

Recherches sociographiques



La structure spatiale des caractéristiques socio-économiques de Montréal, 1961

Jean-Pierre Thouez

Volume 14, Number 1, 1973

URI: <https://id.erudit.org/iderudit/055603ar>

DOI: <https://doi.org/10.7202/055603ar>

[See table of contents](#)

Article abstract

La structure spatiale des caractéristiques socio-économiques de Montréal, 1961

Publisher(s)

Département de sociologie, Faculté des sciences sociales, Université Laval

ISSN

0034-1282 (print)

1705-6225 (digital)

[Explore this journal](#)

Cite this article

Thouez, J.-P. (1973). La structure spatiale des caractéristiques socio-économiques de Montréal, 1961. *Recherches sociographiques*, 14(1), 81–115. <https://doi.org/10.7202/055603ar>

Tous droits réservés © Recherches sociographiques, Université Laval, 1973

This document is protected by copyright law. Use of the services of Érudit (including reproduction) is subject to its terms and conditions, which can be viewed online.

<https://apropos.erudit.org/en/users/policy-on-use/>

Érudit

This article is disseminated and preserved by Érudit.

Érudit is a non-profit inter-university consortium of the Université de Montréal, Université Laval, and the Université du Québec à Montréal. Its mission is to promote and disseminate research.

<https://www.erudit.org/en/>

LA STRUCTURE SPATIALE DES CARACTÉRISTIQUES SOCIO-ÉCONOMIQUES DE MONTRÉAL, 1961 *

L'approche multidisciplinaire a considérablement favorisé le développement des études portant sur l'espace social.¹ La plupart d'entre elles ont pour objet la reconnaissance de dimensions socio-économiques de la population urbaine.

Bien que l'on puisse remonter à la fin du XIX^e siècle (citons en particulier l'ouvrage de C. Booth sur Londres), ce n'est qu'à partir de 1920 que l'on voit proliférer les études descriptives sous la direction de R.E. Park et W. Burgess qui orienteront l'école dite « de Chicago ».²

Cette période marquée par les recherches de H.W. Zorbaugh, de L. Wirth, N. Anderson, F.M. Thrasher, D. St. Clair et H.R. Cayton souligne le développement de l'écologie humaine. La plupart des travaux des années cinquante seront fondés sur ces idées, bien que leur approche théorique et pratique différera sur de nombreux points.³

Cette orientation marque l'intérêt que l'on attache aux travaux de W. Shevky, en collaboration avec M. Williams puis avec W. Bell, sur

* Cette étude a été réalisée comme exercice de séminaire au Département de géographie de l'Université de Sherbrooke (Québec). Les cinq étudiants : J. Castonguay, R. Deschênes, P. Gagnon, A. Masse et B. Rivard ont sélectionné les variables, résolu les différentes opérations et interprété l'importance et la distribution des dimensions. Les programmes utilisés sont un programme BMD X72, un programme de groupement hiérarchique fondé sur l'algorithme de Ward (1963), enfin un programme d'analyse discriminatoire établi par Casetti (1965), opérationnels à l'Université de Sherbrooke pour le système IBM 360.

¹ Voir : A. BUTTIMER, « Social Space in Interdisciplinary Perspective », *The Geographical Review*, 1960, pp. 418-426.

² F. W. BURGESS et D. BOGUE, (eds), *Contributions to Urban Sociology*, University of Chicago Press, 1964.

³ Voir : W. C. KAUFMAN, *Social Area Analysis. An explanation of Theory, Methodology and Techniques with Statistical Tests of Revised Procedures*, San Francisco and Chicago, 1950, Ph. D. non publié, Northwestern University, 1961.

Los Angeles et San Francisco. À partir d'un certain nombre de postulats sur les conséquences du changement dans la société, ils définissent trois indices : rang social (statut économique), urbanisation (statut familial), ségrégation (statut ethnique) qui doivent décrire la façon dont les populations urbaines se différencient à l'intérieur d'une société industrielle dynamique ; les trois dimensions sont fondées sur un certain nombre de variables élaborées à l'échelle du secteur de recensement (*census tract*).

Indice 1 : statut socio-économique (rang social)

- profession
- éducation
- location (prix du loyer).

Indice 2 : urbanisation

- fécondité
- femmes qui travaillent
- logements unifamiliaux.

Indice 3 : ségrégation (composition ethnique).

L'analyse permet donc la classification des secteurs de recensement en aires dites sociales (*social areas*).

Cette analyse des « aires sociales » a soulevé un certain nombre de controverses.⁴ La première concerne l'homogénéité des unités de recensement ; la seconde, le support théorique et la troisième, l'approche empirique, c'est-à-dire la méthode de construction des indices. Quoi qu'il en soit, un certain nombre d'études se sont largement inspirées de ces travaux dans leur analyse des « aires sociales ». W.C. Kaufman, à l'aide de l'analyse factorielle, revise l'ensemble des variables supportant les indices pour San Francisco et Chicago.⁵ Van Arsdol, Camilleri, Schmid étendent cette méthode à l'analyse de dix villes américaines. Cependant, dans la mesure où

⁴ La plus complète est celle de A. H. HAWLEY et O. D. DUNCAN, « Social Area Analysis : A critical Appraisal », *Land Economics*, XXXIII, 1957, pp. 337-45. Voir aussi : C. M. TULORET, « Hawley and Duncan on Social Area Analysis : A comment », *Id.*, XXXIV, 1958, pp. 182-184 ; W. BELL et S. GREET, « Social Area Analysis and its Cities », *Pacific Sociological Review*, V, 1962, pp. 3-9.

Citons également la controverse DUNCAN et BELL dans le *American Journal of Sociology* (LXI, 1955, pp. 84-85, 260-261 et 261-62).

Enfin les « Communications » de E. GREENWOOD et E. G. ERICKJEN dans *American Sociological Review* (XV, 1950, pp. 108-109 et XX, 1955, pp. 497-98), suivies de : R. W. BUECHLEY, « Review of Social Area Analysis », *Journal of the American Statistical Association*, LI, 1956, pp. 195-97.

⁵ W. C. KAUFMAN, *op. cit.*

deux villes ne se conforment pas au modèle,⁶ les auteurs semblent admettre qu'une approche inductive serait tout aussi souhaitable,⁷ corroborant les travaux de R.C. Tryon sur San Francisco qui avait utilisé la méthode de groupements (*cluster-analysis*) pour réduire trente-trois variables en trois groupes possédant une structure similaire aux trois indices originaux.⁸ Parmi les études réalisées sur les villes non américaines et utilisant le modèle de Shevky et Bell, on trouve celles de McElrath sur Rome,⁹ D.T. Herbert sur Newcastle,¹⁰ G. Gagnon sur Québec¹¹ et J.L. Abu-Lughod sur Le Caire,¹² cette dernière par l'analyse factorielle. D.T. Herbert souligne :

« The Social Area Map is meaningful and accurately differentiates the urban structure of Newcastle thus fulfilling one claim which may be made of the approach in that it summarizes several essential aspects of the social geography of urban area. That social area analysis is a useful comparative tool has perhaps been demonstrated by the comparisons which had been made between the results of this study and those which have been obtained from other parts of the world. The social area map and social space diagram are also held as valuable frames of reference in the context of which samples of selected parts of the urban area may be made. » (P. 55.)

L'utilité d'une telle analyse provient en particulier du fait qu'elle s'intègre dans un certain nombre de modèles descriptifs.¹³ Les indices qui mesurent le statut socio-économique des individus ou des groupes tendent à se distribuer par secteurs ; ceux qui s'attachent aux caractéristiques de la famille et à l'âge de la population varient principalement par zones concentriques ; enfin la structure de l'utilisation du sol s'articule autour de plusieurs *noyaux multiples*. Certes, là encore, les controverses sont

⁶ M. Van ARSDOL, S.F. CAMILLERI et C.F. SCHMID, « An Application of the Shevky Social Area Indexes to a Model of Urban Society », *Social Forces*, XXXVII, 1958, pp. 26-33 ; « The Generality of the Shevky Social Area Indexes », *American Sociological Review*, XXIII, 1958, pp. 227-284.

⁷ M. Van ARSDOL, J.F. CAMILLERI et C.F. SCHMID, « An Investigation of the Utility of Urban Typology », *Pacific Sociological Review*, IV, 1961, pp. 26-32 ; « Further Comments on the Utility of Urban Typology », *Id.*, V, 1962, pp. 9-13.

⁸ R. C. TRYON, *Identification of Social Areas by Cluster Analysis. A General Method with an Application to the San Francisco Bay Region*, Berkeley, Presses de l'Université de Californie, 1955.

⁹ D.C. McELRATH, « The Social Areas of Rome. A Comparative Analysis », *American Sociological Review*, XVII, 1962, pp. 376-91.

¹⁰ D.T. HERBERT, « Social Area Analysis : a British Study », *Urban Studies*, IX, 1967, pp. 41-60.

¹¹ G. GAGNON, « Les zones sociales de l'agglomération de Québec », *Recherches sociographiques*, I, 3, 1960, pp. 255-67.

¹² J.L. ABU-LUGHOD, « Testing the Theory of Social Area Analysis. The Ecology of Caire, Egypt », *American Sociological Review*, 1969, pp. 198-212.

¹³ P.H. REES, « The Factorial Ecology of Metropolitan Chicago », Thèse de maîtrise, Université de Chicago, 1968 ; chap. 10, in B.L.J. BERRY et HORTON, (eds), *Geographic Perspectives on Urban Systems*, Prentice Hall, 1970.

nombreuses, toutes les localisations ne se conforment pas nécessairement aux modèles descriptifs.¹⁴

Le modèle des *zones concentriques*, hérité de l'apport théorique allemand, en particulier des travaux de Von Thünen, doit être compris à la fois comme un modèle statique et dynamique. En ce sens, il doit être attaché aux modèles de croissance urbaine, principalement fondés sur les notions de forces centripètes et centrifuges, sur la formulation mathématique des mécanismes de densité, enfin sur les théories du comportement.

Celui des *secteurs radiants* de Hoyt, inspiré du modèle de Hurd, doit être révisé à la lumière des travaux de Firel,¹⁵ de Gans,¹⁶ de Peterson.¹⁷ Enfin le modèle élaboré par C.D. Harris et E.L. Ullman suggère l'existence de plusieurs schémas écologiques possibles, bien que cette approche flexible soit controversée par B.J.L. Berry et par W. Simmons.¹⁸ La généralité ou les limites des modèles descriptifs soulignent la nécessité d'un renouvellement théorique ou conceptuel. À l'heure actuelle les études les plus récentes « expérimentent » à l'aide de l'analyse factorielle, les dimensions socio-économiques sans considérer les *a priori* des modèles précédents, ce qui ne signifie pas que l'on puisse éliminer les hypothèses de travail relatives aux dimensions du problème envisagé.¹⁹

Cette technique en particulier est propre à fournir une description condensée d'un ensemble de variables associées. Elle peut aussi permettre

¹⁴ T.R. ANDERSON et EGELAND, « Spatial Aspects of Social Area Analysis », *American Sociological Review*, XXVI, 1961, pp. 392-99.

P.H. REES, *op. cit.*

B.J.L. BERRY et P.H. REES, « The Factorial Ecology of Calcutta », *The American Journal of Sociology*, 1969.

J. B. RACINE, « Le modèle urbain américain », *Annales de géographie*, 490, juillet-août, 1971, pp. 397-427.

¹⁵ W. FIREY, *Land Use in Central Boston*, Cambridge, Presses de l'Université d'Harvard, 1947 ; « Sentiment and Symbolism as Ecological Variables », *American Sociological Review*, X, 1945, pp. 140-48. (Rep. in G.A. THEODORSON, (ed.), *Studies in Human Ecology*, Evanston, 1961.)

