figure 41). Cette classe latente est constituée des élèves qui ont une probabilité très élevée (de plus de 0,5 à 1) de trouver la bonne réponse sur la quasi-totalité des items (19 items sur 23) du test de lecture PASEC2014 (Figure 42). Cependant, ils peinent quelque peu, comme les élèves des deux autres classes latentes, sur les items f11, f18, f19 et f20 pour lesquels ils ont des probabilités inférieures à 0,5 de trouver les bonnes réponses (Figure 42). Au regard de leurs prouesses sur la quasi-totalité des items, cette catégorie d'élèves peut être étiquetée de « **très bons lecteurs** »

La deuxième classe latente contient 39% d'élèves de la strate Grand Ouest (Figure 41). Les élèves de cette classe latente présentent des probabilités inférieures à 0,5 de trouver la bonne réponse à 9 items, dont une probabilité quasi nulle (0,09) pour l'item f11 (Figure 42). Sur les 14 autres items où ils ont une probabilité supérieure à 0,5 de réussite, celles-ci sont légèrement inférieures à celles des élèves de la classe 1 excepté aux items f2 et f22 (Figure 42). Nous appellerons ce groupe d'élèves qui ont des probabilités légèrement supérieures à 0,5 de réussir plus de la moitié des items du test de lecture, de « **bons lecteurs** ».

La troisième et dernière classe latente de la strate Grand Ouest regroupe des élèves visiblement en grande difficulté sur l'ensemble du test de lecture (Figure 42). Ces élèves représentent 14% des effectifs de la strate Grand Ouest (Figure 41) et ont des probabilités de réussite aux différents items inférieures à 0,5 excepté à l'item f2 où ils ont une probabilité de 0,53 à réussir à cet item (Figure 42). Nous identifions cette dernière catégorie d'élève, visiblement en difficulté, de « **lecteurs en difficulté** ».

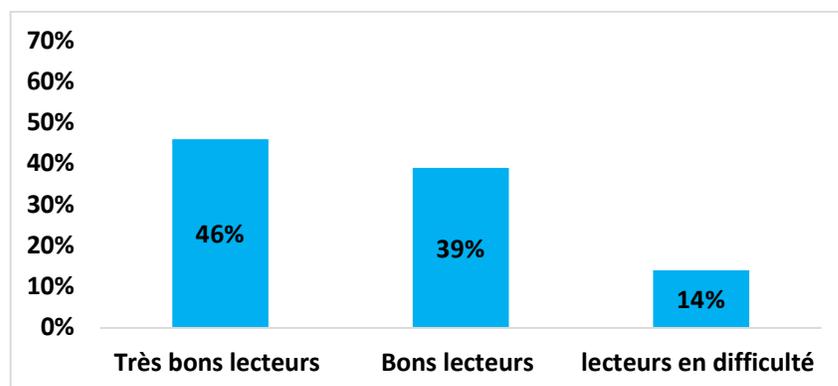


Figure 41 : Prévalences d'appartenance aux classes latentes dans la strate Grand Ouest

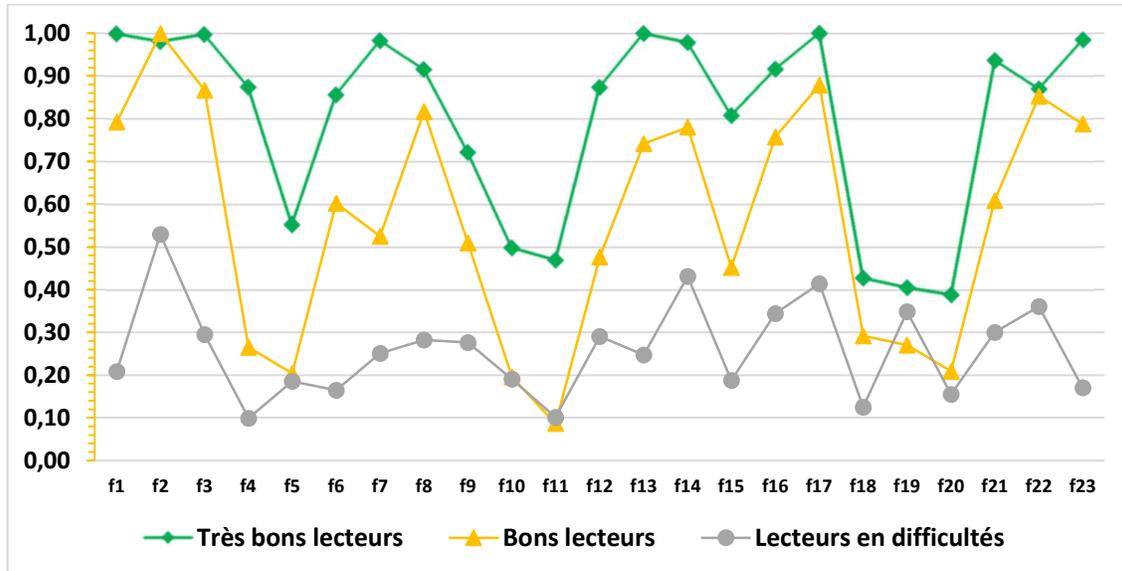


Figure 42 : Probabilités à fournir des réponses justes par item et par classe latente dans la strate Grand Ouest

En reprenant les probabilités de réussite aux 23 items dans un tableau et en regroupant celles supérieures ou égales à 0,5 d'un côté en vert et celles inférieures à 0,5 d'un autre côté en rouge, les résultats permettent d'identifier cinq items communs (f2, f11, f18, f19 et f20) pour les trois profils de performances. Indépendamment des profils de performances « très bons lecteurs », « bons lecteurs » et « lecteurs en difficulté », les élèves ont des probabilités supérieures à 0,5 à réussir l'item f2 tandis qu'ils ont des probabilités inférieures à 0,5 à réussir les items f11, f18, f19 et f20. Nous encadrons ces items communs en bleu dans le tableau 22.

Tableau 22 : Items communs aux profils de performances dans la strate Grand Ouest

Items	Très bons lecteurs	Bons lecteurs	Lecteurs en difficulté
f1	Vert	Vert	Rouge
f2	Vert	Vert	Vert
f3	Vert	Vert	Rouge
f4	Vert	Rouge	Rouge
f5	Vert	Rouge	Rouge
f6	Vert	Vert	Rouge
f7	Vert	Vert	Rouge
f8	Vert	Vert	Rouge
f9	Vert	Vert	Rouge
f10	Vert	Rouge	Rouge
f11	Rouge	Rouge	Rouge
f12	Vert	Rouge	Rouge

Items	Très bons lecteurs	Bons lecteurs	Lecteurs en difficulté
f13	■		■
f14	■		■
f15	■	■	■
f16	■		■
f17	■		■
f18	■		
f19	■		
f20	■		
f21	■		■
f22	■		■
f23	■		■

Profils de performances dans la strate Grand Centre

Deux classes latentes se dégagent des résultats des données de la strate du Grand Centre. La première classe latente regroupe 65% d'élèves (Figure 43) qui ont des probabilités de réussite aux items supérieures à 0,5 sur l'ensemble du test excepté pour 6 items. Ces items où ils affichent des difficultés sont les items f5, f10, f11, f18, f19 et f20 (Figure 44). Cette classe est très similaire à la classe « **bons lecteurs** » de la strate Grand Ouest pour ce qui est des probabilités de réussite aux items. Nous donnons également le nom de « **bons lecteurs** » à cette catégorie d'élèves.

La deuxième classe latente de cette strate contient les 35% d'élèves restants (Figure 43). Les élèves de cette classe latente, à l'instar des « **lecteurs en difficulté** » dans la strate Grand Ouest, ont des probabilités de réussites aux items inférieures à 0,5 sur la quasi-totalité (18 items) du tests (Figure 44). Les cinq items où ils ont des probabilités supérieures à 0,5 de fournir des bonnes réponses sont les items f2, f3, f8, f17 et f22 (Figure 44).

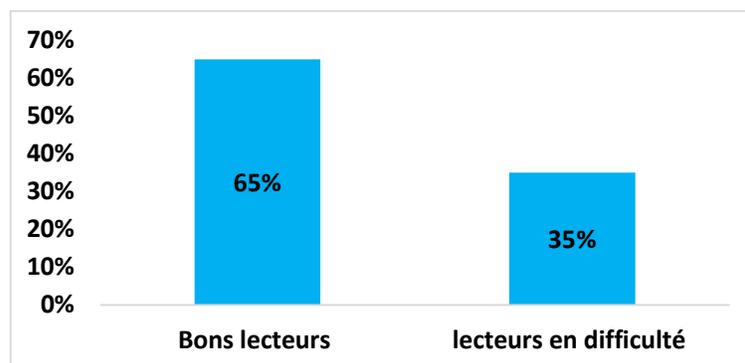


Figure 43 : Prévalences d'appartenance aux classes latentes dans la strate Grand Centre

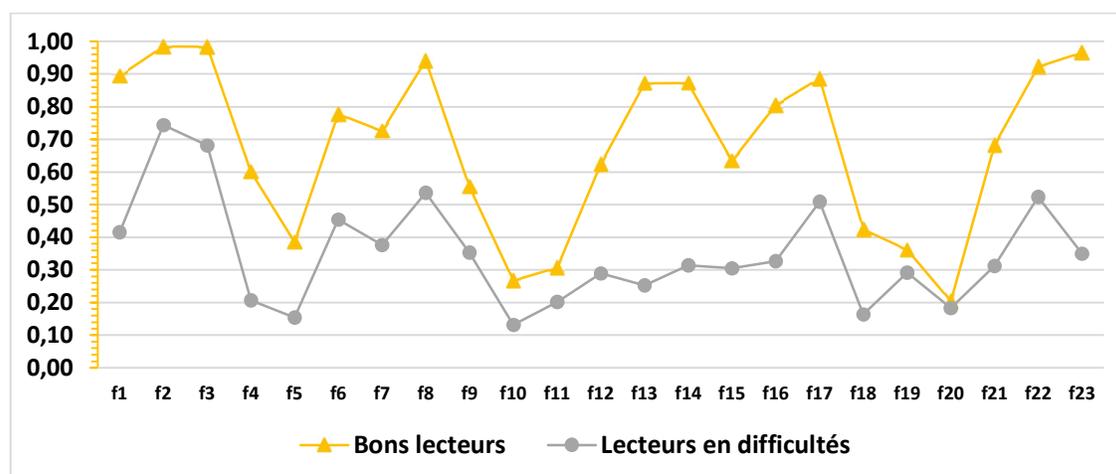


Figure 44 : Probabilités à fournir des réponses justes par item et par classe latente dans la strate Grand Centre

Dans la strate Grand Centre, les items communs aux deux profils de performances sont au nombre de onze (f2, f3, f5, f8, f10, f11, f17, f18, f19, f20 et f22). Les élèves des classes latentes « bons lecteurs » et « lecteurs en difficulté » affichent tous des probabilités supérieures à 0,5 à réussir aux items f2, f3, f8, f17 et f22. Pour les items f5, f10, f11, f18, f19 et f20, ils ont des probabilités inférieures à 0,5 à réussir à ces items. Comme dans le tableau 22, nous reprenons dans le tableau 23, les probabilités de réussite aux 23 items en codifiant celles supérieures ou égales à 0,5 en vert et celles inférieures à 0,5 en rouge et nous encadrons les items communs en bleu.

Tableau 23 : Items communs aux profils de performances dans la strate Grand Centre

Items	Bons lecteurs	Lecteurs en difficulté
f1		
f2		
f3		
f4		
f5		
f6		
f7		
f8		
f9		
f10		
f11		

Items	Bons lecteurs	Lecteurs en difficulté
f12		
f13		
f14		
f15		
f16		
f17		
f18		
f19		
f20		
f21		
f22		
f23		

Profils de performances dans la strate Grand Nord

Les résultats des analyses dans la strate du Grand Nord fournissent également deux classes latentes. La première classe latente se compose de 47% d'élèves de cette strate (Figure 45). Les élèves de cette classe latente ont des probabilités supérieures à 0,5 à réussir la moitié (11 items) du test et des probabilités inférieures à 0,5 à trouver la bonne réponse sur l'autre moitié (12 items) du test (Figure 46). Ce groupe d'élève a des résultats quelque peu inférieurs aux « **bons lecteurs** » des classes latentes Grand Ouest et Grand Centre. Nous donnons donc le nom de « **lecteurs moyens** » à cette classe latente.

La deuxième classe latente regroupe plus de la moitié des effectifs (53%) et présente des caractéristiques très similaires à celles des « **lecteurs en difficulté** » des strates Grand Ouest et Grand Centre (Figure 45). A l'image de ces « **lecteurs en difficulté** », les élèves de cette classe latente éprouvent des difficultés sur la quasi-totalité du test. Ils ont ainsi des probabilités inférieures à 0,5 à fournir la bonne réponse à ces items à l'exception de l'item f2 (Figure 46). Nous désignons également cette classe latente de « **lecteurs en difficulté** ».

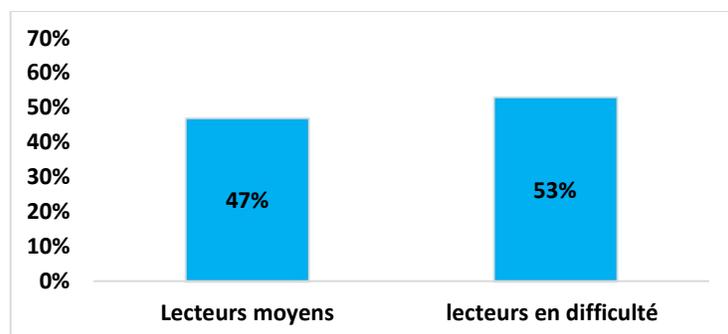


Figure 45 : Prévalences d'appartenance aux classes latentes dans la strate Grand Nord

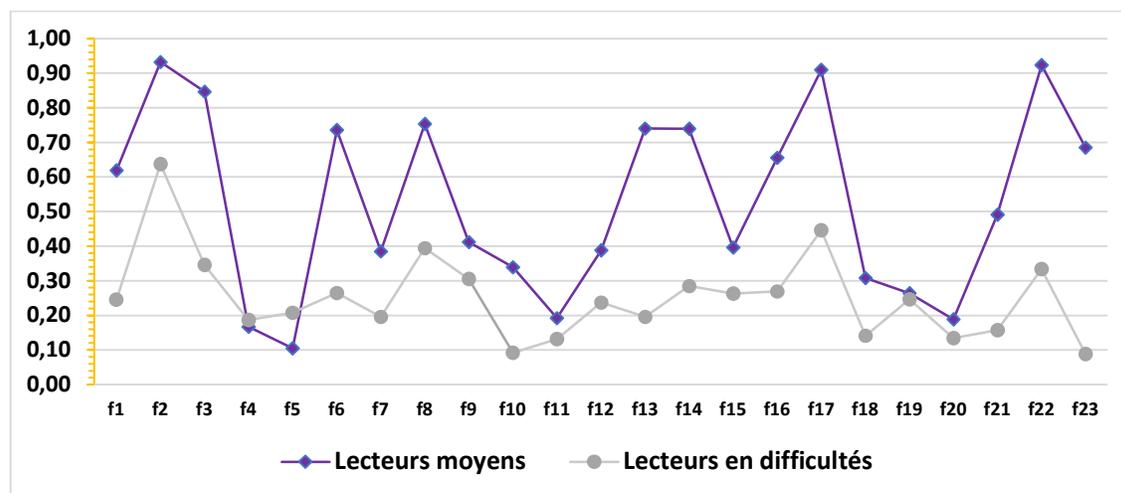


Figure 46 : Probabilités à fournir des réponses justes par item et par classe latente dans la strate Grand Nord

Dans la strate Grand Nord, plus de la moitié (13 sur 23) des items sont communs aux profils de performances «lecteurs moyens» et «lecteurs en difficultés». Les élèves, indépendamment de leur appartenance aux deux profils de performances, ont une probabilité supérieure à 0,5 à réussir à l'items f2. Cependant, ils ont des probabilités inférieures à 0,5 à réussir aux items f4, f5, f7, f9, f10, f11, f12, f15 et f18, f19, f20 et f21. Nous recodons, dans le tableau 24, en vert les probabilités de réussite aux items supérieures à 0,5 et en rouge, celles inférieures à 0,5. Puis, nous encadrons en bleu les items communs aux deux profils de performances.

