

LE TRAVAIL HUMAIN

REVUE TRIMESTRIELLE

Fondateurs

J.-M. LAHY et H. LAUGIER

Directeur honoraire

R. BONNARDEL

COMITÉ DE DIRECTION

J. LEPLAT, directeur à l'EPHE
M. REUCHLIN, professeur à Paris V
J. SCHERRER, professeur à la Faculté de Médecine de Paris (Paris VI)
S. BOUISSET, professeur à l'Université de Paris-Sud (Orsay)

COMITÉ DE RÉDACTION :

Correspondants français

P. ANDLAUER (Méd. du trav., Annecy), A. BISSERET (INRIA, Rocquencourt), F. BRESSON (EHSS-CNRS, Paris), P. FRASSE (Paris V-CNRS-EPHE, Paris), A. GILET (Méd. du trav., Olivet), J. GRISEZ (AFPA, Montreuil), B. METZ (CEB-CNRS, Strasbourg), H. MONOD (CNRS, Paris), M. de MONTMOLLIN (Paris XIII-Villetaneuse), S. MOSCOVICI (EHSS, Paris), J. PAILLARD (INP-CNRS, Marseille), S. PACAUD (CNRS-EPHE, Paris), P. VARENE (Bordeaux II), A. WISNER (CNAM, Paris)

Correspondants étrangers

Belgique J.-M. FAVERGE (ULB, Bruxelles)
V. de KEYSER (Université de Liège)
Canada M. LORTIE (Ecole polytechnique, Montréal)
Hollande J. W. H. KALSBECK (Université de Twente)
Gde-Bretagne L. BAINBRIDGE (University College, London)
K. D. DUNCAN (University of Wailes, Cardiff)
R.F.A. B. WILPERT (Technische Universität, Berlin)
Suisse F. GENDRE (Inst. Psychol. appl., Lausanne)
P. REY (Université de Genève)

Secrétaire de Rédaction

Mlle M. CHESNAIS

EXPERTS CONSULTÉS

F. BACHER, Y. BERNARD, A. BERTHOZ, Cl. BONNET, M.-C. BOTTE, E. CAUZINILLE, X. CUNY, J. CURIE, F. DANIELLOU, A. FERRANDEZ, J. FORET, R. FRANCÉS, Ch. GADBOIS, Ch. GEORGE, G. GHIGLIONE, M.-T. HO, J.-M. HOC, M. HUGON, M. HUTEAU, G. KARNAS, J. LAUTREY, A. LAVILLE, J.-Ch. LEBAHAR, B. LECOUTRE, D. LEFINE, F. LILLE, J. MARCELLIN, G. de MONTMOLLIN, E. MULLET, M. NEBOIT, A. NGUYEN-XUAN, J. PAILHOUX, B. PAVARD, Y. QUEINNEC, R. REBIFFE, J.-F. RICHARD, G. ROUSTANG, A. SAVOYANT, D. SCAPIN, S. SEBILLOTTE, G. de TERSAC, G. TIBERGHIEN, Y. TOUZARD, G. VERGNAUD, P. VERMERSCH, J.-C. VERRIEST, J.-J. VOGT, A. WEILL-FASSINA

RÉDACTION

Laboratoire de Psychologie du Travail de l'EPHE
41, rue Gay-Lussac, 75005 Paris. Tél. 46-33-69-21
Envoyer les articles, notes, informations à cette adresse.
Les manuscrits devront être présentés sous forme dactylographiée (interligne 2) et fournis en triple exemplaire. Comme illustrations, il ne sera publié, en principe, que des graphiques et des dessins. (Cf. Instructions aux auteurs, p. 3 de couverture.)

ADMINISTRATION

Presses Universitaires de France (PUF), 108, bd Saint-Germain, 75006 Paris

ABONNEMENTS

Presses Universitaires de France (PUF), Département des Revues
14, avenue du Bois-de-l'Épine, BP 90, 91003 Evry Cedex. Tél. (1) 60 77 82 05
Télécopie (1) 60 79 20 45. Téléx : PUF 600 474 F. CCP 1302 69 C Paris
Abonnements 1989 : France..... 300 F — Etranger..... 360 F

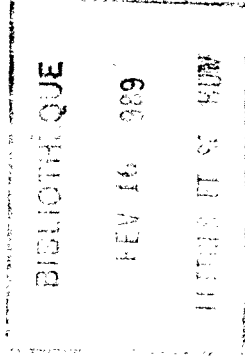
Avis important : Les demandes en duplicata des numéros non arrivés à destination ne pourront être admises que dans un délai maximum de 15 jours après réception du numéro suivant.

DOCUMENT PÉDAGOGIQUE

LES MODÈLES STRUCTURAUX EN PSYCHOLOGIE PRÉSENTATION D'UN MODÈLE : LISREL

Deuxième partie

par F. BACHER*



SUMMARY

Structural models in Psychology. Presentation of a model : LISREL. — After a first part bearing on confirmatory factor analysis, the second part of this paper presents the use of LISREL for the study of hypothesized, eventually causal, relationship between the variables of a system. Possibilities and limits of the model are briefly reviewed and the example of a study analysing cross lag relationships between variables measured at two occasions on the same subjects is used to illustrate the different steps.

(Key words : Cross lag relationships, LISREL, Structural modeling.)

Dans la première partie de cet article (Bacher, 1987) nous avons souligné le développement qu'ont pris les modèles structuraux en psychologie et présenté le plus utilisé d'entre eux, LISREL, en nous centrant sur ses applications dans le domaine de l'analyse factorielle confirmatoire. Cette seconde partie sera consacrée aux applications du modèle à l'étude des relations entre des ensembles de variables dont certaines sont considérées comme indépendantes et d'autres comme dépendantes.

I — PRINCIPE DE MISE EN ŒUVRE

Le modèle LISREL intègre, rappelons-le, un modèle d'analyse factorielle, qui généralise la théorie de la mesure et un modèle d'équations simultanées multiple et l'analyse des méthodes telles que l'analyse de régression d'intervalles ou considérées comme telles.

Ce modèle, linéaire, est utilisable sur des variables ayant le niveau d'échelles

* Laboratoire de Psychologie différentielle, 41, rue Gay-Lussac, 75005 Paris. Ce travail a utilisé des moyens que nous devons à l'École Pratique des Hautes Études, Section Sciences de la Vie et de la Terre, à l'Université René-Descartes, au Conservatoire national des Arts et Métiers (Service de Recherches de l'INETOP), au Centre national de la Recherche scientifique (UA 656).

Il comporte deux parties : un modèle d'équations structurales, qui spécifie les relations entre des variables latentes, et un modèle de mesure, qui indique comment les variables latentes sont mesurées en termes de variables observées.

