



HAL
open science

Bref retour cliométrique sur 50 ans de performances scolaires en lecture et en mathématiques en France : 1970-2020

Nadir Altinok, Claude Diebolt

► **To cite this version:**

Nadir Altinok, Claude Diebolt. Bref retour cliométrique sur 50 ans de performances scolaires en lecture et en mathématiques en France : 1970-2020. *Revue d'économie politique*, 2024, 134 (1), pp.81-104. 10.3917/redp.341.0081 . hal-04796412

HAL Id: hal-04796412

<https://hal.science/hal-04796412v1>

Submitted on 25 Nov 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Bref retour cliométrique sur 50 ans de performances scolaires en lecture et en mathématiques en France : 1970-2020

Nadir ALTINOK¹ et Claude DIEBOLT²

Résumé

Cette contribution vise à retracer l'évolution de longue période des compétences scolaires des élèves français de 1970 à 2020. Partant des résultats fournis par les enquêtes internationales sur les acquis des élèves, nous présentons, à partir d'une base de données inédite, l'originalité de la trajectoire nationale vis-à-vis des autres pays de l'OCDE. Notre approche est historique et comparative. En effet, plus que la variation absolue, c'est aux écarts relatifs vis-à-vis de l'OCDE que nous mesurons la performance française. Au cours des 50 dernières années, nous montrons que, pour le cas de la France, les performances en lecture et en mathématiques augmentent avant d'entamer une stagnation voire une diminution (les scores moyens étant significativement faibles). De manière générale, nous observons une croissance quasi-généralisée de la performance scolaire sur les décennies 1970, 1980 et 1990, même si les taux de croissance diffèrent entre pays. Par ailleurs, nous nous interrogeons sur la trajectoire française en termes d'acquis scolaires qui, pas à pas, tend à diverger de celle des autres pays de l'OCDE. Ce faisant, nos résultats enrichissent les conclusions des travaux issus des enquêtes nationales tout en livrant de nouveaux éléments de preuve associés à des éclairages historiques et comparatifs renouvelés.

Mots-clés : Base de données, Cliométrie, Edumétrie, Qualité de l'éducation, Enquêtes internationales sur les acquis des élèves.

Classification JEL : C8, I2, J24, N3, O15

Brief Cliometric Looking back of 50 years of school performance in reading and mathematics in France: 1970-2020

Abstract

The aim of this contribution is to trace the long-term evolution of French pupils' academic skills from 1970 to 2020. Based on the results of international surveys of pupil achievement, we present the originality of France's trajectory in relation to other OECD countries, using a previously unpublished database. Our approach is both historical and comparative. Indeed, rather than absolute variation, we measure French performance in terms of relative deviations from the OECD. Over the last 50 years, we show that, in the case of France, performance in reading and mathematics increases before stagnating or even declining (average scores being significantly low). Generally speaking, we observe a quasi-generalized growth in school performance over the 1970s, 1980s and 1990s, even if growth rates differ between countries. We also examine the French trajectory in terms of educational achievement, which is gradually diverging from that of other OECD countries. In so doing, our findings enrich the conclusions drawn from national surveys, while providing new evidence combined with renewed historical and comparative insights.

Keywords: Databases, Cliometrics, Edumetrics, Quality of education, International surveys of pupil achievement.

¹ Nadir ALTINOK, BETA/CNRS, Université de Lorraine & Association Française de Cliométrie. Email: nadir.altinok@univ-lorraine.fr (auteur correspondant)

² Claude DIEBOLT, BETA/CNRS, Université de Strasbourg & Association Française de Cliométrie. Email: cdiebolt@unistra.fr

1. Introduction

En 1989, Christian Baudelot et Roger Establet publiaient *Le niveau monte*. Ils y affirmaient qu'il ne faut pas s'alarmer quant à la capacité du système éducatif français à faire progresser les élèves (Baudelot et Establet, 1989). Près de vingt ans plus tard, les conclusions des auteurs sont plus timorées et soulignent l'importance accordée aux acquis scolaires, notamment aux résultats de l'évaluation PISA (*Programme for International Student Assessment*), peu glorieux pour l'hexagone (Baudelot et Establet, 2009).

Certes, la scolarisation universelle est atteinte pour quasiment tous les jeunes entre 6 et 16 ans. Il serait cependant trompeur de se focaliser sur la seule expansion scolaire. La scolarisation n'équivaut pas à l'apprentissage (Pritchett, 2013). Qui plus est, depuis plusieurs décennies, il est convenu d'accepter qu'une année de scolarisation n'apporte pas le même niveau de compétences, quel que soit le pays. Le gain en termes de compétences scolaires est le plus souvent basé sur des évaluations standardisées sur les acquis des élèves. La plus connue, PISA, a mis en avant les faiblesses du système éducatif français. Après quelques années de critiques vis-à-vis d'évaluations supposément non adaptées au contexte français (Mons et Pons, 2013, Cédelle, 2015), force est de constater que l'école française n'est pas la meilleure école du monde. Quant à la désaffection des candidats aux fonctions d'enseignement, quel que soit le niveau scolaire, elle témoignerait d'un malaise au sein du système éducatif français.

Mais l'expansion scolaire a-t-elle vraiment conduit à une diminution des acquis scolaires en France ?

Récemment, un rapport de la DEPP (2022) montre une baisse continue, depuis 30 ans, du niveau en mathématiques. Il reprend et standardise les scores recueillis dans l'ensemble des enquêtes qui évaluent le niveau en mathématiques par le biais d'une échelle de scores. Martin *et al.* (2022) reprennent les résultats de cette étude et l'approfondissent pour souligner un déclassement de la France vis-à-vis des systèmes éducatifs d'autres pays de l'OCDE, avec un niveau général qui semble diminuer, ou tout du moins se stabiliser, et une montée des inégalités à travers un tassement dans les résultats des moins performants.

Notre contribution se situe en prolongement de ces travaux. En recourant à une base de données inédite, construite à partir des enquêtes internationales sur les acquis des élèves, nous comparons l'évolution de la performance, depuis 1970, du niveau des élèves français en mathématiques, sciences et en lecture. Notre approche est historique et comparative. En effet, alors que la plupart des études analysent l'évolution de la performance de la France dans le temps, nous privilégions une approche relative et mesurons les écarts de variation de la France vis-à-vis des autres pays de l'OCDE. Cette comparaison nous permet de mieux mettre en lumière les écarts de performance de la France par rapport à l'OCDE, malgré une hausse apparente de la performance sur le long terme.

2. Données et méthodologie

L'originalité de notre travail réside dans l'actualisation et dans l'emploi d'une base de données contenant des scores comparables dans plusieurs domaines de compétences scolaires, à différents niveaux et sur une période couvrant environ cinq décennies (1970-2020)³.

Pour y parvenir, nous avons dû surmonter deux difficultés majeures : la faible participation des pays à des enquêtes sur les acquis des élèves – notamment avant les années 1990 – mais aussi la grande hétérogénéité des données disponibles depuis les années 1970 (cf. Filmer *et al.*, 2020, Angrist *et al.*, 2021).

2.1. Les évaluations sur les acquis des élèves

Notre base de données internationale sur les acquis des élèves mobilise les résultats issus de diverses enquêtes (TIMSS, PIRLS et PISA⁴). D'autres évaluations se focalisent sur les compétences des adultes (IALS, PIAAC⁵).

Dès 1960, sous l'égide de l'Association Internationale pour l'Evaluation du Rendement scolaire (IEA) une enquête pilote est menée pour 12 pays dont la France, ceci afin d'évaluer les compétences en mathématiques, lecture, géographie, sciences et compétences non verbales (Foshay *et al.*, 1962). Par la suite, plusieurs évaluations ont testé les élèves du primaire et du secondaire en mathématiques et en sciences (*First International Mathematics Study*, *Second International Mathematics Study*, *First International Science Study* et *Second international Science Study*). L'évaluation TIMSS débute en 1995 et se focalise sur les mathématiques et les sciences. Depuis cette date, plusieurs vagues de tests se sont déroulés avec une périodicité de 4 ans⁶. En complément, l'IEA a aussi évalué les compétences en lecture des élèves et ce dès les années 1970. La première étude principale a été menée en 1970 et a inclus 14 pays (*Six Subject Survey: Reading Comprehension*). C'est toutefois l'enquête PIRLS qui a mis en avant les compétences des jeunes en lecture. L'évaluation mesure les compétences des élèves du grade 4 (équivalent CM1 en France) tous les 5 ans depuis 2001⁷. Parallèlement aux enquêtes de l'IEA, c'est l'évaluation PISA qui a suscité un grand écho médiatique, notamment en Europe ; PISA étant une évaluation développée par l'OCDE. Elle a démarré en 2000. Testant les élèves de 15 ans dans trois domaines de compétences (mathématiques, sciences et lecture), PISA est devenue l'enquête recouvrant le plus de pays dans le monde. Organisée tous les 3 ans, sept différentes vagues ont été effectuées

³ Pour une description détaillée, voir notamment Altinok et Diebolt (2024).

⁴ Respectivement *Trends on International Mathematics and Science Study*, *Progress on International Reading Literacy Study* et *Programme for International Student Assessment*.

⁵ Respectivement *International Adult Literacy Survey* et *Programme for the International Assessment of Adult Competencies*.

⁶ La deuxième vague date de 1999, puis 2003, 2007... La dernière vague avec des données disponibles date de 2019.

