

# CONNAISSANCES SCOLAIRES ET COMPÉTENCES MESURÉES PAR PISA

*RÉSULTATS AUX ÉPREUVES CANTONALES ET À PISA :  
QUELLES RELATIONS CHEZ LES ÉLÈVES VAUDOIS ?*

Bruno Suchaut

Avec la collaboration de Ladislav Ntamakiliro

159 / Mars 2014



Unité de recherche pour le pilotage  
des systèmes pédagogiques



## REMERCIEMENTS

Les auteurs remercient vivement Christian Nidegger du Service de la recherche en éducation (SRED) pour la mise à disposition des données PISA 2009 ainsi que la Direction générale de l'enseignement obligatoire (DGEO) pour nous avoir permis de disposer des résultats des élèves aux épreuves cantonales de référence. Les remerciements s'adressent également à Jean Moreau et à Jean-Pierre Abbet (URSP) pour les échanges fructueux que nous avons eus sur cette recherche.

*Dans le cadre des missions de l'URSP,  
ses travaux sont publiés sous l'égide  
du Département de la formation, de la jeunesse et de la culture.  
Les publications expriment l'avis de leurs auteurs  
et n'engagent pas les institutions dont ils dépendent.*

# SOMMAIRE

<b>1</b>	<b>INTRODUCTION</b> .....	<b>5</b>
<b>2</b>	<b>CONTEXTE ET PROBLÉMATIQUE</b> .....	<b>7</b>
<b>3</b>	<b>STRUCTURE DES ANALYSES ET DONNÉES MOBILISÉES</b> .....	<b>11</b>
3.1	LA STRUCTURE DES ANALYSES .....	11
3.2	LES CARACTÉRISTIQUES DE L'ÉCHANTILLON.....	12
3.3	LES RÉSULTATS AUX ÉVALUATIONS (PISA ET ECR) .....	14
3.4	LES RELATIONS STATISTIQUES INTERNES AUX ÉVALUATIONS .....	17
3.4.1	<i>Les relations entre les domaines de compétences dans PISA</i> .....	17
3.4.2	<i>Les relations entre les disciplines aux ECR</i> .....	19
<b>4</b>	<b>LES DÉTERMINANTS DES PERFORMANCES DES ÉLÈVES AUX ÉVALUATIONS</b> .....	<b>21</b>
4.1	L'INFLUENCE DES CARACTÉRISTIQUES INDIVIDUELLES.....	21
4.2	L'INFLUENCE DU CONTEXTE SCOLAIRE.....	27
4.2.1	<i>Caractéristiques des élèves et contexte scolaire</i> .....	27
4.2.2	<i>L'influence du milieu socio-économique dans les filières scolaires</i> .....	29
4.3	ANALYSE GLOBALE DES DÉTERMINANTS DES PERFORMANCES.....	31
<b>5</b>	<b>LES RELATIONS ENTRE LES RÉSULTATS À PISA ET LES RESULTATS AUX ECR</b> .....	<b>37</b>
5.1	UNE ANALYSE GLOBALE DES RELATIONS ENTRE LES SCORES AUX ÉVALUATIONS .....	37
5.1.1	<i>Les relations statistiques entre les scores globaux</i> .....	37
5.1.2	<i>Les relations en fonction des niveaux de compétences à PISA</i> .....	40
5.2	LES DÉTERMINANTS DES ÉCARTS DES RÉSULTATS ENTRE PISA ET LES ECR.....	46
<b>6</b>	<b>ANALYSE SELON LES DOMAINES DE COMPÉTENCES</b> .....	<b>49</b>
6.1	LES RELATIONS ENTRE PISA ET LES DIMENSIONS DES ECR EN LECTURE.....	49
6.2	LES RELATIONS ENTRE PISA ET LES DIMENSIONS DES ECR EN MATHÉMATIQUES .....	52
6.3	L'INFLUENCE DES CARACTÉRISTIQUES DES ÉLÈVES DANS LES DIFFÉRENTES DIMENSIONS DES ECR .....	55
<b>7</b>	<b>CONCLUSION</b> .....	<b>57</b>
<b>8</b>	<b>RÉFÉRENCES</b> .....	<b>59</b>
<b>9</b>	<b>LISTE DES TABLEAUX, GRAPHIQUES ET SCHÉMAS</b> .....	<b>62</b>
	<b>ABRÉVIATIONS ET ACRONYMES</b> .....	<b>65</b>



## 1 INTRODUCTION

Lancé par l'Organisation de coopération et de développement économiques (OCDE), le programme international pour le suivi des acquis des élèves, plus connu sous son acronyme «PISA», a pris une importance croissante dans les débats sur l'éducation; dans de nombreux pays, à chaque vague d'enquête, les résultats de PISA nourrissent les réflexions des décideurs et suscitent les commentaires des acteurs de l'éducation, même si c'est avant tout le classement des pays qui est au centre des débats, notamment au niveau politique. La méthodologie utilisée par PISA lui confère une qualité scientifique incontestable et, à ce titre, PISA peut-être considéré comme très fiable sur le plan de la mesure. Toutefois, au niveau théorique, l'approche par compétence, mobilisée dans cette enquête, pose un certain nombre de questions qui ne sont pas indépendantes de positionnements idéologiques (Planche, 2012). On peut ainsi débattre de la pertinence d'une conception universelle des compétences et de sa dimension utilitaire. Dans le prolongement du débat, on peut s'interroger sur la capacité de PISA de rendre compte de la qualité des systèmes d'enseignement, à cause de la variété de structure et d'organisation de ces derniers et, surtout des différences de curriculums d'un pays à l'autre. Ce n'est donc pas la mesure des compétences qui serait à mettre en cause, mais plutôt la définition même de ces compétences et leur correspondance avec des programmes scolaires.

Si de nombreux travaux ont contribué à alimenter le débat sur les finalités de PISA et son influence sur le pilotage des systèmes éducatifs, il existe peu de recherches qui ont tenté de confronter les résultats de cette enquête à d'autres indicateurs de mesure du niveau des performances des élèves. Or, l'idée de mettre en correspondance PISA avec des mesures provenant d'évaluations établies sur la base des contenus d'enseignement aux élèves n'est pas sans intérêt; cette manière originale d'interroger PISA peut en effet permettre de tester empiriquement la pertinence du programme PISA de l'Organisation de coopération et de développement économiques (OCDE), pour évaluer les résultats des systèmes éducatifs, en fonction de leurs objectifs associés au niveau des acquisitions des élèves.

La variété des programmes d'enseignement, du cycle primaire et du cycle secondaire inférieur, en termes d'objectifs et de contenus, au sein de l'échantillon des nombreux pays concernés par PISA nourrit l'hypothèse principale de la recherche exposée dans ce rapport. C'est en effet dans cette perspective que la présente recherche vise à décrire et à analyser les relations entre les résultats des élèves vaudois à l'enquête PISA et ceux obtenus aux épreuves cantonales de référence (ECR), pour ce même échantillon— les ECR pouvant être considérées comme des évaluations externes standardisées (Ntamakiliro, Tessaro, 2010; Ntamakiliro, Moreau, 2011). La possibilité technique de fusionner les deux bases de données (PISA et ECR) était une condition minimale et indispensable pour explorer empiriquement notre questionnement — ce qui a pu être réalisé. Peu de recherches avec une méthodologie comparable ont été conduites sur le plan international, ce qui confère à notre travail un caractère à la fois exploratoire et original.

L'objectif général de notre démarche est d'apprécier le degré de recouvrement des deux mesures mobilisées: les compétences telles que PISA les envisage d'une part et, d'autre part, les connaissances et compétences de nature scolaire basées sur les plans d'études via les épreuves cantonales (ECR). On imagine clairement que selon l'intensité de la relation entre ces deux mesures, les interprétations des résultats pourront être différentes. Si l'hypothèse d'un fort recouvrement des résultats des élèves est rejetée, on pourrait ainsi exprimer certaines limites à l'utilisation de PISA pour apprécier la qualité du système d'enseignement vaudois et son degré d'équité. Dans ce cas, l'enquête PISA serait considérée principalement comme un indicateur, parmi d'autres, du niveau de

compétences des élèves. Au contraire, dans le cas d'une relation statistique forte entre les scores obtenus de PISA et ceux relevés dans les épreuves vaudoises (ECR), il est clair que les estimations fournies par l'enquête internationale PISA sur le niveau des élèves auraient bien la capacité et la légitimité à instruire des éléments de la politique éducative au niveau cantonal vaudois.

Au-delà de cet objectif, les analyses mobilisées dans notre travail s'attacheront plus précisément à examiner les variations éventuelles de la relation entre les deux mesures de performances selon les disciplines et les domaines évalués, mais aussi selon les caractéristiques des élèves et certains aspects du contexte d'enseignement. Les données réunies pour traiter notre questionnaire concernent l'enquête PISA 2009 sur la base de l'échantillon vaudois d'élèves de 9<sup>e</sup> année. Les résultats aux épreuves cantonales de référence (ECR) de 8<sup>e</sup> année pour ce même échantillon sont également mobilisés<sup>1</sup>. Afin de disposer, pour les deux mesures de performances, de données comparables, nous ferons porter principalement les analyses sur deux domaines disciplinaires, à savoir le français et les mathématiques.

Sur le plan formel, le rapport s'attachera tout d'abord à situer le contexte et à décrire la problématique et les enjeux de la recherche, en s'appuyant sur la littérature existante dans le champ scientifique, les comparaisons internationales en éducation principalement. Les données mobilisées et la structure des analyses seront ensuite présentées avant d'aborder la partie statistique avec des analyses menées en plusieurs phases. Nous nous intéresserons dans un premier temps à l'influence des caractéristiques sociodémographiques des élèves et de leur contexte d'enseignement (les filières d'enseignement) sur les résultats des différentes évaluations à notre disposition. Au cours de la phase suivante, qui constituera le cœur des analyses, nous examinerons les relations globales entre les indicateurs de performances des élèves (PISA, ECR et les notes) en identifiant et en interprétant les écarts éventuels entre les deux mesures. Nous distinguerons ensuite dans les analyses, les domaines de compétences dans les types d'épreuves afin d'identifier les variations potentielles et de mieux comprendre la nature de la relation entre PISA et les ECR construites à partir des objectifs des programmes d'enseignement; l'influence des caractéristiques individuelles sera à nouveau examinée dans les différents domaines. Les notes des élèves seront ensuite mobilisées pour étudier leurs relations avec les épreuves externes (ECR et PISA) et, avec certaines limites, pour mesurer l'impact des caractéristiques individuelles sur cet indicateur de réussite interne au système. Les principaux résultats de la recherche feront enfin l'objet d'une courte synthèse conclusive, ils pourront nourrir la réflexion sur l'utilisation des résultats des prochaines enquêtes PISA et en proposer une autre lecture. Plus généralement, on peut aussi attendre que les enseignements produits par cette recherche puissent contribuer à alimenter le débat scientifique sur la portée et les limites des enquêtes internationales et régionales en éducation, et à mieux connaître leurs spécificités par rapport à des évaluations nationales basées sur les plans d'étude.

---

<sup>1</sup> C'est en effet la mesure des résultats scolaires la plus proche dans le parcours des élèves de l'évaluation PISA 2009 qui a pu être retenue. Il existe donc un écart de plusieurs mois entre les résultats des ECR (mai 2008) et ceux relevant de l'enquête PISA.

## 2 CONTEXTE ET PROBLÉMATIQUE

Avec PISA, les évaluations internationales en éducation ont gagné une visibilité sans précédent, qui se limitait auparavant à la sphère de la recherche (Bottani, Vrignaud, 2005). L'intérêt porté aujourd'hui par les décideurs politiques à PISA est indéniable et, dans de nombreux pays, les politiques éducatives utilisent des résultats de l'enquête pour appuyer et justifier en partie leurs orientations. L'objectif de PISA est clairement de fournir des informations fiables sur les performances des systèmes éducatifs afin de permettre de dégager des perspectives en termes d'amélioration de la qualité et de l'équité (Goldstein, 2008). Il s'agit effectivement de mesurer le rendement des systèmes d'enseignement à l'aune d'indicateurs variés (OCDE, 1999). Fort de l'expérience antérieure en matière d'évaluations internationales, PISA en reprend les caractéristiques classiques sur la base de grands échantillons d'élèves dans un nombre important de pays, avec une ampleur toutefois jamais atteinte dans ces enquêtes précédentes. PISA s'appuie sur des tests standardisés des acquis des élèves et mobilise une méthodologie rigoureuse, tant pour l'élaboration des épreuves et des questionnaires que pour la constitution des échantillons. Par ailleurs, les analyses statistiques utilisent les techniques les plus élaborées en la matière, donnant ainsi aux résultats une grande fiabilité.

Par rapport à d'autres programmes internationaux, PISA présente certaines spécificités qui lui confèrent un intérêt particulier. La première est de tester les élèves à un âge (15 ans) qui correspond, pour de nombreux pays, à la fin de la scolarité obligatoire. La seconde spécificité est de viser trois domaines de compétences (compréhension de l'écrit, culture mathématique et culture scientifique) en donnant à chacun une place plus importante que les autres, à chaque cycle d'évaluation. Rappelons que l'enquête se déroule tous les trois ans et concerne un nombre croissant de pays: de 43 pays en 2000 à 67 pays en 2012 pour la vague la plus récente. L'aspect cyclique de l'enquête constitue sa troisième spécificité; la mesure des évolutions temporelles dans chaque domaine de compétences est donc rendue possible et elle présente un intérêt évident pour suivre les progrès réalisés par les systèmes d'enseignement des différents pays. Outre ces caractéristiques globales, PISA a l'originalité de viser les acquis des élèves au-delà de leur seule dimension scolaire: il s'agit d'évaluer des connaissances et aptitudes dont les jeunes auront besoin tout au long de leur vie d'adulte (*life skills*). C'est donc l'aptitude des jeunes à exploiter leurs savoirs et savoir-faire dans la vie future qui est visée et non la mesure de l'assimilation de telle ou telle matière du programme d'enseignement (OCDE, 2004). Les compétences et connaissances évaluées ne sont pas sélectionnées parce qu'elles appartiennent au même dénominateur commun des programmes scolaires nationaux, mais parce qu'elles sont considérées comme essentielles à la vie future des jeunes (OCDE, 2006).

Il est incontestable que les résultats produits par les vagues successives de PISA ont livré des informations utiles, tant aux décideurs qu'à la communauté des chercheurs. Pour cette dernière, l'exploitation des enquêtes internationales a permis, sur la base d'analyses secondaires, d'apporter une contribution significative à la production scientifique dans plusieurs domaines: l'influence de certains modes d'organisation des systèmes sur les performances des élèves (Duru-Bellat, Suchaut, 2005; Mons, 2004), les inégalités scolaires et sociales (Meuret, Morlaix, 2006; Crahay, Monseur, 2008; Felouzis, 2009; Felouzis, Charmillot, Fouquet-Chauprade, 2011) ou encore sur la ségrégation scolaire (Alegre, Ferrer, 2010). C'est en fait une vision renouvelée des systèmes éducatifs que PISA a permis d'avoir, via la question des différenciations des parcours et des inégalités qui peuvent y être associées (Green, 2008).

En ce qui concerne l'utilisation de PISA par les responsables politiques, les situations sont variables d'un pays à l'autre. Aux Etats-Unis par exemple, les résultats de PISA n'ont pratiquement reçu aucun écho (Martens, Niemann, 2010). Par contre, en Allemagne, PISA a eu un impact très fort et a permis de placer l'éducation au centre des questions politiques (Martens, Niemann, 2010). Au Danemark (Egelund, 2008), au Portugal (Afonso, Costa, 2009), aux Pays-Bas (Kuiper, van den Akker, 2011) et même en Finlande (Rautalin, Alasuutari, 2009), pays pourtant toujours très bien classé dans le palmarès international, PISA a permis aux responsables politiques d'accorder une attention plus particulière aux outils d'évaluation nationale des élèves, voire de prendre des mesures plus globales pour faire évoluer leur système. En France, on a pu constater une prise de distance par rapport aux résultats (au moins jusqu'à PISA 2006), alors qu'en Suisse, PISA a eu une influence qui s'est manifestée notamment lors de la mise en place du projet d'harmonisation de la scolarité obligatoire (HarmoS) et de l'introduction de standards de formation (Moreau, Nidegger, Soussi, 2006). On rappellera par ailleurs que lors de la publication des premiers résultats du programme PISA, en 2000, les acteurs (les responsables politiques et les médias notamment) ont été très surpris de constater que la Suisse était loin de figurer en tête du classement international (Nidegger et al., 2011).

Mais l'enquête PISA a aussi alimenté des débats dans de nombreux pays avec des critiques, parfois fortes, exprimées par certains membres de la communauté scientifique sur les résultats produits et leur utilisation potentielle pour la politique éducative. Toutefois ces critiques n'ont pas porté sur les aspects statistiques (échantillonnage, traitement et analyse des données) caractérisés par une rigueur méthodologique indéniable. Une première critique très générale concerne le choix d'un âge, 15 ans, qui ne confère donc pas à l'échantillon la représentativité d'un niveau scolaire particulier dans les pays comme la France, où les élèves testés sont scolarisés soit dans le secondaire inférieur, soit dans le secondaire supérieur, alors que ce n'est pas le cas dans les pays qui ont recours à la promotion automatique, comme la Finlande ou la Suède par exemple (Grenet, 2008). Mais les débats autour de la validité de PISA ont principalement visé l'aspect de la mesure des acquis des élèves. Ainsi, les biais culturels, toujours possibles dans ce type d'enquêtes, ont été évoqués étant donné le nombre important de pays et de langues concernés. La question de la formulation des items composant les épreuves et des modalités de correction a alors été posée (Robin, 2002); certains chercheurs se sont aussi interrogés sur la conception de la littératie, telle qu'elle est envisagée dans les épreuves PISA (Bautier, Crinon, Rayou, Rochex, 2006). Enfin, et plus largement, c'est le concept théorique de compétence qui peut même être remis en cause avec l'argument selon lequel PISA définit les compétences comme une réalité objective et autonome, alors que celles-ci sont en fait le produit d'un modèle de mesure (Vrignaud, 2006).

Cette critique majeure sur la définition des compétences, mobilisée dans PISA, est à mettre en relation avec l'objectif de l'OCDE qui souhaite, à travers cette enquête, répondre à la question suivante: «*Dans quelle mesure les jeunes adultes sont-ils prêts à relever les défis de demain ?*» (OCDE, 2004, p. 7). Cette interrogation sur la nature des compétences est bien au cœur d'une problématique qui concerne le domaine scientifique, mais aussi la sphère politique. Le fait que PISA se démarque volontairement des connaissances figurant dans les programmes scolaires peut en effet avoir une incidence sur l'interprétation des résultats en termes de palmarès entre pays ou régions: le classement des pays serait-il le même avec des épreuves plus proches des contenus scolaires ?

Plus largement, il est légitime de s'interroger sur l'appréciation du niveau moyen des élèves tel qu'il est évalué par PISA. A notre connaissance, sans doute limitée par le fait que PISA a donné lieu à énormément de publications à travers le monde, il existe peu de

recherches qui aient tenté de rapprocher empiriquement la mesure des compétences de PISA avec celles provenant d'autres évaluations des acquis des élèves (basées sur les plans d'étude notamment). Toutefois, certaines analyses, réalisées principalement dans le contexte anglo-saxon (Shiel, Kellaghan, Moran, 2010) semblent légitimer l'intérêt de notre questionnement. On peut notamment citer un travail conduit aux Etats-Unis qui a porté sur l'analyse des données de TIMSS<sup>2</sup> et du NAEP<sup>3</sup> (Phillips, 2007) et dont l'objectif était de relier statistiquement les échelles de mesure des deux évaluations. Cette méthodologie a permis (en mathématiques et en sciences) d'utiliser les critères de performances du NAEP pour situer les performances des pays participants à TIMSS et de les rendre ainsi plus lisibles et interprétables pour le public américain. Il existe également des travaux qui, avec des perspectives méthodologiques variées, se sont penchés sur la comparaison entre le contenu des items proposés dans PISA et celui présent dans les curriculums; ainsi, une étude française fait état, dans le domaine des mathématiques, de nettes différences en ce qui concerne les notions abordées et les niveaux taxonomiques ciblés dans PISA et le curriculum français (Bodin, 2003). Cela ne semble pas être le cas en Allemagne pour les sciences et les mathématiques car les épreuves proposées par PISA semblent tout à fait en accord avec les exigences des programmes d'enseignement (Prenzel, Zimmer, 2006).

L'objectif de notre mandat de recherche se situe dans cette logique de comparaison des mesures de performances des élèves: celles produites par PISA et celles qui se basent plus directement sur les programmes d'enseignement. En Suisse, des analyses secondaires des données PISA effectuées dans le domaine de la lecture (Broi, Moreau, Soussi, Wirthner, 2003), montrent que les intentions générales des programmes d'enseignement semblent converger avec le cadre conceptuel de PISA, mais il est néanmoins possible que les mesures effectives des compétences des élèves dans cette enquête divergent des autres mesures qui relèvent plus directement des plans d'étude. La question est alors de savoir quel est ce degré de convergence / divergence entre ces deux sources d'informations sur le niveau des élèves. Selon la réponse, l'interprétation des résultats produits par PISA peut varier et même donner lieu à des conclusions différentes sur la qualité du système éducatif. Ainsi, si cette convergence était forte, la pertinence de PISA s'en trouverait renforcée et les résultats des élèves vaudois pourraient effectivement interroger le fonctionnement du système. Si, au contraire, la convergence était faible, cela témoignerait de la distance entre l'approche par compétences telle que PISA la conçoit et le degré de maîtrise des programmes tel qu'il peut s'apprécier par les évaluations locales et scolaires. Dans cette logique, les résultats de PISA mettraient moins en cause la qualité du système que le fait que PISA mesurerait d'autres aspects que ceux visés par les programmes.

Il est toutefois probable que la réponse à cette question ne soit pas de nature dichotomique, mais plus nuancée et que, selon les disciplines et dimensions testées, l'intensité de la relation entre les deux types d'épreuves varie. Dans tous les cas, une telle analyse peut être heuristique car elle interroge le lien entre une évaluation prospective (des compétences utiles tout au long de la vie) et une évaluation bilan (le niveau des acquis des élèves à un moment de la scolarité). En fait, cette étude devrait permettre de tester empiriquement le point de vue de l'OCDE qui annonçait clairement que PISA «*porte sur l'aptitude des jeunes à exploiter leurs savoirs et savoir-faire pour faire face aux défis de la vie réelle et qu'elle ne cherche pas à déterminer dans quelle mesure les élèves ont assimilé une matière spécifique du programme d'enseignement*» (OCDE, 2004, p. 20). Au niveau plus local, pour le Canton de Vaud, notre recherche s'inscrit dans la continuité et la complémentarité de travaux précédents. Il s'agit plus précisément de recherches qui,

<sup>2</sup> Trends in International Mathematics and Science Study.

<sup>3</sup> National Assessment of Educational Progress.

d'une part ont pour objet l'analyse des résultats des élèves aux épreuves cantonales de référence (Ntamakiro, Longchamp, 2009; Ntamakiro, Moreau, 2011) et, d'autre part, sont ciblées sur l'exploitation des résultats vaudois de l'enquête PISA (Moreau, 2004; Moreau, 2007).

