

# L'analyse du changement, la causalité et le devis longitudinal : trois modèles élémentaires et un exemple

## The Analysis of Change, Causality, and the Longitudinal Protocol: Three Elementary Models and an Example

François BÉLAND

Volume 18, Number 2, octobre 1986

Travail, santé, prévention

URI: <https://id.erudit.org/iderudit/001032ar>

DOI: <https://doi.org/10.7202/001032ar>

[See table of contents](#)

Publisher(s)

Les Presses de l'Université de Montréal

ISSN

0038-030X (print)

1492-1375 (digital)

[Explore this journal](#)

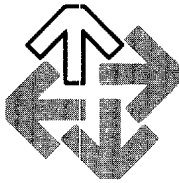
### Article abstract

Establishing causal relationships between variables has been one of the most common operations in sociology. However, after it became evident that transversal analysis did not make it possible to validate causal relationships, recourse to longitudinal analysis seemed for many to be a way out. There too, however, it was established that longitudinal analysis could only partially remove the limits on our capacity to prove causal relationships, without eliminating them altogether. Three models for relationships between variables are defined here and their capacity for generating survey data verified. Even though a formal procedure for investigating these models is used, it is clear that choosing one of the three contradictory models remains difficult. The intention of this paper is not to discredit the use of longitudinal data, but to emphasize a few of the difficulties in applying different methods of analysis to these types of data.

### Cite this article

BÉLAND, F. (1986). L'analyse du changement, la causalité et le devis longitudinal : trois modèles élémentaires et un exemple. *Sociologie et sociétés*, 18(2), 151–164. <https://doi.org/10.7202/001032ar>

# L'analyse du changement, la causalité et le devis longitudinal: trois modèles élémentaires et un exemple\*



FRANÇOIS BÉLAND

---

## 1. L'ANALYSE LONGITUDINALE ET LA CAUSALITÉ

Une relation causale implique trois conditions (Cook et Campbell, 1979): 1) la cause précède temporellement l'effet; 2) il existe une relation fonctionnelle entre la cause et l'effet; 3) la relation entre la cause et l'effet n'est pas une fonction d'autres phénomènes.

Le devis longitudinal semble particulièrement bien adapté à l'étude de relation causale puisque 1) l'ordre temporel entre la cause et l'effet peut être établi par l'observation simultanée et différée de chacun d'entre eux, 2) des méthodes d'analyse sont disponibles pour étudier la distribution conjointe de phénomènes dans le temps auprès des mêmes individus et 3) l'observation d'autres phénomènes ou l'introduction de variables latentes permettent d'étudier les rapports nets entre la cause et l'effet.

Lorsque deux variables X et Y sont observées à deux moments, l'effet causal de X sur Y suppose que seul le rapport qui va de X au temps 1, soit  $X_1$ , à Y au temps 2, soit  $Y_2$ , est légitime. Le lien causal de Y au temps 1, soit  $Y_1$ , à X au temps 2, soit  $X_2$ , doit être nul. La recherche de relations causales en sciences sociales pourrait se coupler à cette démarche très opérationnaliste qui y verrait l'occasion d'appliquer des techniques statistiques relativement connues, telles la comparaison de la paire d'intercorrélations diachroniques (connue en anglais sous le nom de *Cross-lagged panel correlation analysis* — CLPC — voir Crano, Kenny et Campbell, 1972; Pelz et Andrew, 1964; Rozelle et Campbell, 1969) ou l'analyse des cheminements de causalité (Duncan 1975). Nous verrons dans ce qui va suivre que ces deux techniques d'analyse ne sont pas en soi fautives d'opérationnalisme. Le CLPC est une seule des formalisations possibles d'un modèle longitudinal satisfaisant les exigences d'une hypothèse précise à propos de relations entre variables. L'analyse des cheminements de causalité est plus souple puisqu'elle se prête à l'écriture formelle de modèles qui traduisent des hypothèses mutuellement incompatibles à propos des relations entre variables observées longitudinalement. L'espoir d'un chercheur étant de pouvoir choisir, sur la base des résultats de tests d'hypothèse, l'un de ces modèles.

L'objectif général de ce texte est d'illustrer par un exemple que l'analyse longitudinale n'offre pas en soi une réponse à des problèmes de recherche conçus à l'aide des rapports de causalité, mais qu'elle permet de formuler des hypothèses qui autrement seraient inaccessibles. Dans ce but deux modèles d'analyse relativement connus dans la littérature seront présentés (Kenny,

---

\* L'auteur est chercheur-boursier du Fonds de la recherche en santé du Québec.

1973, 1975) et un troisième qui mérite plus d'attention que celle qui lui a été réservée jusqu'ici sera introduit. Les trois modèles seront incompatibles en ce sens que les rapports causaux qu'ils définissent entre les variables se contredisent. Dans une première partie, les trois modèles seront définis et dans une seconde partie un exemple d'utilisation des trois modèles sera présenté.

## 2. TROIS MODÈLES POUR L'ANALYSE DE DONNÉES LONGITUDINALES

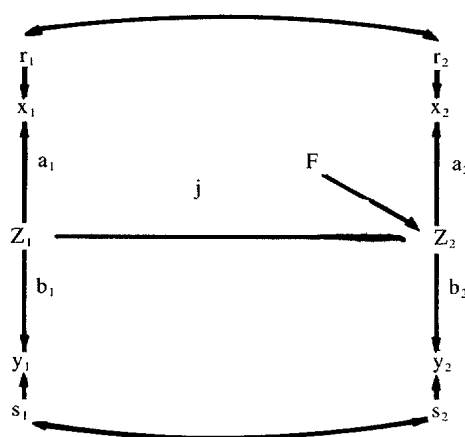
La vérification empirique de la capacité de données longitudinales de prédire la direction des relations causales, comme premier moment de l'analyse, est une démarche prématurée. Une question importante a été occultée: existe-t-il des liens de causalité entre les deux variables? En effet, l'étude de la causalité à l'aide de données longitudinales se développe en deux étapes. La première étape vérifie s'il est raisonnable de soutenir l'hypothèse de rapports de causalité dans le système, la deuxième étape propose, quant à elle, une direction aux rapports de causalité. Ces deux étapes ont donné lieu, dans la littérature, au développement de deux modèles relativement connus. Aussi leur présentation sera-t-elle brève et se limitera au cas où deux variables sont observées à deux moments dans le temps. Ces modèles seront schématisés en utilisant la représentation habituelle des modèles de cheminement de causalité. Cette procédure est donc utilisée ici pour formaliser la représentation d'hypothèses diverses sur des données longitudinales.

### 2.1 ÉTABLIR LA PRÉSENCE DE LA CAUSALITÉ

La première question à se poser lors de l'étude d'hypothèses causales ne concerne pas la direction de la causalité, mais la possibilité même qu'il existe des rapports causaux dans un système de rapports entre variables. La figure 1 présente un contre-modèle, où les rapports entre les variables X et Y sont médiatisés par une variable latente Z. Ce modèle rejette donc la troisième condition pour qu'existe des liens de causalité dans un système longitudinal de rapports entre variables: les liens entre les variables ne sont pas causés par un autre ensemble de variables.

FIGURE 1

Contre-modèle pour rejeter l'hypothèse  
de relation causale entre deux variables  
dans un étude longitudinale



Depuis le réseau de relations causales de la figure 1, il est aisé d'obtenir des équations qui permettent d'estimer les principaux paramètres du modèle, et en particulier de vérifier si le rapport entre les deux variables latentes  $Z_1$  et  $Z_2$  est différent de zéro.