Nous considérerons, pour illustrer les principes de mise en œuvre de la partie structurale du modèle, le sous-modèle $y = Bx + \Gamma x + \zeta$ dans lequel il n'y a pas de variables ζ et η , de sorte qu'il n'y a pas lieu de faire intervenir de modèle de mesure.

Jöreskog et Sörbom (1985, p. III-33 à III-40) réanalysent, par exemple, des corrélations rapportées par Kerchoff (1974, p. 46) obtenues sur 767 garçons en 12^e année de scolarité et déjà réanalysées par Kenny (1979, p. 47-73). Ces corrélations portent sur les variables suivantes :

- x_1 intelligence ;
- x_2 nombre de frères et sœurs ;
- x_3 éducation du père ;
- x_4 profession du père ;
- y_1 notes scolaires ;
- y_2 études envisagées ;
- y_3 niveau d'aspiration professionnelle.

L'objectif de l'étude est d'analyser les relations qui s'établissent entre réussite scolaire (y_1), études envisagées (y_2) et niveau d'aspiration professionnelle (y_3), en faisant l'hypothèse que la première de ces variables est susceptible d'influencer les deux suivantes, que la deuxième est susceptible d'influencer la troisième, mais qu'aucune de ces variables n'est susceptible d'influencer celles qui la précèdent (modèle récuratif). On se propose en outre d'examiner l'effet de plusieurs variables indépendantes en corrélation, x_1 , x_2 , x_3 , x_4 , dont on fait l'hypothèse qu'elles sont susceptibles d'avoir une influence, soit directement sur chacune des variables dépendantes, soit indirectement (en influençant, par exemple, les notes scolaires qui, à leur tour, influenceront les études envisagées et le niveau d'aspiration professionnelle).

A / Spécification du modèle en fonction des hypothèses

On peut, tout d'abord, traduire les hypothèses à l'aide d'un *graphe* (fig. 1). Le modèle fait appel ici à 4 variables x , 3 variables y et 3 variables ζ . Les variables x sont en corrélation, mais on ne cherchera pas à expliquer ces corrélations qui sont considérées comme prédéterminées par rapport au modèle. Des doubles flèches courbes représentent les corrélations de ce type. L'hypothèse que chaque variable x influence chacune des variables y est traduite par des flèches allant des premières aux deuxièmes. Des flèches allant de y_1 vers y_2 et y_2 vers y_3 traduisent l'hypothèse que y_1 est susceptible d'influencer y_2 et que y_2 est susceptible d'influencer y_3 . Par contre, on ne fait pas figurer de flèche allant de y_3 vers y_1 et y_2 , ou de y_2 vers y_1 puisqu'on fait l'hypothèse que les effets correspondants n'existent pas.

On suppose en outre que chacune des variables y est influencée par une variable résiduelle (ζ_1 pour y_1 , ζ_2 pour y_2 , ζ_3 pour y_3) qui représente la part de chaque variable y non expliquée par les variables antérieures du système. D'après les hypothèses générales du modèle, les ζ sont indépendants des x .

On ne fait pas figurer de flèche reliant les x et les ζ . On fait en outre ici l'hypothèse que les ζ ne sont pas liés entre eux. Cette hypothèse correspond à une hypothèse de clôture : elle consiste à considérer qu'il n'y a pas de cause commune de variation des y en dehors des variables (x et y) incluses dans le système ; on n'a donc pas fait figurer de doubles flèches courbes reliant les ζ .

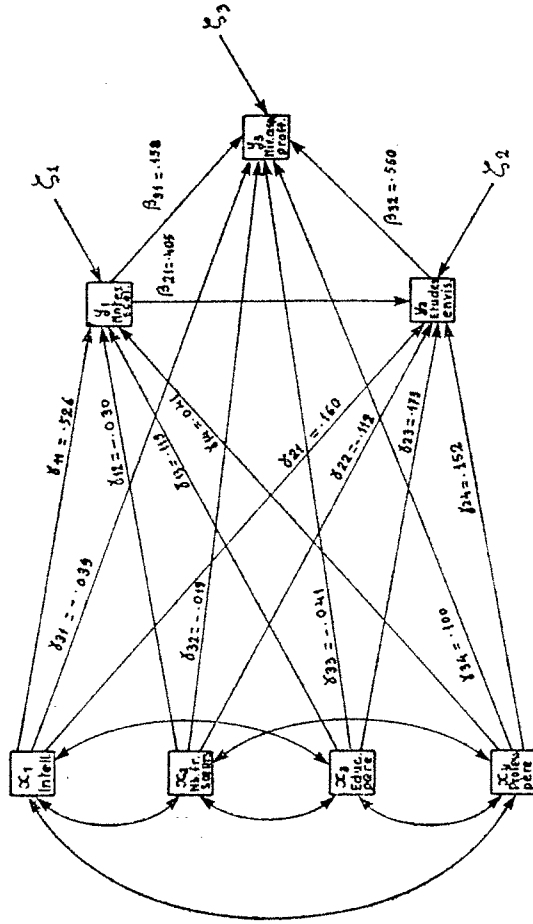


Fig. 1. — Graphe d'un modèle récuratif : réussite scolaire, études envisagées et niveau d'aspiration professionnelle. Les valeurs obtenues pour les coefficients β et γ à l'issue de l'analyse ont été portées sur le graphe.

(Path diagram of a recursive model : school achievement, educational expectation and occupational aspiration. Values obtained for the β and γ coefficients as a result of the analysis are shown on the graph.)

Les hypothèses peuvent aussi être traduites par le système d'équations reliant les variables dépendantes aux autres variables du système :

$$y_1 = \gamma_{11} x_1 + \gamma_{12} x_2 + \gamma_{13} x_3 + \gamma_{14} x_4 + \zeta_1$$

$$y_2 = \beta_{21} y_1 + \gamma_{22} x_2 + \gamma_{23} x_3 + \gamma_{24} x_4 + \zeta_2$$

$$y_3 = \beta_{31} y_1 + \beta_{32} y_2 + \gamma_{33} x_3 + \gamma_{34} x_4 + \zeta_3$$

Ces équations peuvent être obtenues à partir du graphe de la façon suivante. Une équation est établie pour chaque variable dépendante. Cette variable figure à gauche de l'équation ; à droite, on écrit autant de termes que de flèches orientées aboutissant à la variable considérée. La convention habituelle, en ce qui concerne les indices, est de mettre en premier celui de la variable à laquelle aboutit la flèche et en deuxième celui de la variable d'où elle vient.

On remarque que certains termes ne figurent pas dans les deux premières équations (termes en y_2 et y_3 dans la 1^{re} équation, terme en y_3 dans la 2^e). En effet, on a fait l'hypothèse que y_1 n'était pas influencé par y_2 et y_3 et que y_2 n'était pas influencé par y_3 . Les coefficients β correspondants sont donc fixés égaux à 0.

