

F. GAUCHET

J.-J. BOULANGER

La recherche des facteurs indépendants et son rôle dans la schématisation et l'explication des phénomènes à caractères multiples

Journal de la société statistique de Paris, tome 94 (1953), p. 277-286

http://www.numdam.org/item?id=JSFS_1953__94__277_0

© Société de statistique de Paris, 1953, tous droits réservés.

L'accès aux archives de la revue « Journal de la société statistique de Paris » (<http://publications-sfds.math.cnrs.fr/index.php/J-SFdS>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
<http://www.numdam.org/>

VII VARIÉTÉ

La recherche des facteurs indépendants et son rôle dans la schématisation et l'explication des phénomènes à caractères multiples.

APPLICATION A L'ÉTUDE RATIONNELLE DU TEST DE RORSCHACH.

La recherche de facteurs ou éléments indépendants, permettant la schématisation d'un phénomène à caractère multiple, ne constitue pas un problème récent.

Déjà M. Michel Huber, dans l'un de ses ouvrages, a signalé que la composition d'un indice de prix exige toujours un choix délicat des divers éléments présidant à sa composition et que ces éléments doivent être choisis aussi indépendants que possible. On peut dire cependant que le problème présente un caractère plus vaste ayant une portée quasi universelle.

Toute collectivité d'objets définis par plusieurs caractères, tout groupe d'individus considérés du point de vue anthropométrique, physiologique, psychologique ou à tout autre point de vue, tout phénomène économique en évolution, phénomène général (économie générale) ou particulier (une industrie ou une entreprise particulière) posent le problème d'une schématisation à l'aide d'éléments aussi indépendants que possible, qui permettra une analyse mathématique du schéma ou modèle ainsi obtenu (1).

Il n'est pas dans nos intentions de faire le procès des méthodes d'analyse factorielle, dont les résultats obtenus en psychologie et psychotechnie présentent la meilleure défense. Signalons cependant les résultats moins fructueux obtenus en anthropologie et en physiologie (absence de facteur général) ainsi qu'en biométrie (corrélations non linéaires et surtout concentrations souvent anormales de points dans les espaces à deux dimensions qui conduisent à la notion de type au sens statistique du terme). C'est d'ailleurs ce qui nous a conduit il y a quelques années à reprendre l'analyse des répartitions de points dans l'espace à deux dimensions sans autre fil directeur que celui indiqué par l'ensemble des généralités « cliniques » normalement admises et d'un point de vue beaucoup moins théorique que pratique.

Et le fait important est que la statistique permet de retrouver les conclusions « cliniques » généralement admises.

(1) R. KHÉRUMIAN et J.-J. BOULANGER, *Essai de caractéristique somatique de l'individu par un ensemble d'indices indépendants*. Biotypologie, juillet 1949.

L.-C. SOULA et J.-J. BOULANGER, *Le test de la glycémie post-prandiale*. Biotypologie, juillet 1949.

J.-J. BOULANGER, *Chances d'une économie appliquée*, Revue d'Économie politique, mars-avril 1953.

Ainsi il est reconnu en anthropologie que la longueur de tête, la largeur de tête et le diamètre bi-zygomatique sont les mesures les plus importantes pour la description des races (avec la taille) : en fait, le coefficient de variation $Cv = 100 \sigma/m$ est généralement de 3,5 à 4 pour la longueur et la largeur de tête, de 3,5 à 4 pour le diamètre bi-zygomatique, de 4,5 à 6 pour la hauteur de la face et de 6 à 10 pour la largeur et la hauteur du nez.

Ainsi il est souvent reconnu en biométrie que la taille debout et la taille assis, les diamètres biacromial et bicrête sont les mesures les plus importantes pour caractériser somatiquement l'individu, le poids n'ayant de véritable rôle qu'une fois le bâti osseux défini : en fait, la variabilité de ces éléments est très faible (Cv de l'ordre de 4,0 — 4,0 — 5,0 et 5,5) par rapport à celle des autres mesures.

