Mesure et évaluation en éducation

# Les extensions des modèles multiniveau et leur application pour l'évaluation en éducation

Pascal Bressoux, Christine Leroy-Audouin et Paul Coustère

Volume 21, numéro 1, 1998

URI : https://id.erudit.org/iderudit/1091357ar
DOI : https://doi.org/10.7202/1091357ar

Aller au sommaire du numéro

Éditeur(s)

ADMEE-Canada - Université Laval

ISSN

0823-3993 (imprimé)
2368-2000 (numérique)



Citer cet article

Bressoux, P., Leroy-Audouin, C. & Coustère, P. (1998). Les extensions des modèles multiniveau et leur application pour l'évaluation en éducation. *Mesure et évaluation en éducation*, *21*(1), 39–59.
https://doi.org/10.7202/1091357ar

Résumé de l'article

Les modèles multiniveau ont été conçus pour modéliser des structures hiérarchisées, où les individus (par exemple, des élèves) sont groupés dans des unités d'un niveau supérieur (par exemple des classes). Ils permettent d'analyser les effets de l'environnement sur les comportements des individus. Nous nous attacherons ici plus spécifiquement aux effets de l'environnement scolaire sur les acquis des élèves dans le cadre des recherches sur ce qu'il est convenu d'appeler l'évaluation des effets-maîtres et des effets-écoles. Cet article, après avoir rappelé les fondements des modèles multiniveau, présente, sur la base de cas empiriques, de nouveaux champs d'application des modèles multiniveau. Ces champs d'application apparaissent particulièrement importants pour l'évaluation en éducation : il s'agit de la modélisation des courbes de croissance et de celle des effets aléatoires croisés.

# Les extensions des modèles multiniveau et leur application pour l'évaluation en éducation [1]

**Pascal Bressoux**
*Université Pierre-Mendès-France, Grenoble*

**Christine Leroy-Audouin**
*Université de Bourgogne, Dijon*

**Paul Coustère**
*Université de Bourgogne, Dijon*

MOTS-CLÉS : Modèles multiniveau, structures hiérarchisées, courbes de croissance, effets aléatoires croisés, effets-maîtres, effets-écoles

*Les modèles multiniveau ont été conçus pour modéliser des structures hiérarchisées, où les individus (par exemple, des élèves) sont groupés dans des unités d'un niveau supérieur (par exemple des classes). Ils permettent d'analyser les effets de l'environnement sur les comportements des individus. Nous nous attacherons ici plus spécifiquement aux effets de l'environnement scolaire sur les acquis des élèves dans le cadre des recherches sur ce qu'il est convenu d'appeler l'évaluation des effets-maîtres et des effets-écoles. Cet article, après avoir rappelé les fondements des modèles multiniveau, présente, sur la base de cas empiriques, de nouveaux champs d'application des modèles multiniveau. Ces champs d'application apparaissent particulièrement importants pour l'évaluation en éducation : il s'agit de la modélisation des courbes de croissance et de celle des effets aléatoires croisés.*

KEY WORDS : Multilevel models, hierarchical structures, growth curves, random crossed effects, teacher effects, school effects

*Multilevel models have been conceived to analyse hierarchical structures in which individuals (e.g., students) are nested within higher level units (e.g., classrooms). Multilevel models allow one to analyse the effects of the environment on individuals' behaviour. In this article, we focus on schooling environments and their effects on students' achievement within the framework of teacher effectiveness and school effectiveness research. After reviewing the principles of multilevel modelling, this article presents — on the basis of empirical data — new applications of multilevel models. These new applications are particularly fruitful for evaluation in education. They involve modelling of growth curves and of random crossed effects.*

## Introduction : modéliser des structures hiérarchisées

Beaucoup de recherches en sciences sociales prennent pour objet la question de l'effet de l'environnement dans lequel vivent les individus sur leurs comportements. C'est ainsi, par exemple, que se posent la question de l'effet de la famille sur le comportement des individus, la question de l'influence des groupes de pairs sur le comportement des jeunes, et ainsi de suite. Les données ont alors généralement une structure hiérarchisée dans laquelle les individus sont groupés dans des unités situées à un « niveau » supérieur (famille, groupe des pairs, etc.).

Les recherches en éducation qui tentent d'évaluer les effets-maîtres ou les effets-écoles font partie de cette catégorie de recherches : on se demande dans quelle mesure, et par quels processus, l'environnement scolaire d'un élève est susceptible d'influencer ses résultats scolaires. La structure hiérarchisée y est particulièrement apparente, puisque les élèves (niveau 1) apprennent dans des classes (niveau 2), qui sont elles-mêmes groupées dans des écoles (niveau 3)[2].

Or, l'analyse des effets de l'environnement sur les individus pose depuis longtemps de nombreux problèmes. On doit en effet mettre en relation des caractéristiques de l'environnement avec des variables qui caractérisent des individus. Il s'agit donc de prendre en compte des données qui appartiennent à des « niveaux » différents : un niveau que nous appellerons « agrégé » (celui de l'environnement) et un niveau que nous appellerons « individuel » (celui des personnes).

Dès lors se pose le problème de l'unité d'analyse à retenir. Deux solutions opposées s'offrent alors : ou bien on prend l'élève comme unité d'analyse et on « désagrège » les données de niveau supérieur au niveau des élèves. En ce cas, tous les élèves d'une même classe se voient attribuer une même valeur, puisqu'ils sont exposés à un facteur commun ; par exemple, l'ancienneté de leur maître ou encore la taille de la classe. Ou bien on choisit comme unité d'analyse celle qui relève du niveau agrégé (la classe ou l'école). En ce cas, on doit agréger les données individuelles à ce niveau. Les variables retenues pour l'analyse caractérisent donc la classe ; elles peuvent ne pas avoir de contrepartie au niveau individuel (taille de la classe, âge de l'enseignant, etc.), mais elles peuvent également résulter de l'agrégation de caractéristiques individuelles (niveau scolaire moyen de la classe, pourcentage d'enfants de cadres, etc.). La variable à expliquer change aussi évidemment de niveau et ce ne sont plus les résultats individuels qu'on explique, mais le résultat moyen de la classe.

Chacune de ces solutions comporte d'évidentes lacunes. Ainsi, dans la première, les variables qui caractérisent les classes ou les écoles se trouvent démultipliées au niveau des élèves. De ce fait, on ne respecte pas le nombre de degrés de liberté attaché à chacun des niveaux de la structure hiérarchisée, ce qui conduit à sous-estimer les erreurs-types de ces variables. On risque ainsi d'attribuer un effet significatif à tel élément du système scolaire qui, en fait, ne joue aucun rôle. Quant à la seconde solution, elle conduit au risque bien connu de biais d'agrégation : nombre de chercheurs ont démontré que la relation agrégée peut être très différente de la relation individuelle (*e.g.*, Hammond, 1973 ; Robinson, 1950). On ne peut donc déduire sans risque d'erreurs la relation individuelle sur la seule base d'une relation agrégée.