L'équation matricielle $y = \Gamma x + B y + \zeta$ résume le système. Les hypothèses relatives à la récursivité se traduisent, comme on vient de le voir, par des éléments de la matrice **B** fixés égaux à 0. Les hypothèses de clôture (absence de corrélation entre les ζ) se traduisent par la condition imposée que la matrice des variances et covariances des ζ , Ψ , soit diagonale. Comme il n'y a pas de variables latentes, la matrice Φ est la matrice des corrélations observées entre variables x . On trouvera au tableau I, les matrices **B**, Γ , Ψ , avec l'indication des paramètres fixés et de ceux qui doivent être estimés.

TABLEAU I

Matrices correspondant au graphe de la figure 1 (modèle récursif de la réussite scolaire, des études envisagées et du niveau d'aspiration professionnelle). Les éléments marqués (*) sont fixés. Les éléments diagonaux de **B** sont nuls par définition

(Matrices corresponding to figure 1 path diagram (recursive model of school achievement, educational expectation and occupational aspiration). Elements marked (*) are fixed. Diagonal elements of **B** are 0 by definition.)

	B			Ψ		
	y_1	y_2	y_3	ζ_1	ζ_2	ζ_3
y_1	0	0*	0*	Ψ_{11}	0*	0*
y_2	β_{21}	0	0*	0*	Ψ_{22}	0*
y_3	β_{31}	β_{32}	0	0*	0*	Ψ_{33}
	Γ					
	x_1	x_2	x_3	x_4		
y_1	γ_{11}	γ_{12}	γ_{13}	γ_{14}		
y_2	γ_{21}	γ_{22}	γ_{23}	γ_{24}		
y_3	γ_{31}	γ_{32}	γ_{33}	γ_{34}		

Dans l'exemple que nous présentons, le modèle est un modèle récursif complet (sans autres contraintes introduites sur les coefficients β et γ que celles qui assurent la récursivité). Les modèles de ce type sont juste identifiés. L'estimation a été faite par maximum de vraisemblance!

B / Examen des résultats

Le tableau II donne la matrice de corrélations analysée, **R**, et les résultats de l'analyse.

L'analyse ayant porté sur des corrélations, la solution est standardisée : les variables y ont une variance égale à 1 ; les coefficients β et γ sont réduits.

L'examen de la matrice Ψ montre que, pour chacune des variables dépendantes, une part relativement importante de la variance n'est pas expliquée par les variables précédentes du modèle (.651 pour les notes scolaires, par exemple).

L'examen des coefficients de la matrice **B** (qui sont peu différents de coefficients de régression partiels réduits)³ montre que les notes scolaires ont une

1. Elle ne pourrait être faite par moindres carrés ordinaires car le modèle LISREL impose l'indépendance de chaque ζ avec les x mais pas avec les y qui le précèdent dans le modèle comme l'exigerait l'utilisation de cette méthode (voir Fox, 1984, p. 251-260).

2. Il faut prendre garde au fait que les coefficients relatifs à y_1 , y_2 et y_3 ne sont pas de même ordre : pour chacun, ce sont les variables précédentes dans le modèle qui sont tenues constantes. Par exemple le coefficient $\beta_{31} = .405$ est estimé lorsque x_1, x_2, x_3, x_4 sont tenus constants. Le coefficient $\beta_{21} = .550$ est estimé lorsque x_1, x_2, x_3, x_4 et y_1 sont tenus constants. Ces coefficients mesurent l'effet direct (c'est-à-dire non attribuable aux variables précédentes du modèle) d'une variable sur l'autre.

TABLEAU II

Analyse des relations entre notes scolaires, études envisagées et niveau d'aspiration professionnelle. Ψ est diagonale. Les valeurs marquées (*) ont été fixées

(Analysis of relationships between school achievement, educational expectation and occupational aspiration. Ψ is diagonal. Values marked (*) were fixed.)

	R					
	x_1	x_2	x_3	x_4	y_1	y_2
x_1	1.000					
x_2	-.100	1.000				
x_3	.277	-.152	1.000			
x_4	.250	-.108	.611	1.000		
y_1	.572	-.105	.294	.248	1.000	
y_2	.489	-.213	.446	.410	.597	1.000
y_3	.335	-.153	.303	.331	.478	.651
	B					
	y_1	y_2	y_3	y_4	y_5	y_6
	Notes scolaires			Etudes envisagées		Niveau aspiration professionnelle
y_1	0	0*	0*	0*	0*	0*
y_2	.405	0	0	0	0*	0*
y_3	.158	.550	0	0	0	0
	\Gamma					
	x_1	x_2	x_3	x_4	x_5	x_6
	Intelligence		Nombre frères sœurs		Education Père	
y_1	-.526	-.030	-.119	.041		
y_2	.160	-.112	.173	.152		
y_3	-.039	-.019	-.041	.100		
	\Psi					
ζ_{11}	.651					
ζ_{22}	.517					
ζ_{33}	.557					

$\chi^2_0 = 0$ (P = 1.00)
GFI = 1
RMR = 0

forte influence directe sur la durée des études envisagées, la durée des études envisagées, une forte influence directe sur le niveau d'aspiration professionnelle. Les notes scolaires ont une assez faible influence directe sur le niveau d'aspiration professionnelle ; mais elles ont une influence indirecte par l'intermédiaire des notes scolaires. On peut estimer cette influence en faisant le produit des coefficients rencontrés sur le chemin allant de y_1 à y_3 en passant par y_2 : $.495 \times .550 = .273$. L'effet total des notes scolaires sur le niveau d'aspiration professionnelle est la somme de ces effets direct et indirect, .381.

On remarque, en examinant les coefficients de la matrice Γ que ceux qui mettent en jeu la variable x_2 (nombre de frères et sœurs) sont négatifs ; ceci tient à l'orientation de la variable x_2 (les familles nombreuses sont, en moyenne, un peu moins propices à la réussite scolaire, à la poursuite des études, à un

haut niveau d'aspiration professionnelle). Les coefficients γ_{31} et γ_{32} négatifs de façon inattendue, le sont assez faiblement pour pouvoir être négligés³.

L'examen des autres coefficients de la matrice Γ montre une forte influence de l'intelligence sur les notes scolaires. Les autres effets directs sont relativement faibles, mais il existe des effets indirects non négligeables. Par exemple x_1 influence y_2 par l'intermédiaire de y_1 : $.526 \times .405 = .213$; x_1 influence y_3 par l'intermédiaire de y_1 ($.526 \times .158$), par l'intermédiaire de y_2 ($.160 \times .550$) et par l'intermédiaire de y_1 et y_2 ($.526 \times .405 \times .550$), soit au total pour les effets indirects de x_1 sur y_3 .288.

