

**Marcel-Paul Schützenberger**

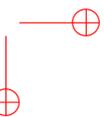
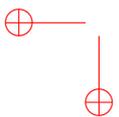
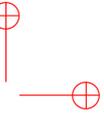
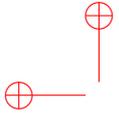
# **ŒUVRES COMPLÈTES**

éditées par  
**Jean Berstel, Alain Lascoux et Dominique Perrin**

\*

**Tome 1 : 1943–1949**

**Institut Gaspard-Monge, Université Paris-Est  
2009**



## Préface

Les treize tomes de cette édition contiennent l'ensemble des œuvres de Marcel-Paul Schützenberger qui ont fait l'objet d'une publication dans une revue scientifique, les actes d'un colloque ou un livre. Ses travaux couvrent une période de plus de 50 ans, puisque sa première note aux Comptes Rendus, *Sur la théorie des structures de Dedekind* [1943-1], date de 1943, et son dernier article, *Pour le monoïde plaxique* [1997-3], est paru en 1997.

Dans la préparation de cette édition, nous avons à chaque fois choisi la version finale telle qu'elle a été publiée. Nous avons donc écarté les versions préliminaires, les manuscrits, les épreuves annotées. Nous avons aussi réprimé notre envie de corriger des erreurs manifestes, comme celle dans le titre de *A propos du relations rationnelles fonctionnelles* [1973-2], ou de présenter une édition critique de certains travaux fondateurs : on trouvera ailleurs une nouvelle édition du livre avec Dominique Foata *Théorie géométrique des polynômes eulériens* [1970-1] et une édition critique, par Dominique Perrin, du manuscrit de la fameuse école de Royan *Codes à longueur variable* [1965-8].

La présentation est chronologique. Il peut exister une différence de plusieurs années entre la date d'un colloque et la date de publication de ses actes. La date retenue est, comme d'usage, la dernière.

En plus des travaux purement scientifiques, comme les articles de journaux scientifiques et les communications à des colloques, nous avons réuni, dans le dernier tome, un certain nombre d'écrits de nature diverse, polémiques ou de vulgarisation, comme l'article sur le darwinisme, ou les canevas de deux émissions radiophoniques avec François LeLionnais, et divers autres interviews.

Chaque tome est précédé d'une courte introduction qui essaie d'éclairer certains des travaux, à la fois par leur côté mathématique intrinsèque que par l'écho rencontré et les développements qu'ils ont suscités. Nous remercions Jacques Besson, Luc Boasson, Robert Cori, Dominique Foata, Jean-Éric Pin, Antonio Restivo, Christophe Reutenauer et Dominique Séébold qui, à des degrés divers, ont contribué à ces présentations.

**Remerciements.** L'édition présente des écrits de Marcel-Paul Schützenberger est l'aboutissement d'un projet qui a commencé à voir le jour il y a longtemps. De nombreuses personnes y ont contribué, qu'elles soient remerciées ici :

- Brigitte Cuisinier n'a pas compté son temps pour la réalisation des fichiers,
- Catherine Morge a suivi le projet et a aidé à réunir les documents de médecine,
- Geoffroy Beauquier a fait les copies à Orsay,
- Patrice Séebold a fouillé dans les archives de la bibliothèque de Montpellier,
- Antonio Restivo, Aldo de Luca, Luigi Petrone ont fourni les documents de l'ère italienne,
- Volker Strehl, Jacques Désarménien et Bernard Robinet ont interprété avec succès des photos,
- Pierre-André Picon a cherché et trouvé le volume philomatique,
- Christophe Reutenauer a donné des documents visuels et sonores,
- Paulette Le Louët a trouvé des documents d'archives,
- Dominique Foata a participé activement à l'entreprise,
- Maurice Nivat a constamment encouragé et soutenu ce projet,
- Gilles Roussel et Marie-Pierre Béal ont efficacement aidé à sa réalisation.

Merci enfin à l'Institut Gaspard-Monge et à l'Université Paris-Est pour leur contribution à ce projet. Et un grand merci, bien entendu, à Hélène Schützenberger. Elle nous a confié, et transporté jusqu'à Marne, ses immenses archives.

Jean Berstel  
Alain Lascoux  
Dominique Perrin

# Introduction

## Tome I :1943–1949

M.-P. Schützenberger a publié à cette époque une série de notes aux Comptes-Rendus sur la théorie des treillis, qu'il a étudiés sous l'impulsion d'Albert Châtelet.

Il y a aussi un premier ensemble de publications de statistique appliquée à la médecine, ou de statistique pure. Son expérience du calcul statistique lui permet, selon ses propres dires, *de remplacer, dans la mesure du possible, les longs calculs par des méthodes expéditives souvent aussi efficaces*. Durant cette période, il a fait des exposés introductifs de statistique mathématique, appliquée à la biologie ou à la psychologie.

Il soutient une thèse en médecine, intitulée *Contribution à l'étude statistique du sexe à la naissance* à Paris le 2 juillet 1948. Ce qui suit est une citation d'un texte de Jacques Besson :

La *Contribution à l'étude statistique du sexe à la naissance*, travail que l'Académie Nationale de Médecine couronna du prix du Baron Larrey, a porté sur 14 230 familles de cinq enfants ou plus, soit un total de 95 875 naissances. En voici les principales conclusions.

- Les familles où il existe un fort excès de garçons ou de filles sont nettement plus nombreuses que ne le voudrait un schéma binomial.
- L'intervalle qui sépare deux naissances consécutives de même sexe est plus court que l'intervalle de deux naissances de sexes différents.
- Tout se passe comme si l'organisme maternel réagissait au sexe du fruit de la conception assez durablement pour influencer sur le sexe de l'enfant suivant.
- Il existe donc des périodes gynophiles favorables à la naissance de filles et des périodes androphiles favorables à la naissance de garçons.

Pour nettes qu'elles soient, ces conclusions n'ont pas attiré l'intérêt des chercheurs quant aux causes des disparités. Sans doute la difficulté de pousser plus avant était-elle trop grande à l'époque, mais l'explosion de la génétique « moléculaire » devrait, sinon apporter un début de réponse, en tout cas relancer la quête sur ce qui demeure une énigme, en particulier au niveau du sex-ratio, toujours en faveur des garçons à la naissance, du moins dans le monde occidental (sauf également à Hiroshima après l'explosion atomique de 1945, mais il semble que là un début de réponse ait été avancé).

Ces recherches se poursuivirent sur les jumeaux, puis dans l'étude des dermatoglyphes d'enfants anormaux. On sait que le Professeur Turpin, convaincu de l'avenir et de l'importance de l'application à la biologie des règles et méthodes

---

## Introduction

de la statistique mathématique, lui confia le soin de diriger les thèses de doctorat en médecine de Mesdames Deroche et Pronque-Levinet et de Jérôme Lejeune. Ce dernier devait quelque temps plus tard découvrir la cause du mongolisme dans cette aberration chromosomique qu'est la trisomie 21.

Bien que technique, l'*analyse séquentielle* relève de la statistique théorique. Ses applications se retrouvent au niveau des essais cliniques, médicaments en particulier, pour lesquels le nombre de cas étudiés n'est pas fixé à l'avance. Par exemple, les patients se présentent l'un après l'autre, l'intérêt est alors de ne poursuivre le traitement que si nécessaire ou, éventuellement, de l'arrêter.

C'est Wald qui, aux États-Unis, a mis au point cette méthode, en 1947. M.-P. Schützenberger en a tout de suite saisi l'intérêt, puisque sa publication sur *Une application de l'analyse séquentielle* [1949-5] est parue dans le numéro 60 de la revue La Semaine des Hôpitaux de Paris en date du 14 août 1949.

Il est donc très probablement le premier à avoir introduit cette technique en France. Sans entrer dans les détails mathématiques et selon une démarche constante chez lui, il en a donné un exemple didactique simple, immédiatement utilisable par les fabricants de médicaments en particulier.

Il en résulte une très grande économie de moyens en temps comme en coût, en particulier lors des contrôles destructifs (pilules, ampoules, ...). Son application fut quasi immédiate chez les industriels concernés. Notons que, en mathématicien, au-delà de cet exemple premier, médical et didactique, il rend compte de l'analyse de Wald, très rigoureusement, dans sa thèse de mathématiques, soutenue un peu plus tard, en 1953.

