Revue française de pédagogie
Recherches en éducation

161 | octobre-décembre 2007
Varia

Progrès statistiques et glissements conceptuels dans l'analyse des inégalités sociales à l'école

Progress in statistics and conceptual changes in analysing social inequalities at school

Progresos estadísticos y desplazamientos conceptuales en el análisis de las desigualdades sociales en la escuela

Fortschritte im Bereich der Statistik und konzeptuelle Verschiebungen bei der Analyse sozialen Ungleichgewichts in der Schule

Marie Duru-Bellat

OpenEdition Journals

Édition électronique
URL : https://journals.openedition.org/rfp/793
DOI : 10.4000/rfp.793
ISSN : 2105-2913

Éditeur
ENS Éditions

Édition imprimée
Date de publication : 1 décembre 2007
Pagination : 5-14
ISBN : 978-2-7342-1112-9
ISSN : 0556-7807

Référence électronique
Marie Duru-Bellat, « Progrès statistiques et glissements conceptuels dans l'analyse des inégalités sociales à l'école », *Revue française de pédagogie* [En ligne], 161 | octobre-décembre 2007, mis en ligne le 01 décembre 2011, consulté le 12 mars 2026. URL : http://journals.openedition.org/rfp/793 ; DOI : https://doi.org/10.4000/rfp.793

Le texte seul est utilisable sous licence CC BY-NC-ND 4.0. Les autres éléments (illustrations, fichiers annexes importés) sont susceptibles d'être soumis à des autorisations d'usage spécifiques.

Progrès statistiques et glissements conceptuels dans l'analyse des inégalités sociales à l'école

Marie Duru-Bellat

Dans la recherche, les questions de méthode sont toujours des questions de fond. Le texte pointe l'évolution des méthodes statistiques employées par les chercheurs pour analyser les inégalités sociales à l'école, évolutions qui apparaissent comme un progrès même si les nouveaux outils ont évidemment aussi leurs limites, pour ensuite s'interroger sur les glissements conceptuels dont ils peuvent être le vecteur. Mais être conscient du caractère construit et discutable des instruments d'objectivation statistique ne saurait conduire à s'en priver, ce qui serait extrêmement dommageable pour une sociologie soucieuse d'une explication globale des inégalités et consciente de ses enjeux sociaux.

Descripteurs (TEE) : inégalités sociales-sociologie de l'éducation-technique de mesure-méthode statistique.

INTRODUCTION

La question des inégalités sociales est au cœur de la sociologie de l'éducation française depuis son essor théorique et empirique des années soixante. À cet égard, les chercheurs ont repris à leur compte l'adage de Marcel Mauss selon lequel « tout problème social est un problème statistique » et personne ne conteste qu'il convient de documenter cette question par des analyses chiffrées soigneuses et actualisées. Certes, il est tout aussi classique de questionner les limites de ces approches dites « quantitatives » et les multiples biais inhérents à cette « construction de l'objet » particulière qu'elles imposent. Mais de quelles approches « quantitatives » parle-t-on ? On a assisté depuis une vingtaine d'années à des changements importants dans les méthodes statistiques utilisées et c'est en toute connaissance de cause, face à ces approches d'aujourd'hui, qu'il convient de reposer ces questions.

Après avoir présenté à grands traits les évolutions des méthodes statistiques utilisées pour décrire les inégalités, nous nous demanderons si, comme toute méthode, elles n'induisent pas une manière particulière de penser ce qu'elles mesurent, qui gagne à être explicitée. Ceci est d'autant plus nécessaire que ces débats méthodologiques ne sont pas dénués d'incidences politiques et qu'ils questionnent aussi la façon dont le chercheur conçoit son métier.

L'ÉVOLUTION DES MÉTHODES : PROGRÈS ET LIMITES

Depuis Durkheim, la sociologie se donne pour objectif de mettre au jour les régularités qui structurent le monde social et résister à l'approche statistique serait nier l'existence de relations systématiques exprimant des déterminations proprement sociales. Mais la mise en évidence de ces régularités, comme toute démarche empirique, expose à de multiples biais et artefacts, qui peuvent être méconnus un temps et apparaître à la lueur de nouvelles méthodes qui se présentent alors comme un progrès.

Des progrès…

Ainsi, qu'il s'agisse des inégalités d'accès à un niveau éducatif ou de réussite scolaire, la sociologie s'en est longtemps tenue à des tableaux croisés. Pourtant, on sait combien, à maints égards, cela oriente et limite les analyses.

Tout d'abord, cela produit une certaine image de l'ampleur des inégalités. Par exemple, si l'on construit un tableau croisant les scores à des épreuves de connaissances et l'origine sociale, on met sans peine en évidence des écarts statistiquement significatifs entre les enfants d'ouvriers et les enfants de cadres, sachant qu'il convient de rapporter les différences de scores moyens à l'écart-type de la distribution, pour mieux en apprécier l'échelle et donc l'ampleur. Le plus souvent, face à ce « résultat », on conclura à la force des déterminismes sociaux à l'école. Mais si, analyse statistique plus élaborée, on estime la part de variance des scores expliquée par l'origine sociale, on trouvera en général qu'environ 10 % de cette variance est expliquée par l'origine sociale, ce qui invite non seulement à une vision moins massive de l'influence de l'origine sociale mais aussi et surtout à rechercher d'autres facteurs ou processus susceptibles d'influer sur les scores. Car le meilleur des tableaux ne peut révéler que le solde des variables multiples qui régissent le phénomène étudié. Et ce solde est parfois nul, ce qui peut être extrêmement trompeur. Ainsi, dans les premières enquêtes sur la réussite à l'université conduites dans les années soixante dix à quatre vingts, il était courant d'observer, dans les tableaux croisant réussite et origine sociale, une absence totale de relation. Loin de traduire une démocratisation parfaite de la réussite à ce niveau, cette absence de relation résultait en fait de deux relations sous-jacentes, jouant en sens inverse : d'une part une forte sur-sélection scolaire des étudiants de milieu défavorisé, exerçant un effet positif sur leurs chances de réussite, et d'autre part, chez ces mêmes étudiants, des conditions de vie (notamment l'exercice d'une activité rémunérée) venant au contraire obérer fortement leurs chances. Cet exemple daté, souligne que les mécanismes en jeu dans la production des inégalités sociales impliquent un bien trop grand nombre de variables pour pouvoir être étudiés sur la seule base de tableaux croisés.