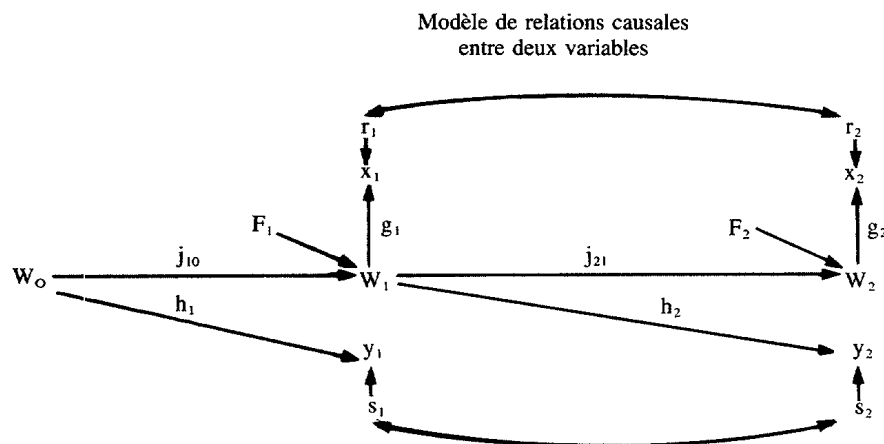
Mais le test d'hypothèse de la stabilité des corrélations synchroniques entre X et Y ( $H_0: [a_1 = a_2; b_1 = b_2]$ ), étant donné le modèle de la figure 1, est le même que celui du test courant et connu du CLPC qui détermine si les intercorrélations diachroniques, soit  $r_{x_1y_2}$  et  $r_{y_1x_2}$ , sont égales. Si cette hypothèse nulle est rejetée, la direction de la causalité est habituellement déterminée par la valeur de l'intercorrélations diachronique la plus élevée.

Il semble donc que la technique de la comparaison de la paire d'intercorrélations diachroniques peut être employée pour vérifier, non seulement une hypothèse générale de causalité, mais aussi la direction de la causalité.

## 2.2 UN PREMIER MODÈLE DE RAPPORT CAUSAL ENTRE VARIABLES: L'EFFET DIFFÉRÉ

Dans un contexte où deux variables sont observées à deux moments différents, un rapport de causalité différée entre  $X$  et  $Y$  peut se schématiser comme à la figure 2. Ce schéma introduit trois variables latentes  $W$  qui représentent l'état d'un système de relations causales à trois moments dont deux sont observés. La variable observée  $X$  dépend de l'état du système de façon contemporaine, tandis que la variable observée  $Y$  dépend de l'état du système à moment antérieur. La variable latente  $W$  représente donc la partie commune de la variation  $X_{t-1}$  et de  $Y_t$ , lorsque le  $t$  indicé indique la période d'observation des variables. Ce modèle est appelé ici le modèle de causalité différée puisqu'il suppose que l'effet d'une variable sur une autre se manifeste à un moment postérieur. Ce modèle représente la façon dont les relations causales ont été conceptualisées de façon implicite dans la plupart des études longitudinales ou latitudinales en sciences sociales. En effet, la plupart des études latitudinales supposent que les estimations des effets causals qu'elles obtiennent sont une conséquence contemporaine d'effets différés dans le temps. Il est possible, avec des contraintes appropriées, d'estimer la valeur des coefficients  $j$  qui mesurent l'impact de l'état du système à un moment  $t-1$  sur le même système au moment  $t$ . Les contraintes pour estimer la valeur des  $j$  consiste à supposer que  $g_1$  est égal à  $g_2$ , et que  $h_1$  est égal à  $h_2$ . Kenny (1973) a cependant démontré que ces contraintes, nécessaires pour estimer les paramètres, peuvent conduire aux mêmes conséquences que le modèle de la figure 1. Autrement dit, l'hypothèse nulle  $H_0$ : [ $a_1 = a_2$ ;  $b_1 = b_2$ ] n'est pas une conséquence unique d'un modèle qui suppose que les rapports entre les variables  $X$  et  $Y$  ne sont pas d'ordre causal.

FIGURE 2



Les modèles représentés aux figures 1 et 2 conceptualisent deux hypothèses contraires, soit l'absence de relations causales dans le premier cas et la présence d'un effet de  $X$  sur  $Y$  dans le second. Mais l'un et l'autre peuvent avoir les mêmes conséquences quant à la structure des rapports entre corrélations, soit l'égalité des intercorrélations diachroniques entre  $X$  et  $Y$ . Donc, quoique conceptuellement différents, les deux modèles peuvent avoir des conséquences instrumentales identiques.

Reste que ces deux modèles peuvent être soumis à un test statistique et être tous deux rejetés, auquel cas les rapports entre les variables  $X$  et  $Y$  devraient être repensés. Mais la probabilité qu'une hypothèse équivalente à  $H_0$  ne soit pas rejetée diminue au fur et à mesure qu'augmente le nombre de périodes d'observation si le modèle de la figure 2 est plausible. La certitude qu'une relation causale existe ne peut donc jamais être acquise.

### 2.3 UN DEUXIÈME MODÈLE DE RAPPORT CAUSAL ENTRE VARIABLES: L'EFFET CONTEMPORAIN

De façon générale, les chercheurs assument que les causes prennent un certain temps avant de manifester pleinement leurs effets. Ils assument aussi que seuls les effets causals qui durent présentent un certain intérêt scientifique. Mais l'un et l'autre de ces choix sont des postulats qui ont rarement été interrogés. L'état actuel du développement des sciences sociales pourrait tout aussi bien soutenir un autre ensemble de postulats: 1) les effets causals dans les sociétés sont quasi instantanés, 2) la structure des effets causals quasi instantanés est relativement stable dans le temps, 3) les changements internes des relations causales se manifestent par une transformation des directions et de la valeur des paramètres qui les mesurent et 4) les études longitudinales doivent chercher à expliquer la stabilité et le changement dans le temps de la structure de relations causales quasi instantanées.

Ces quatre postulats supposent que l'intérêt des chercheurs se déplace de la détermination du sens de relations causales à l'étude des modalités de changement des rapports entre variables dans le temps. Ces rapports peuvent être de nature causale bien sûr, et la confiance dans la directionnalité de la causalité augmentera avec l'observation de la stabilité des valeurs des paramètres définissant les relations causales dans plusieurs cueillettes de données longitudinales auprès d'une même population.

Ce modèle peut être schématisé de la même façon que les deux modèles précédents. À la figure 3, l'absence de relations diachroniques entre variables latentes  $X$  et  $Y$  est notée par l'absence de flèches entre  $X_1$  et  $Y_2$  d'une part et  $Y_1$  et  $X_2$  de l'autre. Les variables du modèle sont mesurées avec erreur tandis que ces erreurs sont corrélées entre périodes d'observation, comme aux modèles précédents. L'hypothèse de stabilité du modèle est de deux ordres: 1) la stabilité de la relation causale dans le temps qui s'exprime par l'égalité des paramètres  $a_1$  et  $a_2$  et 2) la stabilité des erreurs de mesure  $e_i$  dans le temps d'où  $b_1 = b_2$  et  $c = c_2$ . Le respect de cette condition accroît la probabilité d'observer la stabilité du modèle, mais de façon générale le contrôle de l'instabilité des erreurs est une condition qui assure la validité de la stabilité des relations causales puisque leur variation peut n'être qu'une conséquence de la variation de la fiabilité des mesures. Enfin, une troisième condition de stabilité d'un système de relations entre variables peut être imposée si le chercheur dispose de plus de deux périodes d'observation: les effets diachroniques d'une variable sur elle-même sont stables dans le temps. De façon presque naturelle, la stabilité des corrélations diachroniques des erreurs peut aussi être imposée dans ce contexte.