La décomposition des effets totaux en effets directs et indirects présente l'intérêt d'éclairer sur les cheminements par lesquels s'exerce l'influence d'une variable sur une autre. Mais cette décomposition ne vaut que compte tenu des hypothèses qui ont été formulées et ont permis de spécifier le modèle. On obtiendrait une décomposition différente si, par exemple, l'ordre postulé entre y_2 et y_3 était modifié.

Le problème de l'ajustement du modèle ne se pose pas ici : ce modèle étant juste identifié s'adapte exactement aux données. Ceci se traduit par la valeurs des indices d'ajustement : $\chi^2 = 0$, GFI = 1, RMR = 0.

On pourrait tenter de rendre le modèle plus « économique » en posant comme nuls les effets les plus faibles. Par exemple, poser $\gamma_{32} = 0$ et $\gamma_{12} = 0$ reviendrait à supposer que le nombre de frères et sœurs n'a d'influence sur le niveau d'aspiration professionnelle que par l'intermédiaire des études envisagées. Une estimation du nouveau modèle serait alors nécessaire et il serait possible, le modèle devenant suridentifié grâce à l'introduction de ces nouvelles contraintes de voir dans quelle mesure il resterait capable de rendre compte des données malgré cette simplification. On pourrait aussi remettre en cause certaines des hypothèses initiales, par exemple admettre un effet réciproque de y_2 sur y_3 et de y_3 sur y_2 (les aspirations professionnelles étant considérées comme susceptibles d'avoir un effet sur les études envisagées. Le modèle deviendrait non récursif.

II. — EXEMPLE D'UTILISATION DU MODÈLE

Parmi tous les types d'études pouvant utiliser le modèle LISREL, nous en choisirons un qui a fait l'objet d'un assez grand nombre d'applications (voir, par exemple, Kohn et Schooler, 1978 ; Newman, 1984 ; Olweus, 1983) et dans lequel les possibilités offertes par le modèle ont permis des progrès méthodologiques importants.

Il s'agit d'un type d'étude dans lequel on cherche à savoir, de deux variables en covariation, laquelle influence l'autre ou au moins laquelle a le plus d'influence sur l'autre (il n'est pas rare, en effet, qu'il faille admettre la possibilité d'actions réciproques, éventuellement décalées dans le temps, entre les deux variables).

3. Un fort coefficient négatif là où on attendrait une influence positive doit être interprété avec prudence. Il peut être l'indication non pas que la variable indépendante considérée exerce une influence négative mais qu'elle joue le rôle de variable suppressive par rapport à une autre variable indépendante, c'est-à-dire l'épure de ses aspects non liés à la variable dépendante considérée (voir Van de Geer, 1971, p. 125-127 ; Bacher, 1982, II, p. 214, 408-409).

On ne peut guère espérer obtenir de réponse à une telle question, pour des variables non manipulables, lorsqu'on ne dispose sur ces variables que de mesures effectuées à une seule occasion. En revanche, si l'on dispose sur les mêmes sujets de mesures des deux variables à au moins deux occasions, il devient possible de tirer parti du moyen d'ordonner les effets que procure le décalage temporel : une variable mesurée antérieurement peut avoir un effet sur une variable mesurée postérieurement tandis que l'inverse peut être exclu sauf lorsque le moment de la mesure ne coïncide pas avec le moment auquel la variable se réfère (évocation du passé, représentation de l'avenir).

Le plan de base, pour ce type d'étude, consiste à mesurer à deux occasions les deux variables au sujet desquelles on s'interroge (two waves, two variables design).

Une méthode fréquemment utilisée pour obtenir une réponse à la question de l'influence dominante s'exerçant entre deux variables a consisté, suivant une suggestion de Campbell (1963), généralisant une méthode proposée par Lazarfeld (Lipset, Lazarsfeld, Barton, Linz, 1954), à comparer les corrélations croisées entre les deux variables mesurées aux deux occasions. La justification de cette méthode est que l'on peut attendre que la corrélation entre une variable « cause » antérieure et une variable « effet » postérieure soit plus forte que la corrélation entre une variable « effet » antérieure et une variable « cause » postérieure, un changement dans la variable « cause » provoquant un changement dans la variable « effet », alors que l'inverse n'a pas lieu.

Peu à peu, cependant, il apparut que les corrélations croisées étaient sensibles à de nombreux effets parasites par rapport à celui que pouvait exercer l'une des variables sur l'autre : évolution au cours du temps de la corrélation entre les deux variables mesurées au même moment, fidélités inégales, stabilités dans le temps inégales des deux variables, etc.

On trouve, dans Rogosa (1980), une analyse très précise de ces sources d'invalidation des conclusions tirées de la comparaison de corrélations croisées, analyse dont le bien-fondé a été largement reconnu (voir, notamment Kenny et Campbell, 1984).

La formalisation du problème en termes d'analyse de parcours, déjà ancienne (Duncan, 1969), présente au moins deux avantages : elle conduit à comparer les effets directs des variables l'une sur l'autre plutôt que leurs corrélations, qui sont ambigus car elles constituent les résultantes d'ensembles d'effets ; elle met en lumière la nécessité d'hypothèses complémentaires qui restaient, dans le cas de la comparaison de corrélations croisées, implicites.

Par rapport à l'analyse de parcours, le modèle LISREL apporte plus de souplesse dans le choix des hypothèses complémentaires ; il permet, en outre, par l'emploi de variables latentes, de tenir compte des erreurs de mesure et de situer l'étude au niveau de relations entre des variables plus abstraites et plus générales que les variables observées.

La recherche que nous présenterons est celle d'Olweus (1983). Il se propose d'éclairer les mécanismes qui peuvent expliquer l'association régulièrement observée entre comportements agressifs et faibles résultats scolaires chez des adolescents. 444 garçons suédois ont été examinés à deux reprises, en 6^e et en 9^e année de scolarité (13 et 16 ans). Trois variables latentes, mesurées chacune par 2 ou 3 variables observées, ont été utilisées à chaque occasion. Elles sont désignées par un de leurs pôles : mauvaises notes (mesurées par la

moyenne des notes en suédois et en anglais, SE, d'une part, la moyenne des notes en mathématiques et dans les autres matières théoriques, MR, d'autre part); attitude négative à l'égard de l'école (mesurée par les notes à deux parties d'un questionnaire, N1 et N2), comportements agressifs (mesurés à l'aide d'estimations des bagarres, A1, exprime de l'agressivité verbale à l'égard des professeurs, A2, exprime de l'agressivité verbale à l'égard des pairs, A3).

Plusieurs hypothèses ont été formulées dans les travaux antérieurs quant aux mécanismes pouvant expliquer la relation observée.