Certaines techniques statistiques permettent de prendre en compte simultanément plusieurs variables. C'est le cas des diverses méthodes dites d'analyse des données (analyses factorielles, analyses des correspondances) qui proposent sous la forme de « cartes » de lecture aisée, une description de la manière dont s'agencent les multiples facteurs qui structurent un phénomène donné (l'accès aux grandes écoles comme dans *La noblesse d'État* de Pierre Bourdieu). Mais on leur préfère souvent les analyses de variance ou de régression (notamment dans la recherche étrangère, où leur usage est complètement routinisé) ; car si, comme les premières, elles introduisent simultanément, pour expliquer la variance d'un phénomène, de multiples facteurs, elles permettent, de plus, d'estimer leurs effets respectifs, les autres étant constants. On parle alors de modélisation multivariée. Cette modélisation est évidemment une abstraction, puisque, pour expliquer statistiquement un comportement, on ne retient que les variables dont on a des raisons de penser qu'elles l'influencent, soit des raisons théoriques, soit l'accumulation de résultats antérieurs. C'est également une abstraction en ce sens que cela amène à démêler l'influence de variables qui dans la réalité se présentent imbriquées. Mais il suffit parfois d'une modélisation très simple pour renverser la vision qu'on a d'une « inégalité » ; par exemple, si le taux de passage des filles, à un palier d'orientation, apparaît supérieur à celui des garçons, une modélisation incluant les niveaux scolaires montre souvent que ceci ne révèle pas un avantage des filles (ou une discrimination frappant les garçons) mais le fait qu'elles possèdent davantage les atouts scolaires pris en compte lors des décisions d'orientation, et qu'au contraire, si l'on intègre dans les modèles multi-variés ces atouts scolaires, elles passent plutôt moins que les garçons (ou, dit autrement, qu'elles devraient même passer encore un peu plus vu leurs résultats). Les exemples sont nombreux de constats faits sur la base de tableaux qui se trouvent balayés par les modélisations multi-variés. Les tableaux, même s'ils sont nécessaires, sont donc insuffisants.

Un autre apport de la modélisation est de fournir une estimation chiffrée des différents effets. Les coefficients estimés dans les modèles donnent certes un ordre de grandeur (qu'il convient de relativiser, nous y reviendrons), mais ils sont comparables au sein d'un même modèle et permettent d'ordonner raisonnablement l'influence de différentes variables. C'est en soi un résultat : si tout le monde s'accorde pour dire que tant les notes que l'âge ou l'origine sociale jouent sur l'orientation, par exemple, il n'est pas trivial d'ordonner précisément ces influences et d'estimer que telle ou telle de ces variables joue deux fois plus que telle autre. De plus, les modèles multivariés donnent de ces effets une vision nuancée (1) ; ils permettent notamment de déceler des effets de seuil et des effets de saturation. Alors que le temps alloué à une discipline apparaît comme un ingrédient majeur des acquis des élèves, allonger le temps d'enseignement ne s'avère profitable que dans tel ou tel intervalle (pas tant que l'on reste en deçà d'un niveau minimal, et on parlera d'effet de seuil, et pas au-delà d'un certain niveau, et on parlera d'effet de saturation). Ces seuils, il est intéressant de les déterminer et ceci tranche avec la vision dichotomique, en termes de tout ou rien, que l'on a parfois des divers facteurs susceptibles d'affecter les résultats des élèves.

Notons aussi que cette approche permet d'opérer une distinction entre effets découlant des caractéristiques personnelles des élèves et effets strictement contextuels. Lorsque des tableaux font apparaître que les performances sont en moyenne plus faibles dans les établissements de ZEP, il sera intéressant d'analyser ce phénomène par des modèles tenant compte à la fois de l'appartenance sociale ou du niveau académique des élèves d'une part (facteurs qui sont liés aux performances, en général et quel que soit le contexte) et d'autre part du contexte fréquenté, ici ZEP ou non ZEP, ou encore des variables contextuelles comme le pourcentage d'élèves de milieu populaire dans l'établissement. On peut ainsi estimer si des élèves comparables réalisent des performances différentes selon qu'ils sont ou non en ZEP et/ou selon qu'ils fréquentent un établissement plus ou moins populaire. Les modèles permettent aussi de déceler des effets d'interaction : ainsi, les effets du contexte scolaire, comme la composition sociale du public, apparaissent en général plus importants chez les élèves faibles et/ou défavorisés que sur les élèves sans problème.

