

Construction d'un indice de position sociale des élèves

Thierry Rocher

▶ **To cite this version:**

Thierry Rocher. Construction d'un indice de position sociale des élèves. Éducation & formations, 2016, 90, pp.5-27. 10.48464/hal-01350095 . hal-01350095

HAL Id: hal-01350095
https://hal.science/hal-01350095

Submitted on 29 Jul 2016

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.



Distributed under a Creative Commons Attribution - NoDerivatives 4.0 International License

CONSTRUCTION D'UN INDICE DE POSITION SOCIALE DES ÉLÈVES

Thierry Rocher
MENESR-DEPP, bureau de l'évaluation des élèves.

L'origine sociale des élèves est le plus souvent appréhendée par la profession et catégorie sociale (PCS) de leurs parents. L'idée de construire un indice dérivé de la PCS, qui puisse mesurer la position sociale des élèves répond à deux besoins issus des études statistiques des performances scolaires des élèves. En premier lieu, la PCS des parents est certainement la variable la plus utilisée dans le champ des études sur les inégalités sociales à l'école, mais elle peut montrer certaines limites à remplir son rôle de repérage central des disparités de réussite scolaire. Cet indice aurait pour vocation de synthétiser davantage de dimensions (sociales, économiques, culturelles). En second lieu, le profil social d'ensembles plus larges que le simple individu se laisse difficilement appréhender à l'aide de la nomenclature des groupes socioprofessionnels. Une mesure quantitative permet de passer plus facilement du niveau de l'élève à celui de la classe ou de l'établissement scolaire. Un indice peut ainsi servir d'instrument pour la mesure des effets de contexte qui sont au cœur des problématiques actuelles de la sociologie de l'éducation. En outre, il peut donner une mesure plus fiable et plus robuste de la mixité sociale dans les établissements. En réponse à ces deux préoccupations, la solution envisagée et présentée dans cet article est la suivante : une transformation de la PCS des parents en un indice de position sociale.

Les résultats tirés des évaluations nationales [Arzoumanian et Dalibard, 2015] et internationales [OCDE, 2013] alertent sur l'ampleur des inégalités sociales dans notre système éducatif et sur leur augmentation au cours des dernières années. Les débats récents sur la mixité sociale ont de leur côté révélé le niveau important de ségrégation sociale entre établissements scolaires [Ly et Riegert, 2015]. Dans ce contexte, la réflexion sur les outils de mesure de la position sociale des élèves et des établissements apparaît comme un enjeu important, afin d'objectiver l'état des inégalités sociales à l'école. C'est l'objet de cet article qui présente un indice de position sociale permettant d'appréhender le statut social des élèves à partir de la profession et catégorie sociale (PCS) de leurs parents.

Il s'agit en réalité d'une refonte d'un indice proposé il y a quelques années [Le Donné et Rocher, 2010], baptisé « indice de synergie socio-scolaire ». Depuis sa publication, cet indice a suscité un réel intérêt, selon des axes d'analyse beaucoup plus variés que ceux envisagés lors de sa construction. L'indice s'est ainsi révélé être un outil pertinent de description et de comparaison des disparités sociales au sein du système éducatif : par exemple, il a été utilement mobilisé dans le cadre de comparaisons interacadémiques des écarts sociaux entre collèges

[MENESR-DEPP, 2014] ou bien pour étudier l'évolution des performances des élèves selon le niveau social des établissements [Arzoumanian et Dalibard, 2015].

À l'origine, la construction de cet indice était motivée par deux besoins principaux :
– disposer d'une variable qui synthétise plusieurs dimensions (sociale, économique, scolaire), à partir de la PCS, variable centrale de description de l'origine sociale dans le champ de l'éducation ;
– disposer d'une variable quantitative pour décrire des groupes d'élèves et pouvoir ainsi facilement passer du niveau de l'élève à celui de la classe, de l'établissement, etc.

Nous présentons plus en détail ces aspects dans la partie introductive.

INTRODUCTION

Un indice dérivé de la PCS

La profession et catégorie socioprofessionnelle (PCS) est une variable centrale de description de l'origine sociale des élèves. Elle est mobilisable dans les bases de données des élèves et elle est utilisée dans de nombreuses études et recherches en éducation. L'indice de position sociale se présente donc comme une variable dérivée de la PCS.

La nomenclature des PCS utilisée dans le système d'information du ministère de l'Éducation nationale comprend 32 postes (catégories). Il s'agit d'un mélange de deux niveaux de nomenclature de la PCS de l'Insee, entre la nomenclature en 24 postes et la nomenclature en 42 postes[1]. La DEPP a ainsi adapté la nomenclature « officielle » de l'Insee aux spécificités du domaine de l'éducation.

En outre, la DEPP utilise régulièrement, pour des raisons pratiques, un regroupement des PCS en quatre catégories, décrites brièvement ici :
– favorisée A : cadres et assimilés, chefs d'entreprise, professeurs des écoles et assimilés ;
– favorisée B : professions intermédiaires ;
– moyenne : employés, agriculteurs, artisans, commerçants ;
– défavorisée : ouvriers, inactifs.
Cette classification, en agrégeant les PCS en seulement quatre groupes, présente des avantages pratiques certains en matière d'analyses et de publications[2]. Par ailleurs, ce regroupement constitue en réalité une hiérarchisation des PCS selon ce que l'on pourrait appeler leur « distance » à l'école. De ce point de vue, un indice quantitatif apparaît comme un prolongement naturel de cette approche.

Un indice quantitatif

Quel que soit leur niveau d'agrégation, les classifications des PCS présentent des limites dès qu'il s'agit de décrire la tonalité sociale de groupes d'élèves (classes, établissements, académies). Quel est le niveau social d'un établissement ? Quelles sont les disparités sociales au sein d'un département ?

1. Ce sont les niveaux 2 et 3 de la nomenclature des PCS, qui comporte quatre niveaux emboîtés (www.insee.fr).

2. Cependant, les critères de regroupement n'ont pas été explicités lors de sa création il y a plus de vingt ans. Et la question de revoir cette catégorisation se pose aujourd'hui, l'indice de position sociale pouvant être utile à ce travail.

Concernant le niveau social moyen, un indicateur usuel consiste à considérer un pourcentage d'élèves d'une certaine catégorie sociale (par exemple, le pourcentage d'élèves défavorisés). Cependant, cette perspective n'est pas très satisfaisante, car, d'une part, la catégorie retenue peut se révéler hétérogène (en l'occurrence, les défavorisés regroupent par exemple des ouvriers qualifiés et des chômeurs n'ayant jamais travaillé), d'autre part, elle se focalise sur une partie seulement des élèves. S'agissant maintenant des indicateurs de disparités, les variables qualitatives conduisent à considérer des indices statistiques plus complexes et plus difficiles à interpréter [Givord, Guillerm *et alii*, 2015] que ne l'est un écart-type ou bien une corrélation.

Un indice quantitatif permet ainsi de décrire plus facilement des groupes d'élèves et constitue de ce point de vue un outil pratique d'analyse. L'idée est simple : il s'agit de transformer la variable PCS, c'est-à-dire d'attribuer une valeur numérique à chaque PCS. Dès lors que les PCS sont disponibles (base de données, enquêtes), il suffit d'appliquer les valeurs de référence et de considérer cette nouvelle variable comme quantitative. Il est alors très simple de calculer le niveau social d'un établissement, avec l'indice moyen par exemple, ou les disparités sociales au sein d'une académie, avec l'écart-type.

Cette idée n'est pas nouvelle[3], mais elle s'inscrit plutôt dans une tradition anglo-saxonne et renvoie à la notion de statut social (*status*). Cette approche consiste en effet à construire une mesure unidimensionnelle de la stratification sociale. En théorie, c'est une vision discutable de l'univers social — en tout cas opposée à la tradition sociologique française [voir par exemple Bourdieu, 1979]. En pratique, un tel indice constitue un outil d'analyse très utile.

Concrètement, pour déterminer les valeurs associées aux PCS, nous considérons tout simplement la moyenne pondérée de caractéristiques (ou combinaison linéaire) par PCS. L'indice d'une PCS est ainsi le résumé quantitatif d'un certain nombre d'attributs socio-économiques. Cette approche implique donc un modèle sociologique. Par exemple, nous pourrions considérer l'indice comme le revenu moyen par PCS, ou bien le nombre d'années d'études par PCS, ou encore le pourcentage de bacheliers par PCS.

Deux exemples d'indices existants et décrits par Le Donné et Rocher [2010] adoptent ainsi chacun une démarche spécifique : le *Socio-Economic Index* [Ganzeboom, De Graaf, Treiman, 1992] et le statut social des professions [Chambaz, Maurin, Torelli, 1998]. Pour faire simple, le premier consiste à considérer le revenu par profession, à niveau d'éducation donné, tandis que le second traite du statut social, c'est-à-dire de l'importance sociale que les individus accordent aux différentes professions. De notre côté, nous nous plaçons dans le cadre scolaire. Notre modèle sociologique implique donc de considérer des variables en lien avec l'éducation.