⁷ Les différentes vagues de tests sont 2001, 2006, 2011, 2016 et 2021. Les données de PIRLS 2021 n'étaient pas disponibles lors de la rédaction de cet article.

jusqu'en 2018 avec plus de 70 pays participants lors de la dernière vague. En parallèle des évaluations sur les acquis des élèves, les enquêtes sur les compétences des adultes (IALS et PIAAC) peuvent aussi s'avérer être très utiles pour mieux évaluer les compétences des élèves (voir le Tableau 1 pour le détail de chaque enquête). En utilisant une méthodologie spécifique, certains auteurs ont, en effet, obtenu des données de panel relatifs au niveau de compétences des jeunes adultes (Schwerdt et Wiederhold, 2019).

Tableau 1. Enquêtes internationales sur les acquis des élèves

No	Années	Organisation	Abrév.	Matière	Nombre de Pays ou Régions	Grade	Incluse
1	1959-60	IEA	Enq. pilote	M,S,L	12	7/8	
2	1964	IEA	FIMS	M	12	7/FS	■
3	1970-71	IEA	SSS-RC	L	15	4/8/FS.	
4	1970-72	IEA	FISS	S	19	4/8/FS.	■
5	1980-82	IEA	SIMS	M	19	8/FS	■
6	1983-84	IEA	SISS	S	23	4/8/ FS	■
7	1988/1990-91	NCES	IAEP	M,S	6-19	4/7-8	
8	1990-91	IEA	RLS	L	32	3-4/7-8	■
9	1995/1999/2003/2007/2011/2015/2019	IEA	TIMSS	M,S	45-38-26-48-66-65-64	3-4/7-8/ FS	■
10	1992-97	UNESCO	MLA	M,S,L	72	6-8	
11	1997/2006/2013/2019	UNESCO	LLECE	M,S,L	13-16-15-16	3-6	■
12	1999/2002/2007	UNESCO	SACMEQ	M,L	7-15-16	6	■
13	2000/2010/2014/2019	CONFEMEN	PASEC	M,L	22-22-10-14	2/5 puis 3-6	■
14	2001/2006/2011/2016	IEA	PIRLS	L	35-41-55-50	4	■
15	2000/2003/2006/2009/2012/2015/2018	OECD	PISA	M,S,L	43-41-57-74-65-71-79	15 ans	■
16	2010/2015/2019	USAID/RTI	EGRA	L	29-40-5	1 to 6	■
17	2008-2019	ASER	ASER	L	2	1 to 6	■
18	1994-1998	OCDE	IALS	L	22	Adultes	
19	2011-2017	OCDE	PIAAC	L	39	Adultes	■

Note: IEA = "International Association for the Evaluation of Educational Achievement", NCES = "National Center for Education Statistics", CONFEMEN = "Conférence des Ministres de l'Éducation des États et Gouvernements de la francophonie", OCDE = "Organisation de coopération et de développement économiques", USAID = "United States Agency for International Development", RTI = "Research Triangle Institute", ASER = "Australian Council for Educational Research", FIMS = "First International Mathematics Study", SSS-RC = "Six Subject Survey: Reading Comprehension", FISS = "First International Science Study", SIMS = "Second International Mathematics Study", IAEP = "International Assessment of Educational Progress", RLS = "Reading Literacy Study", TIMSS = "Trends on International Mathematics and Science Study", MLA = "Monitoring Learning Achievement", LLECE = "Latin American Laboratory for Evaluation of the Quality of Education", SACMEQ = "The Southern and Eastern Africa Consortium for Monitoring Educational Quality", PASEC = "Programme d'Analyse des systèmes éducatifs de la Confemén", PIRLS = "Progress in International Reading Literacy Study", PISA = "Programme of International Student Assessment", EGRA = "Early-Grade Reading Assessment", IALS = "International Adult Literacy Survey", PIAAC = "Programme for International Assessment of Adult Competencies". Matières : M=mathématiques ; S=sciences ; L= lecture. FS = "fin du secondaire, équivalent au lycée".

2.2. Méthodologie générale

Notre recherche repose sur un regroupement des différentes enquêtes internationales (présentées dans la section précédente) sur des échelles comparables. Puisque ces enquêtes sont de nature diverses, il nous a fallu élaborer une méthodologie permettant de les ajuster entre elles, ceci afin d'éviter tout biais potentiel. En effet, inutile de mentionner ici que le score obtenu, en données brutes, pour un pays aurait aisément pu être plus élevé dans une enquête par rapport à une autre, du simple fait que la première enquête, par exemple, aurait eu un niveau de difficulté supérieur. Ce faisant, si l'on comparait naïvement les résultats entre les deux enquêtes, il serait hautement probable que l'interprétation première aurait conduit à une baisse du niveau des élèves. Nous avons donc procédé à un ajustement des enquêtes, ceci afin d'envisager des comparaisons dans le temps et dans l'espace. L'idée générale qui a servi de fil conducteur à notre démarche consiste à utiliser les résultats des pays ayant participé simultanément à plusieurs enquêtes à la même période (Angrist *et al.*, 2021). En effet, en combinant l'ensemble des pays ayant participé à plusieurs évaluations, il est possible de calculer des taux de conversion, comme on le ferait pour un taux de change entre des devises par exemple. Cette méthode est utilisée pour ajuster les scores des enquêtes réalisées à partir des années 1990 et a déjà été mise en œuvre par de précédentes recherches (Angrist *et al.*, 2013, Angrist *et al.*, 2021, Gust *et al.*, 2022). Une approche additionnelle a aussi été implémentée pour ancrer les enquêtes entre elles dans le temps, et surtout avant 1990. Suivant l'approche pionnière de Hanushek et Kimko (2000), il est ainsi possible d'ancrer les données des enquêtes internationales en se référant aux résultats des Etats-Unis dans l'enquête nationale sur les acquis des élèves, à savoir le *National Assessment of Educational Progress* (NAEP)⁸.

L'évaluation NAEP a débuté dès les années 1970. Désormais, elle permet une comparaison des scores en lecture, mathématiques et sciences sur près de 50 ans. Jusqu'en 1990, les enquêtes de l'IEA ont été ancrées, en réajustant les scores, de telle sorte à correspondre aux scores des Etats-Unis dans l'enquête NAEP. Ainsi, si le score des Etats-Unis augmente de 1% entre 1970 et 1980, l'enquête de l'IEA de 1980 sera ajustée de telle manière que le score des Etats-Unis soit égal à cette même variation. L'ajustement du score des autres pays s'effectue de la même manière, mais le taux de variation de leur performance ne sera pas forcément de 1 % car l'ancrage que nous effectuons préserve en partie la variation initiale observée dans les enquêtes. Nous supposons ainsi que l'ancrage réalisé sur les scores de l'enquête NAEP est suffisamment robuste pour réévaluer le degré de difficulté de chaque enquête antérieure aux années 1990 et que la variabilité des scores est stable entre les pays participants aux enquêtes pionnières. Cette dernière hypothèse est plutôt crédible car les pays participants aux enquêtes antérieures aux années 1990 sont principalement des pays développés. Ce travail d'ancrage est toutefois facilité depuis les années 1990 puisque le recours à l'utilisation de la méthode IRT (*Item Response Theory*) permet dorénavant

⁸ Le détail de la méthodologie pour l'obtention d'une base de données regroupant plus de 130 pays est présentée dans Altinok et Diebolt (2024). Le présent article se focalise sur les résultats de la France, en comparant l'hexagone à un échantillon de 20 pays de l'OCDE. Pour mémoire, rappelons ici que la base de données créée contient bien sûr un échantillon de données plus larges et plus détaillées. D'autres travaux ont déjà utilisé cette approche, voir notamment Hanushek et Kimko, 2000, Hanushek et Woessmann, 2012, Altinok *et al.*, 2014, Hanushek et Woessmann, 2015.

d'avoir des scores comparables entre les différentes séries d'évaluation au sein des enquêtes. Aussi, il est possible de comparer les résultats de PISA 2009 avec ceux de PISA 2015 sans ajustement spécifique. Ce faisant, l'ajustement via l'enquête NAEP devient quasi inutile pour la plupart des enquêtes réalisées après 1990.

Quant à PISA et TIMSS, le calcul d'un taux de conversion général suffit pour obtenir une base de données comparable sur le plan des acquis scolaires⁹. Une nouveauté réside toutefois dans l'implémentation de la méthode d'imputation utilisée dans Lim *et al.* (2018). En recourant à des données fortement corrélées avec les acquis scolaires, notre recherche permet ainsi d'élargir à la fois la période couverte par les évaluations sur les acquis, mais aussi de prédire les scores pour des pays n'ayant pas participé à toutes ces évaluations. Cette procédure d'imputation utilise le paquet "Amelia II" (Honaker *et al.*, 2011) et s'avère utile pour obtenir des données de panel sur près d'un demi-siècle (1970-2020). En effet, la base de données préparée par Angrist *et al.* (2021) ne couvre que la période 2010-2020. Le recours à de l'imputation multiple et aux enquêtes les plus anciennes, en utilisant l'ancrage par le taux de conversion et par le biais de l'évaluation NAEP, nous permet donc d'obtenir des scores comparables pour les principaux pays de l'OCDE entre 1970 et 2020¹⁰.