Certes, dans la réalité, ces diverses variables sont fortement associées, les élèves de milieu populaire sont plus fréquemment en ZEP et dans des établissements à tonalité populaire. Mais précisément, un des apports les plus intéressants de la modélisation est de fonctionner comme un substitut aux expérimentations impossibles, grâce aux simulations que les modèles permettent de calculer (2). Dans une recherche visant à évaluer l'influence spécifique de la composition sociale de la classe fréquentée sur les progressions (Duru-Bellat & Suchaut, 2005), des simulations ont permis d'évaluer si l'effet contextuel de la composition sociale de la classe sur les progressions (qui s'était avéré significatif pour les élèves d'origine modeste) était susceptible, pour ces élèves, de combler le handicap tenant à leur propre origine sociale. En d'autres termes, pouvait-t-il y avoir compensation entre le poids des caractéristiques personnelles et des facteurs contextuels au cours d'une année scolaire donnée ? Les conclusions vont dans le sens d'une certaine compensation : dès lors que le groupe classe est composé d'au moins 20 % d'enfants de milieu favorisé (3), les enfants de milieu modeste réalisent (toutes choses égales par ailleurs) des progressions très légèrement meilleures que les enfants de cadres (sachant que c'est seulement dans la fourchette de 40 à 60 % d'enfants de milieu favorisé que l'écart de progression devient important). Cet écart est le solde de deux effets de sens inverse : l'influence négative sur les progressions de l'origine sociale individuelle de l'élève d'une part, de l'effet positif associé chez les élèves de milieu populaire à la fréquentation d'une classe de tonalité sociale plutôt favorisée d'autre part. Évidemment, ces simulations sont à interpréter en fonction des situations réelles : en l'occurrence, la probabilité pour un élève de milieu populaire de se trouver dans une situation dans laquelle la composition sociale de la classe compense le handicap de progression lié à son origine sociale n'est pas très élevée, vu la ségrégation sociale « ordinaire » qui prévaut ; néanmoins, ce résultat souligne la portée compensatoire de certaines situations de contexte.

Enfin, un autre apport de la modélisation est qu'elle permet de tenir compte du fait que les inégalités se « fabriquent » selon une séquence temporelle que les tableaux figent. Ainsi, face aux inégalités sociales d'acquis telles qu'observées à un instant t, en fin de cours préparatoire par exemple, il n'est pas trivial de distinguer les inégalités déjà présentes en début de période et celles qui se sont mis en place spécifiquement pendant cette année. Ce déroulement temporel correspond parfois à des phénomènes de nature différente. Si on s'intéresse aux inégalités d'accès à un niveau scolaire donné, le seul croisement entre origine sociale et orientation (taux de passage de 3^{e} en

2de par PCS) est insuffisant, même s'il fait apparaître des écarts très significatifs. Car on sait aujourd'hui qu'ils résultent de deux processus bien distincts : les inégalités de réussite scolaire d'une part, les inégalités d'orientation à valeur scolaire donnée de l'autre (valeur « contrôlée » dans un modèle), deux types d'inégalités qui relèvent d'interprétations différentes. En introduisant dans un même modèle à la fois la demande d'orientation des élèves et leur niveau scolaire (et d'autres caractéristiques, dont bien sûr l'origine sociale), on pourra estimer le poids respectif, dans les inégalités sociales d'accès à un niveau, des inégalités sociales qui affectent la réussite d'une part, des inégalités sociales qui découlent strictement des phénomènes d'orientation (en l'occurrence de la plus forte auto-sélection des enfants de milieu populaire, même quand le niveau de réussite est identique) d'autre part. Il est aujourd'hui admis, dans la littérature internationale (pour la Suède, Erikson & Jonsson, 2000) que les inégalités tenant spécifiquement à l'orientation pèsent d'un poids comparable aux inégalités de réussite. Ce type de résultat, dont les incidences tant théoriques que politiques sont importantes n'aurait pas pu être produit sur la seule base de tableaux croisés.

Les modélisations constituent sans aucun doute un progrès dans la compréhension des processus qui génèrent les inégalités sociales de carrières scolaires. Mais des risques d'artéfacts existent, comme avec toute mesure.

Des limites

Ces risques d'artefacts, les chercheurs qui modélisent y sont assurément très sensibles. Ainsi, les chercheurs qui ont étudié l'évolution des inégalités sociales d'accès à l'éducation tout au long du siècle polémiquent (Maurin, 2002 ; *versus* Vallet, 2005) – la tendance de longue période est-elle à la stabilité ou à une baisse légère de ces inégalités ? – en invoquant le type de modélisation statistique effectuée, ou la taille des échantillons des enquêtes prises en compte (puisqu'il est d'autant plus facile de déceler une évolution statistiquement significative que les échantillons sont de taille importante).

De plus, dans les modèles, l'effet spécifique d'une variable est sensible à son mode de catégorisation (la variable est-elle bien ou mal mesurée ?) ainsi qu'à celui des autres variables prises en compte : tous les « effets » mis en évidence sont relatifs aux autres variables incluses dans le modèle, dont il faut connaître les inter-relations pour comprendre les résultats, ce qui suppose d'avoir au préalable construit et analysé des tableaux ou des indices simples comme des coefficients de corrélations entre variables. Lorsque l'on croise la réussite scolaire des enfants avec la profession et le niveau d'instruction des deux parents, il arrive qu'on se rallie à une mesure grossière de l'activité maternelle (active ou non), pour des raisons d'effectifs ; dans ce cas, on observe souvent un effet de l'instruction maternelle plus fort que celui de l'instruction paternelle, que l'on renvoie intuitivement à l'investissement plus fort des mères dans le suivi scolaire des enfants. En fait, ce « résultat » est peut-être en partie un artefact dû aux modes de catégorisation des variables : chez les pères, une partie de l'effet de l'instruction est « absorbé » par l'effet de leur profession, alors que cela vaut moins chez les mères, dès lors que l'activité est appréhendée moins finement.