### 3 STRUCTURE DES ANALYSES ET DONNÉES MOBILISÉES

La possibilité de disposer de plusieurs évaluations ayant des statuts différents pour un même échantillon d'élèves n'est généralement pas fréquente. Nous avons pu toutefois la saisir dans le contexte vaudois grâce aux épreuves cantonales de référence (ECR) administrées à différents niveaux de la scolarité<sup>4</sup>. Ces épreuves basées sur les contenus des plans d'études, ont pour finalité de contribuer à la qualité du système scolaire en assurant l'harmonisation des exigences de l'enseignement au niveau cantonal, tout en mettant à disposition des enseignants des repères extérieurs permettant d'apprécier objectivement la progression des élèves. Les épreuves cantonales de référence combinent les deux fonctions prévues dans la loi et le règlement scolaires: le monitoring et pilotage de l'enseignement et l'évaluation certificative ou pronostique des élèves<sup>5</sup> (Ntamakiro, Tessaro, 2010). Ce sont donc les résultats à ces épreuves cantonales externes que nous avons rapprochés de ceux de l'enquête PISA<sup>6</sup>. Par ailleurs, les notes scolaires obtenues par les mêmes élèves ont été collectées; elles permettront de compléter l'image des performances scolaires en fournissant une mesure, certes moins fidèles que les évaluations externes, mais en lien direct avec les pratiques des enseignants.

Nous avons mobilisé dans ce travail les données disponibles les plus récentes, à savoir celles provenant de l'enquête PISA 2009 qui concernent les élèves de 9<sup>e</sup> année<sup>7</sup>, testés dans les trois dimensions habituelles: compréhension de l'écrit, culture mathématique et culture scientifique, sachant que, comme en 2000, c'est la compréhension de l'écrit qui était particulièrement ciblée en 2009. Les résultats des mêmes élèves aux ECR de 2008, ont été également collectés sachant que ces élèves étaient scolarisés en 8<sup>e</sup> année. Ce décalage de plusieurs mois entre les dates des évaluations constitue bien sûr un handicap pour la comparaison, mais il reste faible dans le sens où les deux mesures de performances sont relativement proches au niveau temporel quand on considère l'ensemble du parcours scolaire des élèves et même, encore plus largement, tous les facteurs du contexte scolaire et de l'environnement personnel des élèves, qui ont pu jouer sur le développement de leurs compétences. Ce décalage entre les deux mesures de performances ne peut être alors considéré comme un biais important pour notre démarche.

En résumé, pour traiter la problématique, la recherche rassemble les données suivantes pour les élèves vaudois testés lors de l'enquête PISA 2009 : les résultats PISA dans les trois disciplines, les résultats aux épreuves communes de référence (ECR) en fin de 8<sup>e</sup> année (en 2008), les caractéristiques individuelles des élèves contenues dans la base de données PISA.

#### 3.1 LA STRUCTURE DES ANALYSES

Les analyses statistiques à conduire pour traiter la problématique sont potentiellement très nombreuses, certaines d'entre elles peuvent toutefois être considérées comme secondaires dans la mesure où elles correspondent à des étapes intermédiaires pour

<sup>4</sup> En 2<sup>e</sup> année primaire (fin de CYP1) en français, en 4<sup>e</sup> année primaire (fin de CYP2) en français et en mathématiques, en 6<sup>e</sup> année (fin du CYT) en français et en mathématiques et en 8<sup>e</sup> année en français et en mathématiques.

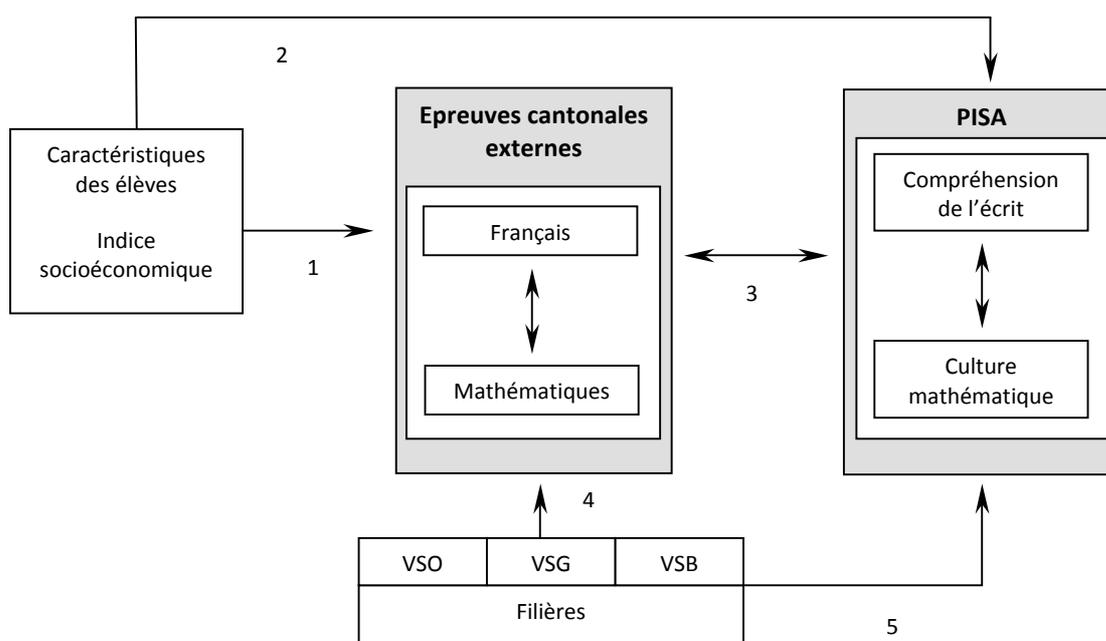
<sup>5</sup> A partir de 2013, les ECR de 6<sup>e</sup> année (8<sup>e</sup> Harmos) ont un objectif différent puisqu'elles participeront à l'orientation des élèves et compteront dans la moyenne à la hauteur de 30%.

<sup>6</sup> Cette opération a été rendue possible grâce à l'identifiant commun aux deux bases de données: l'identifiant cantonal.

<sup>7</sup> Rappelons que la Suisse a mobilisé jusqu'à présent pour PISA un échantillon complémentaire d'élèves de 9<sup>e</sup> année. Cet échantillon permet de conduire des analyses au niveau régional et cantonal et enrichit ainsi celles réalisées au niveau international sur l'échantillon suisse des élèves de 15 ans.

nourrir des analyses plus fondamentales. Il ne s'agit donc pas ici de lister l'ensemble de ces analyses mais plutôt d'en présenter la logique en termes de relations à tester. A cet égard, le schéma suivant présente la structure générale des analyses à conduire au niveau statistique (corrélations et analyses de régression principalement). Les flèches symbolisent les relations qui seront testées, les plus importantes étant numérotées.

Schéma 1 : Structures des analyses



Les analyses peuvent être distinguées selon la logique suivante:

- 1) La comparaison de l'influence des caractéristiques individuelles des élèves en fonction de la mesure des résultats concernée: PISA / Evaluations externes (relations 1 et 2). Ces analyses pourraient mettre en évidence des variations dans la capacité des différents types d'évaluation à expliquer l'influence des différences sociodémographiques entre élèves.
- 2) L'analyse des relations entre les évaluations externes (ECR) et les scores obtenus à PISA (relation 3). Cette phase sera décomposée en plusieurs parties pour étudier précisément les relations qui peuvent varier en fonction de la discipline considérée et des différents domaines.
- 3) La comparaison de l'influence du contexte scolaire (voie fréquentée) en fonction de la mesure concernée: PISA / Evaluations externes (relations 4 et 5). Les analyses testeront si cet environnement scolaire a une influence plus ou moins forte en fonction des mesures mobilisées: compétences à PISA ou résultats aux ECR.

## 3.2 LES CARACTÉRISTIQUES DE L'ÉCHANTILLON

Le tableau 1 présente les principales caractéristiques des élèves sur la base du fichier fusionné qui intègre donc les scores aux ECR de français et de mathématiques aux données complètes de l'enquête PISA 2009. Nous avons limité cette description aux

quelques variables individuelles dont on sait qu'elles exercent une influence notable sur les performances des élèves en général, mais aussi, et bien évidemment, sur celles des élèves vaudois (Moreau, 2012). Ce sont uniquement ces variables qui seront prises en compte dans les estimations intégrant les caractéristiques des élèves. L'enquête PISA comporte bien sûr de très nombreuses autres informations sur les élèves, via les questionnaires individuels administrés pendant l'enquête, mais celles-ci ne présentent pas vraiment d'intérêt pour la problématique retenue dans cette recherche<sup>8</sup>.

Une caractéristique individuelle nécessite toutefois d'être retenue pour les analyses, il s'agit de l'environnement familial dans lequel évolue l'élève. Le milieu social et culturel joue en effet un rôle important dans les parcours scolaires et il sera intéressant de voir si cet effet est de la même intensité selon la mesure des performances considérée (PISA ou ECR). L'indicateur disponible dans PISA, et le plus adapté à notre questionnement, est celui nommé *index of economic, social and cultural status* (IESCS/ESCS). Il prend en compte les dimensions suivantes: le statut professionnel le plus élevé des parents, leur niveau de formation également le plus élevé et le patrimoine familial. L'échelle de cet indice correspond à un score centré réduit dont la valeur attribuée à la moyenne de l'OCDE est de 0 (l'écart type étant de 1). Le tableau 2 présente la valeur moyenne et les paramètres de dispersion de l'indicateur socio-économique et culturel sur l'échantillon vaudois. La moyenne relevée sur cet échantillon est supérieure à celle des pays de l'OCDE (11% d'écart type de plus), mais il existe une dispersion notable des valeurs individuelles autour de cette moyenne qui traduit la variété des environnements familiaux des élèves.

Tableau 1 : Principales caractéristiques individuelles des élèves de l'échantillon<sup>9</sup>

	Effectifs	Pourcentages
<i>Sexe</i>		
Garçon	492	49,6
Fille	499	50,4
<i>Langue parlée à la maison</i>		
Français	805	81,2
Autre	186	18,8
<i>Pays de naissance</i>		
Né en Suisse	649	65,5
Né dans un autre pays	325	32,8
Données manquantes	17	1,7
<i>Retard scolaire</i>		
Pas de retard	673	67,9
Retard d'au moins un an	318	32,1
Total	991	100

Tableau 2: Indicateur socio-économique et culturel (IESCS)

	Moyenne	Ecart-type	Minimum	Maximum
IESCS	+0,11	0,95	-2,93	+2,46

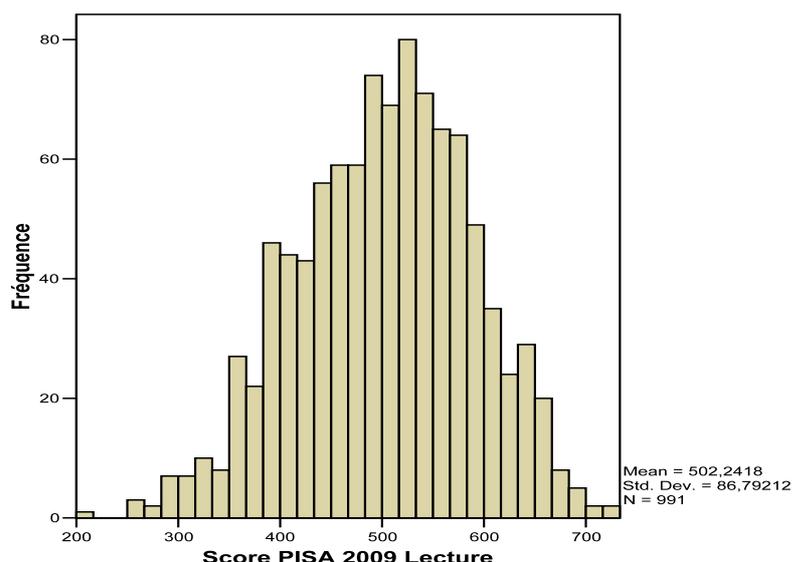
<sup>8</sup> Les élèves répondent à un questionnaire portant sur leur milieu familial, leurs stratégies d'apprentissage et leurs attitudes face à la lecture, leur motivation, etc. Par ailleurs, les directions d'établissement scolaire remplissent un questionnaire qui fournit des informations sur les caractéristiques de l'environnement scolaire de l'élève (ressources et fonctionnement de l'établissement scolaire).

<sup>9</sup> Les effectifs d'élèves de notre échantillon diffèrent légèrement de ceux de l'échantillon vaudois de PISA 2009 (Nidegger et al., 2011) ; la différence est due au fait que pour certains élèves ayant passé PISA, nous ne disposons pas de leurs résultats aux ECR (élèves absents ou ayant déménagé). Néanmoins, la perte d'effectifs n'engendre pas de problème de représentativité au niveau de l'échantillon.

### 3.3 LES RÉSULTATS AUX ÉVALUATIONS (PISA ET ECR)

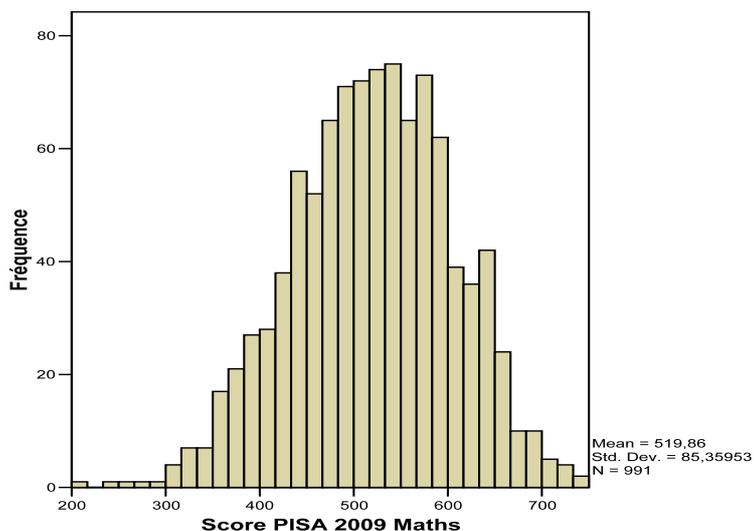
Nous allons dans un premier temps nous intéresser aux caractéristiques des distributions des scores obtenus par les élèves aux différentes épreuves, celles-ci constituant les variables centrales pour nos analyses. Les caractéristiques des notes scolaires seront présentées ultérieurement de manière spécifique. Elles présentent en effet des biais importants sur le plan de la mesure, du fait de leur forte dépendance au contexte scolaire (établissement scolaire, filière, classe et enseignant) et une comparaison directe avec les épreuves externes n'aurait pas vraiment de sens. Les graphiques suivants (graphiques 1, 2 et 3) présentent les scores obtenus aux tests PISA dans les trois domaines de compétences<sup>10</sup>. Les scores moyens reflètent le niveau des élèves vaudois en référence aux performances internationales et cantonales, à savoir une position en lecture équivalente à la moyenne romande et voisine de la moyenne internationale, des résultats jugés corrects en mathématiques mais faibles en sciences et inférieurs à la moyenne romande dans cette dernière dimension (Moreau, 2012). Rappelons que pour PISA 2009, les scores moyens des pays de l'OCDE étaient respectivement de 493, 496 et 501 en compréhension de l'écrit, culture mathématique et culture scientifique (OCDE, 2010).

Graphique 1 : Distribution des scores en lecture à PISA 2009

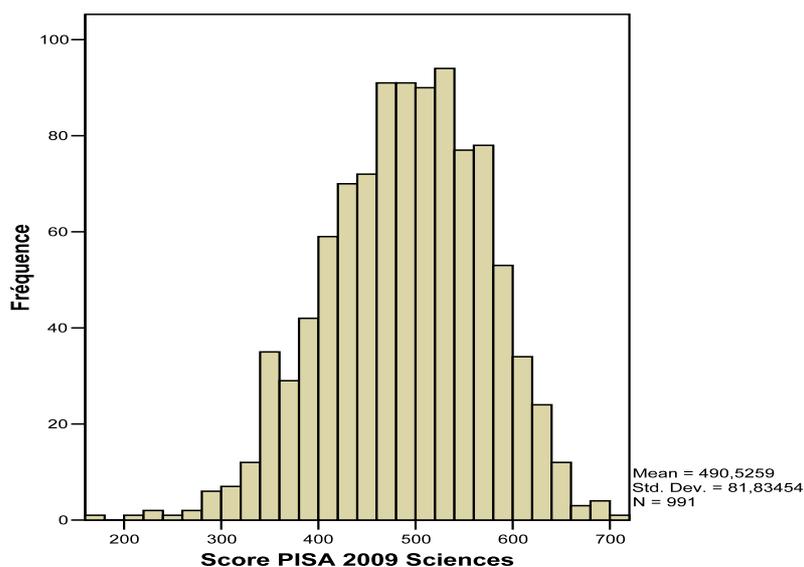


<sup>10</sup> Ces scores ont été estimés en calculant la moyenne des cinq valeurs plausibles attribuées à chaque élève.

Graphique 2 : Distribution des scores en mathématiques à PISA 2009

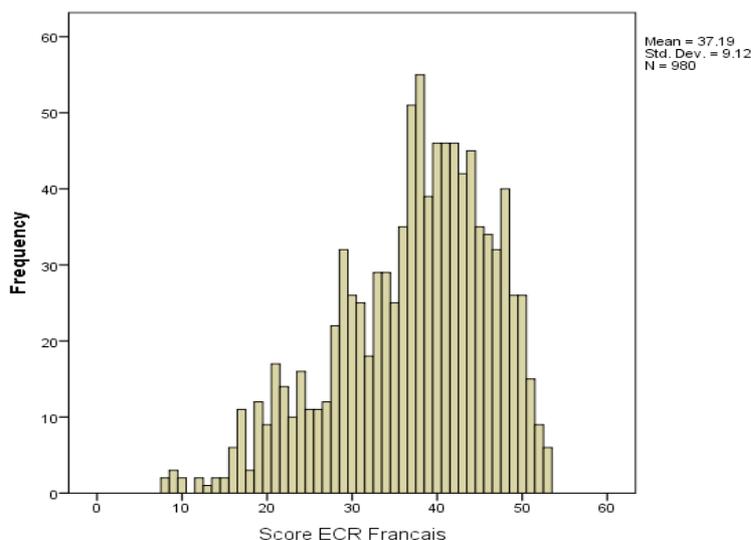


Graphique 3 : Distribution des scores en sciences à PISA 2009

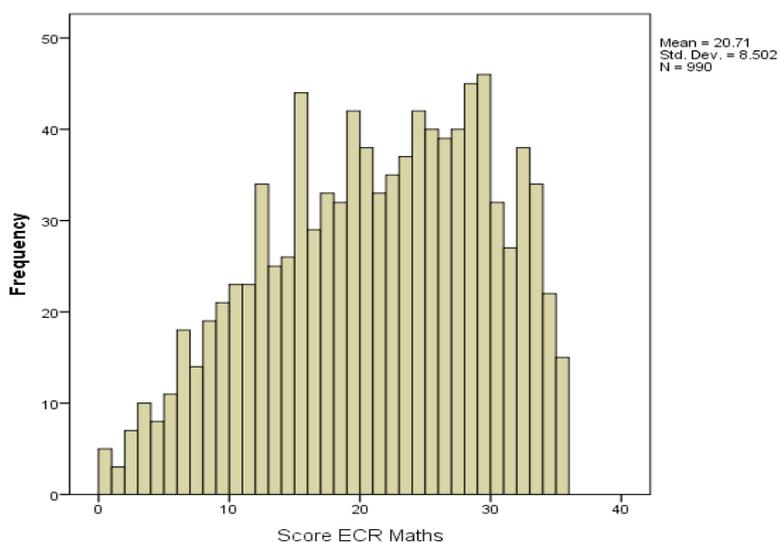


Les distributions des scores aux ECR figurent dans les graphiques 4 et 5. Ces scores sont exprimés dans leur échelle d'origine qui dépend des critères de cotation utilisés dans chacune des épreuves.

Graphique 4 : Distribution des scores de français aux ECR



Graphique 5 : Distribution des scores de mathématiques aux ECR



Si les distributions des scores à PISA sont gaussiennes, il n'en est pas de même pour celles des scores obtenus aux ECR, les valeurs pouvant se répartir de manière dissymétrique autour de la moyenne; cela est particulièrement flagrant pour l'épreuve de mathématiques qui apparaît moins discriminante que celle de français puisqu'une proportion importante d'élèves obtient des résultats élevés. Ce constat n'est pas étonnant dans la mesure où l'épreuve n'est pas conçue selon les mêmes critères psychométriques que ceux appliqués dans les épreuves internationales; l'objectif des ECR est bien de permettre d'estimer le niveau des connaissances et compétences acquises dans le cadre du programme, en référence à la maîtrise des objectifs fondamentaux (Ntamakiro, Moreau, 2011). Les moyennes n'ont donc pas de sens statistique particulier dans le sens où elles ne correspondent pas à des échelles normalisées (Moreau, 1981), mais la variété des valeurs autour de ces moyennes permettront de conduire les analyses prévues selon notre problématique en termes d'analyse de variance notamment. Ce ne se sont pas en effet les

scores en eux-mêmes qui présentent un intérêt pour notre questionnement, mais bien leur variabilité.

Pour fournir au lecteur une image plus parlante de ces distributions, on peut pour l'occasion exprimer les scores en pourcentages de points obtenus avec une échelle qui varie donc dans l'absolu de 0 à 100, cela peut en effet permettre une interprétation plus aisée des résultats des élèves à ces épreuves externes<sup>11</sup> et davantage comparable entre les deux disciplines testées dans les ECR de 8<sup>e</sup> année (tableau 3).

Tableau 3 : Pourcentages de points obtenus aux ECR de 8<sup>e</sup> (2008)

	Mini.	Maxi.	Moyenne	Ecart type
Français	15	100	70,2	17,2
Mathématiques	0	100	59,1	24,3

Il semblerait donc, sur la base des pourcentages moyens de réussite, que les élèves affichent une meilleure maîtrise des compétences dans le domaine du français qu'en mathématiques. Par ailleurs, on peut constater la forte dispersion des résultats dans cette dernière matière, la valeur de l'écart type est en effet très élevée en référence à la moyenne. Le coefficient de variation affiche une valeur de 0,24 en français et de 0,41 en mathématiques<sup>12</sup>.

### 3.4 LES RELATIONS STATISTIQUES INTERNES AUX ÉVALUATIONS

Une première analyse peut être effectuée au sein de chaque type d'évaluation; elle consiste à examiner les relations entre les scores des différentes dimensions évaluées, à savoir les compétences en langue française (français pour les ECR ou lecture pour PISA) et celles en mathématiques (ainsi que les sciences pour PISA). La manière la plus simple de tester ces relations est de calculer des corrélations entre les scores et de les présenter également de manière graphique afin de pouvoir les visualiser.

#### 3.4.1 LES RELATIONS ENTRE LES DOMAINES DE COMPÉTENCES DANS PISA

Le tableau 4 présente les coefficients de corrélations entre les trois domaines de compétences testés par l'enquête PISA<sup>13</sup>, les graphiques 6, 7 et 8 permettent de visualiser ces relations. Les valeurs de ces coefficients (toutes supérieures à +0,80) peuvent surprendre tant elles témoignent d'une forte liaison statistique entre les performances des élèves dans les trois domaines évalués. Les relations entre, d'une part la lecture et les sciences et, d'autre part, les mathématiques et les sciences sont même très élevées puisque les coefficients sont proches de +0,90<sup>14</sup>.

<sup>11</sup> Nous avons calculé ce pourcentage de réussite de la manière la plus simple possible: nombre de points obtenus / nombre de points maximal possible x 100.

<sup>12</sup> Ce coefficient est considéré comme une mesure de dispersion relative qui permet de comparer la variabilité des valeurs de plusieurs distributions, c'est le rapport de l'écart type sur la moyenne.