Par ailleurs, rien n'assure que les relations entretenues au niveau individuel soient identiques d'un contexte à l'autre, ce qui revient à dire que le contexte scolaire peut jouer un effet d'interaction avec les caractéristiques individuelles. Ainsi, son effet peut être plus marqué sur certaines catégories d'élèves ; les travaux montrent en général que ce sont les élèves en difficulté qui sont les plus sensibles aux effets des facteurs scolaires (Bressoux, 1994 ; Mingat, 1991).

Ces problèmes, connus des chercheurs depuis fort longtemps, ont été l'objet de nombreuses discussions (Blau, 1960 ; Boudon, 1963 ; Burstein, 1980 ; Cronbach & Webb, 1975 ; Erbring & Young, 1979 ; Firebaugh, 1978 ; Hammond, 1973 ; Hauser, 1970, 1974 ; Robinson, 1950), mais ce n'est que depuis l'avènement des modèles multiniveau[3] qu'on y a apporté une solution satisfaisante. Ces modèles ont été conçus au milieu des années 80 par quatre équipes de recherche qui, bien qu'ayant travaillé de façon indépendante, ont élaboré des techniques d'estimation qui produisent des résultats similaires (Aitkin & Longford, 1986 ; Goldstein, 1986 ; Mason, Wong & Entwisle, 1983 ; Raudenbush & Bryk, 1986). Les modèles multiniveau permettent d'introduire ensemble dans l'analyse les variables individuelles et les variables agrégées. Ils rendent de ce fait caduque la question du choix de l'unité d'analyse et, par là-même, permettent d'éviter le dilemme agrégation *versus* désagrégation. On peut en effet y intégrer directement et simultanément les différentes variables explicatives, chacune au niveau qui est le sien dans la structure hiérarchisée des données (les élèves au niveau 1, les classes au niveau 2, les écoles au niveau 3, etc.). On rend compte ainsi du fait que chaque élève apprend dans un milieu particulier et que celui-ci peut influer sur lui.

À l'origine, les modèles multiniveau ont été élaborés pour analyser les effets de l'environnement, conçu comme une structure hiérarchisée dans laquelle les individus sont groupés dans des unités qui se situent à un niveau d'agrégation supérieur. Par exemple, s'il s'agit de modéliser les effets respectifs de l'académie (niveau 4), l'école (niveau 3) et la classe (niveau 2) sur la probabilité d'obtention du baccalauréat par les élèves (niveau 1), une telle structure peut être représentée par la figure 1.

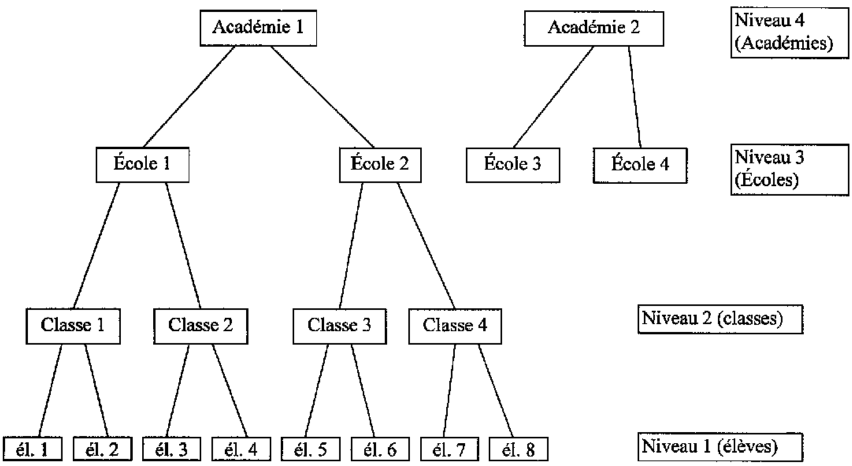


Figure 1. ***Structure hiérarchisée « classique » à quatre niveaux***

Au-delà de ce cas, certains phénomènes, en apparence très éloignés, peuvent cependant être conçus et modélisés comme des structures à plusieurs niveaux. Après avoir rappelé les principes des modèles multiniveau, notre propos sera précisément de montrer, à l'aide de cas empiriques, que ces modèles sont d'une grande souplesse et qu'ils permettent de traiter de façon satisfaisante deux autres grandes catégories de problèmes qu'on rencontre fréquemment dans l'évaluation en éducation : la modélisation des courbes de croissance et la modélisation des effets aléatoires croisés.

## Les principes des modèles multiniveau

### *Rappel sur les moindres carrés ordinaires et leurs limites*

Les procédures statistiques classiquement mises en œuvre dans l'évaluation des effets-maîtres ou des effets-écoles sont des modèles multivariés basés sur la régression par les moindres carrés ordinaires (MCO) dans lesquels la variable-réponse est exprimée comme une fonction des variables explicatives, selon l'équation suivante :

$$y_{ij} = \beta_0 + \beta_1 x_{ij} + e_{ij} \qquad [1]$$

où

$x_{ij}$ représente le score initial de l'individu *i* de la classe *j* ;

$y_{ij}$ représente le score final de l'individu *i* de la classe *j* ;

$\beta_0$ représente la constante (la valeur de y pour une valeur nulle de x) ;

$\beta_1$ représente la pente de la droite de régression ;

$e_{ij}$ représente une erreur aléatoire (appelée résidu) associée à chaque individu *i* de la classe *j* de moyenne nulle et de variance $\sigma^2$.

Un des problèmes du modèle MCO est qu'il repose sur des hypothèses statistiques qui sont souvent contradictoires avec les hypothèses de recherche. Ainsi, le modèle MCO repose-t-il sur l'hypothèse de l'indépendance des résidus. Or, si l'on cherche à évaluer les effets-classes, c'est bien qu'on émette l'hypothèse que les résidus peuvent varier d'une classe à l'autre et, par conséquent, qu'il ne sont pas indépendants les uns des autres. De fait, les résultats empiriques sur l'effet-maître contredisent les hypothèses statistiques des modèles qui ont permis de mettre en lumière ces résultats (Bressoux, 1994, 1995, 1996 ; Bru, 1991 ; Duru-Bellat & Leroy-Audouin, 1990 ; Grisay, 1993 ; Mingat, 1984, 1991). Or, la violation de l'hypothèse d'indépendance des résidus a pour effet d'entraîner un biais dans les erreurs-types des coefficients des variables, ce qui peut affecter leur significativité. On risque donc de conclure à l'effet d'une variable là où, en fait, il n'y en a pas.