Une refonte

L'utilisation pratique de la première version de l'indice a soulevé un certain nombre de questions, portant notamment sur les modalités de construction de l'indice. Une refonte de cet indice a donc été engagée pour tenir compte de ces questions, ainsi que des suggestions d'amélioration.

3. Par exemple, l'évaluation PISA utilise quantité d'indices, notamment concernant le milieu socio-économique [Keskpaik et Rocher, 2011].

Les aspects considérés pour la refonte de l'indice sont les suivants :
– La méthode retenue pour calculer l'indice : si le principe d'un indice correspondant à la moyenne pondérée de caractéristiques sociales par PCS est conservé, différentes méthodes permettent d'établir cette moyenne. Nous avons donc appliqué et comparé plusieurs approches en ce sens.
– Le choix des variables prises en compte dans sa construction : c'est sans doute le point le plus revisité par rapport à la première version. En particulier, la question est celle de la place respective des variables socio-économiques décrivant le foyer d'une part, et des variables de résultats scolaires des élèves d'autre part. Nous envisageons une perspective nouvelle ici, plus cohérente, de notre point de vue.
– La catégorisation sociale : elle peut être appréhendée de façon différente. En particulier, qui considérer pour apprécier la position sociale d'un élève : la PCS du responsable, de la mère, du père, des deux parents ?
– Les données analysées : quelle est l'influence de l'échantillon utilisé, du traitement des non-réponses, etc. ?
– Les caractéristiques de l'indice, en particulier l'échelle de ses valeurs et sa standardisation.

L'article se présente de façon relativement classique. Nous décrivons tout d'abord les données utilisées, puis la méthodologie employée. Les résultats principaux sont présentés et des compléments d'analyse permettent de répondre à certaines questions s'agissant notamment de la robustesse de l'indice.

DONNÉES

Le panel sixième 2007

Depuis les années 1970, la DEPP met régulièrement en place des « panels », c'est-à-dire des suivis longitudinaux de larges échantillons d'élèves. L'objectif de ce type de dispositif est d'éclairer le système éducatif sur les parcours des élèves, leurs performances scolaires et les processus d'orientation [voir par exemple, Caille, 2014].

À la rentrée 2007, un panel d'environ 35 000 élèves entrant en sixième a été tiré au sort. Il s'agit d'élèves entrés pour la première fois en classe de sixième dans un collège public ou privé sous contrat, en France métropolitaine ou dans un département d'outre-mer. Outre les données sur les parcours scolaires, des informations concernant les familles des élèves ont été recueillies à travers une enquête réalisée en 2008, puis en 2011. Enfin, nous disposons également de données sur les acquis cognitifs des élèves, ainsi que sur des variables affectivo-motivationnelles (conatives), à travers des questionnaires d'évaluation spécifiques passés en 2008 puis en 2011, et en 2012 pour les élèves ayant redoublé.

L'échantillon analysé dans cet article est celui des 29 544 élèves de sixième qui ont passé les évaluations standardisées et dont les familles ont répondu au questionnaire famille de 2008[4]. Bien que les données utilisées pour l'indice soient uniquement celles issues du questionnaire famille, nous nous sommes cependant restreints aux élèves ayant passé les évaluations, afin de comparer des variantes, pouvant inclure les scores cognitifs.

4. L'ensemble des données a été repondéré pour assurer la représentativité de cet échantillon restreint.

Notons enfin que les mêmes analyses ont été conduites sur les données du questionnaire famille passé en 2011 et que les résultats obtenus ne diffèrent que très marginalement. Le choix du questionnaire 2008 a été préféré, car il est plus riche : il comprend notamment des questions sur les pratiques culturelles et sur l'implication des parents. C'est d'ailleurs ce type de variables qui créent des différences de valeurs d'indices (plutôt que la cohorte considérée), ainsi que nous l'évoquons plus loin dans la partie sur les analyses complémentaires.

Les professions et catégories sociales (PCS)

Les PCS des parents sont appréhendées de deux façons différentes dans le panel sixième. La première repose sur le questionnaire famille avec un codage selon la nomenclature des PCS de l'Insee. La seconde consiste à récupérer les PCS des parents telles qu'elles existent dans les bases élèves académiques (BEA), selon la nomenclature utilisée à la DEPP.

Il est clair que la première source d'information est de meilleure qualité, car le questionnement est précis et le codage standardisé. Les PCS présentes dans les BEA sont quant à elles plus fragiles, car elles reposent sur de simples déclarations et leur codage n'est pas contrôlé de manière serrée [Soulié, 2000]. En outre, elles incluent la notion de « responsable » de famille, au lieu de distinguer d'emblée et explicitement la mère et le père.

Cependant, l'utilisation pratique de l'indice, par les statisticiens de la DEPP, par ceux des services statistiques académiques ou par les chercheurs, concerne principalement les bases de données du ministère (Scolarité, Ocean, BCP, etc.) qui s'appuient sur les BEA. C'est pourquoi nous considérons ici les variables issues des BEA pour construire l'indice, en l'occurrence la PCS du responsable (PCS1), mais également les PCS des deux parents[5]. Les analyses croisées présentées dans la suite de l'article montrent en effet qu'il est intéressant de pouvoir distinguer les deux parents.

MÉTHODOLOGIE

Analyse des correspondances multiples (ACM)

Comme nous l'avons décrit en introduction, l'indice se présente comme un résumé de différentes caractéristiques familiales de l'élève, pour chaque PCS.

Pour calculer l'indice, nous procédons à l'analyse des correspondances multiples (ACM) d'un ensemble de caractéristiques familiales et nous projetons les PCS sur la première composante. Pour une PCS donnée, la valeur de l'indice correspond donc à la moyenne du premier score factoriel des élèves concernés. L'indice est ainsi la résultante d'une combinaison linéaire de variables familiales, dont les valeurs sont ensuite moyennées par PCS. Le poids d'une variable dans la combinaison linéaire est déterminé par l'inertie qu'elle représente parmi l'ensemble des variables retenues.

Cette méthode est identique à celle utilisée pour la version initiale de l'indice [Le Donné et Rocher, 2010]. Notons au passage que l'ACM est ici utilisée pour réduire les données (trouver

5. En effet, il est possible de transformer les deux variables renseignées (PCS1 et PCS2 de la base Scolarité) pour obtenir la PCS de la mère et celle du père.

le « meilleur » résumé de l'ensemble des données), alors que traditionnellement, l'ACM est utilisée pour décrire un espace multidimensionnel.

Nous avons retenu cette méthode, mais certaines variantes peuvent être envisagées, toujours selon ce principe d'un indice construit comme une combinaison linéaire de caractéristiques familiales. Ces variantes conduisent en pratique à modifier les poids des variables. Elles sont présentées plus loin dans l'article.

Choix des variables

Pour la refonte de l'indice, la principale modification a consisté à repenser les procédures de sélection des variables qui composent l'indice, après avoir identifié quelques faiblesses dans la première version[6].

Nous mobilisons le concept de « médiation » pour sélectionner les variables entrant en compte dans l'ACM. L'idée est que les PCS n'agissent pas directement sur les résultats scolaires, mais *via* un certain nombre de caractéristiques. Le fait d'avoir un père cadre ne procure pas *en soi* un avantage, en termes scolaires, par rapport au fait d'avoir un père ouvrier. La catégorie sociale est en réalité le reflet de capitaux (économiques, éducatifs, culturels), de pratiques et d'attitudes qui tendent à plus ou moins favoriser les conditions d'apprentissage, la réussite scolaire, etc. Par exemple, la PCS est associée à un niveau de revenu qui permet d'assurer les conditions matérielles d'une bonne scolarité ; de la même manière, la PCS distingue des niveaux de diplômes qui signent un certain bagage de connaissances susceptibles d'être transmises dans le cadre scolaire ; la PCS est associée à des niveaux différents d'implication, d'attente par rapport à l'école, etc.

Ainsi, on pourrait dire que « l'effet » de la PCS des parents sur la réussite scolaire des enfants transite par cet ensemble de variables dites « médiatrices ». Dans l'**encadré 1**, nous donnons quelques éléments sur ce concept de médiation.

Cette perspective très générale se place sur le plan pratique, à travers l'étude des corrélations entre variables. En réalité, sur le plan statistique, une variable est médiatrice dès lors qu'elle capte une partie de l'effet de la PCS sur les résultats scolaires. C'est précisément ce qui nous intéresse : les variables constitutives de l'indice sont liées à la fois aux résultats scolaires et à la PCS, car elles expliquent une partie du lien observé entre PCS et résultats scolaires.