Tableau 24 : Items communs aux profils de performances dans la strate Grand Nord

Items	Lecteurs en difficulté	Lecteurs en difficulté
f1		
f2		
f3		
f4		
f5		
f6		
f7		
f8		
f9		
f10		
f11		

Items	Lecteurs en difficulté	Lecteurs en difficulté
f12		
f13		
f14		
f15		
f16		
f17		
f18		
f19		
f20		
f21		
f22		
f23		

4.1.1 Éléments actifs dans l'écologie de réponse au test de lecture

Les résultats des analyses ACL avec covariables nous ont permis d'identifier parmi les variables contextuelles retenues pour nos analyses, celles qui contribuent à la formation des différents profils de lecteurs que nous venons de présenter par strate. Les résultats ainsi obtenus permettent de dégager les éléments actifs des différentes écologies de réponse au test de lecture en fonction des strates. Avant de présenter ces résultats, nous notons de nouveau que l'analyse ACL avec covariables est une opération qui nécessite de préciser une classe latente de référence. Ce choix est arbitraire et dépend de la facilité d'interprétation qu'il peut offrir au chercheur (Collins & Lanza, 2010). Dans ce sens, nous avons sélectionné la classe latente « bons lecteurs » comme classe latente de référence pour la strate Grand Ouest. Ce profil nous a semblé être le candidat idéal qui permettrait une interprétation plus aisée des différents coefficients de régression. Dans les strates Grand Centre et Grand Nord qui regroupent chacun deux profils de performances, nous avons retenu la classe latente « lecteurs en difficulté » comme classe latente de référence.

Les tableaux 25, 26 et 27 présentent les différents résultats des analyses ACL avec covariables dans les différentes strates. L'interprétation de l'ordonnée à l'origine (β_0) pour ces résultats reste encore un grand défi pour les chercheurs surtout dans une situation où le modèle intègre plusieurs covariables (Collins & Lanza, 2010). Les coefficients (β) associés aux covariables et les cotes associées s'interprètent entre la classe latente qui contient ces valeurs et la classe latente de référence (Collins & Lanza, 2010). Ces valeurs se lisent alors de la façon

suivante : (i) une valeur positive du coefficient (β) et une cote supérieure à 1 indique une chance plus élevée pour l'individu d'appartenir à cette classe latente comparativement à la classe latente de référence lorsque la covariable prend la valeur indiquée ou a une valeur moyenne (ii) une valeur négative du coefficient (β) et une cote inférieure à 1 indique par contre une chance plus élevée pour l'individu d'appartenir à la classe latente de référence comparativement à cette classe latente lorsque la covariable prend la valeur indiquée ou a une valeur moyenne et (iii) une valeur de la cote à 1 indique une égalité de chances à appartenir aux deux classes latentes indépendamment de la valeur que prend la covariable (Collins & Lanza, 2010).

Dans une logique binaire, nous nous serons intéressés à interpréter les différents coefficients et les cotes obtenus. Cependant, la perspective écologique se détache d'une structure binaire dans laquelle, chaque variable contextuelle aurait de façon isolée un effet sur des performances à un test (Zumbo et al, 2015). Elle invite plutôt le chercheur à s'intéresser aux effets d'un contexte dans son ensemble sur les performances auxquelles il s'intéresse (Zumbo et al, 2015). Ainsi, nous ne nous préoccupons pas de regarder les liens de façon isolée entre les covariables et les différents profils de performances. Ainsi, nous recensons uniquement les différentes covariables statistiquement significatives dans la formation de ces profils de performances et nous les considérons comme des éléments actifs au sein des différentes écologies.

➤ *Strate Grand Ouest*

Dans la strate Grand Ouest, sur les seize (16) covariables retenues pour les analyses, dix (10) se sont avérées statistiquement significatives dans la formation des différents profils de performances. Il s'agit du statut socioéconomique, de l'âge de l'élève, du redoublement, de la fréquentation du préscolaire, du sexe du directeur, du statut de l'école, de la zone où est située l'école, de l'indice de ressources pédagogiques de la classe, de la langue française parlée à la maison et de l'indice d'aménagement du territoire. Le tableau 25 présente les différents éléments actifs dans l'écologie de la strate Grand Ouest.

Tableau 25 : Covariables participant à la formation des classes latentes dans la strate Grand Ouest

Couches	Covariables actives		Très bons lecteurs	Bons lecteurs	Lecteurs en difficulté
Couche 2 : « Caractéristiques personnelles »	Ordonnée à l'origine	β_0	-0,58	REF	-1,85
		Cote	0,56		0,16
	Indice socioéconomique (p<0,01)	β	0,72	REF	-0,40
		Cote	2,05		0,67
	Âge (p<0,01)	β	-1,24	REF	0,43
		Cote	0,29		1,53
	A redoublé (p<0,01)	β	-1,80	REF	0,71
		Cote	0,17		2,03
A fait le préscolaire (p<0,01)	β	1,04	REF	-0,42	
	Cote	2,83		0,66	
Couche 3 : « Environnement scolaire, Enseignant, classe, ... »	Le directeur est un homme (p<0,01)	β	-1,08	REF	0,06
		Cote	0,34		1,06
	L'école a un statut public (p<0,01)	β	-1,02	REF	0,25
		Cote	0,36		1,28
	L'école est située en zone urbaine (p<0,01)	β	1,86	REF	-1,98
		Cote	6,45		0,14
	Indice de ressources pédagogiques de la classe (p=0,03)	β	-0,03	REF	-0,73
		Cote	0,97		0,48
Couche 4 : « Écologie extra-scolaire, Famille, ... »	L'élève parle français à la maison (p<0,01)	β	2,18	REF	-0,89
		Cote	8,81		0,41
Couche 5 : « Caractéristiques de la communauté, voisinage, état, ... »	Indice d'aménagement du territoire (p<0,01)	β	0,53	REF	-1,13
		Cote	1,70		0,32

Strate Grand Centre

Les résultats dans la strate du grand centre révèlent 9 covariables sur 16 qui contribuent de façon significative à la formation des deux profils de performances dans cette strate (Tableau 26). Il s'agit du statut socioéconomique, du redoublement, de la fréquentation du préscolaire, du niveau scolaire du directeur, du statut de l'école, de la zone où est située l'école, de l'indice de ressources pédagogiques de la classe, de la langue française parlée à la maison Indice d'aménagement du territoire. Comparativement aux covariables statistiquement significatives dans la strate du Grand Ouest, l'âge de l'élève et le sexe du directeur ne le sont pas pour les données de la strate du Grand Centre. En revanche, le dernier niveau académique atteint par les responsables des écoles a un effet significatif dans la formation des profils de performances dans cette strate ; ce qui n'était pas le cas pour les résultats dans la strate du Grand Ouest.

Tableau 26 : Covariables participant à la formation des classes latentes dans la strate Grand Centre

Couches	Covariables actives		Bons lecteurs	Lecteurs en difficulté
Couche 2 : « Caractéristiques personnelles »	Ordonnée à l'origine	β_0	0,00	REF
		Cote	1,00	
	Indice socioéconomique (p<0,01)	β	0,77	REF
		Cote	2,15	
	A redoublé (p=0,02)	β	-0,85	REF
		Cote	0,43	
Couche 3 : « Environnement scolaire, Enseignant, classe, ... »	A fait le préscolaire (p<0,01)	β	0,18	REF
		Cote	1,20	
	Le directeur a un niveau universitaire (p<0,01)	β	1,16	REF
		Cote	3,18	
	L'école a un statut public (p<0,01)	β	-1,61	REF
		Cote	0,20	
Couche 4 : « Écologie extra-scolaire, Famille, ... »	L'école est située en zone urbaine (p<0,01)	β	1,92	REF
		Cote	6,80	
	Indice de ressources pédagogiques de la classe (p=0,01)	β	0,43	REF
Couche 5 : « Caractéristiques de la communauté, voisinage, état, ... »	L'élève parle français à la maison (p<0,01)	β	1,58	REF
		Cote	4,84	
	Indice d'aménagement du territoire (p<0,01)	β	0,65	REF
		Cote	1,92	

Strate Grand Nord

Dans la strate Grand Nord nous avons retiré 5 covariables des analyses pour des problèmes liés à la taille des échantillons. C'est le cas des covariables sexe de l'enseignant, niveau de l'enseignant, sexe du directeur, niveau du directeur et zone de l'école pour les analyses des données dans la strate Grand Nord. L'échantillon dans cette strate ne contenait par exemple que des enseignants hommes ou de très faibles proportions de certaines modalités pour les autres covariables. Sur les 11 covariables que nous avons utilisées dans nos analyses seules 4 se sont avérées statistiquement significatives dans la formation des deux profils de performances « lecteurs moyens » et « lecteurs en difficulté » (Tableau 27). Il s'agit du statut socioéconomique, de la fréquentation du préscolaire, de la langue française parlée à la maison Indice d'aménagement du territoire.

Tableau 27 : Covariables participant à la formation des classes latentes dans la strate Grand Nord

	Grand Nord		Lecteurs moyens	Lecteurs en difficulté
	Ordonnée à l'origine	β_0	-0,60	REF
		Cote	0,55	
Couche 2 : « Caractéristiques personnelles »	Indice socioéconomique (p<0,01)	β	0,45	REF
		Cote	1,57	
	A fait le préscolaire (p=0,04)	β	1,29	REF
		Cote	3,62	
Couche 4 : « Écologie extra-scolaire, Famille, ... »	L'élève parle français à la maison (p<0,01)	β	1,27	REF
		Cote	3,55	
Couche 5 : « Caractéristiques de la communauté, voisinage, état, ... »	Indice d'aménagement du territoire (p<0,01)	β	0,45	REF
		Cote	1,57	

CHAPITRE 5 : INTERPRETATION ET DISCUSSION

Dans ce chapitre, nous discutons des résultats en lien avec notre question générale de recherche et les deux questions spécifiques. Pour rappel, la question générale de notre recherche est :

Comment la mesure des performances qui intègre une approche écologique influence-t-elle les résultats du test en lecture du PASEC2014 ?

Estimer les performances des élèves à un test dans le cadre d'une ELE, consiste généralement à déterminer un score ou une échelle sur la base des réponses cognitives que fournissent les candidats dans le but de les classer en deux principales catégories qui refléteraient la compétence et la non-compétence ou la maîtrise et la non-maîtrise des contenus enseignés (Blais, 2008). Comme nous le développons dans cette partie, une analyse écologique sur les résultats d'un test à une ELE nous amène par contre à nous détacher de ces performances standards focalisées uniquement sur des facteurs cognitifs et a pour but de nous intéresser également aux différentes hypothèses contextuelles, sociales et culturelles qui sous-tendent le sens que nous donnons aux performances (McNamara, 2007 ; McNamara et Roever, 2006 ; Zumbo et al, 2015). Un tel regard sur les performances à un test permettrait non seulement une interprétation moins surreprésentée des résultats, mais aiderait également à fournir des pistes pertinentes pour des prises de décisions politiques en matière d'éducation (Zumbo et al, 2015).

Nous déroulons donc progressivement notre argumentaire en répondant à nos deux questions spécifiques de recherche ; la première question étant :

(i) Comment interpréter les résultats du test en lecture PASEC2014-Cameroun lorsqu'on les analyse à partir de l'approche écologique de Zumbo et ses collègues (2015) en utilisant des analyses des classes latentes avec covariables ?

La lentille écologique invite les praticiens, à travers les différentes couches de son modèle, à s'intéresser non seulement aux dimensions cognitives, mais également aux caractéristiques et aux différences individuelles, à l'environnement scolaire et extra-scolaire et aux caractéristiques de la communauté pour lire des performances à un test. De telles performances deviennent inéluctablement des performances liées à des contextes bien définis (Zumbo et al, 2015) à l'exemple de celles qui se dégagent de nos trois différentes strates (Grand Ouest, Grand Centre et Grand Nord).

Notre première question de recherche élucidée, nous nous pencherons ensuite sur la seconde. Pour rappel, la deuxième question est intitulée :

(ii) Comment la compréhension des résultats des tests de lecture PASEC-Cameroun2014 est-elle influencée par l'usage de l'approche écologique de Zumbo et ses collègues (2015) en comparaison avec l'approche utilisée par le PASEC ?

A l'exemple des résultats des autres ELE, les résultats PASEC2014 visent principalement à comparer les performances des candidats ou celles des systèmes éducatifs des pays ou des établissements scolaires (Blais, 2008 ; Loye, 2011 ; Pons, 2011). De telles visées impliquent des prises de décisions holistiques pour tout le système éducatif et donc des échelles de mesure qui s'appliquent à tous les candidats (Pons, 2011). L'approche écologique que nous avons utilisée pour analyser nos données nous révèle que chaque contexte dans nos trois strates, est assez singulier et produit des résultats particuliers. A côté des résultats PASEC2014, nous esquissons donc quelques recommandations et pistes d'actions qui tiennent compte de ces particularités. Les sections suivantes sont consacrées à une discussion plus détaillée de ces questions de recherche.

5.1 Le regard écologique sur les résultats du test de lecture PASEC2014

Nous commençons donc notre discussion en portant un regard sur les différents profils de performances qui se dégagent dans nos trois strates. Ensuite, nous évoquons les caractéristiques des élèves et du milieu scolaire et extrascolaire dans laquelle ils évoluent. Nous terminons cette partie en dressant un portrait des différentes écologies que nous pouvons retenir de nos analyses.

5.1.1 Des marqueurs propres à chaque strate

Dans la strate Grand Ouest, comme nous l'avons vu, les profils de performances sont au nombre de trois : le profil « **très bons lecteurs** », « **bons lecteurs** » et « **lecteurs en difficulté** ». Dans la strate Grand Centre en revanche, seuls deux profils émergent : les « **bons lecteurs** » et les « **lecteurs en difficulté** ». Ces deux classes latentes présentent des caractéristiques très proches de celles de la strate Grand Ouest auxquelles nous avons donné les mêmes noms. Cependant, les résultats des élèves de la strate Grand Centre ne permettent pas de voir émerger un profil de performance qui serait proche de celui de « **très bons lecteurs** ». Dans la strate Grand Nord, deux profils de performances émergent également. Toutefois, contrairement à ceux de la strate Grand Centre, les différents profils sont limités à ceux de

« **lecteurs moyens** » et de « **lecteurs en difficulté** ». Dans cette dernière strate, les élèves semblent limités à produire des résultats qui ne pourraient les placer sur des profils similaires à ceux de « **très bons lecteurs** » et « **bons lecteurs** » des strates Grand Ouest et Grand Centre.

A côté de ces particularités propres à chaque strate, les résultats révèlent également des caractéristiques communes aux différents profils de performances à l'intérieur de chacune des strates. En effet, en regardant de plus près les profils de performances au sein des strates, nous remarquons que les élèves affichent des probabilités similaires à trouver la bonne réponse sur certains items ; cette identité commune aux profils de performances à l'intérieur des strates évoque pour nous l'idée de « **marqueurs** ».

Dans la strate Grand Ouest, les élèves, indépendamment de leur appartenance aux trois profils, ont une probabilité supérieure à 0,5 à réussir l'item **f2**. Ils ont cependant des probabilités inférieures à 0,5 à trouver la bonne réponse aux items **f11, f18, f19** et **f20**. Alors que l'item **f2** semble facile pour tous les élèves de cette strate, les items **f11, f18, f19** et **f20** s'avèrent inaccessibles pour l'ensemble des élèves et ce, quel que soit leur profil de performance. A côté, dans la strate Grand Centre, les élèves ont des chances élevées (probabilités > 0,5) à réussir aux items **f2, f3, f8, f17, et f22** et de faibles chances (probabilités < 0,5) à réussir les items **f5, f10, f11, f18, f19** et **f20** quel que soit leur appartenance aux profils de performance « **bons lecteurs** » et « **lecteurs en difficulté** ». Dans la strate Grand Nord où se trouvent des « **lecteurs moyens** » et des « **lecteurs en difficulté** », les élèves, indépendamment de leur appartenance à un profil donné ont une chance élevée (probabilités > 0,5) à réussir l'item **f2**. Ces élèves peinent (probabilités < 0,5) toutefois à réussir aux items **f4, f5, f7, f9, f10, f11, f12, f15, f18, f19, f20** et **f21** soit, la moitié du test. Dans le tableau 28, nous présentons donc les caractéristiques linguistiques des élèves des différentes strates en lien avec le test de lecture PASEC2014.