Plus fondamentalement, il est clair que toutes les « choses » ne sont jamais contrôlées et que la quête de l'« effet pur » d'une variable est forcément chimérique. La clause « toutes choses égales par ailleurs » expose à un risque de « sociologie fiction » redoutable (discuté notamment par Passeron, 1991). C'est une fiction de raisonnement expérimental, souvent « limite », précisément parce que ce raisonnement expérimental est évidemment très éloigné de la réalité. Un exemple grossier (et donc parlant) en est l'introduction simultanée dans un modèle de l'origine sociale et de l'origine ethnique, pour isoler un effet de l'origine ethnique « toutes choses égales par ailleurs » ; parce qu'évidemment, rien n'est égal par ailleurs entre des personnes d'origine ethnique différente, dans la France d'aujourd'hui. Comme le disait Simiand, les comparaisons « toutes choses égales par ailleurs » reviennent à se demander « comment vivrait un chameau si, restant chameau, il était transplanté dans les régions polaires et comment vivrait un renne si, restant renne, il était transplanté dans le Sahara ». Car, quelle que soit la sophistication des analyses, les résultats sont nécessairement contextualisés (4) : ils sont produits dans un contexte caractérisé par telles ou telles « conditions initiales » ou telle ou telle association entre les variables, qu'il faut décrire, ce que l'on ne fait jamais assez, en associant des tableaux ou des descriptions « qualitatives » aux analyses modélisées les plus complexes.

De plus, cette notion d'« effet net » d'une variable peut être discutée, et les tenants des analyses multivariées globales de type analyse des correspondances rappellent avec raison l'intérêt qu'il y a à prendre en compte des configurations de variables. Mais

il n'en demeure pas moins que certains processus fondamentaux pour qui s'intéresse aux inégalités ne peuvent être appréhendés qu'avec des méthodes analytiques comme les régressions. Ainsi, en permettant d'appréhender l'« effet net » d'une variable (qui est certes une abstraction), seuls des modèles multivariés permettent de penser la notion de discrimination : concrètement, je ne peux conclure à une discrimination envers les élèves étrangers sur la seule base du constat d'une moindre orientation vers telle filière, mais seulement si cela s'observe quand on tient compte de tous les facteurs qui ont un impact sur cette orientation (valeur scolaire identique, demande identique, milieu social identique, etc.). La discrimination est en effet de l'ordre de quelque chose qui se passe différemment selon une caractéristique donnée (le sexe, l'origine ethnique, etc.), toutes choses égales par ailleurs, parmi les mécanismes connus à l'instant t.

DE QUELQUES GLISSEMENTS CONCEPTUELS DANS L'ANALYSE DES INÉGALITÉS

Quels que soient les outils statistiques mobilisés, ce sont toujours des hypothèses, des idées, que l'on teste. Il est trivial de souligner qu'aucune mesure n'est « neutre », en ce qu'elle repose sur des classifications et des mises en correspondances qui ne sont jamais aléatoires. Comme les tableaux croisés, la modélisation multi-variée véhicule une façon de penser les phénomènes scolaires et les inégalités afférentes.

La « production » de la réussite

Les modèles multi-variés diffusent certaines grilles de lecture des phénomènes scolaires ; en l'occurrence « expliquer » statistiquement une performance (ou un taux de passage) en la mettant en regard avec un certain nombre de paramètres caractérisant l'élève et/ou son environnement scolaire, c'est susciter une lecture en termes de « fonction de production » où les « outputs » résultent des « inputs ». On se centre ainsi sur l'« efficacité interne » des systèmes de formation et cette efficacité est évaluée dans un temps donné à l'aune des progressions réalisées au cours d'une année scolaire et qui s'avèrent mesurables.

Ceci éclaire certaines questions cruciales (comme celle des effets visibles des pratiques pédagogiques) et en laisse d'autres dans l'ombre, comme celle des effets d'accumulation avec le temps. C'est le cas des effets, sur la réussite des élèves, de la tonalité sociale du milieu scolaire. Ces effets sont faibles dès lors que l'on se centre sur les progressions enregistrées une année donnée (on sera alors tenté de dire que les caractéristiques personnelles importent bien plus que le contexte) ; mais ils sont bien plus forts si l'élève ne change pas de contexte, ce qui est souvent la règle. Il reste que si ce genre de limitation « technique » est au demeurant relativement facile à contrer, il en va tout autrement de l'orientation globale inhérente à cette perspective modélisante dominée par l'analogie avec une « production ». Un certain nombre de chercheurs (Derouet, 2006 ; Normand, 2006 ; Poupeau, 2003) dénoncent, pour le déplorer, le fait qu'elle a contribué à déplacer le regard des chercheurs vers le rôle propre des phénomènes scolaires dans la genèse des inégalités, avec ces « effets », discutés, que sont les effets maîtres, les effets des pratiques pédagogiques. Mais s'il est vrai que la montée en puissance des analyses de la « production » du système éducatif peut avoir pour effet pervers une polarisation sur les effets les plus facilement mesurables de l'action éducative, on peut rétorquer qu'elle donne aussi des outils pour une vision moins « fataliste » de l'échec scolaire et des inégalités afférentes. Et rien n'empêche de situer ces analyses de la « production » des inégalités dans le contexte concret de l'école dans un cadre d'analyses plus « macro », comme le montre la thématique de la démocratisation.

L'analyse de la démocratisation

La façon dont les sociologues de l'éducation posent la question de la démocratisation a beaucoup évolué, les outils se sont également profondément renouvelés. Mais là encore, s'agit-il d'un simple progrès technique ? Oui, à certains égard.