Il y aurait beaucoup à dire sur les répartitions asymétriques obtenues en pratique, d'autant plus asymétriques que le coefficient de variation est élevé (cf. Les distributions de χ et de Poisson) et sur leur ajustement par des courbes classiques : Binomiale, Gauss, Poisson, χ (1), ou Gibrat, chaque type d'ajustement conduisant à une interprétation particulière des facteurs supposés déterminer le phénomène observé. Nous arrivons ainsi à une deuxième notion de facteurs indépendants qui sont en quelque sorte les « atomes » d'un phénomène déterminé et seront représentés par exemple par les degrés de liberté d'une courbe de χ .

L'ajustement par une courbe de χ permet, en effet, de déclarer que « tout se passe comme si » le nombre de facteurs indépendants conduisant à une taille était de l'ordre de 300, ceux conduisant à un poids de l'ordre de 40 à 50, ceux conduisant à un taux de glucose sanguin à jeun de l'ordre de 50.

Mais ce n'est là qu'un mode primaire d'explication.

Par suite de circonstances diverses, déterminées (hérédité) ou fortuites (milieu), les divers facteurs ne se comportent pas d'une façon indépendante mais sont liés ensemble, groupés en « paquets »... bref, tout se passe comme s'ils étaient moins nombreux. Ce qui oblige à conclure qu'en réalité le nombre de facteurs influant sur la taille par exemple est certainement beaucoup plus élevé que le chiffre avancé précédemment.

Les conclusions de l'analyse factorielle ont d'ailleurs suivi la même évolution au cours des récentes années.

Il est certain que le développement des méthodes statistiques comme interprétation et explication des phénomènes ne fait que commencer et que l'avenir offre à la technique statistique un rôle infiniment plus vaste et plus important que celui occupé jusqu'à ce jour.

Dès à présent, on pourrait songer par exemple à des familles de courbes reliant, entre autres, la moyenne et le coefficient de variation et traduisant des types d'organisation similaires dans les domaines de l'hérédité, du physio-

(1) Distribution de χ : $dP = 2 Cv \chi^{v-1} e^{-\frac{\chi^2}{2}} d\chi$ avec $Cv = \frac{1}{2^v \Gamma(\frac{v}{2})}$, v étant le nombre de

degrés de liberté.

Quand $v > 30$, $\chi \sqrt{2}$ suit approximativement une loi de Gauss de moyenne $\sqrt{2v-1}$ et de variance 1 (Fisher).

logique, du psychologique, du social ou de l'économique. Mais il s'agit plutôt de perspectives d'avenir. C'est pourquoi on peut préférer donner aux quelques pages qui suivent le titre d'Introduction.

J.-J. BOULANGER.

INTRODUCTION A L'ÉTUDE RATIONNELLE DU TEST DE RORSCHACH

Le test de Rorschach se présente comme une série de dix images non structurées (taches d'encre). On montre séparément chacune de ces images au sujet en lui demandant d'indiquer ce qu'elle pourrait représenter, ce qu'il y voit. On l'avertit qu'il n'y a pas de solution à trouver, qu'il est donc entièrement libre de ses réponses.

Celles-ci sont classées suivant la partie de la tache interprétée (ensemble, détails fréquemment interprétés, détails insolites), le déterminant de l'interprétation (formes, mouvement, couleur), le contenu enfin de l'interprétation (ceci répond respectivement aux questions où? pourquoi? quoi?). Et c'est en structurant l'informe, en lui prêtant parfois vie et couleur, en le peuplant de ses propres pensées que le sujet projette dans sa façon de voir une façon d'être.

La méthode d'interprétation du test de Rorschach, comme celle, du reste, de tous les tests projectifs ne peut être ici exposée, même de façon succincte, en raison de son extrême complexité et de sa richesse.

L'interprétation d'un test à détermination multiple, comme le Rorschach pose des problèmes très variés parmi lesquels la liaison entre les « localisations », les « déterminants » et les « contenus », ainsi que l'étude des corrélations entre les divers « déterminants ».

L'étude préliminaire qui suit s'efforce de jeter les bases d'une analyse systématique et statistique des déterminants, de leurs similitudes et différences, et de leur groupement, indépendamment de leur interprétation psychologique.