Tableau 28 : Marqueurs des élèves de 6e année du primaire au Cameroun selon les strates

Strates	Profils de performances	Items communément réussis	Items communément manqué
Grand Ouest	- Très bons lecteurs		
	- Bons lecteurs	f2	f11, f18, f19 et f20
	- Lecteurs en difficulté		
Grand Centre	- Bons lecteurs	f2, f3, f8, f17, et f22	f5, f10, f11, f18, f19 et f20
	- Lecteurs en difficulté		
Grand Nord	- Lecteurs moyens	f2	f4, f5, f7, f9, f10, f11, f12, f15, f18, f19, f20 et f21
	- Lecteurs en difficulté		

Cette caractérisation rejoint le point de vue de McNamara (2007) pour qui, tout test linguistique serait un test d'identité. Les tests de langue auraient en effet cette particularité de faire ressortir différents traits caractéristiques des individus en matière de pratique de la langue (McNamara, 2007). C'est souvent le cas avec des tests de langue seconde passés aux immigrés qui, bien analysés, permettraient de souligner les caractéristiques communes des différents groupes de participants en lien avec ces tests (McNamara, 2007). Le point de vue de Mislévy (2018) à ce sujet est de portée plus générale. Pour cet auteur en effet, toute évaluation peut être considérée comme une occasion permettant d'apprendre sur des aspects particuliers des individus et sur leurs capacités à agir en termes de pratiques ciblées en lien avec cette évaluation (Mislévy, 2018). Dans la suite de nos discussions nous regardons donc les différents profils des élèves et les milieux scolaire et extrascolaire dans lesquels émergent les différents profils de performances que nous venons d'analyser.

5.1.2 Des élèves et des milieux hétérogènes

Dans le modèle écologique de Zumbo et ses collègues (2015), les facteurs liés à l'individu correspondent aux éléments de la couche 2 de ce modèle tandis que les facteurs liés à son milieu scolaire et extrascolaire correspondent aux éléments des couches 3 à 5. Pour rappel, au niveau des caractéristiques individuelles des élèves, nous avons retenu cinq covariables : le sexe, l'indice socioéconomique, l'âge, le redoublement, et la fréquentation du préscolaire. Le mode de vie et certaines expériences particulières façonnent l'individu et finissent par lui donner son caractère unique (Papalia et al, 2010). A ce titre, en plus du sexe et de l'âge des individus, nous considérons toute valeur avantageuse en lien avec l'indice socioéconomique, le redoublement, et la fréquentation du préscolaire d'un groupe d'individus comme étant des prédispositions de leur part à fournir de bons résultats. Un tel groupe aurait par exemple un âge moyen plus bas, un statut socioéconomique moyen plus élevé, le minimum d'individus ayant redoublés ou encore le maximum d'individus ayant fréquenté le préscolaire comparativement à un autre groupe d'individu auquel il serait comparé.

Les analyses nous révèlent ainsi que les élèves de la strate Grand Ouest auraient de meilleures prédispositions à fournir de bons résultats scolaires comparativement à ceux des deux autres strates. Viennent ensuite ceux de la strate Grand Centre dont les valeurs sur les facteurs individuels sont très proches de la moyenne nationale. Les élèves de la strate Grand Nord de leur côté ont des valeurs sur ces facteurs qui évoquent pratiquement un « handicap » en comparaison avec celles de leurs camarades des deux autres strates.

Ces résultats peuvent nous permettent donc de dresser un portrait « **d'élèves à fortes prédispositions** » dont les caractéristiques seraient meilleures que celles des caractéristiques moyennes des élèves de la strate Grand Ouest. De même, nous pouvons dresser un portrait des « **élèves très vulnérables** » dont les caractéristiques seraient pires que celles des caractéristiques moyennes des élèves de la strate Grand Nord. Le tableau 29 présente un portrait possible de ces deux groupes extrêmes d'élèves.

Tableau 29 : Portraits potentiels des élèves à fortes prédispositions et des élèves très vulnérables

Caractéristiques	Élèves à fortes prédispositions	Élèves très vulnérables
Age (années)	Inf. ou égal à 12	Sup. ou égal à 13
SES	Sup. ou égal à 56,6	Inf. ou égal à 44,4
Probabilité redoublement	Inf. à 0,5	Sup. à 0,6
Probabilité fréquentation préscolaire	Sup. à 0,6	Inf. à 0,07

Inf. : Inférieur(e) ; Sup. : Supérieur(e)

Dans la strate Grand Ouest, un élève sur trois (35%) présente les caractéristiques des « **élèves à fortes prédispositions** » mais ils ne sont que 3% à avoir des caractéristiques similaires à celles des « **élèves très vulnérables** ». Dans la strate Grand Centre, c'est 11% d'élèves qui sont dans la catégorie « **élèves à fortes prédispositions** » contre 9% d'élèves dans la catégorie « **élèves très vulnérables** ». Dans la strate Grand Nord, ils représentent seulement 1%, les élèves pouvant être identifiés dans le groupe « **élèves à fortes prédispositions** ». Cependant, un quart des élèves (24%) peut être classé dans le groupe « **élèves très vulnérables** » dans la strate Grand Nord. Le tableau 30 récapitule la répartition de ces deux catégories d'élèves dans les différentes strates

Tableau 30 : Répartition des élèves à fortes prédispositions et des élèves très vulnérables dans les différentes strates

% dans les strates	Élèves à fortes prédispositions	Élèves très vulnérables
Grand Ouest	35%	3%
Grand Centre	11%	9%
Grand Nord	1%	24%

Ces différents portraits soulignent les disparités individuelles qui existent entre les élèves des différentes strates. Ces disparités sont très prononcées entre la strate Grand Ouest qui regorge « **d'élèves à fortes prédispositions** » et la strate Grand Nord qui compte une forte proportion « **d'élèves très vulnérables** ». Entre les deux, se trouve la strate Grand Centre qui contient des proportions raisonnables (11% et 9%) de ces deux catégories d'élèves.

Au niveau des caractéristiques des milieux scolaires et extrascolaire des élèves, nous notons également des disparités entre les strates. La strate Grand Ouest semble avoir plus de conditions environnementales favorables comparativement aux autres strates. En effet, elle est celle qui accueille plus d'élèves en zone urbaine ($p < 0,01$ et V de Cramer = 0,5) et ses écoles sont les mieux dotées en ressources pédagogiques ($p < 0,01$ et $r = 0,6$). Cette meilleure dotation peut s'expliquer par le fait que c'est dans cette strate que la proportion d'écoles privées est plus élevée ($p < 0,01$ et V de Cramer = 0,2) ; les écoles privées en Afrique subsaharienne étant généralement mieux dotées que les écoles publiques (PASEC, 2015). Au niveau de l'encadrement de ces élèves, les enseignants de la strate Grand Ouest sont plus expérimentés comparativement à leurs collègues des autres strates ($p < 0,01$ et $r = 0,2$). Cependant, c'est dans la strate Grand Centre que se trouvent les plus grandes proportions d'enseignants et de directeurs d'écoles ayant un niveau universitaire ($p < 0,01$ et V de Cramer = 0,3).

Pour ce qui est de l'environnement extrascolaire, le milieu familial des élèves de la strate Grand Ouest est légèrement plus stimulant pour la maîtrise de la langue d'enseignement à travers notamment la pratique plus régulière de la langue dans les foyers ($p < 0,01$ et V de Cramer = 0,2). Les localités de la strate Grand Ouest sont également mieux aménagées que celles des autres strates : l'indice moyen d'aménagement dans cette strate est plus élevé ($p < 0,01$ et $r = 0,4$). Rappelons que cet indice tient compte de la présence dans la localité de certains biens et services que sont les routes goudronnées, l'électricité, les collèges, les lycées, les hôpitaux, les centres de soins ou de santé, les postes de gendarmerie ou de police, les banques, les caisses d'épargne, le bureau de poste, les centres culturels ou les bibliothèques (PASEC, 2015). Aucune communauté ne semble cependant se détacher de celles des autres strates en ce qui concerne son lien avec le milieu scolaire et les divers appuis qu'elle offre aux écoles.

Rappelons également que les covariables contextuelles que nous venons de présenter ont des valeurs supérieures à la moyenne nationale dans la strate Grand Ouest, des valeurs proches de la moyenne nationale dans la strate Grand Centre et des valeurs inférieures à la moyenne nationale dans la strate Grand Nord. Nous pouvons traduire ces situations à travers un environnement scolaire et extrascolaire aux « **conditions favorables** » pour la strate Grand Ouest, aux « **conditions modérées** » pour la strate Grand Centre et aux « **conditions défavorables** » pour la strate Grand Nord. Le tableau 31 récapitule les caractéristiques des milieux scolaires et extrascolaires dans nos différentes strates

Tableau 31 : Caractéristiques des milieux scolaires et extrascolaires dans les différentes strates

Strates	Caractéristiques des milieux scolaire et extrascolaire
Grand Ouest	Conditions favorables
Grand Centre	Conditions modérées
Grand Nord	Conditions défavorables

La strate Grand Ouest qui regroupe une proportion élevée (35%) d'élèves à fortes prédispositions est celle dont les milieux scolaire et extrascolaire présentent des conditions « favorables ». La strate Grand Nord, où 1 élève sur 4 présente des caractéristiques de vulnérabilité, présente des milieux scolaire et extrascolaire dont les conditions sont « défavorables ». La strate Grand Centre dont les milieux scolaire et extrascolaire affichent des conditions « modérées » regroupe des proportions modestes de ces deux catégories d'élèves (respectivement 11% et 9% d'élèves à « fortes prédispositions » et d'élèves « très vulnérables »). Dans la section suivante, nous regroupons tous ces éléments ensemble et regardons les caractéristiques des différentes écologies dans les strates.

5.1.3 Des écologies différenciées dans les strates

Comme nous l'avons mentionné dans les chapitres précédents, la perspective écologique se détache de l'idée d'une structure binaire dans laquelle, chaque variable contextuelle aurait de façon isolée un effet sur des performances à un test (Zumbo et al, 2015). Elle se positionne plutôt pour des performances contextualisées (Zumbo et al, 2015). Dans ce sens, la principale préoccupation du chercheur qui adopte une telle perspective est de regarder l'effet du contexte dans son ensemble sur les performances auxquelles il s'intéresse. En effet, un individu qui évolue dans un environnement stimulant est susceptible de produire de bons résultats même s'il présente un certain handicap (Absil et al, 2012 ; Malo, 2000 ; Papalia et al, 2010). A l'opposé, un autre individu ayant des prédispositions à fournir de bons résultats peut manifester certaines limites s'il évolue dans un environnement pauvre en opportunités (Absil et al, 2012 ; Malo, 2000 ; Papalia et al, 2010).

Dans nos analyses, nous avons utilisé 11 covariables pour caractériser les milieux scolaires et extrascolaires des élèves des différentes strates. Au Grand Ouest et au Grand Centre, 6 de ces covariables participent significativement à la formation des différents profils de performances. Comme nous l'avons souligné dans le chapitre des résultats, nous avons retiré 5 covariables des analyses dans la strate Grand Nord pour des faibles proportions de certaines modalités. C'est le cas par exemple de la covariable « sexe des enseignants » ; l'échantillon

dans la strate Grand Nord ne contenant que des enseignants hommes. Ainsi, sur les 11 covariables analysées dans cette strate, seules 2 se sont avérées statistiquement significatives dans la formation des profils de performances. Le tableau 32 synthétise les résultats obtenus de l'ACL avec covariables.

Tableau 32 : Récapitulatif des résultats obtenus de l'ACL avec covariables

No	Covariables	Grand Ouest	Grand Centre	Grand Nord
1	Sexe de l'enseignant	NS	NS	
2	L'enseignant a un niveau universitaire	NS	NS	
3	Ancienneté de l'enseignant	NS	NS	NS
4	Sexe du directeur	S	NS	
5	Le directeur a un niveau universitaire	NS	S	
6	Statut de l'école	S	S	NS
7	Zone où est située l'école	S	S	
8	Indice de ressources pédagogiques de la classe	S	S	NS
9	Langue française parlée à la maison	S	S	S
10	Indice d'implication de la communauté	NS	NS	NS
11	Indice d'aménagement du territoire	S	S	S

S : Participe significativement à classer les élèves dans un profil de performance.

NS : Ne participe pas à classer les élèves dans un profil de performance.

■ : Non pris en charge dans l'analyse pour des raisons liées à l'échantillon.

Pour les strates Grand Ouest et Grand Centre, 5 des 6 covariables statistiquement significative à la formation des profils de performance sont communes : le statut de l'école, la zone où est située l'école, l'indice de ressources pédagogiques de la classe, la pratique de la langue d'enseignement à la maison et l'indice d'aménagement du territoire. La 6^e covariable qui participe à la formation des profils de performances est le sexe du directeur pour la strate Grand Ouest et le niveau universitaire du directeur pour la strate Grand Centre. Dans la strate Grand Nord cependant, seules les covariables « pratique de la langue d'enseignement à la maison » et « indice d'aménagement du territoire » participent à la formation des profils des performances.

Dans une logique binaire d'analyse de ces effets, nous aurions pu discuter des effets des covariables une à une sur les performances en lecture de nos élèves. Dans une telle logique, nous aurions pu chercher à comprendre pourquoi, dans la strate Grand Ouest, les élèves qui fréquentent une école dirigée par une femme ont plus de chances d'obtenir de meilleures performances en lecture que leurs camarades des écoles dirigées par des hommes. Nous regardons plutôt les caractéristiques de chaque contexte dans son ensemble. Ainsi, dans les strates Grand Ouest et Grand Centre les éléments des deux milieux (scolaire et extrascolaire) concourent à la formation des profils de performances. Nous ne nous attarderons pas dans la

petite nuance qui porte sur le sexe du directeur pour la strate Grand Ouest et le niveau universitaire du directeur pour la strate Grand Centre. Dans la strate Grand Nord par contre, seuls les éléments du milieu extrascolaire participent à la formation des profils de performances.

Nous pouvons donc déduire que nous avons en face de nous deux types d'environnements : un environnement à facteurs « **scolaire et extrascolaire déterminants** » dans les régions Grand Ouest et Grand Centre et un environnement à facteur « **extrascolaire déterminant** » dans la strate Grand Nord. En complétant cette classification avec les analyses descriptives des conditions environnementales dans les strates que nous avons présenté dans la section 5.1.2, nous obtenons donc une représentation complète des trois écologies de nos strates. Ces trois écologies sont donc : une écologie à « conditions **favorables** et à facteurs **scolaire et extrascolaire déterminants** » dans la strate Grand Ouest, une écologie à « conditions **modérées** et à facteurs **scolaire et extrascolaire déterminants** » dans la strate Grand Centre et une écologie à « conditions **défavorables** et à facteur **extrascolaire déterminant** » dans la strate Grand Nord. Le tableau 33 regroupe les trois écologies qui se dégagent de nos analyses.

Tableau 33 : Écologies de performances dans les différentes strates

Strates	Écologies à :
Grand Ouest	- Conditions favorables - Facteurs scolaire et extrascolaire déterminants
Grand Centre	- Conditions modérées - Facteurs scolaire et extrascolaire déterminants
Grand Nord	- Conditions défavorables - Facteur extrascolaire déterminant

Nous pouvons donc à l'aide de ce portrait, de nos profils de performance présentés à la section 5.1.1 et de la description faites des élèves à la section 5.1.2 vérifier l'hypothèse écologique des performances des individus dans des écologies différentes. Pour rappel, cette hypothèse stipule qu'un individu qui évolue dans un environnement stimulant est susceptible de produire de bons résultats même s'il présente un certain handicap (Absil et al, 2012 ; Malo, 2000 ; Papalia et al, 2010). De même, un autre individu ayant des prédispositions à fournir de bons résultats peut manifester certaines limites s'il évolue dans un environnement pauvre en opportunités (Absil et al, 2012 ; Malo, 2000 ; Papalia et al, 2010). Les tableaux 34, 35, et 36 présentent les proportions des « **élèves à fortes prédispositions** » et des « **élèves très vulnérables** » qui font partie des différents profils de performances en fonction des écologies.