L'étude des effets de médiation nous permet de sélectionner des variables, d'un point de vue statistique, sur des bases objectivables. Cette approche pragmatique permet ainsi d'intégrer différents modèles sociologiques, notre article n'ayant clairement pas pour ambition de se

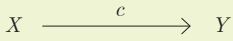
6. Pour la première version de l'indice, la sélection des variables entrant en compte dans le calcul de l'indice s'était appuyée sur les résultats de la régression sur le score cognitif total des variables familiales. Seules les variables dont le coefficient de régression était significatif avaient été retenues pour l'ACM. Cette approche a l'avantage d'être explicite sur le choix des variables. Mais elle présente l'inconvénient d'une certaine forme de double analyse : la première composante de l'ACM est une combinaison linéaire des variables familiales, combinaison linéaire qui, d'une certaine manière, est déjà approchée dans la régression multiple. En outre, elle conduit à évincer certaines variables clés comme le revenu, car son effet sur le score n'est plus significatif, lorsque l'on contrôle d'autres variables. Enfin, le score cognitif était inclus dans l'ACM, suivant l'idée de synergie socio-scolaire. Ce choix pose question. D'une part, le score est employé doublement : pour sélectionner les variables puis comme composante de l'indice. D'autre part, si l'indice est voué à être analysé en lien avec des résultats scolaires, un effet de circularité est théoriquement possible. Pour toutes ces raisons, nous avons retenu une autre approche pour sélectionner les variables constitutives de l'indice.

Encadré 1

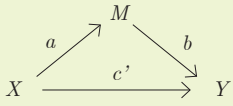
ÉLÉMENTS SUR L'EFFET DE MÉDIATION

Il existe une littérature très abondante sur la notion de médiation, en particulier en psychologie sociale [Baron et Kenny, 1986]. Nous donnons ici quelques éléments de base.

Supposons qu'une variable dite causale X explique une variable de résultat Y. L'effet de X sur Y est noté c. Typiquement, il s'agit du coefficient de la régression associé à X dans la régression simple de X sur Y.



L'effet de X sur Y peut transiter (être « médiatisé ») par une variable dite médiatrice M, comme dans le diagramme ci-dessous. Dans ce cas, on parle d'effet direct de X sur Y, noté c' (typiquement le coefficient de régression associé à X dans la régression multiple de X et M sur Y).



Généralement, quatre conditions sont discutées pour établir un effet de médiation :

1. La variable causale X est corrélée à la variable de résultat Y.
2. La variable causale X est corrélée à la variable médiatrice M.
3. La variable médiatrice M affecte la variable de résultat Y, lorsque la variable causale X est contrôlée.
4. L'effet de X sur Y est atténué, voire annulé, lorsque l'on contrôle par M. Autrement dit, la différence $(c - c')$ est positive et significative.

Dans notre cas, nous avons vérifié ces conditions pour différentes variables médiatrices issues du questionnaire famille, en considérant la PCS comme variable causale et le score cognitif total comme variable de résultat.

Techniquement, la significativité de la différence $(c - c')$ a été testée *via* des procédures de *bootstrap* (1 000 tirages avec remise), à défaut de disposer d'une méthode d'estimation faisant consensus.

positionner dans le débat théorique sur les mécanismes à l'œuvre dans la transmission des inégalités sociales à l'école. Concrètement, en amont, les variables retenues pour l'ACM ont tout d'abord été sélectionnées parmi celles disponibles dans le questionnaire famille du panel, sur la base d'un choix *a priori*[7], puis d'un examen empirique de leur effet de médiation[8].

L'indice est donc envisagé comme la résultante de variables familiales médiatrices des PCS sur les acquis cognitifs, c'est-à-dire de capitaux, de pratiques et d'attitudes caractérisant les PCS des parents et susceptibles d'avoir une influence sur le niveau de compétence de l'enfant.

Les variables retenues au final sont présentées dans le **tableau 1** (et plus en détail dans le **tableau 5** en annexe, p. 26). Nous pouvons catégoriser ces variables de la manière suivante : les diplômes des parents, les conditions matérielles du foyer, la composition familiale, le capital culturel, le niveau d'ambition et d'implication des parents et les pratiques culturelles de la famille et de l'élève.

7. En partant de la sélection déjà opérée dans le cadre d'analyses menées à la DEPP, par exemple Ben Ali et Vourc'h [2015].

8. Notons que l'effet de médiation des variables est testé séparément pour chacune des variables et non pas de manière simultanée (évitant une sélection reposant sur un effet « toutes choses égales par ailleurs », comme cela était le cas pour la régression multiple).

↘ Tableau 1 Variables du questionnaire famille incluses dans l'analyse des correspondances multiples (ACM)

Variables	Nombre de modalités
Diplômes des parents	
Diplôme de la mère	9
Diplôme du père	9
Conditions matérielles	
Revenus mensuels du foyer	10
Nombre de pièces du logement	6
Chambre seul	3
Ordinateur	2
Internet	2
Capital culturel	
Nombre de livres à la maison	4
Télévision dans la chambre	3
Regarde la télévision régulièrement	3

Variables	Nombre de modalités
Ambition et implication	
Aspiration	6
Diplôme le plus utile	5
Implication des parents	5
Conversations (vie scolaire)	3
Conversations (avenir scolaire)	3
Pratiques culturelles (6 derniers mois)	
Événement sportif	4
Concert	4
Théâtre	4
Cinéma	4
Musée	4
Activités extra-scolaires	3

Par rapport au corpus disponible, certains choix ont été opérés[9]. Tout d'abord, certaines variables n'ont pas été prises en compte *a priori* : c'est le cas de la structure familiale ou du statut d'immigration. Le statut d'immigration est pourtant lié à la PCS et aux résultats scolaires, mais de notre point de vue, cette variable ne peut être véritablement considérée comme une médiatrice, c'est-à-dire comme expliquant une partie de l'effet des PCS sur les résultats scolaires. C'est d'ailleurs plutôt l'inverse qui est généralement considéré : l'étude de l'influence du statut d'immigration sur la réussite, à PCS donnée [voir par exemple, Vallet et Caille, 1996].

D'autres variables n'ont pas été retenues, car l'effet médiateur n'a pas été empiriquement révélé. Par exemple : la possession d'un ordinateur personnel n'est pas incluse en raison d'un faible effet de médiation, alors que la présence d'un ordinateur dans le foyer a un effet médiateur significatif.

Enfin, se pose la question délicate de l'endogénéité de certaines variables. Par exemple, le fait de faire appel à des cours de soutien payants n'est pas que le reflet de pratiques familiales, mais renvoie (potentiellement) à des difficultés scolaires rencontrées par les élèves. Cette variable n'a pas été retenue, d'autant que son effet de médiation (si l'on passe outre le problème d'endogénéité) est faible.

Néanmoins, cette question de l'endogénéité des variables concerne également des variables que nous avons décidé d'inclure, telles que le niveau d'ambition et d'implication des familles, ainsi que les pratiques culturelles. Mais le caractère médiateur de ces variables ressort significativement à l'analyse. Nous avons considéré que, d'une part, l'effet d'endogénéité pouvait être relativisé (en comparaison des cours de soutien payants notamment) et que, d'autre part, ces variables renvoient à des facteurs sociologiques bien identifiés, en termes de capital culturel, de

9. Notons que les scores cognitifs n'ont pas été intégrés à l'ACM, tel que cela était fait dans la première version. En effet, selon le principe retenu de variables médiatrices de la PCS sur les résultats cognitifs, il n'y a plus de raison de considérer les scores comme des variables constitutives de l'indice. Pour évaluer l'impact concret de ce choix, nous avons réalisé le même exercice en ajoutant aux variables retenues pour l'ACM les scores cognitifs issus des évaluations menées en sixième. La comparaison entre les deux versions montre une correspondance *quasi* parfaite entre les deux séries par PCS. En pratique, l'intégration des scores à l'ACM ne change donc pas les résultats.

distance par rapport à l'univers scolaire, de stratégies familiales, etc. De manière pratique, nous testons plus loin quel est l'impact sur les valeurs de l'indice de considérer ou non ces variables.

Traitement des valeurs manquantes

Nous analysons l'ensemble des données, y compris les non-réponses aux différentes questions du questionnaire famille. Si l'on supprimait les observations pour lesquelles il y a au moins une valeur manquante dans la liste des variables retenues, cela reviendrait à éliminer 10 435 élèves sur 29 544, ce qui n'est pas envisageable.

La question est alors de savoir s'il faut intégrer ou non les indicatrices de valeurs manquantes dans l'ACM. Pour la version initiale de l'indice, les indicatrices de non-réponses n'ont pas été introduites dans l'ACM. Nous retenons ici l'autre perspective qui consiste à les y inclure. En pratique, une corrélation de 0,97 est observée entre les deux scores factoriels et l'indice des PCS est quasi-identique dans les deux cas.

Notre choix se justifie dans la mesure où les non-réponses sont signifiantes : elles contribuent fortement à créer un second facteur mais elles renvoient généralement à des valeurs négatives sur le premier axe. Autrement dit, la non-réponse est associée à un faible statut social.