Il est possible, dans un modèle MCO, d'évaluer les effets-classes en introduisant dans le modèle *n* – 1 variables muettes représentant chacune une classe, la *é*nième classe servant de référence. En agissant de la sorte, le coefficient de régression attaché à chaque classe traduit la différence entre le score moyen final de la classe considérée et la classe de référence. Toutefois, en traitant le problème de la sorte, on considère que les effets-classes sont des effets fixes, ce qui revient à dire que les effets sont attribués à un nombre défini de modalités de la variable, modalités choisies parce que c'est à elles qu'on s'intéresse spécifiquement (Searle, Casella & McCulloch, 1992). Autrement dit, la classe de M. Untel nous intéresserait justement parce c'est la classe de M. Untel.

Il est pourtant difficile de soutenir une telle assertion dans la recherche en éducation; nous n'étudions pas les classes pour elles-mêmes. Une classe nous intéresse en ce sens qu'elle est un élément sélectionné dans une population plus vaste dont la distribution des performances nous intéresse. Il s'agit pour nous d'une variable aléatoire, dont les effets sont donc aléatoires. Cette distinction a des conséquences importantes car on ne fait pas les mêmes inférences sur les effets aléatoires et sur les effets fixes. Tout d'abord, on n'est plus intéressé par les moyennes attribuées à chacune des modalités de la variable mais par la variance du phénomène. Or, dans un modèle de régression MCO (ou dans une analyse de la variance classique), les tests de significativité portent sur des moyennes; le test $F$ a pour hypothèse nulle $H_0$: $\beta_1 = \beta_2 = \beta_3 = \ldots = \beta_k = 0$. Mais nous sommes ici intéressés par $H_0$: $\sigma^2{}_\beta$ (où représente la variance des effets-classes).

Par les modèles MCO, on modélise des moyennes, tandis que dans les modèles multiniveau, on modélise des variances. Il faut bien comprendre que ces variances (d'une population infinie de classes) sont estimées et non pas observées. Elles sont estimées à partir d'observations provenant d'un échantillon qui n'est qu'une des multiples réalisations possibles. Cela revient à dire que le hasard intervient aussi dans le choix des classes et que, dès lors, la variance des paramètres est composée d'une variance «vraie» (seule potentiellement explicable) et d'une variance d'échantillonnage (par définition inexplicable). Les modèles multiniveau «débarrassent» les estimations de leur variance d'échantillonnage, ce qui produit un effet de «réduction» (*shrinkage effect*); *i.e.*, les estimations sont ramenées vers la moyenne des observations. Les modèles MCO ne prenant pas en compte cette variance d'échantillonnage, l'ampleur de l'effet-classe y est, de fait, généralement surestimée. Cette surestimation peut être plus ou moins forte; elle l'est d'autant plus que la variance d'échantillonnage est élevée (*i.e.*, que le nombre d'élèves par classe est faible). Par exemple, Bressoux, Coustère et Leroy-Audouin (1997) ont montré, dans une recherche portant sur 145 classes d'élèves de CE2 (troisième année élémentaire) ayant en moyenne 16 élèves, que la part de variance brute inter-classes des scores de ces élèves en mathématiques s'élevait à

30,0 % dans les modèles MCO, mais qu'elle n'était estimée qu'à 26,7 % par un modèle multiniveau.

De plus, un modèle MCO tel que [1] ne permet pas d'explorer d'éventuelles différences de relation entre *x* et *y* en fonction de la classe. On contraint de la sorte la relation à être identique d'une classe à l'autre. Or, cette hypothèse devrait être empiriquement testée[4]. C'est tout l'intérêt des modèles multiniveau que de résoudre ces problèmes.

### *Présentation des modèles multiniveau*

Une manière simple de présenter les modèles multiniveau consiste à séparer les équations correspondant aux différents niveaux de l'analyse, bien qu'en fait ces équations soient résolues simultanément[5]. On obtient :

au niveau 1 (celui des élèves) :

$$y_{ij} = \beta_{0j} + \beta_{1j}x_{ij} + e_{ij} \qquad [2]$$

au niveau 2 (celui des classes) :

$$\beta_{0j} = \gamma_{00} + \gamma_{01}z_j + u_{0j} \qquad [3]$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}z_j + u_{1j} \qquad [4]$$

Dans l'équation [3], l'erreur aléatoire $u_{0j}$ indique que les constantes $\beta_{0j}$ varient d'une classe à l'autre (d'où l'indice *j* qui leur est maintenant affublé dans [2]). La constante moyenne dans l'échantillon est représentée par $\gamma_{00}$ ; $u_{0j}$ représente l'écart de chaque classe par rapport à cette moyenne. Plus la constante d'une classe est élevée (*i.e.*, plus la valeur de $u_{0j}$ est forte), meilleur est le niveau moyen final des élèves de cette classe. On autorise ainsi des différences d'efficacité entre les classes. L'efficacité doit donc être entendue ici comme la capacité des classes à élever le niveau moyen des élèves.

Dans l'équation [4], l'erreur aléatoire $u_{1j}$ indique que les pentes $\beta_{1j}$ varient d'une classe à l'autre. La pente moyenne dans l'échantillon est représentée par $\gamma_{10}$ ; $u_{1j}$ représente l'écart de chaque classe par rapport à cette pente moyenne. Cela permet de modéliser l'hétérogénéité des relations d'une classe à l'autre. Dans les classes où la pente est forte, les écarts entre élèves se creusent (relativement parlant), tandis qu'ils se réduisent dans celles où la pente est faible. Cette dimension de l'analyse est généralement appelée l'efficacité différentielle, ou bien l'équité. Il s'agit donc là d'une dimension importante de l'analyse puisqu'elle porte sur la capacité des classes à réduire les écarts initiaux entre élèves.

Dans les équations [3] et [4], $z_j$ représente une variable de niveau 2, caractérisant la classe $j$ (*e.g.*, la taille de la classe, la formation de l'enseignant, etc.). Elle est introduite comme une variable susceptible d'expliquer les variations de constantes et de pentes entre les classes. Il est important de noter à ce stade que les constantes et les pentes sont exprimées comme une fonction, d'une part, d'effets fixes représentés par les coefficients gammas et, d'autre part, d'effets aléatoires représentés par et. Ces effets aléatoires ont respectivement pour variance $\mathrm{var}(u_{0j}) = \sigma^2_{u0}$ (*i.e.*, la variance de l'efficacité) et $\mathrm{var}(u_{1j}) = \sigma^2_{u1}$ (*i.e.*, la variance de l'équité). Leur covariance est notée $\mathrm{cov}(u_{0j}, u_{1j}) = \sigma_{u01}$ et elle s'interprète comme la relation entre équité et efficacité. Si l'on intègre [3] et [4] dans [2], on obtient l'équation suivante :

$$y_{ij} = \gamma_{00} + \gamma_{01} z_j + \gamma_{10} x_{ij} + \gamma_{11} z_j x_{ij} + (u_{0j} + u_{1j} x_{ij} + e_{ij}) \qquad [5]$$

Le modèle comporte une partie fixe (les gammas) et une partie aléatoire (présentée entre parenthèses)[6]. C'est cette partie aléatoire qui le distingue fondamentalement du modèle MCO puisqu'elle est constituée de plusieurs termes résiduels, dont les variances et la covariance deviennent des paramètres à estimer.