Tableau 34 : Pourcentages d'élèves des profils extrêmes à faire partie des profils de performances dans l'écologie à « conditions favorables et facteurs scolaire et extrascolaire déterminants » (Grand Ouest)

Profils	Très bons lecteurs	Bons lecteurs	Lecteurs en difficulté
Élèves à fortes prédispositions	70%	30%	0.0
Élèves très vulnérables	ND	ND	ND

ND : Non pris en charge dans l'analyse pour des raisons liées à l'échantillon (pourcentages négligeables)

Tableau 35 : Pourcentages d'élèves des profils extrêmes à faire partie des profils de performances dans l'écologie à « conditions modérées et facteurs scolaire et extrascolaire déterminants » (Grand Centre)

Profils	Bons lecteurs	Lecteurs en difficulté
Élèves à fortes prédispositions	90%	10%
Élèves très vulnérables	60%	40%

Tableau 36 : Pourcentages d'élèves des profils extrêmes à faire partie des profils de performances dans l'écologie à « conditions défavorables et facteur extrascolaire déterminant » (Grand Nord)

Profils	Lecteurs moyens	Lecteurs en difficulté
Élèves à fortes prédispositions	ND	ND
Élèves très vulnérables	30%	70%

ND : Non pris en charge dans l'analyse pour des raisons liées à l'échantillon (pourcentages négligeables)

Le risque est quasi nul de voir un élève qui présente des caractéristiques similaires à celles des « **élèves à fortes prédispositions** », avoir des difficultés en lecture dans le cas où il évoluerait dans une écologie similaire à celle de la strate Grand Ouest. L'écologie dans cette strate est à « conditions **favorables** et facteurs **scolaire et extrascolaire déterminants** ». Cependant, ce risque n'est pas nul pour un élève qui présente les mêmes caractéristiques mais qui évoluerait dans une écologie dont les conditions seraient moins favorables que dans celle de la strate Grand Ouest. En effet, Un élève sur dix ayant le profil « **d'élèves à fortes prédispositions** » pourrait développer des difficultés en lecture s'il venait à évoluer dans l'écologie de la strate Grand Centre où les conditions propices à la réussite sont modérées.

D'un autre côté, un élève dont le profil est similaire à celui des « **élèves très vulnérables** » et qui fréquenterait la strate Grand Nord dont les conditions propices à la réussite sont **défavorables** et où seul le facteur **extrascolaire** est **déterminant**, aurait 70% de chances d'avoir des difficultés en lecture. Seuls 3 élèves sur dix auraient des chances d'être des « **lecteurs moyens** » ; une classe de lecteurs qui n'est cependant pas fameuse. Dans la strate Grand Centre dont l'écologie dispose de conditions plus favorables, le risque de voir un élève de même profil dans la catégorie de « **lecteurs en difficultés** » est pratiquement réduit de

moitié. Cet élève aurait en effet, une probabilité de 0,4 de faire partie de la classe de « **lecteurs en difficultés** » et une probabilité de 0,6 d'être parmi les « **bons lecteurs** ».

Les très faibles proportions « **d'élèves très vulnérables** » dans la strate Grand Ouest d'une part et « **d'élèves à fortes prédispositions** » dans la strate Grand Nord d'autre part, ne permettent pas de comparer les éléments de ces deux strates. La comparaison entre ces deux strates extrêmes aurait probablement révélé des effets contextuels plus élevés. Nous aurions ainsi pu estimer les chances pour des élèves ayant des caractéristiques « **élèves très vulnérables** » à faire partie des classes latentes « **très bons lecteurs** » et « **bons lecteurs** » dans une écologie à « conditions **favorables** et facteurs **scolaire et extrascolaire déterminants** » (Grand Ouest). De même, nous aurions pu estimer les risques pour des « **élèves à fortes prédispositions** » à faire partie des profils de performances « **lecteurs moyens** » et « **lecteurs en difficultés** » dans une écologie à « conditions **défavorables** et **facteur extrascolaire déterminant** » (Grand Nord).

Rendu à ce niveau de nos discussions, nous disposons donc d'assez d'éléments pour répondre à notre première question spécifique de recherche. Dans ce sens, nous pouvons affirmer que l'approche écologique nous invite à situer les performances des individus à un test dans un contexte social bien défini. Toute interprétation des résultats doit donc se faire à travers un regard sur les caractéristiques individuelles des candidats, du milieu scolaire et, du milieu extrascolaire de leurs possibles effets sur les performances. Comme nous venons de le voir, d'un point de vue écologique, nous ne devons pas regarder les différences de performance des élèves en termes de différences qui seraient liées à un quelconque bagage intellectuel que chacun d'eux posséderait. Nous devons plutôt les voir comme un concours de circonstances lié au parcours scolaire, aux caractéristiques familiales, à un mode de vie, à la configuration du milieu scolaire et extrascolaire. Certains élèves réussissent parce qu'ils ont non seulement de très « fortes prédispositions » à réussir mais également et surtout à cause d'une écologie dont les conditions sont favorables et dont beaucoup de facteurs sont déterminants pour la réussite. D'autres par contre présentent des difficultés à cause d'une part, de leur « vulnérabilité » mais aussi et surtout à cause d'une écologie dont les conditions sont plus ou moins défavorables et dont peu de facteurs sont déterminants pour la réussite.

Dans les sections suivantes, nous discutons des implications d'une telle analyse sur les données d'un résultat à un test standardisé. La première implication est avant tout conceptuelle. Faire le choix d'une approche écologique pour analyser nos données implique en effet un regard différent sur la performance. Un regard qui se détache de l'idée qui résumerait des performances

à un ensemble de scores ou d'échelles générales applicables à l'ensemble des candidats. Nous terminerons ensuite nos discussions en suggérant quelques pistes de réflexions pour améliorer la performance en lecture des élèves de 6^e année du primaire au Cameroun.

5.2 Apports de l'utilisation de l'approche écologique

Comme annoncé, cette partie de nos discussions tourne autour de notre deuxième question spécifique de recherche :

Comment la compréhension des résultats des tests de lecture PASEC-Cameroun2014 est-elle influencée par l'usage de l'approche écologique de Zumbo et ses collègues (2015) en comparaison avec l'approche utilisée par le PASEC ?

Les évaluations en générale et les tests de langue en particulier peuvent être déterminants en matière d'opportunités, d'accès, de privilèges et de discrimination (Mislevy, 2018). Ces implications nous invitent à faire un choix judicieux du prisme à travers lequel nous décidons d'analyser les données de ces évaluations. En effet, chaque modèle de mesure que nous utilisons pour nous faire une représentation d'un concept donné à l'instar des performances à un test, raconte une histoire particulière (Magnani et al, 2017).

En optant pour une approche écologique, matérialisée par une analyse ACL avec covariables, nous faisons donc le choix de donner un aperçu différent des résultats du test de langue PASEC2014. Ce faisant, nous adoptons une vision locale et contextualisée de la performance dans laquelle nous ne pouvons démêler les performances des individus de leurs caractéristiques individuelles et de l'environnement scolaire et extrascolaire dans lequel ils évoluent.

5.2.1 D'une performance uniforme vers une performance contextualisée

Les évaluations standardisées évoquent généralement, et à raison, l'idée de scores et d'échelles qui font office de performance ou de catégories de performances qui s'appliqueraient à tous les candidats (Blais, 2008 ; Loye, 2011 ; Pons, 2011). Ces échelles et scores permettent ainsi de comparer des individus, des établissements scolaires ou plus largement des systèmes éducatifs (Loye, 2011 ; Pons, 2011). Les figures 47 et 48 montrent par exemple les résultats du test de lecture tel que présenté dans les rapports PASEC 2014.

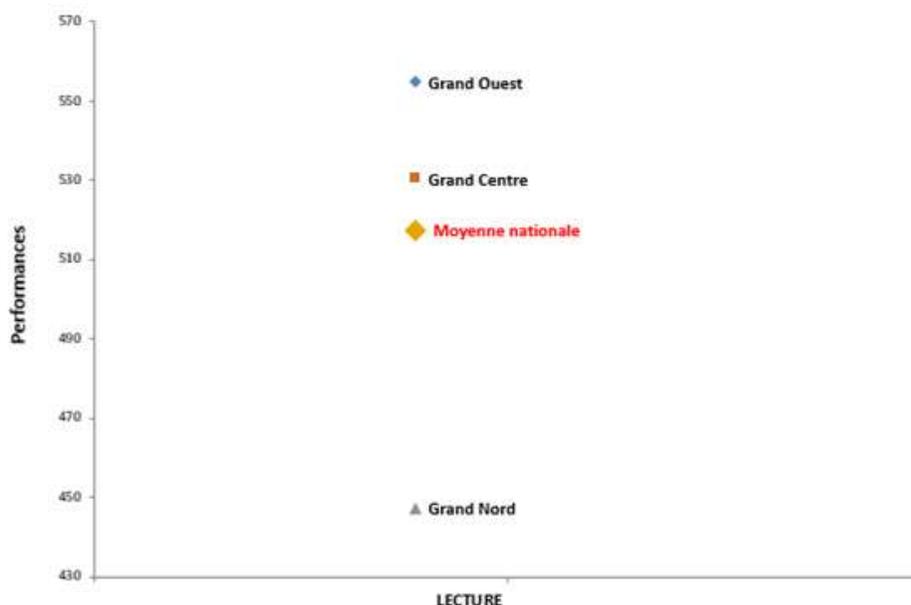


Figure 47 : Performances des strates par rapport à la moyenne des scores nationaux en lecture – Fin de scolarité, Sous système francophone (PASEC, 2016, p. 53)



Figure 48 : Pourcentage d'élèves de 6e année primaire par niveau de compétences en lecture, sous-système francophone (PASEC, 2016, p.52) ¹⁶

Dans la figure 47, les strates sont comparées à l'aide de leurs scores moyens et du score moyen national. Dans la figure 48 cependant, sur une échelle de niveau 4, les strates sont comparées à l'aide de pourcentages de leurs élèves qui se retrouvent dans chaque niveau de

¹⁶ La barre verticale en rouge représente le seuil « suffisant » de compétence. Les niveaux de compétences sont décrits dans l'annexe 5.

compétence ou de part et d'autre du niveau « suffisant » de compétence représenté par la barre verticale en rouge. Que ce soit dans la figure 47 ou dans la figure 48, les scores et les échelles ont la même signification dans les trois strates. Cette visée comparative impose ainsi une représentation uniforme de la performance qui occulte le contexte dans lequel se réalise ces performances (Pons, 2011). L'approche écologique par contre nous invite à porter un regard local sur le concept de performance.

En effet, chaque communauté présente des configurations identifiables et récurrentes de thèmes, de structures, d'activités, des façons de penser et de faire qui prennent le nom de pratiques (Mislevy, 2018). Ces pratiques sont généralement façonnées par des régularités renforcées dans l'usage quotidien, ou dans des pratiques institutionnelles ou encore éducatives. Il est donc courant de voir des individus d'une même aire géographique partager des types de jeux, des rituels ou des modèles sociaux communs, une même conception de la notion de maladie ou du mariage, ou encore partager des granularités langagières (Mislevy, 2018).

Ainsi, en fonction des lieux et des situations, la pratique d'une langue peut prendre des colorations spécifiques qui peuvent se décliner en un accent particulier, un usage singulier de la grammaire ou des constructions particulières de phrases (Mislevy, 2018). Les différents profils de performances qui se sont dégagés de nos analyses nous révèlent des performances en lecture propres à chacune des strates. De même, chaque strate accueille des profils d'élèves particuliers et présente des caractéristiques contextuelles particulières. Tout se passe en effet comme si le fait pour un élève d'être dans une strate donnée le prédispose à faire partie d'un ensemble de profils de performance en lecture donné.

Nous présentons donc les performances en lecture des élèves dans leurs écologies respectives à l'aide du modèle écologique de Zumbo et ses collègues (2015) en retenant uniquement pour chaque strate, les covariables qui participent à la formation des profils de performances. Dans la couche 1, nous rappelons les sous-domaines, les processus cognitifs et le nombre d'items qui constituent le test de lecture PASEC2014. La couche 2 constituée des éléments sur les caractéristiques personnelles et les différences individuelles contient donc : (i) l'indice socioéconomique, (ii) l'âge, (iii) le redoublement et (iv) la fréquentation du préscolaire. La couche 3 qui renferme les éléments sur l'environnement scolaire, et de la classe contient : (i) le sexe du directeur ; (ii) le statut de l'école ; (iii) la zone où est située l'école et (iv) l'indice de ressources pédagogiques de la classe. La couche 4 qui regroupe les éléments sur l'écologie extra-scolaire et la famille contient la covariable sur la langue française parlée à la maison tandis que la couche 5 qui porte sur les caractéristiques de la communauté, du voisinage, de l'état

contient l'indice d'aménagement de la localité. Dans les figures 49, 50 et 51, présentent les différentes écologies de réponse au test de lecture dans nos trois strates.

↑ Covariables contextuelles dont les valeurs sont supérieures à la moyenne nationale

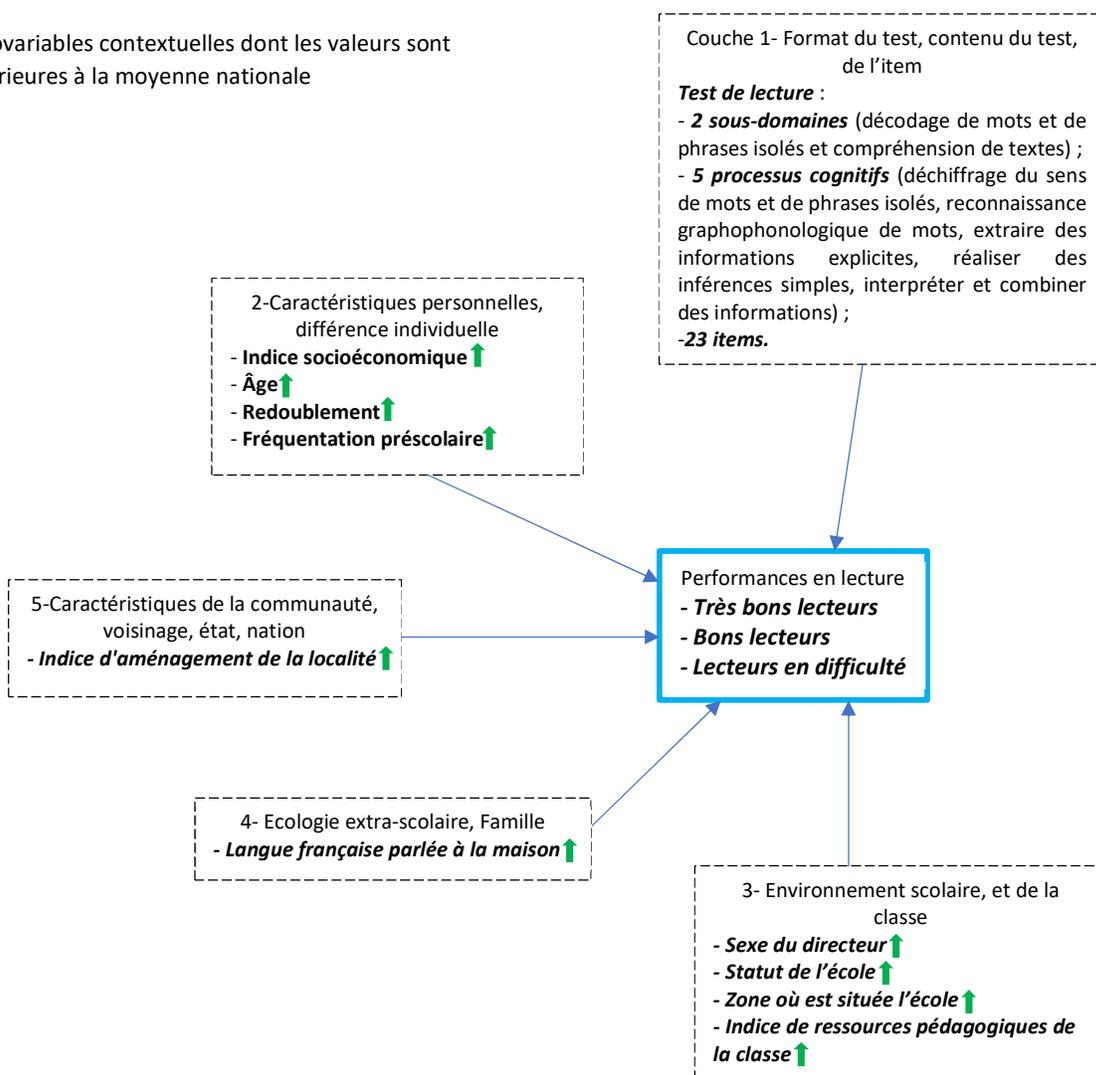


Figure 49 : L'écologie de performances au test de lecture dans la strate Grand Ouest