RÉSULTATS

Principaux résultats de l'ACM

La **figure 1** p. 17 représente les projections des modalités des variables retenues indiquées dans le **tableau 1** sur le premier plan factoriel de l'ACM. Il apparaît clairement que le deuxième axe doit sa création aux indicatrices de non-réponse (points noirs sur la figure). Comme indiqué précédemment, ces indicatrices renvoient à des valeurs négatives sur le premier axe, témoignant du lien entre non-réponse et faible position sociale[10].

Au-delà des non-réponses, les projections des autres indicatrices laissent apparaître un espace assez fortement unidimensionnel. Les indicatrices renvoyant à de faibles positions sociales apparaissent dans les valeurs négatives de la première dimension, tandis que les indicatrices de positions sociales élevées sont associées à des valeurs élevées sur le premier axe. C'est le cas du revenu et du diplôme, deux variables représentées sur le nuage en guise d'illustration.

Les contributions des différents groupes de variables à la création du premier axe de l'ACM sont relativement équilibrées entre les diplômes, les conditions matérielles, l'ambition et l'implication et les pratiques culturelles ↘ **Figure 1**. En revanche, les variables relatives à la composition familiale et au capital culturel pèsent moins, ce qui est possiblement dû à un effet du nombre de modalités par groupes.

Valeurs de l'indice

Le **tableau 2** donne les valeurs de l'indice pour les PCS du responsable, de la mère et du père. Il s'agit donc des moyennes du score factoriel par PCS.

10. Par conséquent, le nuage de la première version qui faisait apparaître un effet guttmanien (opposition sur le deuxième axe des modalités extrêmes avec les modalités médianes) n'est plus visible ici car les indicatrices de non-réponses sont conservées ↘ **Figure 1**.

↘ Tableau 2 Indice de position sociale selon la PCS des parents (standardisé)

PCS	Libellé	Responsable	Mère	Père
10	Agriculteurs exploitants	104	99	104
21	Artisans	95	95	95
22	Commerçants et assimilés	101	101	102
23	Chefs d'entreprise de dix salariés ou plus	142	128	140
31	Professions libérales et assimilés	158	155	158
33	Cadres de la fonction publique	148	154	148
34	Professeurs, professions scientifiques	159	163	160
35	Professions de l'information, des arts et des spectacles	143	143	147
37	Cadres administratifs et commerciaux d'entreprises	147	145	148
38	Ingénieurs et cadres techniques d'entreprises	158	164	157
42	Professeurs des écoles, instituteurs et assimilés	145	154	149
43	Professions intermédiaires de la santé et du travail social	112	120	125
44	Clergé, religieux	134	147	132
45	Professions intermédiaires de la fonction publique	112	115	116
46	Professions intermédiaires en entreprise	121	125	123
47	Techniciens	116	125	115
48	Contremaîtres, agents de maîtrise	110	122	109
52	Employés civils et agents de service de la fonction publique	90	95	95
53	Policiers et militaires	118	110	115
54	Employés administratifs d'entreprises	105	113	115
55	Employés de commerce	87	88	93
56	Personnels des services directs aux particuliers	73	82	86
61	Ouvriers qualifiés	79	78	79
66	Ouvriers non qualifiés	62	65	64
69	Ouvriers agricoles	58	67	60
71	Anciens agriculteurs exploitants	68	76	84
72	Anciens artisans, commerçants, chefs d'entreprise	83	101	93
73	Anciens cadres et professions intermédiaires	125	129	123
76	Anciens employés et ouvriers	60	97	59
81	Chômeurs n'ayant jamais travaillé	48	51	54
82	Personnes diverses sans activité professionnelle	55	75	61
99	Non renseignée (inconnue ou sans objet)	70	91	70

Note : l'indice est standardisé sur l'échantillon des élèves du panel sixième, à une moyenne de 100 et un écart-type de 30. La colonne responsable renvoie à la PCS du responsable telle qu'elle est présente dans la base scolarité (PCS1).

Notons tout d'abord que l'indice a été standardisé de manière à obtenir une moyenne de 100 et un écart-type de 30. Deux raisons au moins motivent le fait de sortir de la standardisation statistique classique de moyenne 0 et d'écart-type 1. Tout d'abord, il semble préférable d'éviter des valeurs négatives pour une PCS donnée, car la négativité pourrait être mal interprétée, comme une nocivité en termes d'effets sur la réussite. Deuxièmement, il peut être plus pratique de considérer (au moins en matière de communication, pas forcément de calcul) des valeurs entières plutôt que décimales. La présentation de l'indice étant jusqu'à présent faite à la deuxième décimale, un écart-type de 30 semble être un minimum (distinguant ainsi des écarts de 0,03 sur l'échelle de départ). Dans ce cas, une moyenne de 100 permet d'obtenir des valeurs positives pour chaque PCS[11].

S'agissant du responsable de famille, les valeurs obtenues s'étalent de 48 pour les « chômeurs n'ayant jamais travaillé » (code PCS : 81) à 159 pour les « professeurs, professions scientifiques » (34). Nous retrouvons des résultats très cohérents avec les différences sociales bien connues dans le système éducatif. Cependant, des résultats intéressants apparaissent dans les catégories considérées comme moyennes et défavorisées dans la nomenclature habituellement utilisée à la DEPP.

Les valeurs élevées de l'échelle (au-delà de l'indice 140) concernent de façon attendue les « cadres et professions libérales » (31 à 38) ainsi que les « instituteurs et assimilés » (42) et les « chefs d'entreprise de dix salariés ou plus » (23). Les « professions intermédiaires » (42 à 48) se situent ensuite dans une fourchette de 110 à 145, avec également les « policiers et militaires » (53).

Ce dernier résultat est intéressant, car il renvoie à l'hétérogénéité importante des professions de la catégorie des employés (PCS 52 à 56). Avec un indice de 118, les « policiers et militaires » (53) occupent la plus haute position, tandis qu'à l'inverse les « personnels de services directs aux particuliers » (56) ont un indice faible, de 73, entre l'indice des ouvriers qualifiés et celui des ouvriers non qualifiés. En position intermédiaire, mais également dispersées, nous retrouvons les trois autres catégories d'employés : d'entreprise avec un indice de 105, de la fonction publique avec un indice de 90 et de commerce avec un indice de 87. Rappelons ici que la codification des PCS, réalisée localement dans les établissements scolaires, peut être soumise à des biais. Ainsi, concernant la catégorie « policiers et militaires » (53), il est possible que l'ensemble des policiers et militaires soient regroupés dans cette catégorie, indépendamment de leurs grades. Autour de la moyenne se situent également les « agriculteurs exploitants » (10) avec un indice de 104, les « commerçants et assimilés » (22) avec un indice de 101, ainsi que les « artisans » (21) avec un indice de 95.

Les catégories dites défavorisées ont en réalité des indices très dispersés. Avec un indice de 79, les « ouvriers qualifiés » (61) se distinguent nettement des « ouvriers non qualifiés » (66) ou des « ouvriers agricoles » (69) qui ont respectivement des indices de 62 et de 58. Les inactifs se situent quant à eux à des valeurs encore bien inférieures : les « personnes diverses sans activité professionnelle » (82) obtiennent un indice de 55 et les « chômeurs n'ayant jamais travaillé » (81) un indice de 48.

Concernant les autres catégories, les PCS « non renseignées (inconnues ou sans objet) » (99) obtiennent un indice très faible, de 70. Par ailleurs, les catégories concernant les retraités (71

11. Des valeurs plus élevées conduiraient potentiellement à des confusions avec les scores de Cedre (standardisation 250/50) ou de PISA (standardisation 500/100).

à 76) sont également assez dispersées, bien que ces valeurs soient relativement fragiles, car établies sur de petits effectifs.

Notons que l'indice est une moyenne de score factoriel par PCS. Il rend donc compte de situations différentes, à PCS donnée. La **figure 5** en annexe montre ainsi les dispersions du score factoriel observées au sein de chaque catégorie, ainsi que les concentrations d'effectifs sur certaines PCS.

Pour finir, la démarche employée ici concerne le milieu familial de l'enfant, incluant des caractéristiques sur son foyer, mais aussi sur chacun de ses deux parents à travers leurs diplômes. Dès lors, l'indice peut être différent pour les deux parents, car une PCS donnée peut renvoyer à des réalités différentes, selon qu'elle concerne la mère ou le père. Il apparaît que la hiérarchie des indices est globalement respectée entre les trois PCS (responsable, père, mère). Les indices concernant le père et le responsable sont très corrélés, les deux étant d'ailleurs majoritairement confondus. On observe quelques modifications concernant la mère, en particulier les « chefs d'entreprise de dix salariés ou plus » qui obtient un indice de 128 pour la mère, contre 142 lorsqu'il s'agit du responsable. Autre exemple, la catégorie « Personnes diverses sans activité professionnelle » apparaît moins défavorisée, s'agissant de la mère (indice de 75, contre 51 pour le responsable), car elle peut regrouper des mères au foyer ayant des caractéristiques socio-économiques favorables.