¹⁶ H. GANS, *The Urban Villagers*, Glencoe, Free Press, 1962 ; *The Levittowners*, New York, Pantheon Books, 1967.

¹⁷ G.L. PETERSON, « A Model of Preference. Quantitative Analysis of the Perception of the Social Appearance of Residential Neighbourhoods », *Journal of Regional Science*, VII, 1967, pp. 19-31.

¹⁸ B.J.L. BERRY, « Cities as Systems within System of Cities », in J. FREIDMANN et N. ALONSO, (eds), *Regional Development and Planning*, Presses du M.I.T., 1964 ; « Internal Structure of the City », *Law and Contemporary Problems*, XXX, 1965, pp. 111-119.

J.W. SIMMONS, « Descriptive Models of Urban Use », *The Canadian Geographer*, IX, 1965, pp. 170-174.

¹⁹ Carl JANSON GUNNAR, « The Spatial Structure of Newark, New Jersey, I. The Central City », *Acta Sociologica*, XI, 1968, pp. 145-69.

J. B. RACINE, *op. cit.*, pp. 397-427.

R.A. MURDIE, *Factorial Ecology of Metropolitan Toronto*, Département de géographie, Université de Chicago, 1969, pp. 26-37.

D.W.G. TIMMS, *The Urban Mosaic*, Presses de l'Université de Cambridge, 1971, pp. 47-54.

après une interprétation des résultats de découvrir des relations ou des arrangements complexes tant parmi les variables analysées que parmi les individus notés dans ces variables.

Si l'on souhaite factoriser les variables (analyse dite du R mode), on aboutit en général à remplacer les variables originales par des facteurs que ces variables ont soit en commun, soit en groupe, soit d'une manière spécifique. En d'autres termes, on détermine un nombre limité de dimensions ou facteurs qui rendent compte de la plus grande part possible de la variance des variables de la matrice d'informations.²⁰

Parmi les facteurs les plus souvent cités, nous trouvons le statut socio-économique ou rang social (*social rank*) et le statut familial, du moins un certain nombre de caractéristiques attachées au cycle familial; enfin les caractéristiques liées à la composition ethnique de la population étudiée et des variables de mobilité de la population. Bien que ce schéma soit général, nous devons tenir compte des spécificités suivantes: aire d'étude, données d'entrée, technique employée, nombre de caractéristiques; poids de ces caractéristiques sur les facteurs; de la structure des pointages factoriels (*factor scores*).

L'analyse factorielle permet donc de déterminer le nombre et la nature de dimensions latentes fondamentales et indépendantes de l'espace social pour en tester cartographiquement et mathématiquement (grâce à des analyses de variance différentielle) les configurations spatiales des centres urbains²¹ ou en vue de l'analyse longitudinale.²² Les dimensions extraites permettent, à la lumière de caractéristiques envisagées, d'utiles comparaisons. Le but de cette étude est d'appliquer l'analyse en *composantes principales* aux

²⁰ H.H. HARMAN, *Modern Factor Analysis*, Presses de l'Université de Chicago, 1960. R.J. RUMMEL, « Understanding Factor Analysis », *Journal of Conflict Resolution*, 1967, pp. 444-480; *Applied Factor Analysis*, Evanston, Northwestern University Press, III, 1970.

²¹ J.K. HADDEN et E.F. BORGOTTA, Rand McNalley, Chicago, 1965. Étude complétée par K.J. JONES et W.C. JONES, « Toward a typology of American Cities », *Journal of Regional Science*, X, 1970, pp. 217-224.

C.A. MOSER et W. SCOTT, *British Towns. A Statistical Study of their Social and Economic Differences*, Londres, 1961.

M. ALDSKOGUIS, *Dimensions of Urban Structure in Sweden. An Exploratory Factor Analysis Study*, Forskin, Université d'Upsala, 1969.

P.P. GRIGER, M.R. DA SILVA GUIMARES et E.R. STEFFAN, « Application de la technique des composants principaux à l'étude des villes de la région du centre ouest du Brésil », *Cahiers de géographie de Québec*, 35, 1971, pp. 191-212.

J.P. THOUÉZ, « L'Analyse socio-économique et politique des villes moyennes de la rive sud du St-Laurent (Québec) », Département de géographie, Université de Sherbrooke. (À paraître.)

²² P.O. PEDERSON, « An Empirical Model of Urban Population Structure. A Factor Analytical Study of the Population Structure in Copenhagen. 1950 and 1960 », *Proceedings of the first Scandinavian-Polish Regional Science Seminar*, Varsovie, Polish Scientific Pub., 1967.

données socio-économiques sur la ville de Montréal²³ en essayant dans la mesure du possible de se référer aux études réalisées sur les métropoles canadiennes de Winnipeg,²⁴ de Toronto,²⁵ de Montréal.²⁶

I. LES COMPOSANTES PRINCIPALES

a) Méthodologie

Nous présenterons seulement l'essentiel de l'analyse en *composantes principales* ; d'excellents ouvrages sont apparus sur cette méthode. De plus notre discipline s'enrichit d'un nombre appréciable de travaux depuis quelques années.²⁷

Le processus tend essentiellement à réduire sous certaines conditions une matrice d'informations (au besoin normalisée) de dimension (n, m) avec n : unités d'observations et m : attributs ou variables. Comme l'indique le schéma ci-dessous, on élabore une matrice d'informations de dimension X (n, m) que l'on ramène à une échelle réduite de manière à rendre comparables les valeurs. La méthode de standardisation consiste en effet à remplacer chaque valeur par le rapport existant entre les deux écarts fondamentaux qui caractérisent sa position dans la distribution d'où elle est issue : l'écart à la moyenne de la distribution et l'écart-type de la distribution. Grâce à ce changement d'échelle toutes les distributions ont donc une moyenne égale à zéro et une variance égale à un. La réduction des valeurs à l'échelle réduite (et leur normalisation) permet alors de transformer cette matrice standardisée X (n, m) en une matrice de corrélation R (m, m) ou (variables, variables). Matrice symétrique dont la diagonale exprimant la corrélation de chaque variable avec elle-même est remplie de valeurs égales à l'unité. La décomposition de la matrice des corrélations Z consiste alors à extraire les dimensions ou facteurs expliquant le maximum de la variance

²³ Service d'urbanisme de Montréal, *Cahier des données. Occupation du sol, Zone métropolitaine*, 1964.

²⁴ T.G. NICHOLSON et M.H. YEATES, « The Ecological and Spatial Structure of the socio-economic Characteristics of Winnipeg, 1961 », *Canadian Review of Sociology and Anthropology*, VI, 1969, pp. 162-78.

²⁵ B.J.L. BERRY et R.A. MURDIE, *Prepared under Contract for the Urban Renewal Study of the Metropolitan Toronto Planning Board: Socio-economic Correlated of Housing Condition*, août 1965.

²⁶ B. GREER WOOTTEN, *Cross Sectorial Social Area Analysis, Montréal 1951-1961*, Département de géographie, Université McGill, 1968, pp. 9-46.

K.E. HAYNES, « Spatial Change in Urban Structure. Alternative Approaches. Ecological Dynamics », *Economic Geography*, XLII, pp. 324-335.

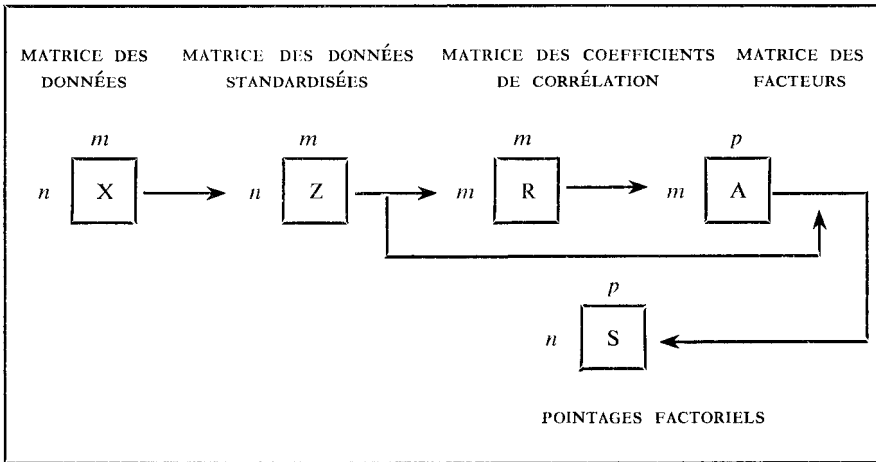
²⁷ Voir : D. THOMPSON, « Some comments on the Relevance of Multivariate Analysis to Geography, A Methodology Review », *Geographica Polonica*, XVIII, 1970, pp. 158-175.

P. LAVALLEE, M. MC CONNELL et R.G. BROWN, « Certain Aspects of the Expansion of Quantitative Methodology in American Geography », *Annals of the Association of American Geography*, LVII, 1967, pp. 423-436.

totale de l'information. En général, il y a autant de facteurs que de variables mais on ne retient que ceux qui expliquent la plus grande partie de la variation totale, c'est-à-dire ceux dont les valeurs propres sont égales ou supérieures à 1. Nous aboutissons aussi à deux matrices, la première dite matrice des facteurs ou matrice des saturations de chacune des variables avec chaque facteur A (m, p) et la seconde matrice des pointages factoriels S (n, p) exprimant la saturation de chacune des unités d'observation sur chacun des facteurs.

Il existe de nombreuses options offrant une telle transformation. Le programme en *composantes principales* a été choisi pour trois raisons : orthogonalité des composantes extraites par ordre, solution finale unique ; enfin, il ne requiert pas une distribution normale des données. Cependant, la matrice d'information nécessite la standardisation des données : on ramène toutes les moyennes à zéro et tous les écarts-types à l'unité.

Schéma d'analyse en composantes principales.



Le modèle en composantes principales peut être formulé de la façon suivante :

$$Z_j = a_{j1} F_1 + a_{j2} F_2 + \dots + a_{jn} F_n \quad (j = 1, 2, \dots, n)$$

dans lequel chacune des m variables observées s'exprime linéairement en p nouvelles composantes non corrélées F_1, F_2, \dots, F_n . Chaque composante extrait le maximum de la somme des variances des m variables. En général, on retient que les composantes les plus explicatives en terme de variance.

Rappelons que la matrice d'information de départ comporte les unités de recensement en lignes et les caractéristiques en colonnes (analyse dite « R. Mode » c'est-à-dire que les caractéristiques - colonnes seront factorialisées plutôt que les unités d'observation). De plus, l'approche en composantes principales suppose que nous choisissons d'inscrire l'unité sur la diagonale principale de la matrice de corrélation. La première étape a pour fondement la sélection des coefficients sur les premiers facteurs, de façon à ce que la somme des contributions de ce facteur dans la communauté totale soit maximum. Une fois ce facteur F dégagé, on analyse les résidus de la matrice de corrélation en vue d'élaborer un deuxième facteur (F_2) indépendant du premier, solutionné lui aussi sur sa contribution maximum à la communauté totale résiduelle. Le processus reste le même pour les facteurs subséquents.