↑ Covariables contextuelles dont les valeurs sont supérieures à la moyenne nationale

→ Covariables contextuelles dont les valeurs sont proches à la moyenne nationale

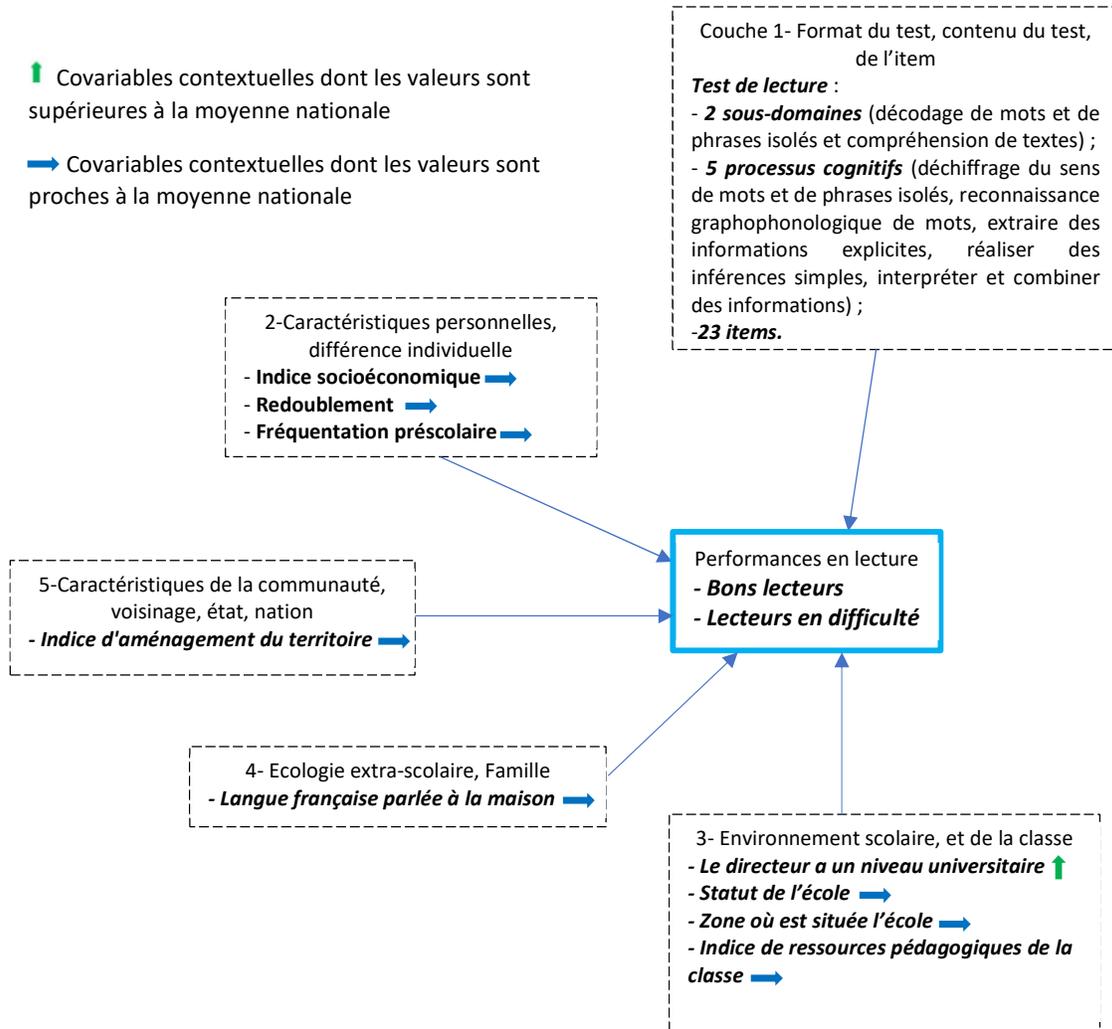


Figure 50 : L'écologie de performances au test de lecture dans la strate Grand Centre

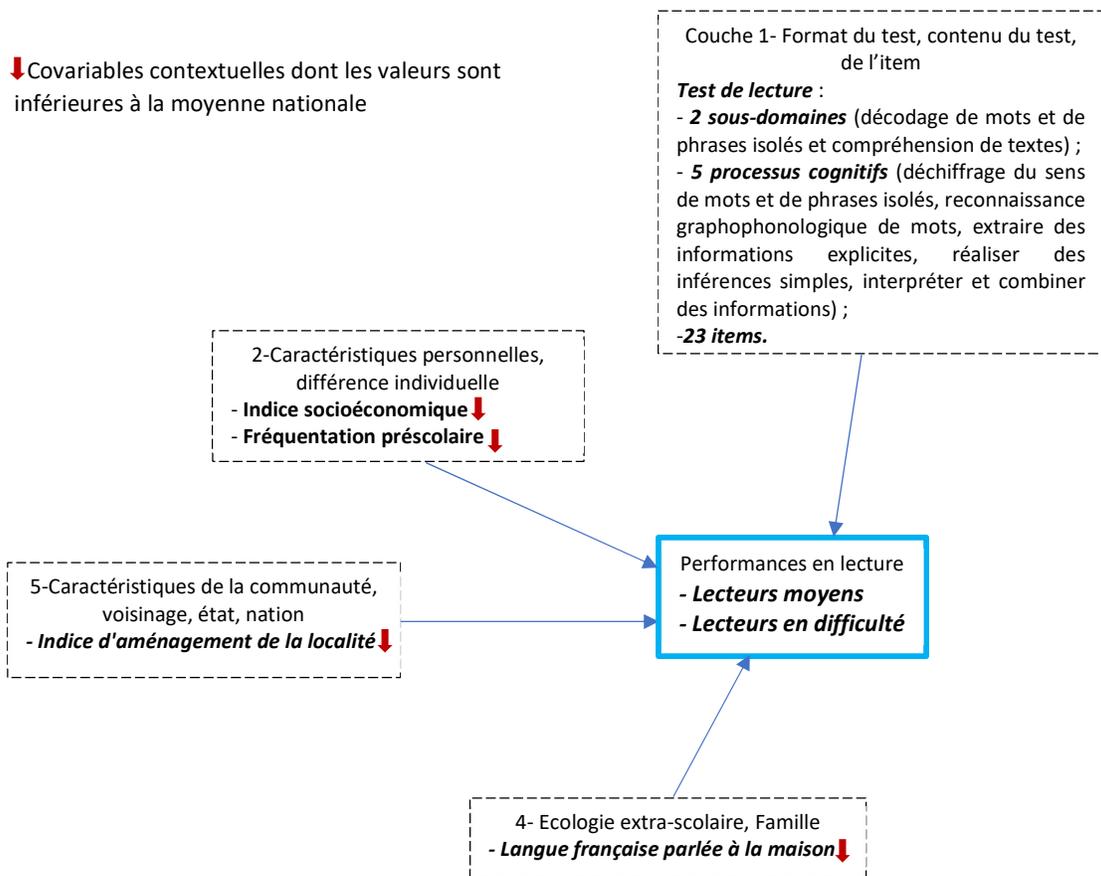


Figure 51 : L'écologie de performances au test de lecture dans la strate Grand Nord

Ces présentations écologiques des résultats au test de lecture PASEC2014 que nous venons de réaliser proposent donc des lectures contextualisées des résultats en fonction des strates. Chaque performance est située en fonction de son contexte ce qui permet ainsi une lecture plus équitable de ces performances. Dans la strate Grand Ouest par exemple, nous comprenons ainsi que les trois profils de performances (« très bons lecteurs », « bons lecteurs » et « lecteurs en difficulté ») émergent en raison des caractéristiques des élèves qui les prédisposent à fournir de bons résultats mais également d'un environnement scolaire et extrascolaire à conditions « favorables » et « déterminants ». Dans la strate Grand Nord par contre, les deux profils de performances (« lecteurs moyens » et « lecteurs en difficulté ») sont tributaires des caractéristiques de ces élèves qui les rendent vulnérables et d'un environnement extrascolaire à conditions « défavorables » mais « déterminants ». Cette lecture rejoint fortement le point de vue de Mislevy (2018) selon qui, lorsque nous évaluons, nous observons

des candidats agir dans une situation particulière et nous devons donc interpréter cette situation et les actions dans un contexte social, et tirer des conclusions à propos de ces candidats.

Notons cependant que la prise en compte de données supplémentaires, notamment individuelles et contextuelles, lors de l'estimation des performances des individus n'est pas une problématique totalement absente dans les ELE (Von Davier et Sinharay, 2013). Au contraire, elle est au centre des recherches en cours dans les programmes tel que le PISA, le TIMSS et le PIRLS (Von Davier et Sinharay, 2013). Actuellement, le PISA utilise une version multidimensionnelle du modèle Rasch comme modèle analytique de base tandis que le TIMSS et le PIRLS utilisent respectivement des modèles IRT logistique à deux paramètres (2PL) et à trois paramètres (3PL) (Von Davier et Sinharay, 2013). Le choix de ces modèles de mesure traduit déjà le souhait de ces programmes d'exploiter les nombreux avantages de l'IRT associés à une approche explicative basée sur des informations supplémentaires autre que les réponses aux items pour tenir compte de certaines caractéristiques des sous-groupes d'individus (Von Davier et Sinharay, 2013). Des discussions sur l'insertion ou pas de données supplémentaires dans les modèles lors de l'estimation des performances des individus sont également en cours (Von Davier et Sinharay, 2013).

Les voix les plus concordantes vont dans le sens d'encourager les recherches qui visent à étendre les modèles IRT actuelles pour les rendre moins restrictifs (Von Davier et Sinharay, 2013). Cette option privilégie ainsi une évolution du modèle IRT vers un modèle de plus en plus souple au lieu d'une révolution qui le remplacerait par des modèles alternatifs à l'exemple de l'ACL avec covariables que nous avons utilisé dans le cadre de nos travaux. Ce qui est important de souligner ici est le consensus grandissant autour de l'idée d'intégrer des données supplémentaires, en plus des réponses aux items, lors de l'estimation des performances des individus à un test. Dans le cadre des analyses ACL qui nous concernent, la majeure partie des recherches existantes sur le sujet s'accorde à souligner que le modèle ACL avec covariables présente plus d'avantages que le modèle ACL de base sans covariables (Muthén, 2004 ; Reboussin et al, 2008 ; Vermunt et Magidson, 2005 ; Wurpts et Geiser, 2014). Pour ce qui des modèles IRT couramment utilisées dans les ELE, les recherches à venir éclaireront davantage sur l'apport des modèles qui incluent des covariables versus les modèles actuels (Von Davier et Sinharay, 2013). Les résultats de nos analyses ne remettent donc pas en cause les cadres d'analyses des performances actuels dans les ELE, mais viennent plutôt enrichir les réflexions qui sont déjà en cours au sein de ces différents programmes d'évaluations. Et ces résultats, comme nous en avons discuté, apportent des éclairages nouveaux et proposent une lecture

contextualisée des performances à un test. Opter pour un tel point de vue sur la performance a des implications sur l'utilisation des résultats qui se dégagent. Dans la section suivante, nous esquissons donc quelques recommandations à la lumière des différentes écologies de performances que nous avons obtenu.

5.2.2 Constats et recommandations au regard des écologies de réponses au test de lecture

Le choix d'une méthode d'analyse des données d'une évaluation standardisée va au-delà des aspects techniques et scientifiques (Pons, 2011 ; Wagemaker, 2014). La méthode d'analyses serait en effet un outil puissant de problématisation de l'action publique (Pons, 2011 ; Wagemaker, 2014). Le choix par les ELE des modèles de mesure qui aboutissent à des scores agrégés par discipline et des échelles de performance communes pour tous les candidats est souvent perçu comme une volonté de la part de ces programmes d'imposer un modèle unique et uniforme d'éducation (Pons, 2011). Ainsi, les décisions prises autour de ces résultats tendent le plus souvent à proposer des actions dont le but est d'avoir des systèmes éducatifs où les conditions de réussite seraient les mêmes pour chacun. Le tableau 37 présente les différentes recommandations formulées auprès des pouvoirs publics camerounais à la suite des tests PASEC2014.

Tableau 37 : Piste de réflexions du PASEC à l'issus des évaluations en lecture et en mathématiques des élèves de 6e année primaire (PASEC, 2015, p. 149-154)

Recommandations	Actions
Améliorer le niveau d'acquisition des élèves dans les disciplines fondamentales, notamment en lecture et en mathématiques	<ul style="list-style-type: none"> - Identifier et accompagner les élèves en difficulté d'apprentissage dès leur entrée au primaire afin d'éviter que ces difficultés ne se traduisent en échecs scolaires ; - Tenir compte, lors des formations continues des enseignants, des domaines dans lesquels les élèves ont le plus de difficulté (niveaux sous le seuil « suffisant ») ; - Inciter les enseignants à développer des stratégies spécifiques en faveur des élèves les plus faibles (niveau 1 et sous le niveau 1) ; - S'assurer de la prise en compte des difficultés d'apprentissage des élèves lors de la révision du curriculum officiel ; - Donner la priorité aux zones (notamment la zone Grand Nord dans le sous-système francophone), où les difficultés scolaires sont plus fortes, dans la mise en œuvre des actions d'amélioration de la qualité des apprentissages ; - Renforcer les actions déjà entreprises dans le cadre du Document de Stratégie Sectorielle de l'Éducation et de la Formation.
Réduire les disparités de contexte et de réussite entre les zones et les groupes d'élèves	<ul style="list-style-type: none"> - Mettre en œuvre des actions d'accompagnement en faveur des élèves en difficulté d'apprentissage, notamment pour les plus défavorisés ou ceux se trouvant en milieu rural, afin de leur assurer les mêmes chances de réussite ; - Renforcer les actions en faveur de la réduction des disparités entre les filles et les garçons, notamment en ce qui concerne l'enseignement des mathématiques dans le sous-système francophone ; - Veiller à une politique équitable de dotation en ressources scolaires entre les zones et entre les écoles au sein des zones.
Repenser la politique de redoublement	<ul style="list-style-type: none"> - L'efficacité des mesures de remédiation mises en place pour accompagner les redoublants et limiter le redoublement, afin de les renforcer et de les rendre plus efficaces. Pour rappel, l'une de ces mesures porte sur l'enseignement compensatoire à travers lequel des

Recommandations	Actions
Promouvoir l'utilisation des données sur les apprentissages dans le suivi des politiques éducatives pour assurer un meilleur pilotage du système éducatif	<p>cours de soutien pour remédier aux problèmes sont organisés pour les élèves en difficulté ;</p> <ul style="list-style-type: none"> - La mise en place de mesures innovantes et adaptées pour accompagner les élèves le plus tôt possible dans leur scolarité, tout en renforçant les connaissances des acteurs du milieu de l'éducation sur les difficultés des élèves et les mécanismes de remédiation, tant au niveau des décideurs politiques que des personnels au niveau local. - Renforcer le système national d'évaluation à tous les niveaux pour assurer une mesure dans le temps basée sur des objectifs nationaux ; - Promouvoir l'utilisation des données des évaluations des acquis scolaires parmi les indicateurs de suivi du plan sectoriel de l'éducation ; - Faire interagir les différents types d'évaluation et les articuler dans le temps et dans le cycle de l'éducation de base pour obtenir des données fiables sur les résultats et permettre le suivi de ceux-ci ; - Impliquer toute la communauté éducative, en amont et en aval, dans les évaluations pour faciliter sa compréhension et la prise en charge des recommandations qui pourraient en découler ; - Diffuser largement les résultats des études et évaluations auprès de la communauté éducative.
Mener des analyses thématiques supplémentaires pour comprendre certains résultats restés à l'état descriptif	<ul style="list-style-type: none"> - Mener une étude relative aux contenus et aux bénéfices de l'éducation préscolaire ; - Mener une étude du redoublement, notamment sur l'effectivité des mesures de réduction ainsi que de leur efficacité ; - Analyser l'effet des caractéristiques des enseignants sur les performances scolaires ; - Analyser le contexte des élèves atypiques : il s'agit d'élèves qui arrivent à surmonter des conditions (socioéconomiques) défavorisées et obtiennent de bons résultats.