Dans l'étude de l'évolution des inégalités à l'école, on sait aujourd'hui que le type de mesure retenu (différence entre pourcentages ou rapport de chances) n'est pas sans incidence sur le résultat produit. On sait ainsi que « appliqués à des taux croissants, rapports et taux de variation concluent plus souvent à une réduction des inégalités que les mesures d'écart » (5) (Combessie, 1984). On sait aussi combien ces diverses mesures sont sensibles au niveau des chiffres initiaux. Dès lors, l'évolution des taux bruts (d'accès à tel niveau) et des écarts entre groupes sociaux à cette aune va mêler ce qui relève d'une réelle modification des inégalités et ce qui reflète l'évolution globale de la scolarisation aux niveaux considérés, sachant que toute élévation des taux tend à réduire

les écarts, dès lors que l'on se rapproche de 100 %. Ainsi, si l'on confronte les écarts de taux d'accès au supérieur entre groupes sociaux dans deux pays où le supérieur est très inégalement développé (80 % d'une classe d'âge *versus* 20 %), il sera difficile de comparer spécifiquement les inégalités entre les deux pays, car elles seront de toutes façons plus marquées quand l'accès au supérieur concerne une minorité d'une génération que lorsqu'il est quasi universel. C'est pour cette raison que l'on privilégie, dans les comparaisons des inégalités dans le temps et dans l'espace, des mesures qui se veulent « insensibles aux marges » comme les odds ratios, qui vont de pair avec les analyses modélisées : il s'agit de rapports concurrentiels de chances, par rapport à une « chance » qui peut être l'obtention d'un diplôme ou l'accès à un niveau d'études, rapports ayant l'avantage d'être indépendants du poids numérique et des catégories et des chances elles-mêmes.

Mais comme tout indicateur, les *odds ratios* expriment une conception particulière de l'inégalité. Ils mesurent spécifiquement la compétition entre deux groupes pour l'accès à un bien, et évaluent le caractère juste de cette compétition ; ils sont par là congruents avec la notion d'égalité des chances. Ils ne disent rien sur l'ampleur de l'écart observé au terme de cette compétition. C'est par contre ce que font les différences entre pourcentages (et les autres mesures plus frustres qui intègrent les « marges », c'est-à-dire le niveau des taux eux-mêmes), qui évaluent davantage une égalité de résultat : c'est à l'aune de l'évolution de l'ampleur de l'écart entre groupes que l'on évalue la démocratisation. Si les taux de scolarisation se sont élevés, les inégalités entre groupes, dans l'accès à tel niveau, peuvent s'être resserrées (si les taux du groupe le plus en tête sont proches de 100 %, l'expansion concerne surtout les groupes moins scolarisés, qui se rapprochent des premiers) ; le niveau scolaire atteint par les jeunes peut alors être moins contrasté, même si la compétition reste en elle-même aussi inégalitaire.

La question de savoir s'il faut privilégier les mesures « insensibles aux marges » concerne l'appréhension des inégalités en général : les inégalités de revenu peuvent être stables, alors que le niveau de vie moyen a doublé, ce qui n'est pas sans importance pour le bien-être de chacun. En filigrane, faut-il concevoir la pauvreté comme un concept absolu (avec la définition d'un seuil de pauvreté défini en valeur absolue, parce que correspondant à des « besoins » incompressibles) ou comme un concept relatif ? (par rapport à un niveau de revenu moyen, 50 % de ce revenu moyen par exemple, l'accent étant alors mis sur l'écart avec la situation moyenne qui prévaut dans le pays, avec les risques d'exclusion afférents). Ces questions, indissociablement conceptuelles et méthodologiques sont rarement posées dans le domaine de l'éducation. Pourtant que l'éducation soit tenue pour un bien absolu (avec définition corollaire d'un seuil minimal) ou pour un bien relatif affecte aussi bien les débats sur les inégalités scolaires que la façon dont les chercheurs vont les mesurer.

On peut se demander si privilégier les *odds ratios*, soit les mesures de chances relatives d'atteindre, pour tel groupe social, tel niveau de formation, comme cela tend à devenir la norme, du moins au niveau international (mais aussi, de manière quasi routinisée, par la Direction de l'évaluation, de la prospective et de la performance, DEPP) ne contribue pas à diffuser une conception particulière de la démocratisation et en filigrane de l'éducation elle-même. En l'occurrence, on entérine l'idée d'une compétition qui doit « simplement » être la plus ouverte possible : qu'importe, à la limite, que les uns quittent le système sans diplôme et les autres aillent à Polytechnique, l'essentiel est que l'égalité des chances et la méritocratie aient prévalu. C'est une conception de l'égalité et de l'éducation, cette dernière étant conçue avant tout comme un bien instrumental et positionnel, dont la valeur dépend de ce que possèdent les « concurrents ». Mais c'est une conception de la justice à l'école parmi d'autres possibles (Dubet & Duru-Bellat, 2004).

La problématique dominante de la sociologie de la mobilité sociale semble ici « déteindre » sur l'analyse des inégalités scolaires, quand elle juge une société à sa « fluidité » (avec une concurrence non faussée pour l'accès aux meilleures places) et non aux inégalités entre positions sociales elles-mêmes. En matière éducative, cela conduit à sur-valoriser la fonction de sélection et de classement de l'école, la scolarité étant analysée avant tout comme un canal de mobilité sociale. Cela conduit aussi à une vision particulière de l'évolution des inégalités sociales : dès lors qu'on se centre sur l'évolution de la compétition entre groupes sociaux, on conclura en général à la stabilité des inégalités sociales, qui se produisent « simplement » plus tardivement, puisque les scolarités s'allongent. Dans la littérature internationale, où ces mesures relatives et les modélisations afférentes constituent la manière « normale » d'étudier les inégalités, le constat dominant se résume au titre de l'ouvrage comparatif de Y. Shavit et H.P. Blossfeld paru en 1993 : *Persistent Inequality* (L'inégalité inchangée).