Dans le paragraphe suivant, nous allons plus loin dans l'investigation du croisement des indices entre le père et la mère.

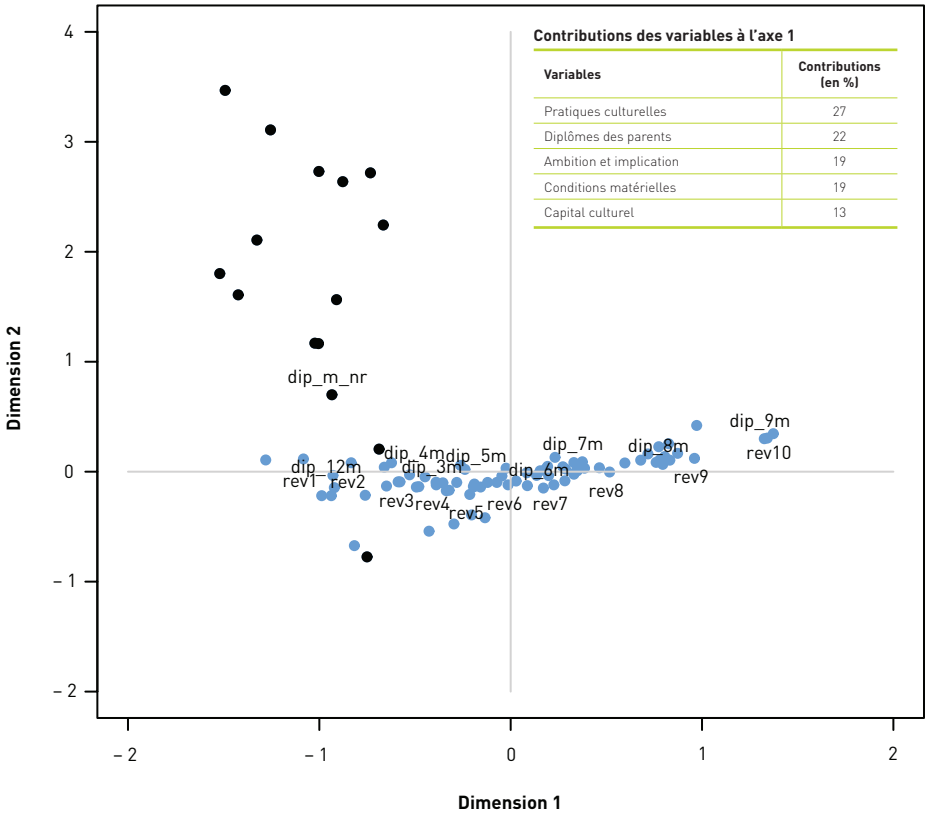
Un indice pour le croisement des PCS des deux parents

Une des interrogations soulevées lors de l'exploitation de l'indice tient à la nature de la variable construite. L'idée est bien d'utiliser l'indice en tant que variable continue représentant la position sociale des familles sur un continuum. Cependant, en pratique, l'indice n'admet qu'une trentaine de valeurs possibles, avec des concentrations importantes, ce qui peut fragiliser le calcul de moyennes ou d'écarts-types pour certaines sous-populations.

Afin de réduire la nature discrète de l'indice et de tenir compte de la dispersion des scores factoriels, nous avons donc considéré la projection du croisement entre la PCS du père et celle de la mère. La **figure 2** montre la relation entre ce nouvel indice et les indices de départ, de la mère et du père. Bien que ces indices soient fortement liés, l'indice croisé permet d'introduire une variabilité intéressante : à valeur fixée pour un des parents, l'indice croisé peut être relativement dispersé. Par exemple, c'est le cas pour de faibles valeurs de l'indice concernant la mère, comme les mères inactives, qui peuvent en fait être associées à de fortes valeurs si l'on tient compte du croisement avec la PCS du père.

En pratique, il s'agit d'utiliser toute l'information présente dans les bases de données, c'est-à-dire les deux PCS à disposition. Se pose cependant la question de la robustesse des valeurs obtenues pour l'indice croisé. D'une part, l'échantillon ne contient pas tous les croisements possibles (744 sur 1 024). D'autre part, avec l'effet d'endogamie, tous les croisements ne sont pas équiprobables. Par conséquent, certaines cases sont vides ou bien très faiblement représentées. C'est pourquoi au final, nous avons calculé l'indice croisé, en tenant compte d'un seuil de trente élèves exigé quant à l'effectif du croisement considéré. Si l'effectif est inférieur à ce seuil, nous avons considéré l'indice le plus élevé entre celui de la mère et celui du père.

↘ Figure 1 Premier plan factoriel de l'analyse des correspondances multiples



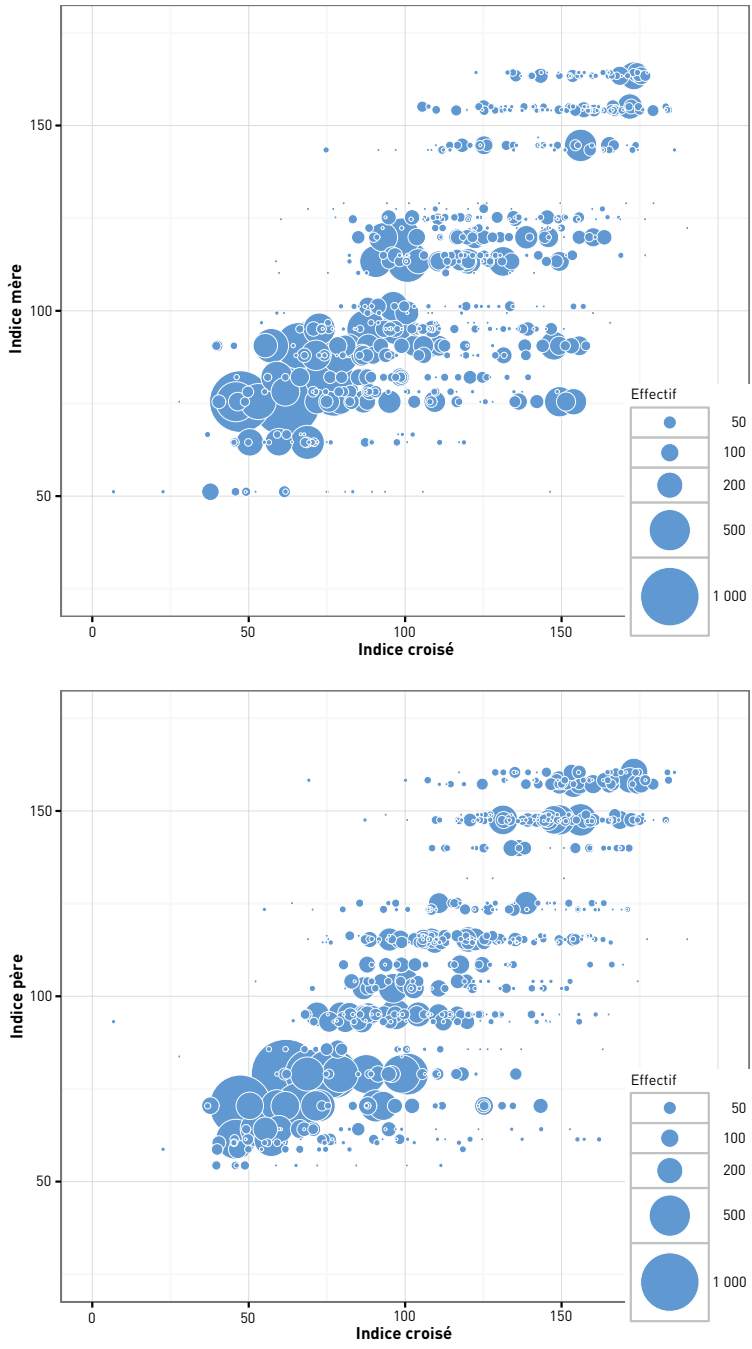
Note : les points représentent les modalités positionnées sur le premier plan factoriel de l'ACM. Les points en noir sont les indicatrices de non-réponse. Les modalités des revenus du foyer (rev1 à rev10) ainsi que les modalités du diplôme de la mère (dip) ont été représentées pour illustrer deux variables parmi l'ensemble des variables retenues.