1 / Une mauvaise réussite scolaire crée de la frustration qui entraîne des réactions agressives compensatoires. Si c'est le cas, on devrait observer que de mauvaises notes en 6^e année conduisent à des comportements agressifs accrus en 9^e année, tandis qu'une forte agressivité en 6^e année n'influence pas les notes en 9^e année.

2 / Un comportement agressif est associé à un faible intérêt pour l'école et peut-être même à un rejet de tout ce qui est scolaire, ce qui entraîne à la fois de mauvaises notes et des réactions négatives de la part des professeurs. Dans ce cas, on devrait observer que des comportements agressifs en 6^e année déterminent de mauvaises notes en 9^e année alors que de mauvaises notes en 6^e année n'ont pas d'effet sur les comportements agressifs en 9^e année.

3 / Mauvaise réussite scolaire et comportements agressifs s'influencent mutuellement au cours du temps. Dans ce cas, on devrait observer simultanément les deux effets : de mauvaises notes en 6^e année entraînent des comportements agressifs en 9^e année ; des comportements agressifs en 6^e année entraînent de mauvaises notes en 9^e année.

4 / D'autres variables, précédant causalement les notes scolaires et les comportements agressifs, sont responsables des relations observées. Dans ce cas, l'introduction de ces variables dans le modèle devrait faire disparaître ces relations.

Une première analyse limitée aux variables de mauvaises notes, attitude négative, comportements agressifs a été effectuée. Elle a porté sur les corrélations entre ces variables.

Le modèle posé est celui de la figure 2.

Grâce à l'emploi d'indicateurs multiples, il n'y a pas de problème d'identification. Les variables latentes à la première occasion peuvent être considérées comme des variables ξ ; celles à la deuxième occasion comme des variables η . Les variables latentes sont reliées aux variables observées (x et y respectivement) par le modèle de mesure

$$\begin{aligned} \eta &= \Lambda_y y + \epsilon \\ \xi &= \Lambda_x x + \delta \end{aligned}$$

Les matrices Λ_x et Λ_y (que nous avons reconstituées tableau III) traduisent les hypothèses relatives à la variable latente mesurée par chaque variable observée.

Les échelles de mesure des variables latentes peuvent être fixées de différentes façons, par exemple en posant $\varphi_{11} = \varphi_{22} = \varphi_{33} = 1$.

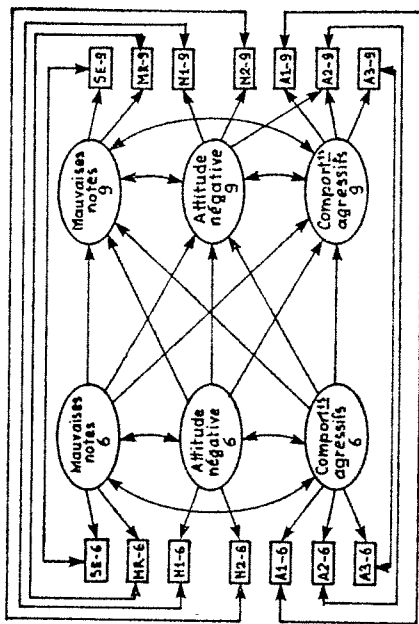


Fig. 2. — Modèle d'Olweus (1983) : relations entre mauvaises notes attitude scolaire négative et comportements agressifs (Olweus' model of relationships between poor grades negative attitude toward school and aggressive behavior)

Le modèle structural

$$\eta = B\eta + I\xi + \zeta$$

permet de spécifier les relations attendues entre ξ et η . Ici, l'auteur admet la possibilité que chaque variable ξ influence chaque variable η , la valeur obtenue pour les paramètres γ devant permettre ensuite de savoir si certains de ces effets peuvent être considérés comme nuls. L'auteur n'envisage pas d'effets instantanés entre les variables mesurées en une même occasion. Il admet seulement entre elles des corrélations (non orientées). Il n'y a donc pas de paramètres β à estimer. Les corrélations entre variables ξ sont calculées par le programme (matrice Φ); elles correspondent aux corrélations désatténuées entre variables observées x^4 . En ce qui concerne les corrélations entre variables η , l'auteur n'indique pas comment il a procédé. Une possibilité est de les reporter sur les variables résiduelles ζ affectant chaque variable η , en admettant des variables ζ en corrélation⁵. C'est ce que nous supposons, tableau III.

Le modèle LISREL permet aussi d'envisager des corrélations entre erreurs de mesure. Dans le cas de données longitudinales, de telles corrélations sont très probables, lorsqu'un même instrument est utilisé aux différentes occasions, entre les erreurs de mesure affectant cet instrument. Ces corrélations, compte

4. Une corrélation désatténuée est une corrélation corrigée pour tenir compte de l'atténuation qu'exerce l'infidélité des variables sur la valeur de la corrélation « vraie », r_{xy} . La corrélation désatténuée est :

$$r_{xy} / \psi_y = \frac{r_{xy}}{\sqrt{r_{xx}} \sqrt{r_{yy}}}$$

où r_{xy} est la corrélation observée, r_{xx} et r_{yy} les fidélités des variables.

5. Des corrélations entre variables ζ signifient qu'il existe des sources de variation communes (en dehors des variables explicitement incluses dans le modèle) affectant simultanément plusieurs variables η .

TABLEAU III. — Matrices correspondant au modèle de la figure 2
(Matrices corresponding to the model of figure 2)

x	Λ_x			ξ_3 (comportements agressifs, 6 ^e)
	ξ_1 (mauvaises notes, 6 ^e)	ξ_2 (attitude négative, 6 ^e)	ξ_3	
1. SE-6	λ_{11}	0	0	0
2. MR-6	λ_{41}	0	0	0
3. NI-6	0	λ_{32}	0	0
4. N2-6	0	λ_{42}	0	0
5. AI-6	0	0	λ_{53}	λ_{63}
6. A2-6	0	0	0	λ_{65}
7. A3-6	0	0	0	λ_{75}

y	Λ_y			γ_3 (comportements agressifs, 9 ^e)
	γ_1 (mauvaises notes, 9 ^e)	γ_2 (attitude négative, 9 ^e)	γ_3	
8. SE-9	λ_{81}	0	0	0
9. MR-9	λ_{91}	0	0	0
10. NI-9	0	λ_{102}	0	0
11. N2-9	0	λ_{112}	0	0
12. AI-9	0	0	λ_{123}	λ_{133}
13. A2-9	0	0	0	λ_{133}
14. A3-9	0	0	0	λ_{143}

γ_1	γ_2	γ_3	Φ		
			ξ_1	ξ_2	ξ_3
γ_{11}	γ_{12}	γ_{13}	Φ_{11}	Φ_{12}	Φ_{13}
γ_{21}	γ_{22}	γ_{23}	Φ_{21}	Φ_{22}	Φ_{23}
γ_{31}	γ_{32}	γ_{33}			