De plus, afin de combiner quantité et qualité de l'éducation, nous proposons d'associer le nombre moyen d'années de scolarisation avec les scores des élèves aux évaluations sur les acquis. Filmer *et al.* (2020) ont été les premiers à effectuer une telle conversion. Cette combinaison apporte ainsi une mesure relative de la quantité d'éducation, en prenant en compte le différentiel de qualité éducative entre les pays. L'hypothèse principale se base sur l'idée qu'une année d'éducation dans un pays donné n'apporte pas nécessairement le même rendement scolaire en termes d'acquis dans un autre pays. Certains pays s'avèrent être plus efficaces que d'autres, comme le soulignent les évaluations telles que PISA. Dans cette perspective, l'indicateur hybride que nous mettons en œuvre se nomme "*Learning-Adjusted Years of Schooling*" (Filmer *et al.*, 2020). Afin de calculer cet indicateur relatif, nous divisons le score obtenu par 700 (étant considéré comme un plafond théorique), ce qui nous donne une valeur comprise entre 0 et 1. Puis, ce ratio est multiplié par le nombre moyen d'années de scolarisation, issu de la base de données de Barro et Lee (2013). La nouvelle base de données ainsi obtenue fournit des informations au sujet de trois variables complémentaires : la qualité des systèmes éducatifs (via les scores aux différentes

⁹ Supposons que la moyenne des scores des pays participant à TIMSS 1999 et PISA 2000 en mathématiques soit égale à 520 dans TIMSS et 495 dans PISA, alors le taux de conversion entre PISA et TIMSS sera de $525/495 = 1,06$. Les scores de PISA seront ainsi réévalués à la hausse d'environ 6,06 %.

¹⁰ Le détail de la procédure d'imputation multiple peut être retrouvé dans Altinok et Diebolt (2024). Contrairement à Lim *et al.* (2018), nous utilisons cette méthode d'imputation multiple en suivant différentes étapes afin de minimiser l'erreur de mesure. Plus précisément, huit étapes sont implémentées, allant de la prédiction des scores pour les pays ayant le plus d'observations sur une période courte (entre 1995 et 2020, étape numéro 1) à celle de pays ayant très peu de données sur la qualité de l'éducation et sur une période plus longue (entre 1970 et 2020, étape numéro 8). Cette procédure d'imputation utilise des facteurs explicatifs préalablement sélectionnés via une régression multiple et inclut les variables présentées dans l'annexe A (Tableau A.1). On pourra également trouver le taux d'imputation pour chaque croisement entre l'année et la dimension des acquis scolaires (Tableau A.2). Il est à noter que les pourcentages concernant les pays de l'OCDE sont moins élevés que ceux des pays en développement.

enquêtes sur les acquis et les seuils de performance), la quantité d'éducation (via le nombre moyen d'années de scolarisation) et le nombre d'années de scolarisation ajusté par la qualité de l'éducation (en anglais "*Learning-Adjusted Years of Schooling*" ou LAYS). La principale innovation de notre travail réside dans la mise à disposition de la communauté scientifique d'une base de données en panel pour un très grand nombre de pays¹¹. Les travaux antérieurs concernent soit des données en coupe instantanée (Hanushek et Kimko, 2000, Altinok et Murseli, 2007, Angrist *et al.*, 2021), soit des données de panel non cylindré (Angrist *et al.*, 2013, Altinok *et al.*, 2014), ou alors des données de panel sur une période plus restreinte (Lim *et al.*, 2018). De manière originale, nous proposons d'utiliser une méthode proche du travail de Angrist *et al.* (2021) en utilisant en partie l'approche proposée par Lim *et al.* (2018) afin d'obtenir une base de données qui a une approche similaire mais plus précise des travaux de Barro et Lee (2013) en ce qui concerne l'éducation.

De manière très concrète, les données sur la qualité de l'éducation sont réparties sur plusieurs niveaux scolaires (primaire, secondaire) et domaines de compétences (mathématiques, lecture, sciences) et disponibles entre 1970 et 2020 pour la plupart des pays de l'OCDE. En moyenne sur la période étudiée, la performance de la France est inférieure sur le plan de la qualité de l'éducation (16 points de différence), mais aussi sur la scolarisation : le nombre moyen d'années de scolarisation est de 8,6 années contre 10,2 pour la moyenne des pays de l'OCDE-20 (Tableau 2). Ces deux retards entraînent, à leur tour, un déficit d'environ 1,3 année de scolarisation ajustée par la qualité de l'éducation (indicateur LAYS).

Tableau 2. Base de données sur l'éducation (1970-2020)

	Moyenne	Ecart-type	Minimum	Maximum
OCDE-20				
Qualité de l'éducation	507	24	435	561
Années de scolarisation	10,2	2,0	3,2	13,8
Indicateur mixte (LAYS)	7,4	1,7	2,0	10,3
France				
Qualité de l'éducation	491	23	452	517
Années de scolarisation	8,6	2,1	5,5	11,3
Indicateur mixte (LAYS)	6,1	1,7	3,5	8,2

Notes : Les données regroupent toutes les observations disponibles entre 1970 et 2020 pour une sélection de 20 pays de l'OCDE. Au total 220 observations sont disponibles pour l'échantillon complet et 11 observations pour la France (données disponibles tous les 5 ans). Source : Altinok et Diebolt (2024)

3. Une évolution des acquis scolaires contrastée

Comme indiqué dans l'introduction de cet article, la principale innovation de notre travail réside dans la mise à disposition de données comparables pour la qualité de l'éducation sur près d'un demi-siècle

¹¹ Mentionnons ici que les données proposées dans nos travaux ne sont pas des données de cohorte. Elles mesurent les acquis scolaires ainsi que les taux d'achèvement d'une population donnée à un instant t, sans toutefois pouvoir suivre l'évolution des acquis scolaires de ces mêmes individus dans le temps. Notre approche étant davantage macroéconomique que microéconomique, le suivi de cohorte a, toute choses égales par ailleurs, moins d'utilité opérationnelle, puisque nous souhaitons d'abord mesurer le stock de capital humain à un instant t pour un pays donné.

(1970-2020). Notre méthode est comparative et se focalise ici sur des pays similaires à la France, les pays de l'OCDE et plus précisément les 20 pays qui, *ceteris paribus*, font le cœur de l'OCDE¹² et que nous nommons OCDE-20. Un avantage de se restreindre à ces pays est la disponibilité des données sur le long terme et le fait de placer la France autour de pays qui ont quasiment tous achevé la scolarisation primaire et secondaire universelle dès le début des années 1980. On peut ainsi éviter d'inclure des pays qui observent une sélection de leur public scolaire (comme le Mexique ou la Turquie), ce qui aurait tendance à surestimer leur performance dans les différents domaines de compétences¹³.

3.1. Une performance moindre par rapport à son niveau économique

Nous commençons par comparer la qualité de l'éducation en France en coupe instantanée pour l'année 2020. Le niveau de compétences que nous relevons pour le cas français est l'un des plus faibles des pays de l'OCDE-20, notamment pour le niveau primaire, que ce soit en lecture ou en mathématiques (Figure 1). Alors que le score moyen de la France est de 482 points en mathématiques au niveau primaire, il avoisine 549 points en Irlande. En supposant qu'une année d'acquis est similaire à environ 30 points (Avvisati et Givord, 2023), cela signifierait que la France est en retard de près de deux ans sur l'Irlande. Les résultats sont plus élevés au niveau secondaire, la France réalisant une performance qui est proche de la moyenne des pays de l'OCDE.

Le Tableau 3 présente l'évolution de la performance pour chaque domaine de compétence (mathématiques, lecture et sciences) et chaque niveau scolaire analysé (primaire et secondaire). La période étudiée s'étale sur 50 ans, de 1970 à 2020. Partant d'une moyenne internationale en 2000 qui est de 500 et d'un écart-type de 100, il apparaît que, depuis 1970, la performance scolaire en France a augmenté. Afin d'apprécier la significativité de la variation de la performance, nous avons calculé l'erreur de mesure pour chaque score moyen¹⁴. Le niveau en lecture pour le primaire est passé de 441 points à 502 points entre 1970 et 2020. Cette hausse représente environ 60 points, soit 12 points par décennie ou l'équivalent de 2 années scolaires d'acquis sur un demi-siècle. Ces variations sont presque

¹² Nous qualifions ces pays de "membres historiques de l'OCDE" et sélectionnons principalement des pays ayant de fortes similitudes avec la France : Allemagne, Australie, Autriche, Belgique, Canada, Danemark, Espagne, États-Unis, Finlande, France, Grèce, Irlande, Italie, Nouvelle-Zélande, Norvège, Pays-Bas, Portugal, Royaume-Uni, Suède et Suisse. Les pays asiatiques ont des performances plus élevées que les pays occidentaux aux enquêtes sur les acquis. Cependant, certaines spécificités culturelles pouvant expliquer ces différences (voir notamment Boman, 2022). Dans le cadre de cet article, nous avons donc préféré centrer notre attention sur les pays non asiatiques de l'OCDE.

¹³ A titre d'exemple, Spaul (2018) a montré que les résultats de la Turquie dans l'enquête PISA semblaient surestimés au début des années 2000 du fait d'une scolarisation des jeunes de 15 ans plus faible que dans les autres pays de l'OCDE. La France et la plupart des autres pays européens ne sont pas spécifiquement concernés par le problème de sélection qui viendrait biaiser la mesure de la qualité de l'éducation. Notons toutefois qu'il est possible d'observer un changement de la structure de la population, suite à des vagues d'immigration ou un éventuel sursaut des naissances, mais il n'y a aucune raison de penser que ces changements concerneraient exclusivement la France.