# Années 1943–1946

## Bibliographie

- [1943-1] Marcel-Paul Schützenberger. Sur la théorie des structures de Dedekind. *C. R. Acad. Sci. Paris*, 216 :717–718, 1943. Séance du 6 janvier 1943.
- [1944-1] Marcel-Paul Schützenberger. Sur les structures de Dedekind. *C. R. Acad. Sci. Paris*, 218 :818–819, 1944.
- [1945-1] Marcel-Paul Schützenberger. Sur certains axiomes de la théorie des structures. *C. R. Acad. Sci. Paris*, 221 :218–220, 1945.
- [1946-1] Marcel-Paul Schützenberger. Méthodes typologiques en sélection professionnelle. *Biotypologie*, 8(1–1) :106–107, juillet 1946.

SÉANCE DU 31 MAI 1943.

717

THÉORIE DES GROUPES. — *Sur la théorie des structures de Dedekind.*

Note de M. MARCEL-PAUL SCHUTZENBERGER, présentée par M. Gaston Julia.

I. Un ensemble ordonné par une relation d'infériorité transitive désignée par le signe  $\subset$  est une structure au sens de Ore (cf. Glivenko, *Théorie générale des structures*; Ore, *Annals of Mathematics*, 1935 et 1936) si, pour deux êtres quelconques  $a$  et  $b$ , il existe des bornes inférieure et supérieure uniques  $a \cap b$  et  $a \cup b$ . Une structure finie a des absolus supérieur et inférieur  $\bar{\mathfrak{U}}$  et  $\bar{\Omega}$  entre lesquels sont compris tous ses êtres  $x$ .

Deux êtres comparables  $a$  et  $b$  définissent une structure quotient  $a|b$ , qui est l'ensemble des êtres compris entre  $a$  et  $b$  et les admettent par suite pour absolus inférieur et supérieur.

II. Je m'occupe dans cette Note des structures d'un nombre fini d'êtres vérifiant la condition de Dedekind  $(a \cap c) \cup (b \cap c) = c \cap [b \cup (a \cap c)]$ .

Cette condition peut être mise sous diverses formes équivalentes entraînant notamment une extension du théorème de Jordan sur les suites de composition; dans une structure quotient  $ab$  on peut intercaler entre les bornes une suite d'êtres différents  $a_i$  (de  $a_0 = a$  à  $a_h = b$ ) tels que chacun d'eux soit immédiatement inférieur au suivant, c'est-à-dire  $a_i \subset a_{i+1}$  et  $a_i|a_{i+1}$  formé des seuls êtres extrêmes. Les diverses suites ainsi obtenues ont le même nombre  $H$  d'intervalles qui est la hauteur de la structure quotient  $a|b$ .

Dans les propriétés des structures, il y a une dualité (ou corrélation) réalisée en permutant les signes  $\cap$  et  $\cup$ ,  $\subset$  et  $\supset$ .

III. J'ai été amené à définir « élément générateur inférieur » (ou supérieur), un être  $E$  qui n'a qu'un seul être  $O$  immédiatement inférieur (ou supérieur) appelé *origine*.

J'ai énoncé à leur sujet les propriétés corrélatives suivantes : les origines  $O$  de la structure quotient  $m|\bar{\mathfrak{U}}$  sont constituées par les  $m \cup O$  ( $O$  désignant une origine de la structure  $\bar{\Omega}|\bar{\mathfrak{U}}$ ).

J'appelle entière une structure qui n'a d'autre origine que les absolus, ce sont les seules à posséder au moins un inverse  $a'$  pour chacun de leurs êtres  $a$ , tel que  $a$ , tel que  $a \cap a' = \bar{\Omega}$  et  $a \cup a' = \bar{\mathfrak{U}}$ .

Toute structure quotient d'une structure entière est entière.

Le produit direct au sens habituel de ce terme de deux structures entières est entier.

IV. J'appelle *simplex* une structure (de Dedekind) entière indécomposable par produit direct; et j'ai obtenu à ce sujet les résultats suivants :

a. Toute structure quotient d'un simplex est un simplex.

b. Un simplex  $S_H$  de hauteur  $H$  est formé d'un absolu supérieur  $\mathfrak{O}$  et d'un certain nombre  $N^{H-1}/N-1$  simplex  $S_{H-1}$  ayant en commun  $N+1$  à  $N+1$  un simplex de hauteur  $H-2$ .

Dans un simplex ainsi caractérisé par sa hauteur  $H$  et sa densité  $N \geq 2$ , le nombre de structures quotient  $\Omega | x$  ou  $x | \mathfrak{O}$  de hauteur  $h$  est

$$C_H^h(N) = \prod_0^{h-1} \left( \frac{N^{H-i}-1}{N^{i+1}-1} \right).$$

Pour  $N=1$  ou  $0$ , l'on n'obtient plus des simplex, mais des structures distributives, et les  $C_H^h(1)$  sont les coefficients du binôme. Pour  $N$  premier, le simplex est isomorphe à la structure des sous-groupes d'un groupe abélien engendré par  $H$  générateurs d'ordre  $N$ .

Je n'ai pu établir l'existence de simplex pour des valeurs quelconques de la hauteur et de la densité, mais j'ai pu montrer que toute structure entière finie de Dedekind était un produit direct de simplex.

Ces propriétés se généralisent pour des structures non entières.

ALGÈBRE. — *Sur les structures de Dedekind.*

Note (1) de M. MARCEL-PAUL SCHÜTZENBERGER.

Je me propose d'appliquer aux structures d'ordre quelconque les résultats précédemment signalés (2).

Pour tout élément  $a$  de  $\Omega/\mathfrak{I}$  il existe un *nœud supérieur*  $a^0$ , borne inférieure des  $x$  tels que  $a/x$  soit entière, et, corrélativement, un *nœud inférieur*  $a_0$  qui est en général différent de  $(a^0)_0$ , mais il existe les relations

$$((a^0)_0)^0 = a^0, \quad ((a_0)_0)^0 = a_0.$$

Les structures entières  $(a^0)_0/a^0$  et  $a_0/(a_0)_0$  sont les mailles *supérieure* et *inférieure* de  $\Omega/\mathfrak{I}$  attachées à  $a$ . L'ensemble des mailles *inférieures* est identique à l'ensemble des mailles *supérieures* attachées à tous les éléments de  $\Omega/\mathfrak{I}$ .

J'ai démontré que l'on pouvait obtenir d'une manière unique, à l'ordre près des opérations, toutes les structures de Dedekind, à partir des *simplex*, en introduisant dans une structure d'ordre inférieure, à partir d'une nouvelle origine convenable, des éléments générateurs, et les éléments nécessités par les axiomes et la condition de Birkhoff. L'on passe ainsi d'une structure à une autre d'un ordre plus élevé par une ou plusieurs des quatre opérations (*dilatations*) corrélatives suivantes effectuées à partir d'une origine nouvelle  $a$  quelconque.

1° Produit direct généralisé : si  $a/a^0 = \Sigma$  l'on peut obtenir  $a/(a^0)' = \Sigma \times \Sigma'$  ( $\Sigma$  et  $\Sigma'$  étant des structures de Dedekind quelconques); cette opération est valable d'ailleurs pour n'importe quelle structure.

2° Dilatation simpliciale : si  $a/a^0 = \Sigma \times S^H$  ( $S^H$  simplex) l'on peut obtenir  $a/(a^0)' = \Sigma \times S^{H'}$  avec  $H' > H$ .

3° Si  $a/a^0 = \Sigma \times S^H(n) \times S^{H'}(n)$  et si  $S^H$  et  $S^{H'}$  sont *induits* en  $a$  par deux dilatations précédentes l'on peut obtenir  $a/(a^0)' = \Sigma \times S^{H+H'}(n)$ .

4° Sous des conditions analogues et plus restrictives, si  $a/a^0 = \Sigma \times S^1 \times S^1$ , l'on peut obtenir  $a/(a^0)' = \Sigma \times S^2(n)$ ; ( $n$  quelconque).