Documents

L'échantillon analysé comprenait 100 sujets à niveau mental élevé (Test Progressive Matrices de Raven, moyenne : 47,6) et formait à ce point de vue un groupe beaucoup plus homogène que n'en procure un simple échantillonnage.

Le test est interprété suivant la technique de Klopfer et Kelley (« The Rorschach Technique » — World Book Company, 1942).

Les déterminants considérés étaient les suivants :

1. M — Interprétations dans lesquelles des figures humaines sont vues en mouvement.
2. FM — Interprétations dans lesquelles des figures animales sont vues en mouvement.
3. m — Interprétations dans lesquelles des objets sont vus en mouvement.

TESTS DE RORSCHACH

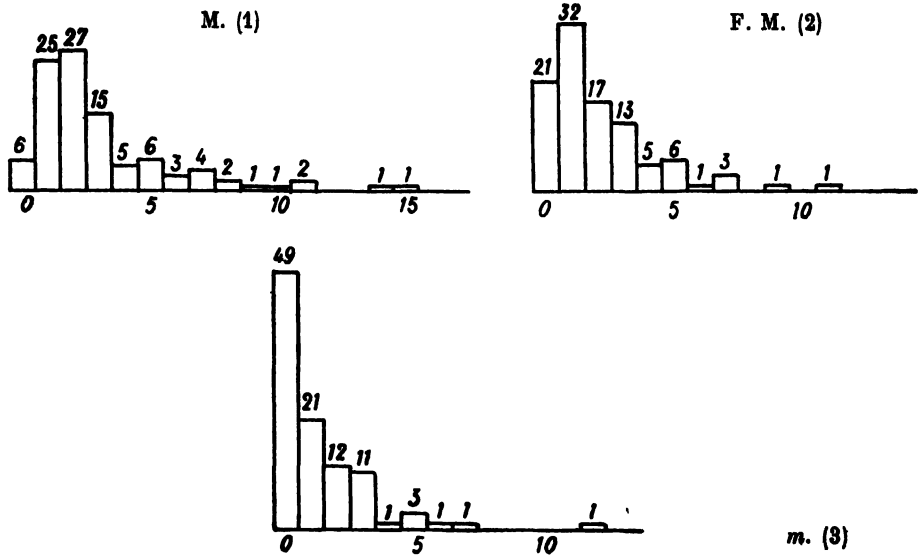


Fig. 1. — Déterminants M - F.M - m.