D'un point de vue empirique, c'est également sur la partie aléatoire que les estimations diffèrent le plus de celles des MCO. Kreft (1996) a fait une revue des simulations (procédures Monte Carlo) faites pour comparer les estimations de modèles multiniveau et de modèles MCO. Les résultats montrent que, dans les deux cas, les coefficients fixes ont des valeurs proches (leurs valeurs ne sont biaisées dans aucun des deux types de modèles) ; il n'en est pas de même toutefois des erreurs-types des coefficients, souvent sous-estimées dans les MCO. Mais c'est sur la partie aléatoire du modèle qu'on trouve les différences fondamentales. Par exemple, Bressoux, Coustère et Leroy-Audouin (1997) ont montré que la variance des pentes peut diminuer considérablement quand on passe d'une estimation MCO à une estimation multiniveau.

# Extensions du champ d'application des modèles multiniveau

## *La modélisation des courbes de croissance*

Les études longitudinales (définies comme des études portant sur des données à mesures répétées pour chaque individu) représentent une des voies de recherche les plus fructueuses en sciences sociales et dans la recherche en éducation en particulier. Du fait que les mesures sont répétées, les données incluent une dimension temporelle qui est nécessaire pour analyser des évolutions. L'éducation se déroulant dans le temps selon un processus cumulatif, l'évolution des acquis est, de fait, un domaine de recherche particulièrement important.

L'analyse des données longitudinales connaît toutefois de sérieuses limitations lorsqu'on utilise des modèles traditionnels par les MCO. Dans de tels modèles, une mesure, considérée comme initiale, est utilisée comme variable explicative d'une mesure finale; par exemple, les scores de début d'année scolaire sont inclus dans le modèle pour expliquer les scores de fin d'année. Ces modèles sont donc généralement contraints à n'utiliser que deux mesures, dont une seule peut être analysée comme variable-réponse (Bryk & Raudenbush, 1992).

Or, il peut être beaucoup plus riche de disposer de plusieurs mesures répétées et de considérer cet ensemble de mesures comme constituant la variable-réponse. Cela permettrait de déterminer des courbes de croissance. On pourrait alors estimer quels sont les facteurs qui influent sur ces courbes de croissance (et non plus sur la seule mesure finale). Une telle analyse ne peut être réalisée par des modèles MCO, mais les modèles multiniveau sont tout à fait bien adaptés à un tel cas. Ils permettent donc de modéliser des courbes de croissance, ce qui est un apport tout à fait fondamental et d'autant plus crucial qu'on s'intéresse à des effets à long terme.

On peut se demander en quoi des structures de ce genre peuvent être analysées à l'aide de modèles multiniveau, puisque nous avons jusqu'à présent considéré des structures dans lesquelles l'individu représentait toujours le premier niveau de la hiérarchie. Cependant, les mesures peuvent être conçues comme étant emboîtées dans les individus de la même façon que les élèves peuvent être conçus comme étant emboîtés dans une classe. C'est ce qui est représenté dans la figure 2 (dans laquelle les mesures sont notées «mes»), où l'on peut voir que la structure est formellement la même que celle d'une structure hiérarchisée «classique» telle que celle qui a été présentée dans la figure 1.

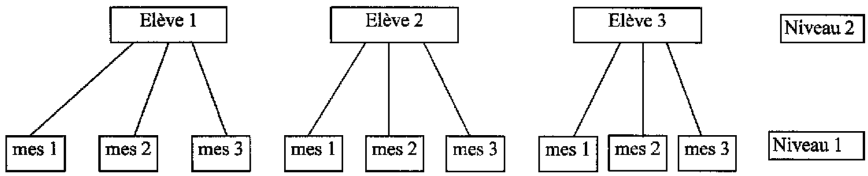


Figure 2. ***Structure à deux niveaux d'un modèle longitudinal de croissance***

C'est donc bien un modèle hiérarchique qu'on construit, modèle dans lequel les mesures composent le niveau 1 et les individus le niveau 2. On décompose ainsi la variance totale des acquis en une variance intraindividuelle et une variance interindividuelle. Il s'agit de rendre compte du fait que la courbe de croissance a une dimension générique (commune à tous les élèves), mais également que chaque élève peut avoir sa propre courbe d'acquisitions, qui diffère plus ou moins de la moyenne. En ce sens, et par analogie à une structure hiérarchisée classique, l'élève constitue «l'environnement» qui influe sur les résultats obtenus à chacune des mesures. On peut évidemment complexifier le modèle en passant à une structure à trois niveaux ou plus, pour peu qu'on veuille étudier l'effet que des niveaux supérieurs seraient également susceptibles d'exercer sur la courbe de croissance (l'école, la famille, etc.).

Les modèles multiniveau offrent au chercheur l'avantage de pouvoir intégrer dans l'analyse un grand nombre de mesures (un nombre indéterminé de mesures peut être pris en compte), ce qui conduit à améliorer la confiance dans les estimations. En outre, les modèles multiniveau sont d'une grande souplesse. Ils permettent que tous les individus ne soient pas mesurés à chaque occasion: s'il y a $t$ mesures, il n'est pas nécessaire que chaque individu soit évalué $t$ fois. Il n'est pas indispensable non plus que les mesures soient effectuées au même moment pour tous les individus.

Afin d'illustrer la modélisation multiniveau des courbes de croissance, nous utiliserons des données qui portent sur 608 élèves, suivis de la grande section de maternelle (GS: dernière année d'école préélémentaire) jusqu'à la fin du cours préparatoire (CP: première année d'école élémentaire) (Suchaut, 1996). Les élèves ont été testés à l'entrée en grande section, à la fin de cette même classe, et puis à la fin de la classe de CP. On dispose ainsi de trois mesures d'acquisitions[7]. En ce cas, le niveau 1 correspond aux différentes mesures et le niveau 2 correspond aux élèves. Les tests n'étant pas constitués des mêmes items, la moyenne des scores a été arbitrairement fixée comme étant égale à l'âge moyen des élèves (en nombre de mois) à chacune des trois occasions[8]. De ce fait, une progression des scores a été fixée, dont l'ampleur est arbitraire, mais qui traduit un impact positif de l'âge sur les acquisitions, ce qui est conforme à ce qu'on observe dans la réalité à ce niveau d'enseignement. La variance a été maintenue constante à chacune des occasions ce qui, là encore, est arbitraire. Il ne s'agit donc pas de s'intéresser à la hausse des scores en valeur absolue, mais plutôt à l'évolution relative de groupes d'élèves (définis en fonction de leur appartenance sociale ou de leur sexe par exemple) au sein des distributions des scores.