Réciproquement toute structure obtenue au moyen des dilatations 1°, 2° ou 3° peut être complétée en une structure *entière* et de *même hauteur*.

Signalons une particularité des structures complètes de diviseurs d'entiers

(1) Séance du 15 mai 1944.

(2) *Comptes rendus*, 216, 1943, p. 717.

( 2 )

et de sous-groupes de groupe abélien. Toute *origine* est l'*absolu* ou un *élément générateur*.

Dans le cas général, la construction précédente m'a permis d'établir le théorème suivant :

*Dans toute structure finie de Dedekind, le nombre des éléments générateurs inférieurs est égal au nombre des éléments générateurs supérieurs.*

Ce nombre est aussi dans les *structures distributives* toujours obtenues par la dilatation  $r^\circ$ , égal à la *hauteur* et au nombre de valeurs *indépendantes* dont on dispose pour *normer* la structure.

On en déduit aussi des extensions, des théorèmes de Remack (signalées par Birkhoff, *Proc. of the Cambr. Ph. Soc.*, 1933) et de Zassenhaus, où l'isomorphie doit être entendue au sens restreint d'isomorphie structurelle.

1° THÉORÈME DE REMACK. — *a, b et c étant trois éléments quelconques d'une structure de Dedekind, les trois structures quotients  $(a \amalg b) \amalg (a \amalg c) / a \amalg (b \amalg c)$ ,  $(b \amalg c) \amalg (b \amalg a) / b \amalg (a \amalg c)$  et  $(c \amalg a) \amalg (c \amalg b) / c \amalg (a \amalg b)$  sont isomorphes.*

2° THÉORÈME DE ZASSENHAUS. — *a, a', b et b' étant quatre éléments quelconques d'une structure de Dedekind, les trois structures quotients*

$$(a \amalg a') \amalg (a \amalg b \amalg b') / (a \amalg a') \amalg (a \amalg b), \quad (b \amalg b') \amalg (a \amalg a' \amalg b) / (b \amalg b') \amalg (a \amalg b)$$

et

$$(a \amalg a' \amalg b') \amalg (a \amalg b \amalg b') / (a \amalg b)$$

sont isomorphes.

(Extrait des *Comptes rendus des séances de l'Académie des Sciences*,  
t. 218, p. 818-819, séance du 22 mai 1944.)

Dépôt légal d'éditeur. — 1944. — N° d'ordre 25.

Dépôt légal d'imprimeur. — 1944. — N° d'ordre 42.

GAUTHIER-VILLARS, IMPRIMEUR-LIBRAIRE DES COMPTES RENDUS DES SÉANCES DE L'ACADÉMIE DES SCIENCES.  
121195-44 Paris. — Ouv. des Grands-Augustins, 55.

## CORRESPONDANCE.

THÉORIE DES NOMBRES. — *Sur certains axiomes de la théorie des structures.* Note (\*) de M. MAURICE-PAUL SCHÜTZENBERGER, présentée par M. Gaston Julia.

Nous indiquerons dans cette Note un certain nombre de résultats relatifs aux *lois universelles* en théorie des treillis (structures, lattices), c'est-à-dire aux lois définies par une ou plusieurs égalités pour toutes valeurs des variables entre deux polynômes en  $\cap$  et  $\cup$  (que nous remplaçons par  $\cdot$  et  $+$ ). A chaque loi  $\mathcal{U}_z$  correspond pour toute valeur de  $K$  une structure  $\mathfrak{M}(K)$  homomorphe, au moyen des relations d'équivalence fournies par  $\mathcal{U}_z$  au treillis libre  $\mathfrak{M}_z(K)$  (cf. Whitmann, *Ann. of Mathematics*, 1941 et 1942) engendré par  $K$  générateurs (treillis de rang  $K$ ). Réciproquement, à tout treillis fini  $\mathfrak{u}_z$  il est possible d'attacher une loi  $\mathcal{U}_z$  unique, plus forte que toutes les autres lois compatibles avec  $\mathfrak{u}_z$  et telle que  $\mathfrak{M}(K)$  soit finie pour toute valeur finie de  $K$ .

1° La loi distributive  $\mathcal{U}_0 \ni a(b+c) = ab+ac$  est la plus forte de toutes les lois universelles non dégénérées. En effet toute loi  $\mathcal{U}_z$  qui n'est pas plus faible que la loi  $\mathcal{U}_0$  entraîne l'égalité de tous les éléments de  $\mathfrak{M}_z(K)$ . L'on pourrait se servir de ceci pour démontrer que, si un polynôme de treillis libre  $\mathfrak{M}_0(K)$  a pour valeur  $\mathfrak{M}_0(K)$  un monome, il a effectivement cette valeur dans  $\mathfrak{M}_0(K)$ .

2° Il existe deux et seulement deux lois  $\mathcal{U}_1$  et  $\mathcal{U}_2$  immédiatement plus faibles que  $\mathcal{U}_0$  :

$\mathcal{U}_1$  est définie par

$$a_1[a_2 + a_3(a_4 + a_5)] = a_1[a_2 a_4 + a_3 a_5] + a_1[a_2 + a_1 a_3 a_4] + a_1[a_2 a_4 + a_3 a_5] + a_1[a_2 + a_1 a_3 a_5] + a_1 a_3[a_2 a_4 a_5 + a_3] + a_1 a_3[a_2 + a_1 a_3 a_4]$$

son *sous-treillis typique* est réalisé par les 5 éléments

$$\Omega = a_1 a_2, \quad x = a_2, \quad y = a_1 a_2 + a_1 a_3, \quad z = a_1(a_2 + a_1 a_3), \quad \mathcal{V} = a_2 + a_1 a_3$$

entraînant

$$xz = \Omega \subset y \subset z \subset \mathcal{V} = x + y, \quad \Omega \subset x \subset \mathcal{V}$$

dans  $\mathfrak{M}_0(3)$ , la loi  $\mathcal{U}_2$

$$a(b+ac) = ab+ac$$

(\*) Séance du 23 juillet 1945.

SÉANCE DU 20 AOUT 1945.

219

et

$$a_1[a_2 + a_3(a_4 + a_5)] = a_1(a_2 + a_3a_4) + a_1(a_2 + a_3a_5) + a_1a_3(a_4 + a_5) + a_1a_3(a_4 + a_5a_6) + a_1a_3(a_5 + a_2a_4)$$

est satisfaite par les éléments du treillis de Dedekind  $\mathfrak{U}_d(3)$ . Ces deux lois donnent des treillis  $\mathfrak{U}_1(\mathbb{K})$  et  $\mathfrak{U}_2(\mathbb{K})$  finis si leur rang  $\mathbb{K}$  est fini.

Il n'en est pas de même en ce qui concerne  $\mathfrak{U}_d(\mathbb{K})$ , déduite de la loi de Dedekind  $\mathfrak{U}_d$

$$a_1[a_2 + a_1a_3] = a_1a_2 + a_1a_3,$$

qui pour  $\mathbb{K} \cong 4$  est non seulement infini, mais a ses chaînes infinies comme, l'on peut l'établir facilement en considérant les expressions

$$\tilde{z}_{i+1} = [(\tilde{z}_i + a_1)a_2 + a_3]a_4, \quad \tilde{z}_{i+1} \supset \tilde{z}_i, \quad \tilde{z}_{i+1} \not\subseteq \tilde{z}_i$$

(par réalisation convenable dans un espace projectif).

Signalons ici que toute loi strictement plus faible que  $\mathfrak{U}_d$ , par exemple

$$\mathfrak{U}_l : a_1a_2 + a_1a_3 = a_1(a_2 + a_1a_3)(a_3 + a_1a_2),$$

est nécessairement plus faible que  $\mathfrak{U}_1$  et que toute loi plus forte que  $\mathfrak{U}_d$  et différente de  $\mathfrak{U}_d$  est  $\mathfrak{U}_2$  ou plus faible que  $\mathfrak{U}_2$ .