¹⁴ L'erreur de mesure ou "standard error" est ici calculée de telle façon à englober à la fois l'erreur de mesure inhérente à l'estimation d'une performance *macro* à partir d'une enquête *micro*, mais aussi en prenant en compte l'erreur de mesure provenant de la méthode d'imputation multiple. En supposant que les deux erreurs soient indépendantes, nous pouvons calculer l'erreur standard comme suit : $erreur\ standard = erreur\ de\ mesure_{enquête} + erreur\ de\ mesure_{imputation}$. L'erreur de mesure de l'imputation est réalisée par le biais d'une méthode de *bootstrapping* avec 30 répliquations sur les résultats issus de l'imputation, qui inclut elle-même environ 80 prédictions de scores.

toutes significatives au regard des erreurs standard calculées¹⁵. Si l'on observe la colonne relative à la variation standardisée de long terme (colonne 7), les gains sont tous positifs sur le long terme : l'école française a donc globalement amélioré sa capacité à apprendre aux élèves depuis 1970. Cette tendance haussière se retrouve dans presque tous les pays de l'OCDE, confirmant les résultats trouvés aux Etats-Unis, notamment en mathématiques¹⁶. Les gains sont les plus importants en lecture, tant au niveau primaire qu'au niveau secondaire ; ils sont moindres en mathématiques. Signalons ici que les résultats de l'enquête PISA ont, de manière générale, souligné une baisse de niveau des élèves depuis quelques années (OCDE, 2019).

Par-delà l'analyse de long terme (50 ans), nous avons également calculé les variations sur le court terme (sur les 20 dernières années). Ici, les résultats sont clairement plus contrastés et soulignent une baisse absolue de la performance dans tous les domaines et niveaux depuis 2000 (colonne 8). La performance en lecture a diminué d'environ 11 points par décennie sur les 20 dernières années, tandis que la diminution est d'environ 6 points par décennie en mathématiques dans le secondaire. Si l'on rapporte cette variation à l'équivalence en termes d'années d'acquisition, les effets sont assez importants. Alors que, sur le long terme, la France a gagné l'équivalent de plus d'une année d'acquis en lecture, une perte d'une année d'acquis scolaires a été observée sur la période 2000-2020. La baisse est moindre en mathématiques (respectivement 0,70 et 0,43 pour les niveaux primaire et secondaire).

Nous savons que le niveau des compétences scolaires est étroitement lié au niveau économique des nations (Hanushek et Woessmann, 2015). Même si la relation est moins évidente, on pourrait donc logiquement attendre de meilleurs résultats au sein des écoles les mieux dotées en ressources financières (Hanushek, 2019). Qui plus est, afin de tester dans quelle mesure la performance de la France se situe en-dessous de ce qu'on pourrait attendre au vu de son niveau économique ou de l'ampleur de ses dépenses d'éducation, nous opérons différents tests. Nous effectuons, tout d'abord, une analyse graphique des corrélations entre chaque domaine de compétence et niveau scolaire et l'indicateur de richesse économique. Plus concrètement, nous établissons une corrélation entre les performances scolaires et le PIB par habitant en dollars ajustés par la méthode des parités de pouvoir d'achat. Nous observons une corrélation assez forte, avec un coefficient proche ou supérieur à 60% dans la plupart des cas. En même temps, ce calcul met en lumière l'écart de performance de la France, notamment au niveau primaire, aussi bien en lecture qu'en mathématiques.

¹⁵ La plupart des variations de court et de long terme sont significatives au seuil de 5 %. Pour conduire le test de significativité, nous avons combiné les erreurs standard ("E.S.") et calculé le seuil en utilisant la formule suivante : $seuil = \sqrt{E.S.^2_{France} + E.S.^2_{OCDE}} \times 1,96$. Dès lors que la différence entre la France et l'OCDE était supérieure à cette valeur, nous en déduisons que la différence est significativement différente de 0 au seuil de 5 %. Les données significatives sont indiquées en gras.

¹⁶ Les résultats de l'évolution des compétences scolaires sur le long terme aux Etats-Unis peuvent être consultés en suivant ce lien : <https://www.nationsreportcard.gov/ltr/?age=9> (accès effectué le 21 juin 2023).

La Figure 2 illustre l'écart significatif de la France vis-à-vis des autres pays en mathématiques au niveau primaire¹⁷. Alors que le score escompté serait d'environ 517 points, le score réel n'est que de 482 points.

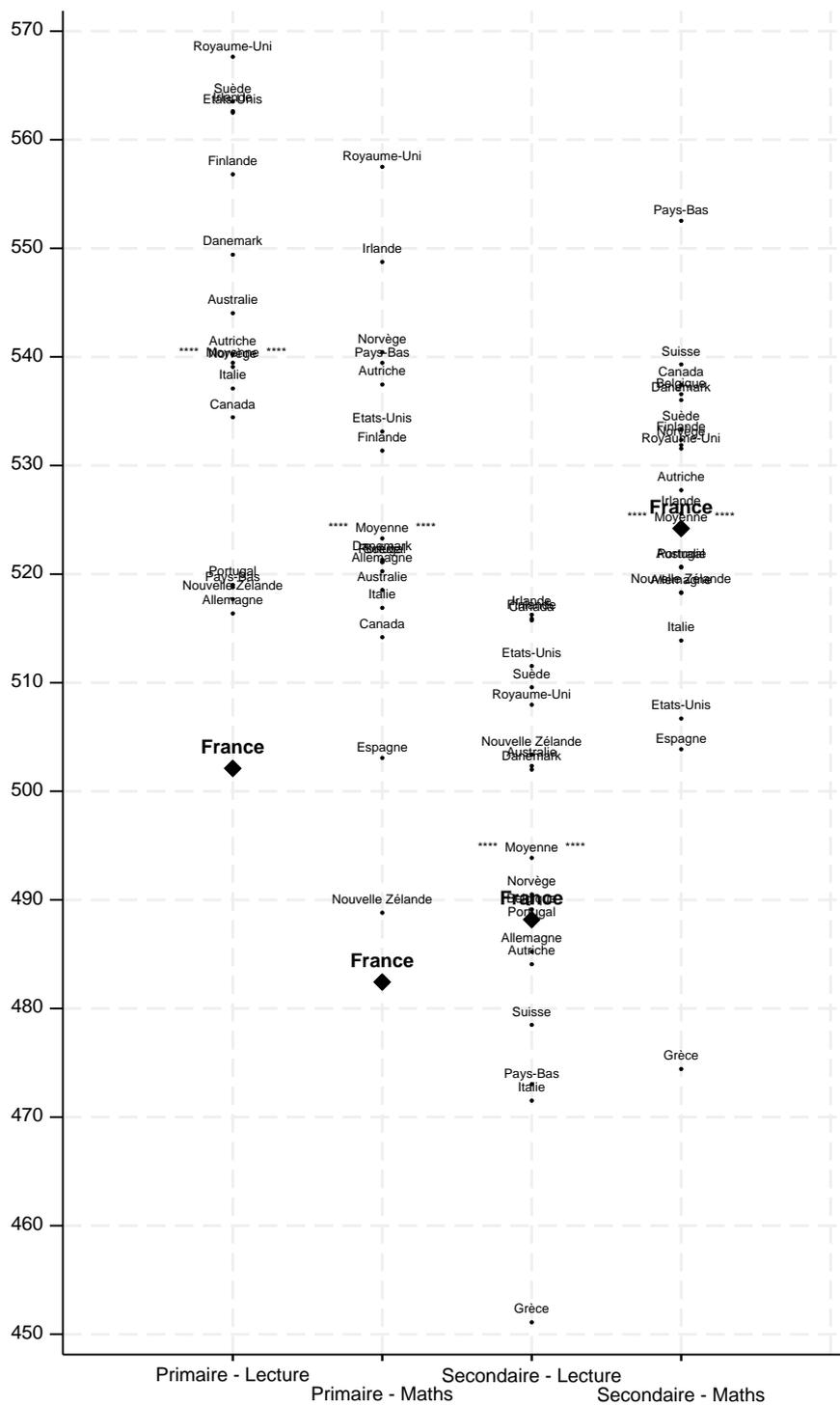
Tableau 3. Evolution de la performance scolaire et effets de productivité

	1970	1980	1990	2000	2010	2020	Variation standardisée		Effet de productivité			
									Long terme		Court terme	
							Score	Score	Score	Score	Score	Score
Moyenne	452 (6,3)	460 (2,7)	495 (2,9)	517 (3,3)	508 (3,0)	501 (2,6)	9,87	-7,95	1,64	2,47	-0,53	-0,80
Primaire	446 (3,3)	451 (3,3)	500 (3,3)	510 (3,3)	500 (3,3)	490 (3,3)	8,86	-9,57	1,48	2,22	-0,64	-0,96
Littérature	441 (2,7)	457 (2,7)	522 (2,7)	525 (2,4)	520 (2,7)	502 (2,0)	12,21	-11,29	2,03	3,05	-0,75	-1,13
Mathématiques	457(3,3)	450 (3,3)	493 (3,3)	504 (3,3)	489 (3,3)	482 (3,3)	5,04	-10,54	0,84	1,26	-0,70	-1,05
Sciences	440 (3,1)	446 (3,1)	484 (3,1)	501 (3,1)	491 (3,1)	487 (3,3)	9,34	-6,87	1,56	2,33	-0,46	-0,69
Secondaire	457 (6,3)	470 (2,7)	490 (2,9)	524 (3,3)	517 (3,0)	511 (2,6)	10,88	-6,34	1,81	2,72	-0,42	-0,63
Littérature	427 (2,7)	438 (2,7)	471 (2,7)	505 (2,7)	499 (3,2)	488 (2,2)	12,19	-8,28	2,03	3,05	-0,55	-0,83
Mathématiques	488 (6,3)	490 (2,7)	516 (2,9)	537 (3,3)	523 (3,0)	524 (2,6)	7,23	-6,38	1,21	1,81	-0,43	-0,64
Sciences	455 (2,7)	481 (2,7)	482 (2,7)	530 (3,8)	529 (3,4)	521 (2,5)	13,21	-4,36	2,20	3,30	-0,29	-0,44