Chaque facteur est constitué par la combinaison linéaire des caractéristiques et peut être interprété à partir de la note des variables (c'est-à-dire de la contribution de la variable) à l'intérieur de sa structure. Nous ne considérons que la note supérieure ou inférieure à 0.4 par ordre d'importance (sauf la matrice des coefficients de corrélation où nous indiquerons le signe du coefficient à l'intérieur des limites choisies). Les facteurs extraits étant « indépendants », ils peuvent s'additionner et être ainsi reliés aux modèles descriptifs décrits ci-dessus. Nous cartographions les pointages factoriels (*factor scores*) de façon à dégager les aires d'homogénéité en tenant compte de la contrainte géographique.²⁸

b) Analyse

Ce travail a pour aire d'étude la ville de Montréal et les municipalités de Westmount, Mont-Royal et Outremont (y compris les unités de recensement de Rivière-des-Prairies et Saint-Jean-de-Dieu). La population de cet espace en 1961 s'élevait à 1,268,009 habitants ; en 1956, elle était de 1,171,219, soit un accroissement approximatif de 1%.²⁹ Après avoir éliminé un certain nombre d'unités de recensement non représentatives, nous avons considéré 229 secteurs de recensement, soit approximativement

²⁸ La rotation des facteurs a été établie par la méthode dite du « Varimax » établie par M. F. KAISER, « Computer Program for Varimax Rotation in Factor Analysis », *Educational and Psychological Measurement*, XIX, 1959, pp. 413-420. Le but de ces rotations est de placer les axes des composants dans une position telle que chaque facteur absorbe une proportion de variance la plus importante possible (ce qui simplifie le nombre de caractéristiques par facteur général). Quant à la méthode utilisée pour le calcul de la matrice Z , elle fut établie en 1933 par HOTELLING, « Analysis of a Complex of Statistical Variables into Principal Components », *Journal of Educational Psychology*, XXIV, pp. 417-441 et 498-520.

²⁹ L'accroissement de l'aire métropolitaine est de 12.6% entre 1956 et 1961, de 43% entre 1951-1961. (*Recensement du Canada, 1961*, « Caractéristiques de la population et du logement par secteur de recensement », BFS 95-519.

N.W. LITHWICK, *Urban Canada Problems and Prospects*, Ottawa, 1970.)

98.30% de la population de l'aire délimitée ci-dessus.³⁰ Les cinquante-quatre variables choisies ont été déterminées à partir des différentes études publiées sur le sujet³¹ et de notre connaissance de Montréal. La liste des variables ainsi que leur moyenne, écart-type et communauté apparaît au tableau 1.

TABLEAU 1

VARIABLE	MOYENNE	ÉCART-TYPE	COMMUNAUTÉ
<i>Accroissement de la population</i>			
1. % d'accroissement de la population, 1956-61	5.39	30.34	.511
<i>Taille de la famille</i>			
2. % d'enfants de moins de 6 ans par famille	0.51	0.65	.863
3. % des familles de 0-2 enfants	77.06	7.81	.753
4. % des familles de 5 enfants et plus	5.78	4.077	.516
<i>Taille des ménages</i>			
5. % des ménages de 6 personnes et plus	14.864	6.00	.854
6. % des ménages de 1 famille	77.75	12.49	.756
<i>Structure par âge des familles</i>			
7. % des chefs de famille de moins de 25 ans	4.45	4.56	.102
8. % des chefs de famille de moins de 65 ans	81.90	4.26	.630

³⁰ La taille moyenne du secteur est de 4,949 habitants. Nous avons éliminé les secteurs de recensement inférieurs à l'écart type (2,700) soit les numéros 35, 38, 42, 43, 46, 51, 52, 53, 55, 57, 60, 63, 65, 79, 111, 125, 138, 156, 160, 206, 211, 226, pour la ville de Montréal; 260, pour la ville de Mont-Royal; 271, pour la cité d'Outremont et 274, pour celle de Westmount. Cependant, le fait de traiter les caractéristiques en pourcentage et en ratio élimine ce problème de taille entre unités de recensement.

³¹ On peut trouver une liste de références dans un numéro spécial de *Economic Geography*, XLVII, 2, juin 1971 : *Comparative Factorial Ecology*.

Voir aussi l'excellent article : J.A. KAHL et J.A. DAVIS, « A Comparison of Indexes of Socio-economic Status », *American Sociological Review*, XX, 1955, pp. 317-325.

TABLEAU 1 (suite)

VARIABLE	MOYENNE	ÉCART-TYPE	COMMUNAUTÉ
<i>Taux démographiques</i>			
9. Sexe ratio ^a	107.08	13.47	.612
10. Taux de fécondité ^b	4.64	3.88	.706
<i>Structure par âge de la population</i>			
11. % d'hommes de 0-15 ans	13.34	3.22	.899
12. % de femmes de 0-15 ans	12.94	3.05	.897
13. % d'hommes de plus de 65 ans ..	3.28	1.81	.299
14. % de femmes de plus de 65 ans..	4.27	2.09	.784
<i>Composition sociale</i>			
15. % de célibataires de plus de 15 ans	23.48	6.72	.961
16. % de gens mariés de plus de 15 ans	44.43	4.39	.878
<i>Lieu de naissance</i>			
17. % de gens nés à l'extérieur du Canada	16.45	14.55	.961
18. % d'immigrés, 1946-1961	10.058	11.35	.922
<i>Structure ethnique</i>			
19. % d'anglais	12.85	15.77	.588
20. % de français	67.38	27.56	.950
21. % de parlant français seulement	40.28	20.67	.917
22. % d'italiens	5.61	8.11	.377
23. % de juifs	5.54	12.53	.549
<i>Éducation</i>			
24. % de ceux qui ont quitté l'école sans aucun diplôme	60.33	14.65	.928
25. % de ceux qui ont suivi un enseignement secondaire	21.11	14.63	.853
26. % de ceux qui ont fréquenté l'école après le secondaire	0.31	0.23	.726
<i>Structure de l'emploi</i>			
27. % des femmes de plus de 15 ans qui travaillent	35.24	8.031	.887
28. % des femmes administrateurs, professions libérales, techniciennes	14.10	8.56	.708

a) Nombre de femmes pour mille hommes.

b) Nombre de naissances pour mille femmes de quinze à vingt-quatre ans.

TABLEAU 1 (suite)

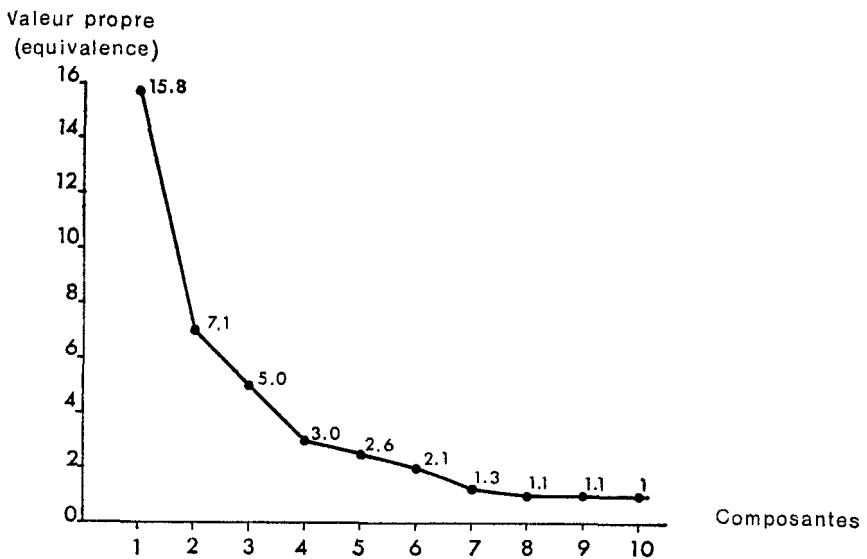
VARIABLE	MOYENNE	ÉCART-TYPE	COMMUNAUTÉ
29. % des femmes ouvrières et manœuvres	27.102	14.29	.809
30. % des hommes administrateurs, professions libérales, techniciens..	19.21	16.33	.933
31. % des hommes ouvriers et manœuvres	39.01	13.58	.904
32. % de la main-d'œuvre travaillant à son compte	8.05	6.3	.342
33. % de la main-d'œuvre en chômage	3.70	4.45	.088
<i>Revenus et salaires</i>			
34. % des revenus familiaux de moins de \$4,000.	37.67	15.00	.879
35. % des revenus familiaux de \$10,000 et plus	10.50	14.089	.874
36. Moyenne des traitements et salaires masculins	\$3,688.03	1101.13	.907
<i>Structure résidentielle</i>			
37. Logements ayant besoin de réparations importantes	1.788	6.189	.498
38. % de logements simples détachés	5.98	15.47	.842
39. % de logements par appartements	94.95	11.45	.572
40. % de logements occupés par le propriétaire	8.82	19.36	.894
41. % de logements occupés par un locataire	92.74	29.78	.881
42. % de logements surpeuplés.....	15.033	10.04	.705
43. Taux moyens de loyers	\$64.51	22.45	.716
<i>Indicateurs socio-économiques</i>			
44. % de congélateurs par famille..	0.36	0.07	.051
45. % de personnes par automobile..	9.43	3.83	.807
<i>Indicateurs de stabilité résidentielle</i>			
46. % des gens de 5 ans et plus ayant occupé leur logement depuis 6 ans et plus	37.73	10.90	.677
47. % des gens de 5 ans et plus ayant déménagé de la ville	41.90	11.68	.737

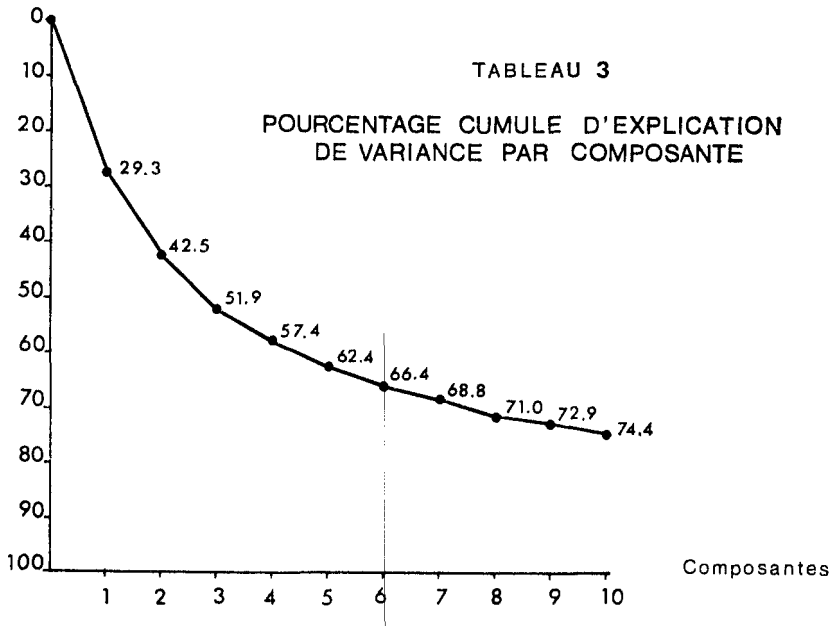
TABLEAU 1 (suite)

VARIABLE	MOYENNE	ÉCART-TYPE	COMMUNAUTÉ
48. % des gens de 5 ans et plus ayant déménagé de banlieue	3.46	7.47	.482
49. % des gens de 5 ans et plus ayant déménagé dans la même province	3.41	1.89	.246
<i>Utilisation du sol</i>			
50. % des terrains vacants par rapport à la surface nette	3.22	7.59	.632
51. % des parcs par rapport à la surface nette	5.63	10.17	.054
52. % des des bâtiments publics par rapport à la surface nette	9.93	11.66	.339
53. % des commerces par rapport à la surface nette	15.17	12.47	.466
54. % des industries par rapport à la surface nette	6.63	11.03	.142

TABLEAU 2

VALEUR PROPRE (ÉQUIVALENCE) X COMPOSANTES





Les tableaux 2 et 3 illustrent graphiquement le nombre de composantes qui sont significatives. Ils montrent qu'au-delà du quatrième facteur (57.46% d'explication cumulée) le gain de facteurs additionnels est faible (68.88% d'explication cumulée jusqu'au septième facteur).