Pourtant, si on considère l'éducation comme un bien en soi, et que l'on mobilise des indicateurs non « épurés » des marges, qu'il s'agisse de taux bruts d'accès ou de scores à des tests (comme ceux passés il y a quelque années encore par les conscrits), on montre de manière toute aussi correcte que la diffusion de l'éducation a eu des effets égalisants sur les acquis de l'ensemble d'une classe d'âge. L'ouverture du système éducatif apparaît alors comme une voie pour combler certains écarts, même si elle ne fait que différer l'apparition des inégalités (ou de modifier leur forme, ces dernières devenant plus qualitatives) ; mais cela s'accompagne aussi d'un niveau éducatif plus élevé pour tous, ce qu'on peut considérer comme non négligeable. À des instruments de mesure différents correspondent donc deux facettes différentes des inégalités face à l'éducation et de l'éducation tout court.

ABANDONNER OU MIEUX MAÎTRISER LA SOCIOLOGIE DE L'ÉDUCATION « QUANTITATIVE » ?

Si personne ne soutiendrait, aujourd'hui, en France, qu'on puisse se passer d'une sociologie « quantitativement » armée des inégalités sociales à l'école, il reste que le décalage entre la sociologie de l'éducation française et la communauté européenne est tout à fait important. Structurée dans plusieurs réseaux d'excellence dont le réseau EQUALSOC, centré sur l'analyse des inégalités et qui compte un groupe sur les inégalités en éducation (6), elle produit de nombreuses recherches, le plus souvent comparatives, dans le temps et/ou dans l'espace, à partir de bases de données solides, qui conjuguent une grande technicité statistique et la volonté de tester empiriquement les grands cadres théoriques existants.

Ne pas abandonner les statistiques aux « experts »

Sur ces divers aspects, un certain nombre de difficultés se présentent en France. Tout d'abord, concernant les données, la situation française se caractérise par une abondance trompeuse. Il est certain que depuis une trentaine d'années, les données disponibles se sont considérablement développées, et permettent d'instruire de nouvelles questions. Ainsi, le développement d'épreuves standardisées de connaissance, qui commencent à enrichir les panels d'élèves, constitue un changement profond dans la manière d'analyser ce que produit l'école et les inégalités afférentes. Mais les fichiers administratifs, produits essentiellement par la DEPP, ne peuvent s'intéresser à tout, et dès lors occultent nécessairement certaines questions. C'est ainsi qu'ils ne permettent pas encore aujourd'hui d'appréhender de manière satisfaisante les phénomènes contextuels. En particulier, les panels échantillonnent les élèves sur la base de leur date de naissance, mode de sondage dit atomique qui extrait des élèves dispersés dans des milliers de classes et d'établissements. Seul un sondage en grappe, observant les élèves de classes et/ou d'établissements entiers permettrait d'explorer d'éventuels phénomènes de nature contextuelle et d'évaluer le rôle du contexte sur le déroulement des scolarités et les inégalités afférentes (7). Certes, quand les premiers panels ont été conçus (au début des années soixante dix), ce mode d'échantillonnage atomique s'est imposé car il était en phase à la fois avec une représentation du système scolaire comme centralisé et uniforme d'un lieu à l'autre, d'une part, d'autre part avec un modèle explicatif des carrières scolaires dominé par la notion d'héritage culturel et considérant corrélativement comme négligeable le rôle du contexte scolaire. Aujourd'hui, alors que l'on connaît l'importance des facteurs de contexte, ce type d'échantillonnage devient éminemment discutable. Par conséquent, s'il est de plus en plus tentant de pratiquer des analyses secondaires sur les données de la DEPP, cela limite les interrogations possibles. Ceci vaut aussi pour les données de l'INSEE (enquêtes emploi, enquête FQP, etc.). Si elles caractérisent bien l'origine sociale des personnes et leur situation, elles ne disposent que de mesures très frustres de la réussite scolaire (en général, seulement le retard scolaire) ou encore ne prennent jamais en compte le contexte scolaire, ce qui affecte évidemment les conclusions des travaux, par ailleurs sophistiqués sur le plan statistique, que réalisent les économistes de l'INSEE.

Dans l'idéal, il conviendrait de construire des données spécifiques, pour répondre aux questions que l'on se pose. Mais la taille des équipes de recherche en éducation et aussi la compétence statistique des chercheurs le permettent rarement. Le risque est alors que ces derniers se cantonnent dans l'analyse des « processus fins » qui prennent place au niveau « micro » ou « méso », en laissant aux économistes ou aux chercheurs de la DEPP les analyses « macro » des carrières scolaires, sans que ces divers niveaux d'analyse, tous légitimes, ne soient jamais intégrés. Mais il est sans doute trop facile d'accuser le développement des « statistiques d'État » de « déposséder »

le milieu des sociologues de l'usage des statistiques (comme le suggère Poupeau, 2003). Car rien n'empêche les chercheurs d'exploiter ces données, à condition d'en être capable techniquement.