Notes : Les douze premières colonnes indiquent les scores standardisés de la France dans notre base de données. La variation standardisée représente ainsi la variation absolue de la performance pour chaque dimension sur une base décennale. Les chiffres entre parenthèses indiquent les erreurs standard d'estimation (E.S.). Voir texte pour le détail des calculs. Les données de long terme concernent la période 1970-2020, tandis que les données de court terme sont relatives à la période 2000-2020. Les données en gras soulignent une différence significative de l'évolution de la performance française (voir note de bas de page n°15). Les quatre dernières colonnes reprennent les résultats de l'OCDE sur les gains potentiels obtenus par une année de scolarisation. Nous considérons deux effets extrêmes, en utilisant les résultats de l'Allemagne. La limite faible est calculée à partir d'un effet de 20 points pour chaque année de scolarisation, tandis que la limite haute est basée sur une hypothèse de 30 points. Voir Avvisati et Givord (2023) pour les analyses de productivité. Source des données : Altinok et Diebolt (2024)

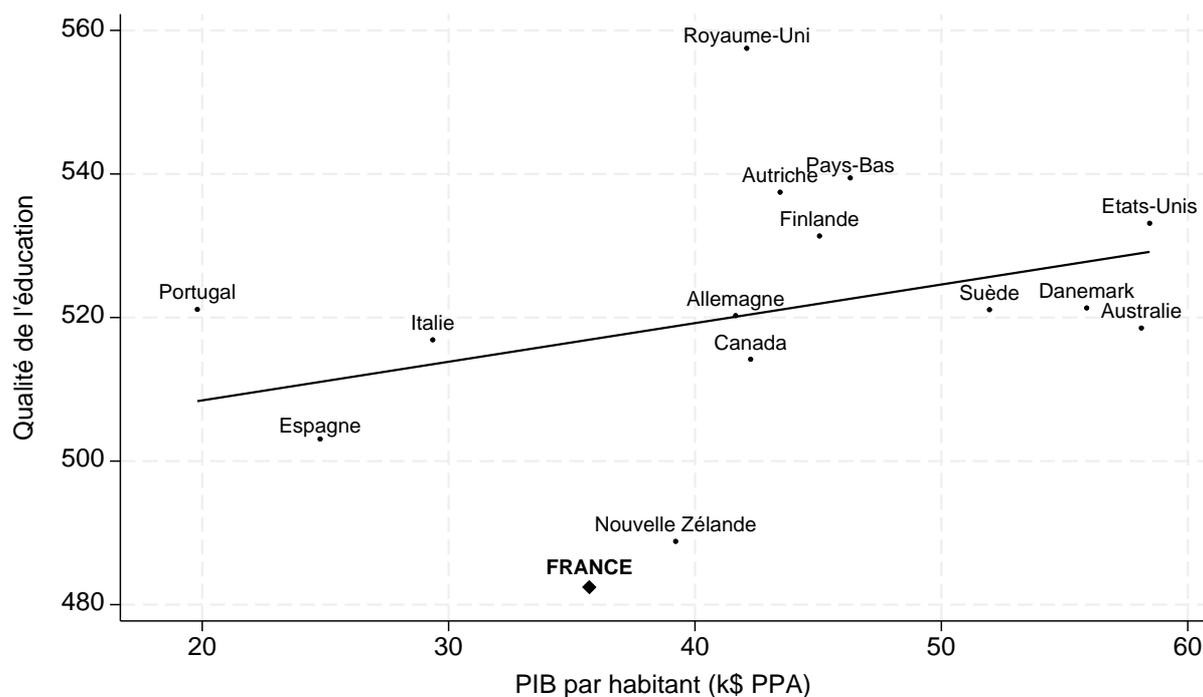
¹⁷ Cette analyse a été effectuée pour chaque domaine de compétence et chaque niveau scolaire. L'ensemble des résultats est disponible sur demande auprès des auteurs.

Figure 1. Performance moyenne des pays de l'OCDE en 2020



Notes : Seuls 20 pays de l'OCDE sont utilisés pour établir ce classement. Les scores représentent une moyenne des résultats des pays à chaque domaine de compétence (mathématiques et lecture) pour l'année 2020. La méthodologie utilisée pour construire les données est présentée en section 2. Pour des raisons de clarté, les résultats obtenus pour le Japon sont omis. Source : Altinok et Diebolt (2024)

Figure 2. Niveau économique et scores moyens en mathématiques au primaire



Notes : Cette figure relie le niveau économique des pays (mesuré par le PIB par habitant, exprimé en milliers de dollars des Etats-Unis, ajusté par la méthode des parités de pouvoir d'achat [PPA]) avec la qualité de l'éducation, mesurée à partir des résultats standardisés en mathématiques, au niveau primaire. Source : Altinok et Diebolt (2024).

Afin de mieux appréhender ce potentiel écart de performance, nous calculons une régression de la performance sur le niveau économique des pays. Au-delà du coefficient associé à cette dernière variable, c'est le résidu attaché à chaque pays qui retient notre attention (Tableau 4).

Dans chacune des régressions, nous associons une dimension particulière des compétences éducatives avec le PIB par habitant. Puisque les données sont en panel, cela nous permet d'obtenir des effets fixes pays qui ne sont pas expliqués par l'indicateur de richesse économique. Sur les 20 pays inclus dans l'échantillon¹⁸, le classement de la France est souvent en dernière position, soulignant le score anormalement faible lorsque l'on prend en compte son niveau économique. En effet, rappelons ici que le classement des résidus permet de mesurer ce qui n'est pas pris en compte dans le modèle, ici le niveau économique.

Dans presque toutes les régressions, c'est la Finlande qui présente le résidu le plus élevé. Ce pays, souvent mis en lumière dans les études PISA, affiche en effet une performance bien supérieure à ce qu'il devrait avoir en tenant compte de son niveau économique. Pour la France, c'est le cas inverse qui se produit. A titre d'illustration, les résultats de la régression entre les mathématiques et le niveau économique laissent supposer une sous-performance d'environ 14 points (colonne 3), une fois le niveau économique pris en compte. Comme souligné précédemment (Avvisati et Givord, 2023), cette différence

¹⁸ Pour certaines régressions, le nombre de pays est inférieur à 20.

renvoie à l'équivalent de la moitié d'année de scolarisation perdue pour la France. Notons toutefois que même si la « sous-performance » de la France n'est apparue que dans les années 2000, elle semble être latente depuis au moins 1970. En effet, le classement des résidus révèle des résultats assez proches si l'on compare le court terme et le long terme (voir les lignes "Période 1970-2020" et "Période 2000-2020"). Notons, par ailleurs, que ces différences de résidus sont basées sur des effets fixes et contrôlées par le niveau économique. Il nous semble donc hautement probable que les explications de l'écart de la France reposent sur des causes plus structurelles que conjoncturelles.

Une approche similaire avec un indicateur alternatif – les dépenses éducatives exprimées en pourcentage du PIB par habitant – permet d'aboutir à des conclusions similaires. En effet, lorsque l'on compare des pays selon leur niveau de dépenses d'éducation, la France accuse un retard vis-à-vis des autres pays de l'OCDE dans le niveau primaire, tandis que la situation est proche du score escompté au niveau secondaire¹⁹.

Tableau 4. Régression de la performance scolaire avec le niveau économique des pays

	(1)	(2)	(3)	(4)	(5)	(6)
PIB/hab	18,867*** (3,205)	23,903*** (2,111)	24,053*** (3,417)	23,279*** (3,293)	18,814*** (2,728)	21,521*** (2,058)
Constante	425,518*** (10,362)	424,870*** (7,465)	424,966*** (11,151)	417,161*** (10,449)	428,934*** (9,081)	426,864*** (6,778)
Niveau	Primaire	Secondaire	Tous	Tous	Tous	Tous
Matière	Toutes	Toutes	Maths	Sciences	Littérature	Tous
R ²	0,344	0,540	0,298	0,348	0,643	0,324
Observations	573	659	408	408	416	1232
Nb pays	19	20	20	20	20	20
Période 1970-2020						
Classement	17	18	18	18	18	18
Résidu	-15,31	-12,29	-14,52	-19,18	-7,50	-13,93
Meilleur classement	Finlande	Finlande	Finlande	Finlande	Belgique	Finlande
Période 2000-2020						
Classement	17	18	17	17	19	18
Résidu	-17,61	-9,14	-12,53	-17,09	-9,54	-13,27
Meilleur classement	Finlande	Finlande	Finlande	Finlande	Finlande	Finlande

Notes : Nous effectuons une régression multiple du niveau économique des pays sur la performance scolaire en données de panel, en incluant des variables indicatrices pour les années. Les résidus pour chaque pays sont ensuite calculés et ordonnés. Le classement des résidus est effectué selon deux périodes distinctes : 1970-2020 et 2000-2020. * significatif à 10%, ** significatif à 5%, *** significatif à 1%. Les erreurs-types sont présentées entre parenthèses.

3.2. Une convergence de la performance éducative inachevée

Les divers résultats des évaluations nationales laissent toutes supposer une diminution de la performance en mathématiques depuis près d'une décennie (Chabanon et Pastor, 2019). Afin de

¹⁹ Ces résultats ne sont pas publiés dans cet article, mais disponibles sur demande auprès des auteurs.

confirmer ou d'infirmer cette tendance, nous nous sommes focalisés sur l'évolution de la performance en mathématiques au niveau primaire, depuis 1970, pour les principaux pays de l'OCDE.