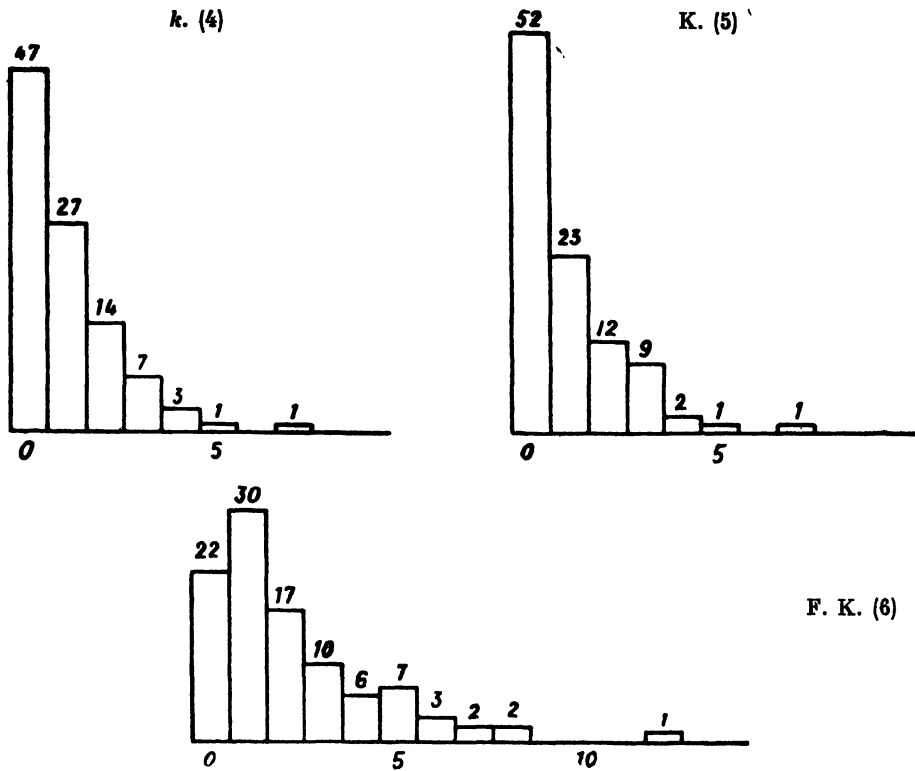


Fig. 2. — Déterminants k - K - F.K

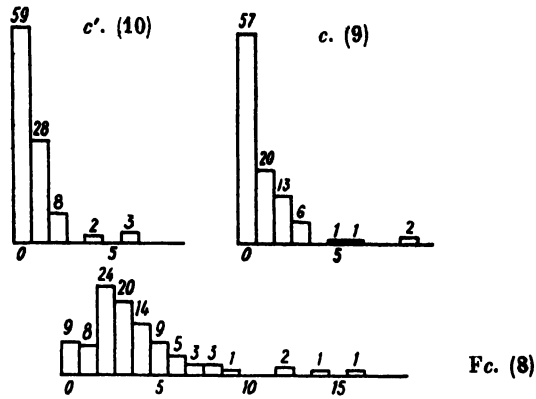


Fig. 3. — Déterminants $Fc - c - C'$.

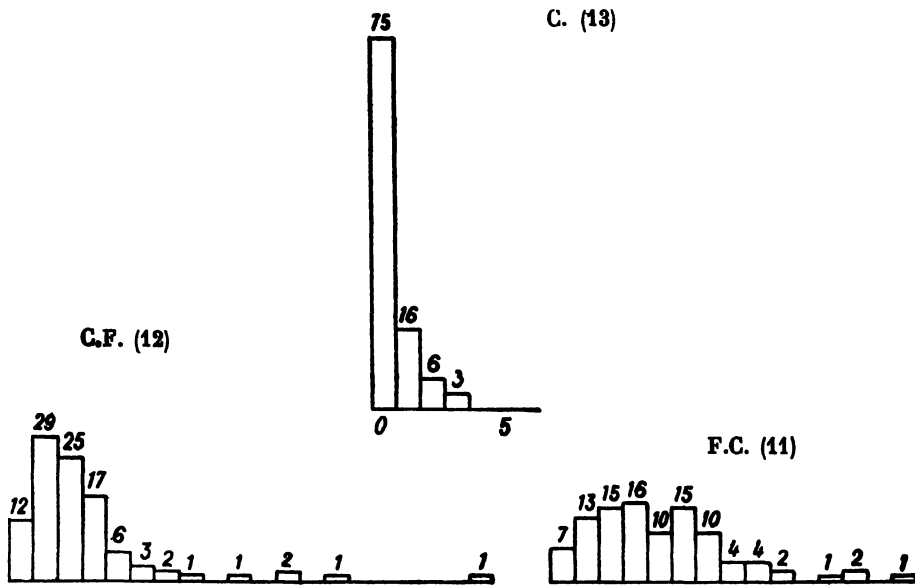


Fig. 4. — Déterminants $F.C. - C.F. - C.$

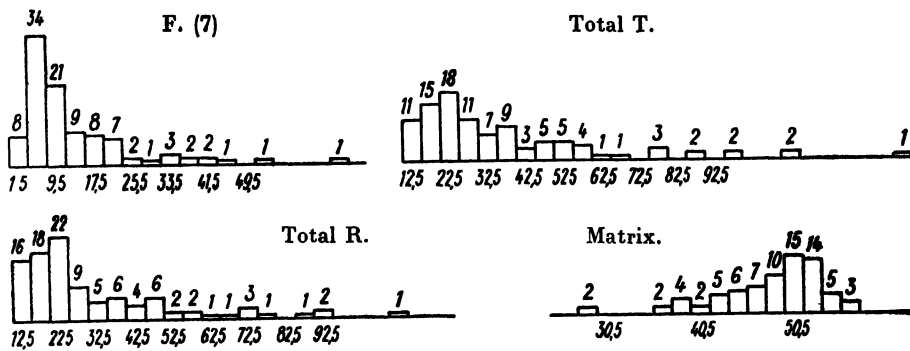


Fig. 5.