<sup>13</sup> Par souci de simplicité, nous ne mentionnerons pas toujours les intitulés exacts des domaines testés par PISA (compréhension de l'écrit, culture mathématiques, culture scientifique), il nous arrivera fréquemment de réduire ces intitulés à: lecture, mathématiques et sciences.

<sup>14</sup> Rappelons que dans l'absolu, le coefficient de corrélation peut varier de -1 (pour une corrélation négative parfaite) à +1 (pour une corrélation positive parfaite).

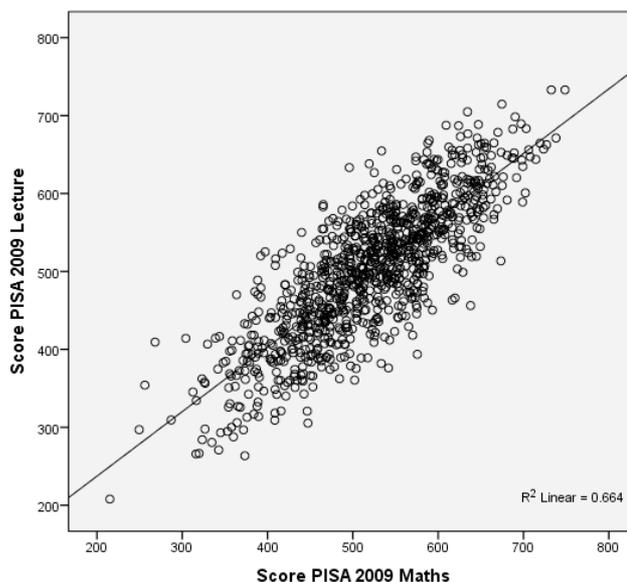
Tableau 4 : Corrélations entre les scores de lecture, mathématiques et sciences (PISA 2009)

	Lecture	Mathématiques
Mathématiques	+ 0,81 ***	
Sciences	+0,88 ***	+ 0,88 ***

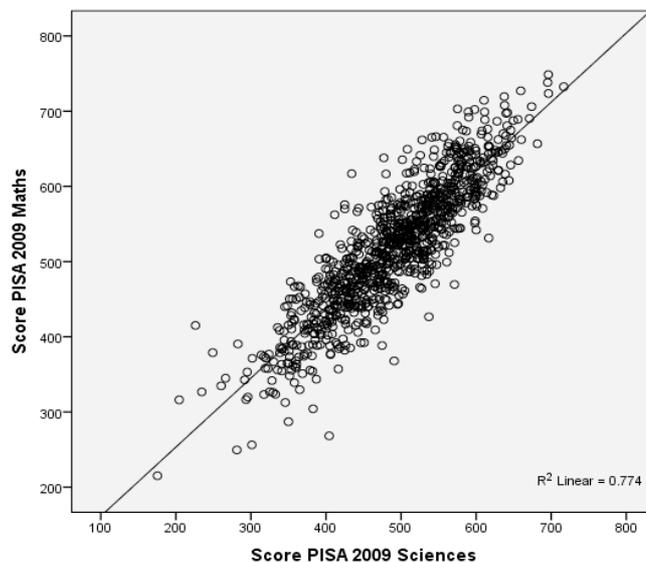
\*\*\* : significatif au seuil de 1%

Cela signifie clairement, qu'en moyenne, un élève qui réussit dans un domaine affiche également de bons résultats dans l'autre. A l'inverse, ce constat est bien évidemment vrai pour les élèves qui obtiennent des performances faibles. On signalera aussi que, quel que soit le niveau des élèves testés, la relation statistique est de nature linéaire, ce qui signifie qu'elle ne varie pas selon le niveau de compétences des élèves. Par ailleurs, on signalera que ce même phénomène d'une forte relation entre les résultats de deux disciplines existe aussi avec la même intensité sur des données agrégées au niveau des pays, à savoir que les pays ont des rangs très proches dans le classement international, quel que soit le domaine de compétences testé (Duru-Bellat, Mons, Suchaut, 2004). Autrement dit, et sur la base de ces constats établis à deux niveaux d'analyse statistique différents (élève et pays), la forte relation entre les scores obtenus dans les différents domaines testés par PISA au niveau individuel est aussi vraie dans les autres contextes nationaux. Si l'on s'interroge sur la signification de ces très fortes corrélations, on peut faire l'hypothèse, sans prendre beaucoup de risque, que la conception des compétences dans l'enquête PISA dépasse les champs disciplinaires. L'approche mobilisée par PISA dans la définition des compétences se révèle donc pertinente sur le plan empirique puisque celles-ci peuvent s'exprimer transversalement, au-delà des domaines évalués, ce qui reflète d'ailleurs bien les intentions de l'OCDE avec les items de PISA qui visent surtout la capacité à exploiter des connaissances, dans l'ensemble du programme d'enseignement (OCDE, 2006).

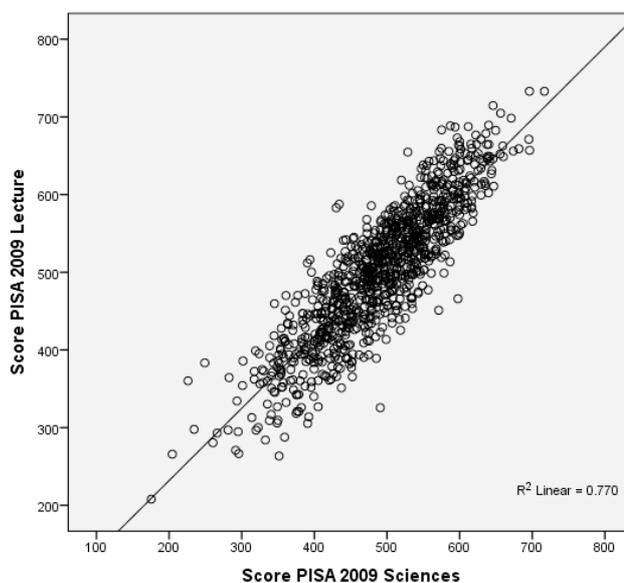
Graphique 6 : Relation entre les scores en mathématiques et en lecture (PISA 2009)



Graphique 7 : Relation entre les scores en mathématiques et en sciences (PISA 2009)



Graphique 8 : Relation entre les scores en lecture et en sciences (PISA 2009)



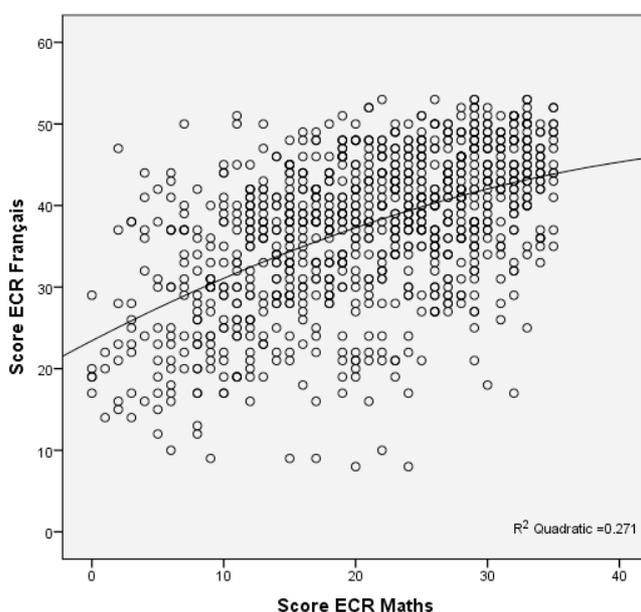
### 3.4.2 LES RELATIONS ENTRE LES DISCIPLINES AUX ECR

La situation est bien différente en ce qui concerne les ECR puisque la corrélation entre les scores de français et de mathématiques est  $+0,52$  ( $P < \alpha .01$ ), soit une corrélation qui peut être qualifiée de moyenne. Ce chiffre témoigne de la distinction plus forte des performances des élèves entre ces deux disciplines quand on se situe à des niveaux élevés du cursus scolaire (Morlaix, Suchaut, 2007). Les profils de réussite peuvent donc varier pour un même élève selon la discipline évaluée dans le cadre des épreuves cantonales de référence. Le graphique 9 met bien en évidence la dispersion des situations pour un nombre important d'élèves. On notera également que la relation est mieux appréhendée (meilleure qualité statistique de l'estimation) avec une spécification quadratique qui rend

compte d'une plus faible relation entre les scores des deux disciplines pour les élèves faibles (le nuage de points étant davantage dispersé dans la partie gauche du graphique). La corrélation est en revanche plus marquée pour les élèves qui affichent des scores élevés dans les deux matières évaluées par les ECR.

Cette première analyse de nature surtout descriptive permet déjà de se familiariser avec les données qui seront exploitées plus en profondeur ultérieurement. Elles laissent imaginer, sur la base des relations statistiques constatées, que PISA et les épreuves cantonales de référence risquent de ne pas rendre compte du même classement entre les élèves. D'une part les fortes corrélations entre les domaines de compétences testés dans PISA et, d'autre part, la relation tout à fait moyenne entre les disciplines évaluées par les ECR suggèrent en effet qu'au moins l'une des deux matières (français ou mathématiques) ne soit pas reliée fortement aux épreuves de PISA.

Graphique 9 : Relation entre les scores en lecture et en mathématiques aux ECR



## 4 LES DÉTERMINANTS DES PERFORMANCES DES ÉLÈVES AUX ÉVALUATIONS

### 4.1 L'INFLUENCE DES CARACTÉRISTIQUES INDIVIDUELLES

Il s'agit ici d'identifier les différences éventuelles liées à l'influence des caractéristiques individuelles des élèves sur leurs performances. Au-delà des seuls liens statistiques entre les évaluations, il est aussi important de voir en quoi les différentes évaluations peuvent s'avérer être plus ou moins discriminantes par rapport à certaines caractéristiques des élèves. Cela peut permettre de se prononcer sur la capacité que peuvent avoir certaines épreuves à marquer, plus que d'autres, les inégalités interindividuelles. Avant de proposer une analyse globale de la variabilité des scores des élèves, il peut être utile de présenter quelques statistiques bivariées qui permettent déjà de dégager certaines tendances. Le tableau suivant présente les scores moyens en fonction des caractéristiques des élèves, les scores aux ECR étant exprimés ici en pourcentages de réussite.

Tableau 5 : Scores moyens à PISA et aux ECR selon les caractéristiques des élèves

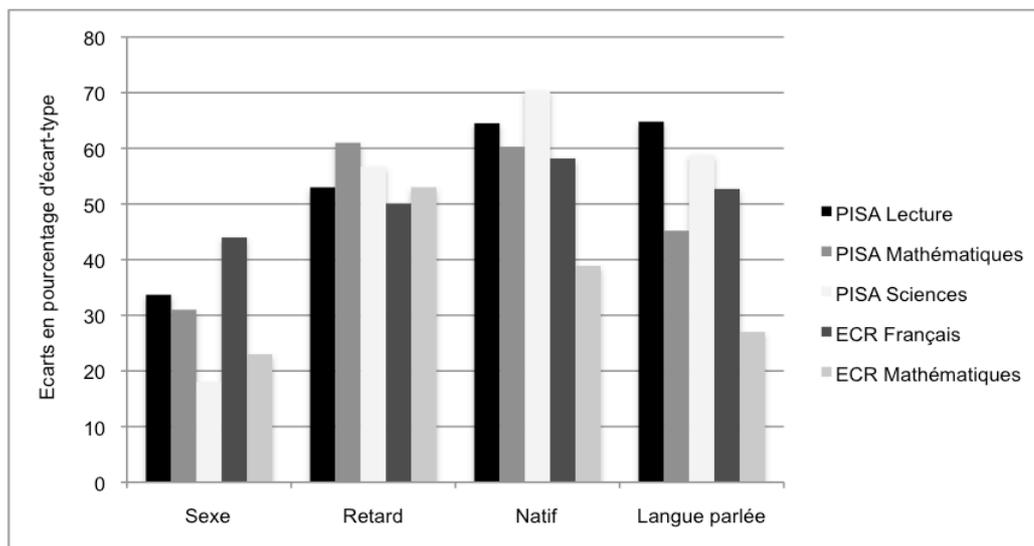
	PISA			ECR	
	Lecture	Mathématiques	Sciences	Français	Mathématiques
Fille	516	507	483	73,9	56,4
Garçon	487	533	497	66,4	62,0
Age normal	516	536	505	73,0	63,3
Retard scolaire	471	484	459	64,2	50,4
Issu de la migration	466	487	453	63,6	53,1
Autochtone	522	538	511	73,6	62,6
Parle français à la maison	512	527	499	71,9	60,4
Parle une autre langue	456	488	451	62,8	53,8

Nous ne commenterons pas en détail ces différences souvent très marquées entre groupes d'élèves, les résultats à l'enquête PISA chez les élèves vaudois ayant notamment déjà fait l'objet d'analyses régulières (Moreau, 2004, 2007, 2013). On peut simplement rappeler les écarts importants liés au retard scolaire, à la langue parlée à la maison et au statut migratoire. On notera que les scores moyens aux ECR suivent la même logique que ceux produits par PISA, à savoir une supériorité des filles en lecture, et des garçons en mathématiques, une performance moindre des élèves ayant redoublé et des résultats supérieurs pour les élèves nés en Suisse et parlant le français à la maison par rapport aux élèves relevant d'une autre situation. Par ailleurs, on retrouve aussi globalement la même structure des résultats en 8<sup>e</sup> année que celle observée dans les ECR en 2<sup>e</sup>, 4<sup>e</sup> et 6<sup>e</sup> année (Ntamakiro, Moreau, 2011).

Le graphique suivant (graphique 10) permet de visualiser les différences de scores entre groupes d'élèves selon les variables retenues, ces différences sont exprimées dans une même échelle, à savoir en pourcentage d'écart type des scores, ce qui permet de pouvoir les comparer plus directement. On remarque que le retard scolaire exerce une influence semblable, quelle que soit l'épreuve considérée, soit entre 50 et 60% d'écart type du score entre les élèves d'âge normal et ceux ayant redoublé au moins une fois au cours de leur scolarité. Il n'est pas aisé de dégager de fortes régularités en termes d'effets des coefficients selon les évaluations (PISA ou ECR), on peut néanmoins constater que les

performances des élèves aux ECR de mathématiques sont nettement moins discriminantes que les autres mesures en ce qui concerne la langue parlée à la maison, le pays de naissance et, dans une moindre mesure le sexe.

Graphique 10 : Différences de scores entre groupes d'élèves (en % d'écart type)



Examinons à présent les différences liées au milieu socio-économique et culturel de l'élève en utilisant l'indice élaboré dans l'enquête PISA<sup>15</sup>. Le tableau 6 présente les corrélations entre cet indice et les scores obtenus aux différentes épreuves. L'influence du milieu familial est identique dans les trois domaines de compétences de PISA et assez marquée (coefficient de +0,44). La situation est différente pour les ECR puisque les corrélations sont nettement moins élevées, particulièrement pour les mathématiques (coefficient de +0,29). Il semblerait donc que les items des ECR soient moins discriminants que ceux de PISA en ce qui concerne l'influence de l'environnement familial des élèves.

Ces relations statistiques peuvent également être présentées graphiquement, tout d'abord en ce qui concerne PISA. Compte tenu de l'équivalence des relations dans les trois domaines, nous présenterons seulement celle relative au score en compréhension de l'écrit (graphique 11). Comme dans les autres dimensions testées dans PISA, le milieu socio-économique de l'élève explique à lui seul 19% de la variance des scores de compréhension de l'écrit (la relation étant par ailleurs de nature linéaire).

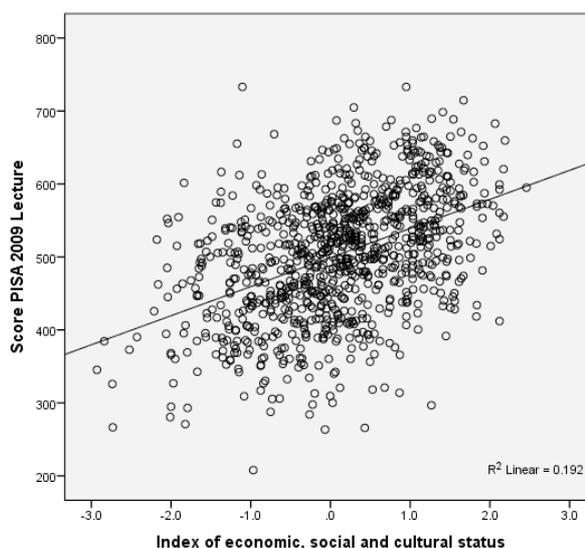
Tableau 6 : Corrélation entre l'indice socio-économique et culturel et les scores aux épreuves

Coefficient de corrélation	PISA			ECR	
	Lecture	Mathématiques	Sciences	Français	Mathématiques
	+0,44 ***	+0,44 ***	+0,44 ***	+0,39 ***	+0,29 ***

\*\*\* : significatif au seuil de 1%

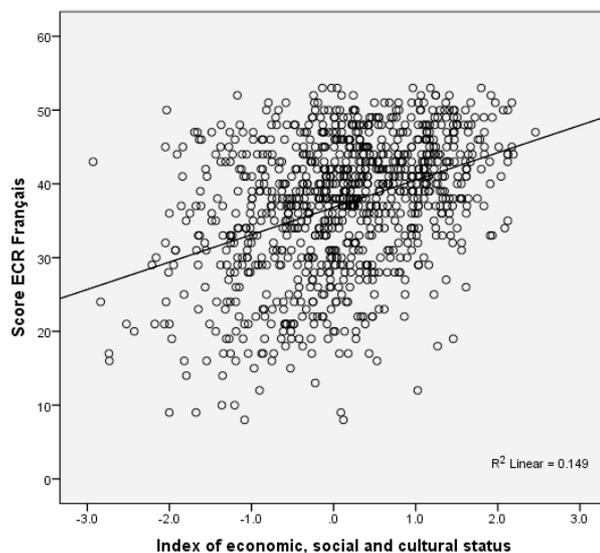
<sup>15</sup> Nous avons préféré conserver cette variable sous sa forme première, à savoir avec une échelle d'intervalle.

Graphique 11 : Relation entre l'indice socio-économique et culturel de l'élève et ses performances en lecture (PISA 2009)

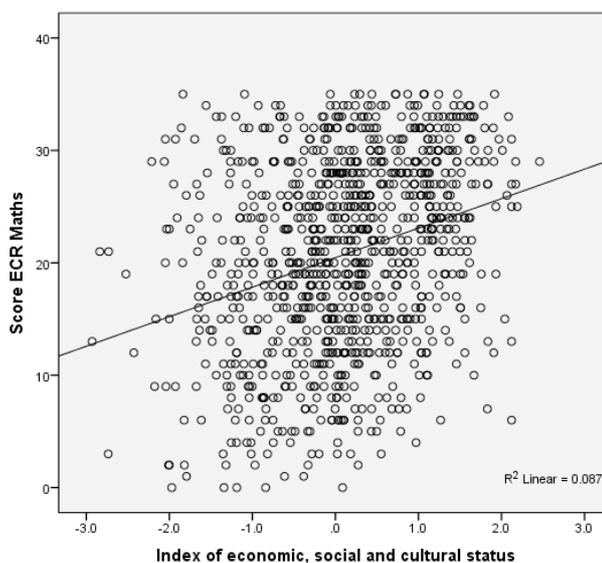


Les graphiques 12 et 13 visualisent quant à eux les relations nettement plus faibles entre le même indice et les scores aux ECR. Le pourcentage de variance expliquée est proche de 15% en français, soit 4 points de moins que celui relevé pour le test PISA dans le même domaine. En mathématiques, on constate visuellement une forte dispersion du nuage de points (graphique 13). L'indice socio-économique et culturel explique en effet moins de 9% de la variance des scores, ce qui est très nettement plus faible que ce qui a été observé avec PISA. Il est aussi possible de raisonner en termes d'impact du milieu social sur les performances des élèves avec les valeurs des coefficients des régressions linéaires expliquant la variance des scores par l'IESCS. Avec PISA, un point de plus à l'IESCS produit une augmentation du score de compréhension de l'écrit de l'ordre de 40 points (les valeurs sont équivalentes pour les mathématiques et les sciences). Ce chiffre renvoie directement à l'importance du milieu socio-économique sur le niveau d'acquisition des élèves. A cet égard, et sur un plan comparatif au niveau de la Suisse romande, les inégalités sociales de compétences sont particulièrement marquées dans le canton de Vaud par rapport aux autres cantons quand on se base sur les données de PISA (Kaiser, Moreau, Nidegger, 2011).

Graphique 12 : Relation entre l'indice socio-économique et culturel de l'élève et ses performances à l'ECR de français



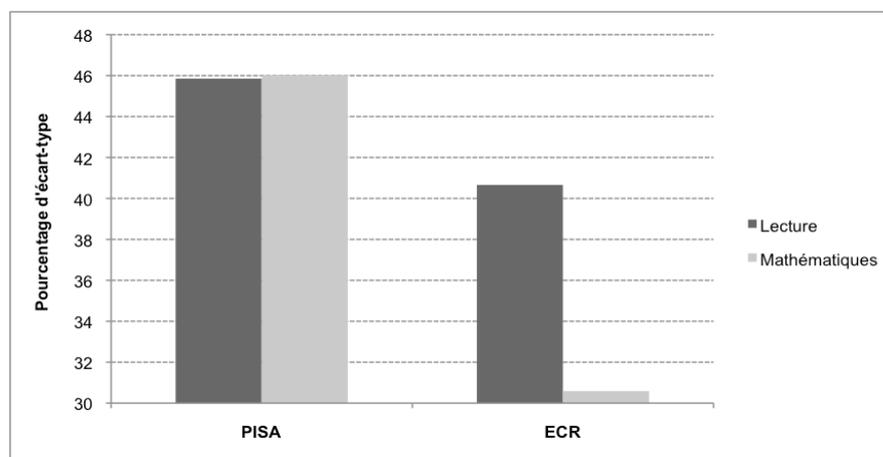
Graphique 13 : Relation entre l'indice socio-économique et culturel de l'élève et ses performances à l'ECR de mathématiques



Pour effectuer une comparaison de cet effet du milieu social sur les performances entre PISA et les ECR, il est préférable de raisonner en termes relatifs en exprimant les coefficients de régression de l'IESCS en pourcentage d'écart type des scores. Cela donne pour PISA, respectivement 45,8%, 46% et 46,4% en lecture, mathématiques et sciences (soit près d'un demi écart type des distributions). Pour les ECR, les effets sont de 40,6% en français et seulement de 30,6% en mathématiques (graphique 14). Les inégalités sociales s'expriment donc différemment selon la mesure des compétences choisie et il apparaît clairement à cet égard que les évaluations basées sur les programmes d'enseignement vaudois sont nettement moins sensibles aux différences socio-économiques de la population d'élèves que celles utilisées dans le programme PISA. Cet

écart dans l'intensité de l'effet de l'origine sociale entre les deux types d'épreuves est particulièrement marqué dans le domaine des mathématiques.