↘ Tableau 3 Corrélations entre les indices et les scores

Indices	F6	M6	S1	S2	S4	S6	S7	S8
ACM 1	0,48	0,49	0,32	0,49	0,50	0,39	0,46	0,33
Indice mère	0,29	0,30	0,19	0,31	0,31	0,24	0,29	0,20
Indice père	0,34	0,35	0,22	0,35	0,36	0,27	0,33	0,22
Indice croisé (IC)	0,40	0,40	0,26	0,41	0,42	0,32	0,39	0,26
IC30	0,39	0,39	0,25	0,40	0,40	0,31	0,38	0,26

Note : le niveau le plus fin de l'indice est la valeur de la première composante – score factoriel – de l'ACM (ACM 1). Les valeurs de l'indice de la mère et de l'indice du père correspondent au tableau 3. L'indice croisé est obtenu en projetant le croisement des deux PCS sur le premier axe de l'ACM. L'indice IC30 renvoie aux valeurs de l'indice croisé selon les règles de recodage définies dans le texte. Les scores F6 et M6 sont les scores obtenus aux évaluations diagnostiques de début de sixième en français et en mathématiques. Les scores S1, S2, S4, S6, S7 et S8 sont les scores calculés à partir des séquences cognitives de l'évaluation standardisée du panel sixième.

↘ Figure 2 Comparaison des indices des parents avec l'indice croisé



Note : le graphique du haut croise les valeurs de l'indice obtenues pour la projection du croisement de la PCS de la mère et du père (en abscisse), avec celles obtenues pour la projection de la PCS de la mère. Le graphique du bas est similaire, mais pour la PCS du père.

Nous avons comparé ces différents indices (simples ou croisés) en termes de corrélations avec des variables externes. Le **tableau 3** donne les corrélations entre les indices et les scores cognitifs disponibles. De manière attendue, tous les niveaux d'indices retenus (mère, père, croisés) conduisent à une sous-estimation des corrélations observées entre le score factoriel issu de l'ACM et les scores cognitifs. Cependant, les indices croisés rendent mieux compte de ces corrélations que les indices simples selon la mère et le père. Enfin, l'indice croisé calculé avec la règle du seuil de trente élèves ne dégrade que très peu les corrélations, en comparaison de l'indice croisé « brut ».

Une table[12] contient les indices de position sociale pour les 1 024 croisements possibles (32 PCS de la mère x 32 PCS du père).

ANALYSES COMPLÉMENTAIRES

Influence du choix des variables

Nous avons comparé les valeurs de l'indice, selon les variables retenues. En particulier, nous nous sommes intéressés au cas des variables les plus « endogènes », à savoir l'ambition et l'implication déclarées des parents, et les pratiques culturelles. L'ACM a ainsi été réalisée de deux autres manières, partant de l'ensemble des variables du **tableau 1**, mais d'abord sans les variables concernant l'ambition et l'implication des parents, puis sans les variables concernant à la fois l'ambition et l'implication, mais aussi sans les pratiques culturelles.

La **figure 6** présentée en annexe donne les résultats de ces deux nouvelles versions, en comparaison avec l'indice de base. Il apparaît que l'omission des variables qui concernent l'ambition et l'implication des parents n'affecte que très marginalement les valeurs de l'indice par PCS. En revanche, l'omission des variables portant sur les pratiques culturelles conduit à une modification des indices pour certaines catégories. En particulier, les agriculteurs, les artisans, les ouvriers agricoles voient leurs indices augmenter de 7 à 8 points. À l'inverse, les professions de l'information, des arts et des spectacles subissent une baisse de 11 points d'indice. Ces résultats traduisent les différences d'intensité des pratiques culturelles dans ces différentes catégories. Malgré ces différences, les deux versions restent très corrélées.

Au final, l'indice montre de manière générale une forme de robustesse. En effet, même la comparaison avec l'ancienne version [Le Donné et Rocher, 2010] montre une forte corrélation, alors que ni la méthode, ni les variables n'étaient les mêmes[13].

Comparaison avec d'autres méthodes

Nous comparons maintenant différentes méthodes de calcul de l'indice. Le principe retenu consiste à synthétiser dans les PCS un ensemble de caractéristiques familiales *X*, en lien avec les résultats scolaires *Y*. Selon ce principe, au-delà de l'ACM des variables *X*, deux autres possibilités ont été envisagées :

12. Téléchargeable sur : www.education.gouv.fr/revue-education-et-formations-numero-90

13. Une modification notable est cependant l'apparente réduction des écarts entre PCS avec la nouvelle version qui peut s'expliquer par le fait que les PCS projetées ici sont issues des BEA et donc plus « bruitées » que celles recueillies dans le questionnaire famille (et projetées dans première version).

– l'analyse discriminante : déterminer la composante des *X* qui discrimine le plus les PCS[14] ;
– la régression PLS (*Partial Least Square*) : déterminer la composante des *X* qui rend le mieux compte des covariances entre *X* et *Y*.

D'une certaine manière, en comparaison de l'ACM, on pourrait dire que l'analyse discriminante tient compte des PCS tandis que la régression PLS tient compte des scores. L'analyse discriminante prolonge l'ACM : les poids des variables sont calculés de manière à discriminer au maximum les PCS et non plus seulement en fonction leurs inerties respectives. La régression PLS, quant à elle, prend en compte le pouvoir explicatif des variables sur les scores aux épreuves cognitives.

Comme le montre le **tableau 4**, les indices tirés de l'ACM ou de la régression PLS sont très proches, montrant à nouveau que la prise en compte des scores cognitifs n'a que peu d'impact sur les valeurs de l'indice. En revanche, les indices issus de l'analyse discriminante sont moins bien corrélés, entre eux et avec les deux premiers. Visiblement, l'analyse discriminante produit un « éclatement » des valeurs pour certaines catégories. La **figure 3** montre que les distributions des indices issus de l'analyse discriminante sont bimodales, contrairement à celles des indices tirés de l'ACM ou de la régression PLS qui sont quasi-normales. Les moyennes par PCS révèlent que les valeurs de l'indice pour les PCS des cadres sont globalement très supérieures pour l'analyse discriminante. Une analyse des contributions des variables dans les différents indices montre que les poids des variables sont relativement proches, à l'exception du diplôme universitaire du deuxième ou troisième cycle qui est plus important dans l'indice issu de l'analyse discriminante, ce qui pourrait expliquer ces décalages.

Au final, l'influence importante d'une seule variable sur la distribution de l'indice dans le cadre de l'analyse discriminante nous conforte dans le choix de la méthode de l'ACM.

↘ Tableau 4 Selon les méthodes – corrélations des indices

	ACM	PLS	DISCM	DISCP
ACM		0,98	0,84	0,86
PLS	0,98		0,84	0,86
DISCM	0,84	0,84		0,83
DISCP	0,86	0,86	0,83	

Note : il y a deux indices pour l'analyse discriminante (DISCM et DISCP) car la discrimination concerne soit la variable PCS de la mère, soit celle du père.

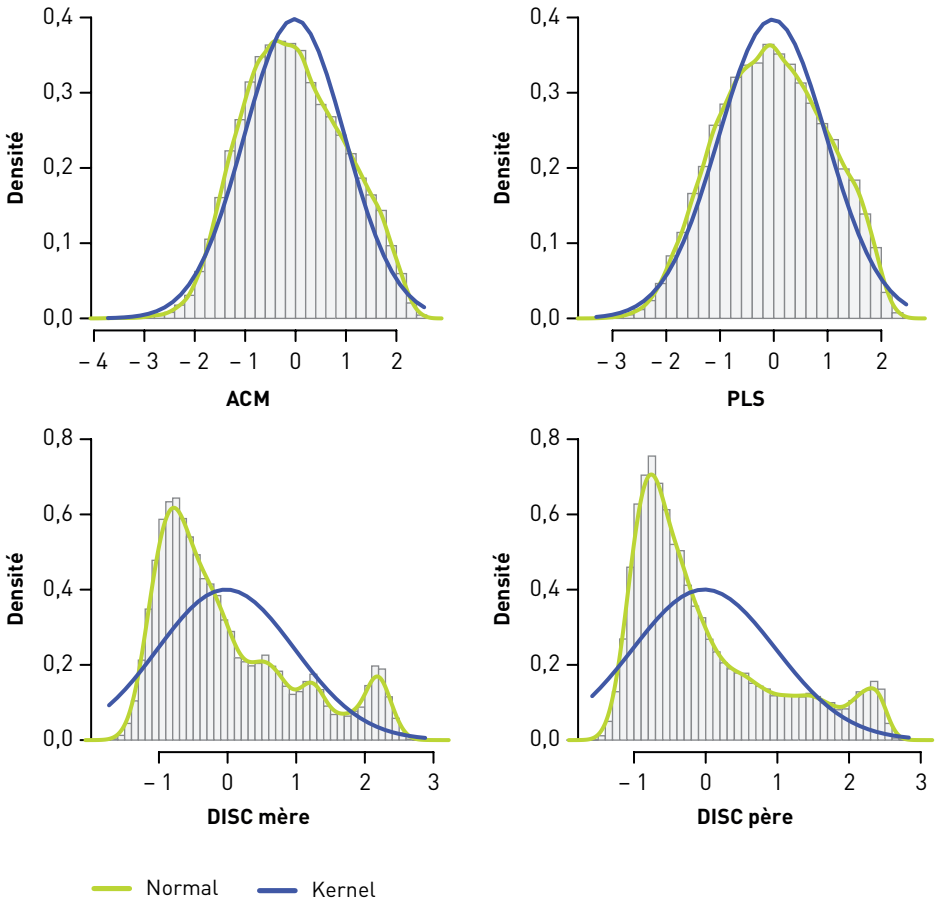
Relation avec des indicateurs de résultats

Pour terminer ces analyses complémentaires, nous donnons quelques éléments concernant le lien observé entre cet indice et quelques résultats scolaires. La **figure 4** compare les valeurs de l'indice avec quatre indicateurs de résultats, en fonction de la PCS du responsable : taux de retard à l'entrée en sixième, score aux évaluations standardisées du panel en sixième, taux de réussite au DNB et taux de réussite au baccalauréat.