Le tableau 4 indique successivement les saturations des notes sur les facteurs, la valeur propre (*eigenvalue*), c'est-à-dire la proportion absolue d'explication de variance pour chaque composante, et la proportion cumulée de variance totale pour chaque étape. Les cinquante-quatre valeurs propres analysent effectivement la variance totale ; cependant il apparaît que là encore, l'augmentation de variance expliquée après les quatre premiers facteurs est insignifiante. Il semble donc possible d'interpréter la structure socio-économique des unités de recensement de notre aire géographique à la lumière des quatre premiers facteurs plutôt que de trois.³² L'application

³² Nous avons tour à tour utilisé la matrice de corrélations avec l'unité sur la diagonale et le coefficient de corrélation multiple. La somme de variance expliquée est différente mais la structure reste la même. Il semble que le mieux serait d'utiliser la différence entre l'unité et le coefficient de corrélation multiple (CATELL) ou l'analyse en image (GUTTMAN).

Nous avons élaboré une seconde liste de variables après avoir éliminé 10 des 54 variables dont la communauté semblait trop faible. Cette nouvelle analyse extrait 6 facteurs expliquant 71.2% de la variance totale (les 5 premiers absorbant 66.8%).

TABLEAU 4

Analyse en composantes principales, 1961 : saturations.

VARIABLE	FACTEUR			
	I	II	III	IV
1				
2		-.788		.906
3			-.427	
4				.432
5	-.647			
6				
7				
8		.430		
9	.553	.514		
10				
11		-.799		
12		-.790		
13				
14		.639		
15		.952		
16	.407	-.622		
17			-.953	
18			-.930	
19	.481		-.421	
20			.892	
21	-.438		.798	
22	-.808			
23	.420		-.607	
24				
25	.719	.412		
26	.619			
27		.842		

TABLEAU 4 (suite)

Analyse en composantes principales, 1961 : saturations.

VARIABLE	FACTEUR			
	I	II	III	IV
28	.561	.575		
29	-.773			
30	.786			
31	-.731	-.412		
32				
33	.501			
34	-.847			
35	.618			
36	.783			
37				.627
38				
39				
40				
41				.683
42	-.580			.431
43	.709	-.427		
44				
45				
46				
47				
48				
49				
50				.662
51				
52		.447		
53	-.503			
54				

TABLEAU 4 (Suite)

	FACTEUR			
	I	II	III	IV
Valeur propre	15.82	7.17	5.02	3.01
Variance totale (%)	29.3	13.3	9.3	5.6
Pourcentage cumulé	29.30	42.59	51.89	57.46

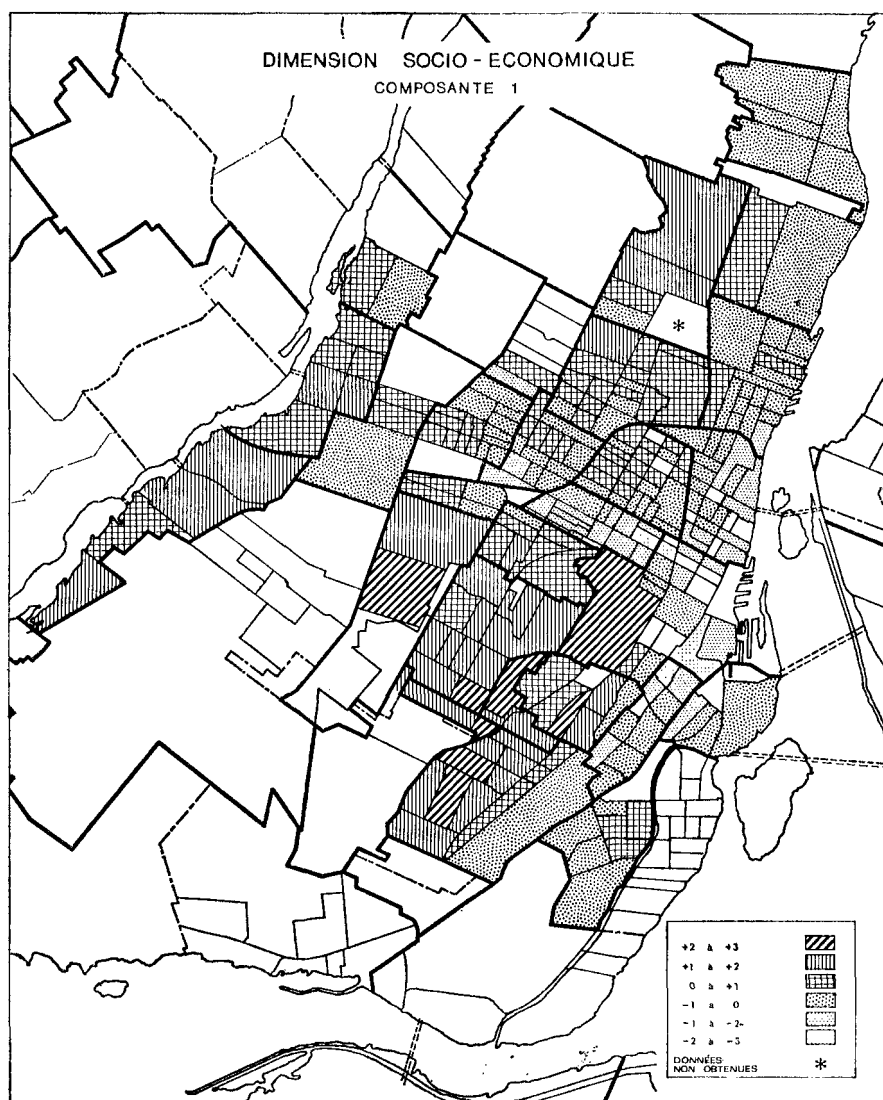
d'un test de signification au coefficient de corrélation nous a permis de constater qu'une saturation de 0.3 n'est pas significative, par contre 0.4 est significativement valable. Par conséquent, nous ne retiendrons que les saturations supérieures à ± 0.4 .

L'analyse du tableau 1 révèle que le nombre et le type de variables sont en partie satisfaisants. En effet, si l'on examine la communauté de chaque variable, c'est-à-dire la proportion de variance expliquée par la variable sur les quatre premiers facteurs ($h^2_j : a^2_{j_1} + a^2_{j_2} + a^2_{j_3} + a^2_{j_4}$), on constate que neuf d'entre elles ont une communauté supérieure à 90%, vingt-deux ont leur communauté supérieure à 80% (une communauté de 80% indique un coefficient de corrélation multiple de 0.89%). D'autres variables ont par contre une communauté faible inférieure à 40% (coefficient de corrélation multiple de 0.63%). Les saturations sont équivalentes aux coefficients de corrélation, c'est de leur interprétation que dépendra le titre du facteur. On peut s'occuper soit des variables qui ont la même direction dans le facteur étudié, soit étudier l'opposition combinée des saturations positives et négatives (dans la majorité des cas, on a affaire à un facteur bipolaire).

La *première composante* (29.3% de la variance totale) distingue les secteurs selon leur statut socio-économique. On note que les saturations positives se rapportent aux revenus élevés (.618); aux professions (administrateurs, professions libérales, techniciens : .619, femmes et .561, hommes; main-d'œuvre travaillant à son compte : .501), à l'instruction (secondaire : .719; supérieure : .619). Les associations de variables semblent traduire à la fois l'importance des groupes ethniques de statut élevé (anglais : .481; juifs : .420) et l'élément féminin (sexe ratio : .553). Les saturations négatives comprennent les bas revenus (— .847), la population unilingue française (— .438), les ouvriers et manœuvres (femmes : — .773, et hommes :

— .731), l'instruction élémentaire (— .808), les secteurs d'utilisation du sol mixtes (— .503).

La cartographie des pointages factoriels permet d'observer la position centrale des aires de statut inférieur. Les poids ont été divisés en six classes. Les notes positives de plus forte valeur indiquent un statut économique élevé et vice-versa.



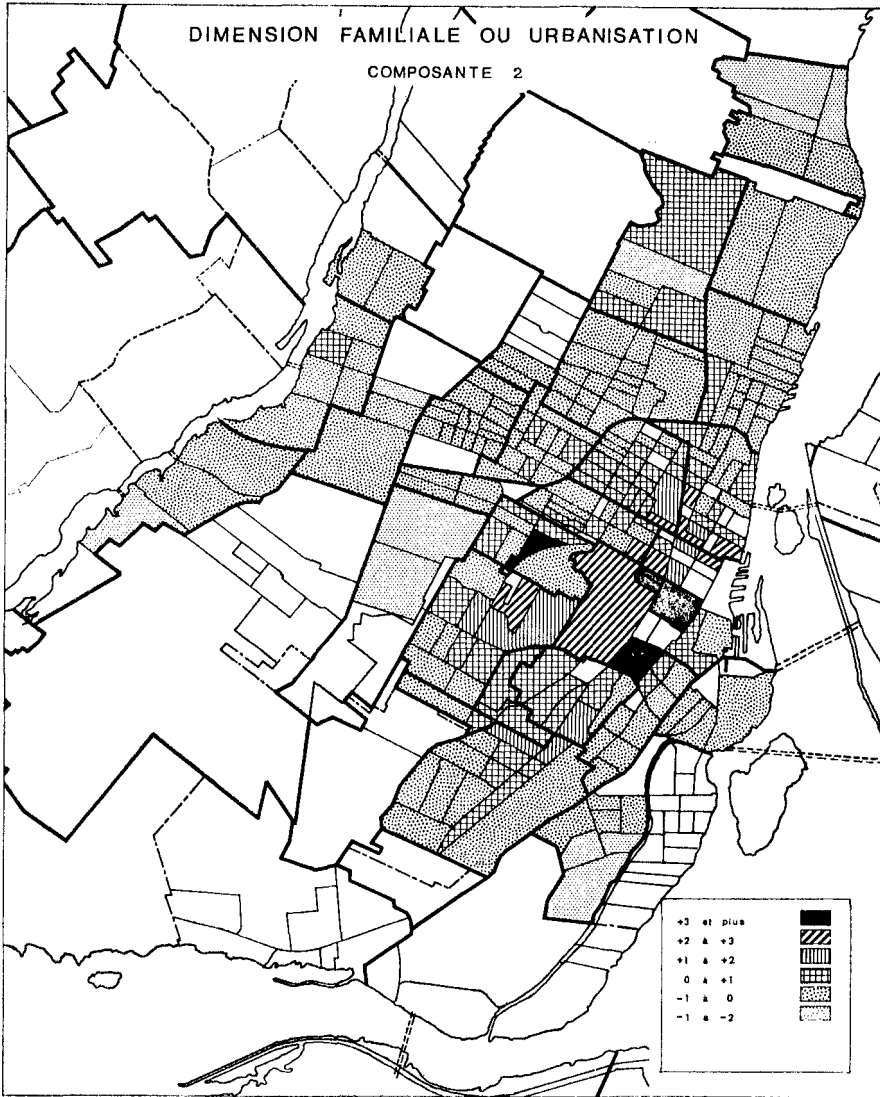
On peut dégager plusieurs zones géographiques présentant une configuration sectorielle. La première est fondée sur l'axe formé par le boulevard St-Laurent. La seconde et la troisième partent du centre-ville mais ont une direction sud-est (quartier et classes pauvres et zones de rénovation) et centre-ouest se prolongeant par ailleurs vers l'ouest au sud du Canal Lachine. Ces secteurs de recensement présentent un poids local relativement faible à très faible pour la dimension socio-économique. Par contre, les secteurs de recensement qui présentent un poids élevé ou très élevé se localisent autour du Mont-Royal (ville d'Outremont et de Westmount), sur la rive nord de l'île de Montréal (Lakeshore), au nord et à l'est de la rue St-Denis et au sud du boulevard métropolitain. Il est difficile surtout pour les quartiers de statut élevé de dégager une disposition sectorielle. Il semble à cet égard que les modalités historiques et culturelles de développement spatial ont une importance primordiale dans la structuration de l'espace social de Montréal.