Nous remarquons qu'un net décrochage semble effectivement s'opérer en France depuis le début des années 2000, alors que son classement était déjà faible pour les périodes antérieures (voir Tableau 1). Une baisse de la performance est également observée au niveau secondaire sur les vingt dernières années après une hausse assez forte depuis les années 1970. De manière plus significative, il apparaît que la France était nettement au-dessus de la moyenne de l'OCDE-20 en 2000, tandis qu'en 2020, son classement retrouve un niveau plus proche de la moyenne.

Précédemment, nous avons observé une hausse tendancielle sur le long terme de la performance scolaire en France, mais aussi une baisse sur les deux dernières décennies pour tous les domaines de compétence (section 3.1). Il nous semble important de questionner ces variations au regard des autres pays de l'OCDE-20 (Tableau 5). Pour ce faire, nous calculons pour chaque domaine de compétences, ainsi que pour chaque année, l'écart entre la performance de la France et celle de la moyenne des 20 pays de l'OCDE inclus dans notre échantillon. La différence est standardisée et exprimée en termes d'écart-type. Puisque notre base de données a un écart-type de 100, le calcul effectué pour la standardisation s'effectue en divisant l'écart de performance par 100. Alors que la France a un score de 457 en mathématiques au primaire en 1970, contre une moyenne de 500 points pour l'OCDE, l'écart de performance est de 43 points, soit 0,43 écart-type en défaveur de la France. De manière quasi-systématique, la performance de la France est presque toujours inférieure à celle de la moyenne de l'OCDE, sauf en mathématiques au niveau secondaire pour les décennies 1980 et 1990. La significativité des écarts de moyenne peut être approchée en estimant le seuil à partir duquel l'écart deviendrait significatif. Le seuil est calculé en cumulant les erreurs standards et en les pondérant par le seuil d'erreur de 5 %²⁰. Ainsi, dans le Tableau 5, les écarts relatifs de performance qui sont significatifs au seuil de 5 % sont indiqués en gras.

Nous calculons également la variation de la performance pour le court (20 années) et le long terme (50 années). Les quatre dernières colonnes de notre Tableau 5 soulignent que la variation de la performance française est assez proche de la moyenne de l'OCDE sur le long terme (deux dernières colonnes), même si les progrès sont légèrement moindres en mathématiques. Sur le court terme, la différence de trajectoire est toutefois nette : alors que la performance diminue dans tous les domaines en France, la variation est plus contrastée au sein des pays de l'OCDE. Le niveau primaire s'améliore en moyenne pour les pays de l'OCDE d'environ 6 points pour décennie, alors qu'une baisse de 10 points est observée en France. Nous constatons que les écarts de performance de la France sont plus importants au niveau primaire, par rapport au niveau secondaire. La recherche de l'explication de ces différences selon le niveau scolaire demanderait des recherches supplémentaires qui sortent du cadre de notre étude.

²⁰ Pour conduire le test de significativité, nous avons combiné les erreurs standard ("E.S.") et calculé le seuil en utilisant la formule suivante : $seuil = \sqrt{E.S.^2_{France} + E.S.^2_{OCDE}} \times 1,96$. Dès lors que la différence entre la France et l'OCDE était supérieure à cette valeur, nous en déduisons que la différence est significativement différente de 0 au seuil de 5 %.

Cependant, il est possible d'émettre quelques hypothèses qui nous paraissent crédibles comme le fait d'avoir un système éducatif primaire relativement défaillant vis-à-vis du niveau secondaire, une allocation des enseignants du primaire non optimale (notamment dans le domaine des mathématiques), mais aussi un effet "enquêtes" propre aux données disponibles : comme la France participe davantage aux enquêtes PISA plutôt qu'aux évaluations de l'IEA, l'écart moindre observé au secondaire renverrait à la différence de l'approche utilisée dans les enquêtes. PISA se focalisant davantage sur les compétences, tandis que l'IEA privilégie les connaissances basées sur des curriculums communs.

Dans la Figure 2, nous affinons la variation de la performance scolaire. Nous posons l'hypothèse que les pays de l'OCDE ayant un déficit de compétence initial auront une hausse plus soutenue de leur performance dans les décennies suivantes. Cette hypothèse de convergence peut déjà être en partie vérifiée grâce au Tableau 5, puisque le retard relatif de la France s'amenuise pour la plupart des domaines de compétence²¹. Pour vérifier cette intuition, nous croisons le niveau initial des mesures éducatives avec la variation de celles-ci sous forme standardisée (i.e. exprimée en écart-types). L'hypothèse de convergence est plutôt confirmée, la pente de la droite de régression étant négative dans la totalité des quadrants. De manière plus précise, rappelons que nous testons ici la convergence de trois indicateurs de l'éducation : un indicateur quantitatif de l'éducation, à savoir le nombre d'années de scolarisation ; l'indicateur d'acquis scolaires et enfin l'indicateur d'années de scolarisation ajusté par le niveau des acquis scolaires (que nous appelons LAYS, voir section 2.2).

Tableau 5. Evolution comparative de la performance scolaire

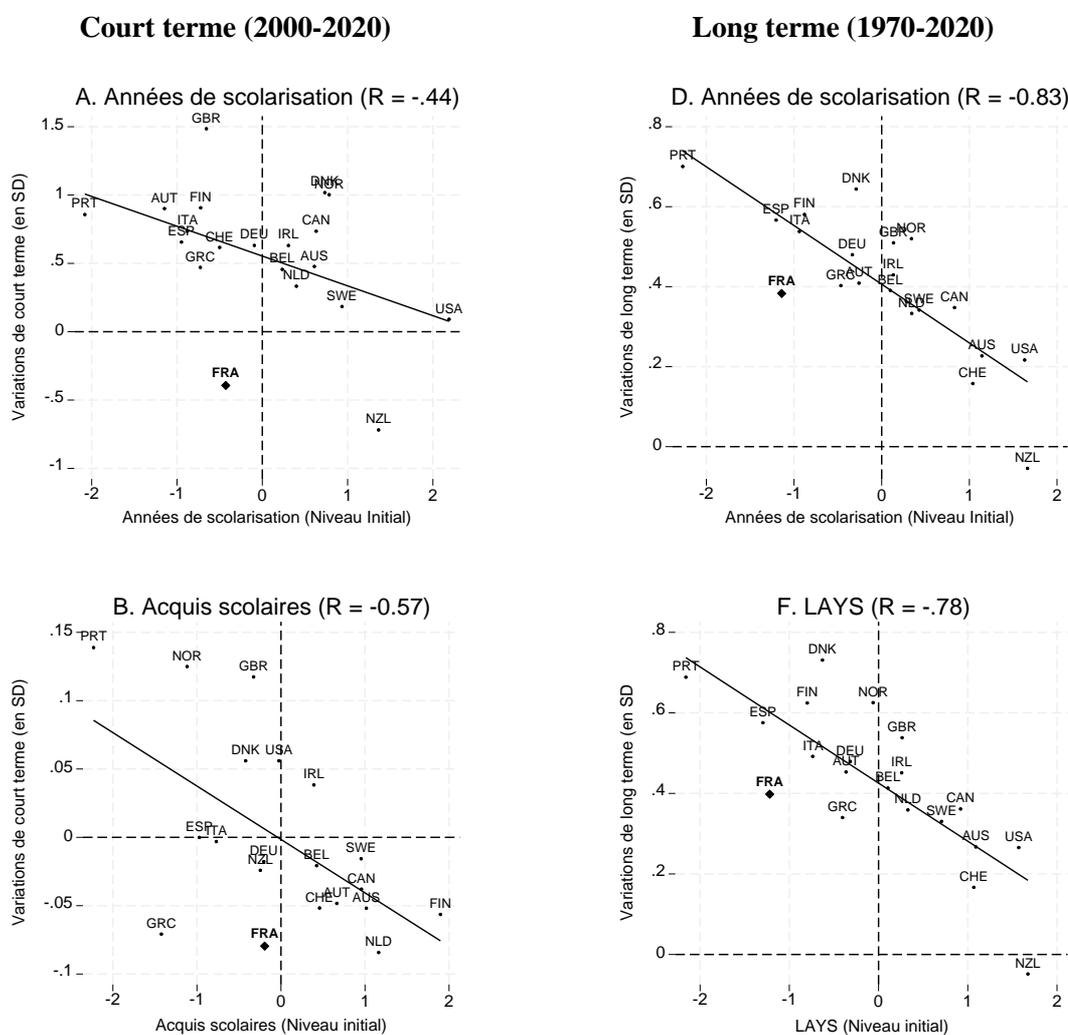
	1970	1980	1990	2000	2010	2020	Valeur initiale		Valeur finale		Variation court terme		Variation long terme	
							France	OCDE	France	OCDE	France	OCDE	France	OCDE
Moyenne	-0,36	-0,29	-0,06	-0,03	-0,17	-0,19	452 (6,3)	488 (3,4)	501 (2,6)	520 (2,5)	-8	0	10	6
Primaire	-0,41	-0,38	0,02	-0,07	-0,26	-0,38	446 (3,3)	487 (4,7)	490 (3,3)	528 (2,5)	-10	6	9	8
Lecture	-0,37	-0,36	0,10	-0,08	-0,18	-0,37	441 (2,7)	478 (2,8)	502 (2,0)	539 (2,0)	-11	4	12	12
Mathématiques	-0,43	-0,33	0,09	-0,01	-0,27	-0,41	457 (3,3)	500 (4,8)	482 (3,3)	523 (2,5)	-11	10	5	5
Sciences	-0,45	-0,45	-0,16	-0,14	-0,32	-0,36	440 (3,1)	485 (3,7)	487 (3,3)	522 (2,7)	-7	4	9	7
Secondaire	-0,30	-0,18	-0,13	0,02	-0,06	-0,02	457 (6,3)	486 (3,4)	511 (2,6)	513 (2,5)	-6	-4	11	5
Lecture	-0,29	-0,26	-0,10	-0,01	-0,02	-0,06	427 (2,7)	456 (2,8)	488 (2,2)	494 (2,4)	-8	-6	12	8
Mathématiques	-0,17	-0,07	0,07	0,10	-0,06	0,00	488 (6,3)	505 (3,4)	524 (2,6)	524 (2,5)	-6	-1	7	4
Sciences	-0,43	-0,21	-0,34	-0,02	-0,10	0,00	455 (2,7)	498 (3,2)	521 (2,5)	522 (2,7)	-4	-5	13	5