4. \bar{k} — Interprétations dans lesquelles les différences de nuances de la tache jouent un rôle et donnent au sujet une impression de tri-dimensionnalité atténuée (photographies, cartes de géographie).
5. K — Interprétation dans lesquelles les nuances de la tache jouent un rôle et donnent au sujet une impression diffuse (nuages, fumée).
6. FK — Interprétations dans lesquelles les fines nuances de la tache jouent un rôle et donnent au sujet une impression de tri-dimensionnalité réelle (paysages, châteaux).
7. F — Interprétations suscitées exclusivement par la forme de la tache.
8. Fc — Interprétations où les fines nuances de la tache jouent un rôle secondaire par rapport à la forme et donnent au sujet une impression de texture (fourrure, velours).
9. c — Interprétations où les fines nuances de la tache jouent un rôle plus important que la forme et donnent au sujet une impression de texture.
10. C' — Interprétations suscitées par les couleurs blanches et noires de la tache.
11. FC — Interprétations où les couleurs de la tache jouent un rôle secondaire par rapport à la forme.
12. CF — Interprétations où la couleur de la tache joue le rôle principal, la forme n'étant point exclue.
13. C — Interprétations suscitées exclusivement par la couleur de la tache.

Il leur était ajouté le total T des déterminants, le total R égal au total T diminué des « Additionnels » et le résultat d'un test d'intelligence : le Matrix.

Résultats

Les résultats sont rassemblés dans le tableau n° 1 qui donne pour chaque déterminant la moyenne (m), la dispersion (σ) et le coefficient de variation. ($CV = \frac{100 \sigma}{m}$). Les différences entre déterminants sont très notables tant du point de vue moyenne (variant de 0,37 pour C à 3,55 pour Fc — et 13,2 pour F, soit 40 % du total T) que du point de vue variabilité (CV variant de 81 à 198), mis à part les totaux T et R et le test Matrix (fig. 1, 2, 3, 4, 5).

Plusieurs groupements ont été envisagés parmi lesquels le groupement :

— Mo	Mouvement	1 — 2 — 3
— λ		4 — 5 — 6
— F	Forme	7
— μ		8 — 9 — 10
— Co	Couleur	11 — 12 — 13

le groupement	}	A 1-2-3-4-5 B 6-7-8 C 9-10-11-12-13	et le groupement	}	Mo + λ F Co + μ
---------------	---	---	------------------	---	-----------------------------------

Ces trois groupements conduisent aux résultats suivants :

GROUPEMENT	Mo	λ	F	μ	Co	A	B	C	Mo+λ	F	Co+μ
Moyenne . . .	6,86	4,07	18,2	5,10	6,84	8,66	18,7	8,48	10,4	18,2	12,0
Dispersion . .	5,15	3,29	12,0	4,51	5,12	6,94	14,5	6,64	8,25	12	8,64
CV	80	81	91	88	75	80	77	79	79	91	71

On voit que les groupements estompent les différences entre déterminants. Un phénomène de compensation conduit, au total, à une variabilité souvent inférieure à la plus petite du groupe et les groupements obtenus sont davantage comparables du point de vue moyenne et surtout du point de vue coefficient de variation.

Influence du nombre total de réponses

Malheureusement ces groupements ne sont pas indépendants. Quand le total des réponses est plus élevé que la moyenne, il a normalement tendance à être plus élevé pour chacun des déterminants et il en résulte normalement des corrélations positives entre les déterminants.