ψ_1	ψ_2	ψ_3	Ψ		
			ζ_1	ζ_2	ζ_3
Ψ_{11}	Ψ_{12}	Ψ_{13}	Ψ_{11}	Ψ_{12}	Ψ_{13}
Ψ_{21}	Ψ_{22}	Ψ_{23}			
Ψ_{31}	Ψ_{32}	Ψ_{33}			

θ_1	θ_2	θ_3	θ_4	θ_5	θ_6	θ_7	θ_8	θ_9	θ_{10}	θ_{11}	θ_{12}	θ_{13}	θ_{14}	Θ		
														ψ_1	ψ_2	ψ_3
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
θ_{22}	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	θ_{33}	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	θ_{44}	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	θ_{55}	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	θ_{66}	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	θ_{77}	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	θ_{88}	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	θ_{99}	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	θ_{1010}	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	θ_{1111}	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	θ_{1212}	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	θ_{1313}	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	θ_{1414}	0	0	0	0

Les modèles structuraux en psychologie

tenu du modèle de mesure, sont attribuables non seulement aux erreurs de mesure proprement dites mais à la spécificité (commune aux deux occasions) de l'instrument par rapport à la variable latente. Les hypothèses de corrélations de ce type formulées par Olweus font que la matrice θ comporte certains éléments supposés non nuls en dehors de la diagonale (tableau III).

Certains des paramètres estimés étant très faibles, l'auteur modifie en un second temps le modèle en posant égaux à 0 les coefficients non significatifs

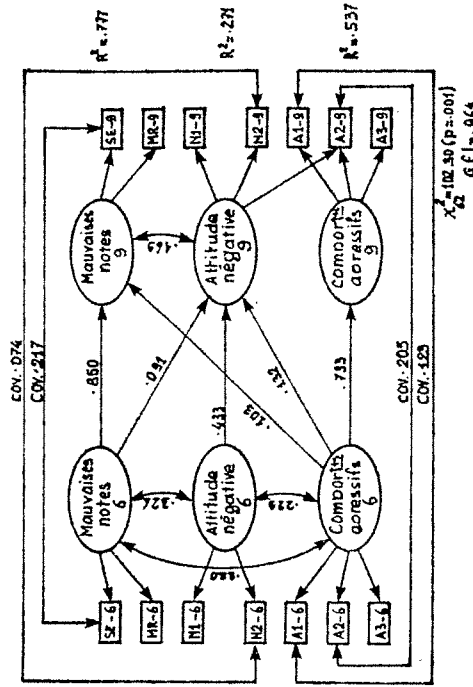


Fig. 3. — Résultats de la première analyse d'Olweus (1983) Relations entre mauvaises notes, attitudes scolaires négatives et comportements agressifs en 6^e et 9^e année de scolarité (Results of Olweus' first analysis Relationships between poor grades negative attitude toward school and aggressive behavior)

au seuil de .05. La figure 3 donne les résultats de l'estimation du modèle modifié. L'ajustement du modèle aux données est satisfaisant ($\chi^2_{68} = 102.30$, GFI = .964).

On observe, en considérant les coefficients γ , une forte stabilité des notes scolaires (.850) et des comportements agressifs (.733) entre 6^e et 9^e année. L'attitude négative est moins stable (.433). En ce qui concerne les effets croisés, on remarque que les notes scolaires en 6^e ont une influence faible (.091) sur l'attitude négative en 9^e année et que les comportements agressifs en 6^e année ont une influence (relativement faible), sur les mauvaises notes et sur l'attitude négative en 9^e année (.103 et .132). En revanche, les mauvaises notes en 6^e année n'ont pas d'influence sur les comportements agressifs en 9^e année. Ces résultats sont en faveur de l'hypothèse 2 plutôt que des hypothèses 1 et 3 : ce sont les comportements agressifs qui semblent déterminer de mauvaises notes plutôt que l'inverse, et on peut aussi écarter l'hypothèse d'effets réciproques.

Reste l'hypothèse 4. La première analyse a été effectuée sans tenir compte de variables antécédentes susceptibles d'influencer à la fois mauvaises notes et comportements

agressifs en 6^e et en 9^e année de scolarité et ainsi de rendre compte de leurs covariations.

Pour éprouver cette hypothèse, l'auteur procède à une seconde analyse en fonction d'un modèle dans lequel figure un ensemble de variables familiales et démographiques considérées comme prédéterminées par rapport aux variables mauvaises notes, attitude négative, comportements agressifs en 6^e et 9^e année. Ces variables, désignées par l'un de leur pôle sont les suivantes :

- Age faible des parents, mesuré par deux variables observées (âge du père, âge de la mère).
- Niveau social faible, mesuré par deux variables observées (revenu, statut professionnel).
- Enfant occupant un rang intermédiaire dans la fratrie.
- Parents divorcés.
- Enfant né hors mariage.

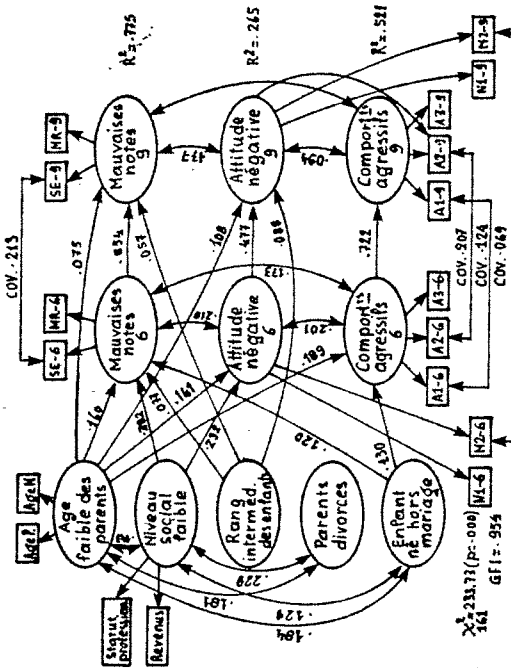


Fig. 4. — Résultats de la seconde analyse d'Olweus (1983) incluant des variables explicatives familiales et démographiques (Results of Olweus' second analysis including familial and demographic explanatory variables)

Leur introduction annule les effets croisés observés lors de la première analyse. Les résultats sont donc finalement favorables à l'hypothèse 4 : mauvaises notes et comportements agressifs sont sous la dépendance de variables familiales et démographiques qui suffisent à expliquer leurs covariations. On trouvera le détail des résultats, qui pourraient susciter de nombreux autres commentaires, fig. 4.