Dans le cadre de notre travail, nous ne portons pas de jugement critique à l'endroit des différentes recommandations formulées par le PASEC auprès des pouvoirs publics camerounais. Nous présentons ces recommandations pour juste illustrer cet idéal d'équité que visent les ELE pour tous les élèves. C'est le cas par exemple des actions la première recommandation « **améliorer le niveau d'acquisition des élèves dans les disciplines fondamentales, notamment en lecture et en mathématiques** » ou encore de la deuxième recommandation « **réduire les disparités de contexte et de réussite entre les zones et les groupes d'élèves** ». Nous n'insistons donc que sur les constats et les recommandations supplémentaires que nos analyses suggèrent.

Une grande inconnue : l'environnement extrascolaire

Nos analyses nous ont révélé trois types d'écologies de performances au test de lecture au Cameroun : (i) une écologie à conditions « **favorables** » et aux facteurs « **scolaire et extrascolaire déterminants** » ; (ii) une écologie à conditions « **modérées** » et aux facteurs « **scolaire et extrascolaire déterminants** » et (iii) une écologie à conditions « **défavorables** » et à facteur « **extrascolaire déterminant** ». Dans les trois écologies nous notons un dénominateur commun : un environnement « **extrascolaire déterminant** ». Les leviers d'actions sur lesquels les décideurs du secteur éducatif peuvent actionner sont cependant limités

autour de l'environnement scolaire (allocation de ressources, choix de méthodes pédagogiques, ...). L'environnement extrascolaire, qui peut échapper au contrôle des décideurs du secteur éducatif joue pourtant un rôle déterminant dans la formation des différents profils de performance. Cette situation pourrait en partie expliquer l'absence de résultats dans le cadre des actions menées par les pouvoirs publics pour améliorer l'équité et la qualité de l'éducation en l'occurrence dans la strate Grand Ouest.

En effet, au Cameroun, des programmes ont vu le jour dans le but de relever les insuffisances des élèves suite aux constats alarmants fournis par les ELE (MINEPAT, 2013). C'est le cas du Programme d'Amélioration de la qualité de l'Éducation de Base (PAQUEB) et du Programme d'Amélioration de l'Équité et de la Qualité de l'Éducation (PAEQUE). À travers ces programmes, l'état camerounais a pu recruter plus d'enseignants dans des zones défavorisées, renforcer le réseau scolaire à travers la construction de nouvelles salles de classe et distribuer gratuitement des manuels essentiels aux élèves du primaire et du matériel pédagogique aux enseignants (MINEPAT, 2013). Cependant, le Document de Stratégie du Secteur de l'Éducation et de la Formation (DSSEF) élaboré en 2013 déplore des résultats mitigés notamment dans les zones de la strate Grand Nord (MINEPAT, 2013). Ce constat est appuyé par la dernière évaluation PASEC2014 qui souligne les mêmes déficits et les mêmes disparités que ceux relevés dans les précédentes évaluations (PASEC, 2016).

Les résultats de nos analyses montrent justement une strate Grand Nord dont le milieu scolaire est non stimulant et un milieu extrascolaire « **déterminant** » mais dont les conditions sont « **défavorables** ». Ce constat met malheureusement les pouvoirs publics dans les pays en voie de développement face à des problématiques difficiles pour l'amélioration de la qualité de l'éducation compte tenu de tels paramètres qui leurs échappent. Les attentes envers les élèves en matière de performances devraient donc être revues en fonction du fait que ceux-ci soient dans des localités dont les conditions sont « **favorables** » ou « **défavorables** ».

La nécessité de réviser les standards de performances

Les résultats de nos analyses montrent qu'il existe des « **plafonds** » de performances en fonction des différentes strates. En effet, pendant que dans la strate Grand Ouest, le meilleur profil de performances est celui des « **très bons lecteurs** », dans la strate Grand Centre il est caractérisé par « **bons lecteurs** » et seulement par « **lecteurs moyens** » dans la strate Grand Nord. Comme nous l'avons présenté, les différentes écologies semblent avoir une emprise sur les éventuelles performances en lecture des élèves. Dans le système éducatif camerounais

actuel, les critères d'évaluation et les attentes sont identiques pour tous les élèves (MINEDUB, 2018). Pourtant, comme le souligne si bien Blais (2008), le « seuil suffisant » de maîtrise d'un contenu ne devrait pas être figé dans le temps et dans l'espace. La délimitation d'une compétence minimale pourrait tenir compte en plus des tâches à réaliser par les candidats, du contexte dans lequel ils produisent ces résultats (Blais, 2008). C'est pourquoi, certains processus d'établissement de standards en termes de niveaux de compétences attendus ressemblent à des contrats sociaux (Blais, 2008). Ainsi, dans ces ententes, ce qui constitue la performance attendue pour chacun des niveaux établis est arrêté de façon consensuelle en tenant compte de certaines réalités contextuelles (Blais, 2008).

L'uniformité du « seuil suffisant » de maîtrise de contenus dans le système éducatif camerounais est probablement à l'origine des taux élevés de redoublement dans les strates Grand Nord et Grand Centre comparativement à la strate Grand Ouest. Les responsables du secteur éducatif gagneraient éventuellement à réviser ces seuils en fonctions des strates. En effet, plusieurs études présentent le redoublement comme inefficace, socialement injuste, et favorisant le décrochage scolaire (Galand et al, 2019). Cependant, à défaut de procéder à de telles réformes, ces responsables peuvent bonifier les milieux scolaires notamment dans les strates Grand Ouest et Grand Centre ; les résultats ne nous permettant pas d'avancer des recommandations dans ce sens dans la strate Grand Nord.

Bonifier les différentes écologies en agissant sur les éléments actifs du milieu scolaire

Deux éléments actifs du milieu scolaires par strates peuvent être bonifiés dans les différentes écologies de performances en lecture. Pour la strate Grand Ouest, il s'agit des ressources pédagogiques dans les classes et du sexe du directeur de l'école. Les analyses montrent en effet que dans cette strate, les ressources pédagogiques disponibles dans les salles de classes et les écoles dirigées par des femmes favorisent l'émergence des profils « **bons lecteurs** » et « **très bons lecteurs** ». Dans cette strate, la dotation des salles de classes démunies en matériels pédagogiques ou encore la promotion du leadership féminin au sein des directions d'écoles peuvent probablement favoriser la maîtrise de la lecture par les élèves.

Dans la strate Grand Centre d'autre part, les éléments actifs du milieu scolaire sont les ressources pédagogiques disponibles dans les salles de classe et le niveau universitaire du directeur. De ce côté, les décideurs pourraient en plus des dotations des ressources pédagogiques, promouvoir la formation continue des enseignants.

Ces différents constats et recommandations sont cependant à prendre avec quelques réserves. En effet, comme dans toute recherche, la nôtre présente également quelques limites qu'il conviendrait de souligner. Dans la section suivante, nous présentons donc ces différentes limites.

5.3 Principales limites

La première limite de notre recherche, liée à l'exhaustivité, est une limite inhérente à toute recherche. En effet, l'approche écologique essaye de recenser tous les facteurs qui peuvent médier le processus cognitif dans le cadre des réponses aux items d'un test (Zumbo et al, 2015). Cependant, il est généralement très difficile de recenser tous les facteurs en lien avec le sujet d'intérêt de la recherche (Boudon et Fillieule, 2002). De même en raison des moyens souvent limité dans le temps, un chercheur est amené à opérer des choix sur ces facteurs et à retenir ceux qui lui semble plus pertinents pour son objet d'étude (Boudon et Fillieule, 2002). Cependant, ce n'est pas parce qu'un facteur est laissé de côté que ses effets sur notre objet de recherche s'estompent pour autant (Boudon et Fillieule, 2002).

Dans nos travaux, dans un souci de pouvoir comparer nos résultats avec ceux du PASEC, nous n'avons tenu compte que de certains éléments pouvant faire partie de l'écologie de réponse aux items d'un test. Nous n'avons pas ainsi tenu compte du climat social au sein des écoles, des pratiques pédagogiques en cours dans les écoles, de la motivation des enseignants et des conditions de passation du test de lecture entre autres dans ces analyses. Ces données existent cependant dans la base de données PASEC2014. Les effets des facteurs dans nos écologies sont donc à prendre avec précaution car l'introduction d'un ou plusieurs facteurs que nous venons de citer dans notre modèle ACL avec covariables aurait eu pour conséquences de fournir de profils de performances différents dans nos strates mais également des descriptions différentes des composantes de nos écologies.

La deuxième limite de la recherche que nous avons menée est lié au plan d'analyse contextuel du modèle écologique de Zumbo et ses collègues (2015). En effet, le modèle de ces auteurs ne tient pas compte des possibles interactions qui peuvent exister entre les différentes couches du modèle. Dans la réalité pourtant, au-delà de leurs effets sur les phénomènes sociaux, les facteurs interagissent constamment (Holmbeck, 2002). La prise en compte de ces interactions lors des analyses offre donc des angles d'analyses plus intéressants comparativement à celui que nous avons proposé.

Une autre limite de notre recherche, et pas la moindre, est l'inaccessibilité des contenus des items PASEC. Comme nous l'avons souligné dans la partie méthodologique de nos travaux, le PASEC ne met malheureusement pas à la disposition du grand public le contenu de items de ses tests. Cette contrainte ne nous permet pas de faire des inférences au niveau items mais plutôt sur l'ensemble du test ; une situation qui impacte énormément l'interprétation des profils de performances dans nos strates. Le choix de libellé pour ces classes latentes, comme nous l'avons mentionné, est approximatif et peut malheureusement ne pas correspondre à la réalité. De même, cette contrainte nous impose de donner les mêmes étiquettes à des classes latentes dont les caractéristiques sont similaires mais pas identiques. C'est le cas par exemple des profils de performances « **bons lecteurs** » et « **lecteurs en difficulté** » qui, dans nos analyses, se retrouvent dans toutes les strates. Cependant, en regardant de plus près ces profils, nous pouvons constater que par endroit, ces profils présentent des différences en fonction des strates.

Toutes ces limites ne doivent pas cependant occulter la pertinence des résultats que nous venons de présenter. Comme le disent certains auteurs, il n'existe pas de modèle de mesure parfait (Bertrand et Blais, 2004 ; Magnani et al, 2017). Chacun de ces modèles offre au chercheur l'opportunité de voir un objet d'étude sur un angle donné. Le plus important est donc de se munir des instruments qui permettent de nous rapprocher le plus près possible de la réalité (Magnani et al, 2017). L'approche écologique dans ce sens et à travers l'analyse ACL avec covariables, nous rappelle que la performance à un test ne doit pas être liée uniquement aux capacités cognitives des individus mais devrait plutôt être considérée comme un phénomène qui se réalise dans un réseau interconnecté de connaissances, de caractéristiques individuelles et de contextes particuliers (Mislevy, 2018 ; Zumbo et al, 2015). Un tel point de vue appelle à des prises de décisions en matière de politiques en fonction des lieux dans lesquels les individus sont évalués pour des solutions plus proche de leurs réalités.

CONCLUSION

Aujourd'hui, les ELE sont très répandues dans le globe et leurs résultats sont au centre des actions en matière de politiques publiques éducatives (Hogan, 2017 ; Loye, 2011 ; Rutkowski et al, 2013 ; Wagemaker, 2014). Les résultats de ces évaluations servent également de « baromètres » pour comparer des individus, des établissements scolaires ou des systèmes éducatifs (Blais, 2008 ; Loye, 2011 ; Pons, 2011). Cette visée comparative amène généralement les ELE à estimer les performances des candidats uniquement sur la base des réponses qu'ils fournissent aux tests (Blais, 2008 ; Pons, 2011). Une telle approche met plus d'emphasis sur les aspects cognitifs des candidats occultant quelque peu la dimension sociale de l'évaluation (McNamara, 2007 ; Mislevy, 2018 ; Zumbo et al, 2015). Une telle approche implique également des prises de décisions holistiques pour tous les acteurs du système éducatif (Pons, 2011)

L'approche écologique que nous avons utilisé pour notre recherche par contre, soutient une vision locale et contextuelle de la performance (McNamara, 2007 ; Zumbo et al, 2015). Cette approche peut certes, sembler difficile à concilier avec le besoin des ELE de généraliser les scores mais elle essaye d'appréhender la performance à un test dans toute sa complexité. De plus en plus de recherches soutiennent en effet que la performance à un test se réalise dans un réseau interconnecté de connaissances, de caractéristiques individuelles et de contextes particuliers (McNamara, 2007 ; Zumbo et al, 2015). Une telle approche a pour mérite de mettre en exergue les différentes écologies dans lesquelles émergent les profils de performance et de souligner la nécessité de mener des politiques différenciées en fonction des milieux dans lesquels évoluent les différents candidats (McNamara, 2007 ; Zumbo et al, 2015).

Notre travail, implémenté grâce au modèle ACL avec covariables, est le premier à utiliser une approche écologique sur des données d'une ELE, notamment sur les données du test de lecture PASEC2014 au Cameroun. Nous avons pu montrer que les élèves de 6^e année du primaire au Cameroun évoluent dans différentes écologies de performances au test de lecture conduisant ainsi à des performances plus variées de la part de ces apprenants. Ces résultats soulignent ainsi la nécessité de mener des politiques différenciées auprès des élèves en fonction des écologies dans lesquelles ils se trouvent pour des actions plus efficaces et plus proches de la réalité des apprenants.

Notre choix d'utiliser le modèle ACL avec covariables pour implémenter une approche écologique sur nos données n'est cependant pas un appel à une révolution contre les modèles

de mesures, robustes et aux pouvoirs déjà établis, utilisés actuellement par les ELE pour estimer les performances aux différents tests. Comme le soulignent Von Davier et Sinharay (2013), l'alternative la plus viable serait une évolution plus souple des modèles présentement utilisées par les ELE, notamment le modèle de RASH, qu'une alternative plus complexe et moins utilisée des praticiens du domaine de la mesure.

D'autres recherches empiriques devraient donc se pencher sur le sujet pour explorer ou développer des modèles de mesures qui implémenteraient au mieux l'approche écologique. Ces recherches devraient également explorer la possibilité de tenir compte des différentes interactions entre les facteurs au moment d'estimer les performances. Ces réflexions seront d'autant plus importantes compte tenu de la nature complexe des phénomènes étudiés en sciences sociales et en sciences de l'éducation (Bertrand & Blais, 2004). Des domaines où notre façon de penser les modèles peut nous éloigner de la réalité ou nous en rapprocher (Mislevy, 2018).