Tableau 1
***Modèles de croissance maternelle/CP***

| Paramètres | ML1 | ML2 | ML3 | ML4 |
|---|---|---|---|---|
| *Effets fixes* : | | | | |
| – constante | 73,64 (0,52) | –0,32 (1,68) | –0,32 (1,71) | –2,83 (1,79) |
| – âge | | 1,00 (0,02) | 1,00 (0,02) | 1,00 (0,02) |
| – cadre | | | | 8,47 (1,23) |
| – fille | | | | 1,60 (0,95) |
| *Effets aléatoires* : | | | | |
| – Niveau 2 | | | | |
| Variance inter-individuelle des constantes $\sigma^2_{u0}$ | 106,3 (9,77) | 127 (8,59) | 302,6 (129,9) | 281,5 (128,7) |
| Covariance constantes/pentes $\sigma_{u01}$ | | | –2,78 (1,71) | –2,69 (1,70) |
| Variance des pentes âge/score $\sigma^2_{u1}$ | | | 0,04 (0,02) | 0,04 (0,02) |
| – Niveau 1 | | | | |
| Variance intra-individuelle $\sigma^2_e$ | 176,7 (7,17) | 65,68 (2,66) | 60,83 (3,43) | 60,88 (3,43) |
| Déviance | 15241,5 | 13975,0 | 13969,4 | 13922,0 |

Entre parenthèses figurent les erreurs-types des paramètres.

Le modèle ML1 montre que la part de variance intraindividuelle (niveau 1) représente 62,4% [176,7/(106,3 + 176,7)], soit près des deux tiers de la variance totale. C'est dire qu'il y a plus de variabilité dans les acquisitions d'une mesure à l'autre pour un même individu, qu'il n'y en a entre les individus. Le modèle ML2 montre l'impact significatif de l'âge sur les acquisitions, notamment par une forte réduction de la variance au niveau 1. Rappelons toutefois qu'il s'agit là d'une progression construite, attribuable aux moyennes arbitrairement fixées des scores de chacune des épreuves. De même, le coefficient fixe de l'âge est égal à 1 en raison de l'échelle utilisée (le score moyen étant égal à l'âge moyen à chaque occasion).

Ce sont les modèles ML3 et ML4 qui nous intéressent tout spécialement ici puisqu'ils apportent, quant à eux, des données qui ne dépendent pas de l'échelle de mesure choisie. Le modèle ML3 autorise une pente aléatoire de l'âge. On teste ainsi le fait que le rythme de croissance (précisément, la relation entre l'âge et les scores) puisse être différent d'un individu à l'autre. La décroissance de la déviance par rapport au modèle précédent est faiblement significative [9]. La spécification du modèle ML3 apporte donc peu par rapport à celle du modèle ML2, ce qui signifie qu'il n'y a pas de variation considérable entre les élèves dans leur rythme de croissance d'acquisitions [10]. La variance intraindividuelle (niveau 1) n'est d'ailleurs que légèrement diminuée par le fait qu'on autorise les pentes à être aléatoires, c'est-à-dire à varier d'un individu à l'autre. La covariance entre les constantes et les pentes n'est pas significative, ce qui amène à conclure que le rythme de progression ne dépend pas du niveau initial des élèves.

Le modèle ML4 intègre des variables explicatives autres que l'âge ; il montre que le sexe de l'élève n'influe ni sur le rythme, ni sur le niveau des acquisitions. L'origine sociale en revanche a un effet très significatif, diminuant sensiblement la variance de la constante au niveau 2 : les enfants de cadres ont en moyenne des acquis supérieurs à ceux des autres élèves mais leur rythme de progression n'est pas supérieur (la variance des pentes n'est pas du tout affectée par la prise en compte de l'origine sociale).

Nous nous sommes contentés ici d'un modèle linéaire de croissance. On peut pourtant considérer, d'un point de vue théorique, que les progressions ne sont pas linéaires. Si l'on étudiait les progrès en lecture jusqu'à l'âge adulte, la courbe aurait sans soute une forme de type quadratique ou logarithmique, indiquant que les individus font, à partir d'un certain seuil, de moins en moins de progrès. Modéliser de telles non-linéarités nécessite cependant un nombre suffisant de mesures (au moins quatre). Si une telle structure est spécialement visible sur du long terme, la forme linéaire est souvent considérée comme une modélisation convenable sur une durée réduite telle que celle considérée dans notre exemple.

### *La modélisation des effets aléatoires croisés*

Nous avons construit jusqu'à présent des modèles strictement hiérarchisés ; nous entendons par là le fait que les unités des niveaux inférieurs n'appartenaient qu'à une seule unité de niveau supérieur. Par exemple, les élèves (niveau 1) de la même classe (niveau 2) sont tous scolarisés dans la même école (niveau 3). Il arrive bien souvent cependant qu'on ne soit pas dans la situation d'une hiérarchie parfaite. Les individus peuvent être caractérisés, de façon croisée, par leur appartenance à différentes structures qui ne sont pas strictement emboîtées les unes dans les autres. Les exemples de ce type sont nombreux dans le domaine social en général et éducatif en particulier: les enfants (niveau 1) d'un même quartier (niveau 2) peuvent fréquenter des écoles différentes (niveau 2); les élèves (niveau 1) ayant été scolarisés dans la même école préélémentaire (niveau 2) ne vont pas nécessairement dans la même école élémentaire (niveau 2).

Ainsi, dans nos exemples, la hiérarchie croisée apparaît au niveau 2. Si l'on reprend le dernier cas, les écoles préélémentaires et les écoles élémentaires appartiennent au même niveau car les premières ne sont pas emboîtées dans les secondes. La modélisation multiniveau permet de prendre en compte l'influence des deux écoles, sans les confondre. Une telle structure est représentée dans la figure 3.

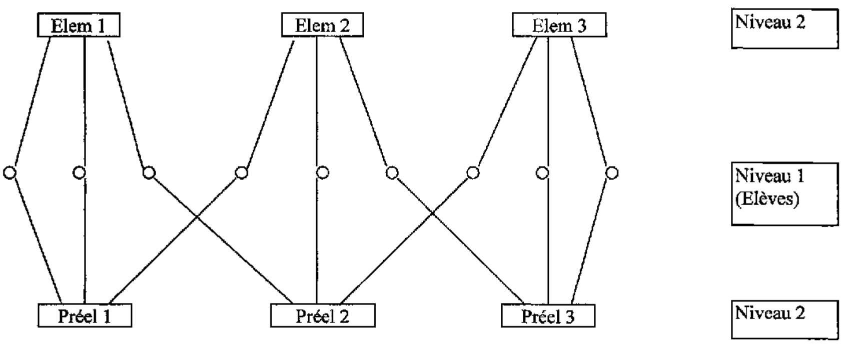


Figure 3. ***Structure à effets aléatoires croisés***

On voit que la structure ne correspond plus strictement à celle d'une hiérarchie «classique». Le modèle multiniveau permet toutefois de tenir compte de cette hiérarchie croisée pour estimer les effets respectifs des différents environnements. Le modèle est le suivant (pour une plus grande simplicité de présentation, nous considérons seulement un modèle avec constantes aléatoires et nous n'y intégrons pas de variable de niveau 2):

$$y_{i12} = \beta_0 + \beta_1 x_{i12} + u_{1} + u_{2} + e_{i12} \qquad [6]$$

où l'indice $i$ fait référence au niveau 1 et les indices $j1$ et $j2$ au niveau 2 (la hiérarchie croisée). Apparaissent donc deux termes aléatoires au niveau 2 : $u_{j1}$ et $u_{j2}$.