Graphique 14 : Impact du milieu socio-économique sur les performances en lecture et mathématiques (PISA et ECR)



L'influence de l'origine sociale a pour l'instant été mesurée de manière autonome par des régressions univariées et il est utile d'en fournir une image plus précise en prenant en compte les autres caractéristiques des élèves qui sont, elles aussi, associées aux performances scolaires comme nous venons de le voir. Il suffit pour cela d'estimer des modèles de régression multivariés permettant de dégager l'effet net de chacune des variables explicatives des scores individuels, dont celui du milieu socio-économique de l'élève<sup>16</sup>. A côté de l'IESCS, nous intégrerons dans les modèles les variables présentées auparavant dans la description de l'échantillon, à savoir le sexe, le retard scolaire, l'origine migratoire et la langue parlée à la maison. Pour ces variables, nous avons utilisé une spécification dichotomique en opposant pour l'origine migratoire les élèves issus de la migration (dont les deux parents sont nés à l'étranger) aux autochtones (tous les autres élèves). Pour la langue parlée à la maison, nous opposons les élèves qui ont déclaré dans le questionnaire que la langue la plus fréquemment parlée à la maison était la langue du test (soit le français) à ceux qui ont déclaré qu'ils parlaient le plus fréquemment une autre langue. Le tableau 7 présente les résultats des trois modèles de régression qui concernent PISA. Les trois modèles présentent une grande similitude, aussi bien pour leurs pouvoirs explicatifs (les  $R^2$  sont quasiment identiques) que pour la valeur des coefficients. On notera toutefois les différences évoquées précédemment sur les performances différenciées des filles et des garçons selon la dimension évaluée. On remarquera que l'effet du milieu socio-économique sur les scores a diminué (du fait de la prise en compte des autres caractéristiques des élèves) par rapport aux estimations univariées précédentes, il est à présent de l'ordre de 36 à 38% d'écart type selon le domaine de compétences considéré.

<sup>16</sup> Les modèles de régression permettent d'estimer une équation de la forme :  $Y = a_0 + b_1 X_1 + b_2 X_2 + \dots + b_n X_n + \varepsilon$ , Y étant la variable dépendante ou à expliquer,  $X_1, X_2, \dots, X_n$  les variables indépendantes ou explicatives. A chaque variable indépendante est associé un coefficient (b) qui permet de mesurer son effet sur la variable dépendante. Par ailleurs, chaque modèle de régression fournit un indicateur ( $R^2$ ) qui permet de rendre compte de la qualité de ce modèle en termes de pourcentage de variance expliquée (de la variable dépendante) par les variables indépendantes intégrées dans l'équation de régression.

Tableau 7 : Modèles d'analyse de la variance des scores à PISA en fonction des caractéristiques des élèves

Variables		PISA					
		Lecture		Mathématiques		Sciences	
Référence	Active	Coef.	Sign.	Coef.	Sign.	Coef.	Sign.
Garçon	Fille	+27,8	***	-28,6	***	-16,6	***
Age normal	Retard scolaire	-28,7	***	-35,3	***	-29,4	***
Issu de la migration	Autochtone	+20,4	***	+23,4	***	+28,3	***
Parle une autre langue	Parle français à la maison	+11,3	n. s.	-8,6	n. s.	+2,1	n. s.
Indice socio-économique (IESCS)		+31,4	***	+32,5	***	+29,8	***
Constante		+473	***	+536	***	+486	***
R <sup>2</sup>		0,26		0,27		0,26	

n. s. : non significatif , \*\*\* : significatif au seuil de 1%

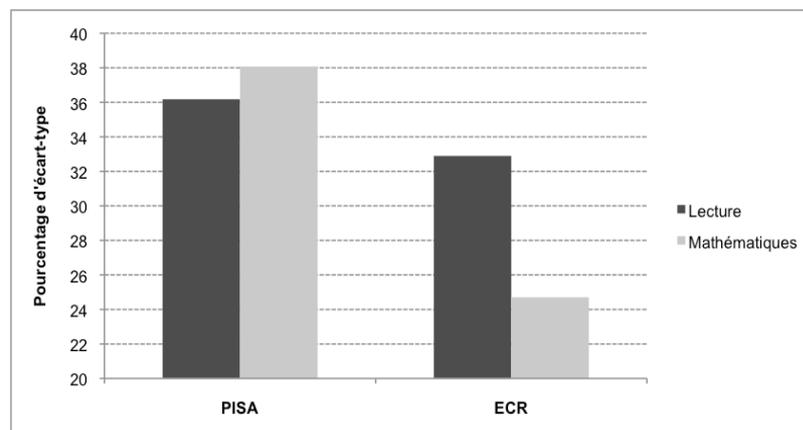
Le tableau 8 expose la réplique de ces analyses multivariées pour les ECR de français et de mathématiques (les scores sont exprimés ici en pourcentages de points).

Tableau 8 : Modèles d'analyse de la variance des scores des ECR en fonction des caractéristiques des élèves

Variables		ECR			
		Français		Mathématiques	
Référence	Active	Coef.	Sign.	Coef.	Sign.
Garçon	Fille	+7,3	***	-6,0	***
Age normal	Retard scolaire	-6,0	***	-9,7	***
Issu de la migration	Autochtone	+3,9	***	+4,4	***
Parle une autre langue	Parle français à la maison	+1,4	n. s.	-1,3	n. s.
Indice socio-économique (IESCS)		+5,6	***	+6,0	***
Constante		+64,1	***	+64,3	***
R <sup>2</sup>		0,24		0,13	

n. s. : non significatif , \*\*\* : significatif au seuil de 1%

Graphique 15 : Effets nets de l'indice socio-économique sur les performances à PISA et aux ECR

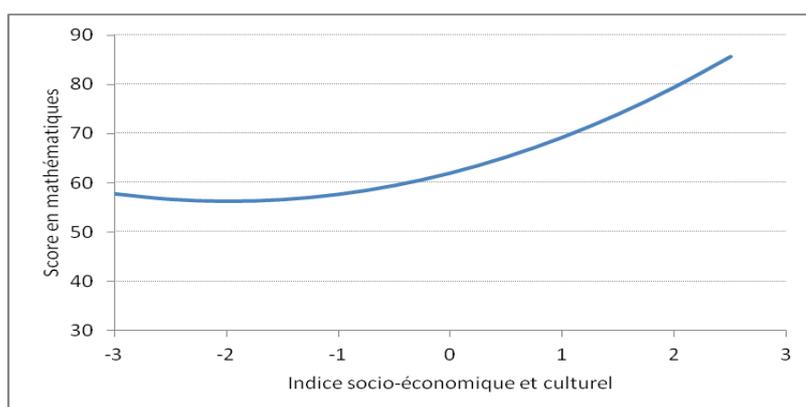


La comparaison entre les cinq modèles des tableaux 7 et 8 montre aussi que la langue parlée à la maison n'est plus une variable discriminante puisqu'aucun des coefficients n'est significatif sur le plan statistique, les différences constatées auparavant lors de

simples comparaisons de moyennes masquaient donc des effets de composition avec les variables intégrées dans les modèles de régression, le pays de naissance notamment. Pour revenir sur l'impact de l'environnement familial de l'élève sur ses performances, on peut avancer l'idée qu'en prenant en compte d'autres caractéristiques (genre, retard scolaire, pays de naissance et langue parlée à la maison), les ECR offrent un contexte d'évaluation des performances plus équitable que PISA et cela est particulièrement significatif dans le domaine des mathématiques.

Des analyses plus poussées de l'effet du milieu socio-économique sur les résultats des élèves dans cette discipline appuient cette affirmation. En effet, il apparaît que l'impact de la variable de l'IESCS n'est pas proportionnel, celui-ci étant plus fort pour les élèves de milieu favorisé et inexistant pour les élèves issus d'un milieu plus modeste. Le graphique 16 permet d'illustrer ce constat, il est établi grâce à une simulation effectuée sur la base du modèle précédent (tableau 8) mais en ajoutant une variable représentant le carré de la variable de l'IESCS (ce qui revient à une spécification quadratique). On remarque l'absence d'effet pour environ la moitié de la population d'élèves (indice inférieur à 0) et un impact de plus en plus fort pour les élèves les plus favorisés socialement<sup>17</sup>.

Graphique 16 : Effet de l'indice de statut socio-économique sur le score en mathématiques (ECR)



## 4.2 L'INFLUENCE DU CONTEXTE SCOLAIRE

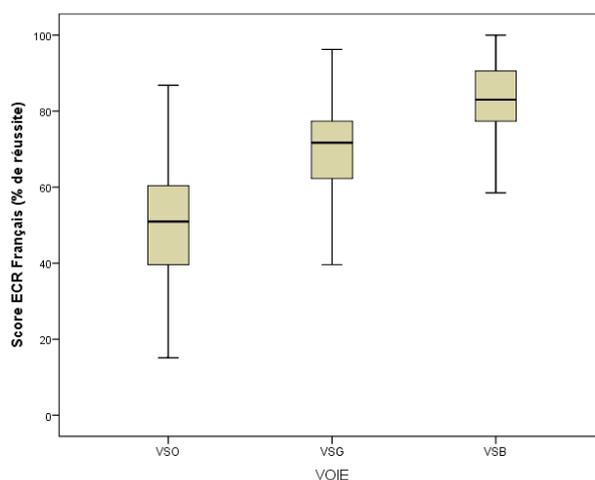
### 4.2.1 CARACTÉRISTIQUES DES ÉLÈVES ET CONTEXTE SCOLAIRE

Jusqu'à présent, aucune variable du contexte scolaire n'a été prise en compte, ni même mentionnée dans les analyses. Il est pourtant nécessaire de considérer aussi certains éléments de l'environnement scolaire de l'élève, ceux-ci pouvant influencer sur les résultats des élèves aux différentes évaluations. Cette dimension n'étant pas centrale dans notre problématique, nous appréhenderons toutefois ce contexte dans sa globalité et non pas sur la base des nombreuses informations disponibles dans la base de données de l'enquête PISA. Compte tenu de la structure de l'échantillon et, plus précisément, de la répartition des élèves dans les filières et dans les établissements scolaires, il n'est pas pertinent de se

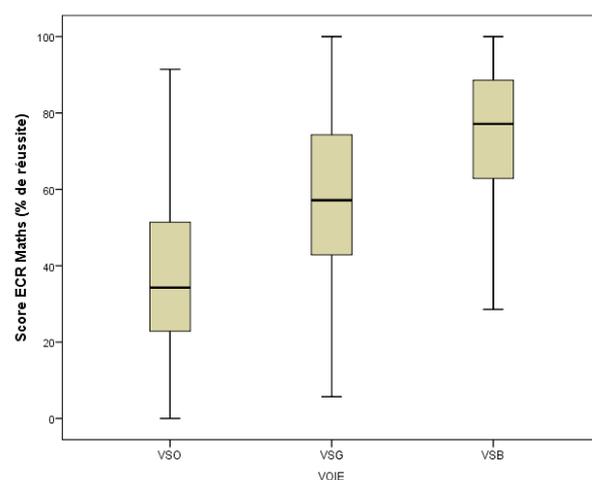
<sup>17</sup> Les scores sont exprimés en pourcentages de réussite au test; ils correspondent à ceux d'un garçon d'âge normal, issu de l'immigration et qui parle chez lui une autre langue que le français (ce qui correspond aux modalités de référence des variables dans le modèle).

focaliser sur le niveau de l'établissement scolaire<sup>18</sup>. En revanche, il paraît intéressant de savoir si la fréquentation des différentes filières du secondaire I peut expliquer la variabilité des résultats selon que l'on se base sur PISA ou sur les épreuves cantonales de référence. L'une des particularités de l'enseignement secondaire vaudois est son organisation en filières conduisant à des parcours scolaires différenciés. Après la 6<sup>e</sup> année, les élèves sont orientés dans l'une des trois filières suivantes: VSB (voie secondaire de baccalauréat), VSG (voie secondaire générale) et VSO (voie secondaire à options)<sup>19</sup>. Environ un tiers des élèves fréquente chacune de ses filières qui sont académiquement très hiérarchisées comme le montrent les graphiques suivants (graphiques 17 et 18) représentant les distributions des scores des élèves aux ECR de français et de mathématiques dans les trois voies (les scores sont exprimés en pourcentages de réussite).

Graphique 17 : Distribution des scores aux ECR de français selon la filière



Graphique 18 : Distribution des scores aux ECR de mathématiques selon la filière

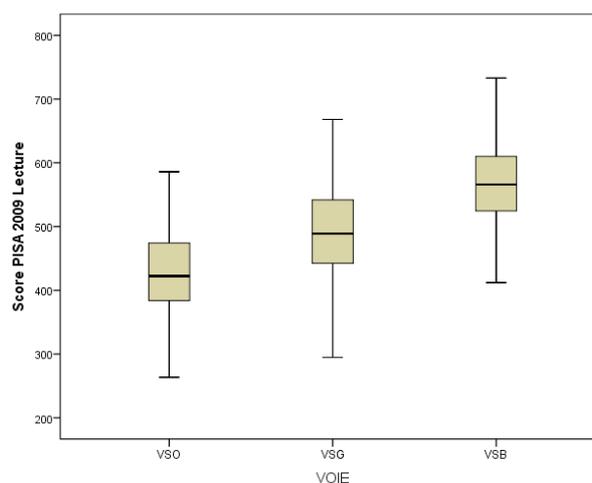


<sup>18</sup> En effet, dans l'échantillon vaudois PISA des élèves de 9<sup>e</sup> année, certaines filières ne sont pas représentées dans certains établissements scolaires.

<sup>19</sup> A partir de la rentrée 2013, l'application de la nouvelle loi sur l'enseignement obligatoire (LEO) conduit à une organisation progressive du secondaire en deux filières: voie pré-gymnasiale (VP) et voie générale (VG).

Les répartitions des performances selon les voies semblent conformes aux objectifs du système puisque les orientations s'effectuent selon les résultats scolaires (sur la base des notes obtenues) avec des degrés d'exigence variés selon les filières (plus grande exigence pour la VSB). La hiérarchisation des performances des élèves est très nette dans le domaine du français pour lequel on constate, outre de fortes différences de moyennes entre les voies, une homogénéité plus forte des résultats dans la voie la plus exigeante. Cette hiérarchie académique est aussi constatée pour les ECR en mathématiques mais avec moins d'intensité. On retrouve également ces performances différenciées des élèves selon les voies avec les scores obtenus à PISA, mais les écarts sont moins forts, comme on peut le voir en compréhension de l'écrit (graphique 19).

Graphique 19 : Distribution des scores à PISA en lecture selon la filière

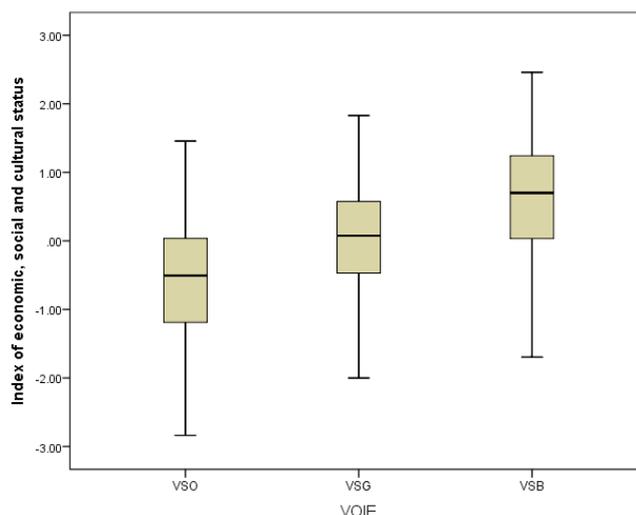


Le constat est par ailleurs identique en ce qui concerne les mathématiques et les sciences. Cela laisse supposer que les compétences telles que PISA les conçoit (donc moins liées aux programmes scolaires) se distribuent plus également entre les élèves fréquentant des filières différentes que les connaissances évaluées sur la base du programme d'enseignement, qui, lui, est différent selon les filières.

#### 4.2.2 L'INFLUENCE DU MILIEU SOCIO-ÉCONOMIQUE DANS LES FILIÈRES SCOLAIRES

Cette hiérarchie académique entre les voies est comparable à celle qui existe aussi sur le plan socio-économique. Il y a en effet une ségrégation sociale importante entre les trois filières comme le montre le graphique suivant (graphique 20). Cette constatation rejoint les observations réalisées pour les précédentes vagues d'enquête PISA (Moreau, 2004, 2007). Entre les deux voies les plus opposées (VSO et VSB), la différence moyenne de l'indice IESCS est de l'ordre d'un écart type, ce qui est considérable. Bien sûr, si l'on considère les filières «voisines» dans la hiérarchie (VSO et VSG d'une part et VSG et VSB d'autre part) le recouvrement des populations est plus fort puisque la différence d'IESCS est d'environ un demi-écart type entre les voies. D'autres caractéristiques des élèves se répartissent inégalement entre les voies. Ainsi, la filière VSO est composée, pour plus de la moitié, d'élèves nés hors de Suisse et, pour un tiers environ, d'élèves ne parlant pas le français à la maison (Moreau, 2013). La situation est bien différente dans les autres voies puisque l'on compte plus de trois-quarts de natifs en VSB, la proportion étant d'environ deux-tiers en VSG.

Graphique 20 : Indice socio-économique et filière fréquentée



Si l'on s'intéresse à présent à l'influence du milieu socio-économique, du fait de cette ségrégation sociale et scolaire présente entre les filières, la situation diffère de ce qui a été constaté précédemment sur l'échantillon complet. Ainsi, pour PISA, l'indice IESCS exerce un effet limité puisqu'il explique une faible proportion de la variance des scores, quelle que soit la filière considérée. Au niveau des ECR, la situation est encore plus marquée puisque l'influence du milieu social est inexistante, sauf en lecture pour les élèves de VSO où l'on relève un faible effet puisque le pourcentage de variance expliquée est de l'ordre de 2,2% ( $R^2 = 0,02$ ).

Tableau 9 : Pourcentage de variance des scores expliquée par l'IESCS par filière

	PISA		ECR	
	Lecture	Mathématiques	Français	Mathématiques
VSO (N = 241)	4,7 %	2,2 %	2,2	0,0 %
VSG (N = 368)	4,0 %	7,8 %	0,0 %	0,0 %
VSB (N = 375)	2,7 %	2,8 %	0,0 %	0,0 %

Cette absence d'influence des caractéristiques socio-économiques des élèves sur leurs performances scolaires (dans des contextes scolaires déjà ségrégués) peut interroger car nous avons quand même mentionné auparavant un certain recouvrement des populations d'élèves du point de vue de leurs caractéristiques sociales pour les filières «voisines»: VSB et VSG d'une part, VSG et VSO d'autre part. Autrement dit, il existe au sein de chaque voie une relative diversité sociale (sur la base de l'indice IESCS), on peut alors faire l'hypothèse que le système est assez égalitaire en termes de traitement des élèves à l'intérieur des filières. Plus largement, on pourrait s'interroger aussi sur l'influence que peut avoir la représentation des enseignants, du degré de mixité sociale de leur classe; en étant conscients de la ségrégation sociale liée aux filières, les enseignants n'auraient-ils pas des pratiques plus égalisatrices dans un contexte qu'ils imaginaient unifié alors qu'une diversité existait pourtant ?

### 4.3 ANALYSE GLOBALE DES DÉTERMINANTS DES PERFORMANCES

Pour terminer cette section sur l'influence différenciée des facteurs influençant les performances des élèves selon le type d'évaluation considérée, nous intégrerons dans une même analyse statistique les facteurs individuels et contextuels. Le tableau 10 présente pour PISA les estimations des modèles intégrant les caractéristiques des élèves et la filière fréquentée. Les pourcentages de variance expliquée augmentent nettement par rapport aux estimations précédentes qui ne prenaient en compte que les caractéristiques individuelles. C'est la moitié de la variance des scores obtenus à PISA qui est à présent expliquée ( $R^2$  entre 0,47 et 0,50). La hiérarchisation académique et sociale des filières est très visible dans les modèles, puisque de forts écarts de scores moyens apparaissent entre les trois voies, qui se chiffrent à plus d'un 1,3 écart type de la distribution en considérant les voies les plus opposées (VSB versus VSO). Bien évidemment, la prise en compte des filières dans la modélisation multivariée a pour conséquence de diminuer l'influence spécifique des caractéristiques des élèves. On observe notamment une baisse de l'impact de l'origine sociale sur les résultats aux tests.

Tableau 10 : Modèles d'analyse de la variance des scores obtenus à PISA en fonction des caractéristiques des élèves et de la filière fréquentée

Variables		PISA					
		Lecture		Mathématiques		Sciences	
Référence	Active	Coef.	Sign.	Coef.	Sign.	Coef.	Sign.
Garçon	Fille	+21,8	***	-34,6	***	-22,3	***
Age normal	Retard scolaire	-14,7	***	-21,2	***	-16,0	***
Issu de la migration	Autochtone	+9,7	*	+12,5	***	+18,0	***
Parle une autre langue	Parle français à la maison	+15,0	n. s.	-5,1	n. s.	+5,7	n. s.
Indice socio-économique (IESCS)		+10,9	***	+11,9	***	+10,1	***
VSO	VSG	+48,9	***	+51,3	***	+47,3	***
	VSB	+116,2	***	+117,4	***	+111,5	***
Constante		+415	***	+477	***	+430	***
$R^2$		0,47		0,50		0,49	

n. s. : non significatif, \* : significatif au seuil de 10%, \*\*\* : significatif au seuil de 1%

Les mêmes analyses ont été réalisées sur les scores aux ECR (tableau 11). Celles-ci produisent des résultats comparables à ceux observés avec PISA, en ce qui concerne la plus faible influence des variables individuelles; ainsi, seuls le retard scolaire et le sexe sont des variables significatives. Le lieu de naissance, la langue parlée à la maison et le milieu socio-économique, quant à eux, n'exercent plus d'effets significatifs sur les performances en français et en mathématiques. L'essentiel des différences de scores entre élèves est dû à la fréquentation de telle ou telle filière.

Tableau 11 : Modèles d'analyse de la variance des scores obtenus aux ECR en fonction des caractéristiques des élèves et de la filière fréquentée

Variables		ECR			
		Français		Mathématiques	
Référence	Active	Coef.	Sign.	Coef.	Sign.
Garçon	Fille	+5,8	***	-7,9	***
Age normal	Retard scolaire	-2,6	***	-5,3	***
Issu de la migration	Autochtone	+1,0	n. s.	+0,82	n. s.
Parle une autre langue	Parle français à la maison	+1,8	n. s.	-2,2	n. s.
Indice socio-économique (IESCS)		+0,6	n. s.	+0,26	n. s.
VSO	VSG	+18,0	***	+20,7	***
	VSB	+29,7	***	+37,1	***
Constante		+47,7	***	+44,6	***
R <sup>2</sup>		0,57		0,40	

n. s. : non significatif, \*\*\* : significatif au seuil de 1%

Les modèles présentés ne sont toutefois pas totalement satisfaisants sur le plan statistique dans la mesure où ils mobilisent des régressions estimées au niveau des élèves, alors que les variables contextuelles demanderaient à être appréhendées à un niveau supérieur<sup>20</sup>. Le faible nombre d'observations à ce niveau (il n'y a que trois filières) rend cette analyse impossible. Nous avons évoqué précédemment une autre difficulté liée à la nature des données, qui est de ne pas pouvoir mobiliser le niveau de l'établissement scolaire dans des modèles multiniveaux. En effet, les trois filières n'étant pas toutes présentes dans chaque établissement scolaire (tableau 12), il serait impossible de distinguer l'effet filière de l'effet établissement, dans les analyses.

<sup>20</sup> Le fait que les élèves fréquentant une même filière ont des caractéristiques sociales et scolaires beaucoup plus proches que les élèves fréquentant des filières différentes ne peut en effet pas être totalement pris en compte avec nos analyses effectuées au seul niveau «élève».