La *deuxième composante* (13.8% de la variance totale) accentue l'importance de l'élément féminin (sexe ratio : .514 ; femmes de plus de quinze ans qui travaillent : .842 ; femmes dans l'administration, professions libérales, techniciennes : .575 ; femmes de plus de soixante-cinq ans : .639). Elle semble refléter le concept *urbanistique* tel qu'employé par les sociologues d'aires sociales. Nous l'avons dénommé dimension « urbanistique ». Les classes d'âges zéro à quinze ans (hommes : — .799 ; femmes : — .790), le taux de mariage (— .622), le nombre d'enfants en bas âge (— .788), composent la partie négative de cette composante. Les variables s'appliquent aux secteurs les plus peuplés à proximité du centre ville (pourcentage des bâtiments par rapport à la surface nette : .447). Le dessin de ce facteur varie d'une manière concentrique avec l'accroissement de la distance du centre ville de même que les attributs qui en dépendent. La corrélation est positive (significative à 0.5). Ce facteur reflète l'opposition entre les secteurs périphériques, à accroissement rapide, et les secteurs centraux.³³ L'étude de R. Murdie sur Toronto offre pour ces deux premières dimensions un schéma plus classique.³⁴

³³ N. LACOSTE, *Les caractéristiques sociales de la population du grand Montréal*, Faculté des sciences sociales, Université de Montréal, 1958, pp. 83-101 : « Dimensions de la famille et conditions de logement ».

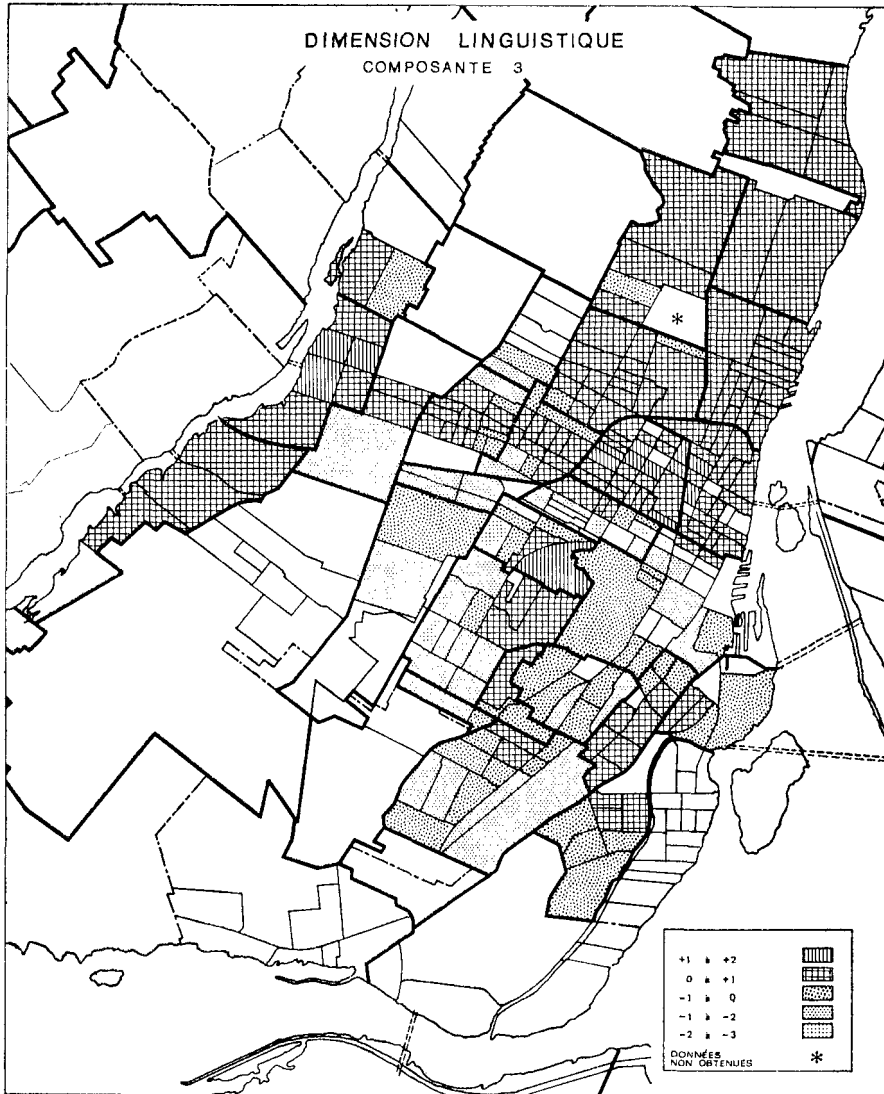
À part les secteurs d'attraction cités ci-dessus, les autres n'ont guère changé depuis cette étude (cf. pp. 105-168). On remarquera cependant la localisation géographique des secteurs 270, 272, 116, 117, et 101, dont l'image socio-économique est élevée.

³⁴ R. MURDIE, *op. cit.*



La troisième composante (9.3% de la variance totale) oppose les zones françaises aux zones non françaises (français : .892 ; unilingues français : .798 ; anglais : — .421 ; juifs : — .607). Cependant, l'examen de la carte des pourcentages des variables unilingues français et bilingues, telles que définies par le recensement, nous a poussé à considérer cette composante comme la dimension linguistique. Le boulevard Saint-Laurent limite nette-

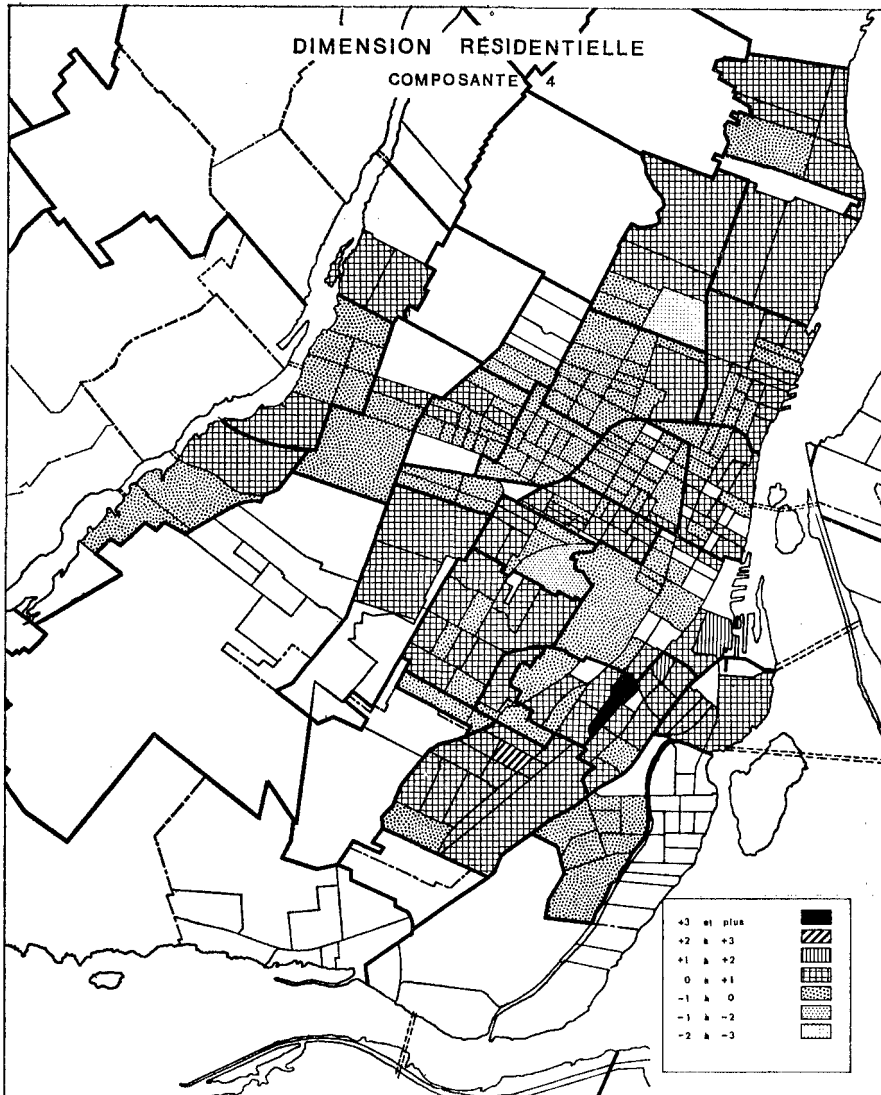
ment l'élément français de l'élément non-français. Au nord de la ville d'Outremont s'étend le quartier juif intégré, du moins par la langue, à l'élément anglais.³⁵



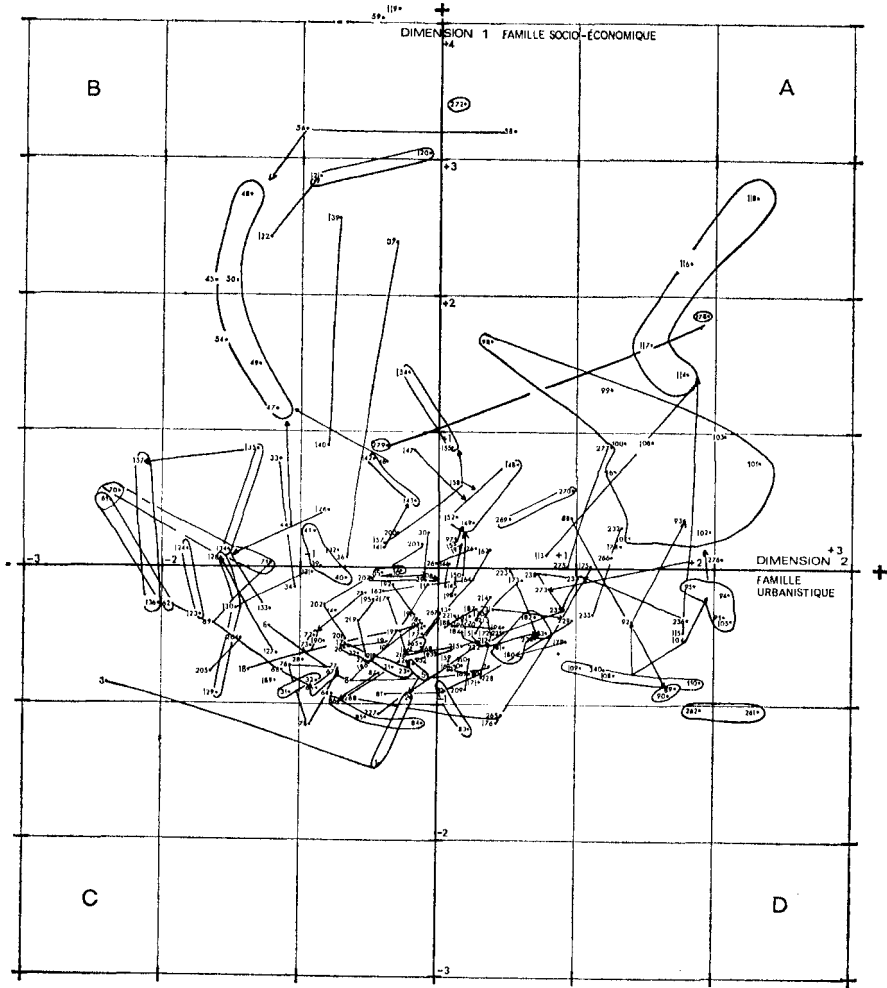
³⁵ LIEBERSON, *Linguistic and Ethnic Segregation in Montreal*, Commission royale d'enquête sur le bilinguisme et le biculturalisme, Ottawa, Imprimeur de la Reine, 1967.