BIBLIOGRAPHIE

- Abarda A. (2013). *Utilisation des modèles à classes latentes pour la classification des élèves en niveaux scolaires* [mémoire de maîtrise, Institut National de Statistique et d'Economie Appliquée, Maroc].
https://www.researchgate.net/profile/Pierre_Varly/publication/278673287_Utilisation_des_modeles_a_classes_latentes_pour_la_classification_des_eleves_en_niveaux_scolaires/data/55829acf08ae6cf036c18eb9/Rapport-FINAL2013.pdf
- Absil, G., Vandroorne, C., et Demarteau, M. (2012). Bronfenbrenner, écologie du développement humain. *Réflexion et action pour la promotion de la santé*.
- Agresti, A., Mehta, C. R., et Patel, N. R. (1990). Exact inference for contingency tables with ordered categories. *Journal of the American Statistical Association*, 85(410), 453-458.
- Alidou, H. (2006). Utilisation des langues Africaines et l'alphabétisation: conditions, facteurs et processus (Bénin, Burkina Faso, Cameroun, Tanzanie et Zambie). *Document de travail pour la Biennale de l'Association pour le développement de l'éducation en Afrique (Programmes d'alphabétisation efficaces, session parallèle A-3), Libreville, Gabon (27-31 mars)*.
- American Educational Research Association, American Psychological Association, et National Council on Measurement in Education. (2003). Normes de pratique de testing en psychologie et en éducation (Trad. par G. Sarrazin). *Montréal: Institut de recherches psychologiques*.
- Asil, M., et Brown, G. T. (2016). Comparing OECD PISA reading in English to other languages: Identifying potential sources of non-invariance. *International Journal of Testing*, 16(1), 71-93.
- Becker, H. S. (2002). *Les ficelles du métier : comment conduire sa recherche en sciences sociales*. Paris: Paris : La Découverte.
- Bertrand, R., et Blais, J. G. (2004). *Modèles de mesure: l'apport de la théorie des réponses aux items*. PUQ.
- Blais, J. G. (2008). Les standards de performance en éducation. *Mesure et évaluation en éducation*, 31(2), 93-105.
- Blay, M. (2007). Comte et Duhem ou la construction d'une optique positive. *Revue philosophique de la France et de l'étranger*, 132(4), 493-504.
- Boote, D. N., et Beile, P. (2005). Scholars before researchers: On the centrality of the dissertation literature review in research preparation. *Educational researcher*, 34(6), 3-15.
- Boudon, R., et Fillieule, R. (2002). *Les méthodes en sociologie (12e éd.)*. Paris: Presses universitaires de France.
- Bronfenbrenner, U. (1979). *The ecology of human development*. Harvard university press.
- Bronfenbrenner, U. (1994). Ecological models of human development. *Readings on the development of children*, 2(1), 37-43.
- Carley, K. M. (2001). Learning and using new ideas: A sociocognitive perspective. Diffusion processes and fertility transition: Selected perspectives, 179-207.
- Chalhoub-Deville, M. (2003). Second language interaction: Current perspectives and future trends. *Language testing*, 20(4), 369-383.
- Chinapah, V. (2003). Monitoring learning achievement (MLA) project in Africa. *Association for the Development of Education in Africa (ADEA)*. Paris: International Institute for Educational Planning.
- CIEP. (2014). Synthèse de la conférence internationale. *Réussir l'éducation en Afrique : l'enjeu des langues*: CIEP.
- Clark, S., & Muthén, B. (2009). Relating latent class analysis results to variables not included in the analysis. Récupéré de <http://www.statmodel.com/download/relatinglca.pdf>
- Cohen, L., Manion, L., et Morrison, K. (2002). *Research methods in education*. routledge.

- Collins, L. M., et Lanza, S. T. (2009). *Latent class and latent transition analysis: With applications in the social, behavioral, and health sciences* (Vol. 718). John Wiley et Sons.
- Damon, J. (2009). La fièvre de l'évaluation. *Sciences humaines*, 208, 18-23.
- Demeure, B. (2016). La sexualité dans la philosophie positiviste d'Auguste Comte, et ses effets dans le discours nationaliste de l'Action française. *Topique*, (1), 139-145.
- Engelhard Jr, G. (2008). Historical perspectives on invariant measurement: Guttman, Rasch, and Mokken. *Measurement*, 6(3), 155-189.
- Engelhard Jr, G. (1992). Historical views of invariance: evidence from the measurement theories of Thorndike, Thurstone, and Rasch. *Educational and psychological measurement*, 52(2), 275-291.
- Ercikan, K. (1998). Translation effects in international assessments. *International Journal of Educational Research*, 29(6), 543-553.
- Ercikan, K. (2002). Disentangling sources of differential item functioning in multilanguage assessments. *International Journal of Testing*, 2(3-4), 199-215.
- Ercikan, K., et Koh, K. (2005). Examining the construct comparability of the English and French versions of TIMSS. *International Journal of Testing*, 5(1), 23-35.
- Felouzis, G., et Hanhart, S. (2011). *Gouverner l'éducation par les nombres?. Usages, débats et controverses*. De Boeck.
- Fox, J. D. (2003). From products to process: An ecological approach to bias detection. *International Journal of Testing*, 3(1), 21-47.
- Galand, B., Lafontaine, D., Baye, A., Dachet, D., & Monseur, C. (2019). Le redoublement est inefficace, socialement injuste, et favorise le décrochage scolaire. *Cahiers des Sciences de l'Education*, (38), 1-33.
- Gbati, K. Y. (2006). fréquentation préscolaire et performances scolaires des élèves du cours préparatoire première année à Lomé. *Revue du Cameroun-Nouvelle série B*, 1, 51-62.
- Gelfert, A. (2017). The ontology of models. In *Springer handbook of model-based science* (pp. 5-23). Springer, Cham.
- Giasson, J., et Vandecasteele, G. (2011). *La lecture: apprentissage et difficultés*. De Boeck.
- Giasson, J., et Escoyez, T. (2013). *La lecture : de la théorie à la pratique* (4e éd, Ser. Outils pour enseigner. français). De Boeck.
- Gingras, Y. (2008). Du mauvais usage de faux indicateurs. *Revue d'histoire moderne et contemporaine*, 5(55-4bis), 67-79.
- Hagenaars, J. A., et McCutcheon, A. L. (Eds.). (2002). *Applied latent class analysis*. Cambridge University Press.
- Hayes, N., O'Toole, L., et Halpenny, A. M. (2017). *Introducing Bronfenbrenner: A guide for practitioners and students in early years education*. Routledge.
- Heyneman, S. P., et Lee, B. (2014). The impact of international studies of academic achievement on policy and research. *Handbook of international large-scale assessment: Background, technical issues and methods of data analysis*, 37-72.
- Hogan, T. P. (2017). *Introduction à la psychométrie*. Chenelière éducation.
- Holmbeck, G. N. (2002). Post-hoc probing of significant moderational and mediational effects in studies of pediatric populations. *Journal of pediatric psychology*, 27(1), 87-96.
- Hox, J. J., Moerbeek, M., et Van de Schoot, R. (2017). *Multilevel analysis: Techniques and applications*. Routledge.
- Huang, X., Wilson, M., et Wang, L. (2016). Exploring plausible causes of differential item functioning in the PISA science assessment: language, curriculum or culture. *Educational Psychology*, 36(2), 378-390.

- Kody, Z. D. (2005). *La Dynamique des Langues Camerounaises en Contact avec le Français* (Doctoral dissertation, University of Yaounde I).
- Koné, A. S. (2007). *L'influence de trois facteurs familiaux sur la réussite scolaire au primaire et au secondaire d'élèves arabophones, créolophones et francophones de Montréal*.
- Lafortune, L. (2009). Évolution des profils d'état de santé et utilisation des services sociosanitaires chez les personnes âgées.
- Landry, R., Becheikh, N., Amara, N., Ziam, S., Idrissi, O., et Castonguay, Y. (2008). La recherche, comment s'y retrouver. *Revue systématique des écrits sur le transfert de connaissances en éducation*, 47.
- Lazarsfeld, P. F., et Henry, N. W. (1968). *Latent structure analysis*. Houghton Mifflin Co.
- Legendre, R. (2005). Dictionnaire de l'éducation (3e éd.). *Montréal, Canada: Guérin*.
- Lord, H. G. (1973). *Ex Post Facto Studies as a Research Method*. Special Report No. 7320.
- Loye, N. (2011). Panorama des programmes actuels d'enquêtes à large échelle. *Mesure et évaluation en éducation*, 34(2), 3-24.
- Lubke, G., et Muthén, B. O. (2007). Performance of factor mixture models as a function of model size, covariate effects, and class-specific parameters. *Structural Equation Modeling*, 14(1), 26-47.
- Luke, D. A. (2004). *Multilevel modeling* (Vol. 143). Sage
- Magnani, L., et Bertolotti, T. (Eds.). (2017). *Springer handbook of model-based science*. Springer.
- Malo, C. (2000). Le modèle écologique du développement humain: conditions nécessaires de son utilité réelle. *Atelier présenté dans le cadre du Psycho-stage*.
- Martinková, P., Drabinová, A., Liaw, Y. L., Sanders, E. A., McFarland, J. L., et Price, R. M. (2017). Checking equity: Why differential item functioning analysis should be a routine part of developing conceptual assessments. *CBE—Life Sciences Education*, 16(2), rm2.
- Mc Andrew, M., Garnett, B., Ledent, J., Ungerleider, C., Adumati-Trache, M., et Ait-Said, R. (2008). La réussite scolaire des élèves issus de l'immigration: une question de classe sociale, de langue ou de culture?. *Éducation et francophonie*, 36(1), 177-196.
- McCutcheon, A. L. (1987). *Latent class analysis* (No. 64). Sage.
- McNamara, T. F. (1997). 'Inter action' in second language performance assessment: Whose performance?. *Applied linguistics*, 18(4), 446-466.
- McNamara, T., et Roever, C. (2006). *Language testing: The social dimension* (Vol. 1). C. Roever.
- McNamara, T. (2007). Language testing: A question of context. *Language testing reconsidered*, 131-137.
- Millsap, R. E., et Kwok, O. M. (2004). Evaluating the impact of partial factorial invariance on selection in two populations. *Psychological methods*, 9(1), 93.
- MINEDUB (2018). *Curriculum de l'enseignement primaire francophone camerounais*. MINEDUB, Cameroun. [http://minedub.cm/uploads/media/03-Programme Niveau 03 2019 01.pdf](http://minedub.cm/uploads/media/03-Programme_Niveau_03_2019_01.pdf)
- MINEPAT (2013). *Document de Stratégie du Secteur de l'Éducation et de la Formation*. Ministère de l'Économie, de la Planification et de l'Amenagement du Territoire. https://planipolis.iiep.unesco.org/sites/planipolis/files/ressources/cameroon_education_sector_plan_2013_2020.pdf
- Mislevy, R. J. (2018). *Sociocognitive foundations of educational measurement*. Routledge.
- Mondada, L., et Pekarek Doehler, S. (2000). Interaction sociale et cognition située: quels modèles pour la recherche sur l'acquisition des langues?. *Acquisition et interaction en langue étrangère*, (12).
- Mons, N., et Crahay, M. (2011). L'évaluation des performances scolaires des élèves: un instrument d'évaluation des politiques éducatives?
- Muthén, B. O. (2004). Latent variable analysis: Growth mixture modeling of individual differences in experimental designs: A latent variable framework for analysis and power estimation. *Psychological Methods*, 2, 371-402.

- Nylund-Gibson, K., & Masyn, K. E. (2016). Covariates and mixture modeling: Results of a simulation study exploring the impact of misspecified effects on class enumeration. *Structural Equation Modeling: A Multidisciplinary Journal*, 23(6), 782-797.
- OCDE (2016). *PISA 2015, résultats à la loupe*. OCDE. <https://www.oecd.org/pisa/pisa-2015-results-in-focus-FR.pdf>
- Oehmichen, H., et Viedrov, O. (2017). L'usage comparé des statistiques par Gabriel Tarde et Emile Durkheim. *Statistique et Société*, 4(3), 65-82.
- Oliveri, M. E., Olson, B. F., Ercikan, K., et Zumbo, B. D. (2012). Methodologies for investigating item- and test-level measurement equivalence in international large-scale assessments. *International Journal of Testing*, 12(3), 203-223.
- Papalia, M. D., Olds, M. S., et Feldman, M. R. (2010). *Psychologie du développement humain*. De Boeck Supérieur.
- PASEC (2015). *Performances des systèmes éducatifs en Afrique subsaharienne : Compétences et facteurs de réussite au primaire francophone*. PASEC, CONFEMEN, Dakar. <http://www.pasec.confemen.org/publication/pasec2014-rapport/>
- PASEC (2016). *Performances du système éducatif camerounais : Compétences et facteurs de réussite au primaire*. PASEC, CONFEMEN, Dakar. <http://www.pasec.confemen.org/publication/pasec2014-cameroun/>
- PASEC (2017a). *Cadre de référence des tests PASEC2014 de lecture et de mathématiques de fin de scolarité primaire PASEC2014*. PASEC, CONFEMEN, Dakar. <http://www.pasec.confemen.org/publication/cadre-de-reference-des-tests-pasec2014-de-lecture-et-de-mathematiques-en-fin-de-scolarite-primaire/>
- PASEC (2017b). *Manuel d'exploitation des données : Évaluation internationale PASEC2014*. PASEC, CONFEMEN, Dakar. <http://www.pasec.confemen.org/manuel-des-donnees/>
- PASEC (2017c). *Rapport technique de l'évaluation internationale PASEC2014*. PASEC, CONFEMEN, Dakar. <http://www.pasec.confemen.org/rapport-technique-pasec2014/>
- Patrick, M. E., Bray, B. C., et Berglund, P. A. (2016). Reasons for marijuana use among young adults and long-term associations with marijuana use and problems. *Journal of studies on alcohol and drugs*, 77(6), 881-888.
- Pons, X. (2011). Les méthodes des enquêtes internationales et leurs fonctions politiques. L'exemple de la France face à PISA (1995-2008). *Mesure et évaluation en éducation*, 34(2), 57-85.
- Poulin, R., Beaumont, C., Blaya, C., et Frenette, E. (2015). Le climat scolaire: un point central pour expliquer la victimisation et la réussite scolaire. *Canadian Journal of Education*, 38(1), 1.
- Quivy, R., et Van Campenhoudt, L. (2006). *Manuel de la Recherche en Sciences sociales*, Coll. *Psychologie sociale*, 3ème édition, Édition Dunod, Paris.
- Reboussin, B. A., Ip, E. H., & Wolfson, M. (2008). Locally dependent latent class models with covariates: an application to under-age drinking in the USA. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 171(4), 877-897.
- Rutkowski, L., von Davier, M., et Rutkowski, D. (Eds.). (2013). *Handbook of international large-scale assessment: Background, technical issues, and methods of data analysis*. CRC Press.
- Ryan, K. E., et Chiu, S. (2001). An examination of item context effects, DIF, and gender DIF. *Applied Measurement in Education*, 14(1), 73-90.
- Schmiege, S. J., Masyn, K. E., et Bryan, A. D. (2018). Confirmatory latent class analysis: Illustrations of empirically driven and theoretically driven model constraints. *Organizational Research Methods*, 21(4), 983-1001.
- Sireci, S. G. (2011). Evaluating test and survey items for bias across languages and cultures. *Cross-cultural research methods in psychology*, 216-240.