Tableau 12 : Répartition du nombre d'élèves de l'échantillon par voie et par établissement

	Voie			Total
	VSO	VSG	VS	
Etablissement 1	7	13	24	44
Etablissement 2	15	20	16	51
Etablissement 3	5	19	18	42
Etablissement 4	13	14	21	48
Etablissement 5	6	2	43	51
Etablissement 6	4	12	36	52
Etablissement 7	7	13	13	33
Etablissement 8	6	11	10	27
Etablissement 9	9	12	26	47
Etablissement 10	7	19	17	43
Etablissement 11	6	16	15	37
Etablissement 12	5	24	21	50
Etablissement 13	11	12	23	46
Etablissement 14	17	12	11	40
Etablissement 15	12	13	18	43
Etablissement 16	6	9	29	44
Etablissement 17	6	15	13	34
Etablissement 18	11	11	20	42
Etablissement 19	7	14	1	22
Etablissement 20	15	24	0	39
Etablissement 21	8	18	0	26
Etablissement 22	15	16	0	31
Etablissement 23	26	13	0	39
Etablissement 24	13	17	0	30
Etablissement 25	11	19	0	30
<b>Total</b>	<b>248</b>	<b>368</b>	<b>375</b>	<b>991</b>

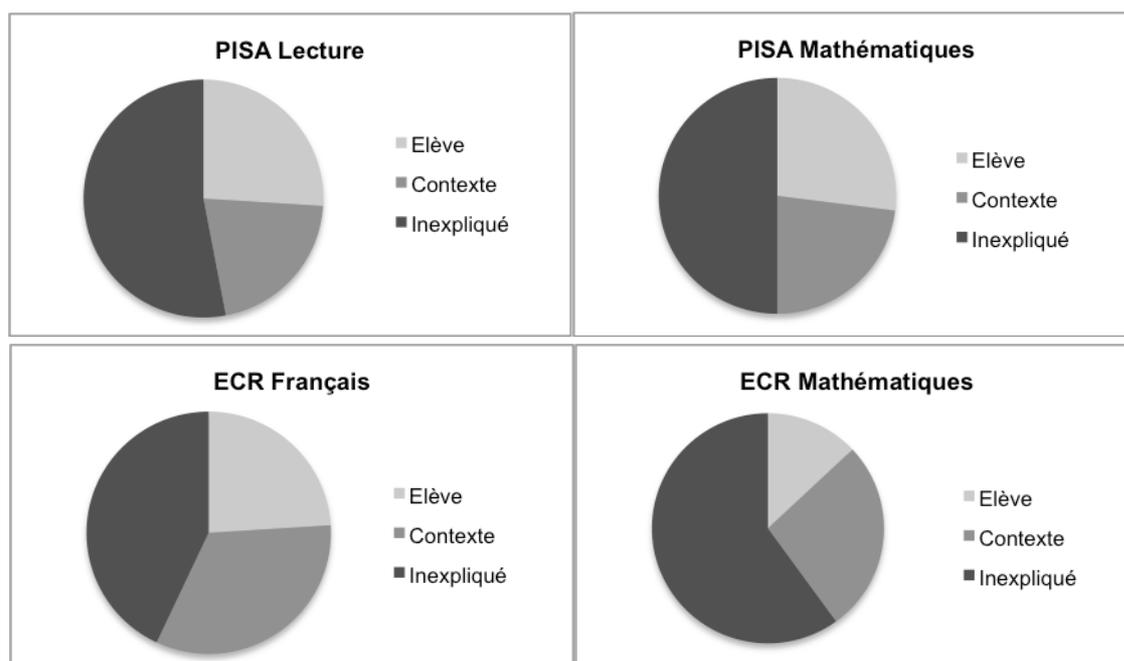
Cette imperfection des estimations n'est toutefois pas un problème majeur, dans le sens où la mesure précise des effets contextuels sur les résultats des élèves n'est pas un point central pour notre questionnement. Nous avons surtout comme objectif de mettre en évidence des différences entre les épreuves d'évaluation quant à l'influence des caractéristiques des élèves. Le tableau 13 et les diagrammes en secteurs qui suivent (graphique 20) donnent une image globale des différences entre PISA et les ECR en ce qui concerne l'influence des caractéristiques individuelles des élèves, et de leur contexte scolaire sur les performances.

Tableau 13 : Décomposition de la variance des scores des évaluations (PISA et ECR)

	PISA		ECR	
	Lecture	Mathématiques	Français	Mathématiques
Elève	26%	27%	24%	13%
Contexte	21%	23%	33%	27%
Inexpliqué	53%	50%	43%	60%

L'ECR de mathématiques se démarque des autres épreuves avec, d'une part une faible proportion de variance du score expliquée par les caractéristiques des élèves (seulement 13%) et, d'autre part, un pourcentage de variance inexpliquée importante (60%). L'ECR de français se caractérise quant à elle par un pouvoir explicatif élevé des caractéristiques de l'élève, une faible part de variance inexpliquée et un poids important donné au contexte d'enseignement (filières) dans l'explication des différences de performances entre élèves. On constate à nouveau la similitude de la structure des estimations entre les deux dimensions testées dans PISA (la situation étant identique pour les sciences).

Graphique 21 : Décomposition de la variance des scores aux évaluations (PISA et ECR)



Au final, les différentes analyses exposées dans cette partie mettent en évidence des différences assez marquées entre PISA et les ECR quant à leur capacité à rendre compte des caractéristiques des élèves avec leur contexte de scolarisation. C'est dans doute au niveau de l'influence du milieu social que la distinction entre les deux formes d'évaluation est la plus nette. Les épreuves cantonales de référence mettent moins à jour les inégalités dues à l'environnement familial de l'élève que PISA; ce constat d'ailleurs est renforcé quant on considère plus particulièrement les mathématiques. Ce résultat sur l'influence différenciée des facteurs socio-économiques est important dans la mesure où c'est le caractère plus ou moins inégalitaire d'un système scolaire qui peut être apprécié avec une intensité variable selon la mesure sur laquelle on se base pour l'évaluation des compétences et connaissances des élèves. Les compétences évaluées à travers PISA marquent donc davantage les inégalités interindividuelles que les évaluations réalisées sur

la base des programmes, ces dernières se révélant moins sensibles aux différences interindividuelles. L'évaluation de matières auxquelles les élèves ont été exposés via les programmes scolaires (ECR) est moins sensible à la culture d'origine. A l'inverse, l'évaluation des compétences dont l'acquisition provient de diverses sources, dont extrascolaire et familiale (PISA) est plus sensible à cette culture d'origine. Ce point est à souligner car il fournit une autre vision du système éducatif vaudois que celle véhiculée par PISA.

La partie suivante va concerner le cœur de la problématique, à savoir la relation entre les résultats des élèves aux deux formes d'épreuves, nous limiterons l'analyse à la lecture et aux mathématiques, les deux disciplines évaluées à la fois par PISA et les ECR, sachant que les sciences, au regard des liens très forts entretenus au sein de PISA avec les deux autres disciplines n'apparaissent pas utiles à analyser.



## 5 LES RELATIONS ENTRE LES RÉSULTATS À PISA ET LES RESULTATS AUX ECR

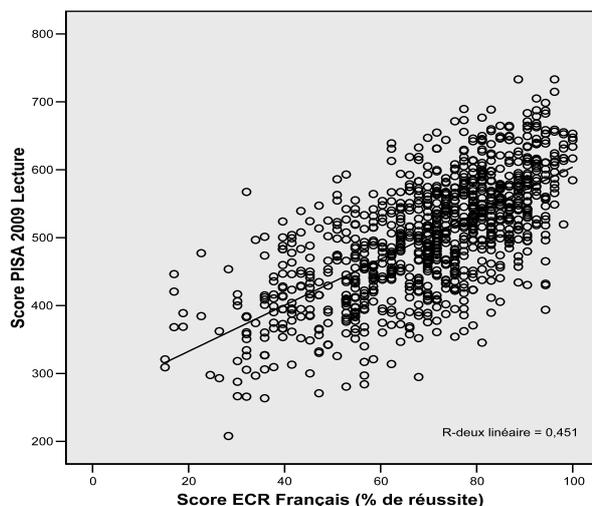
### 5.1 UNE ANALYSE GLOBALE DES RELATIONS ENTRE LES SCORES AUX ÉVALUATIONS

#### 5.1.1 LES RELATIONS STATISTIQUES ENTRE LES SCORES GLOBAUX

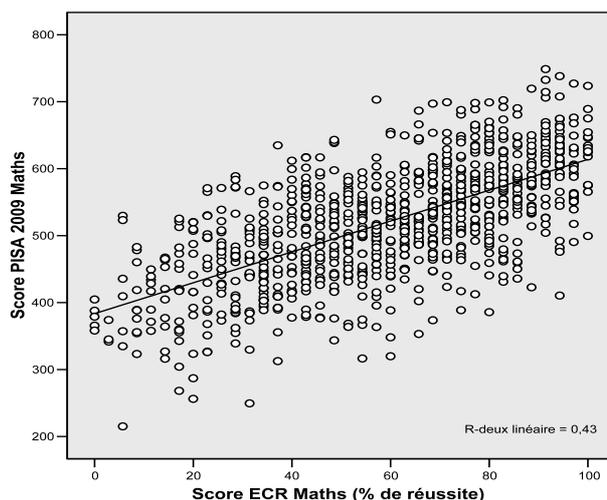
La démarche qui découle de notre raisonnement consiste dans un premier temps à observer la relation entre les scores d'un même champ disciplinaire produits par les deux évaluations externes qui sont de nature et de finalité différentes (PISA et ECR). Comme nous l'avons fait précédemment pour les relations entre les disciplines d'une même évaluation, nous avons calculé des corrélations pour traduire l'intensité de ces relations. La corrélation entre les scores de lecture est de +0,67 ( $P < .01$ ) et celle entre les scores de mathématiques est de +0,65 ( $P < .01$ ), soit des chiffres très voisins qui témoignent donc du même degré de liaison entre PISA et les ECR pour les deux disciplines. Les valeurs des coefficients peuvent donner lieu à une double interprétation. En premier lieu, il est évident que les classements des élèves dans les deux types d'épreuves sont bien liés et que, en moyenne, ceux qui obtiennent de bons résultats à PISA réussissent également bien les ECR. Mais, et en second lieu, les relations ne sont pas d'une forte intensité et on observe, pour de nombreux élèves, un degré de réussite différent selon l'évaluation considérée. Rappelons à cet égard que les corrélations entre les scores obtenus à PISA dans les trois domaines (compréhension de l'écrit, culture mathématique et culture scientifique) étaient bien plus élevées (de +0,81 à +0,88) que celles constatées pour un même domaine de discipline entre PISA et les ECR. On peut alors affirmer que les deux évaluations, pour chaque discipline évaluée, ne fournissent pas les mêmes informations sur les performances des élèves.

Les graphiques suivants (graphiques 22 et 23) visualisent les relations et permettent des commentaires plus précis. Les diagrammes de dispersion ont fait l'objet d'un choix en ce qui concerne les variables à sélectionner pour chacun des axes. Nous avons opté pour faire figurer sur l'axe des ordonnées les scores à PISA pour deux raisons; la première est que les ECR ont été administrées aux élèves avant les tests PISA et il est donc logique que la variable dépendante (le phénomène que l'on cherche à expliquer) corresponde à un événement qui s'est déroulé à un moment plus récent que celui lié à la variable indépendante ou explicative (axe des abscisses). La seconde raison tient au fait que les tests PISA évaluent des connaissances et des aptitudes utiles pour la vie future alors que les ECR mesurent davantage l'assimilation du programme d'enseignement à un instant de la scolarité, en l'occurrence ici en fin de 8<sup>e</sup> année. Sur la base des régressions univariées expliquant la variance des scores de PISA par les résultats aux ECR, on constate des coefficients de détermination ( $R^2$ ) de 0,45 pour la lecture et de 0,43 en mathématiques. Les scores obtenus aux ECR n'expliquent donc que deux cinquièmes des différences des performances à PISA, la part de variance non expliquée (de l'ordre de 60%) nous autorise à considérer les deux types d'épreuves comme porteuses de caractéristiques spécifiques dans la mesure du niveau d'acquisition des élèves.

Graphique 22 : Relation entre les scores en français (ECR) et de lecture (PISA 2009)



Graphique 23 : Relation entre les scores de mathématiques (ECR et PISA 2009)



On peut examiner plus en détail les similitudes et les différences dans la répartition des élèves selon leurs scores relatifs dans les deux évaluations. Nous avons opté pour examiner cette répartition en fonction des quartiles des distributions. Le tableau 14 présente la valeur des trois quartiles pour PISA et les ECR (sur la base du score exprimé en pourcentage de réussite).

Tableau 14 : Valeurs des quartiles des distributions des scores de PISA et des ECR

	PISA		ECR	
	Lecture	Mathématiques	Français	Mathématiques
Q1 (premier quartile)	441	463	58,5	42,1
Q2 (médiane)	506	523	71,7	60,0
Q3 (troisième quartile)	564	579	83,0	80,0

Les tableaux suivants (tableaux 15 et 16) croisent les pourcentages d'élèves se situant par rapport aux quartiles des deux distributions (PISA et ECR) dans chacune des disciplines. La comparaison entre les distributions des scores est bien de nature relative puisque l'on s'intéresse à la position des élèves par rapport à des paramètres statistiques classiques (en l'occurrence ici les quartiles). C'est donc ici une logique normative que l'on mobilise (par rapport à une norme statistique) et non pas une logique critériée établie sur la base d'un niveau de score correspondant à tel ou tel niveau de maîtrise des compétences. Les chiffres en gras renvoient à des classements identiques à PISA et aux ECR. Par exemple, en lecture, 14,8% des élèves obtiennent des scores inférieurs au premier quartile (Q1) de chacune des deux distributions. Pratiquement le même pourcentage est observé pour les meilleurs élèves, à savoir ceux qui affichent des scores supérieurs au troisième quartile, à la fois pour PISA et les ECR. Mais ce qui retiendra notre attention, ce sont surtout les cas de discordance dans les répartitions. Ainsi, en lecture, un dixième des élèves (9,9%) ont un score inférieur au premier quartile de la distribution de PISA et, en même temps, un score supérieur au premier quartile de la distribution de l'ECR (soit 5,6% + 3,3% + 1,0%). On relève aussi 12,8% d'élèves qui obtiennent un score supérieur au troisième quartile de l'ECR alors qu'ils ont des performances inférieures au troisième quartile de l'évaluation PISA (soit 1,0% + 3,5% + 8,3%). En regroupant les données, il y a au final près un quart des élèves (24,1%) qui affichent un score relatif plus élevé à PISA qu'à l'ECR et 29% des élèves qui ont un score relatif plus élevé aux ECR qu'à PISA. La situation est comparable pour les mathématiques et, toujours en termes de scores relatifs par rapport aux quartiles des distributions, on constate que plus d'un quart des élèves (26,4%) ont un score à PISA supérieur à celui de l'ECR et que 28,7% sont dans une situation inverse.

Tableau 15 : Répartition des élèves en fonction des quartiles des distributions des scores en lecture à PISA et au ECR (en %)

		PISA lecture			
		<Q1	Q1-Q2	Q2-Q3	>Q3
ECR Français	<Q1	<b>14,8</b>	5,7	2,2	0,3
	Q1-Q2	5,6	<b>8,3</b>	5,2	2,6
	Q2-Q3	3,3	7,3	<b>9,3</b>	8,1
	>Q3	1,0	3,5	8,3	<b>14,6</b>

Tableau 16 : Répartition des élèves en fonction des quartiles des distributions des scores en mathématiques à PISA et au ECR (en %)

		PISA Mathématiques			
		<Q1	Q1-Q2	Q2-Q3	>Q3
ECR Maths	<Q1	<b>14,4</b>	7,0	2,9	0,6
	Q1-Q2	6,7	<b>7,1</b>	6,3	2,0
	Q2-Q3	2,5	8,2	<b>8,6</b>	7,6
	>Q3	1,2	2,8	7,3	<b>14,8</b>

Ces chiffres, qui illustrent les corrélations relevées précédemment entre PISA et les ECR dans les deux disciplines, montrent bien qu'une proportion non négligeable d'élèves se classe différemment selon le type d'évaluation considérée. On peut explorer davantage cette relation en mobilisant, non plus le score global à PISA, mais les critères de performances définis selon des seuils définis par l'échelle de mesure.

### 5.1.2 LES RELATIONS EN FONCTION DES NIVEAUX DE COMPÉTENCES À PISA

Le tableau suivant présente, pour la compréhension de l'écrit, les différents niveaux de compétences et les scores qui y sont associés.

Tableau 17 : Niveaux de compétences en compréhension de l'écrit à PISA 2009

<b>Définition des niveaux de compétences de PISA 2009</b>
<b>Niveau de compétence 6 : plus de 698 points</b> Être capable de comprendre de façon fine et détaillée un ou plusieurs documents, d'émettre des hypothèses d'analyse ou de juger d'un regard critique un texte complexe portant sur un thème peu familier.
<b>Niveau de compétence 5 : de 625 à 697 points</b> Être capable d'évaluer l'information et d'élaborer des hypothèses, en faisant appel à des connaissances spécialisées, en développant des concepts contraires aux attentes.
<b>Niveau de compétence 4 : de 552 à 624 points</b> Être capable de réussir des tâches de lecture complexes comme retrouver des informations enchevêtrées, interpréter le sens à partir de nuances de la langue et évaluer de manière critique un texte.
<b>Niveau de compétence 3 : de 480 à 551 points</b> Être capable de réussir des tâches de lecture de complexité modérée, telles que repérer plusieurs éléments d'information et les relier avec des connaissances familières et quotidiennes.
<b>Niveau de compétence 2 : de 407 à 479 points</b> Être capable d'effectuer des tâches de base en lecture telles que retrouver des informations linéaires et en dégager le sens en se référant à des connaissances extratextuelles.
<b>Niveau de compétence 1a : de 335 à 406 points</b> Être capable de reconnaître les thèmes principaux et l'objectif d'un texte portant sur un sujet familier.
<b>Niveau de compétence 1b : de 262 à 334 points</b> Être capable de repérer un élément d'information explicitement formulé dans un texte court et facile portant sur un sujet familier et de faire des connexions simples.

Source : OCDE, 2010

Six niveaux ont été définis par les concepteurs des épreuves, allant du niveau de compréhension le plus simple aux tâches les plus complexes traduisant un excellent niveau de compréhension fine de l'écrit. On notera que le niveau 1 est subdivisé en deux : 1b (niveau le plus faible) et 1a; le niveau minimal permettant de participer à la vie courante est le niveau 2. Selon la méthodologie utilisée dans l'enquête, les élèves atteignant un niveau donné sont censés être capables de résoudre toutes les tâches des niveaux inférieurs et au moins 50% des tâches correspondant au niveau atteint.

La répartition des élèves de l'échantillon vaudois selon ces niveaux de compétences est présentée dans le tableau 18. Ce tableau met aussi en correspondance cette répartition avec celle valant pour l'ensemble des élèves des pays de l'OCDE testés dans PISA. On constate une grande similitude entre la répartition vaudoise et celle des pays de l'OCDE.

Par exemple, dans les deux échantillons, les niveaux élevés de compétences ne sont atteints que par une faible proportion d'élèves (moins de 1% d'élèves atteignent le niveau 6).

Tableau 18 : Répartition des élèves de l'échantillon et de l'OCDE selon le niveau de compétences en lecture à PISA 2009 (en %)

Niveaux	Echantillon	OCDE
Niveau 6	0,5	0,8
Niveau 5	7,3	6,8
Niveau 4	22,2	20,7
Niveau 3	31,9	28,9
Niveau 2	23,4	24,0
Niveau 1a	11,5	13,1
Niveau 1b	3,1	4,6
Inférieur niveau 1b	0,1	1,1

Si l'on considère le seuil nécessaire pour qu'un élève soit considéré comme un bon lecteur (selon les éléments composant les critères de performance), on peut retenir le niveau 4, soit un score minimal de 552. A ce niveau, l'élève est capable de réussir des tâches de lecture complexes, d'interpréter le sens à partir de nuances de la langue et d'évaluer de manière critique un texte. Dans l'échantillon vaudois de 9<sup>e</sup> année, 30% des élèves atteignent ce seuil (et une proportion presque équivalente pour l'échantillon de l'OCDE) et peuvent donc, selon les critères de PISA, être considérés comme de bons lecteurs, voire d'excellents pour ceux qui atteignent les niveaux supérieurs (niveaux 5 et 6). La question est alors de savoir quel niveau de performances à l'ECR en français les élèves doivent atteindre pour appartenir à cette catégorie de bons lecteurs selon le programme PISA. Pour répondre à la question, une possibilité est de raisonner en termes de probabilités d'atteindre ce niveau 4, en estimant un modèle logistique utilisant, comme variable dépendante, une variable dichotomique qui vaut 1 pour les élèves ayant un score supérieur ou égal à 552 et 0 pour les autres élèves; la variable explicative étant le score obtenu à l'ECR de français, exprimé en pourcentage de réussite. Ce modèle logistique permettra alors de prédire la probabilité la plus élevée d'atteindre le niveau 4 de compétences à PISA en fonction du niveau de performance à l'ECR. Concrètement, cette probabilité se traduit par l'expression:

$$P(Y) = 1 / (1 + e^{-g})$$

Dans l'équation, «e» correspond au logarithme népérien et «g» correspond à l'équation de régression linéaire conventionnelle, soit:  $Y = a_0 + b_1 X_1 + \text{erreur}$ . Le tableau 19 présente les estimations du modèle.

Tableau 19 : Modèle logistique analysant la probabilité d'obtenir un score de niveau 4 à PISA en lecture et en fonction du score à l'ECR de français

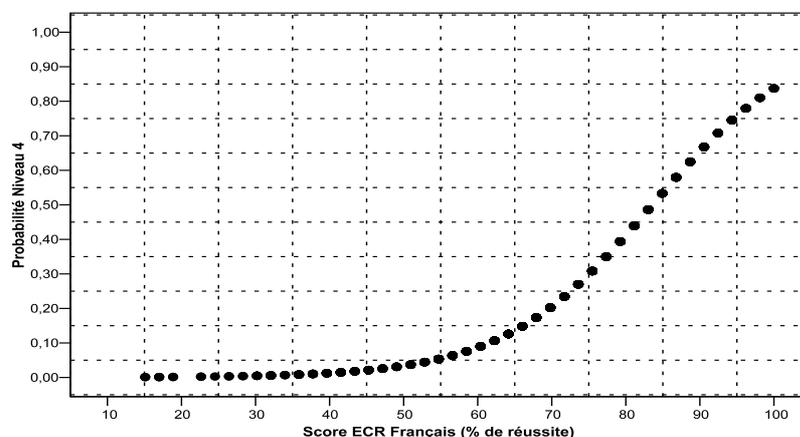
	Coef.	Sign.	Exp(B)
Score à l'ECR de français	+0,10	***	1,105
Constante	-8,33	***	0,000
R <sup>2</sup> de Nagelkerke		0,36	

significatif au seuil de 1%

Sans surprise (compte tenu du lien statistique entre les scores des deux épreuves), on constate que le score à l'ECR est bien prédicteur de l'appartenance au groupe d'élèves ayant atteint le niveau 4; en effet, le coefficient de la variable «score à l'ECR de français» est «positif et significatif». Pour interpréter de manière plus concrète l'impact du coefficient, on s'intéressera à l'indicateur  $\text{Exp}(B)$  qui permet de raisonner en termes de «chances» (*odds*) d'appartenir au groupe des élèves de niveau 4. Ainsi, dans le modèle, la valeur 1,105 signifie que pour 1 point de plus à l'ECR de français, les chances d'atteindre le niveau 4 à PISA augmentent de 1,105. Donc, pour 10 points de plus à l'ECR, les chances augmentent de  $10 \times 1,105$ , soit de 11. Autrement dit, un élève, qui a 10 points de plus à l'ECR qu'un autre élève, a 11 fois plus de chances (que cet autre élève) d'atteindre le niveau 4 de compétences à l'évaluation internationale. Pour simuler cet effet du score à l'ECR sur la probabilité d'obtenir le niveau 4 à PISA, nous avons calculé (sur la base du modèle logistique), les probabilités individuelles d'atteindre ce niveau et nous les avons mises en relation avec les scores aux ECR (graphique 24).