Il semblerait que le système social et culturel français soit fermé si l'on considère les poids négatifs attachés aux variables suivantes : pourcentage des gens nés à l'extérieur du Canada : $-.953$ et pourcentage des gens immigrés entre 1946 et 1961 : $-.930$.

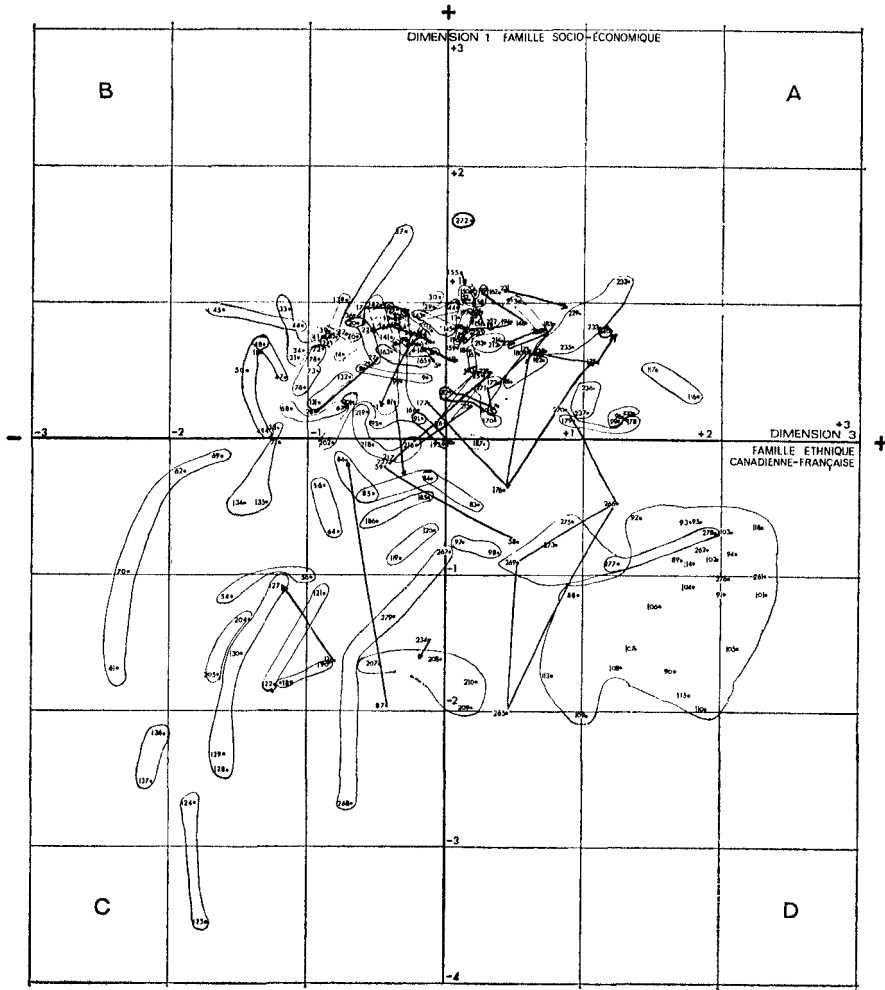
Enfin la *quatrième composante* (5.57% de la variance totale) est nettement reliée au concept de densité (pourcentage des familles de cinq enfants



Distribution des secteurs de recensement par rapport aux dimensions socio-économique et urbanistique.



Distribution des secteurs de recensement par rapport aux dimensions socio-économique et ethnique canadienne-française.



et plus : .432 ; logements surpeuplés : .431). Nous l'avons interprétée comme dimension résidentielle (pourcentage des logements occupés par un locataire : .683 ; logements ayant besoin de réparations importantes : .627). La répartition des appartements pour cette aire d'étude ne présente pas de concentration majeure, en fait cette composante n'est pas claire. (Voir par exemple l'espace entre l'autoroute des Laurentides et le boulevard Saint-Denis.)

Ainsi quatre facteurs peuvent exprimer la matrice d'information de cinquante-quatre variables. Nous avons obtenu trois dimensions, à savoir le statut socio-économique, l'urbanistique et le clivage linguistique, et une quatrième, moins nette, représentant la structure résidentielle. Les deux premières composantes approchent les composantes classiques élaborées par les auteurs cités dans le texte, par contre la dimension linguistique est nettement marquée par les modalités historiques de la croissance spatiale. Enfin, la dernière composante semble refléter la transformation des secteurs périphériques et surtout de certains secteurs centraux.

II. ROTATION OBLIQUE VERSUS ROTATION ORTHOGONALE

La rotation oblique a été délaissée par les géographes ; une des raisons tient à sa complexité. Cependant les travaux de C.W. Harris et H.F. Kaiser, de même que ceux de R.I. Jennrich et P.F. Simpson, présentent une nette amélioration par rapport aux alternatives classiques. L'auteur développe et compare les résultats obtenus par l'analyse factorielle oblique et orthogonale. De plus il soulève un certain nombre de réflexions sur le traitement des données, sur la philosophie et la logique de l'analyse factorielle, tout en affirmant que cet outil mathématique ne saurait se substituer à l'effort méthodologique du chercheur.

On peut considérer la rotation des facteurs comme une solution indéterminée. Une des raisons provient que l'on ne connaît pas en général pour un problème donné le nombre de facteurs à extraire de même que les variables spécifiques qui devraient comporter une saturation égale à zéro sur les facteurs. Il s'agit alors de prévoir une autre alternative. Cattell remarque que l'approche orthogonale entraîne une solution biaisée du fait qu'on ne tient pas compte du comportement des variables qui sont le plus souvent en corrélation (obliques) :

* Experience in locating naturally-existing hyperplanes, as soon as sufficient scientific data became available, showed, however, what would have been

theoretically expected, namely, that in most scientific fields factors are oblique. »³⁶

R.J. Johnston remarque que la majorité des études portant sur l'analyse socio-économique des villes nord-américaines, particulièrement le modèle de Shevky et Bell, dévoilent plus que trois facteurs. Bien que la ou les dimensions supplémentaires n'invalident pas le modèle, certains résultats semblent contredire la théorie.³⁷ Dans leur étude sur Toledo, Anderson et Bean identifient une dimension intitulée urbanisation qui émerge du statut familial ; de plus la caractéristique fécondité est en corrélation à la fois avec sa propre dimension et avec la dimension urbanisation.³⁸ Il apparaît que les dimensions extraites ne sont pas indépendantes, ce qui sous-entend que l'approche semble inappropriée.³⁹

a) *Méthodologie*

La rotation oblique apparaît plus complexe que la rotation orthogonale. L'analyse comporte une matrice de corrélations entre facteurs et une matrice factorielle double. On trouve dans Harman et dans Rummell les notions essentielles à la fois sur les alternatives et sur les techniques de rotation oblique.⁴⁰ Comme dans l'analyse orthogonale, on tend à ce que chaque facteur possède un nombre de saturations qui sont aussi proches de l'unité que possible tandis que le reste des saturations accordées « — » avec ces facteurs aurait une corrélation proche de zéro.⁴¹ Dans la rotation oblique, on peut mettre en évidence deux possibilités de projeter une variable que l'on représentera sous forme de point. On peut en premier lieu tracer du point en question deux parallèles aux axes obliques. Les différentes projections (variables) sur les axes obliques (facteurs) déterminent le schéma des saturations (*pattern loadings*). Celui-ci peut s'interpréter comme une matrice de saturations mesurant la dépendance des variables sur les différents facteurs dans la mesure où ils représentent la table des coefficients de corrélations entre les variables et les facteurs obliques. En second lieu, on peut projeter à partir de points représentant une variable non plus des

³⁶ R.B. CATTELL, « Factor Analysis. An introduction to Essentials », *Biometrics*, XXI, pp. 190-215 et 403-435 : p. 209.

³⁷ R.J. JOHNSTON, « Some Limitation of Factorial Ecologies and Social Area Analysis », *Economic Geography*, XLVII, 2, juin 1971, pp. 314-323.

³⁸ T.R. ANDERSON et L.L. BEAN, « The Shevky - Social Areas. Confirmation of Results and Reinterpretation », *Social Forces*, 40, 1961, pp. 119-124.

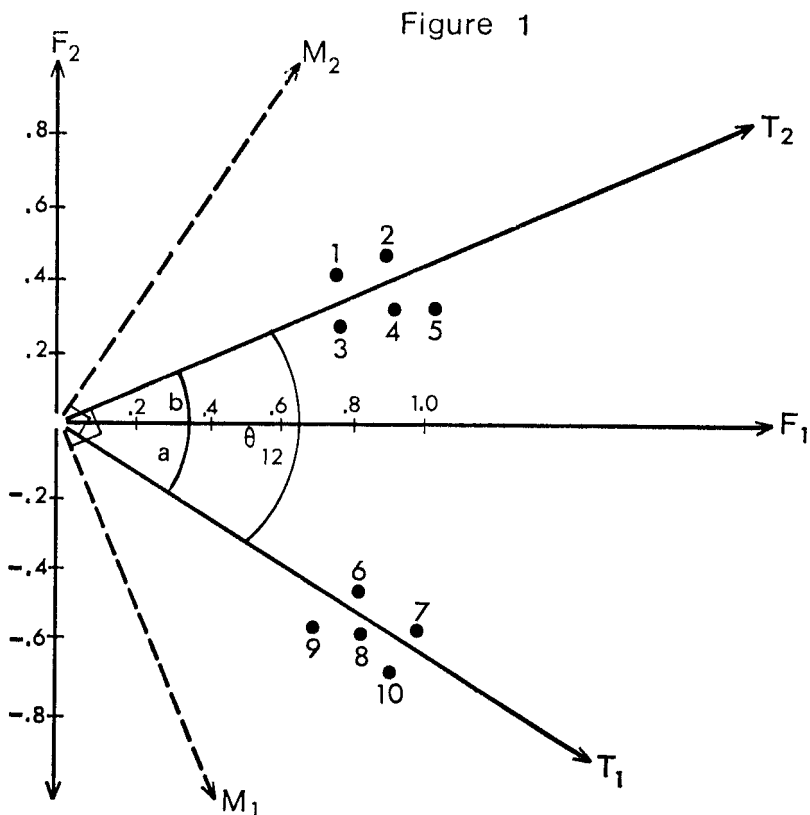
³⁹ B.J.L. BERRY, « Introduction : The Logic and Limitations of Comparative Ecology », *Economic Geography*, XLVII, 2, pp. 209-219.

⁴⁰ H. HARMAN, *Modern Factor Analysis*, Presses de l'Université de Chicago, 2^e édition, 1967, chapitres 13-15, pp. 237-245.