Pour illustrer la modélisation de ce type de phénomènes, les données présentées sont celles qui ont été utilisées dans les modèles de croissance. Tous les élèves d'une même classe de CP ne proviennent pas de la même classe maternelle ; nous sommes donc dans le cas d'une hiérarchie croisée. On ne retiendra que les scores obtenus par les élèves à l'entrée en grande section de maternelle et en fin de CP. On cherche à savoir quel est le poids respectif de chacune de ces deux classes dans les acquisitions de fin de CP.

Ce type de problème pose la question de la rémanence des effets-classes : les acquis mesurés à un temps $t$ ne sont pas uniquement fonction des conditions de scolarité au temps $t$, mais ils sont également le fruit de conditions antérieures dont les effets peuvent être plus ou moins durables et plus ou moins marqués. Existe-t-il ainsi des moments de la scolarité plus importants que d'autres ? Le fait d'avoir bénéficié d'un enseignement performant une année donnée constitue-t-il un avantage durable ou bien éphémère pour la suite de la scolarité ? Malgré l'intérêt évident de telles questions, elles n'ont été que très rarement traitées dans les écrits, sans doute à cause de la difficulté à les appréhender avec des techniques de modélisation classique. Notre exemple correspond à une interrogation de ce type : la classe de grande section maternelle a-t-elle un poids encore observable en fin de CP sur les acquis des élèves ? Si oui, ce poids est-il important ?

Tableau 2
***Modélisation d'effets aléatoires croisés***

| Paramètres | ML5 | ML6 | ML7 |
|---|---|---|---|
| *Effets fixes :* | | | |
| -- constante | 100,2 (1,1) | 32,15 (3,03) | 32,32 (3,03) |
| -- test initial | | 0,68 (0,03) | 0,68 (0,03) |
| *Effets aléatoires :* | | | |
| -- Niveau 2 | | | |
| Variance inter-CP $\sigma^2$ | 27,48 (14,06) | 25,58 (9,62) | 33,62 (8,08) |
| Variance inter-GS $\sigma^2_{u0}$ | 25,21 (14,84) | 11,68 (8,42) | |
| -- Niveau 1 | | | |
| Variance inter-individuelle $\sigma^2_e$ | 165,1 (10,31) | 84,34 (5,29) | 86,86 (5,23) |

Entre parenthèses figurent les erreurs-types des paramètres.

Le modèle ML5 (qui n'inclut pas de variable explicative) est un modèle dit « vide », où il s'agit de décomposer la variance des acquis de fin de CP en trois composantes : au niveau 1, une variance interindividuelle ; au niveau 2, une variance attribuable au CP et une variance attribuable à la GS. Ce modèle montre que la variance brute du score d'acquisitions en fin de CP réside principalement au niveau 1, c'est-à-dire entre les individus mais que, au niveau 2, la part de variance attribuable au CP est quasiment égale à celle qui est attribuable à la grande section de maternelle. Il faut toutefois rester assez prudent dans l'interprétation de l'ampleur de ces variances car il y a beaucoup plus de classes maternelles que de classes de CP dans l'échantillon. Lorsqu'à un même niveau, une des classifications (en l'occurrence les classes maternelles) contient plus d'unités que l'autre (les classes de CP) et donc moins d'individus du niveau inférieur, cela peut conduire à accroître la part de variance attribuée à la première (Goldstein 1995).

Il est en revanche très instructif de voir comment se « comportent » les variances lors de la prise en compte d'autres variables dans la modélisation, en particulier avec l'introduction des scores au test initial, en début de GS. C'est ce que réalise le modèle ML6. On constate tout d'abord un coefficient de la variable qui est très significatif (partie fixe du modèle), qui indique que le score au test de fin de CP dépend bien du score au test initial, et cela de façon importante puisque la réduction des parts de variance est très marquée, en particulier au niveau 1. Mais nous sommes surtout intéressés par l'ampleur des variances au niveau 2, qui révèlent les effets de la hiérarchie croisée : l'introduction du test initial réduit la part de variance attribuable à la classe de grande section de plus de la moitié, tandis qu'elle n'affecte que très peu la part de variance attribuable à la classe de CP. D'ailleurs, l'examen des erreurs-types montre que la variance attribuable à la grande section n'est plus significative : on peut donc conclure, sur la base de cette modélisation croisée, que ce n'est pas au cours de la grande section que se sont générées les différences d'acquisitions observées en fin de CP, mais bien davantage au cours du CP. Une grande partie de la variance inter-GS provenait en fait de différences générées avant même l'entrée dans cette classe de GS.

Par contraste, le modèle ML7 est un modèle hiérarchique classique et considère donc les données comme strictement emboîtées les unes dans les autres. Par conséquent, il ne rend pas compte de la structure croisée que représente le passage de la classe de grande section de maternelle à celle de CP. Il est donc impossible, à partir d'un tel modèle, de dissocier les effets de ces deux classes sur les acquisitions de fin de CP. Il traite en effet d'une durée d'acquisitions de deux ans (celle qui sépare l'évaluation initiale de l'évaluation finale), mais il ne prend explicitement en compte que la deuxième année scolaire (celle du CP). La comparaison des modèles ML6 et ML7 montre que la part de variance entre les classes de CP est légèrement surestimée dans ce dernier modèle. Cela s'explique par le fait que, dans le modèle ML7, on « oublie » qu'une (petite) part de la variance des acquis de fin de CP est attribuable à la GS.

Dans ce cas précis, la non-prise en compte de la structure croisée ne conduit pas à des résultats fondamentalement différents; cela dit, si la variance attribuable à la classification omise était significative, les erreurs d'interprétation seraient substantielles, et le seraient d'autant plus que cette part de variance serait relativement élevée (pour un exemple, voir Goldstein, 1995).

## Conclusion

Les modèles multiniveau représentent une avancée importante dans la recherche des effets qu'exerce l'environnement sur les individus. C'est ce type de problème qui est bien souvent posé dans le cadre de l'évaluation en éducation. La recherche sur les effets-maîtres et les effets-écoles est en ce sens exemplaire, où l'on tente d'estimer l'impact que produit la fréquentation de diverses écoles ou classes sur les comportements, les attitudes, les acquis, etc., des élèves. Dans ce cadre, la structure hiérarchisée des données est particulièrement apparente (élèves au niveau 1, classes au niveau 2, écoles au niveau 3, etc.). C'est spécifiquement pour l'analyse de ce type de structure que les modèles multiniveau ont été conçus, mais ils ont également été adaptés récemment à d'autres champs d'applications, dont la présentation a constitué l'objet de cet article.