A chaque score de français à l'ECR est donc associée une probabilité, pour un élève, d'être considéré comme un bon lecteur à PISA. Par exemple, un élève obtenant un score de 70% de réussite à l'ECR (ce qui correspond à la moyenne sur notre échantillon) a une probabilité de 20% (0,20) d'atteindre le niveau 4 à PISA (soit une probabilité de 0,20), ce qui est faible. Pour que cette probabilité passe à 50%, le pourcentage de réussite à l'ECR doit être égal à 85. Si l'on souhaite obtenir un degré de certitude plus élevé, soit une probabilité de 75% (0,75), le score minimal nécessaire à l'ECR est de 95% de réussite. Or, le pourcentage d'élèves atteignant ce score de 95% est inférieur à 5% dans notre échantillon.

Graphique 24 : Probabilité d'atteindre le niveau 4 à PISA en lecture en fonction du score obtenu à l'ECR de français



Les résultats sont encore plus marqués si l'on utilise les critères habituellement utilisés dans PISA pour identifier les élèves forts. En effet, pour faire partie de cette catégorie les élèves doivent atteindre les niveaux 5 et 6. Ces niveaux correspondent notamment à une compréhension fine et détaillée de plusieurs documents, à la capacité d'émettre des hypothèses et un jugement critique, en faisant appel à des connaissances spécialisées. Sur l'échantillon vaudois, les niveaux 5 et 6 sont atteints par 8% des élèves. Nous avons conduit des estimations similaires aux précédentes sur la base de ce critère caractérisant les très bons lecteurs (tableau 20). Les résultats de la régression logistique indiquent (sur la base de la valeur de  $\text{Exp}(B)$ ) que pour 10 points de plus à l'ECR, les chances d'atteindre les niveaux 5 ou 6 augmentent de 11. Si l'on retrouve le même chiffre que précédemment

pour le niveau 4, les simulations sur la probabilité d'appartenir au groupe des très bons lecteurs sont très sélectives dans la mesure où elles exigent un score très élevé à l'ECR pour avoir une probabilité élevée d'obtenir un niveau 4 ou 6 à PISA.

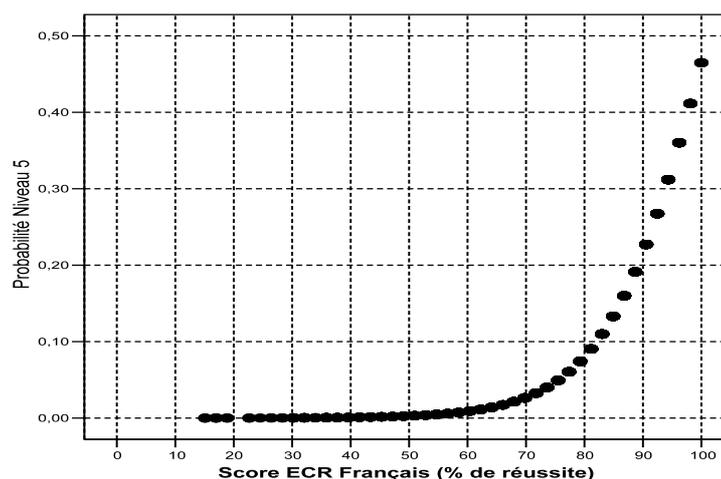
Tableau 20 : Modèle logistique analysant la probabilité d'obtenir un score de niveau 5 ou 6 à PISA en lecture et en fonction du score à l'ECR de français

	Coef.	Sign.	Exp(B)
Score à l'ECR de français	-11,6	***	1,12
Constante	-8,33	***	0,000
R <sup>2</sup> de Nagelkerke		0,26	

significatif au seuil de 1%

Sur le graphique 25, qui illustre les simulations conduites, on voit nettement l'allure exponentielle de la courbe matérialisant la relation entre le score à l'ECR de français et la probabilité d'être un très bon lecteur. Même avec un score de 90% de réussite à l'ECR, la probabilité d'atteindre les niveaux 5 ou 6 n'est que de 20%.

Graphique 25 : Probabilité d'atteindre les niveaux 5 ou 6 à PISA en lecture en fonction du score obtenu à l'ECR de français



En mathématiques, l'échelle de compétences est subdivisée en six niveaux. Selon les concepteurs de l'épreuve, les élèves qui n'atteignent pas le niveau de compétences 2 sont considérés comme très faibles, car ils ne maîtrisent pas les compétences considérées comme le minimum de base. A l'inverse, les élèves qui se situent aux niveaux 5 et 6 sont considérés comme très compétents dans le domaine. Plus précisément, au niveau 2, il s'agit seulement d'interpréter et de reconnaître des situations dans des contextes qui demandent tout au plus d'établir des inférences directes. Au niveau 5, les élèves peuvent notamment élaborer et utiliser des modèles dans des situations complexes pour construire des hypothèses. Le tableau 21 présente plus en détail les différents niveaux de compétences en culture mathématique et les limites des scores qui y sont associés. Le tableau 22 permet de comparer les répartitions vaudoises avec celles de l'OCDE selon les différents niveaux de compétences en mathématiques. On remarque que notre échantillon comporte moins d'élèves très faibles que celui de l'OCDE et une proportion équivalente d'élèves très forts (niveau 6); par ailleurs, même si des différences existent, il n'y a pas de différence structurelle majeure dans les répartitions des élèves dans les différents niveaux de compétences.

Tableau 21 : Niveaux de compétences en culture mathématique, PISA 2009

**Description succincte des six niveaux de compétence de l'échelle de culture mathématique**

Niveau	Score minimum requis	Capacités caractéristiques
<b>6</b>	669.3	<b>Au niveau 6</b> Les élèves sont capables de conceptualiser, de généraliser et d'utiliser des informations sur la base de leurs propres recherches et de la modélisation de problèmes complexes. Ils peuvent établir des liens entre différentes représentations et sources d'information et passer de l'une à l'autre sans difficulté. Ils peuvent se livrer à des raisonnements et à des réflexions mathématiques difficiles. Ils peuvent s'appuyer sur leur compréhension approfondie et leur maîtrise des relations symboliques et des opérations mathématiques classiques pour élaborer de nouvelles approches et de nouvelles stratégies à appliquer lorsqu'ils sont face à des situations qu'ils n'ont jamais rencontrées. Ils peuvent décrire clairement et communiquer avec précision leurs actes et les fruits de leur réflexion – résultats, interprétations, arguments – qui sont en adéquation avec les situations initiales.
<b>5</b>	607.0	<b>Au niveau 5</b> Les élèves peuvent élaborer et utiliser des modèles dans des situations complexes pour identifier des contraintes et construire des hypothèses. Ils sont capables de choisir, de comparer et d'évaluer des stratégies de résolution de problèmes mathématiques leur permettant de s'attaquer à des problèmes complexes en rapport avec ces modèles. Ils peuvent aborder les situations sous un angle stratégique en mettant en œuvre un grand éventail de compétences pointues de raisonnement et de réflexion, en utilisant les caractérisations symboliques et formelles et les représentations y afférentes et en s'appuyant sur leur compréhension approfondie de ces situations. Ils peuvent réfléchir à leurs actes et formuler et communiquer leurs interprétations et leur raisonnement.
<b>4</b>	544.7	<b>Au niveau 4</b> Les élèves sont capables d'utiliser des modèles explicites pour faire face à des situations concrètes complexes qui peuvent leur demander de tenir compte de contraintes ou de construire des hypothèses. Ils peuvent choisir et intégrer différentes représentations, dont des représentations symboliques, et les relier directement à certains aspects de situations tirées du monde réel. Ils peuvent mettre en œuvre un éventail de compétences pointues dans ces situations et raisonner avec une certaine souplesse en s'appuyant sur leur compréhension de ces contextes. Ils peuvent formuler des explications et des arguments sur la base de leurs interprétations et de leurs actions et les communiquer.
<b>3</b>	482.4	<b>Au niveau 3</b> Les élèves peuvent appliquer des procédures bien définies, dont celles qui leur demandent des décisions séquentielles. Ils peuvent choisir et mettre en œuvre des stratégies simples de résolution de problèmes mathématiques. Ils peuvent interpréter et utiliser des représentations basées sur différentes sources d'information et construire leur raisonnement directement sur cette base. Ils peuvent rendre compte succinctement de leurs interprétations, de leurs résultats et de leur raisonnement.
<b>2</b>	420.1	<b>Au niveau 2</b> Les élèves peuvent interpréter et reconnaître des situations dans des contextes qui leur demandent tout au plus d'établir des inférences directes. Ils ne peuvent puiser des informations pertinentes que dans une seule source d'information et n'utiliser qu'un seul mode de représentation. Ils sont capables d'utiliser des algorithmes, des formules, des procédures ou des conventions élémentaires. Ils peuvent se livrer à un raisonnement direct et interpréter les résultats de manière littérale.
<b>1</b>	357.8	<b>Au niveau 1</b> Les élèves peuvent répondre à des questions s'inscrivant dans des contextes familiers, dont la résolution ne demande pas d'autres informations que celles présentes et qui sont énoncées de manière explicite. Ils sont capables d'identifier les informations et d'appliquer des procédures de routine sur la base de consignes directes, dans des situations explicites. Ils peuvent exécuter des actions qui vont de soi et qui découlent directement du stimulus donné.

Source : OCDE, 2010

Tableau 22 : Répartition des élèves de l'échantillon et de l'OCDE selon le niveau de compétences en culture mathématique à PISA 2009 (en %)

Niveaux	Echantillon	OCDE
Niveau 6	3,0	3,1
Niveau 5	12,3	9,6
Niveau 4	24,7	18,9
Niveau 3	27,4	24,3
Niveau 2	19,9	22,0
Niveau 1	9,4	14,0
Inférieur au niveau 1	3,2	8,0

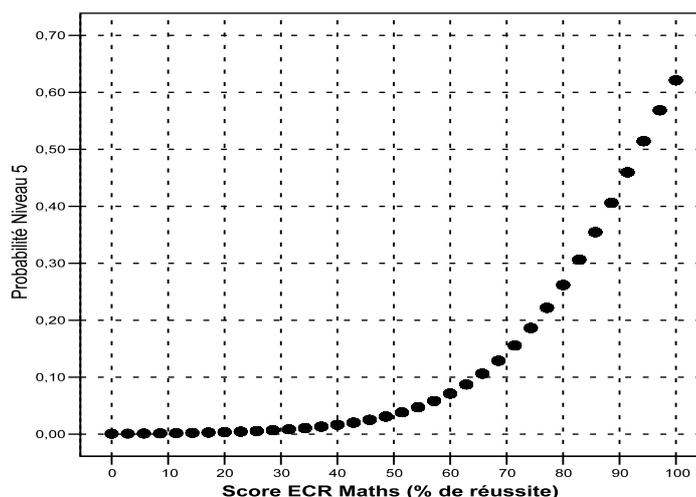
Comme nous l'avons fait pour le français, nous avons estimé la probabilité qu'un élève soit considéré comme fort à PISA, en fonction de son score à l'ECR. Le tableau 23 présente les résultats de la régression logistique et le graphique 26 les simulations liées à la probabilité d'atteindre le niveau 5 ou 6 selon le score obtenu à l'épreuve cantonale. Le coefficient associé à l'ECR implique un effet proche de celui constaté en français puisque le rapport de chances (*odds*) est de 1,08, soit, pour dix points de plus à l'ECR, près d'onze fois plus de chances d'être considéré comme fort à PISA (coefficient de 1,08). Le graphique montre par ailleurs qu'un élève de niveau moyen à l'ECR (le pourcentage moyen de réussite est de 59%) a une probabilité très faible (inférieure à 10%) d'atteindre le niveau 5 à PISA. Il faut que le taux de réussite à l'ECR soit supérieur à 90% pour qu'un élève ait une chance sur deux d'atteindre le niveau 5 de compétences à l'évaluation PISA en mathématiques.

Tableau 23 : Modèle logistique analysant la probabilité d'obtenir un score de niveau 5 ou 6 à PISA en mathématiques et en fonction du score à l'ECR de mathématiques

	Coef.	Sign.	Exp(B)
Score à l'ECR de mathématiques	+0,08	***	1,08
Constante	-7,1	***	0,001
R <sup>2</sup> de Nagelkerke		0,34	

significatif au seuil de 1%

Graphique 26 : Probabilité d'atteindre les niveaux 5 ou 6 à PISA en mathématiques en fonction du score obtenu à l'ECR de mathématiques

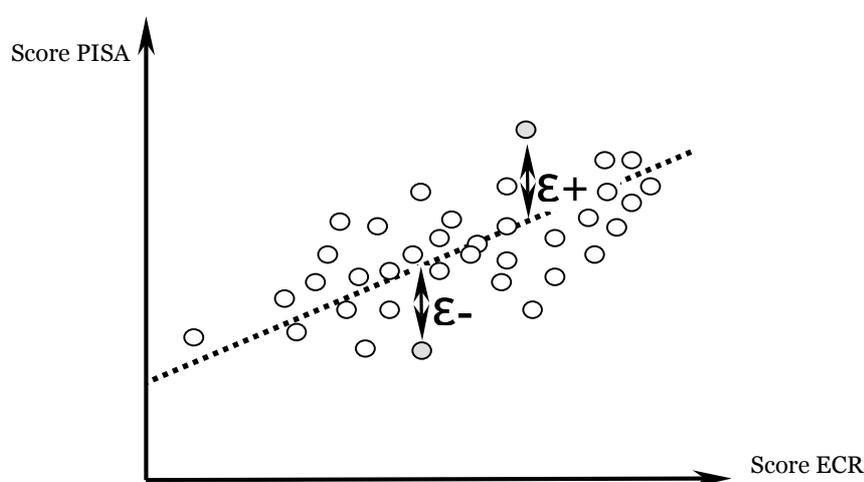


Au final, ces simulations réalisées dans les deux disciplines montrent clairement que l'évaluation PISA s'avère très sélective pour les critères de compétences les plus élevés (niveaux 5 et 6) quand on se base sur les épreuves cantonales. La probabilité d'être considéré comme un très bon lecteur, ou un élève très compétent en mathématiques, est faible pour la grande majorité des élèves de notre échantillon. Ce qui retient surtout l'attention ce sont les allures exponentielles des courbes qui témoignent d'une forte sélectivité quand on se base sur les scores des ECR pour prédire l'appartenance à l'élite de PISA. Après cette première analyse des liaisons entre PISA et les ECR, il peut être instructif de voir si les différences constatées dans les classements respectifs des élèves dans chacune des évaluations peuvent s'expliquer par certaines caractéristiques individuelles et par le contexte d'enseignement.

## 5.2 LES DÉTERMINANTS DES ÉCARTS DES RÉSULTATS ENTRE PISA ET LES ECR

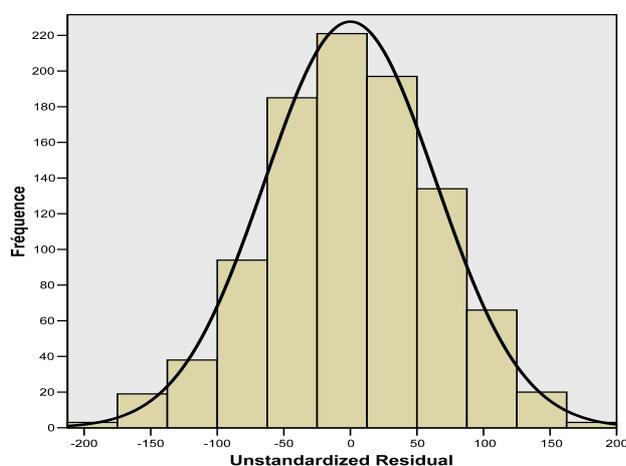
Il s'agit ici de mesurer l'effet des caractéristiques des élèves sur les écarts de scores entre PISA et les ECR. Plusieurs possibilités sont offertes au niveau des analyses statistiques pour explorer cette question qui fournissent des solutions comparables. Nous avons choisi une méthode qui soit porteuse de sens en matière d'interprétation des résultats. Les graphiques précédents qui rendent compte des relations entre les scores à PISA et ceux obtenus aux ECR (graphiques 22 et 23) ont mis en évidence des situations d'élèves assez distantes des relations moyennes estimées par les équations de régression. Il s'agit concrètement des points qui s'écartent sensiblement de la droite de régression, soit situés au-dessus (dans le cas d'un score observé supérieur à celui estimé), soit situés en dessous (dans le cas d'un score observé inférieur à celui estimé). Ce sont précisément ces écarts à la moyenne que nous cherchons à expliquer par les caractéristiques des élèves déjà mobilisées précédemment (sexe, retard scolaire, pays de naissance, langue parlée à la maison et indice socio-économique et culturel). Pour cela, nous estimerons des modèles de régression qui utilisent comme variable dépendante les résidus des régressions précédentes effectuées entre les scores des deux évaluations (PISA et ECR). Ces résidus correspondent aux écarts positifs ou négatifs entre les valeurs observées et celles prédites par les équations de régression précédentes (graphique 27).

Graphique 27 : Illustration des résidus de la régression des scores de PISA sur les scores aux ECR



Le graphique 28 présente la distribution gaussienne de ces résidus en ce qui concerne les épreuves de lecture. Pour faciliter l'interprétation, nous avons choisi d'exprimer ces résidus avec leurs valeurs non standardisées, ils correspondent donc à des valeurs directement interprétables dans l'échelle de PISA. De part et d'autre de la moyenne (égale à 0) des écarts entre scores prédits et observés, il existe une variété assez grande des valeurs, l'écart type étant de 64. Cette dernière valeur signifie que les deux tiers des élèves ont un score observé qui peut varier de - 60 points à + 60 points par rapport à la situation théorique (scores prédits par le modèle).

Graphique 28 : Distribution des résidus de la régression entre les scores aux évaluations de lecture (PISA et ECR)



L'objectif est donc de tester si certaines caractéristiques des élèves (dont la plupart exercent, par ailleurs, une influence sur les scores) peuvent expliquer ces écarts entre scores observés et scores prédits. Le tableau 24 présente les résultats des modèles analysant la variance des résidus en lecture (en français pour les ECR) et en mathématiques avec les caractéristiques individuelles des élèves comme variables indépendantes. Les modèles fournissent des résultats assez différents dans le sens où davantage de variables sont significatives en mathématiques. En lecture, seul le milieu socio-économique et, dans une moindre mesure, le retard scolaire exercent une influence sur la valeur des résidus. En revanche, en mathématiques, l'ensemble des variables, à l'exception de la langue parlée à la maison, sont significatives. Ainsi, sur la base de leurs performances aux ECR, les filles ont dans ce domaine un résidu inférieur de 14,7 points à celui des garçons. A l'inverse, toujours en mathématiques, les élèves autochtones affichent un résidu supérieur de 12,9 points à ceux des élèves nés à l'étranger. Le milieu socio-économique, dans les deux disciplines, est la variable la plus discriminante dans les modèles, elle explique en effet à elle seule une grande part de la variance des résidus (96% pour la lecture et 78% pour les mathématiques). En lecture, un écart type de plus à l'IESCS accroît les résidus de 13 points, le chiffre est de 18,7 points en mathématiques.

Tableau 24 : Modèles d'analyse de la variance des résidus de la régression entre les scores des évaluations (PISA et ECR) en fonction des caractéristiques des élèves

Variables		Lecture		Mathématiques	
Référence	Active	Coef.	Sign.	Coef.	Sign.
Garçon	Fille	+2,7	n. s.	-14,7	***
Age normal	Retard scolaire	-8,0	*	-12,7	***
Issu de la migration	Autochtone	+6,3	n. s.	+12,9	***
Parle une autre langue	Parle français à la maison	+6,9	n. s.	-2,0	n. s.
Indice socio-économique (IESCS)		+12,7	***	+18,7	***
Constante		-8,0	n. s.	+4,0	n. s.
R <sup>2</sup>		0,06		0,14	

n. s. : non significatif, \* : significatif au seuil de 10%, \*\*\* : significatif au seuil de 1%

Le tableau 25 complète ces analyses en intégrant aux modèles nos variables contextuelles classiques, en l'occurrence la fréquentation des filières. L'effet des filières est fort, particulièrement en mathématiques en ce qui concerne la fréquentation de la VSB. On pourra être surpris du coefficient associé à la VSG pour la lecture qui traduit un effet négatif de la fréquentation de cette filière sur les résidus, en référence à la fréquentation de la VSO; ce résultat s'éloigne en effet des précédents qui traduisaient la hiérarchie habituelle des trois voies. En mathématiques, les élèves scolarisés en VSB affichent des écarts positifs et élevés (+31,9 points) par rapport à la relation moyenne entre PISA et l'ECR sur l'échantillon.

Tableau 25 : Modèles d'analyse de la variance des résidus de la régression entre les scores des évaluations (PISA et ECR) en fonction des caractéristiques des élèves et de la filière fréquentée

Variables		Lecture		Mathématiques	
Référence	Active	Coef.	Sign.	Coef.	Sign.
Garçon	Fille	+1,9	n. s.	-16,4	***
Age normal	Retard scolaire	-6,2	n. s.	-8,9	*
Issu de la migration	Autochtone	+5,8	n. s.	+10,6	**
Parle une autre langue	Parle français à la maison	+9,4	n. s.	+0,0	n. s.
Indice socio-économique (IESCS)		+8,9	***	+12,5	***
VSO	VSG	-12,0	**	+3,6	n.s.
	VSB	+16,1	***	+31,9	***
Constante		-11,7	*	-9,3	n.s.
R <sup>2</sup>		0,10		0,18	

n. s. : non significatif, \* : significatif au seuil de 10%, \*\* : significatif au seuil de 5%  
 \*\*\* : significatif au seuil de 1%

Cette analyse de la variance des résidus montrent que les caractéristiques socio-économiques des élèves et la fréquentation des filières expliquent une partie des différences de classements des élèves dans les deux types d'évaluations. En effet, ces variables exercent un impact sur les écarts entre les scores observés à PISA sur la base des ECR et ceux prédits sur cette même base. Autrement dit, un milieu socio-économique de l'élève favorisé, de même que la fréquentation de la VSB, a tendance à surestimer le classement des élèves à PISA. L'inverse est évidemment vrai pour les élèves de milieu social défavorisé et ceux fréquentant la VSO: sur la base de leurs scores aux ECR, ces élèves sont moins bien classés à PISA que les autres. Le fait que ces caractéristiques individuelles jouent un rôle dans la relation statistique entre les scores aux deux évaluations confirme d'une certaine manière les constatations précédentes qui ont mis en évidence l'impact différencié de ces variables sur les performances des élèves selon la mesure des performances prise en compte (PISA ou ECR). Si le classement des élèves dans les deux épreuves est naturellement lié à leurs caractéristiques sociales et scolaires, celles-ci n'interviennent pas de manière similaire dans les deux types d'épreuves.