Il ressort une grande cohérence entre l'indice et ces indicateurs, en particulier s'agissant des scores aux évaluations standardisées du panel, ce qui peut s'expliquer par le fait que ces scores entrent en compte pour la sélection des variables.

14. Les variables à l'étude étant catégorielles, la méthode consiste en pratique à réaliser une ACM de ces variables puis à utiliser les composantes obtenues comme variables continues pour réaliser une analyse discriminante sur les PCS.

↘ Figure 3 Distributions des indices selon les méthodes

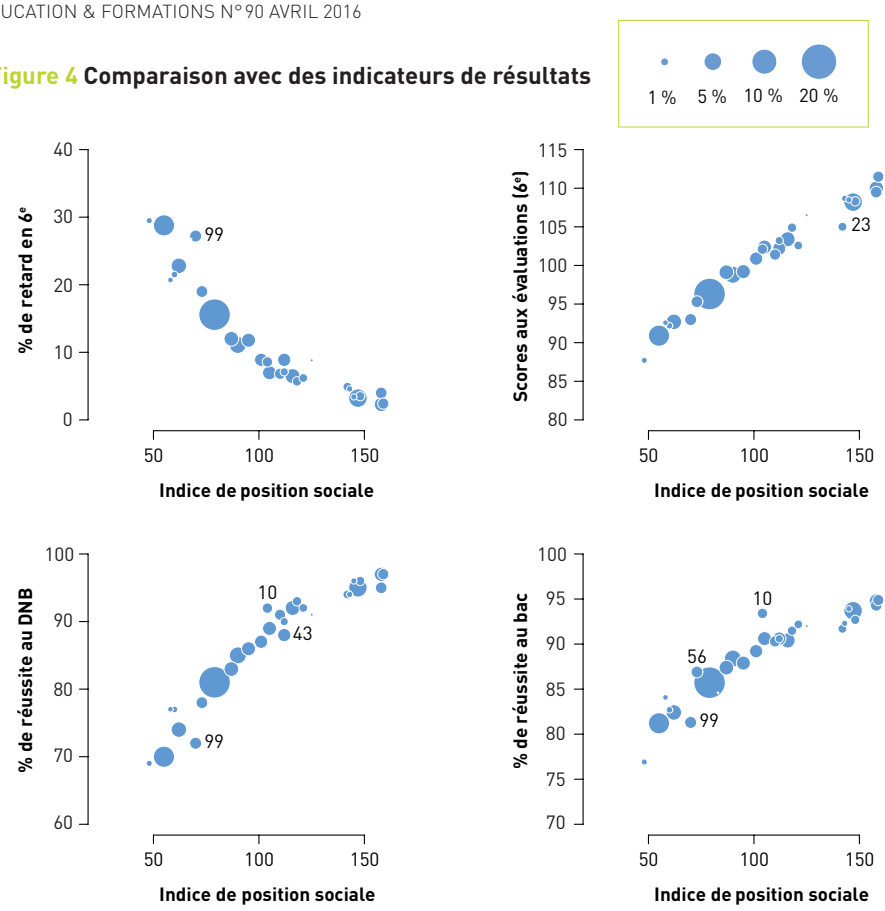


Note : de gauche à droite et de haut en bas : distribution de l'indice issu de l'ACM, de l'indice issu de la régression PLS, de l'indice issu de l'analyse discriminante sur la PCS de la mère et de l'indice issu de l'analyse discriminante sur la PCS du père.

Quelques catégories affichent cependant des écarts entre les valeurs de l'indice et les résultats. C'est le cas de la catégorie qui regroupe les non-réponses (99) : les résultats des élèves de cette catégorie sont moins bons que ne le laisserait présager la valeur de l'indice (tout au moins pour ce qui concerne le retard, le DNB et le baccalauréat, pas les scores aux évaluations).

Il est possible que les variables de l'indice ne captent pas ici certaines caractéristiques potentiellement négatives pour la réussite scolaire que révèle le fait de ne pas renseigner sa profession. À l'inverse, les enfants d'agriculteurs réussissent mieux aux examens (DNB, baccalauréat), en comparaison de leur position sociale. C'est particulièrement vrai pour le baccalauréat. Des études complémentaires mériteraient d'être menées sur cette catégorie d'élèves.

↘ Figure 4 Comparaison avec des indicateurs de résultats



Note : les figures croisent les valeurs de l'indice avec des indicateurs de résultats, selon la PCS du responsable. Les points représentent les PCS, de taille proportionnelle à leur effectif dans l'échantillon du panel.

POUR CONCLURE

Cette nouvelle version de l'indice de position sociale s'est donné pour objectif d'adopter des principes de construction plus cohérents. Cependant, le classement des PCS selon l'indice ne subit pas de bouleversements. Plus généralement, les valeurs de l'indice se révèlent relativement robustes aux changements de méthodes ou de variables, même s'il conviendrait encore, dans de futurs travaux, de poursuivre cette démarche de validation, en reproduisant cette analyse à d'autres jeux de données (par exemple, les données du panel d'élèves entrant en CP en 2011, ou encore des panels plus anciens).

L'indice devrait procurer un avantage pratique selon au moins deux axes d'analyse : décrire les inégalités sociales à l'école ; étudier les relations entre la performance scolaire et l'origine sociale. Dans tous les cas, l'utilisation de l'indice doit tenir compte de ses modalités de construction, afin d'assurer une interprétation correcte des résultats. Désormais, c'est l'application de cet indice à des problématiques concrètes qui permettra d'asseoir sa robustesse et son utilité.