3° Parmi les lois comprises entre  $\mathfrak{U}_d$  et  $\mathfrak{U}_2$  signalons les deux lois intéressantes  $\mathfrak{U}_l$  et  $\mathfrak{U}_L$ , obtenues en adjoignant à  $\mathfrak{U}_d$  les égalités  $\mathfrak{U}'_l$  et  $\mathfrak{U}'_L$ ; posons

$$\begin{aligned} M_{12} = M_{21} = M_{34} = a_1a_2 + a_3a_4; & \quad F_{12} = F_{21} = a_1a_2 + M_{13}M_{14}, \\ \mathfrak{U}'_l = F_{13}F_{23} + F_{23}F_{31} + F_{31}F_{12} = F_{12}F_{23}F_{31} = A_4, \\ a_2a_3 + a_3a_6 = B'_1, & \quad B'_1 + B'_2B'_3 = B_1; \\ a_3a_4 + a_6a_4 = B'_2, & \quad B'_2 + B'_3B'_1 = B_2; \\ a_1a_2 + a_4a_6 = B'_3, & \quad B'_3 + B'_1B'_2 = B_3; \end{aligned}$$

par exemple

$$\mathfrak{U}_L \ni \alpha = a_1a_4 + a_2a_3 + a_3a_6 = (a_2B_2 + a_3B_3)(a_5B_2 + a_6B_3) + \alpha;$$

$\mathfrak{U}_L$  assure dans tous les cas et  $\mathfrak{U}_l$  dans les cas de  $\mathbb{K} \cong 4$  (cas de Möbius) la validité du théorème de Desargues. L'on peut montrer (Veblen) que  $\mathfrak{U}_L$  est plus fort que  $\mathfrak{U}_d$ , mais nous n'avons pu trouver d'exemple effectif de treillis satisfaisant à  $\mathfrak{U}_L$  sans satisfaire à  $\mathfrak{U}_1$ .

Les lois  $\mathfrak{U}_{m(p)}$ , exprimant que les géométries projectives que l'on peut déduire des treillis  $\mathfrak{U}_{m(p)}(\mathbb{K})$  ont pour corps de coefficient un corps modulo  $p$ , s'obtiennent en remarquant que les 4 expressions  $A_i$  se comportent entre elles comme 4 points d'un plan quand  $\mathfrak{U}_l$  est satisfaite. Pour les premières valeurs de  $p$  nous avons trouvé des expressions plus simples; citons

$$\begin{aligned} M_{12}M_{13} + M_{13}M_{14} + M_{14}M_{12} = M_{12}M_{13}M_{14} & \quad (\rightarrow \text{modulo } 2), \\ a_1F_{23} = a_1F_{34} = a_1F_{42} & \quad (\rightarrow \text{modulo } 3); \end{aligned}$$

ces relations sont des identités dans les groupes abéliens principaux d'ordre  $p^n$ .

Il existe aussi un ensemble de lois  $\mathfrak{U}_{n(i)} = \mathfrak{U}_d$  et  $\mathfrak{U}'_{n(i)}$ , que l'on eût utiliser

INSTITUT HENRI POINCARÉ

pour limiter le nombre de dimensions des espaces projectifs. On peut leur donner, par exemple, la forme simple ( $d = h - 2$ )

$$\mathfrak{U}_{h(h)} \Leftrightarrow a_0 \left[ \sum_{i=1}^{h-1} a_i \right] = \sum_{i=1}^{h-1} a_0 \sum_{\substack{j=1 \\ j \neq i}}^{h-1} a_j.$$

Remarquons que les points et les droites du complexe linéaire satisfont à  $\mathfrak{U}_{h(h)}$ .

Toute loi plus forte que  $\mathfrak{U}_h$  est nécessairement plus forte qu'une loi  $\mathfrak{U}_m$  ou  $\mathfrak{U}_n$ .

# BIOTYPOLOGIE

Revue de la Société de Biotypologie  
Vu Redaction

## Quelques problèmes généraux de typologie

### SOMMAIRE

#### I. COMMUNICATION ET MÉMOIRES ORIGINAUX

H. PIÉAPORTE. — Question des types psycho-physiologiques... P. des types.....	1
M. J. DUBLINEAU. — La résistance biologique en... typologique.....	24
M. HIRSCH. — Conceptions générales d'une école française de morphologie. Une application pratique.....	31
C. BURT. — L'étude factorielle des types physiques.....	42
† D. WEINBERG. — Contrôle expérimental des méthodes d'ana- lyse factorielle.....	50
Dr BÉNASSY et C. CHAUFFARD. — Comment vérifier la validité d'un test de tempérament.....	75

#### II. NOTES ET INFORMATIONS AVR 1947

Etudes de Sheldon sur les types physiques et les types de tem- pérament.....	80
Rapprochement entre les typologies de différents auteurs.....	90
Une thèse sur la typologie des adolescents.....	101
Une nouvelle typologie des adolescents criminels.....	104
Méthodes typologiques en sélection professionnelle.....	106
• • •	
Conférence internationale de génétique.....	107
Activité de la Société de Biotypologie.....	111
Nécrologie.....	114

#### III. BIBLIOGRAPHIE

(Voir la suite du sommaire page 4 de la couverture.)

HERMANN & C<sup>ie</sup>, ÉDITEURS  
6, RUE DE LA SORBONNE, PARIS

Publication périodique trimestrielle

vol nous apparait comme une activité régressive par rapport au niveau que nous attendons d'un sujet normal. Qu'il entre dans une conduite infantile ou primitive, ou qu'il soit la réaction inadaptée à un conflit affectif, la métaphore s'impose de le situer « en arrière » d'une conduite normale.

Chez le pervers, la notion de régression est encore plus actuelle : le sujet veut être délinquant, il veut avoir un rôle antisocial et antimoral. Nous avons suffisamment insisté au cours de notre étude sur le double critère qui nous permet de définir un pervers; il assume sa « régression » et se montre imperméable à la conscience commune.

Pour mieux opposer ces différents types, formulons enfin brièvement leurs attitudes respectives devant le délit.

Le voleur infantile joue, et se laisse suggérer la manière de voir de toute personne qui sait l'influencer.

Le primitif n'assume ni la propriété ni la culpabilité du vol : il accuse les autres ou les circonstances et ignore le sentiment de la faute.

Le névrotique en assume la culpabilité, souvent il s'y complait ; suivant la nature du conflit et son type de névrose il en assume la propriété ou la rejette sur les circonstances. C'est le type le plus variable et le plus mobile.

Le pervers en assume la propriété, non sans vanité, et en rejette la culpabilité en niant toute faute. C'est le type le plus fixe et le plus conscient de lui. »

D. W.

---

## MÉTHODES TYPOLOGIQUES EN SÉLECTION PROFESSIONNELLE

---

### NOTE SUR UN PROJET DE RECHERCHE.

A la suite des études récentes sur la typologie on a pensé qu'il pourrait être intéressant d'en essayer les méthodes à la sélection professionnelle.

NOTES ET INFORMATIONS

107

En effet, et bien qu'aucune recherche précise n'ait encore été effectuée, il semble qu'il n'existe pas un type unique de « bon ouvrier ». Autrement dit, étant donné un lot de profils (comprenant d'une part des notations professionnelles non soumises à l'effet de halo avec trop d'importance et d'autre part des mesures somatiques et psychiques), ceux-ci se grouperaient dans certains cas selon plusieurs types.

Plus précisément encore, en employant les méthodes de corrélations entre sujets on constaterait que la matrice est multifactorielle et non hiérarchique comme on le suppose toujours pour calculer la note stannine.

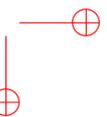
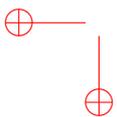
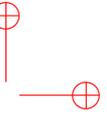
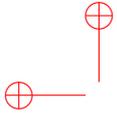
On se propose donc, soit en calculant les corrélations entre sujets, soit en évaluant l'écart des profils par  $\Sigma(x_i - y_i)^2$ , et en utilisant ensuite la méthode de Tryon, de vérifier les hypothèses précédentes.

Si on pouvait effectivement mettre en évidence plusieurs « facteurs sujets » il serait alors possible, d'effectuer la sélection en se référant aux profils types de « bon ouvrier » dans un poste de travail donné.