Ce type de modèle qui suppose que les relations causales dans un réseau de variables sont contemporaines, mais sans impact temporel mutuel, n'est pas souvent vérifié par les chercheurs. Il s'attarde moins à vérifier la structure causale entre des variables qu'à étudier la structure du changement dans ce système. En effet, les modèles des figures 1, 2 et 3 peuvent se concevoir comme des hypothèses concurrentes sur la façon dont un système évolue, change ou est stable dans le temps.

Dans chacun de ces cas, la causalité est un instrument conceptuel qui facilite l'étude du changement dans le cas de données longitudinales.

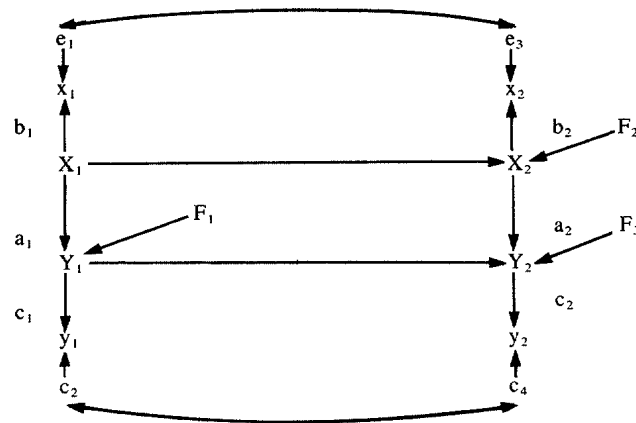
## 3. UN EXEMPLE D'UTILISATION DES TROIS MODÈLES: LE STATUT SOCIAL

### 3.1 CONCEPTUALISATION DE LA NOTION DE STATUT SOCIAL

Les trois modèles seront illustrés à l'aide de données empiriques provenant d'une étude d'un échantillon de 1 482 travailleurs du Québec de 35 ans et plus. Renaud, Bernard et Berthiaume (1980) ont mesuré rétrospectivement l'emploi en début de carrière et à 35 ans de cet échantillon de travailleurs.

Les variables utilisées ici sont la formation générale (FG) et la formation professionnelle (FP); le statut socio-économique en début de carrière (SSE1) et à 35 ans (SSE35) calibrés en fonction de l'échelle de Blishen et McRoberts (1976); la qualification professionnelle en début de carrière (PPS1) et à 35 ans (PPS35) mesurées par la durée de la préparation professionnelle spécifique établie par la classification canadienne descriptive des professions. Enfin, les revenus en début de carrière (REV1) et à 35 ans (REV35) sont établis de la façon habituelle par une question dans l'interview.

FIGURE 3

Modèle de relations causales contemporaines  
entre variables

Le statut social ou socio-économique a longtemps été conçu comme un concept unidimensionnel, du moins de façon implicite, puisque la plupart des études ont employé soit le prestige de l'occupation pour le représenter, tel l'indice de Blighen et McRoberts (1976), soit un classement des individus dans une liste de catégories occupationnelles ou autres (Hauser et Featherman, 1978; Wright et Peronne, 1977). Des chercheurs ont émis des critiques sévères des plus récents indices de statut social en insistant précisément sur les qualités conceptuellement différentes des dimensions qui le composent (Pawson, 1982). Selon cette conception, la question des rapports entre les différentes dimensions du statut social est problématique plutôt que déterminée par définition dans la mesure où elles servent à saisir la place d'un individu selon les multiples facettes de son statut social. Mais dans ce contexte, la tâche de la classification qui consiste à placer un individu dans une catégorie n'est plus centrale. Il s'agit plutôt de conceptualiser les rapports entre les facettes des dimensions qui définissent le statut. L'opération d'attribution d'un statut social change donc complètement de nature, elle devient une opération de conceptualisation des rapports entre dimensions théoriquement autonomes, même si elles restent empiriquement corrélées.

Trois modèles conceptualisent différemment les rapports entre les dimensions du statut social. Un modèle qui suppose que le statut social peut être représenté par un concept latent unique se résume par un modèle semblable à celui de la figure 1. Un modèle qui suppose que les rapports entre dimensions conceptuellement autonomes du statut social sont causals dans un sens classique peut être représenté par un schéma dérivé de celui de la figure 2. Tandis qu'un modèle qui propose une hypothèse de relations causales uniquement contemporaines entre les dimensions du statut social peut être représenté par un schéma qui ressemble à celui de la figure 3. Ces trois modèles seront ici adaptés aux types de données disponibles dans l'étude de Renaud, Bernard et Berthiaume (1980). Ces trois modèles permettront d'étudier 1) la multidimensionnalité du concept de statut social tel que composé du prestige de l'occupation, de la qualification de l'occupation et du revenu, et 2) les changements des rapports entre ces dimensions pendant la carrière professionnelle d'un échantillon de travailleurs du Québec.

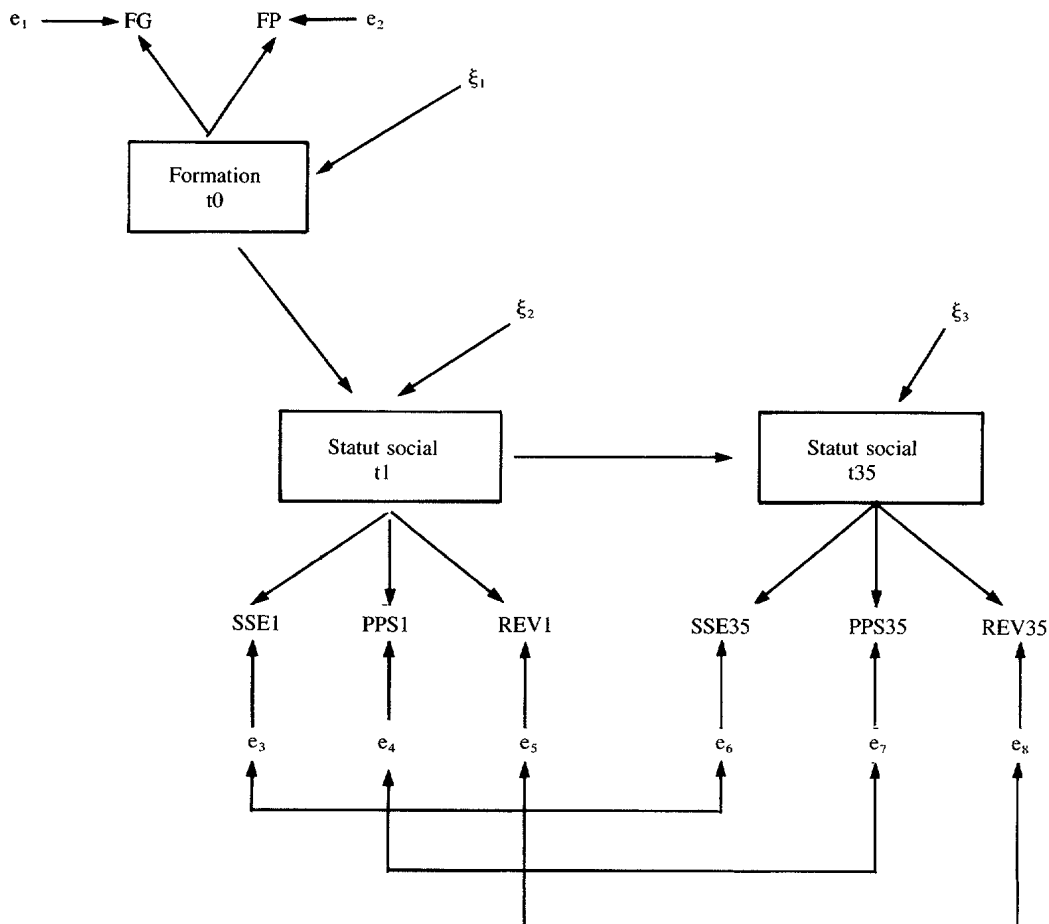
### 3.2 DESCRIPTION DES MODÈLES

Les trois modèles sont schématisés aux figures 4, 5 et 6 de la façon habituelle. L'écriture des modèles a suivi quelques principes généraux: 1) les modèles ont été représentés en fonction de la théorie qui les sous-tendent plutôt qu'en fonction des contraintes nécessaires pour obtenir une estimation de leurs paramètres. L'estimation des paramètres a été faite en tenant compte des contraintes d'égalité définies par les modèles de causalité contemporaine ou différée. 2) la formation générale et la formation professionnelle sont certes premières parmi les variables disponibles de cette étude dans la séquence temporelle impliquée par le devis de l'étude. Mais il est aisé d'imaginer qu'elles ne sont pas les sources premières de la détermination de l'état du système de relations