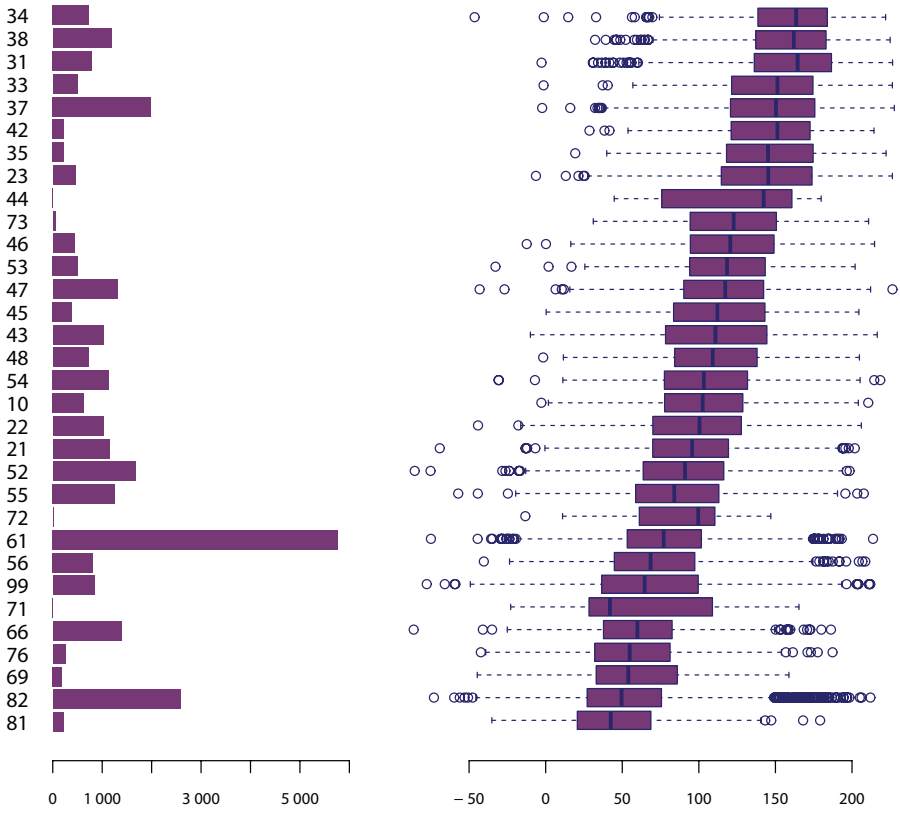
Annexes

↘ Tableau 5 Variables retenues

Variables	Indicatrices	Libellé
Diplômes des parents		
Diplôme de la mère	dip_12m	Sans diplôme ou CEP
	dip_3m	BEPC, brevet élémentaire, brevet des collèges
	dip_4m	CAP ou CAPA
	dip5m	BEP ou BEPA
	dip6m	Baccalauréat technologique ou professionnel
	dip7m	Baccalauréat général
	dip8m	Diplôme universitaire de premier cycle
	dip9m	Diplôme universitaire de deuxième ou troisième cycle
	dipmnr	Non-réponse
Diplôme du père	dip12p	Sans diplôme ou CEP
	dip3p	BEPC, brevet élémentaire, brevet des collèges
	dip4p	CAP ou CAPA
	dip5p	BEP ou BEPA
	dip6p	Baccalauréat technologique ou professionnel
	dip7p	Baccalauréat général
	dip8p	Diplôme universitaire de premier cycle
	dip9p	Diplôme universitaire de deuxième ou troisième cycle
	dippnr	Non-réponse [4mm]
Conditions matérielles		
Revenus	rev1 à rev10	Revenus mensuels du foyer (par déciles)
Nombre de pièces du logement	pieces123	3 pièces et moins
	pieces4	4 pièces
	pieces5	5 pièces
	pieces6	6 pièces
	pieces7p	7 pièces et plus
	piecesnr	Non-réponse
Chambre	partagchamb	Partage sa chambre
	chambseul	A une chambre seul
	partagchambnr	Non-réponse
Ordinateur	ordi	Ordinateur au domicile
Internet	internet	Accès Internet dans le foyer
Capital culturel		
Nombre de livres	nblivmoins30	Moins de 30 livres
	nbliv3099	Entre 30 et 99 livres
	nbliv100plus	Plus de 100 livres
	nblivnr	Non-réponse
Présence d'une télévision dans la chambre	televcham	Télévision dans la chambre
	ntelevcham	Non
	televchamnr	Non-réponse
Temps passé devant la télévision	televreg	Regarde régulièrement la télévision
	televnreg	Non
	televnr	Non-réponse

Variables	Indicatrices	Libellé
Ambition et implication		
Aspiration	aspidip1	Diplôme de niveau inférieur au baccalauréat
	aspidip2	Baccalauréat technologique ou professionnel
	aspidip3	Baccalauréat général pas forcément scientifique
	aspidip4	Baccalauréat scientifique
	aspidip5	Ne sait pas
	aspidipnr	Non-réponse
Diplôme le plus utile	diputil1	Diplôme de niveau inférieur au baccalauréat
	diputil2	Baccalauréat technologique ou professionnel
	diputil3	Baccalauréat général
	diputil4	Diplôme enseignement supérieur
	diputil5	Ne sait pas
Implication des parents	implic0 à 4	Indicateur : relations parents-enseignants, parents délégués, aide des parents
Conversations (vie scolaire)	conversviescolts	Les 5 sujets scolaires régulièrement abordés
	nconversviescolts	Non
	conversviescoltsnr	Non-réponse
Conversation (avenir scolaire)	conversav	Les 2 sujets avenir scolaire régulièrement abordés
	nconversav	Non
	conversavnr	Non-réponse
Pratiques culturelles		
Événement sportif au cours des 6 derniers mois	match0	Jamais
	match1	Une seule fois
	match2p	Plusieurs fois
	matchnr	Non-réponse
Concert au cours des 6 derniers mois	concert0	Jamais
	concert1	Une seule fois
	concert2p	Plusieurs fois
	concertnr	Non-réponse
Théâtre au cours des 6 derniers mois	theatre0	Jamais
	theatre1	Une seule fois
	theatre2p	Plusieurs fois
	theatrenr	Non-réponse
Cinéma au cours des 6 derniers mois	cinema0	Jamais
	cinema1	Une seule fois
	cinema2p	Plusieurs fois
	cinemanr	Non-réponse
Musée au cours des 6 derniers mois	musee0	Jamais
	musee1	Une seule fois
	musee2p	Plusieurs fois
	museenr	Non-réponse
Activités extra-scolaires de l'élève	activpeu	1 activité ou moins
	activmoy	2 activités
	activbcp	Plus de 2 activités

↘ Figure 5 Dispersion des scores factoriels par PCS (responsable)



Note : le graphique de gauche représente l'effectif de chaque PCS du responsable dans l'échantillon du panel. Les « boîtes à moustaches » à droite représentent la répartition des scores factoriels selon les PCS. Les PCS sont classées de manière décroissante selon le score factoriel moyen (c'est-à-dire l'indice).

↘ Figure 6 Indices selon l'inclusion ou non des variables sur l'ambition et l'implication des parents

PSC1	Indice	Indice sans les variables sur l'ambition	Différence	Indice sans les variables sur l'ambition ni celles sur l'implication	Différence
10	104	106	2	110	6
21	95	97	2	98	3
22	101	100	0	100	- 1
23	142	142	0	145	2
31	158	156	- 1	160	2
33	148	147	0	149	2
34	159	158	- 1	158	- 1
35	143	143	- 1	135	- 8
37	147	145	- 2	147	0
38	158	157	- 1	161	3
42	145	145	0	143	- 2
43	112	111	- 1	110	- 2
44	134	133	- 1	144	10
45	112	111	- 1	108	- 4
46	121	119	- 2	117	- 3
47	116	116	- 1	117	1
48	110	110	0	112	2
52	90	91	0	89	- 1
53	118	117	- 1	120	2
54	105	103	- 2	101	- 4
55	87	86	- 1	85	- 2
56	73	73	0	69	- 4
61	79	81	1	81	2
66	62	63	2	62	0
69	58	61	3	62	4
71	68	69	1	89	22
72	83	81	- 3	74	- 9
73	125	121	- 3	124	- 1
76	60	60	0	58	- 2
81	48	47	0	42	- 6
82	55	55	0	52	- 3
99	70	71	1	69	- 2

↘ BIBLIOGRAPHIE

Arzoumanian P., Dalibard E., 2015, « Cedre 2014. Mathématiques en fin de collège : une augmentation importante du pourcentage d'élèves de faible niveau », *Note d'information,* n° 15.19, MENESR-DEPP.

Baron R. M., Kenny D. A., 1986, "The Moderator-Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations", *Journal of Personality and Social Psychology,* n° 51, p. 1173-1182.

Ben Ali L., Vourc'h R., 2015, « Évolution des acquis cognitifs au collège au regard de l'environnement des élèves. Constat et mise en perspective longitudinale », *Éducation & formations,* n° 86-87, MENESR-DEPP, p. 211-234.

Bourdieu P., 1979, *La distinction. Critique sociale du jugement,* Paris, Éditions de Minuit.

Caille J.-P., 2014, « Les transformations des trajectoires au collège : des parcours plus homogènes mais encore très liés au passé scolaire et à l'origine sociale », *Éducation & formations,* n° 85, MENESR-DEPP, p. 5-30.

Chambaz C., Maurin E., Torelli C., 1998, « L'évaluation sociale des professions en France : Construction et analyse d'une échelle des professions », *Revue française de sociologie,* vol. 39, n° 1, p. 177-226.

Ganzeboom H. B. G., De Graaf P. M., Treiman D. J., 1992, "A Standard International Socio-Economic Index of Occupational Status", *Social Science Research,* vol. 21, n° 1, p. 1-56.

Givord P., Guillerm M., Afsa C. Monso O., Murat F., 2015, « Quels outils pour mesurer la ségrégation dans le système éducatif ? Une application à la composition sociale des collèges français », *XII^e Journées de méthodologie statistique,* Paris.

Keskpaik S., Rocher T., 2011, « La mesure de l'équité dans PISA : pour une décomposition des indices statistiques », *Éducation & formations,* n° 80, MENJVA-DEPP, p. 69-78.

Le Donné N., Rocher T., 2010, « Une meilleure mesure du contexte socio-éducatif des élèves et des écoles. Construction d'un indice de position sociale à partir des professions des parents », *Éducation & formations,* n° 79, MENJVA-DEPP, p. 103-115.

Ly S.-T., Riegert A., 2015, *Mixité sociale et scolaire et ségrégation inter- et intra-établissement dans les collèges et lycées français,* Rapport au Conseil national d'évaluation du système scolaire.

MENESR-DEPP, 2014, *Géographie de l'école,* n° 14, Paris, MENESR-DEPP.

OCDE, 2013, *Résultats du PISA 2012 : savoirs et savoir-faire des élèves, volumes 1 à 5,* Paris, OCDE.

Soulié C., 2000, « L'origine sociale des collégiens et des lycéens en France : une analyse des conditions sociales de production de la statistique », *Population,* vol. 55, n° 1, p. 169-181.

Vallet L.-A., Caille J.-P., 1996, « Les élèves étrangers ou issus de l'immigration dans l'école et le collège français. Une étude d'ensemble », *Les dossiers d'Éducation & formations*, n° 67, MEN-DEP.