## 6 ANALYSE SELON LES DOMAINES DE COMPÉTENCES

Jusqu'à présent, les scores des élèves ont été appréhendés sous leur forme globale en prenant en compte l'ensemble des items composant les épreuves sans différencier les différents domaines qui ont été définis par les concepteurs des épreuves. L'objectif de ce chapitre est de mettre à jour d'éventuelles spécificités des relations entre les deux outils d'évaluation selon leurs domaines de compétences. Une difficulté majeure, mais qui est inhérente à notre travail de comparaison, tient au fait que ces domaines de compétences sont définis de manière différente selon que l'on considère PISA ou les ECR. La nature même des épreuves, leurs formes, leurs finalités, les méthodologies utilisées pour la production des items et l'échelle de mesure sont autant d'éléments qui rendent la comparaison directe des domaines de compétences impossible. Il ne s'agit donc pas ici de tenter de construire des domaines équivalents dans les deux évaluations, mais seulement de mesurer et d'apprécier les relations entre ces domaines et de voir si les relations identifiées sur les scores globaux (en termes d'intensité, notamment) se confirment quand les analyses ne portent que sur des dimensions plus spécifiques des acquisitions des élèves. Une première question pour laquelle on peut livrer des informations utiles à notre problématique est celle du lien entre les différents domaines des ECR et le score global obtenu à PISA par les élèves. Au-delà de la relation moyenne observée précédemment entre les deux scores globaux des évaluations, il peut être particulièrement intéressant d'identifier les domaines des ECR qui entretiennent les plus fortes relations avec PISA et, par complémentarité, ceux qui apparaissent les plus indépendants de l'évaluation internationale. Nous traiterons cette question en deux temps en examinant tout d'abord la situation en lecture, puis celle en mathématiques<sup>21</sup>.

### 6.1 LES RELATIONS ENTRE PISA ET LES DIMENSIONS DES ECR EN LECTURE

Nous allons dans un premier temps présenter succinctement les différents domaines des épreuves cantonales de référence de 8<sup>e</sup> année de français. Cette épreuve est sous un format très différent des tests PISA puisqu'elle prend appui sur deux notices biographiques qui constituent le contexte de tous les exercices. Il y a donc une unité thématique qui donne une certaine cohérence et un sens commun aux items. Le tableau 26 fournit une description des compétences associées aux objectifs fondamentaux visés par cette épreuve, quatre domaines de compétences ont été définis. Le premier, nommé «structurer» porte sur le fonctionnement de la langue en évaluant la maîtrise des règles d'orthographe d'usage et grammaticales principalement par le biais de la dictée et de la production écrite. Le second domaine, intitulé «lire» mobilise des compétences en matière de sélection, d'analyse et de compréhension de l'information. Le troisième domaine, intitulé «écrire», cible la production d'écrit avec le respect de différentes contraintes propres à la situation choisie (texte informatif). Le dernier domaine porte le même intitulé que le premier, à savoir «structurer», mais se différencie dans la mesure où les compétences visées sont évaluées à travers une rédaction et ciblent de nombreux aspects du fonctionnement et de l'utilisation de la langue écrite: syntaxe, vocabulaire, orthographe, ponctuation, utilisation des formes verbales, utilisation d'indicateurs chronologiques...

<sup>21</sup> Précisons déjà qu'en aucune manière nous nous prononcerons sur la validité ou la fidélité des épreuves cantonales; ce point, par ailleurs intéressant en soi, est éloigné de nos préoccupations dans cette recherche: ce sont les épreuves sous leur forme actuelle qui nous intéressent avec leurs qualités comme leurs imperfections.

Tableau 26 : Domaines et compétences de l'ECR de français de 8<sup>e</sup> année en 2008<sup>22</sup>

Compétences visées	Compétences associées
<p><b>STRUCTURER</b></p> <ul style="list-style-type: none"> <li>• Raisonner sur le fonctionnement de la langue</li> </ul>	<ul style="list-style-type: none"> <li>• Appliquer les règles d'accord</li> <li>• Appliquer les règles d'orthographe d'usage</li> <li>• Recourir efficacement aux ouvrages de référence</li> <li>• Gérer l'orthographe dans toute production écrite</li> <li>• Utiliser les signes de ponctuation</li> </ul>
<p><b>LIRE</b></p> <ul style="list-style-type: none"> <li>• Lire de façon autonome</li> <li>• Chercher de l'information et en tirer parti</li> <li>• Analyser et comprendre, en situation, de manière critique et sélective, un message à visée informative</li> </ul>	<ul style="list-style-type: none"> <li>• Tirer des informations de sources diverses</li> <li>• Identifier la source d'information</li> <li>• Sélectionner des informations</li> <li>• Repérer les marques d'organisation</li> <li>• Décrire la situation de communication</li> </ul>
<p><b>ÉCRIRE</b></p> <ul style="list-style-type: none"> <li>• Produire un message à visée informative: traiter l'information et la mettre en forme en respectant les contraintes propres à la situation et au genre choisis</li> </ul>	<ul style="list-style-type: none"> <li>• Ecrire une biographie courte</li> <li>• Appliquer les caractéristiques du texte informatif dans la production d'un article de presse</li> <li>• Utiliser des marques d'organisation</li> <li>• Intégrer des illustrations</li> <li>• Adapter la forme et le contenu au destinataire</li> </ul>
<p><b>STRUCTURER</b></p> <ul style="list-style-type: none"> <li>• Raisonner sur le fonctionnement de la langue</li> </ul>	<ul style="list-style-type: none"> <li>• Mobiliser les connaissances textuelles, syntaxiques et orthographiques acquises</li> <li>• Repérer et produire les diverses possibilités de désigner un personnage ou un objet</li> <li>• Repérer et analyser des mots ou expressions qui participent à l'organisation d'un texte</li> <li>• Choisir dans le dictionnaire l'information qui correspond au contexte</li> <li>• Gérer la conjugaison et l'emploi des temps dans toute production écrite</li> <li>• Utiliser les signes de ponctuation</li> </ul>

Source : Direction générale de l'enseignement obligatoire (DGEO)

Il n'est pas inutile d'examiner comment ces quatre domaines peuvent être liés au niveau statistique. Le tableau 27 présente les corrélations établies sur notre échantillon d'élèves. Sans surprise, on constate que les domaines les plus corrélés sont ceux qui sont les plus proches par les compétences visées; toutefois, la corrélation entre les deux domaines intitulés «structurer», même si elle est élevée (+0,66), montre une certaine spécificité de chacun de ces domaines. Les autres corrélations oscillent entre +0,45 et +0,55, elles peuvent donc être considérées comme moyennes et elles témoignent aussi de la spécificité de chaque domaine.

<sup>22</sup> Le tableau est repris de la présentation officielle des épreuves réalisée par la DGEO.

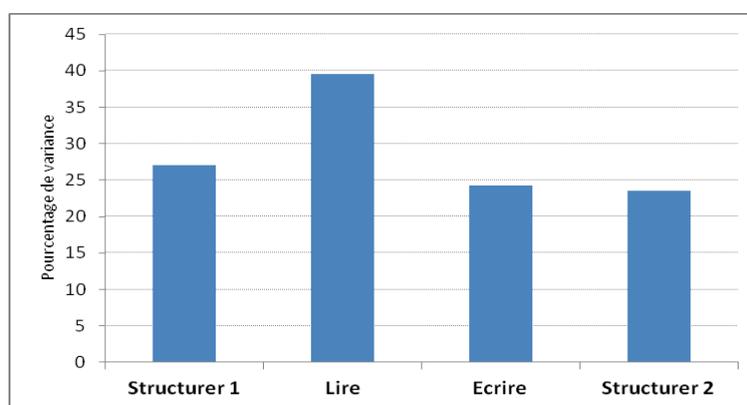
Tableau 27 : Corrélations entre les domaines de l'ECR de français

	Structurer 1	Lire	Ecrire
Lire	+ 0,55 ***		
Ecrire	+0,45 ***	+ 0,48 ***	
Structurer 2	+0,66 ***	+0,48 ***	+0,54 ***

\*\*\* : significatif au seuil de 1%

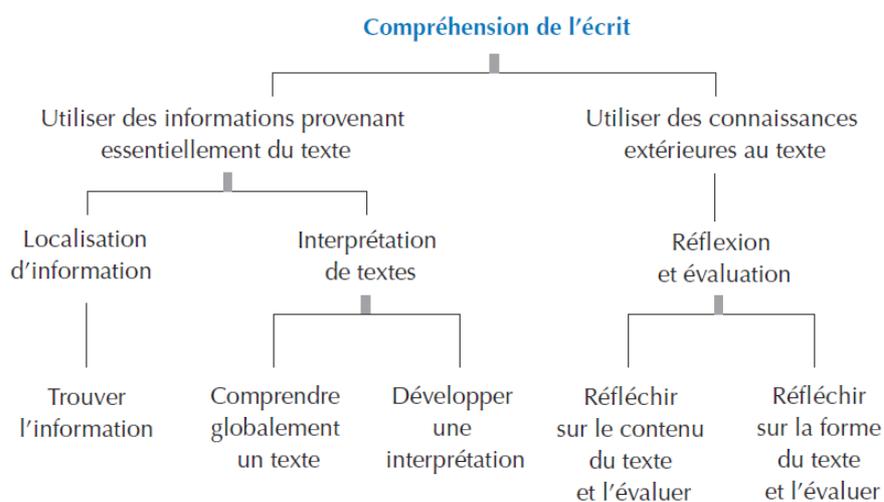
Une première estimation de la liaison des domaines de compétences avec PISA peut être réalisée en estimant, pour chacun d'entre eux, une régression expliquant la variance du score de lecture par rapport à l'évaluation internationale. Le graphique 29 présente les pourcentages de variance expliquée dans les quatre domaines.

Graphique 29 : Pourcentages de variance expliquée du score PISA de lecture par les domaines de compétences de l'ECR de français



Une dimension se détache nettement des autres, il s'agit du domaine intitulé «lire» ce qui est conforme aux attentes. En effet, les compétences visées: «chercher de l'information et en tirer parti, lire de façon autonome, tirer des informations, sélectionner des informations, analyser et comprendre [...]» sont très voisines de celles évaluées par PISA comme on peut le voir sur le schéma suivant, extrait du cadre d'évaluation de PISA 2009.

Schéma 2 : Cadre d'évaluation de PISA 2009 en compréhension de l'écrit



Source OCDE, 2012, p. 35

Ce domaine «lire» explique à lui seul près de 40% de la variance du score à PISA (soit une corrélation de +0,63); ce chiffre est tout à fait important quand on le rapproche de celui constaté précédemment concernant le score global de français qui est de 45%. Ce seul domaine a donc un pouvoir explicatif qui correspond au neuf dixièmes de celui du score global. Les autres domaines ont, quant à eux, une contribution plus faible de l'ordre de 23,5 à 27%. Pour mieux saisir la contribution spécifique des dimensions de l'ECR à l'explication de la variance du score de PISA, on peut estimer une régression «pas à pas» (ou *stepwise*). Cette régression de type itérative inclut d'abord dans le modèle la variable qui propose le meilleur coefficient de détermination ( $R^2$ ), puis celle qui améliore le plus ce coefficient et ainsi de suite<sup>23</sup> (tableau 28). Les valeurs des coefficients indiquent une nette hiérarchie des domaines dans leur capacité à expliquer les différences de résultats des élèves à PISA. Les scores ayant été centrés réduits dans cette régression, on peut comparer directement les valeurs des coefficients associés aux domaines. Ainsi, un écart type de plus au score du domaine «lire» entraîne une augmentation du score de PISA de plus de 36 points, alors que le chiffre est de moins de 8 points pour la dimension «structurer 2».

Tableau 28 : Modèles de régression «stepwise» d'analyse de la variance du score de PISA en lecture en fonction des domaines de compétences des ECR de français

Variabes	Coef. standardisé	Sign.
Score Lire	+36,4	***
Score Ecrire	+15,3	***
Score Structurer 1	+12,9	***
Score Structurer 2	+7,9	***
Constante	+235	***
$R^2$	0,47	

\*\*\* : significatif au seuil de 1%

## 6.2 LES RELATIONS ENTRE PISA ET LES DIMENSIONS DES ECR EN MATHÉMATIQUES

En mathématiques, l'ECR de 2008 vise plusieurs compétences regroupées en trois domaines comportant chacun plusieurs objectifs de la discipline (tableau 29).

<sup>23</sup> Il faut préciser que la seconde variable qui entre dans le modèle n'est pas obligatoirement celle qui présente, à elle seule, le second meilleur coefficient de détermination du modèle, car dans le cas d'une forte corrélation entre  $X_1$  et  $X_2$ , l'accroissement de la variance dû à l'introduction de  $X_2$  serait trivial. C'est donc la variable qui contribue le plus à réduire la variabilité résiduelle du modèle en voie d'élaboration qui est sélectionnée à chaque étape.

Tableau 29 : Domaines et compétences de l'ECR de mathématiques de 8<sup>e</sup> année en 2008

OBJECTIFS NUMÉRIQUES ET ALGÈBRIQUES	<b>Résoudre des problèmes numériques</b> Calculer en situation et porter un regard critique sur le résultat
	<b>Résoudre des problèmes par voie algébrique</b> Résoudre une équation
	<b>Représenter des situations de la vie courante</b> <ul style="list-style-type: none"> <li>• Traduire un ensemble de données par une représentation appropriée</li> <li>• Interpréter des données chiffrées</li> </ul>
	<b>Modéliser des situations de la vie courante</b> Reconnaître et utiliser les propriétés des fonctions
OBJECTIFS GÉOMÉTRIQUES	<b>Mobiliser la mesure pour comparer des grandeurs</b> <ul style="list-style-type: none"> <li>• Calculer des grandeurs</li> <li>• Utiliser les théorèmes métriques pour calculer des longueurs, des aires</li> </ul>
	<b>Modéliser le plan et l'espace</b> Reconnaître et utiliser des transformations géométriques
<b>Communiquer sa démarche et ses résultats avec un vocabulaire précis</b>	

Source : Direction générale de l'enseignement obligatoire (DGEO)

Le premier domaine cible des objectifs numériques et algébriques (*résoudre des problèmes numériques, représenter les situations de la vie courante, etc.*). Le second domaine est axé sur des objectifs géométriques (*mobiliser la mesure pour comparer des grandeurs, modéliser le plan et l'espace, etc.*). Un troisième domaine, davantage transversal, porte sur la communication de la démarche et des résultats. Le tableau suivant montre les corrélations entre les trois domaines. On relève des valeurs élevées, particulièrement entre les items relevant de la démarche et ceux relatifs aux objectifs numériques.

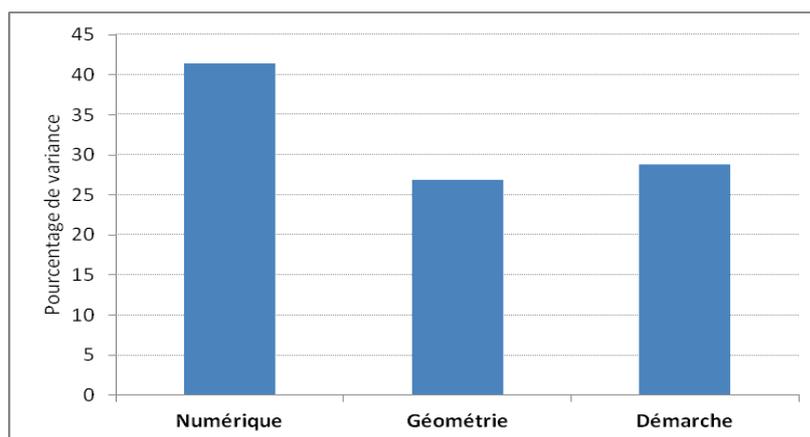
Tableau 30 : Corrélations entre les domaines de l'ECR de mathématiques

	Géométrie	Démarche
Numérique	+0,63 ***	+0,77 ***
Démarche	+0,59 ***	

\*\*\* : significatif au seuil de 1%

Quand on examine séparément la contribution respective des domaines de l'ECR à l'explication des différences de performances à PISA, on remarque qu'un domaine se démarque avec un pourcentage de variance expliquée de 41,4%, il s'agit du domaine numérique (graphique 30). Notons que cette seule dimension explique la même part de variance du score de PISA que l'ensemble des items de l'ECR. Les deux autres composantes ont, quant à elles, une contribution inférieure à 30%.

Graphique 30 : Pourcentages de variance expliquée du score PISA de mathématiques par les domaines de compétences de l'ECR de mathématiques



Compte tenu des fortes corrélations entre les trois domaines, et pour avoir une idée plus précise de leur contribution spécifique au score obtenu à PISA, nous avons estimé, comme précédemment pour le français, une régression «pas à pas» qui permet d'identifier le domaine le plus prédictif du score à PISA (tableau 30). Une variable a été exclue des résultats des estimations, il s'agit de celle représentant la dimension «démarche» du fait de sa très faible contribution autonome à l'explication de la variance du score à PISA. Au-delà de la contribution du domaine «numérique» (41,4% de la variance), le domaine «géométrie» apporte une part spécifique de seulement 2,1%. C'est donc essentiellement les items relevant des objectifs numériques et algébriques qui expliquent les différences de performances des élèves à PISA en mathématiques.

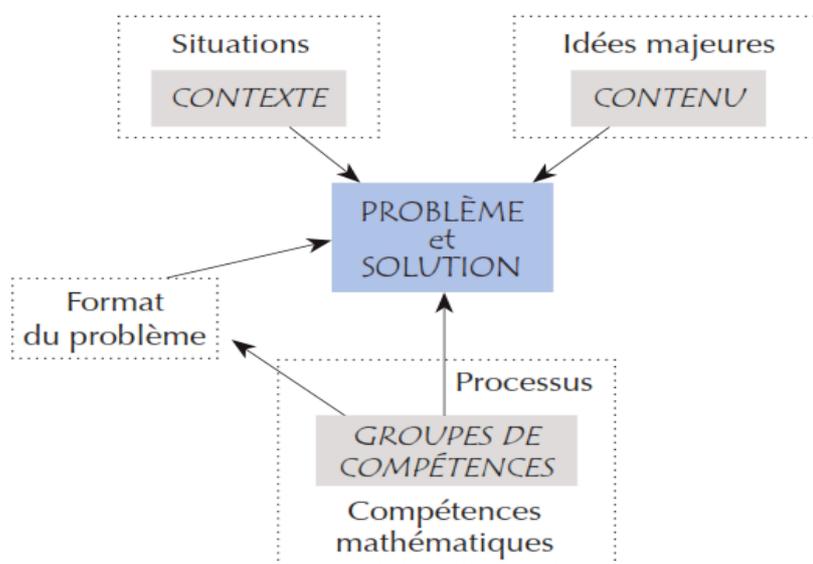
Tableau 31 : Modèles de régression «stepwise» d'analyse de la variance du score de PISA en mathématiques en fonction des domaines de compétences de l'ECR de mathématiques

Variabiles	Coef. standardisé	Sign.
Score numérique	+0,52	***
Score géométrie	+0,19	***
Constante	387	***
R <sup>2</sup>	0,44	

\*\*\* : significatif au seuil de 1%

Ce constat reflète une certaine logique quand on examine la manière dont PISA conçoit l'évaluation des mathématiques (schéma 3).

Schéma 3 : Les composantes du domaine PISA 2009 de mathématiques



Source OCDE, 2012, p. 86

L'accent est mis sur la résolution de problèmes et non sur les connaissances elles-mêmes: «*La culture mathématique d'un individu se reflète, par exemple, dans la manière dont il utilise ses connaissances et compétences mathématiques pour résoudre des problèmes. Les problèmes (et leurs solutions) peuvent se présenter dans des situations et des contextes très divers lors des expériences vécues par l'individu.*» (OCDE, 2010, p.86). Or, les ECR font aussi appel majoritairement à la résolution de problèmes, notamment ceux issus de la vie courante. Les différentes composantes des ECR n'ont donc pas toutes le même poids pour expliquer les performances des élèves à PISA. On notera que deux domaines, «lire» et «numérique», apportent une contribution très importante, en termes de variance expliquée, aux scores PISA. Cela signifie que les compétences visées dans les items de ces deux domaines sont beaucoup plus proches des épreuves PISA que les autres compétences visées dans les ECR.

### 6.3 L'INFLUENCE DES CARACTÉRISTIQUES DES ÉLÈVES DANS LES DIFFÉRENTES DIMENSIONS DES ECR

Le fait d'avoir mobilisé les différentes composantes des ECR nous invite à explorer un point déjà abordé précédemment au niveau des scores globaux, à savoir l'influence des caractéristiques individuelles sur les acquis des élèves. Il est en effet possible que, selon les domaines considérés, l'influence des facteurs soit variable. Dans cette perspective, nous avons estimé un modèle de régression pour chaque domaine des ECR (tableaux 31 et 32). Les coefficients sont standardisés afin de permettre une comparaison directe de leur valeur et de leur impact. Dans les deux disciplines, les modèles fournissent des résultats proches.

Tableau 32 : Modèles d'analyse de la variance des scores des domaines de l'ECR de français en fonction des caractéristiques des élèves

Variables		ECR Français							
		Structurer 1		Lire		Ecrire		Structurer 2	
Référence	Active	Coef.	Sign.	Coef.	Sign.	Coef.	Sign.	Coef.	Sign.
Garçon	Fille	+0,22	***	+0,06	**	+0,21	***	+0,24	***
Age normal	Retard scolaire	-0,17	***	-0,13	***	-0,09	***	-0,11	***
Issu de la migration	Autochtone	+0,11	***	+0,09	**	+0,04	n. s.	+0,09	***
Parle une autre langue	Parle français à la maison	+0,04	n. s.	+0,03	n. s.	+0,03	n. s.	+0,00	n.s.
Indice socio-économique (IESCS)		+0,24	***	+0,29	***	+0,23	***	+0,22	***
Constante		+5,1	***	+15,5	***	+7,2	***	+6,0	***
R <sup>2</sup>		0,21		0,16		0,13		0,15	

n. s. : non significatif, \* : significatif au seuil de 10%, \*\* : significatif au seuil de 5%, \*\*\* : significatif au seuil de 1%

Tableau 33 : Modèles d'analyse de la variance des scores des domaines de l'ECR de mathématiques en fonction des caractéristiques des élèves et de la filière fréquentée

Variables		ECR Mathématiques					
		Numérique		Géométrique		Démarche	
Référence	Active	Coef.	Sign.	Coef.	Sign.	Coef.	Sign.
Garçon	Fille	-0,16	***	-0,07	**	-0,05	*
Age normal	Retard scolaire	-0,17	***	-0,16	***	-0,17	***
Issu de la migration	Autochtone	+0,10	***	+0,02	n. s.	+0,11	n. s.
Parle une autre langue	Parle français à la maison	-0,06	n. s.	-0,04	n. s.	-0,06	*
Indice socio-économique (IESCS)		+0,23	n. s.	+0,19	***	+0,18	***
Constante		+11,4	***	+6,4	***	+4,7	***
R <sup>2</sup>		0,15		0,08		0,09	

n. s. : non significatif, \* : significatif au seuil de 10%, \*\* : significatif au seuil de 5%, \*\*\* : significatif au seuil de 1%

## 7 CONCLUSION

Les enquêtes internationales sur les acquisitions des élèves ont une histoire de plus de cinquante années, mais l'OCDE a donné une nouvelle impulsion à ces études avec le lancement du programme PISA en 2000. Ce programme de grande ampleur concerne un nombre important de pays (65 pays ont participé en 2009, 67 en 2012 et près de 70 participeront en 2015); des moyens financiers substantiels sont mobilisés pour la réalisation des différentes étapes du programme. La richesse des informations collectées sur les élèves et les établissements, la dimension cyclique de l'enquête, administrée tous les trois ans et dans trois domaines de compétences (compréhension de l'écrit, culture mathématique et culture scientifique), sont autant d'éléments qui ont permis une large exploitation des résultats et qui ont contribué, par ailleurs, à la popularité de PISA. Pour les chercheurs en éducation, les données PISA constituent une base de travail précieuse et elles donnent lieu à de nombreuses exploitations secondaires qui produisent des analyses originales sur l'efficacité, l'efficience et l'équité des systèmes éducatifs. Mais l'atout principal du programme PISA est de pouvoir nourrir et enrichir les débats politiques et sociaux sur l'école dans un grand nombre de pays. Les responsables des politiques éducatives font en effet systématiquement référence aux conclusions des analyses menées dans le cadre de PISA et les médias leur donnent une large place, notamment lors de la publication des résultats. Mais l'enquête internationale fait également l'objet de nombreuses critiques dont la principale concerne la mesure des compétences telle que l'envisage PISA. La recherche exposée dans ce rapport contribue à alimenter le débat relatif à cette question en explorant et interrogeant cette mesure des compétences des élèves, de manière empirique.