Or, il est permis de penser qu'aujourd'hui, les sociologues de l'éducation sont souvent de plus en plus dépassés par les techniques statistiques sophistiquées qui se développent. Une certaine division du travail est sans doute, à terme, inévitable. Mais il est fondamental que les plus qualitativistes des chercheurs soient au moins capables de lire et d'évaluer les analyses chiffrées. C'est d'autant plus nécessaire que si chaque nouvelle méthode se présente comme permettant de « mieux » mesurer le phénomène, elle passe le plus souvent par l'introduction de postulats supplémentaires sur la nature des données, et introduit fréquemment une nouvelle dose d'arbitraire, non sans conséquence sur les données obtenues. D'où un certain vertige devant la multiplication des risques d'artefacts. Dans l'idéal, il faudrait être capable d'expliciter les hypothèses associées à chacune de ces mesures et de toujours évaluer, face à la mise en œuvre d'une technique apparemment meilleure, si ce qu'on gagne en technicité n'est pas compensé par les hypothèses ou les contraintes supplémentaires qu'on est amené à introduire. Ceci n'est évidemment pas facile, dès lors que le chercheur en sciences humaines n'est pas en général un statisticien.

En tout cas, ces difficultés ne devraient pas dissuader de recourir aux statistiques, au contraire. Certes, l'essentiel reste, dans le champ de l'éducation comme dans d'autre, la qualité des interrogations. Mais la réflexion méthodologique, à condition qu'elle ne soit pas trop bornée par des ignorances techniques, est une voie des plus heuristiques pour progresser dans l'explicitation et l'affinement des interrogations. De plus, quand on bute sans cesse sur des modèles qui chiffrent votre ignorance (quand on sait, par exemple, que l'on explique 30 % de la variance), cela rend à la fois modeste et infiniment curieux.

Rester avant tout des chercheurs

Une approche quantifiée peut également préserver de prises de position plus idéologiques que scientifiques, qui sont encore légion, dès lors qu'on traite d'éducation. Ainsi, le développement des mesures standardisées des acquis des élèves est-il encore parfois dénoncé dans notre pays comme un terrorisme productiviste ou une dérive managériale. Certes, la statistique a toujours eu maille à partie avec l'action politique et ses orientations sont à ce titre contestables. Ainsi, dans notre pays, la croyance méritocratique fait que l'on s'intéresse aux inégalités sociales dans l'accès aux diplômes plus qu'aux inégalités sociales qui marquent (ensuite) l'accès à l'emploi. Mais il n'est pas impossible de proposer aux politiques d'autres mesures (d'autres statistiques) de l'inégalité, en se fondant sur certaines théories de la justice, comme l'a fait récemment un groupe de chercheurs européens (GERESE, 2005).

Surtout, la statistique est aussi un moyen de connaissance. Le fait de dévoiler les processus concrets par lesquels l'école exerce au quotidien cette violence pédagogique que dénonçaient Bourdieu et Passeron en 1970 est une affaire sociologiquement trop sérieuse pour se passer de l'outillage statistique. On peut aussi considérer que s'intéresser non pas seulement à la production des inégalités d'accès (à tel ou tel niveau éducatif) mais tout autant aux inégalités d'acquis scolaires est en parfaite cohérence avec la sociologie de l'éducation des années soixante dix qui a amplement montré qu'il ne suffisait pas d'élargir l'accès pour égaliser véritablement les chances scolaires.

Certes, la recherche a vocation à verser ses résultats dans le domaine public, et à s'exposer ainsi à ce qu'ils soient utilisés. Certains sociologues proches du monde enseignant craignent que la diffusion large de statistiques sur les résultats des élèves vienne « accuser » ces derniers de quelque responsabilité en la matière. Ce souci n'est pas infondé. Il est incontestable que « l'évaluation par les standards » est la manière dont l'administration entend réguler le système, et c'est son droit, même si les fonctionnaires qui y oeuvrent peuvent estimer que « l'essentiel est ailleurs » (Derouet, 2006). Mais il est discutable de réduire l'administration au « management libéral » alors que l'intérêt général serait à l'évidence porté par ce groupe particulier (et impliqué) que sont les enseignants. La définition des critères d'évaluation est à l'évidence un enjeu, relevant du débat démocratique, mais peut-on s'opposer à une appréhension publique de ce qu'acquièrent les enfants à l'école ? Refuser l'objectivation chiffrée (ce qui n'interdit pas la nuance), publique et donc critiquable, renverrait à un obscurantisme préférant le huis clos entre enseignants et élèves où l'ineffable qui y est produit n'est pas nécessairement positif et juste.

C'est une vraie question de savoir si « les mêmes indicateurs peuvent répondre à la fois aux questions de l'administration centrale et à celle des acteurs » (Derouet, 2006). Il faudrait ajouter, des chercheurs, qui, au-delà de leur objectif d'élucidation des processus

sous-jacents aux « résultats », s'intéressent aussi à la construction sociale des normes en matière de résultats elles-mêmes. Mais une chose est sûre, il ne faut pas jeter les statistiques avec l'eau nécessairement conflictuelle de l'évaluation.

Pour rester sur le terrain de la recherche sociologique, on voit mal pourquoi, le souci d'évaluer spécifiquement ce qui se joue à l'école entraînerait fatalement à réduire l'échec scolaire à des causes scolaires, qui viendrait du même coup nier les déterminismes sociaux plus globaux (comme le craint, voire le dénonce, Poupeau, 2003). Comme si, au sein d'une communauté scientifique diversifiée, il y avait lieu d'opposer les explications par les structures aux explications par les processus ou les variables. Il est tout à fait possible de se focaliser sur des processus « micro » tels que les choix scolaires, en se référant aux théories de Boudon, tout en prolongeant ces analyses par des mises en perspective plus macro sociologiques (ce qu'on peut juger plus « bourdieusien »), cherchant à mettre en relation les stratégies scolaires avec les structures des systèmes éducatifs et des évolutions sociales plus générales (comme l'évolution du marché du travail).