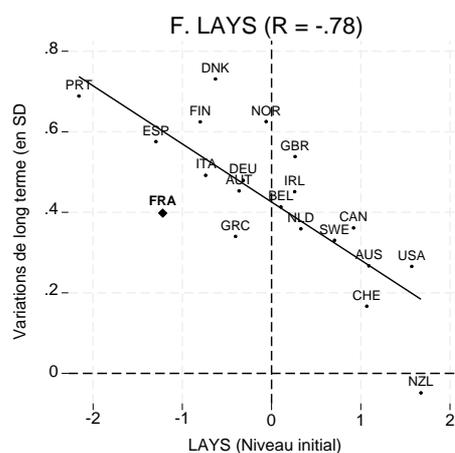
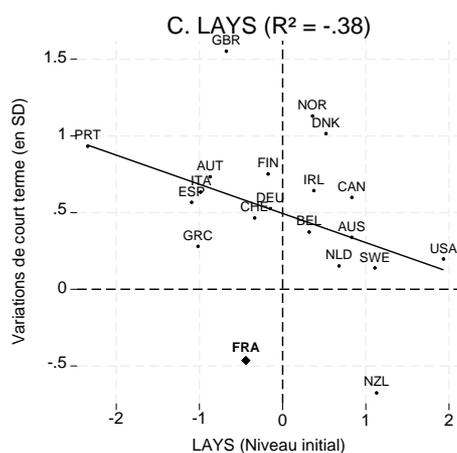
Notes : L'évolution est exprimée en termes d'écart-type de différence de la performance de la France par rapport à la moyenne des 20 pays de l'OCDE inclus dans l'échantillon. On constate qu'en 1970 la performance de la France était de 26 points inférieure à celle des 20 pays de l'échantillon en lecture au niveau primaire (424 contre 450 points). Comme l'écart-type est de 100 points, cela représente une différence de performance d'environ 0,26 écart-type. Les données en gras soulignent une différence significative entre la performance française et celle de l'OCDE avec une erreur de 5 %. Pour conduire le test de significativité, nous avons combiné les erreurs standard ("E.S.") et calculé le seuil en utilisant la formule suivante : $seuil = \sqrt{E.S.^2_{France} + E.S.^2_{OCDE}} \times 1,96$. Dès lors que la différence entre la France et l'OCDE était supérieure à cette valeur, nous en déduisons que la différence est significativement différente de 0 au seuil de 5 %. Les valeurs entre parenthèses correspondent aux erreurs standards ("E.S.") calculées en combinant les erreurs de mesure obtenues des enquêtes et celles calculées en effectuant l'imputation multiple via *bootstrapping* (voir note de bas de page n°14 pour plus d'informations).

²¹ Ce rattrapage peut s'abord s'expliquer par un pur effet de convergence, mais d'autres raisons pourraient l'expliquer. La recherche des facteurs explicatifs demanderait un travail plus minutieux sur les variables entrant dans la fonction de production éducative et la production de modèles dits multiniveaux.

La France est marquée par un retard initial dans toutes les dimensions (elle se situe dans la partie négative de l'axe des abscisses). Sous l'hypothèse de convergence, un tel retard devrait conduire à la hausse la performance éducative. Or, cela n'est pas systématiquement le cas, surtout pour les vingt dernières années (2000-2020). A titre d'exemple, le retard français en ce qui concerne les années de scolarisation est loin d'être comblé entre 2000 et 2020, puisque la scolarisation semble diminuer d'environ 0,4 écart-type en France, si l'on prend en compte la variation moyenne des pays de l'OCDE (quadrant A). La performance quantitative ne semble donc pas être à la hauteur de celle observée dans les autres pays de l'OCDE. De manière additionnelle, le retard observé dans le niveau des acquis scolaires en début de période ne semble se résorber que si l'on utilise les données de long terme (valeur des ordonnées positive dans les quadrants D, E et F). Qui plus est, ce retard relatif dans les acquis scolaires semble se renforcer dans le temps depuis les années 2000 (ordonnée négative dans le quadrant B). Notre comparaison entre niveau initial et variation au sujet de l'indicateur hybride LAYS renforce ces résultats : le faible niveau initial est loin de se résorber et a même tendance à se creuser sur le court terme (quadrant C).

Figure 3. Variation de la performance des indicateurs d'éducation





Notes : Le niveau initial représente la valeur de l'indicateur au début de la période indiquée (1970 ou 2000). Les variations sont exprimées en écart-type (*Standard Deviation* = SD) calculés à partir de l'hypothèse que l'écart-type de la base de données est égal à 100 points. Ainsi, un écart de 50 points est équivalent à 0,5 d'écart-type. Source : Altinok et Diebolt (2024).

4. Conclusion

Les décisions d'allocation de ressources à l'éducation reposent, presque toujours, sur des choix politiques tiraillés entre des exigences socio-culturelles d'une part, et des exigences économiques d'autre part. Ces choix composent avec d'autres choix portant sur d'autres affectations des ressources nationales. Le caractère concurrentiel de ces choix politiques permet de penser, qu'à un moment ou à un autre, les charges économiques liées à l'éducation rencontrent un plafond. Par ailleurs, au plan économique, on peut admettre, presque intuitivement, l'existence d'une limite au-delà de laquelle développer l'éducation risque de se faire au détriment des possibilités mêmes de la croissance économique. De nos jours, cette limite économique fait, plus que jamais peut-être, figure d'évidence. Cela dit, il n'est pas certain qu'elle ait été pleinement assimilée. Quoi qu'il en soit, nous soutenons l'hypothèse que l'éducation joue un rôle important et positif dans l'activité économique de la France. Ce faisant, l'éducation est elle-même un processus utilisateur de ressources économiques. Ainsi, les pressions démographiques et socio-politiques qui, depuis la fin de la Seconde Guerre mondiale, ont entraîné un développement considérable des systèmes éducatifs ont également atteint un niveau tel que l'éducation figure en bonne place parmi les parties prenantes aux ressources nationales. Dès lors, l'éducation est soumise à la contrainte économique majeure, celle de la rareté relative des ressources susceptibles d'emplois alternatifs. En fait, l'éducation est soumise à une exigence d'efficacité. Cela dit, il ne s'agit nullement d'ériger en principe le caractère utilitariste de l'éducation. Nous pensons uniquement que quelles que soient les finalités assignées à l'éducation, il convient d'utiliser au mieux les ressources disponibles pour les satisfaire.

Comme tel, l'allocation des ressources n'est que partiellement réalisée par des mécanismes d'ajustements par les prix traduisant la rencontre d'une offre et d'une demande sur un marché. Parallèlement, les décisions de production y sont plus ou moins décentralisées et les objectifs particuliers ne rencontrent pas nécessairement l'objectif collectif. Enfin, les besoins en termes d'éducation sont socialement hétérogènes, ainsi que dans le temps et dans l'espace. En ces termes, l'absence de signal par les prix peut affecter l'efficacité interne (à production donnée les ressources ne sont pas nécessairement

affectées au moindre coût réel) et l'efficacité externe (les produits ne sont pas toujours adaptés aux besoins).

Afin d'appréhender correctement production et coût de l'éducation, il est donc nécessaire d'élaborer avec précision des indicateurs permettant de décrire ce que l'on produit et avec quels moyens, ceci afin de définir des indices d'efficacité comparée (dans le temps et dans l'espace). Ceci étant, même si l'on suppose le problème des indicateurs résolu, se pose la question de l'élaboration d'outils d'analyse pertinents, susceptibles de fournir une évaluation qualitative de la productivité des dépenses publiques d'éducation. D'une manière générale, nous pensons qu'il est quasi impossible de définir des indicateurs d'efficacité absolus. En revanche, il paraît tout à fait concevable de définir des indicateurs relatifs mesurant des écarts entre quantité et qualité d'éducation. Ce faisant, le problème n'en est pas pour autant résolu, car il conviendra de développer une méthode d'analyse permettant de distinguer les écarts d'efficacité dus à des phénomènes exogènes, de ceux résultant de comportements stratégiques ou enfin, de ceux imputables à de mauvaises allocations de ressources qui favoriseraient la quantité au détriment de la qualité de l'éducation.

Dans le cadre de cet article, notre recherche, consacrée à l'évolution des acquis scolaires en France, a produit de nouveaux éléments de preuve, sur un demi-siècle et en comparaison avec la plupart des pays de l'OCDE, pour aboutir à la conclusion que les performances éducatives, en lecture et en mathématiques diminuent, les scores moyens étant significativement faibles. Par ailleurs, nous avons montré que la trajectoire française en termes d'acquis scolaires tend à diverger de celle des autres pays de l'OCDE. D'une manière générale, nos résultats enrichissent les conclusions des travaux issus des enquêtes nationales tout en livrant des éclairages historiques et comparatifs renouvelés. Enfin, ils invitent à réinvestir, par-delà les considérations centrées sur des indicateurs de pure quantité, l'importance majeure de la dimension qualitative de l'éducation.