entre variables. La formation est donc influencée par d'autres facteurs, non mesurés ici. 3) l'hypothèse que l'effet de la formation sur le statut social à 35 ans est entièrement transmis par le statut social en début de carrière n'est pas toujours essentiel dans chacun des modèles. Mais elle a été introduite systématiquement, nous laissant le loisir d'introduire une relation directe entre la formation et les dimensions du statut social à 35 ans lorsque appropriée dans le processus de vérification empirique du modèle. 4) la formation n'est pas à proprement parler un des éléments d'un statut social. Il s'agit plutôt de ressources qu'un individu acquiert pour accéder à un statut social attribué sur le marché du travail par le prestige de l'occupation, sa qualification professionnelle et son revenu. La formation conserve donc un caractère particulier et non assimilable au statut social. Elle décrit l'état du système au temps d'observation zéro ( $t_0$ ). L'introduction des deux variables de formation aura donc des conséquences pour l'écriture du modèle de causalité différée en particulier puisque l'état du système au temps zéro ( $t_0$ ) est décrit par l'intermédiaire d'indicateurs observés plutôt que par les variables latentes du Temps 1 comme à la figure 2. Ce fait a aussi pour conséquence que la corrélation entre les erreurs de mesure des mêmes variables à deux moments différents n'est pertinent qu'avec les mesures du statut socio-économique, de la qualification professionnelle et du revenu.

FIGURE 4

Schéma représentant le modèle du concept latent unique de statut social



Chaque modèle est aussi écrit en fonction de ses particularités propres. Le modèle du concept latent unique s'opérationnalise par le schéma causal de la figure 4. Le modèle est simple et classique. Il utilise la façon habituelle de représenter des variables latentes par des indicateurs. Les variables latentes réfèrent ici à des étapes différentes de la carrière d'un individu.

Le modèle de causalité différée suppose que le statut social comprend une dimension contemporaine, soit le SSE1, mesuré par le prestige de l'occupation selon l'échelle de Blighen et McRoberts (1976), et qu'il influence deux dimensions de façon différée, soit la qualification professionnelle (PPS35) et le revenu (REV35). La qualification professionnelle se définit de façon contemporaine par le PPS1 et influence de façon différée le revenu (REV35). La figure 5 représente ce modèle où chacune des composantes du statut social conserve son indépendance conceptuelle.

Enfin, le modèle de la causalité contemporaine entre les trois dimensions du statut social est représenté à la figure 6. Chacune de ces trois dimensions est représentée par une variable latente. Elles sont en relation directe entre elles synchroniquement seulement, tandis qu'elles ne s'influencent pas diachroniquement au-delà d'une autorelation causale.

FIGURE 5

Schéma représentant le modèle de relations causales entre le statut socio-économique, la qualification professionnelle et le revenu

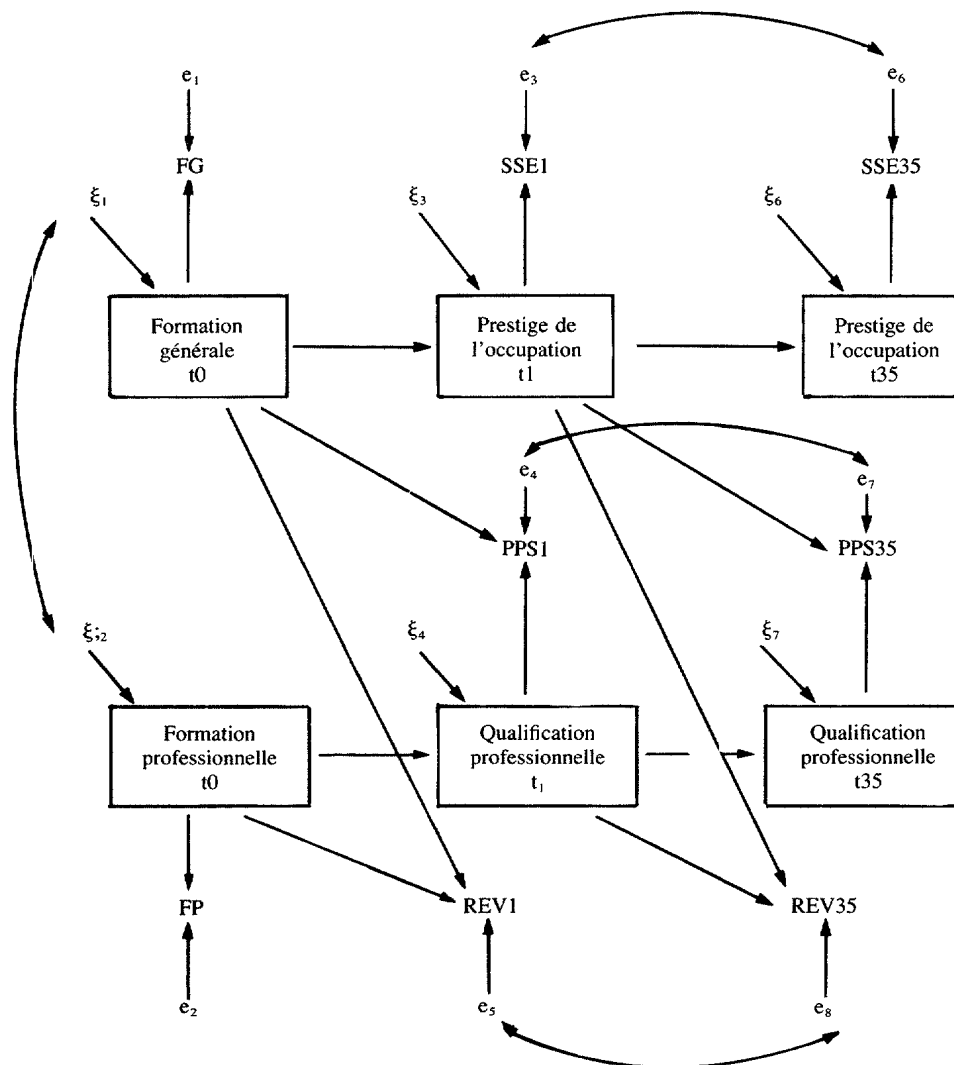
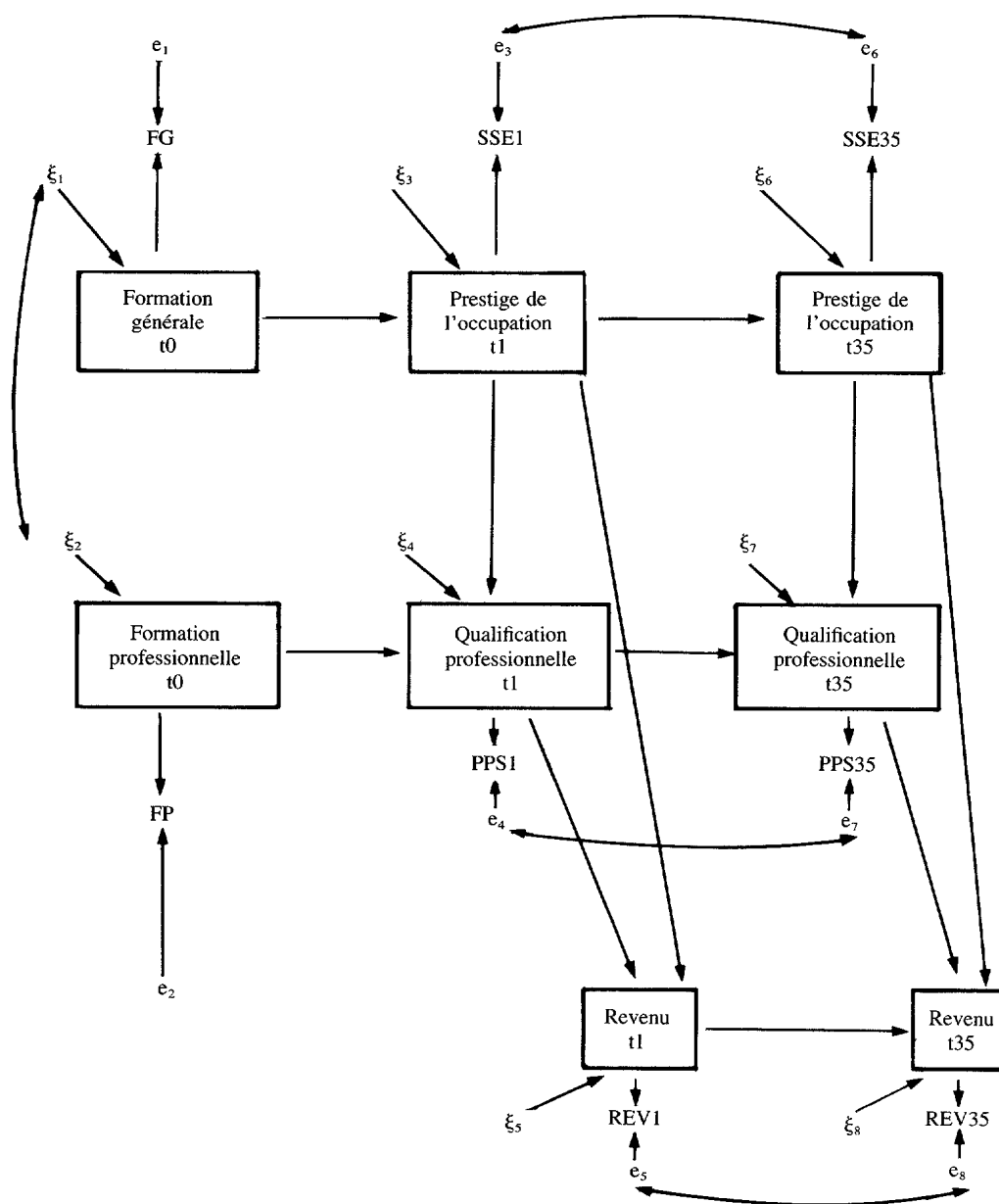