Parmi les avantages que présenterait cette méthode, il faut noter la possibilité de l'utiliser avec moins de risques d'erreurs dans le cas, fréquent dans la pratique, où le nombre des sujets de contrôle n'atteint pas la centaine et où les erreurs à craindre sur les coefficients de régression sont particulièrement élevées.

M. SCHÜTZENBERGER.

---



# Année 1947

## Bibliographie

- [1947-1] Marcel-Paul Schützenberger. Remarques sur la notion de clivage dans les structures algébriques et son application aux treillis. *C. R. Acad. Sci. Paris*, 224 :512–514, 1947. Séance du 20 janvier 1947.
- [1947-2] Marcel-Paul Schützenberger. Sur certains treillis gauches. *C. R. Acad. Sci. Paris*, 224 :776–778, 1947.
- [1947-3] Marcel-Paul Schützenberger. Remarques sur des relations d’ordre entre variables aléatoires indépendantes. *C. R. Acad. Sci. Paris*, 224 :878–880, 1947.
- [1947-4] Pierre Gavaudan, Hélène Poussel et Marcel-Paul Schützenberger. L’excitation des chimiorécepteurs de la langue par les substances du groupe des narcotiques indifférents et la règle thermodynamique de la narcose. *C. R. Acad. Sci. Paris*, 224 :1525–1527, 1947.
- [1947-5] Marcel-Paul Schützenberger. Sur certains paramètres caractéristiques des systèmes d’événements compatibles et dépendants et leur application au calcul des cumulants de la répétition. *C. R. Acad. Sci. Paris*, 225 :277–278, 1947. Séance du 28 juillet 1947.
- [1947-6] Marcel-Paul Schützenberger. Axiomatisation de la géométrie dans un complexe linéaire de droites. *Revue Sci.*, 85e année(Fasc. 13) :782–784, 1er août 1947. No 3278.

512

ACADÉMIE DES SCIENCES.

## CORRESPONDANCE.

L'ACADÉMIE AUTRICHIENNE DES SCIENCES invite l'Académie à se faire représenter aux Cérémonies qui auront lieu à Vienne, du 11 au 16 mai 1947, à l'occasion du centième anniversaire de sa fondation.

MM. **I. M. VINOGRADOV**, élu Correspondant pour la Section de Géométrie, **WALTER DAVIS LAMBERT**, élu Correspondant pour la Section de Géographie et Navigation, et **DAVID KEILIN**, élu Correspondant pour la Section d'Anatomie et Zoologie, adressent leurs remerciements à l'Académie.

**M. RENÉ BOURRET** adresse des remerciements à l'Académie pour la distinction accordée à ses travaux.

**M. K. SINELNIKOV**, Directeur de l'INSTITUT PHYSICO-TECHNIQUE DE L'ACADÉMIE DES SCIENCES DE L'UKRAINE, adresse les condoléances de cet Institut à l'occasion de la mort de *M. Paul Langevin*.

M. le **SECRETÉIRE PERPÉTUEL** signale parmi les pièces imprimées de la Correspondance :

**N. B. MEDVEDEVA**. *Endocrinologie expérimentale* (en langue russe).

ALGÈBRE. — *Remarques sur la notion de clivage dans les structures algébriques et son application aux treillis*. Note <sup>(1)</sup> de **M. MAURICE PAUL SCHÜTZENBERGER**, présentée par **M. Gaston Julia**.

On rappellera quelques résultats dont certains ont déjà été exposés dans une Note antérieure <sup>(2)</sup> que l'on complétera et systématisera à l'aide d'un lemme très simple de la théorie générale des relations d'équivalence dans les structures algébriques. Enfin on appliquera ces résultats au domaine particulier des treillis.

*Lois universelles*. — Étant donnée une structure algébrique  $S$ , on appellera loi universelle de  $S$  l'ensemble  $U$ ,  $U = \mathfrak{U}(S)$  des égalités entre polynômes généralisés identiquement vérifiées quand les variables prennent indépendamment toutes les valeurs possibles dans  $S$ . Réciproquement, on appellera *structure libre à  $k$  générateurs de  $U$*  la structure  $S' = \mathcal{L}_k(U)$ , isomorphe au quotient de la structure des polynômes à  $k$  variables par les relations de  $U$ . On appellera *noyau de  $U$* , quand elle existe, la sous-structure  $S''$  de  $\mathcal{L}(U)$ ,  $\mathfrak{U}(U)$  invariante

<sup>(1)</sup> Séance du 20 janvier 1947.

<sup>(2)</sup> *Comptes rendus*, 221, 1945, p. 218.

SÉANCE DU 24 FÉVRIER 1947.

513

par toutes les substitutions de polynômes à des variables qui laissent  $U$  invariant. Par définition,  $\mathfrak{U}[\mathcal{G}(U)] = U$ .

Parmi les lois universelles, on distinguera les *lois simples* définies par une égalité unique dans l'un des membres de laquelle chaque variable ne figure qu'une fois. Il est à remarquer que, lorsque toutes les opérations possèdent une inverse, toute loi universelle peut s'exprimer par la composition de lois simples. Il en est d'ailleurs de même pour les treillis.

Ordonnées par la relation  $U_1 \supset U_2$  ( $U_1$  plus fort que  $U_2$ ,  $U_2$  plus faible que  $U_1$ ), les lois universelles forment un treillis *fermé* pour les deux opérations  $(.)$  et  $(+)$ ; si  $S_1$  et  $S_2$  satisfont à  $U$ , il en est de même du produit direct  $S_1 \times S_2$  et réciproquement,

$$\mathfrak{U}(S_1) + \mathfrak{U}(S_2) = \mathfrak{U}[\mathcal{G}(S_1) \times \mathcal{G}(S_2)].$$

**LEMME DE CLIVAGE.** — Si dans une structure algébrique libre  $S$  les polynômes  $\pi_i$  constituent une sous-structure  $S'$ , on a nécessairement pour toute loi universelle  $U$  plus forte que  $\mathfrak{U}(S)$ .

- 1° soit  $U$  plus faible que  $\mathfrak{U}(S')$ ;
- 2° soit  $U$  plus fort que l'une des lois  $u_{ij}$  obtenues en posant  $\pi_i = \pi_j$  pour toutes valeurs des variables.

Généralisant une notion de Whitman (\*), on dira que le treillis des lois universelles plus fortes que  $\mathfrak{U}(S)$  est *clivé* entre  $\mathfrak{U}(S')$  et l'ensemble des  $u_{ij}$ .

En effet, si 1° n'était pas vérifié, il y aurait au moins deux être  $\pi_i$  et  $\pi_j$  égaux, d'où 2°, et d'autre part 1° et 2° ne peuvent être simultanément vrais car  $u_{ij} \cdot \mathfrak{U}(S')$  est différent de  $\mathfrak{U}(S')$ .

Donc  $\mathfrak{U}(S')$  est un élément irréductible pour  $(+)$  du treillis des  $U$  plus forts que  $\mathfrak{U}(S)$ , et de même pour  $(.)$  l'élément le plus fort des  $u_{ij}$ , quand il existe. Par exemple, le treillis des lois universelles de la structure de groupe est clivé entre la loi  $xy = yx$  et l'ensemble des lois  $x_p = 1$  ( $p$  premier), car le groupe formé des puissances d'un seul élément est abélien.

*Applications à la structure algébrique de treillis commutatif.* — On peut, ici, préciser encore la notion de clivage en faisant usage du théorème simple suivant :

Si, dans  $\mathcal{L}(U)$  le polynôme  $\pi_i$  est irréductible, alors, pour toute loi universelle  $U'$  plus faible que  $U$ , il existe un polynôme  $\pi'_i$  irréductible dans  $\mathcal{L}(U')$  et tel que  $\pi_i = \pi'_i$  dans  $\mathcal{L}(U)$ .

On peut alors énoncer :

$\mathcal{G}(U')$  peut être réalisé comme sous-structure dans toute  $\mathcal{L}(U' + U'')$ , s'il est réalisé *tautologiquement* dans  $\mathcal{L}(U' \cdot U'')$ , donc nécessairement si  $U''$  n'est pas plus faible que  $U'$ . Par conséquent, dans ces conditions [vérifiées notamment

(\*) *Splittings of a lattice* (Am. J. of Math., 65, 1943, pp. 79-196).