⁴¹ L.L. THURSTONE, *Multiple Factor Analysis*, Presses de l'Université de Chicago, 1947. La structure simple consiste à roter le facteur autour de l'origine jusqu'à ce que chaque facteur maximise sa variance totale.

parallèles aux axes obliques mais des perpendiculaires. Les différentes projections sur les axes obliques déterminent alors une structure de saturations (*structure loadings*). Celle-ci peut s'interpréter comme la table de corrélations multiples de variables avec les facteurs obliques.⁴²

Ces deux matrices offrent une perspective d'analyse des données différente. Le *schéma des saturations* révèle mieux les groupements de variables déterminés par les facteurs obliques, par contre la *structure des saturations* permet de mesurer la variance de chaque variable sur un ou plusieurs facteurs obliques. Par conséquent, dans le cas d'une structure de saturations d'une variable sur deux facteurs, elle établit la relation directe entre la variable et chaque facteur et l'interaction entre les deux facteurs exprimés par la corrélation entre ceux-ci et la variable.⁴³



⁴² Dans le cas d'axes orthogonaux, il y a identité entre les parallèles et les perpendiculaires menées par un point aux axes, par conséquent on observe qu'une seule matrice peut être interprétée comme la table de corrélations entre les variables et les facteurs orthogonaux.

⁴³ Voir: R.J. RUMMELL, *op. cit.*, pp. 395-405.

Pour illustrer ce développement, considérons la solution orthogonale offrant la saturation de dix variables sur deux facteurs F_i ($i = 1,2$). Le groupement des points dans le premier quadrant représente cinq variables et dans le quatrième quadrant les cinq autres variables. Projetons deux vecteurs T_1 et T_2 sur les deux groupements à partir de l'origine 0. Les deux vecteurs forment entre eux un angle θ . En fait, ils constituent la projection des vecteurs représentant les variables sur quatre axes obliques ; en d'autres termes, ils identifient la corrélation des variables avec les facteurs. En écriture matricielle, soit A la matrice des saturations orthogonales et S la matrice représentant la structure de saturations obliques (*oblique factor structure matrix*).

$$S = AT (1)$$

où T représente la matrice de transformation contenant en colonnes le cosinus des axes obliques par rapport aux axes de référence orthogonaux. Dans l'exemple :

$$T = \begin{bmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{bmatrix}$$

Soit B la matrice de corrélation des facteurs :

$$B = \begin{bmatrix} 1 & R_{12} \\ R_{21} & 1 \end{bmatrix}$$

la relation entre les facteurs communs au schéma oblique (*oblique pattern*) et à la structure oblique (*oblique structure*) est :

$$S = PB$$

où S représente la structure des facteurs. Remarquons que la matrice S est égale au schéma des saturations obliques P multiplié par la matrice B des corrélations des facteurs. La solution pour le schéma des facteurs est alors :

$$P = SB^{-1}$$

Une fois établi le schéma des saturations obliques, on peut dégager la contribution directe et conjointe des facteurs.⁴⁴

⁴⁴ R.K. SEMPLE, *An oblique simple structure factor analysis of viability measures for southern Ontario towns*, Discussion Paper no. 2, Département de géographie, Université de Toronto, octobre 1969. Notre présentation est largement inspirée de cette note de discussion.

Au lieu d'interpréter les axes obliques directement, on peut définir un nouveau système d'axes perpendiculaires au système original. Ce système fut proposé par Thurstone pour répondre aux critères de structure simple, en d'autres termes pour garantir un certain nombre de zéros dans la solution factorielle. Pour distinguer ces axes des premiers, on les divise en axes de référence. Il existe une importante relation entre ces deux formes de rotation : le schéma des premiers axes est la structure des axes de référence et la structure des premiers axes est le schéma des axes de référence.⁴⁵ Cette forme de rotation est largement employée, en particulier par Thurstone et Catell. Elle est utile lorsque les facteurs peuvent être en corrélation étroite. Cependant elle est plus compliquée que la première forme de rotation.

Soit M la matrice de transformation. Comme ce fut le cas pour la matrice T , elle comporte, en colonnes, le cosinus formé par l'angle des axes de référence et des premiers axes.

Ainsi la connaissance de la transformation de la matrice T nous permet de déterminer la matrice M . N'oublions pas d'autre part que la matrice T peut être considérée comme extension de la matrice A du fait que les éléments de T sont les cosinus d'axes passant pas les groupements de variables et sont par conséquent des vecteurs additionnels (comme les variables) du schéma de référence orthogonal. Comme la matrice T comporte, je le répète, les cosinus dans ses colonnes, il est nécessaire de prendre la transposée de T . Dès lors en appliquant la transformation $V = AM$ à T' on a : $T'M = D$, où D est la matrice de corrélation entre les vecteurs obliques de la matrice T et le nouveau système de vecteurs obliques M .

La matrice M transformée devient : $M = (T')^{-1} D$. De la connaissance de T , on peut alors calculer l'inverse de sa transposée. Une fois la matrice transformée M déterminée, on multiplie la matrice A par celle-ci pour obtenir la matrice solution V .

Les corrélations des axes de référence M , peuvent être obtenues par la multiplication matricielle suivante : $M'M = R$ où R est la matrice de corrélation. Pour compléter cette solution, le schéma des facteurs W , en terme d'axes de référence M , s'établit :⁴⁶ $W = VR^{-1}$

Harman montre que la relation entre les deux formes de transformation (rotation) est :

$$P = VD^{-1} \quad \text{ou} \quad V = PD$$

$$\text{et } S = WD \quad \text{ou} \quad W = SD^{-1}$$

⁴⁵ R.J. RUMMELL, *op. cit.*, pp. 406-408.

⁴⁶ R.K. SEMPLE, *op. cit.*, pp. 2-6.

en d'autre termes, la structure d'une forme de rotation est liée au schéma de l'autre solution et vice-versa.⁴⁷

Les techniques de rotation oblique sont particulièrement nombreuses. R.J. Rummell en cite sept qui satisfont les critères de structure simple mis en évidence par Thurstone : oblimax, quartimin, covarimin, biquartimin, binormanin, maxplane, promax. On pourrait ajouter la technique développée par C.W. Harris et H.F. Kaiser qui semble particulièrement intéressante. Ils considèrent que la solution oblique peut être obtenue à partir d'une transformation orthogonale.⁴⁸ Quoi qu'il en soit, aucune technique ne paraît « la plus » satisfaisante. Il est recommandé dans la mesure du possible d'essayer plusieurs ; si on ne constate pas de modification majeure dans les matrices de saturation, ce qui devrait être le cas en principe, on peut considérer la solution comme satisfaisante. En 1957, J.B. Carroll établit une méthode originale fondée sur la comparaison des critères des techniques dite du quartimin (Q) et du covarimin (C). L'objectif est de minimiser la fonction (où M est égal au nombre de variables) :

$$B = Q + \frac{C}{M}$$

L'expérience montre que cette technique dite du « biquartimin » donne une solution plus satisfaisante en terme de simple structure et de corrélations entre facteurs que celles du covarimin ou du quartimin.⁴⁹ Cette fonction généralisée découle du terme « oblimin » où les techniques du covarimin et du quartimin apparaissent comme singulières. En pondérant les fonctions (Q) et (C) par l'addition de paramètres B_1 et B_2 , il est amené à combiner (Q) et (C) de différentes façons :

$$B^* = B_1 Q + \frac{B_2 C}{M}$$

⁴⁷ H. HARMAN, *op. cit.*, pp. 290-291. Les deux systèmes d'axes T et M s'inscrivent dans le même espace. Du fait de leur orthogonalité, il est évident que les cinq points les plus proches de l'axe T_1 , ont une projection sur l'axe M_2 , voisine de zéro. Il en est de même des variables les plus proches de l'axe T_2 , en regard de l'axe M_1 .

⁴⁸ Pour une description mathématique des quatre premières techniques, voir : H. HARMAN, *op. cit.*, chapitre 15 ; C.W. HARRIS et H.F. KAISER, « Oblique Factor Analytic Solutions by Orthogonal Transformations », *Psychometrika*, XXIX, 4, décembre 1964, pp. 347-362 ; A.R. HAKSTRAN, « A Comparative Evaluation of Several Prominent Methods of Oblique Factor Transformation », *Psychometrika*, XXXVI, 2, juin 1971, pp. 175-193.

⁴⁹ J.B. CARROLL, « Biquartimin Criterion for Rotation to Oblique Simple Structure in Factor Analysis », *Science*, CXXVI, pp. 1114-15.

En remplaçant (Q) et (C) par leur fonction et en introduisant $j = B_2 / (B_1 + B_2)$ il détermine la fonction à minimiser :

$$B^* = \sum_{p>q=1}^p \left\{ \sum_{j=1}^m \left(\frac{\alpha_j^{2i}}{n^{2j}} \right) \left(\frac{\alpha_j^{2jq}}{h^{2j}} \right) - j \left[\sum_{j=1}^m \left(\frac{\alpha_j^{2i}}{h^{2j}} \right) \right] \left[\sum_{j=1}^m \left(\frac{\alpha_j^{2j}}{h^{2j}} \right) \right] \right\}$$

où α_{j1} est le coefficient de saturation oblique dans le cas des axes de référence et i et q sont le i^e et q^e facteurs.⁵⁰

Cette approche indirecte par les axes (facteurs) de référence a été par la suite remplacée par l'approche directe fondée sur les axes (facteurs) premiers répondant aux critères de structure simple.⁵¹

b) Analyse

L'aire géographique, la matrice d'information restent les mêmes que ci-dessus. La technique utilisée est celle de la rotation oblique fondée sur les critères de « l'oblimin ». Le résultat est un ensemble de facteurs obliques, donc en corrélation entre eux, représentatifs des dimensions latentes de l'espace social.

Matrice de corrélation inter-facteurs

	1	2	3	4
1	1.000	0.239	0.305	0.150
2	0.239	1.000	-0.033	0.092
3	0.305	-0.033	1.000	-0.206
4	0.150	0.092	-0.206	1.000

Cette matrice montre que les facteurs ont une corrélation relativement faible. Plus les coefficients de corrélation tendent vers zéro plus les facteurs obliques approchent l'orthogonalité. Le schéma des saturations ne montre pas de différences importantes avec la matrice des saturations orthogonales

⁵⁰ Cette expression est équivalente à :

$$B(V) = \sum_{p>q=1}^m \left[n \sum_{j=1}^n V^2_{jp} \quad V^2_{jq} \quad - \quad B \sum_{j=1}^n V^2_{jp} \quad \sum_{j=1}^n V^2_{jq} \right]$$

si l'on conserve les termes et expressions du développement matriciel. (J.B. CARROLL, *Oblimin Rotation Solution in Factor Analysis*, Computing Program for the IBM 704, 1958, mimeo.)

⁵¹ R.L. JENNRICH et P.F. SAMPSON, « Rotations for Simple Loadings », *Psychometrika*, XXXI, 1966, pp. 313-323.

après rotation Varimax. En fait, le « nettoyage » de la matrice ne modifie pas l'interprétation. La première dimension comporte des variables dont la saturation est négative (ménages de six personnes et plus, diplôme élémentaire, revenus de moins de \$4,000, femmes ouvriers et manœuvres) ; la seconde dimension conserve des caractéristiques familiales, plus précisément « urbanistiques », telles qu'employées par les sociologues des aires sociales, (saturation négative : ménages, positive : célibataires et femmes qui travaillent). Par contre, la structure diffère pour la troisième et quatrième dimension. En effet les variables « pourcentage de français » et « unilingue français » se détachent de la dimension ethnique dégagée par la technique orthogonale ; de ce fait cette dimension ne comporte que les caractéristiques suivantes : immigrants récents et pourcentage d'anglais (saturation positive pour ces cinq variables).