Autre fait plus important, un test psychologique est loin d'avoir une fidélité absolue, il dépend des dispositions momentanées du sujet. Aussi peut-on penser obtenir une stabilité plus grande des réponses à un déterminant, en rapportant ces réponses au total des réponses du sujet pour l'ensemble des déterminants. A cet effet, les rapports $\frac{Mo}{T}$, $\frac{\lambda}{T}$, $\frac{F}{T}$, $\frac{\mu}{T}$, $\frac{Co}{T}$, ont été calculés. Ils conduisent au tableau des moyennes et dispersions suivant :

	Mo/T	λ/T	F/T	μ/T	Co/T
$m \times 100$	18,3	10,8	35,7	14,9	19,9
$\sigma \times 100$	9,75	5,34	14,7	7,18	7,85
CV	58	49	41	47	39

En gros, ces rapports sont indépendants de T et les quatre rapports $\frac{Mo}{T}$, $\frac{\lambda}{T}$, $\frac{\mu}{T}$, $\frac{Co}{T}$, peuvent être tenus pour indépendants.

Les coefficients de corrélation sont, en effet, les suivants :

CORRÉLATION	λ/T	μ/T	Co/T	F/T
Mo/T	0,02	— 0,19	— 0,10	— 0,56
λ/T		— 0,11	0,05	— 0,53
μ/T			0,11	— 0,23
Co/T				— 0,49

Aucun coefficient de corrélation n'atteint 0,20 en valeur absolue (seuil de 95 % pour 100 sujets) et le χ^2 de l'ensemble (7,32) correspond à 28 chances sur 100 d'obtenir une telle valeur par simples fluctuations statistiques.

Au contraire $\frac{F}{T}$ est lié de façon étroite à $\frac{Mo}{T}, \frac{\lambda}{T}, \frac{\mu}{T}, \frac{Co}{T}$ et l'on peut se proposer en première schématisation de caractériser un sujet par le total T (ou le total R) et les rapports indépendants $\frac{Mo}{T}, \frac{\lambda}{T}, \frac{\mu}{T}, \frac{Co}{T}$ dont l'interprétation psychologique est simple.

Relations entre T et R

La différence moyenne entre T et R est de 4,9, soit en moyenne près de 15 % de T et l'écart (T — R) semble être en moyenne proportionnel à T ($\frac{T - R}{T}$ indépendant de T). Le choix de l'un ou l'autre est donc secondaire, le remplacement de l'un par l'autre se traduisant simplement par une atténuation des corrélations (dans l'hypothèse où $\frac{T - R}{T}$ est indépendant des autres éléments).

Relation entre le Rorschach et le Matrix

La corrélation entre le Matrix et les T et R est très faible. Ainsi la moyenne du Matrix par classes de valeurs de R est la suivante :

VALEUR DE R	10 — 19	20 — 29	30 — 39	40 et +
Moyenne Matrix . . .	48,4	49,2	49,9	49,6

On a pu établir d'autre part que la corrélation entre le matrix et le total T était imputable pour une très large part (80 % environ) à une corrélation entre le Matrix et le déterminant F, mais ceci aurait besoin d'être précisé sur un échantillon plus important.

Les moyennes de Matrix ont d'autre part été calculées par classes de T et de $\frac{F}{T}$ et les chiffres semblent confirmer les conclusions précédentes (corrélation entre T et le Matrix, absence de corrélation entre $\frac{F}{T}$ et le Matrix). Elles ont également été calculées par classes de $\left(\frac{Mo}{T}, \frac{\lambda}{T}, \frac{\mu}{T}, \frac{Co}{T}\right)$ mais aucune corrélation n'apparaît comme marquante.

Analyse systématique des déterminants

On a vu que le groupement des déterminants s'accompagnait d'une diminution souvent importante de la variabilité, par compensation entre les écarts en plus et en moins. Mais le groupement des déterminants pose un problème plus grave. *A priori*, certains groupements sont plus justifiés que d'autres, parce qu'ils correspondent à une réalité psychologique. Même dans ce cas cependant, on peut se demander si la simple somme de deux ou trois déterminants est la meilleure forme de groupement. Ainsi, par exemple, on calcule

parfois la somme $\frac{FC + 2 CF + 3 C}{2}$ pour la comparer à M. Quels sont les coefficients les plus adéquats? C'est une question à laquelle on répond d'une façon plus satisfaisante quand on connaît les coefficients de variation.