FIGURE 6

Schéma représentant le modèle de causalité contemporaine entre le statut socio-économique, la qualification professionnelle et le revenu



### 3.3 VÉRIFICATION EMPIRIQUE DES TROIS MODÈLES

La matrice des corrélations (Tableau 1), calculée à l'aide des coefficients de cheminements de causalité donnés dans Renaud, Bernard et Berthiaume (1980), a été soumise à la procédure LISREL (Jöreskog et Sörbom 1983) pour l'estimation des paramètres des trois modèles.

Si les paramètres du modèle du concept latent unique (fig. 7) ont pu aisément être estimés, il n'en fût pas de même des deux autres modèles. Le modèle de causalité contemporaine a produit un cas de Heywood, soit des estimations négatives des variances, et par conséquent, une matrice de variance-covariance des erreurs non positive définie, tandis que le modèle de causalité différée

a produit une matrice de variance-covariance non positive définie. Ces résultats indiquent clairement que le nombre de concepts indépendants dans ces deux modèles est trop grand. Bien sûr, il est possible d'utiliser des artifices pour prévenir les cas de Heywood (Rindskopf, 1983), mais l'utilisation de ces procédures est décevante dans la mesure où l'observation de cas de Heywood est un instrument diagnostic précieux de la validité du modèle.

TABLEAU 1

Matrice des corrélations entre huit variables  
observées auprès d'un échantillon de travailleurs  
de 35 ans et plus au Québec

|       | FG    | FP    | SSE1  | PPS1  | REV1  | SSE35 | PPS35 | REV35 |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| FG    | 1,000 |       |       |       |       |       |       |       |
| FP    | 0,589 | 1,000 |       |       |       |       |       |       |
| SSE1  | 0,624 | 0,625 | 1,000 |       |       |       |       |       |
| PPS1  | 0,445 | 0,584 | 0,704 | 1,000 |       |       |       |       |
| REV1  | 0,165 | 0,185 | 0,265 | 0,266 | 1,000 |       |       |       |
| SSE35 | 0,615 | 0,613 | 0,694 | 0,474 | 0,180 | 1,000 |       |       |
| PPS35 | 0,502 | 0,563 | 0,561 | 0,540 | 0,170 | 0,746 | 1,000 |       |
| REV35 | 0,310 | 0,394 | 0,361 | 0,325 | 0,325 | 0,412 | 0,413 | 1,000 |

Sur cette base, il apparaît que le modèle du concept latent unique est plus adéquat que les deux modèles de causalité. Le modèle de concept latent unique propose qu'une seule dimension permet de définir le statut social, tandis que les deux autres modèles causals imposent trois dimensions conceptuellement indépendantes au statut social. Mais les modèles des figures 4, 5 et 6 pourraient être repris en tenant compte du diagnostic indiqué par la présence de cas de Heywood. Le statut social pourrait être conceptualisé en deux dimensions plutôt qu'en trois. Ici la mesure du prestige de l'occupation et la qualification professionnelle ont été réunies dans une seule variable latente nommée statut socio-professionnel, tout en laissant le revenu comme dimension indépendante. Les dimensions générale et professionnelle de la formation ont aussi été réunies dans une seule variable latente. Des modèles modifiés, mais inspirés de ceux des figures 5 et 6, ont été écrits et soumis à la procédure LISREL pour estimer leurs paramètres. Dans chacun des cas, des estimations acceptables des paramètres ont été obtenues (voir fig. 8 et 9).

La procédure LISREL permet d'évaluer le degré de conformité d'un modèle aux données par l'intermédiaire d'un test de  $\chi^2$ . Deux des trois modèles des figures 7 à 9 sont acceptables selon ce critère. En effet, le seuil de signification statistique habituelle de 0,05 est largement satisfait pour les modèles des figures 8 et 9. En somme, l'analyse statistique ne nous fournit pas de critère pour choisir entre deux des hypothèses concurrentes d'explication du mode d'attribution des statuts sociaux au cours des carrières d'un échantillon de travailleurs du Québec. D'autres critères que statistiques doivent être utilisés si un choix doit être fait entre les modèles.