514

ACADÉMIE DES SCIENCES.

quand  $\mathcal{N}(U')$  est engendré par des  $\pi_i$  tels que  $\pi_i = \pi_j$  dans  $\mathcal{L}(U'')$ , le treillis des lois plus fortes que  $U' + U''$  est clivé entre  $U'$  et l'ensemble des lois  $\pi'_i = \pi'_j$ .

En application on citera :

- 1° Le clivage entre la loi triviale  $a = b$  et la loi distributive;
- 2° Le clivage dérivé du treillis de Birkhoff à 5 éléments non dédékindien  $A$  entre la loi de Dedekind et  $\mathfrak{A}(A)$ ;
- 3° Les clivages dérivés des treillis duals de Löwig (\*)  $B$  et  $\bar{B}$  entre la loi  $a(ab + bc + ca) = ab + ab$  et  $\mathfrak{A}(B)$  et de manière duale;
- 4° Le clivage dérivé du treillis à 10 éléments engendré par  $a, b$ , et  $a(b + c) + b(a + c)$  entre la loi  $[a + b(a + c)][b + a(b + c)] = a(b + c) + b(a + c)$  (que le clivage même permet de montrer équivalente à sa forme duale) et  $\mathfrak{A}(C)$ ;
- 5° Dans le treillis des lois universelles plus fortes que la loi de Dedekind, le clivage entre la loi distributive et la loi du treillis dédékindien non distributif à 5 éléments.

CALCUL DES PROBABILITÉS. — *Sur une généralisation d'un théorème de M. B. de Finetti et son application à la théorie collective du risque.* Note (\*) de M. J. DUBOURDIEU, présentée par M. Émile Borel.

Soient  $X_1, X_2, \dots, X_k, \dots$  les gains respectifs d'un joueur  $A$  dans une suite de paris successifs  $P_1, P_2, \dots, P_k, \dots$ . Et soit  $K$  la fortune initiale de  $A$ . Nous désignerons par  $U_k$  la variable aléatoire égale au gain algébrique  $X_k$  de  $A$  dans la partie  $P_k$ , si celle-ci est jouée (ce qui suppose que  $A$  n'ait pas été ruiné par les parties antérieurement jouées), et qui est nulle dans le cas contraire. Et posant  $S_n = K + U_1 + U_2 + \dots + U_n$ , nous désignerons par  $F_n(x)$  la fonction des probabilités totales de  $S_n$ . Pour que  $A$  soit ruiné en au plus  $n$  parties, il faut et il suffit que  $S_n$  soit  $\leq 0$ , et la probabilité de cette éventualité est  $\Pi_n = \Pr\{S_n \leq 0\} = F_n(+0)$ . Soit d'autre part  $R(x)$  une fonction croissante de  $x$ . On a

$$\mathcal{N}\{e^{-R(S_n)}\} \geq \int_{-x}^0 e^{-R(x)} dF_n(x) \geq e^{-R(0)} F_n(+0).$$

D'où l'inégalité

$$(1) \quad \Pi_n = F_n(+0) \leq e^{R(0)} \mathcal{N}\{e^{-R(S_n)}\}.$$

D'autre part, comme  $S_n = S_{n-1} + U_n$ , il vient

$$(2) \quad \mathcal{N}\{e^{-R(S_n)}\} = \mathcal{N}\{e^{-R(S_{n-1})}\} \mathcal{N}_{S_{n-1}}[e^{R(S_{n-1}) - R(S_{n-1} + U_n)}],$$

identité dans laquelle  $\mathcal{N}_{S_{n-1}}$  désigne la valeur probable liée à  $S_{n-1}$ .

Ceci dit, supposons que notre joueur ait la faculté de choisir chaque nouvelle

(\*) *Ann. of Math.*, 42, 1943, pp. 573-579.

(1) Séance du 6 janvier 1947.

776

ACADÉMIE DES SCIENCES.

## PRÉSENTATIONS.

Dans la formation d'une liste de candidats à la Chaire de Théorie des équations différentielles et fonctionnelles du Collège de France, pour la première ligne, M. *Jean Leray* obtient 41 suffrages; il y a deux bulletins blancs.

Pour la seconde ligne, M. *André Lichnerowicz* obtient 30 suffrages; il y a deux bulletins blancs.

En conséquence, la liste présentée à M. le Ministre de l'Éducation Nationale comprendra :

*En première ligne*..... M. JEAN LERAY.

*En seconde ligne*..... M. ANDRÉ LICHNEROWICZ.

## CORRESPONDANCE.

Sir EDWARD JOHN RUSSELL, élu Associé étranger, et M. THÉODORE DE KARMÁN, élu Correspondant pour la Section de Mécanique, adressent leurs remerciements à l'Académie.

M. le SECRÉTAIRE PERPÉTUEL signale parmi les pièces imprimées de la Correspondance :

Union géodésique et géophysique internationale. Comité National français. Section d'Hydrologie scientifique. *Commission du Bassin de la Seine*. Cahier n° 6. *Niveau journalier de la Seine à Vernon (1885-1940)*.

ALGÈBRE. — *Sur certains treillis gauches*. Note (1) de M. MAURICE-PAUL SCHÜTZENBERGER, transmise par M. Gaston Julia.

On étudiera brièvement ici certains treillis non commutatifs. Des résultats déjà connus (KLEIN BARMEN, *Math. Zeitschrift*, 1940, pp. 41, 42, 43) seront approfondis.

*Axiomes des treillis gauches*. — Ce sont des structures algébriques où existe une loi de composition interne partout définie, associative, strictement idempotente ( $aa = a$ ). On examinera ultérieurement la loi plus faible : pour tout  $a$ , il existe un entier  $n$  tel que  $a^n = a$ . On postulera en outre l'existence de  $\gamma$  (pour tout  $a : a\gamma = \gamma a = a$ ) et de  $\Omega$  (pour tout  $a : a\Omega = \Omega a = \Omega$ ).

*Inégalité et isotopie* — On écrira

$$\begin{array}{lll} a \dot{\subset} b & \text{équivalent à} & ab = a; \\ a \subset b & \text{équivalent à} & ba = a; \\ a \subset b & \text{équivalent à} & ab = ba = a. \end{array}$$

(1) Séance du 6 janvier 1947.

SÉANCE DU 17 MARS 1947.

777

Cette relation est *transitive* et a une *réflexivité presque complète (croisée)*. On a en outre, pour tout  $c$ ,

$$\begin{array}{ll} \text{si } a \dot{\subset} b, & \text{alors } ca \dot{\subset} b; \\ \text{si } a \dot{\subset} b, & \text{alors } ac \dot{\subset} b. \end{array}$$

On écrira

$$\begin{array}{ll} a \dot{=} b & \text{équivalent à } a \dot{\subset} b \text{ et } b \dot{\subset} a; \\ a \ddot{=} b & \text{équivalent à } a \dot{\subset} b \text{ et } b \dot{\subset} a. \end{array}$$

$\dot{=}$  sera appelée *isotopie*; si  $a \neq b$ , on dira qu'ils forment un système de Klein Barmen (par exemple  $ba$  et  $aba : ba \cdot aba = ba$ ;  $aba \cdot ba = aba$ ). L'isotopie est *symétrique, transitive, presque complètement réflexive (croisée)*. Si les  $a_i$  sont tels que  $a_i \dot{\subset} a_{i+1}$  et  $a_n \dot{\subset} a_0$ , ils sont isotopes (et de même en changeant les symboles en leurs opposés).