CONCLUSION

Nous n'avons pu donner qu'une faible idée du champ d'application du modèle LISREL. Toutes les méthodes statistiques classiques relevant du modèle linéaire général (analyse factorielle, régression multiple, analyse de la variance et de la covariance, généralisations multivariées de ces méthodes) peuvent être considérées comme des sous-modèles qui peuvent être analysés séparément et surtout combinés. Le modèle présente un intérêt particulier pour l'analyse de données longitudinales, par les hypothèses qu'il permet de faire sur l'existence d'erreurs corrélées dans le temps, et se prête également de façon particulièrement commode à la comparaison de groupes multiples sous l'angle de leurs moyennes, variances et covariances.

Les limites du modèle LISREL tiennent d'abord au fait qu'il s'agit d'un modèle linéaire ne permettant pas de traiter directement des interactions (bien qu'on puisse le faire indirectement en analysant séparément les groupes dans lesquels on pense que les lois régissant un phénomène pourraient jouer différemment).

Elles tiennent aussi à la très grande sensibilité de la méthode d'estimation à certains écarts, par rapport aux hypothèses de départ.

Les travaux consacrés à sa robustesse (Boosma, 1982, 1985 ; Gerbing, Anderson, 1985) montrent que le modèle LISREL ne doit pas être utilisé sur des échantillons d'effectif trop faible : une certaine de sujets est un minimum. La méthode est aussi sensible à l'asymétrie des distributions, à l'infidélité des mesures et, en ce qui concerne l'analyse factorielle, au rapport entre le nombre de variables et le nombre de facteurs et au niveau moyen des saturations. On observe plus souvent des cas de non-convergence ou de valeurs non admissibles des paramètres lorsque certaines de ces conditions sont défavorables.

Lorsque ceci se produit, il est parfois difficile de savoir si le modèle n'était pas identifié, s'il comporte une erreur de spécification importante, ou s'il s'agit d'une des raisons indiquées ci-dessus. La constatation d'une impossibilité d'estimer les paramètres d'un modèle envisagé par un utilisateur est, quoi qu'il en soit, assez fréquemment rapportée dans la littérature.

Nous indiquions, au début de cet article, que le modèle LISREL s'inscrit dans un courant plus général de développement de modèles structuraux. D'autres modèles et d'autres logiciels ont été proposés (voir Bentler, 1980, 1986 ; Kivieri et Speed, 1982 ; Long, 1983b) mais sont, jusqu'ici, beaucoup moins utilisés.

Le logiciel LISREL lui-même a fait l'objet d'un certain nombre d'améliorations au cours des dernières années. Par exemple, l'utilisateur n'a plus à proposer les valeurs initiales à partir desquelles se font les itérations aboutissant à la solution, valeurs dont le choix présentait quelque difficulté.

Dès maintenant, certaines de ses limites sont moins contraignantes qu'on pourrait le penser. Ainsi, par des transformations de variables, on peut normaliser des distributions ; on peut traiter des modèles linéaires dans les paramètres, de la forme :

$$y_2 = \gamma_{21} f(x_1) + \gamma_{22} f(x_2) + \gamma_{23} f(x_3) + \beta_{21} f(y_1) + \zeta_2$$

Cependant, on peut encore s'attendre à une évolution importante de ce modèle et des modèles voisins.

Les développements récents tendent en effet à dépasser les limites imposées par les modèles existants pour permettre, par exemple, d'analyser des variables ayant une distribution quelconque et de tenir compte de statistiques fondées sur des moments d'ordre supérieur (Bentler, 1986).

Si l'on se place, maintenant, du point de vue de l'utilisateur, ce qui est peut-être le plus important pour lui est le choix et la mise en œuvre d'une stratégie lui permettant de progresser dans la connaissance du phénomène qu'il souhaite étudier.

La formalisation des hypothèses, une fois maîtrisés les aspects techniques du modèle, est simple dans certains de ses aspects mais plus délicate dans d'autres. Il est, par exemple, assez difficile de saisir la portée des hypothèses portant sur l'indépendance ou la non-indépendance des erreurs de mesure et des erreurs dans les équations.

L'interprétation des coefficients d'un modèle structural va moins de soi qu'il n'y paraît. Il faut, en effet, bien situer chacun dans l'ensemble afin de repérer quelles sont les variables qui ont été tenues constantes lors de son estimation, variables de plus en plus nombreuses à mesure qu'on passe d'une variable à la suivante dans un modèle dont les variables sont ordonnées.

Une caractéristique importante du modèle LISREL, qui fait l'un de ses intérêts, est sa souplesse. Bien qu'il soit conçu comme moyen de formaliser des hypothèses et de les mettre à l'épreuve, ce modèle permet de laisser certaines questions ouvertes, de ne pas spécifier à l'avance un système complet mais de construire progressivement ce système en fonction d'hypothèses préliminaires et de constatations faites sur les données. Ceci permet la mise en œuvre d'une stratégie de construction d'un système rendant bien compte des données tout en restant acceptable théoriquement.

Cependant, mêler démarche hypothético-déductive et démarche inductive au cours d'une même étude comporte certains risques ; on peut, en particulier, se fier trop exclusivement, dans cette construction, aux indices formels d'adaptation aux données et faire passer le souci d'économie avant celui de fondement théorique ; on peut aussi ne plus très bien savoir ce qui a pu être éprouvé et ce qui est seulement suggéré par les résultats et nécessite une contre-validation sur d'autres données.

D'un point de vue plus général, le recours à la modélisation peut donner un faux sentiment d'accès à une vérité établie. Un modèle doit, nous semble-t-il, être essentiellement conçu comme un instrument provisoire utilisé pour faire évoluer la connaissance que l'on a d'un phénomène, des lois qui le régissent et de la généralité de ces lois.

La possibilité de « faire travailler » le modèle est sans doute l'aspect le plus intéressant de cette démarche. Lorsqu'on fait telle hypothèse, on obtient tel résultat ; que se passera-t-il si on la modifie ? Lorsqu'on analyse un système de variables comme s'il était clos, on évalue un certain nombre d'effets ; quelles conséquences aura l'introduction de nouvelles variables dans le système, c'est-à-dire la remise en cause de cette hypothèse de clôture ?

Par sa généralité et sa commodité d'utilisation le logiciel LISREL permet à de nombreux utilisateurs d'adopter cette démarche. Elle n'est pas sans difficulté, mais cette difficulté vaut d'être maîtrisée.