L'objectif principal de cette recherche était d'examiner le lien statistique entre les performances des élèves vaudois mesurées à l'aune de deux indicateurs: les scores obtenus à l'enquête PISA et ceux obtenus aux épreuves cantonales de référence. Le fait de pouvoir disposer, pour un même échantillon d'élèves, des données relatives à ces deux mesures peut être considéré comme une opportunité à saisir, à la fois pour le chercheur en éducation et pour le décideur politique. Pour le premier, cette recherche présente en effet un intérêt particulier dans le sens où peu de travaux ont exploré empiriquement cette question. Pour le second, l'occasion de confronter les résultats des élèves à PISA avec ceux des épreuves locales, plus proches des objectifs des plans d'étude, permet d'envisager une lecture différente des conclusions de l'enquête internationale. Les analyses conduites dans ce travail de recherche se révèlent pertinentes dans la mesure où elles mettent en évidence des différences nettes entre les deux sources de données (PISA et les ECR); trois constats principaux peuvent être dégagés sur la base de ces analyses.

Le premier constat concerne la relation des statistiques entre les résultats des élèves à l'enquête PISA et ceux obtenus aux épreuves cantonales de référence. On aurait pu s'attendre à une proximité plus grande entre les deux mesures et à ce que les compétences visées par PISA soient davantage reliées aux acquisitions évaluées sur la base des programmes scolaires. Or, la corrélation, dans chacune des disciplines (lecture et mathématiques) est de l'ordre de +0,7; cela signifie qu'une proportion non négligeable d'élèves réalise des performances assez différentes selon l'épreuve concernée. Par ailleurs, la corrélation entre la compréhension de l'écrit (lecture) et les mathématiques dans l'enquête PISA est particulièrement marquée (supérieure à +0,80), cela témoigne d'une proximité nettement plus grande entre les deux champs de compétences de PISA que pour un même champ de compétences issu d'évaluations distinctes (PISA et ECR). La notion de compétences utiles tout au long de la vie (*life skills*), telle que le programme PISA le

conçoit, peut donc être considérée comme transversale et s'éloigne nettement d'une approche disciplinaire dans la mesure des performances des élèves.

Un second constat montre clairement que l'évaluation PISA s'avère très sélective en matière de hiérarchisation des niveaux de compétences des élèves. Dans les deux disciplines testées, la probabilité pour un élève d'être considéré comme particulièrement compétent (niveaux 5 et 6 à PISA) est faible pour la grande majorité des élèves de l'échantillon vaudois; seuls ceux qui affichent des scores très élevés (plus de 90% de réussite aux items) peuvent espérer faire partie de l'élite à PISA.

Un troisième constat a trait à l'influence différenciée du milieu social de l'élève sur ses performances scolaires selon la mesure utilisée. Ainsi, cette influence est plus faible avec les évaluations cantonales qu'avec PISA surtout dans le domaine des mathématiques. Les compétences visées par PISA laissent donc davantage s'exprimer les inégalités sociales de réussite que les acquisitions de nature scolaire liées aux programmes d'enseignement.

Les différentes analyses exposées au cours de cette recherche mettent donc en évidence des différences assez marquées entre PISA et les ECR, quant à leur capacité à rendre compte des caractéristiques des élèves avec leur contexte de scolarisation. C'est sans doute au niveau de l'influence du milieu social que la distinction entre les deux formes d'évaluation est la plus nette. Les épreuves cantonales de référence mettent moins à jour les inégalités dues à l'environnement familial de l'élève que PISA. Ce résultat sur l'influence différenciée des facteurs socio-économiques est important dans la mesure où c'est le caractère plus ou moins inégalitaire d'un système scolaire qui peut être apprécié avec une intensité variable, selon la mesure sur laquelle on se base pour l'évaluation des compétences et connaissances des élèves.

Le lien modéré, dans un même domaine, entre les évaluations locales et PISA interroge quant à lui sur le degré de responsabilité des systèmes d'enseignement dans les classements produits par l'enquête internationale. Peut-on, sur la base des résultats à PISA inférer directement le degré de qualité de l'école, alors que les évaluations des acquis scolaires ne reflètent que partiellement les compétences mesurées par PISA ?

Les résultats présentés dans cette recherche, compte tenu du faible nombre de travaux directement comparables, peuvent être considérés comme exploratoires et il serait sans aucun doute utile d'observer si les tendances observées dans le canton de Vaud se vérifient dans d'autres contextes régionaux et nationaux. Le fait de disposer d'informations comparatives sur ce sujet permettrait de s'interroger plus largement sur ce que PISA mesure vraiment et sur sa capacité à rendre compte de la qualité des systèmes d'enseignement. Cela permettrait aussi aux décideurs de disposer d'autres éléments pour instruire la politique éducative que le seul classement entre les régions et les pays, comme de lui donner une moindre importance (Suchaut, 2013).

## 8 RÉFÉRENCES

- Afonso N., Costa E. (2009). The influence of the Programme for International Student Assessment (PISA) on policy decision in Portugal: the education policies of the 17th Portuguese Constitutional Government. *Sísifo. Educational Sciences Journal*, n° 10, p. 53-64.
- Alegre M. A., Ferrer G. (2010). School regimes and education equity: some insights based on PISA 2006. *British Educational Research Journal*, Vol. 36, n° 3, p. 433-461.
- Bautier E., Crinon J., Rayou P., Rochex J. Y. (2006). Performances en littéracie, mode de faire et univers mobilisés par les élèves; analyses secondaires de l'enquête PISA 2000. *Revue française de pédagogie*, n° 157, p. 85-101.
- Bodin A. (2003). *Ce qui est vraiment évalué par PISA en mathématiques. Ce qui ne l'est pas. Un point de vue français.* [http://www.apmep.asso.fr/IMG/pdf/Com\\_PISA\\_FF\\_Francais-2.pdf](http://www.apmep.asso.fr/IMG/pdf/Com_PISA_FF_Francais-2.pdf)
- Bottani N., Vrignaud P. (2005). *La France et les évaluations internationales*. Paris: Ministère de l'Éducation nationale.
- Broi A. M., Moreau J., Soussi A., Wirthner M. (2003). Les compétences en littératie: rapport thématique de l'enquête PISA 2000. Neuchâtel: OFS; Berne: CDIP. (Monitorage de l'éducation en Suisse). 137 p.
- Crahay M., Monseur C. (2008). Composition académique et sociale des établissements, efficacité et inégalités scolaires: une comparaison internationale. *Revue française de pédagogie*, n° 164, p. 55-65.
- Demeuse M., Baye A. (2008). Indicateurs d'équité éducative. Une analyse de la ségrégation académique et sociale dans les pays européens. *Revue française de pédagogie*, n° 165, p. 91-103.
- Duru-Bellat M., Mons N., Suchaut B. (2004). Caractéristiques des systèmes éducatifs et compétences des jeunes de 15 ans: l'éclairage des comparaisons entre pays. *Cahiers de l'Irédu*, n° 66.
- Duru-Bellat M., Suchaut B. (2005). Organisation and context, efficiency and equity of educational systems. What PISA tells us. *European Educational Research Journal*, 4 (3), p. 181-194.
- Egelund N. (2008). The value of international comparative studies of achievement - a Danish perspective. *Assessment in Education: Principles, Policy & Practice*, Vol. 15, n° 3, p. 245-251.
- Felouzis G. (2003). La ségrégation ethnique au collège et ses conséquences. *Revue Française de Sociologie*, 44, 3, p. 413-447.
- Felouzis G. (2009). Systèmes éducatifs et inégalités scolaires: une perspective internationale. *Sociologies*.
- Felouzis G., Charmillot S. (2012). Les enquêtes PISA. Que-sais-je ? P.U.F., 128 p.
- Felouzis G., Charmillot S., Fouquet-Chauprade B. (2011). Les inégalités scolaires en Suisse et leurs déclinaisons cantonales: l'apport de l'enquête Pisa 2003. *Swiss Journal of Sociology*, 37(1), 33-55.
- Green A. (2008). Le modèle de l'école unique, l'égalité et la chouette de Minerve. *Revue française de pédagogie*, n° 164, p. 15-26.
- Grenet J. (2008). PISA: une enquête bancaire? *La vie des idées*.

- Goldstein H. (2008). Comment peut-on utiliser les études comparatives internationales pour doter les politiques éducatives d'informations fiables? *Revue française de pédagogie*, n° 164, p. 69-76.
- Kaiser C., Moreau J., Nidegger C. (2011). Résultats généraux de la Suisse romande in Nidegger (coord.) 2011: PISA 2009: Compétences des jeunes romands. Résultats de la quatrième enquête PISA auprès des élèves de 9ème année. p.37-58.
- Kuiper W., van den Akker J. (2011). A curricular analysis of PISA outcomes and trends 2003-2009: the Netherlands case. Berlin: ECER 2011.
- Lafontaine D., Soussi A., Nidegger C. (2009). Evaluations internationales et/ou épreuves nationales: tensions et changement de pratiques in Mottier Lopez L. & Crahay M. (éd.), *Evaluations en tension : entre la régulation des apprentissages et le pilotage des systèmes* Bruxelles: De Boeck, p. 61-80.
- Martens K., Niemann D. (2010). Governance by comparison: How ratings & rankings impact national policy-making in education. Bremen : Univ., Sonderforschungsbereich 597 Staatlichkeit im Wandel.
- Meuret D., Morlaix S. (2006). Origine sociale et performances scolaires. *Revue française de sociologie*, Vol. 47, n° 1, p. 49-79.
- Mons N. (2004). De l'école unifiée aux écoles plurielles: Évaluation internationale des politiques de différenciation et de diversification de l'offre éducative. Thèse de doctorat. Dijon: Université de Bourgogne, Sciences de l'éducation.
- Moreau J. (1981). Diverses modalités de cotation d'épreuves collectives. Avantages et inconvénients. Lausanne: Centre Vaudois de Recherches Pédagogiques.
- Moreau, J. (2004). Compétence et facteurs de réussite au terme de la scolarité. Analyse des données vaudoises de PISA 2000. Lausanne: URSP 04.3
- Moreau, J. (2007). Compétences et contexte des élèves vaudois lors de l'enquête PISA 2003. Comparaison entre cantons, filières et types d'élèves. Lausanne: URSP 07.1
- Moreau, J. (2012). Les élèves vaudois face aux sciences et à l'environnement. Analyse des résultats cantonaux à l'enquête PISA (2006). Lausanne, URSP, 12.2.
- Moreau J. (2013). Analyse des données vaudoises de l'enquête PISA 2009. URSP, (à paraître).
- Moreau J., Nidegger C., Soussi A. (2006). Définition des compétences, choix méthodologiques et retombées sur la politique scolaire en Suisse. *Revue française de pédagogie*, n° 157, p. 43-53.
- Morlaix S., Suchaut B. (2007). Evolution et structure des compétences des élèves à l'école élémentaire et au collège: une analyse empirique des évaluations nationales. *Les cahiers de l'IREDU*, n° 68, mai 2007, 259 p.
- Nidegger C., Moser U., Angelone D., Brühwiler C., Buccheri G, Abt Gürber N., Mariotta M., Moreau J., Roos E. (2011). PISA 2009: résultats régionaux et cantonaux. Berne et Neuchâtel: OFFT/CDIP et Consortium PISA.ch.
- Ntamakiliro L., Longchamp A. L. (2009). Compétences en lecture et en écriture en sixième année. Analyse de copies d'élèves aux épreuves cantonales de référence. Lausanne : URSP, 141.
- Ntamakiliro L., Tessaro W. (2010). Vers plus d'épreuves externes: une standardisation de l'évaluation scolaire? in P. Gilliéron Giroud et L. Ntamakiliro (éds.), *Réformer l'évaluation scolaire: mission impossible ?* Berne: Lang, p. 165-185.

- Ntamakiliro L., Moreau J. (2011). Analyse des résultats des élèves vaudois aux épreuves cantonales de référence, tendances générales, différences individuelles et disparité entre établissements. URSP, 150.
- OCDE (1999). Mesurer les compétences des élèves. Un nouveau cadre d'évaluation.
- OCDE (2004). Apprendre aujourd'hui pour réussir demain – premiers résultats de PISA 2003.
- OCDE (2006). Compétences en sciences, lecture et mathématiques: Le cadre d'évaluation de PISA 2006. OCDE 2006. <http://www.oecd.org/pisa/pisaproducts/38378898.pdf>
- OCDE (2010). Résultats du PISA 2009: Synthèse. OCDE, 2010. <http://www.oecd.org/pisa/46624382.pdf>
- OCDE (2012). Le cadre d'évaluation de PISA 2009: Les compétences clés en compréhension de l'écrit, en mathématiques et en sciences, PISA, Éditions OCDE. <http://dx.doi.org/10.1787/9789264075474-fr>
- Phillips G. W. (2007). Chance Favors the Prepared Mind: Mathematics and Science Indicators for Comparing States. American Institutes for research. November 14, Washington, DC.
- Planche A. (2012). L'imposture scolaire. La destruction organisée de notre système éducatif par la doctrine des (in)compétences. Presses universitaires de Bordeaux, 253 p.
- Prenzel M., Zimmer K. (2006). Études complémentaires de PISA 2003 en Allemagne: principaux résultats et enseignements. *Revue française de pédagogie* 157.
- Rautalin M., Alasuutari P. (2009). The uses of the national PISA results by Finnish officials in central government. *Journal of Education Policy*, Vol. 24, n° 5, p. 539-556.
- Shiel, G., Kellaghan T. & Moran G. (2010). Standardized testing in lower secondary education. Dublin: National Council for Curriculum and Assessment.
- Suchaut, B. (2013). PISA à l'épreuve des notes. Une autre lecture de la comparaison entre les pays. Renens: URSP.
- Vrignaud P.(2006). La mesure de la littéracie dans PISA: la méthodologie est la réponse, mais quelle était la question ? *Revue française de pédagogie*, n° 157, p. 27-41.

## 9 LISTE DES TABLEAUX, GRAPHIQUES ET SCHÉMAS

Schéma 1 : Structures des analyses .....	12
Tableau 1 : Principales caractéristiques individuelles des élèves de l'échantillon .....	13
Tableau 2: Indicateur socio-économique et culturel (IESCS) .....	13
Graphique 1 : Distribution des scores en lecture à PISA 2009 .....	14
Graphique 2 : Distribution des scores en mathématiques à PISA 2009.....	15
Graphique 3 : Distribution des scores en sciences à PISA 2009 .....	15
Graphique 4 : Distribution des scores de français aux ECR.....	16
Graphique 5 : Distribution des scores de mathématiques aux ECR .....	16
Tableau 3 : Pourcentages de points obtenus aux ECR de 8 <sup>e</sup> (2008) .....	17
Tableau 4 : Corrélations entre les scores de lecture, mathématiques et sciences (PISA 2009) ...	18
Graphique 6 : Relation entre les scores en mathématiques et en lecture (PISA 2009).....	18
Graphique 7 : Relation entre les scores en mathématiques et en sciences (PISA 2009).....	19
Graphique 8 : Relation entre les scores en lecture et en sciences (PISA 2009).....	19
Graphique 9 : Relation entre les scores en lecture et en mathématiques aux ECR .....	20
Tableau 5 : Scores moyens à PISA et aux ECR selon les caractéristiques des élèves.....	21
Graphique 10 : Différences de scores entre groupes d'élèves (en % d'écart type) .....	22
Tableau 6 : Corrélations entre l'indice socio-économique et culturel et les scores aux épreuves .....	22
Graphique 11 : Relation entre l'indice socio-économique et culturel de l'élève et ses performances en lecture (PISA 2009) .....	23
Graphique 12 : Relation entre l'indice socio-économique et culturel de l'élève et ses performances à l'ECR de français .....	24
Graphique 13 : Relation entre l'indice socio-économique et culturel de l'élève et ses performances à l'ECR de mathématiques .....	24
Graphique 14 : Impact du milieu socio-économique sur les performances en lecture et mathématiques (PISA et ECR) .....	25
Tableau 7 : Modèles d'analyse de la variance des scores à PISA en fonction des caractéristiques des élèves.....	26
Tableau 8 : Modèles d'analyse de la variance des scores des ECR en fonction des caractéristiques des élèves.....	26
Graphique 15 : Effets nets de l'indice socio-économique sur les performances à PISA et aux ECR .....	26
Graphique 16 : Effet de l'indice de statut socio-économique sur le score en mathématiques (ECR) .....	27
Graphique 17 : Distribution des scores aux ECR de français selon la filière.....	28
Graphique 18 : Distribution des scores aux ECR de mathématiques selon la filière.....	28
Graphique 19 : Distribution des scores à PISA en lecture selon la filière .....	29
Graphique 20 : Indice socio-économique et filière fréquentée .....	30
Tableau 9 : Pourcentage de variance des scores expliquée par l'IESCS par filière .....	30

Tableau 10 : Modèles d'analyse de la variance des scores obtenus à PISA en fonction des caractéristiques des élèves et de la filière fréquentée.....	31
Tableau 11 : Modèles d'analyse de la variance des scores obtenus aux ECR en fonction des caractéristiques des élèves et de la filière fréquentée.....	32
Tableau 12 : Répartition du nombre d'élèves de l'échantillon par voie et par établissement .....	33
Tableau 13 : Décomposition de la variance des scores des évaluations (PISA et ECR).....	34
Graphique 21 : Décomposition de la variance des scores aux évaluations (PISA et ECR).....	34
Graphique 22 : Relation entre les scores en français (ECR) et de lecture (PISA 2009).....	38
Graphique 23 : Relation entre les scores de mathématiques (ECR et PISA 2009).....	38
Tableau 14 : Valeurs des quartiles des distributions des scores de PISA et des ECR.....	38
Tableau 15 : Répartition des élèves en fonction des quartiles des distributions des scores en lecture à PISA et aux ECR (en %).....	39
Tableau 16 : Répartition des élèves en fonction des quartiles des distributions des scores en mathématiques à PISA et aux ECR (en %) .....	39
Tableau 17 : Niveaux de compétences en compréhension de l'écrit à PISA 2009.....	40
Tableau 18 : Répartition des élèves de l'échantillon et de l'OCDE selon le niveau de compétences en lecture à PISA 2009 (en %) .....	41
Tableau 19 : Modèle logistique analysant la probabilité d'obtenir un score de niveau 4 à PISA en lecture et en fonction du score à l'ECR de français .....	41
Graphique 24 : Probabilité d'atteindre le niveau 4 à PISA en lecture en fonction du score obtenu à l'ECR de français .....	42
Tableau 20 : Modèle logistique analysant la probabilité d'obtenir un score de niveau 5 ou 6 à PISA en lecture et en fonction du score à l'ECR de français.....	43
Graphique 25 : Probabilité d'atteindre les niveaux 5 ou 6 à PISA en lecture en fonction du score obtenu à l'ECR de français.....	43
Tableau 21 : Niveaux de compétences en culture mathématique, PISA 2009 .....	44
Tableau 22 : Répartition des élèves de l'échantillon et de l'OCDE selon le niveau de compétences en culture mathématique à PISA 2009 (en %).....	44
Tableau 23 : Modèle logistique analysant la probabilité d'obtenir un score de niveau 5 ou 6 à PISA en mathématiques et en fonction du score à l'ECR de mathématiques.....	45
Graphique 26 : Probabilité d'atteindre les niveaux 5 ou 6 à PISA en mathématiques en fonction du score obtenu à l'ECR de mathématiques.....	45
Graphique 27 : Illustration des résidus de la régression des scores de PISA sur les scores aux ECR .....	46
Graphique 28 : Distribution des résidus de la régression entre les scores aux évaluations de lecture (PISA et ECR).....	47
Tableau 24 : Modèles d'analyse de la variance des résidus de la régression entre les scores des évaluations (PISA et ECR) en fonction des caractéristiques des élèves .....	47
Tableau 25 : Modèles d'analyse de la variance des résidus de la régression entre les scores des évaluations (PISA et ECR) en fonction des caractéristiques des élèves et de la filière fréquentée.....	48
Tableau 26: Domaines et compétences de l'ECR de français de 8 <sup>e</sup> année en 2008 .....	50
Tableau 27 : Corrélations entre les domaines de l'ECR de français.....	51

Graphique 29 : Pourcentages de variance expliquée du score PISA de lecture par les domaines de compétences de l'ECR de français .....	51
Schéma 2 : Cadre d'évaluation de PISA 2009 en compréhension de l'écrit .....	51
Tableau 28 : Modèles de régression «stepwise» d'analyse de la variance du score de PISA en lecture en fonction des domaines de compétences des ECR de français .....	52
Tableau 29 : Domaines et compétences de l'ECR de mathématiques de 8 <sup>e</sup> année en 2008 .....	53
Tableau 30 : Corrélations entre les domaines de l'ECR de mathématiques .....	53
Graphique 30 : Pourcentages de variance expliquée du score PISA de mathématiques par les domaines de compétences de l'ECR de mathématiques .....	54
Tableau 31 : Modèles de régression «stepwise» d'analyse de la variance du score de PISA en mathématiques en fonction des domaines de compétences de l'ECR de mathématiques .....	54
Schéma 3 : Les composantes du domaine PISA 2009 de mathématiques.....	55
Tableau 32 : Modèles d'analyse de la variance des scores des domaines de l'ECR de français en fonction des caractéristiques des élèves.....	56
Tableau 33 : Modèles d'analyse de la variance des scores des domaines de l'ECR de mathématiques en fonction des caractéristiques des élèves et de la filière fréquentée.....	56

## ABRÉVIATIONS ET ACRONYMES

CDIP	Conférence suisse des directeurs cantonaux de l'instruction publique / Conférence des directeurs d'instruction publique
DFJC	Département de la formation, de la jeunesse et de la culture (Vaud)
DGEO	Direction générale de l'enseignement obligatoire (DFJC)
ECR	épreuves cantonales de référence (Canton de Vaud)
ESCS	(index of) Economic, Social and Cultural Status (PISA)
HarmoS	Harmonisierung der obligatorischen Schule / (projet d'harmonisation de la scolarité obligatoire)
IESCS	[voir supra ESCS]
LEO	loi sur l'enseignement obligatoire (Vaud)
LS	[ancienne] loi scolaire (Vaud)
NAEP	National Assessment of Educational Progress
OCDE	Organisation de coopération et de développement économiques
OFFT	Office fédéral de la formation professionnelle et de la technologie
PISA	Programme international pour le suivi des acquis des élèves (OCDE) / Programme for International Student Assessment
TIMSS	Trends in International Mathematics and Science Study
VG	voie générale (LEO)
VP	voie pré-gymnasiale (LEO)
VS	voie secondaire de baccalauréat (LS)
VSG	voie secondaire générale (LS)
VSO	voie secondaire à options (LS)