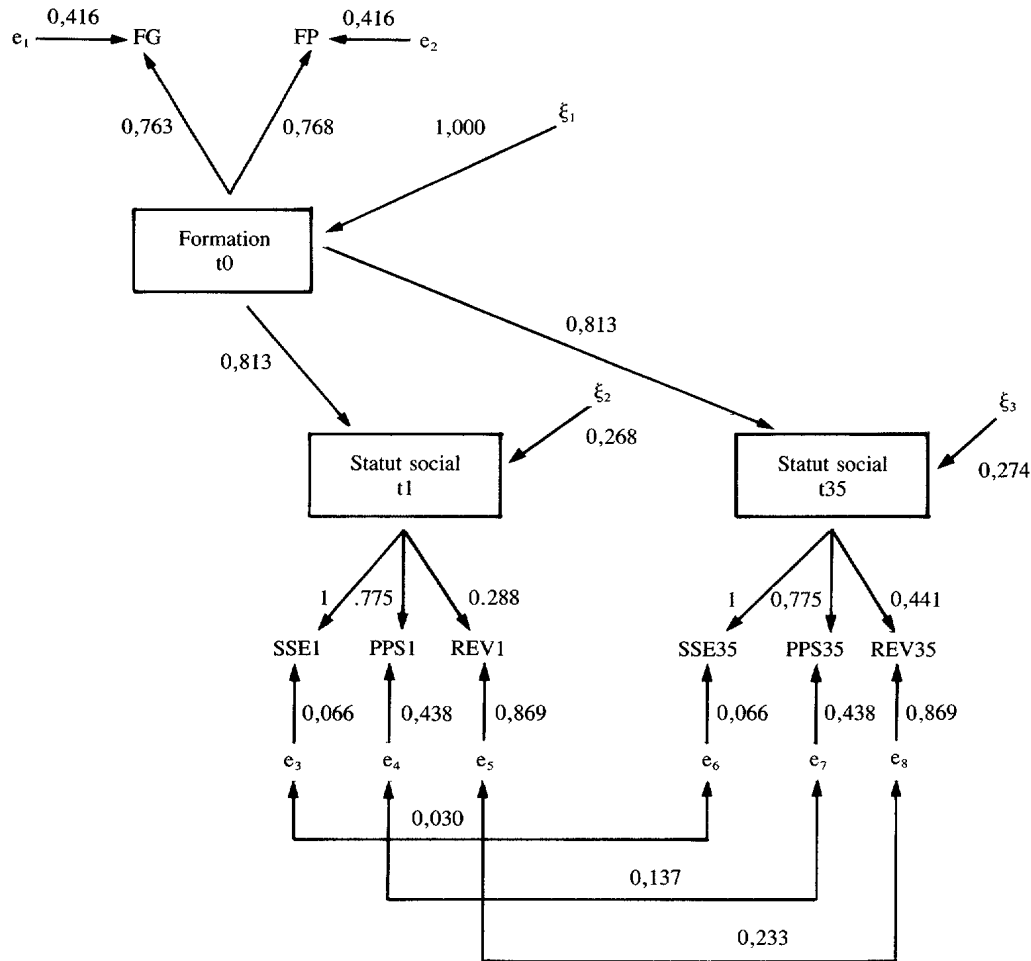
Le nombre de degrés de liberté est un indicateur du nombre d'informations nécessaires pour reproduire les données: plus ce nombre est élevé, moins le nombre d'informations à interpréter est grand. Mais, les variations des nombres de degrés de liberté sont ici insuffisants pour justifier le choix d'un modèle plutôt qu'un autre, sans considérer en même temps la qualité de l'information ajoutée ou retranchée.

Les figures 7 à 9 illustrent les relations importantes des modèles, soit celles qui peuvent être interprétées en fonction des considérations théoriques qui ont présidé à leur élaboration. Les relations qui ont été introduites de façon *ad hoc* dans les modèles pour atteindre un seuil de signification acceptable sont inclus en bas des figures. Ces relations impliquent des corrélations entre erreurs qui sont difficilement interprétables. Plus elles sont nombreuses, moins le modèle est économique et plus il augmente en complexité. Selon ce critère, le modèle de causalité différée avec ses neuf corrélations entre erreurs pourrait être exclu puisque le modèle de relations causales contemporaines n'en compte que 7.

Enfin, un critère quantitatif peut fournir une autre indication sur la capacité des modèles de reproduire la matrice de corrélation du Tableau 1. Les statistiques  $X^2$ , évaluées sur les trois modèles,

FIGURE 7

Résultats de l'étude de la matrice de corrélations  
du Tableau 1 : modèle du concept latent unique



Autres corrélations entre erreurs :

$e_3e_8$  : 0,052

$e_4e_2$  : 0,117

$e_4e_5$  : 0,051

$e_4e_8$  : 0,099

$e_7e_2$  : 0,066

$e_8e_2$  : 0,113

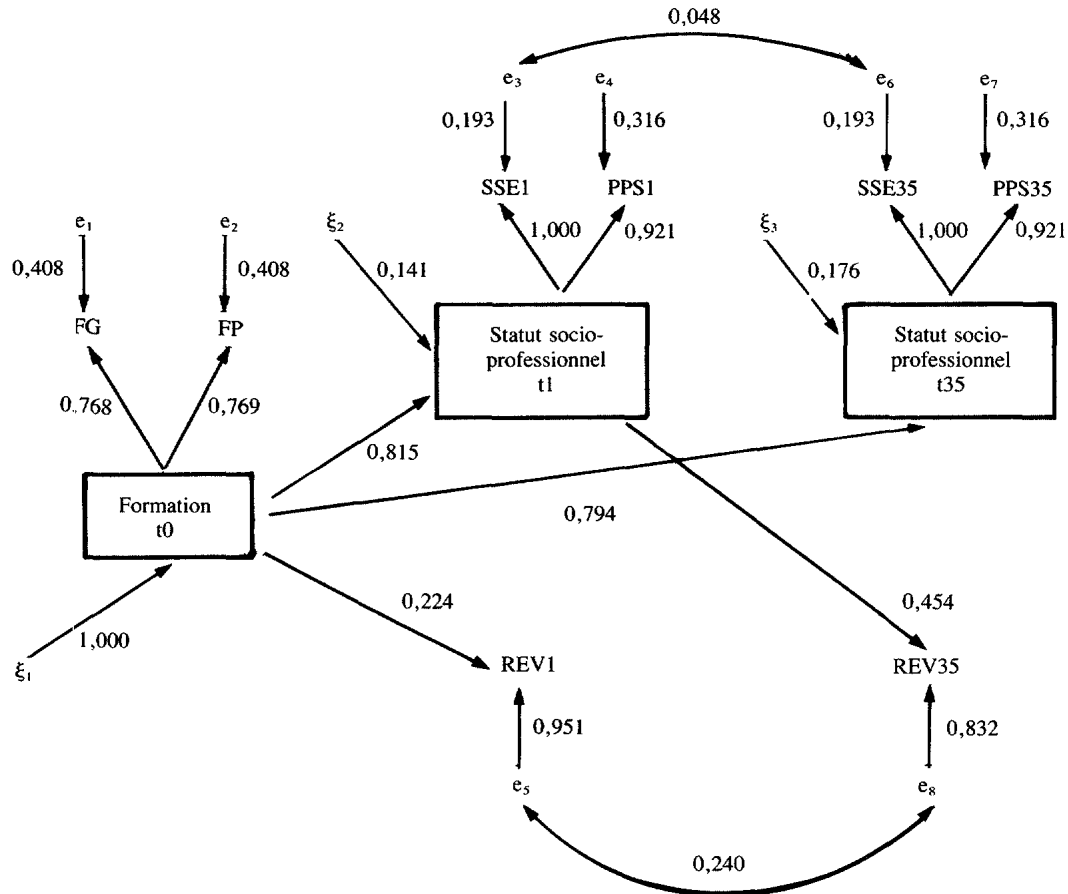
$e_8e_7$  : 0,099

$X^2 = 38,26$  avec 13 degrés de liberté.

en excluant les relations *ad hoc*, donnent une indication de la capacité des relations théoriquement significatives de reproduire les données. Les résultats d'analyse obtenus depuis les relations illustrées aux figures 7 à 9 sont les suivantes : 1) pour le concept latent unique, le  $X^2$  est de 164,87 avec 20 degrés de liberté, 2) pour le modèle de causalité différée, le  $X^2$  est de 354,48 avec 20 degrés de liberté et 3) pour le modèle de causalité contemporaine, le  $X^2$  est de 125,07 avec 18 degrés de liberté. Il est donc clair que les relations *ad hoc* contribuent beaucoup plus au modèle de causalité différée qu'aux deux autres modèles. Selon ce critère, le modèle de causalité différée apparaît peu crédible, même s'il est celui qui traduit l'idée habituelle que les chercheurs se font de la causalité. Il y a donc quelques raisons de choisir le modèle de causalité contemporaine. Mais ce choix ne s'impose que parce que des corrélations *ad hoc* entre erreurs ont été introduites dans le modèle.