En effet, soit un groupement de trois déterminants A, B, C. Supposons que l'on ne considère que le total des déterminants A, B, C, le total des valeurs $X_A + X_B + X_C$ est la valeur à considérer. Supposons que l'on attache la même importance à chacun des déterminants A, B, C, une bonne mesure de l'ensemble est alors $\frac{X_A}{m_A} + \frac{X_B}{m_B} + \frac{X_C}{m_C}$ (m_A, m_B, m_C étant les moyennes des déterminants A, B, C). Enfin, si l'on ne veut pas faire d'hypothèse sur l'importance relative des trois déterminants et que seul intéresse l'écart d'ensemble des trois déterminants, la meilleure mesure est $\frac{X_A - m_A}{\sigma_A} + \frac{X_B - m_B}{\sigma_B} + \frac{X_C - m_C}{\sigma_C}$ ou $\frac{1}{CV_A} \times \frac{X_A}{m_A} + \frac{1}{CV_B} \times \frac{X_B}{m_B} + \frac{1}{CV_C} \times \frac{X_C}{m_C}$ à une constante près, qui redevient (à un facteur près) $\frac{X_A}{m_A} + \frac{X_B}{m_B} + \frac{X_C}{m_C}$ lorsque les déterminants ont la même variabilité (même coefficient de variation).

D'autre part, le graphique n° 6 donne la répartition des déterminants selon la moyenne et le coefficient de variation. Sur ce graphique, la forme pure (F) se classe à part alors que les autres déterminants semblent se répartir selon deux lignes très particulières.

On peut penser que chaque déterminant correspond à une certaine organi-

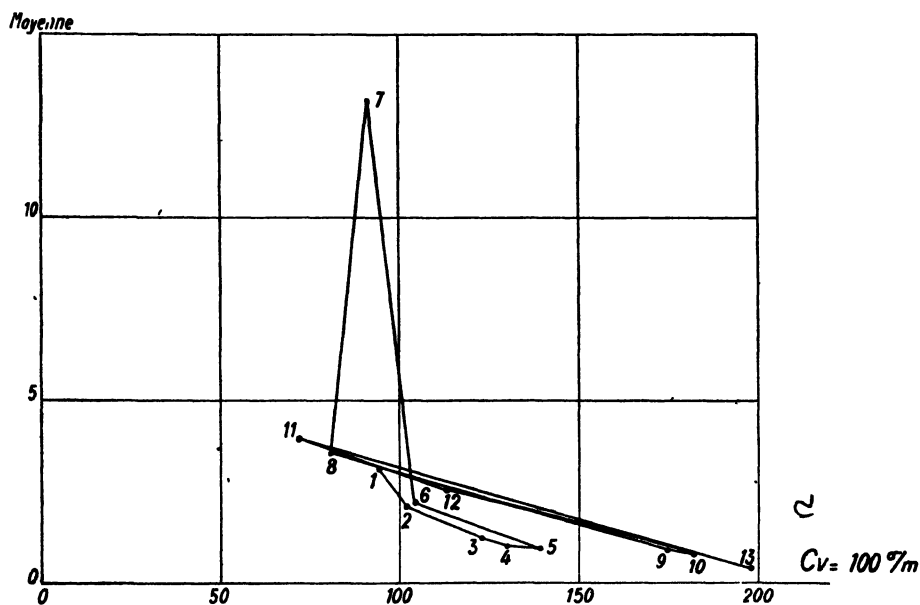


Fig. 6. — Liaisons entre la moyenne et le coefficient de variation.

sation structurelle du mental et que des organisations semblables correspondent à des moyennes et coefficients de variation reliés par une liaison fonctionnelle de la forme $\varphi(\text{moy}, CV) = 0$, des organisations structurelles différentes