BIBLIOGRAPHIE

- Asher, H. B. (1983). — *Causal modeling*. Second edition, Beverly Hills, Sage Publications, Sage University Papers series on Quantitative Applications in the Social Sciences, n° 3, 96 p.
- Bacher, F. (1982). — *Les enquêtes en psychologie, I : Description, II : Explication*, Paris, Presses Universitaires de Lille, 462 p. et 488 p.
- Bacher, F. (1987). — Les modèles structuraux en psychologie, Présentation d'un modèle, LISREL, 1^{re} partie, *Le Travail humain*, 1987, 50, 347-370.
- Bentler, P. M. (1980). — Multivariate analysis with latent variables : causal modeling, *Annual Review of Psychology*, 31, 419-456.
- Bentler, P. M. (1986). — Structural modeling and Psychometrika : an historical perspective on growth and achievements, *Psychometrika*, 51, 35-51.
- Bentler, P. M., Bonett, D. G. (1980). — Significance tests and goodness of fit in the analysis of covariance structures, *Psychological Bulletin*, 88, 588-606.
- Berry, W. D. (1984). — *Nonrecursive causal models*, Beverly Hills, Sage Publications in the Social Sciences, n° 37, 94 p.
- Boomsma, A. (1982). — The robustness of LISREL against small sample sizes in factor analysis models, Part 1, p. 149-173, in Joreskog K. G., Wold H. (eds.), *Systems under indirect observation : causality, structure, prediction*, Amsterdam, North Holland.
- Boomsma, A. (1985). — Non convergence, improper solutions, and starting values in LISREL maximum likelihood estimation, *Psychometrika*, 50, n° 2, 229-242.
- Campbell, D. T. (1963). — From description to experimentation : interpreting trends as quasi-experiments, p. 212-242, in Harris C. (ed.), *Problems in measuring change*, Madison, Univ. of Wisconsin Press, 259 p.
- Duncan, O. D. (1969). — Some linear models for two-wave, two-variable panel analysis, *Psychological Bulletin*, 72, 177-182.
- Fox, J. (1984). — *Linear statistical models and related methods. With applications to social research*, New York, J. Wiley & Sons, 449 p.
- Gerbing, D. W., Anderson, J. C. (1985). — The effects of sampling error and model characteristics on parameter estimation for maximum likelihood confirmatory factor analysis, *Multivariate Behavioral Research*, 20, 255-271.
- Jöreskog, K. G., Sörbom, D. (1985). — LISREL VI. *Analysis of linear structural relationships by maximum likelihood, instrumental variables and least squares methods*, 2nd printing, Uppsala, University of Uppsala.
- Kenny, D. A. (1979). — *Correlation and causality*, New York, J. Wiley & Sons, 277 p.
- Kenny, D. A., Campbell, D. T. (1984). — Methodological considerations in the analysis of temporal data, p. 125-138, in Gergen K. J., Gergen M. M. (eds.), *Historical social psychology*, Hillsdale (NJ), Lawrence Erlbaum Associates.
- Kerchoff, A. C. (1974). — *Ambition and attainment*, Rose Monograph Series.
- Kivert, H., Speed, T. P. (1982). — Structural analysis of multivariate data : a review, *Sociological Methodology*, 209-289.
- Kohn, M. L., Schooler, C. (1978). — The reciprocal effects of the substantive complexity of work and intellectual flexibility : a longitudinal assessment, *American Journal of Sociology*, 84, 24-52.
- Lipset, S. M., Lazarsfeld, P. F., Barton, A. H., Linz, J. (1954). — The psychology of voting : an analysis of political behavior, p. 1124-1175, in Lindzey G. (ed.), *Handbook of social psychology*, Cambridge, Allison Wesley, vol. II, 1126 p.
- Long, J. S. (1983a). — *Confirmatory factor analysis*, Beverly Hills, Sage Publications, Sage University Papers series on Quantitative Applications in the Social Sciences, n° 33, 88 p.
- Long, J. S. (1983b). — *Covariance structure models. An introduction to LISREL*, Beverly Hills, Sage Publications, Sage University Papers series on Quantitative Applications in the Social Sciences, n° 34, 95 p.

- Newman, R. S. (1984). — Children's achievement and self-evaluations in mathematics : a longitudinal study, *Journal of Educational Psychology*, 76, n° 5, 857-873.
- Olweus, D. (1983). — Low school achievement and aggressive behavior in adolescent boys, p. 353-368, in Magnusson D., Allen V. L. (eds.), *Human development. An interactional perspective*, New York, Academic Press, 393 p.
- Rogosa, D. (1980). — A critique of cross lagged correlations, *Psychological Bulletin*, 88, 245-258.
- Van de Geer, J. P. (1971). — *Introduction to multivariate analysis for the social sciences*, San Francisco, Freeman & Co., 293 p.

RÉSUMÉ

Après une première partie consacrée à l'analyse factorielle confirmatoire, la deuxième partie de cet article présente l'utilisation du modèle LISREL pour l'étude de relations supposées, éventuellement causales, entre les variables d'un système. Les possibilités et limites du modèle et du logiciel correspondant sont brièvement rappelées et l'exemple d'une étude analysant les relations croisées entre des variables mesurées à deux occasions sur les mêmes sujets est utilisé pour illustrer la démarche.

(Mots clés : Construction de modèles, Modèles structuraux, LISREL, Relations croisées.)

PSYCHOLOGIE ERGONOMIQUE DE LA PROGRAMMATION INFORMATIQUE

APPROCHE COGNITIVE DE LA PROGRAMMATION INFORMATIQUE

Introduction

par J.-M. Hoc* et W. VISSER**

SUMMARY

Cognitive approach to computer programming. — Before presenting the papers of this special issue, psychological research on programming is defined within the broader domain of human-computer interaction study. Thus it can be shown that a number of questions examined in the restricted domain of programming are also relevant to other human-computer interaction situations. An overview of the evolution of theoretical and methodological frameworks in psychology of programming is given in order to outline the context of the contributions to this issue, which presents : the evolution of the computer science conceptions of programming, the part of programming methods in activity guidance, and illustrations of empirical approaches to the main components of this activity (program designing, understanding, and debugging).

I — LA PROGRAMMATION ET L'ÉTUDE DES SITUATIONS D'INTERACTION HOMME-ORDINATEUR

L'ergonomie de la programmation informatique est habituellement classée sous le thème plus général de l'ergonomie des interactions homme-ordinateur. La raison en est qu'elle a constitué l'un des deux premiers thèmes d'intérêt dans ce domaine, avec l'ergonomie des logiciels de saisie de données. C'est évidemment une interaction particulièrement pauvre qui se manifeste dans une activité de programmation où l'homme contrôle complètement la machine et dans une activité de saisie où, à l'inverse, l'homme est entièrement contrôlé par la machine.

L'évolution des situations de travail informatisées, que nous avons connue au cours de ces vingt dernières années, nous a habitués à rencontrer des situa-

* Directeur de recherche au CNRS, URA, « Psychologie cognitive du traitement de l'information symbolique », Université de Paris VIII, 2, rue de la Liberté, 93526 Saint-Denis.

** INRIA, Groupe de Psychologie ergonomique, Domaine de Voluceau, Rocquencourt, BP 105, 78153 Le Chesnay Cedex.