L'isotopie détermine une partition en classes. On écrira

$$x \in |\dot{a}| \text{ équivalent à } a \dot{=} x; \quad x \in |a| \text{ équivalent à } a \ddot{=} x.$$

- 1° Tout être appartient au moins à une classe de chaque espèce.
- 2° Les classes de même espèce sont deux à deux exclusives :

$$\text{Si } x \in |\dot{a}| \text{ et } x \in |\dot{b}|, \text{ alors } a = ax \text{ et } x = xb,$$

d'où

$$ab = axb = ax = a,$$

donc

$$a \in |\dot{b}| \text{ et } b \in |\dot{a}|, \text{ c'est-à-dire } |\dot{a}| = |\dot{b}|.$$

On a encore les propriétés suivantes :

- 1°  $|\dot{a}| \cap |a| = a$ .
- 2° Si  $a \dot{=} a'$  et  $b \ddot{=} b'$ , alors  $ab \dot{=} a'b'$  et  $ba \ddot{=} b'a'$ .
- 3° Si  $a \dot{=} a'$  et  $a \dot{\subset} b$ , alors  $a' \dot{\subset} b$ ;  
 $a = a'$  et  $a \dot{\subset} b$ , alors  $a' \dot{\subset} b$ .

*L'axiome CJ.* — Pour aller plus loin, il faut introduire un axiome supplémentaire permettant de composer ces classes :

$$\begin{array}{ll} \text{Axiome } \overline{\text{CJ}}, & \text{pour tout } a \text{ et } b, \quad aba = ab; \\ \text{Axiome } \underline{\text{CJ}}, & \text{pour tout } a \text{ et } b, \quad aba = ba. \end{array}$$

On se limitera désormais au cas de  $\overline{\text{CJ}}$  qui entraîne :

- 1° Si  $a \dot{\subset} b$ , alors  $a \dot{\subset} b$ , donc  $a \dot{\subset} b$  équivalent à  $a \dot{\subset} b$ ;
- 2° Tout  $a, b$  :  $ab \dot{=} ba$ ,

d'où le résultat essentiel :

$$\text{si } a \dot{=} a' \text{ et } b \dot{=} b', \text{ alors } a'b' \in |\dot{ab}| \text{ et } b'a' \in |\dot{ab}|$$

778

ACADÉMIE DES SCIENCES.

( $|ab|$  peut comprendre d'autres êtres que les  $a'b'$  et les  $b'a'$ ). Il en découle que le système des classes a même structure qu'un *treillis commutatif*.

3° Les treillis CJ de base finie sont finis;

4° Pour tout *monome* B si  $a \doteq a'$ , alors  $aBa' = aB$ .

Il est maintenant possible d'introduire une deuxième opération « + », duale de la première et jouissant des propriétés habituelles, à condition de faire des hypothèses supplémentaires. Par exemple :

1° On distingue un être dans chaque classe et l'on convient que  $a + b$  désigne l'être distingué de la classe correspondant à l'être  $a + b$  du treillis commutatif associé.

2° S'il existe un isomorphisme entre chaque classe et un système de Klein-Barmen *fixe*, la structure est une *structure produit* (par exemple les domaines convexes orientés) et + compose les classes comme précédemment ; à l'intérieur d'une même classe, + opérant suivant la loi du système de Klein-Barmen. Au contraire de la précédente, cette convention est compatible avec toutes les lois universelles (cf. *infra*) de distributivité au sens large. Toute loi de ce type y implique son opposée (par exemple la distributivité à droite y implique la distributivité à gauche). D'autres conventions sont encore possibles.

*Hiérarchisation des lois commutatives.* — On appellera *loi universelle* une loi algébrique s'exprimant sous la forme de l'égalité de deux (ou plusieurs) monomes pour toutes valeurs des variables.

La loi la plus forte est la loi commutative :  $ab = ba$  équivalente (entre autres) à la composition de  $\overline{CJ}$  et  $\underline{CJ}$ . Il est facile de voir que cette loi est *immédiatement plus forte* que les deux lois  $\underline{CJ}$  et  $\overline{CJ}$  et que ce sont là les trois seules lois universelles à deux êtres : en effet la loi  $aba = a$  n'est pas compatible avec l'existence de  $\Omega$ .

Comme  $aba$  et  $ba$  sont isotopes et ne sont égaux que si  $\underline{CJ}$  est vérifiée, on peut montrer que toute loi universelle à plus de deux êtres est *simultanément plus faible* que  $\overline{CJ}$  et  $\underline{CJ}$  et par conséquent *plus faible* que CM définie par : pour tout  $a, b, c$  :

$$abaca = abca; \quad abacb = abcba; \quad acbab = acab.$$

Tout treillis CM de base finie est fini (la première des trois relations est d'ailleurs suffisante pour cela). Si la base est constituée par trois êtres, la *longueur* maxima des monomes est 5.

ANALYSE MATHÉMATIQUE. — *Système général de coordonnées dans un espace séparable de Hilbert-Hermite*. Note de M. OTTON MARTIN NIKODYM, présentée par M. Arnaud Denjoy.

1. Reprenons les notions et la terminologie de deux Notes précédentes (\*). Soient  $\mathcal{L}$  un espace séparable,  $(L)$  une classe ordonnée et fermée engendrant une

(\*) *Comptes rendus*, 224, 1947, pp. 322 et 628.

ANALYSE MATHÉMATIQUE. — *Remarques sur des relations d'ordre entre variables aléatoires indépendantes.* Note de M. MARCEL-PAUL SCHÜTZENBERGER, présentée par M. Gaston Julia.

En vue d'applications à des problèmes pratiques de statistique mathématique, on s'efforcera de formuler les conditions générales auxquelles doit satisfaire une relation entre variables aléatoires indépendantes pour pouvoir être considérée comme une relation d'ordre.

*Structure d'ensemble préordonné.* — Étant donné un ensemble  $\Phi_I$  de variables aléatoires  $X_i$ ,  $i \in I$  définies sur un même ensemble  $C$ , illimité, totalement ordonné et B-mesurable, on dira que  $\Phi$  est *préordonné* pour  $S, L : o(\Phi, S, L)$ , s'il existe une partie symétrique  $L$  de  $I \times I$  contenant la diagonale, et une application  $S_{ii}$  sur l'ensemble formé par  $-1, 0$  et  $1$  des couples  $X_i, X_i$  faisant partie de  $L$ , telle que :

$o_1$ . — pour tout  $i, i' : S_{i,i'} = -S_{i',i}$ .

$o_2$ . —  $S_{ii} = 0$  entraîne :  $i = i'$ .

$o_3$ . —  $S_{ii} \geq 0$  entraîne : pour tout  $d > 0$ .  $S(X_i, X_i + d) = 1$  même si  $X_i + d$  n'est pas dans  $\Phi$ .

$o_4$ . — Le groupe  $G$  des transformations de  $C$  pour lequel les  $S_{ii'}$  sont invariants a pour sous-groupe le groupe linéaire :  $y = c(x - m)$  où  $-\infty < m < +\infty ; 0 < c < +\infty$ .

*Structure d'ensemble ordonné.* —  $\Phi$  sera dit *ordonné* pour  $S : [O(\Phi, S, L)$  si  $o(\Phi, S, L)$  et même  $O_1 : S_{ii'} S_{i'i} = 1$  entraîne  $S_{ii'} S_{i'i} = 1$ .

*Structure de treillis.* —  $\Phi$  sera dit *treillis* pour  $S : \mathfrak{C}(\Phi, S, L)$  si  $o(\Phi, S, L)$  et  $\mathfrak{C}_1 : S_{ii'} S_{i'i} = 1$  entraîne  $(i, i')$  dans  $L$ ,  $\mathfrak{C}_2$  : pour tout  $ii'$ , il existe  $i''$  telle que  $S_{ii''} \geq 0$  et  $S_{i'i''} \geq 0$ , et  $S_{ii''} \geq 0$  et  $S_{i'i''} \geq 0$  entraîne  $S_{i'i''} \geq 0$ .

*Structure d'ensemble presque totalement ordonné.* —  $\Phi$  sera dit *presque totalement ordonné* pour  $S : t(\Phi, S, L)$  si  $\mathfrak{C}(\Phi, S, L)$  et  $t_1 : (i, i')$  non dans  $L$  et  $(i', i'')$  non dans  $L$  entraîne  $(i, i'')$  non dans  $L$ .

*Structure d'ensemble totalement ordonné.* —  $\Phi$  sera dit *totalement ordonné* pour  $S : T(\Phi, S, L)$ , si  $\mathfrak{C}(\Phi, S, L)$  et  $T_1 : L = I \times I$ .