Les modèles multiniveau sont maintenant adaptés à l'analyse de structures non strictement hiérarchisées, dites aléatoires croisées, ce qui permet de rendre compte du fait que les individus appartiennent simultanément à différents environnements, chacun étant susceptible d'exercer des effets propres. Dans le champ de l'évaluation en éducation, la modélisation des effets aléatoires croisés permet ainsi d'analyser les effets simultanés d'environnements scolaires et extra-scolaires (par exemple, tous les élèves d'une même école n'habitent pas nécessairement le même quartier), les effets simultanés de deux environnements scolaires comme dans le cas de certains décloisonnements (tous les élèves d'une même classe ne font pas forcément partie du même groupe de décloisonnement). Les modèles multiniveau permettent aussi de rendre compte du fait que les individus partagent un même environnement à une période donnée mais qu'ils ont pu fréquenter auparavant des environnements différents. Ainsi, par exemple, les élèves d'un même lycée n'ont pas tous fréquenté le même collège, or ce dernier n'est peut-être pas sans conséquences sur leurs comportements, leurs attitudes ou leurs acquis au cours même du lycée. La modélisation des effets aléatoires croisés s'adapte donc aussi bien à l'analyse de phénomènes transversaux qu'à celle de phénomènes longitudinaux.

Concernant la modélisation des courbes de croissance, la multiplication du nombre de mesures permet d'améliorer la confiance dans les estimations. Ce point est important pour le chercheur en éducation puisque les mesures opérées dans ce domaine, qu'il s'agisse de comportements, d'attitudes ou d'acquisitions, sont généralement entachées d'une part d'erreur substantielle. La compensation qui s'opère entre les diverses erreurs aléatoires permet alors d'obtenir des estimations plus proches de la valeur «vraie». Modéliser des courbes de croissance, c'est aussi la possibilité de s'intéresser à des changements sur du long terme et de prendre en compte, non seulement un niveau d'acquis, mais surtout la forme de l'évolution de ces acquis dans le temps. Une telle modélisation est particulièrement adaptée à des plans d'expérimentation conçus sous la forme prétest, post-test, retests, ces derniers pouvant être nombreux et étalés dans le temps. On peut ainsi, par exemple, évaluer l'influence de diverses séances didactiques, non seulement sur les acquis mesurés immédiatement à l'issue du «traitement» mais aussi, grâce à des mesures ultérieures, déterminer quelles sont les séances qui conduisent à l'évolution la plus favorable de ces acquis.

Les modèles MCO classiques reposent sur des hypothèses très restrictives: indépendance des résidus, relation identique entre $x$ et $y$ d'un contexte à l'autre. Or, ces hypothèses peuvent être empiriquement testées; dans le cas où elles sont démenties par les données, cela affecte la qualité des estimations, ce qui est typiquement le cas dans l'évaluation des effets-maîtres et des effets-écoles. Les modèles multiniveau reposent sur des hypothèses moins restrictives que les modèles MCO et, dans le même temps, ils permettent de les tester empiriquement. De fait, les résultats montrent que ces hypothèses sont souvent vérifiées et que, par là même, les modèles multiniveau sont plus réalistes pour les problèmes qui se posent dans l'évaluation en éducation. En ce sens, on peut dire que les modèles multiniveau sont plus généraux que les modèles MCO.

Ils sont également plus souples car ils s'adaptent à une plus grande variété de structures susceptibles d'être étudiées. Dans une structure hiérarchisée classique, les modèles multiniveau venaient utilement combler les lacunes d'autres modélisations MCO possibles, mais dans la modélisation des courbes de croissance ou dans la modélisation des effets aléatoires croisés, les modèles MCO sont impropres à rendre compte des effets recherchés. Cela rend les modèles multiniveau d'autant plus indispensables et intéressants à étudier.

Il ne s'agit pas pour autant de dire que ces modèles sont une panacée et que leur mise en œuvre suffira à améliorer les connaissances scientifiques. La construction théorique, le choix et la mesure des variables demeurent des éléments primordiaux, en amont de tout problème d'analyse statistique proprement dit. Il n'en demeure pas moins que, par leur souplesse et leur grand champ d'application, les modèles multiniveau offrent la possibilité d'analyser les données empiriques d'une façon qui s'accorde avec les éléments théoriques qui ont présidé à leur collecte et à leur organisation, ce qui n'était pas toujours le cas avec les modèles traditionnels.

## NOTES

1. Nous remercions Bruno Suchaut pour avoir mis à notre disposition les données qu'il avait recueillies.
2. Une telle hiérarchie n'est cependant pas donnée immédiatement au chercheur et celui-ci devra définir l'environnement en fonction de sa question de recherche. L'environnement peut en effet avoir des limites plus ou moins floues (faut-il s'intéresser uniquement à la famille proche par exemple ?), il peut également être plus ou moins perméable (individus qui viennent et s'en vont) et plus ou moins permanent. La recherche sur l'effet-classe ou l'effet-école offre l'avantage d'un découpage clairement délimité puisqu'il correspond à un découpage institutionnel, peu perméable et dont la durée est connue.
3. Ces modèles sont également connus sous d'autres termes dont les plus fréquents sont : modèles hiérarchiques linéaires (Bryk & Raudenbush, 1992), modèles à coefficients aléatoires (Longford, 1993), modèles à composantes de la variance (Searle, Casella & McCulloch, 1992).
4. Il est possible, en MCO, de réaliser une régression par classe et d'observer si la relation entre $x$ et $y$ (*i.e.*, la pente de régression) diffère d'une classe à l'autre, mais on se heurte là à deux écueils majeurs : d'une part, il n'y a pas de test statistique pour vérifier la significativité de la variance et, d'autre part, on surestime très fortement les différences de pente entre les classes (Bressoux, Coustère & Leroy-Audouin, 1997).
5. Pour des raisons de simplicité, on se limite ici à la présentation d'un modèle à deux niveaux mais celui-ci pourrait être aisément étendu à plus de niveaux. Un logiciel tel que MLn (celui qui est utilisé dans le cadre de cet article) ne comporte plus de limite concernant le nombre de niveaux de la structure étudiée.