Références

- Altinok N., Diebolt C. [2024], Cliometrics of learning-adjusted years of schooling: evidence from a new dataset, *Cliometrica*, 18(3), available online first.
- Altinok N., Diebolt C., Demeulemeester J.-L. [2014], A new international database on education quality: 1965–2010, *Applied Economics*, 46(11), 1212-1247.
- Altinok N., Murseli H. [2007], International database on human capital quality, *Economics Letters*, 96(2), 237-244.
- Angrist N., Djankov S., Goldberg P.K., Patrinos H.A. [2021], Measuring human capital using global learning data, *Nature*, 592(7854), 403-408.
- Angrist N., Patrinos H.A., Schlotter M. [2013], An expansion of a global data set on educational quality: a focus on achievement in developing countries, *The World Bank: Policy Research Working Papers*, November 15.
- Avvisati F., Givord P. [2023], The learning gain over one school year among 15-year-olds: an international comparison based on PISA, *Labour Economics*, 84, 102365.
- Barro R.J., Lee J.W. [2013]. A new data set of educational attainment in the world, 1950-2010, *Journal of Development Economics*, 104, 184-198.

- Baudelot C., Establet R. [1989], *Le niveau monte. Réfutation d'une vieille idée concernant la prétendue décadence de nos écoles*. Paris, Seuil.
- Baudelot C., Establet R. [2009], *L'élitisme républicain. L'école française à l'épreuve des comparaisons internationales*. Paris, Seuil.
- Boman B. [2022] Vietnam's exceptional educational achievement: a thematic review of the emerging literature, *Discover Education*, 1(14), 1-16.
- Cédelle L. [2015], PISA à travers la presse en France, *Administration & Education*, 145(1), 63-67.
- Chabanon L., Pastor J.-M. [2019], L'évolution des performances en calcul des élèves de CM2 à trente ans d'intervalle (1987-2017), *Note de la DEPP*, 19(08), 1-4.
- DEPP [2022], *L'état de l'école 2022. Une analyse statistique du système éducatif*, Paris, Direction de l'évaluation, de la prospective et de la performance.
- Filmer D., Rogers H., Angrist N., Sabarwal S. [2020], Learning-adjusted years of schooling (LAYS): Defining a new macro measure of education, *Economics of Education Review*, 77, 101971.
- Foshay A.W., Thorndike R.L., Hotyat F., Pidgeon D.A., Walker D.A. [1962]. *Educational Achievements of Thirteen-year Olds in Twelve Countries: Results of an International Research Project, 1959-61*, Hamburg, Unesco Institute for Education.
- Gust S., Hanushek E.A., Woessmann L. [2022], Global universal basic skills: current deficits and implications for world development, *National Bureau of Economic Research Working Paper Series*, 30566.
- Hanushek E.A. [2019] Addressing Cross-National Generalizability in Educational Impact Evaluation, *National Bureau of Economic Research, Working Paper Series*, 25460.
- Hanushek E.A., Kimko D.D. [2000], Schooling, labor-force quality, and the growth of nations, *American Economic Review*, 90(5), 1184-1184.
- Hanushek, E.A., Woessmann L. [2012], Do better schools lead to more growth? Cognitive skills, economic outcomes, and causation, *Journal of Economic Growth*, 17(4), 267-321.
- Hanushek, E.A., Woessmann L. [2015], *The knowledge capital of nations: Education and the economics of growth*, Cambridge (MA), MIT Press.
- Honaker J., King G., Blackwell M. [2011], Amelia II: A program for missing data, *Journal of Statistical Software*, 45(1), 1-47.
- Lim S.S., Updike R. L., Kaldjian A. S., Barber R.M. and co. [2018], Measuring human capital: a systematic analysis of 195 countries and territories, 1990-2016, *The Lancet*, 392(10154), 1217-1234.
- Martin R., Renault T., Roux B. [2022], Baisse de la productivité en France : échec en «maths»? *Focus*, Paris, Conseil d'Analyse Economique, 091-2022.
- Mons N., Pons X. [2013], Pourquoi n'y a-t-il pas eu de « choc PISA » en France ? Sociologie de la réception d'une enquête internationale (2001-2008), *Revue française de pédagogie*, 182, 9-18.
- OCDE [2019], *PISA 2018 Results (Volume I)*. Paris, OCDE.
- Pritchett L. [2013], *The rebirth of education: schooling ain't learning*. Washington (D.C.), Centre for Global Development.
- Schwerdt G., Wiederhold S. [2019], A Macroeconomic Analysis of Literacy and Economic Performance, *Mimeo*.
- Spaull N. [2018], Who makes it into PISA? Understanding the impact of PISA sample eligibility using Turkey as a case study (PISA 2003-PISA 2012), *Assessment in Education: Principles, Policy & Practice*, 1-25.

Annexe A. Informations supplémentaires sur l'imputation des données

Tableau A.1. Variables utilisées dans le cadre de l'imputation multiple

Variable	Niveau scolaire	Années disponibles	% de données manquantes	% imputation
Groupe 1 – Variables Proxy avec beaucoup de valeurs manquantes				
% des enseignants qui sont des femmes	Tous	2000-2020	20	41
% des enseignants qualifiés	Tous	2000-2020	62	34
% des enseignants ayant suivi une formation pédagogique	Tous	2000-2020	62	25
Groupe 2 - Variables Proxy avec peu de valeurs manquantes				
Taux de redoublement	P – SI	1970-2020	21	45
Ratio enseignants/élèves	Tous	1970-2020	21	35
Dépenses du gouvernement par élève (en % du PIB par habitant)	Tous	1970-2020	18	69
Dépenses du gouvernement à l'éducation (en % des dépenses du gouvernement)	-	1970-2020	23	50
Dépenses du gouvernement à l'éducation (en % du PIB)	-	1970-2020	21	45
Groupe 3 – Variables Proxy en lien avec l'achèvement scolaire				
Taux d'achèvement (%)	P – S	1970-2020	11	50
Taux brut de scolarisation (%)	P – S	1970-2020	10	29
Taux de survie jusqu'au dernier grade du niveau primaire (%)	P	1970-2020	6	64
Taux net de scolarisation (%)	P – S	1970-2020	11	52
Taux d'enfants en dehors du système éducatif pour chaque groupe d'âge (%)	P – S	1970-2020	11	53
Taux net de scolarisation ajusté (%)	P	1970-2020	11	55
Taux de scolarisation (%)	Tous	1970-2020	6	52
Nombre moyen d'années scolaires, groupe d'âge 15-64 ans	Tous	1970-2020	6	27
PIB par habitant	-	1970-2020	13	10
Groupe 4 – Mesures alternatives aux acquis scolaires				
Données sur la littératie dans les pays en développement	P	1970-2005	77	8
Taux d'alphabétisation des adultes	Adultes	1950-2015	0	0

Notes : Les niveaux d'éducation sont « primaire » (P), « secondaire » (S) et « supérieur » (T). Le degré d'imputation est faible pour la variable de littératie, car les données sont manquantes pour tous les pays développés et ne concernent que la période 2005-2020. Les sources et définitions des variables peuvent être trouvées dans Altinok et Diebolt (2024).

Tableau A.2. Part des données imputées pour les acquis scolaires

	01 MOY	02 PRI	03 SEC	04 MATH	05 LECT	06 SCIE	07 PRI+M	08 PRI+L	09 PRI+S	10 SEC+M	11 SEC+L	12 SEC+S
1970	0,83	0,86	0,74	0,81	0,82	0,76	0,82	0,85	0,76	0,74	0,68	0,74
1975	0,84	0,88	0,74	0,81	0,83	0,76	0,84	0,88	0,78	0,74	0,68	0,74
1980	0,80	0,89	0,69	0,77	0,79	0,71	0,84	0,88	0,78	0,68	0,62	0,69
1985	0,73	0,81	0,58	0,69	0,72	0,61	0,74	0,80	0,64	0,58	0,50	0,58
1990	0,70	0,72	0,67	0,66	0,68	0,58	0,64	0,70	0,53	0,67	0,60	0,67
1995	0,51	0,55	0,50	0,44	0,48	0,38	0,43	0,52	0,34	0,50	0,42	0,50
2000	0,42	0,51	0,34	0,33	0,38	0,29	0,38	0,48	0,32	0,34	0,23	0,34
2005	0,31	0,37	0,25	0,21	0,28	0,19	0,23	0,33	0,20	0,23	0,15	0,25
2010	0,11	0,22	0,10	0,07	0,11	0,04	0,15	0,18	0,11	0,07	0,03	0,10
2015	0,08	0,12	0,10	0,07	0,07	0,03	0,08	0,10	0,03	0,07	0,01	0,10
2020	0,20	0,24	0,06	0,32	0,33	0,19	0,22	0,20	0,07	0,06	0,01	0,05
Total	0,49	0,55	0,42	0,44	0,48	0,37	0,46	0,53	0,36	0,41	0,33	0,42

Notes : "Moy" = moyenne des deux niveaux scolaires et des trois domaines de compétence. "PRI" = niveau primaire, "SEC" = niveau secondaire, "MATH" or "M" = mathématiques, "LECT" or "L" = lecture, "SCIE" or "S" = sciences.