- Spearman, C. (1904). Measurement of association, Part II. Correction of 'systematic deviations'. *Am J Psychol*, 15, 88-101.
- Tadadjeu, M., et Mba, G. (1997). *L'utilisation des langues nationales dans l'éducation au Cameroun. Revue Tranel (Travaux neuchâtelois de linguistique)*, 26, 59-75.
- Taylor, C. S., et Lee, Y. (2012). Gender DIF in reading and mathematics tests with mixed item formats. *Applied Measurement in Education*, 25(3), 246-280.
- UNESCO (2005). Éducation pour tous l'exigence de qualité. *Rapport mondial de suivi sur l'EPT*.
- Udry, J. R. (2003). The national longitudinal study of adolescent health (Add Health), Waves I and II, 1994–1996; Wave III, 2001–2002. *Carolina Population Center, University of North Carolina at Chapel Hill*. <http://www.cpc.unc.edu/projects/addhealth/data>.
- Van der Maren, J. M. (1996). *Méthodes de recherche pour l'éducation*. Presses de l'Université de Montréal et de Boeck.
- Vermunt, J.K., and Magidson, J. (2005). Structural equation models : mixture models, in *Encyclopedia of Statistics in Behavioral Science*, eds B.Everittand D.Howell (Chichester:Wiley), 1922–1927.
- Von Davier, M., Gonzalez, E., Kirsch, I., et Yamamoto, K. (Eds.). (2012). *The role of international large-scale assessments: Perspectives from technology, economy, and educational research*. Springer Science et Business Media.
- Von Davier, M., et Sinharay, S. (2013). Analytics in international large-scale assessments: Item response theory and population models. *Handbook of international large-scale assessment: Background, technical issues, and methods of data analysis*, 155-174.
- Wagemaker, H. (2013). International large-scale assessments: From research to policy. *Handbook of international largescale assessment: Background, technical issues, and methods of data analysis*, 11-36.
- Wambach, M. (2003). Chapitre 6. Langue maternelle/Langue étrangère. Dans : Paul Rivenc éd., *Apprentissage d'une langue étrangère/seconde. Vol. 3: La méthodologie* (pp. 135-157). Louvain-la-Neuve, Belgique: De Boeck Supérieur. doi:10.3917/dbu.riven.2003.01.0135.
- Willms, J. D. (2003). Dix hypothèses sur l'impact des gradients socioéconomiques et des différences communautaires sur le développement de l'enfant. *Rapport final*.
- Wurpts, I. C., & Geiser, C. (2014). Is adding more indicators to a latent class analysis beneficial or detrimental? *Results of a Monte-Carlo study. Frontiers in psychology*, 5, 920.
- Zumbo, B. D. (2007). Three generations of DIF analyses: Considering where it has been, where it is now, and where it is going. *Language assessment quarterly*, 4(2), 223-233.
- Zumbo, B. D., Liu, Y., Wu, A. D., Shear, B. R., Olvera Astivia, O. L., et Ark, T. K. (2015). A methodology for Zumbo's third generation DIF analyses and the ecology of item responding. *Language Assessment Quarterly*, 12(1), 136-151.

ANNEXES

Annexe 1 : Variables utilisées pour le calcul du niveau socioéconomique (SES) des familles des élèves (PASEC, 2017b, p. 167)

Variables	Libellés des variables
qe620	Nombre de livres à la maison
qe621a	La maison de l'élève a un téléviseur
qe621b	La maison de l'élève a un ordinateur
qe621c	La maison de l'élève a une radio
qe621d	La maison de l'élève a un lecteur DVD
qe621e	La maison de l'élève a une chaîne HIFI
qe621f	La maison de l'élève a un téléphone portable
qe621g	La maison de l'élève a un congélateur/réfrigérateur
qe621h	La maison de l'élève a un ventilateur
qe621i	La maison de l'élève a un climatiseur
qe621j	La maison de l'élève a une cuisinière (plaque chauffante et four)
qe622a	La maison de l'élève a une table
qe622b	La maison de l'élève a une machine à coudre
qe622c	La maison de l'élève a un fer à repasser électrique
qe622d	La maison de l'élève a une voiture ou un camion
qe622e	La maison de l'élève a un tracteur
qe622f	La maison de l'élève a une moto, mobylette ou scooter
qe622g	La maison de l'élève a un vélo
qe623	Les murs de ta maison sont en ...
qe624	À la maison, il y a des latrines (toilettes sans eau courante)
qe625	À la maison, il y a des toilettes avec chasse d'eau (toilettes avec eau courante)
qe626	À la maison, il y a l'électricité (groupe électrogène / abonnement à l'électricité)
qe627	À la maison, il y a un puits
qe628	À la maison, il y a un robinet

Annexe 2 : Variables utilisées pour le calcul de l'indice de ressources pédagogiques de la classe (PASEC, 2017b, p. 168)

Variables	Libellés des variables
qm624a	Nombre d'élèves par manuel de langue d'enseignement
qm624b	Nombre d'élèves par manuel de mathématiques
qm625a	L'enseignant dispose du manuel de langue
qm625b	L'enseignant dispose du manuel de mathématiques
qm625c	L'enseignant dispose du guide pédagogique en langue
qm625d	L'enseignant dispose du guide pédagogique en mathématiques
qm625e	L'enseignant dispose du programme de langue
qm625f	L'enseignant dispose du programme de mathématiques
qm658a	La classe dispose d'un bureau pour l'enseignant
qm658b	La classe dispose d'une chaise pour l'enseignant
qm658c	La classe dispose d'une armoire (coffre/malle)
qm658d	La classe dispose d'un tableau (mural ou mobile)
qm658e	La classe dispose de craies
qm658f	La classe dispose d'une règle pour tableau
qm658g	La classe dispose d'une équerre pour tableau
qm658h	La classe dispose d'un compas pour tableau
qm658i	La classe dispose d'un dictionnaire
qm658j	La classe dispose d'une carte du monde/globe
qm658k	La classe dispose d'une carte de l'Afrique
qm658l	La classe dispose d'une carte du pays
qm667a	Nombre réel de places assises dans la classe
qm667b	Nombre réel de supports (table ou pupitre) pour écrire dans la classe

Annexe 3 : Variables utilisées pour le calcul de l'indice d'aménagement du territoire (PASEC, 2017b, p. 171)

Variables	Libellés des variables
qd25a	La localité à laquelle est rattachée l'école dispose d'une route goudronnée
qd25b	La localité à laquelle est rattachée l'école dispose de l'électricité
qd25c	La localité à laquelle est rattachée l'école dispose d'un collège
qd25d	La localité à laquelle est rattachée l'école dispose d'un lycée
qd25f	La localité à laquelle est rattachée l'école dispose d'un dispensaire / centre de santé
qd25g	La localité à laquelle est rattachée l'école dispose d'un hôpital
qd25h	La localité à laquelle est rattachée l'école dispose d'un poste de police ou d'une gendarmerie
qd25i	La localité à laquelle est rattachée l'école dispose d'une banque
qd25j	La localité à laquelle est rattachée l'école dispose d'une caisse d'épargne ou d'agence de micro-crédit
qd25l	La localité à laquelle est rattachée l'école dispose d'un centre culturel ou social, d'une bibliothèque

Annexe 4 : Variables utilisées pour le calcul de l'indice d'implication de la communauté (PASEC, 2017b, p. 173)

Variables	Libellés des variables
qd29	Fréquence des réunions avec la collectivité locale (APEE, conseil d'école, comité de gestion ou autre association locale de la société civile) au cours d'une année scolaire
qd40a	Dans l'école, il existe une association de parents d'élèves et d'enseignants (APEE)
qd40b	Dans l'école, il existe une association des mères éducatrices (AME)
qd40c	Dans l'école, il existe une coopérative scolaire
qd40d	Dans l'école, il existe un conseil d'école
qd40e	Dans l'école, il existe un comité de gestion
qd41b	Les activités extrascolaires et sorties ont été réalisées au cours de l'année scolaire en partenariat avec l'association des parents d'élèves (APEE/AME/COGES...) et/ou le comité de gestion
qd41c	Les constructions ou les réhabilitations ont été réalisées au cours de l'année scolaire en partenariat avec l'association des parents d'élèves (APEE/AME/COGES...) et/ou le comité de gestion
qd41d	L'achat d'équipements pédagogiques a été réalisé au cours de l'année scolaire en partenariat avec l'association des parents d'élèves (APEE/AME/COGES...) et/ou le comité de gestion
qd41e	La prise en charge des enseignants a été réalisé au cours de l'année scolaire en partenariat avec l'association des parents d'élèves (APEE/AME/COGES...) et/ou le comité de gestion
qd43a	Le directeur reçoit une aide financière de la communauté locale (village, comité de gestion...)
qd43b	Le directeur reçoit un appui de la communauté locale (village, comité de gestion...) pour la construction, la réhabilitation ou l'entretien des infrastructures (salles de classe, logements des enseignants...)
qd43c	Le directeur reçoit un appui de la communauté locale (village, comité de gestion...) pour l'aménagement, l'entretien ou la réparation des équipements et du mobilier
qd43d	Le directeur reçoit un appui de la communauté locale (village, comité de gestion...) pour la dotation des manuels scolaires
qd43e	Le directeur reçoit un appui de la communauté locale (village, comité de gestion...) pour la dotation de fournitures scolaires
qd43f	Le directeur reçoit un appui de la communauté locale (village, comité de gestion...) pour la prise en charge de la cantine / l'approvisionnement des repas
qd43g	Le directeur reçoit un appui de la communauté locale (village, comité de gestion...) pour la formation pédagogique des enseignants
qd43h	Le directeur reçoit un appui de la communauté locale (village, comité de gestion...) pour l'organisation d'activités extrascolaires (sorties scolaires)
qd43i	Le directeur reçoit un appui de la communauté locale (village, comité de gestion...) pour la prise en charge des frais d'examen
qd43k	Le directeur reçoit un appui de la communauté locale (village, comité de gestion...) pour la prise en charge du personnel enseignant bénévole
qd43l	Le directeur reçoit un appui de la communauté locale (village, comité de gestion...) aux élèves en difficulté (soutien scolaire, cours supplémentaires...)

Annexe 5 : Niveau de compétences, scores minimums et description de compétences au test de lecture en 6^e année primaire au Cameroun (PASEC, 2016, p. 44)

Niveaux	Scores minimums des élèves	Description des compétences
Niveau 4	595,1	Les élèves peuvent effectuer un traitement de texte global pour tirer parti de textes narratifs, informatifs et de documents. Sur ces supports, ils sont capables d'associer et d'interpréter plusieurs idées implicites en s'appuyant sur leurs expériences et leurs connaissances. En lisant des textes littéraires, les élèves sont capables d'identifier l'intention de l'auteur, de déterminer le sens implicite et d'interpréter les sentiments des personnages. En lisant des textes informatifs et des documents, ils mettent en lien des informations et comparent des données pour les exploiter.
Niveau 3	518,4	Les élèves sont capables de combiner deux informations explicites dans un passage de document ou de réaliser des inférences simples dans un texte narratif ou informatif. Ils peuvent extraire des informations implicites de supports écrits en donnant du sens aux connecteurs implicites, aux anaphores ou aux référents. Les élèves localisent des informations explicites dans des textes longs et des documents dont le texte est discontinu.
Niveau 2	441,7	Les élèves mobilisent leur capacité de décodage orthographique pour identifier et comprendre des mots isolés issus de leur vie quotidienne. Ils sont également en mesure de localiser des informations explicites dans des textes courts et moyens en prélevant des indices de repérage présents dans le texte et les questions. Les élèves parviennent à paraphraser les informations explicites d'un texte.
Niveau 1	365,0	Les élèves ont développé des capacités de décodage et sont capables de les mobiliser pour comprendre des mots isolés issus de leur vie quotidienne, mais sont en difficulté pour comprendre le sens de textes courts et simples.
Sous le niveau 1	72,1	Les élèves qui se situent à ce niveau ne manifestent pas les compétences mesurées par ce test en langue d'enseignement. Ces élèves sont en difficulté sur les connaissances et compétences du niveau 1.

Annexe 6 : Statistiques descriptives des données

			Grand Ouest	Grand Centre	Grand Nord	Ensemble
Couche 2 « Caractéristiques personnelles »	Sexe	n	218	178	169	565
		% Garçons	48,2	53,4	62,7	54,2
		% Filles	51,8	46,6	37,3	45,8
	Âge	Moyenne	11,8	12,3	12,9	12,3
		Ecart type	1,6	1,6	1,7	1,7
		Minimum	9	9	9	9
		Maximum	16	16	18	18
	Indice socioéconomique	Moyenne	56,6	50,8	44,4	51,2
		Ecart type	8,0	8,0	8,1	9,4
	À déjà redoublé	% Oui	46,1	63,5	61,1	56,0
		% Non	53,9	36,5	38,9	44,0
	A fait le Préscolaire	% Oui	60,6	51,5	7,5	41,9
		% Non	39,4	48,5	92,5	58,1
	Sexe de l'enseignant	n	50	61	54	165
% Hommes		68,0	78,7	100,0	82,4	
% Femmes		32,0	21,3	0,0	17,6	
L'enseignant a un niveau universitaire	% Oui	16,0	29,5	3,7	17,0	
	% Non	84,0	70,5	96,3	83,0	
Ancienneté de l'enseignant (en années)	Moyenne	13,3	11,5	8,3	11,0	
	Ecart type	8,0	9,2	5,9	8,1	
Couche 3 « Environnement scolaire, Enseignant, classe, ... »	Sexe du directeur	n	50	61	54	165
		% Hommes	72,0	82,0	94,4	83,0
		% Femmes	28,0	18,0	5,6	17,0
	Le directeur a un niveau universitaire	% Oui	18,0	34,4	5,6	20,0
		% Non	82,0	65,6	94,4	80,0
	Statut (% élèves dans les écoles...)	Publiques	58,7	65,7	89,9	70,3
		Privées	41,3	34,3	10,1	29,7
	Zone où est située l'école (% élèves en zone...)	Urbaine	57,3	24,7	7,1	32,0
Rurale		42,7	75,3	92,9	68,0	
Indice de ressources pédagogiques de la classe	Moyenne	50,1	45,6	38,3	44,6	
	Ecart type	6,8	7,8	6,5	8,5	
Couche 4 « Écologie extra-scolaire, Famille, ... »	Langue française parlée à la maison (% élèves)	Toujours le français	12,9	9,6	5,7	9,7
		Parfois le français et une autre langue	81,4	73,1	64,3	73,8
		Jamais le français	5,7	17,4	29,9	16,5
Couche 5 « Caractéristiques de la communauté, voisinage, état, ... »	Indice d'implication de la communauté	Moyenne	47,8	44,5	47,8	46,6
		Ecart type	8,3	10,2	11,6	10,2
	Indice d'aménagement du territoire	Moyenne	52,3	46,3	44,1	47,4
		Ecart type	7,2	8,7	7,4	8,5

Annexe 7 : Distribution des indicateurs de prévalences et probabilités conditionnelles par strates et par profil

	Grand Ouest			Grand Centre		Grand Nord	
	Très bons lecteurs	Bons lecteurs	Faibles lecteurs	Bons lecteurs	Faibles lecteurs	Lecteurs moyens	Faibles lecteurs
Taux de prévalence	0,46	0,39	0,14	0,65	0,35	0,47	0,53
Item_f1	1,00*	0,79	0,21*	0,90	0,42	0,62	0,25
Item_f2	0,98	1,00	0,53	0,98	0,74	0,93	0,64
Item_f3	1,00	0,87	0,30	0,98	0,68	0,85	0,35
Item_f4	0,87	0,26	0,10	0,60	0,21	0,17	0,19
Item_f5	0,55	0,20	0,19	0,39	0,15	0,10	0,21
Item_f6	0,86	0,60	0,16	0,78	0,45	0,74	0,27
Item_f7	0,98	0,53	0,25	0,73	0,38	0,38	0,20
Item_f8	0,91	0,82	0,28	0,94	0,54	0,75	0,39
Item_f9	0,72	0,51	0,28	0,55	0,35	0,41	0,31
Item_f10	0,50	0,20	0,19	0,27	0,13	0,34	0,09
Item_f11	0,47	0,09	0,10	0,31	0,2	0,19	0,13
Item_f12	0,87	0,48	0,29	0,62	0,29	0,39	0,24
Item_f13	1,00	0,74	0,25	0,87	0,25	0,74	0,20
Item_f14	0,98	0,78	0,43	0,87	0,31	0,74	0,28
Item_f15	0,81	0,45	0,19	0,63	0,30	0,40	0,26
Item_f16	0,92	0,76	0,34	0,80	0,33	0,66	0,27
Item_f17	1,00	0,88	0,41	0,89	0,51	0,91	0,45
Item_f18	0,43	0,29	0,12	0,42	0,16	0,31	0,14
Item_f19	0,40	0,27	0,35	0,36	0,29	0,26	0,25
Item_f20	0,39	0,21	0,15	0,20	0,18	0,19	0,13
Item_f21	0,94	0,61	0,30	0,68	0,31	0,49	0,16
Item_f22	0,87	0,85	0,36	0,92	0,52	0,92	0,33
Item_f23	0,98	0,79	0,17	0,97	0,35	0,68	0,09

* En vert, les probabilités de réponses supérieures à 0,5

* En rouge, les probabilités de réponses inférieures à 0,5