6. Nous avons présenté ici directement le modèle multiniveau complet, avec constantes et pentes aléatoires, qui intègre également une variable de niveau 2. En pratique, on procède plutôt par étapes : dans un premier temps, on spécifie un modèle avec constantes aléatoires seulement, puis un autre où l'on autorise en plus des pentes aléatoires. Si, par exemple, la variance des pentes se révèle non significative, alors on se contentera de spécifier un modèle avec constantes aléatoires et pentes fixes. Enfin, on intégrera des variables de niveau 2 pour modéliser les variances et covariances de niveau 2.
7. Modéliser des courbes de croissance d'acquisitions ne suppose pas que les mêmes instruments de mesure soient utilisés à chaque occasion, mais que l'habileté mesurée soit la même. Ce n'est pas tout à fait le cas ici ; le test de fin de CP évalue les acquisitions en lecture et en mathématiques, tandis que les tests de grande section mesurent les compétences générales jugées nécessaires à ces apprentissages fondamentaux. Les résultats doivent donc être considérés avec une certaine prudence.
8. Une telle construction est conseillée par Plewis (1993) dans le cas où les échelles des tests ne sont pas équivalentes.
9. La pertinence d'un modèle s'évalue par la statistique de la déviance : –2 log L (L étant la vraisemblance du modèle). On étudie la décroissance de la déviance qui s'opère quand on passe d'un modèle donné à un autre plus compliqué (*i.e.*, dont le nombre de paramètres à estimer est plus élevé). La décroissance de la déviance suit une loi du ($\chi^2$ à $k$ degrés de liberté ($k$ étant le nombre de paramètres supplémentaires à estimer). Dans le cas présent, quand on passe du modèle ML2 au modèle ML3, la décroissance est égale à 5,6 ; elle suit une loi du ($\chi^2$ à 2 degrés de liberté (une variance et une covariance supplémentaires sont estimées).
10. En autorisant des pentes aléatoires, la variance de la constante varie en fonction des valeurs de $x$ (*i.e.*, en fonction de l'âge). La constante étant la valeur de $y$ quand celle de $x$ est nulle, elle correspond donc, dans le cas présent, aux acquisitions pour un âge zéro, ce qui n'a évidemment pas de sens. La valeur élevée de la variance des constantes dans les modèles ML3 et ML4 n'a donc pas d'interprétation immédiate. Il est parfois conseillé de centrer $x$, ce qui est un moyen de donner une interprétation à la valeur $x = 0$ (en l'occurrence, l'âge moyen). Le centrage des variables a toutefois donné lieu à des controverses (Kreft, De Leeuw & Aiken, 1995).

## RÉFÉRENCES


Aitkin, M., & Longford, N. (1986). Statistical modelling in school effectiveness studies (with discussion). *Journal of the royal statistical society, Series A, 149*, 1-43.

Blau, P. M. (1960). Structural effects. *American sociological review, 25*, 178-193.

Boudon, R. (1963). Propriétés individuelles et propriétés collectives, un problème d'analyse écologique. *Revue française de sociologie, IV* (3), 275-299.

Bressoux, P. (1994). Estimer et expliquer les effets des classes : le cas des acquisitions en lecture. *Mesure et évaluation en éducation, 17* (1), 75-94.

Bressoux, P. (1995). Les effets du contexte scolaire sur les acquisitions des élèves : effets-classes et effet-école en lecture. *Revue française de sociologie, XXXVI* (2), 273-294.

Bressoux, P. (1996). The effects of teachers' training on pupils' achievement : the case of elementary schools in France. *School effectiveness and school improvement, 7* (3), 252-279.

Bressoux, P., Coustère, P., & Leroy-Audouin, C. (1997). Les modèles multiniveau dans l'analyse écologique : le cas de la recherche en éducation. *Revue française de sociologie, XXXVIII* (1), 67-96.

Bru, M. (1991). *Les variations didactiques dans l'organisation des conditions d'apprentissage*. Toulouse : Éditions Universitaires du Sud.

Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear models. Applications and data analysis methods*. Newbury Park, CA : Sage Publications.

Burstein, L. (1980). The analysis of multilevel data in educational research and evaluation. *Review of research in education, 8*, 158-233.

Cronbach, L. J., & Webb, N. (1975). Between-class and within-class effects in a reported aptitude ( treatment interaction : reanalysis of a study by G. L. Anderson. *Journal of educational psychology, 67*, 717-724.

Duru-Bellat, M., & Leroy-Audouin, C. (1990). Les pratiques pédagogiques au CP : structure et incidence sur les acquis des élèves. *Revue française de pédagogie, 93*, 5-15.

Erbring, L., & Young, A. A. (1979). Individuals and social structure : contextual effects as endogenous feedback. *Sociological methods and research, 7*, 396-430.

Firebaugh, G. (1978). A rule for inferring individual-level relationships from aggregate data. *American sociological review, 43*, 557-572.

Goldstein, H. (1986). Multilevel mixed linear model analysis using iterative generalized least squares. *Biometrika, 73*, 43-56.

Goldstein, H. (1995). *Multilevel statistical models*. London : Kendall's Library of Statistics.

Grisay, A. (1993). *Le fonctionnement des collèges et ses effets sur les élèves de sixième et de cinquième* (Les Dossiers d'éducation et formations, n° 32). Paris : Ministère de l'Education nationale.

Hammond, J. L. (1973). Two sources of error in ecological correlations. *American sociological review, 38*, 764-777.

Hauser, R. M. (1970). Context and consex : a cautionary tale. *The American journal of sociology, 75*, 645-664.

Hauser, R. M. (1974). Contextual analysis revisited. *Sociological methods and research, 2*, 365-375.

Kreft, I. G. G. (1996). *Are multilevel techniques necessary ?* Los Angeles : California State University.

Kreft, I. G. G., De Leeuw, J., & Aiken, L. S. (1995). The effect of different forms of centering in hierarchical linear models. *Multivariate behavioral research, 30* (1), 1-21.

Longford, N. T. (1993). *Random coefficient models*. Oxford : Oxford University Press.

Mason, W. M., Wong, G. M., & Entwisle, B. (1983). Contextual analysis through the multilevel linear model. In S. Leinhardt (éd.), *Sociological methodology* (pp. 72-103). San Francisco : Jossey-Bass.

Mingat, A. (1984). Les acquisitions scolaires de l'élève au CP : les origines des différences ? *Revue française de pédagogie, 69*, 49-62.

Mingat, A. (1991). Expliquer la variété des acquisitions au cours préparatoire : les rôles de l'enfant, la famille, l'école. *Revue française de pédagogie, 95*, 47-63.

Plewis, I. (1993). Reading progress. In G. Woodhouse (,d.), *A guide to ML3 for new users* (pp. 98-124). London : Institute of Education.

Raudenbush, S. W., & Bryk, A. S. (1986). A hierarchical model for studying school effects. *Sociology of education, 59*, 1-17.

Robinson, W. S. (1950). Ecological correlations and the behaviour of individuals. *American sociology review, 15*, 351-357.

Searle, S. R., Casella, G., & McCulloch, C. E. (1992). *Variance components*. New York : Wiley.

Suchaut, B. (1996). *Le temps scolaire : allocation et effets sur les acquisitions des élèves en grande section de maternelle et au cours préparatoire*. Thèse inédite, nouveau régime Sciences de l'éducation, Université de Bourgogne, Dijon.