Ce résultat est corroboré par l'association négative entre le facteur 3 et le facteur 4 (- 0.206). La contribution directe et conjointe des quatre facteurs s'élève à 63.77% de la variance totale, ce qui est très peu différent du pourcentage obtenu pour les quatre premiers facteurs par des techniques orthogonales. Par contre, la dimension I passe à 26.96, la dimension II à 19.33, et les dimensions III et IV explicitent seulement 4.8 et 4.5 pour cent de la variance.

c) Commentaires

Nous avons cherché à comparer l'apport de différentes techniques ; il reste à affronter leur apport dans la démarche géographique.⁵² L'accumulation de données souvent redondantes et dont les liaisons statistiques sont difficiles à interpréter sans modèle *a priori*, peut avoir des conséquences importantes sur les résultats de l'analyse factorielle. Joahi a montré que l'inclusion de variables non importantes dans le calcul des pointages factoriels pouvait masquer la contribution des variables importantes⁵³

D'une façon plus générale, R. Murdie a souligné le rôle prépondérant du chercheur dans la composition de la matrice d'information. La plupart des chercheurs explicitent comment et pourquoi l'espace peut être différencié. Comme le souligne P.H. Rees, l'origine de l'analyse factorielle découle de la psychologie, d'études de personnalité ; la théorie portant sur la structure du « psyché » humain étant peu développée, les psychologues essaient d'isoler les composantes fondamentales de la personnalité d'une myriade de données

⁵² J.P. THOUÉZ, « L'analyse typologique et l'espace géographique ». (À paraître.)

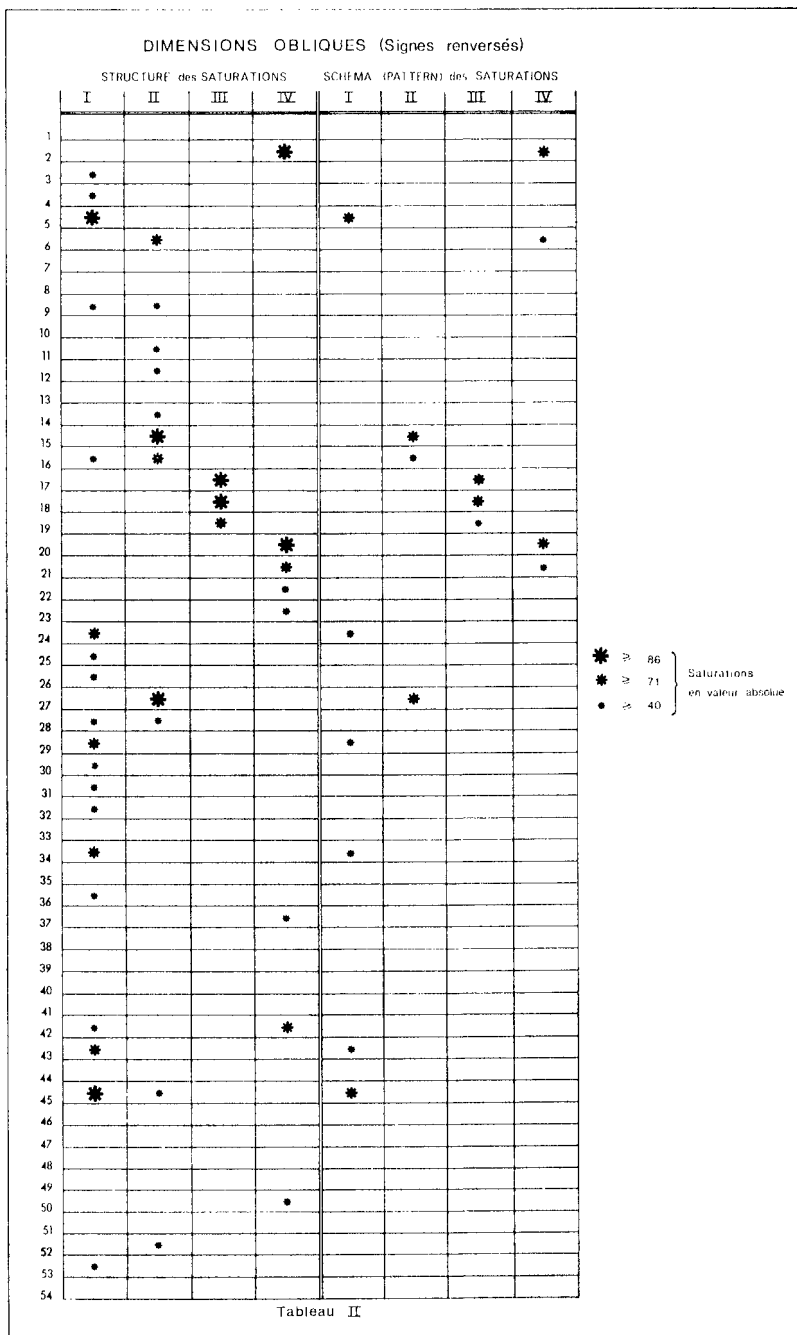
⁵³ T.R. JOAHI, « Toward Computing Factor Scores », in W.P. HELLUNIR, (ed.), *Congrès de géographie internationale*, II, 1972, p. 906.

DIMENSIONS ORTHOGONALES (Signes renversés)

	AVANT ROTATION				APRÈS ROTATION (TECHNIQUE VARIMAX)			
	I	II	III	IV	I	II	III	IV
1				*				*
2								
3	*					*	*	
4	*							*
5	*				*			
6		*						
7								
8						*		
9	*	*			*	*		
10								
11		*				*		
12		*				*		
13						*		
14		*				*		
15	*	*			*	*		
16	*	*			*	*		
17			*				*	
18			*				*	
19					*		*	
20			*				*	
21			*		*		*	
22			*		*			
23			*		*		*	
24	*							
25	*				*	*		
26	*				*			
27		*				*		
28	*	*			*	*		
29	*				*	*		
30	*				*	*		
31	*				*	*		
32	*				*	*		
33					*	*		
34	*				*	*		
35					*	*		
36	*				*	*		
37				*				*
38								
39								
40								
41				*				*
42	*				*			*
43	*				*		*	
44								
45	*	*						
46								
47								
48								
49				*				*
50								
51						*		
52		*			*	*		
53	*				*	*		
54								

* \forall \forall \forall } Saturations
* \forall \forall \forall } en valeur absolue
* \forall \forall \forall }

Tableau I



spécifiques.⁵⁴ Les facteurs dégagés ne sont pas des facteurs explicatifs au sens de causalité, mais des facteurs qui décrivent les schémas d'association entre les attributs sur les observations.⁵⁵ Les saturations dépendent donc des associations entre les données qui répondent à la démarche : hypothèses, observation, traitement statistique des données. Certes, on peut concevoir *a priori* une structure factorielle-type avant d'effectuer l'analyse, mais cette structure décrit des dimensions, elle n'explique pas des influences causales. Ainsi, les trois dimensions classiques dégagées par les sociologues pour les villes américaines ne signifient pas que la population urbaine se différencie par les statuts socio-économique, familial et ethnique mais que les unités d'observation (ou les sous-espaces homogènes dégagés par l'analyse en groupement) varient entre eux quant au statut de leur population respective.

Comme les données sont en général normalisées et standardisées, nous ne connaissons pas l'importance de l'information perdue. De plus, R.J. Johnston observe que peu de chercheurs ont essayé de vérifier si la relation entre deux variables était linéaire. Certes, le coefficient de corrélation représente la relation essentielle entre variables mais la forme de la relation disparaît dans l'analyse factorielle. Le même auteur souligne que l'on devrait utiliser des tables de contingence ou des mesures d'association, telles que gamma ou celles développées par Goodman, dans l'analyse des distributions des données.⁵⁶

La technique de rotation oblique reste difficile surtout à interpréter. De plus, la solution à laquelle on aboutit dépend des contraintes mathématiques qui n'ont pas toujours de signification pour le problème que l'on traite ; en général la solution factorielle oblique ne diffère pas fondamentalement de la solution orthogonale.⁵⁷

Enfin K.E. Haynes analysant la stabilité des dimensions et la stabilité spatiale des pointages factoriels pour 1951 et 1961, à la fois par la technique

⁵⁴ P.H. REES, « Factorial Ecology and Extended Definition. Survey and Critique of the Field », *Economic Geography*, XLVII, 2, juin 1971, pp. 220-233.

⁵⁵ *Ibid.* : « ces relations diffèrent suivant l'échelle d'observation utilisée... » de même les limites géographiques de l'aire d'étude correspondent au système de relations spatiales ou de séparations spatiales que l'on désire expliciter.

Bien que R. FORSTALL, travaillant sur 1761 villes américaines et 97 variables, utilise plus de 170,000 données et 24,000 pointages factoriels, il identifie les mêmes dimensions : statut socio-économique, étapes dans le cycle familial, population de couleur. (« A New Social and Economic Grouping of Cities », *Municipal Yearbook*, Washington, International City on management association, 1970, pp. 102-159.)

⁵⁶ R.J. JOHNSTON, *op. cit.* Voir : L.A. GOODMAN, « How to Ransack Social Mobility Tables and other Kinds of Cross Classification Tables », *American Journal of Sociology*, LXXV, 1966, pp. 1-40 ; L.A. GOODMAN et W. KRUSTAL, « Measures of Association for Cross Classifications », *Journal of the American Statistical Association*, XLIX, 1954, 732-764.

⁵⁷ ANDERSON et BEAN, « The Shevky - Bell Social Areas. Confirmation of Results and Reinterpretation », *Social Forces*, XL, 1961, 119-124, cité par R.J. JOHNSTON, *op. cit.*

orthogonale et oblique, découvre que les facteurs obliques tendent vers une structure orthogonale.⁵⁸

D'autres commentaires pourraient être apportés à notre recherche, en particulier sur l'importance de l'analyse orthogonale ou « image » dans le calcul et la représentation des pointages factoriels. Quoiqu'il en soit, le champ de l'analyse factorielle a connu d'importants développements depuis les années soixante. Un des buts fondamentaux est d'établir des dimensions indépendantes des techniques utilisées. L'emploi de l'analyse oblique reste indispensable à cet égard, si l'on suspecte que les dimensions sont en corrélation. La matrice de corrélations entre les facteurs constitue alors un outil extrêmement important à la connaissance structurelle du milieu à analyser.

Jean-Pierre THOUZ

*Département de géographie,
Université de Sherbrooke.*

⁵⁸ K.E. HAYNES, *op. cit.*

ANNEXE : STRUCTURE SPATIALE DES CARACTÉRISTIQUES
SOCIO-ÉCONOMIQUES, RÉSULTATS COMPARATIFS

L'analyse de Greer Wootten (*Le modèle urbain. Guide d'excursions*, 22^e Congrès international de géographie, Montréal, PUM, 1972, pp. 1-31) porte principalement sur l'île de Montréal. Elle apparaît plus vaste que la nôtre car elle englobe toutes les municipalités de l'île. La matrice d'informations ne comporte que vingt-sept variables ; nous les retrouvons à deux variables près dans notre étude. Par contre, l'auteur emploie une autre forme d'analyse factorielle dite analyse « image ». (H.F. KAISER, « Image Analysis », dans C.W. HARRIS (ed.), *Problems in Measuring Change*, Madison, Presses de l'Université de Wisconsin, pp. 156-166.)

GREER WOOTTEN :

	I	II	III	IV	V
% Variance	23.22	11.40	9.01	9.13	6.59
% Variance accumulée	23.22	34.62	42.63	49.76	56.34

THOUÉZ :

	I	II	III	IV	V
% Variance	29.30	13.55	9.24	5.57	4.28
% Variance accumulée	29.30	42.59	51.89	57.46	62.42

NICHOLSON et YEATES, (Winnipeg, 1961) :

	I	II	III	IV	V
% Variance	35.02	19.52	9.52	6.68	5.04
% Variance accumulée	35.02	54.54	64.60	70.68	75.72

MURDIE (Toronto, 1961) :

	I	II	III	IV	V	VI
% Variance	21.1	15.6	12.2	11.5	7.2	6.9
% Variance accumulée	21.1	36.7	48.9	53.4	61.1	67.9