FIGURE 8

Résultats de l'étude de la matrice de corrélations du Tableau 1 : modèle de relations causales différées entre le statut social et le revenu.



Autres corrélations entre erreurs:

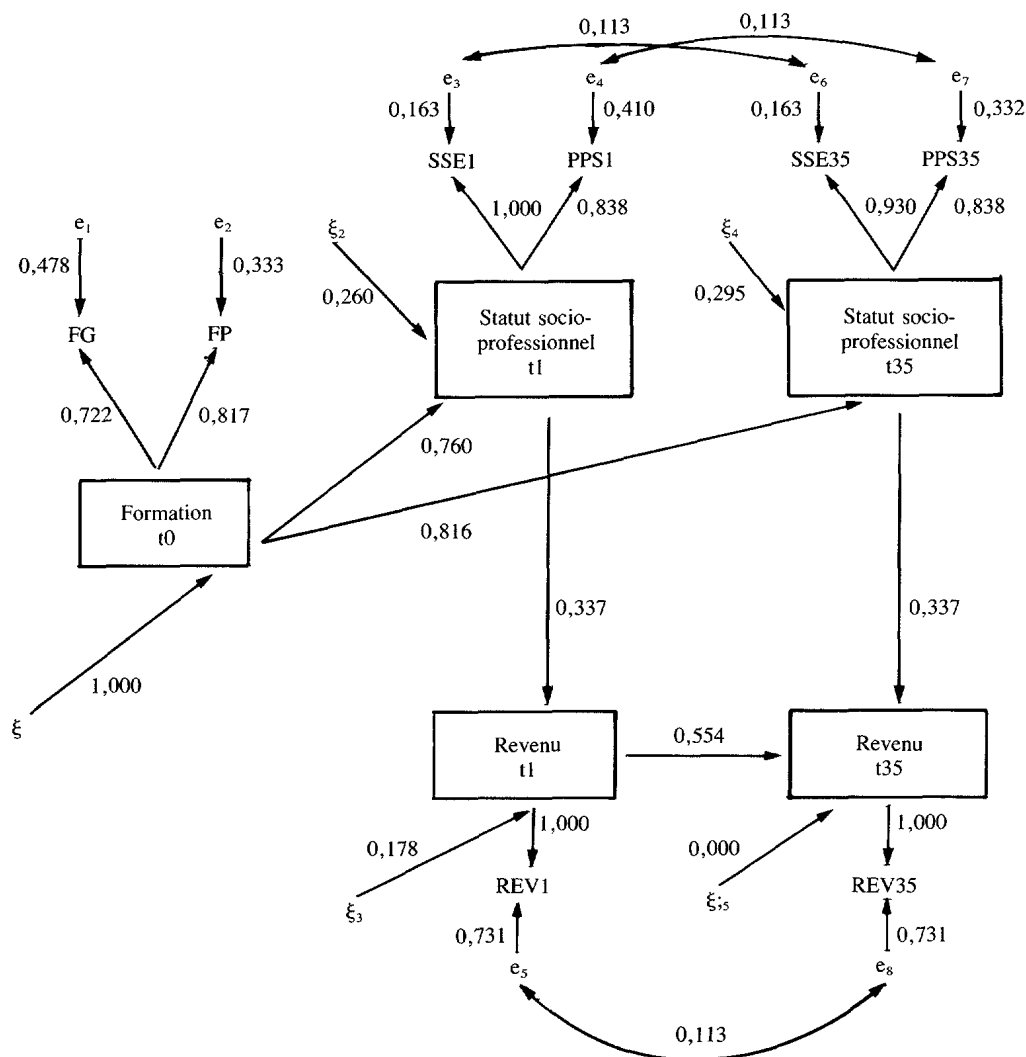
|            |          |            |          |
|------------|----------|------------|----------|
| $e_1e_4$ : | $-0,132$ | $e_7e_3$ : | $-0,031$ |
| $e_1e_7$ : | $-0,038$ | $e_5e_4$ : | $0,096$  |
| $e_2e_8$ : | $0,101$  | $e_8e_4$ : | $-0,121$ |
| $e_5e_3$ : | $0,083$  | $e_8e_5$ : | $0,240$  |
| $e_6e_3$ : | $0,048$  | $e_8e_6$ : | $0,111$  |
|            |          | $e_8e_7$ : | $0,149$  |

$X^2 = 5,33$  avec 10 degrés de liberté.

Sans elles, le modèle du concept latent unique apparaît comme un bon concurrent. Cependant, si les trois modèles proposent des explications très différentes des rapports entre les dimensions du statut social, ils ont tous en commun l'absence de lien direct entre le statut socio-professionnel ou le statut social en début de carrière et à 35 ans. En effet, les statuts à ces deux moments de la carrière sont entièrement expliqués par la formation reçue. Le statut attribué à un individu est donc un élément très stable dans sa carrière, puisqu'il peut être prédit par une caractéristique acquise avant l'entrée sur le marché du travail. Cette absence de relations directes entre les statuts est probablement le résultat le plus intéressant de cette analyse, résultat qui est reproduit par trois modèles qui conceptualisent pourtant de façon différente les rapports entre deux dimensions du statut social.

FIGURE 9

Résultats de l'étude de la matrice de corrélations  
du Tableau 1: modèle de relations causales contemporaines  
entre le statut social et le revenu



Autres corrélations entre erreurs.

$e_3e_1$ : 0,078

$e_7e_3$ : 0,039

$e_4e_2$ : 0,062

$e_8e_2$ : 0,049

$e_5e_4$ : 0,035

$e_8e_7$ : 0,044

$e_6e_1$ : 0,061

$X^2 = ,59$  avec 11 degrés de liberté.

#### 4. DISCUSSION

Les buts poursuivis par l'analyse de l'évolution du statut social en cours de carrière étaient: 1) de déterminer s'il y a autonomie entre trois dimensions du statut social et 2) d'examiner la structure du changement des statuts sociaux en cours de carrière à l'aide d'hypothèses sur les rapports de causalité entre les trois dimensions. Aucun de ces deux objectifs n'a été atteint par l'étude longitudinale des statuts sociaux: 1) En effet la tridimensionnalité de la notion de statut

social a été rejetée non pas sur la base des résultats de l'étude longitudinale mais bien parce que la procédure LISREL a identifié des cas de Heywood et une matrice de variance-covariance des erreurs non positive définie. Ces résultats peuvent être obtenus en utilisant des données latitudinales. 2) Il est apparu que le modèle du concept latent unique ne réussit pas à atteindre un seuil de signification acceptable, cependant, les deux autres modèles atteignent ce seuil au prix de l'introduction de corrélations *ad hoc* entre erreurs de mesure, difficile à interpréter. L'analyse semble cependant favoriser le modèle de causalité contemporaine. Ce résultat est cependant fragile. Il demande reconfirmation avec d'autres données pour être vraiment acceptable. Mais 3) chacun des trois modèles rejettent la relation directe entre le statut socio-professionnel en début de carrière et le statut correspondant à 35 ans. Ce résultat n'a pu être acquis qu'avec l'utilisation de données temporelles.