Mais ce type de démarche, qui combine processus et structures, est indissociable de la manipulation de données d'enquêtes, sous-tendue par des compétences méthodologiques variées, notamment statistiques, et plus fondamentalement par la conviction que seuls les faits permettent de valider ou d'infirmer les théories dont c'est la destinée première que d'être mises à mal par les réalités. Les débats sur les questions d'éducation – par exemple celle cruciale comme l'importance de l'« effet maître » – peuvent ainsi se situer au niveau des faits et non des prises de position a priori. Une chose est sûre, les visées initiales de la sociologie de l'éducation exigent de travailler à un niveau « macro » (et donc statistique), notamment dans le domaine des inégalités scolaires et du rôle de l'école à cet égard. Renoncer à des analyses solides sur le plan statistique pour se tourner vers des approches qualitatives pas moins construites et pas moins entachées d'artefacts amènerait à délaisser certaines interrogations fondatrices de la sociologie de l'éducation, en se résignant à des monographies distrayantes ou à des expertises de moindre portée. Le « choix » des méthodes est alors tout, sauf anodin.

Marie Duru-Bellat
marie.duru-bellat@wanadoo.fr
Institut d'études politiques de Paris, Observatoire sociologique du changement et IREDU (CNRS)

NOTES

(1) Il convient pour cela d'estimer des modèles autres que linéaires (logistiques par exemple) ou de construire des variables spécifiques (une variable mise au carré, permet de tester d'éventuels effets exponentiels).

(2) Les simulations consistent, sur la base des coefficients des modèles qui expriment l'impact spécifique d'une variable, à estimer l'effet de telle ou telle variation d'une ou plusieurs variables, les autres restant constantes ; on peut aussi comparer de cette manière l'impact des changements induits par la variation d'une ou de deux variables du modèle : que devient la variable à expliquer, par exemple, sur une population où deux des variables explicatives prennent telle ou telle valeur.

(3) Soit à partir d'un seuil assez faible, puisque dans l'échantillon étudié, 40 % des enfants ont un père ou une mère cadre ou profession intermédiaire.

(4) En fait, Vallet (2005) montre que le risque de type « renne au Sahara » n'existe que si l'on utilise les coefficients des modèles pour simuler ce que serait la valeur de la variable dépendante pour une sous-population hautement improbable (comme les élèves issus de l'immigration et très favorisés socialement) ; mais en eux-mêmes, les coefficients de régression partiels sont nettement influencés par les situations caractéristiques des sous-populations qui sont les plus fréquemment rencontrées dans la réalité ; ils sont donc relativement réalistes.

(5) Pour une discussion appliquée à une recherche concrète, cf. Duru-Bellat & Merle, 2002.

(6) On consultera sur le site du réseau <www.nuffield.ox.ac.uk/projects/equalsoc/themes/educ> les textes présentés aux journées de ce groupe à Mannheim (décembre 2005).

(7) Certes, des techniques statistiques nouvelles prétendent approcher ces phénomènes avec un petit nombre d'élèves par établissement, et, par ailleurs, des appariements avec d'autres fichiers, des fichiers d'établissements notamment, peuvent permettre de dépasser ce problème.

BIBLIOGRAPHIE

COMBESSIE J.-C. (1984). « L'évolution comparée des inégalités : problèmes statistiques ». *Revue française de sociologie*, vol. XXV, n° 2, p. 233-254.

DEROUET J.-L. (2006). « Entre la récupération des savoirs critiques et la construction des standard du management libéral ». *Revue française de pédagogie*, n° 154, p. 5-18.

DUBET F. & DURU-BELLAT M. (2004). « Qu'est-ce qu'une école juste ? ». *Revue française de pédagogie*, n° 146, p. 105-115.

DURU-BELLAT M. & MERLE P. (2002). « De quelques difficultés à cumuler les savoirs sur les phénomènes éducatifs. L'exemple de la démocratisation de l'enseignement ». *Revue française de pédagogie*, n° 140, p. 65-74.

DURU-BELLAT M. & SUCHAUT B. (2005). « L'approche sociologique des effets du contexte scolaire : méthodes et difficultés ». *Revue internationale de psychologie sociale*, vol. XVIII, n° 3, p. 5-42.

ERIKSON R., JONSSON J. (2000). « Understanding Educational Inequality : The Swedish Experience ». *L'Année sociologique*, vol. L, n° 2, p. 345-382.

GERESE, Groupe européen de recherche sur l'équité des systèmes éducatifs. (2002). L'équité des systèmes éducatifs européens, Projet Socrates.

MAURIN E. (2002). *L'égalité des possibles. La nouvelle société française*. Paris : Seuil.

NORMAND R. (2006). « L'école efficace ou l'horizon du monde comme laboratoire ». *Revue française de pédagogie*, n° 154, p. 33-44.

PASSERON J.-C. (1991). *Le raisonnement sociologique. Le raisonnement non-popperien du raisonnement naturel*. Paris : Nathan.

POUPEAU F. (2003). *Une sociologie d'État. L'école et ses experts en France*. Paris : Éd. Raisons d'agir.

SHAVIT Y., BLOSSFELD H.P. (1993). *Persistent Inequality. Changing Éducational Attainment in Thirteen Countries*. Boulder : Westview Press

VALLET L.-A. (2005). « Concilier intérêt pour la science et intérêt pour la société », Note de synthèse pour l'HDR. Paris : EHESS.