Ces résultats ne discréditent en aucune façon l'utilisation de données longitudinales. Elles illustrent quelques-unes des difficultés auxquelles le chercheur est confronté. Mais, tirer de cette étude cette unique conclusion ne représente pas adéquatement ce qui a été accompli ici. Le fait de disposer de données à deux moments de la carrière de travailleurs a permis d'élaborer des modèles théoriques qui n'auraient pu être vérifiés sur des données latitudinales. L'étude longitudinale nous a certes laissés dans le vague, mais des données latitudinales nous auraient laissés une impression de certitude. Les cas de Heywood se seraient produits, mais il aurait été impossible de distinguer entre les modèles modifiés de causalité contemporaine et différée. Qui plus est, les rapports entre les statuts sociaux à deux moments de la carrière n'auraient pu être éclaircis. L'analyse longitudinale est utile dans la mesure où elle accroît la capacité de concevoir des modèles concurrents d'explication de la dynamique sociale. Rien ne garantit que dans un cas particulier elle permettra de choisir entre ces explications concurrentes. Enfin, même si les résultats obtenus sont clairs, l'analyse longitudinale ne fait qu'accroître la probabilité qu'une relation de causalité d'une certaine direction est valide. Elle ne la confirme jamais.

L'analyse longitudinale se base sur l'observation d'un système à différents moments dans le temps. Elle permet donc d'observer le changement, non pas les rapports de causalité. Ces rapports sont toujours surimposés sur des observations pour mieux comprendre la dynamique du système qu'ils définissent d'un même mouvement. Le raisonnement causal se propose donc d'ordonner le champ d'observation. Cook et Campbell (1979) ont noté que la notion de cause était plus centrale en sciences sociales qu'en sciences naturelles, où elle joue un rôle instrumental mineur. Par exemple, Feynman (1980: 57-58) a présenté trois façons différentes de comprendre le phénomène de la gravité; l'une d'elle seulement appartient à la causalité, les autres étant plutôt structurelles, définissant des rapports mathématiques entre des éléments. En sciences sociales, le raisonnement causal semble faciliter la formulation de modèles explicatifs du changement. Mais, dans la mesure où l'analyse longitudinale est employée pour étudier le changement plutôt que la causalité, peut-être devrions-nous la subordonner à l'étude de l'objet propre de l'analyse longitudinale, le changement, pour nous laisser l'occasion, quand elle se présentera, de formuler autrement nos hypothèses.

#### RÉFÉRENCES

- BLISHEN, B. R. et H. A. McROBERTS, «A Revised Socio-Economic Index for Occupations», *Canadian Review of Sociology and Anthropology*, 13, 1976, 71-80.
- CRANO, W. D., D. A. KENNY, D. T. CAMPBELL, «Does Intelligence Cause Achievement?: A Cross-Lagged Panel Analysis», *Journal of Educational Psychology*, 63, 1972, 258-275.
- COOK, T. D. et D. T. CAMPBELL, *Quasi-Experimentation: Design & Analysis Issues for Field Settings*, Chicago, Rand McNally, 1974.
- DUNCAN, O. D., *Introduction to Structural Equation Models*, New York, Academic Press, 1975.
- FEYNMAN R., *la Nature de la physique*, Paris, Éditions du Seuil, 1980.
- HAUSER, R. M. et D. L. FEATHERMAN, *Opportunity and Change*, New York, Academic Press, 1978.
- JÖRESKOG, K. G. et D. SÖRBOM, *LISREL VI: Analysis of Linear Structural Relationship by the Method of Maximum Likelihood*, Mooresville, Scientific Software, 1984.
- KENNY, D. A., «Cross-Lagged and Synchronous Common Factors in Panel Data», dans A. S. Goldberger et O. D. Duncan (édit.), *Structural Equation Models in the Social Sciences*, New York, Seminar Press, 1973, pp. 153-167.
- KENNY, D. A., «Cross-Lagged Panel Correlation: A Test for Spuriousness», *Psychological Bulletin*, 82, 1975, 887-903.
- PAWSON, R., «Desperate Measure», *British Journal of Sociology*, 33, 1982, 35-63.
- PELZ, D. C. et F. M. ANDREW, «Detecting Causal Priorities in Panel Study Data», *American Sociological Review*, 29, 1964, 836-848.
- RENAUD, J., P. BERNARD et M. BERTHIAUME, «Éducation, qualification professionnelle et carrière au Québec», *Sociologie et sociétés*, 12, 1980, 23-52.

RINDSKOPF, D., «Using Inequality Constraints to Prevent Heywood Cases: The LISREL Parameterization», *Data Analyst*, 1, 1983.

ROZELLE, R. M. et D. T. CAMPBELL, «More Plausible Rival Hypothesis in the Cross-Lagged Panel Correlation Technique», *Psychological Bulletin*, 71, 1969, 74-80.

### RÉSUMÉ

Établir des relations causales entre variables a été l'une des opérations les plus courantes en sociologie. Mais après qu'il fût devenu évident que l'analyse transversale ne permettait pas de valider les relations causales, le recours à l'analyse longitudinale est apparu comme une voie de salut à plusieurs. Là encore, il a été établi que l'analyse longitudinale permettait de restreindre en partie les limites à notre capacité de prouver des relations causales, sans cependant les éliminer. Ici, trois modèles de relations entre variables sont définis et leur capacité à générer des données observées par sondage, vérifiée. Il est clair que malgré qu'une procédure formelle d'examen de ces modèles soit utilisée, le choix d'un, parmi ces trois modèles qui se contredisent, reste difficile. L'utilisation de données longitudinales n'est pas ici discrédité, ce texte entend souligner quelques-unes des difficultés d'application des méthodes d'analyse de ces types de données.

### SUMMARY

Establishing causal relationships between variables has been one of the most common operations in sociology. However, after it became evident that transversal analysis did not make it possible to validate causal relationships, recourse to longitudinal analysis seemed for many to be a way out. There too, however, it was established that longitudinal analysis could only partially remove the limits on our capacity to prove causal relationships, without eliminating them altogether. Three models for relationships between variables are defined here and their capacity for generating survey data verified. Even though a formal procedure for investigating these models is used, it is clear that choosing one of the three contradictory models remains difficult. The intention of this paper is not to discredit the use of longitudinal data, but to emphasize a few of the difficulties in applying different methods of analysis to these types of data.

### RESUMEN

Establecer relaciones causales entre variables ha sido una de las operaciones más corrientes en sociología. Cuando fue evidente que el análisis transversal no permitía la validación de las relaciones causales, apareció el análisis longitudinal como una tabla de salvación para muchos. Aún allí, se estableció que el análisis longitudinal permitía restringir en parte los límites de nuestra capacidad de probar relaciones causales pero sin eliminarlos. En este artículo se definen tres modelos de relaciones entre variables y se verifica su capacidad para generar datos observados en encuestas. Está claro que a pesar de que se utiliza un procedimiento formal de examen de esos modelos, la elección de uno de estos tres modelos que se contradicen, es siempre difícil. La utilización de datos longitudinales no es desacreditada, este texto trata de subrayar las dificultades de aplicación de los métodos de análisis de estos tipos de datos.