Il est facile de voir que ces conditions sont indépendantes. Cependant  $o_1, o_2, \mathfrak{C}_1$  et  $\mathfrak{C}_2$  entraînent  $O_1$ , comme il ressort de ma Note à l'Académie des Sciences du 6 janvier 1947.

D'un point de vue pratique, il faut mentionner deux faits importants, d'ailleurs immédiats :

1° Pour tout  $\Phi$  il existe une infinité de  $S$  pour lesquelles  $\Phi$  est presque

( 2 )

totalelement ordonné : il suffit de prendre une relation d'ordre entre valeurs typiques de position.

2° On peut toujours adjoindre à  $\Phi$  des fonctions lui donnant une structure de treillis presque totalement ordonné; ces treillis ont pour loi universelle (1)

$$(a + b)(c + ab) = ab + c(a + b), \quad ab[c + d(a + b)] = a(bc + bd) + b(ac + ad).$$

Enfin, lorsque  $G$  est le groupe de toutes les transformations conservant la structure d'ordre de  $C$ , il est possible d'exprimer les  $S$  en fonction de la série (infinie) des moments des  $\Gamma_{i\bar{i}}(\lambda)$  définies dans le plan de  $p_i$  et  $p_{\bar{i}}$  par

$$\Gamma_{i\bar{i}}(\lambda) : F_i^{-1}(p_i) = F_{\bar{i}}^{-1}(p_{\bar{i}}) + \lambda,$$

où

$$p_i = F_i(x_i) = \text{Prob}(X_i < x_i),$$

et complétée par des segments de droites là où les fonctions  $F_i$  et  $F_{\bar{i}}$  sont *simultanément* discontinues.

Ces fonctionnelles sont en outre astreintes à être nulles quand  $\Gamma_{i\bar{i}}(0)$  est la diagonale  $p_i = p_{\bar{i}}$  et égales à 1, quand  $\Gamma_{i\bar{i}}(0)$  est la réunion  $\Gamma_{(\infty)}$  des deux segments

$$\Gamma_{(\infty)} \begin{cases} 0 \leq p_i < 1 & \text{et} & p_{\bar{i}} = 0, \\ p_i = 1 & \text{et} & 0 < p_{\bar{i}} \leq 1, \end{cases}$$

vers laquelle  $\Gamma_{i\bar{i}}(\lambda)$  converge quand  $\lambda \rightarrow \infty$

*Treillis des  $S^j$ .* — Étant donné un ensemble  $\Phi$ , l'ensemble des  $S^j, L^j$  tels que au moins  $o(\Phi, S^j, L^j)$ , forme un treillis  $\Theta(\Phi)$  pour la relation «  $S^j$  plus stricte que  $S^{j'}$  » ( $S^{j'} \subset S^j$ ) définie par

$$\text{pour tout } i, i' \quad S_{i\bar{i}'}^j = 1 \quad \text{entraîne} \quad S_{i\bar{i}}^{j'} = 1.$$

Ce treillis a un élément minimal donné par  $S^*$ , correspondant à  $\Gamma_{i\bar{i}}(0) = \Gamma(\infty)$  et la structure de  $\Phi$  est alors une structure presque totalement ordonnée que l'on appellera « structure minimale de  $\Phi$  ».

Il est manifeste que les  $S^j, L^j$  tels que  $\mathcal{O}(\Phi, S^j, L^j)$ ,  $\mathcal{C}(\Phi, S^j, L^j)$ ,  $\mathcal{I}(\Phi, S^j, L^j)$  ou  $\mathcal{T}(\Phi, S^j, L^j)$ , forment des sous-treillis de  $\Theta$  et n'en sont pas *nécessairement* des treillis quotients.

(1) *Comptes rendus*, 224, 1947, p. 512.

(Extrait des *Comptes rendus des séances de l'Académie des Sciences*,  
t. 224, pp. 878-880, séance du 24 mars 1947.)

PHYSIOLOGIE GÉNÉRALE. — *L'excitation des chimiorécepteurs de la langue par les substances du groupe des narcotiques indifférents et la règle thermodynamique de la narcose.* Note (1) de MM. **PIERRE GAVAUDAN**, **MARCEL-PAUL SCHUTZENBERGER** et M<sup>lle</sup> **HÉLÈNE POUSSEL**, transmise par M. Robert Courrier.

Nos recherches sur les différents phénomènes produits par les narcotiques indifférents (2) nous ont conduits à penser avec TAYLOR (3) que l'excitation nerveuse provoquée par les substances sapides n'est qu'une des manifestations de leur activité pharmacodynamique générale et doit obéir aux mêmes règles. Pour vérifier cette hypothèse de travail, il nous a paru intéressant de confronter les phénomènes de narcose, de mito-inhibition et les phénomènes d'excitation que provoquent les narcotiques indifférents. Nous avons groupé dans le tableau suivant, pour une série de substances aliphatiques et aroma-

	Saveur.		Contractions rythmiques du muscle de Sangsue.	Narcose (Tétard).	Mito-inhibition (racine de blé).
Benzène.....	brûlante	0,20	0,05	0,10	0,30
Nitrobenzène.....	sucrée	0,20	—	—	0,29
Anisol.....	amère	0,30	—	0,10	0,25
Acétophénone.....	»	0,20	0,10	—	0,40
Chlorobenzène.....	brûlante	0,50	—	—	0,28
Chloroforme.....	sucrée	0,025	0,06	0,01	0,10
Éther éthylique.....	amère	0,020	0,05(7)	0,13	0,20
Alcool éthylique.....	»	0,015	—	0,020	0,046
Hydrate de chloral.....	»	0,0005	—	0,00025 (poisson)	0,0001
Dichloréthane.....	{ sucrée	0,10	0,030	0,023	—
	{ amère	0,30	—	—	—
Tétrachloréthane.....	{ sucrée	0,20	0,20	0,017	
	{ amère	0,10			

tiques, les seuils d'activité thermodynamique de l'excitation des terminaisons nerveuses chimiosensibles de la langue humaine [essais des auteurs (4)] et les valeurs d'activité pour lesquelles on observe l'excitation du muscle de sangsue [série aromatique : essais des auteurs; série aliphatique : Joachimoglu (5)], la

(1) Séance du 12 mai 1947.

(2) P. GAVAUDAN, M. DODE et H. POUSSEL, *Mémor. des Industr. chim. de l'État*, 31, 1944

(3) *Protoplasma*, 4, 1928, pp. 1-17.

(4) Les valeurs d'activité thermodynamiques sont grossièrement assimilées au quotient : concentration active/concentration saturante (*Bioch. Zeits.*, 156, 1925, pp. 224-235).

(5) *Die Narkose*, Iéna, 1901, p. 82.

( 2 )

narcose du têtard de Grenouille [Overton <sup>(6)</sup>] ou du Poisson [Joachimoglu <sup>(7)</sup>] et la mito-inhibition <sup>(2)</sup>.

Nous résumerons ainsi les résultats : 1° l'excitation sapide se produit pour des valeurs d'activité comprises en général entre 0,01 et 1, c'est-à-dire dans la marge des narcotiques indifférents. L'hydrate de chloral, comme dans la mito-inhibition, montre une valeur anormalement basse; 2° les substances se classent ainsi; *a*, celles ne donnant qu'une sensation sucrée ou amère respectivement à la pointe de la langue ou au V lingual; *b*, celles donnant, avec la même localisation, les deux sensations; *c*, celles ne donnant pas de sensation sapide, mais une sensation d'irritation relevant du sens chimique commun.

De ces constatations on peut tirer les conclusions suivantes destinées à orienter les recherches ultérieures : 1° il existe un net parallélisme entre l'excitation sapide et diverses actions pharmacodynamiques, ce qui laisse supposer une identité de causes; 2° la règle thermodynamique <sup>(2)</sup> indique que, comme dans le cas de la narcose, les lipoides jouent un rôle dans la gustation. La spécificité des actions *dulcigène* et *amarogène* d'un même narcotique indifférent pourrait s'expliquer par des différences de solubilité dans les récepteurs considérés. Une étude thermodynamique s'appliquant à de très nombreuses substances et faite au moyen de données physico-chimiques expérimentales permettant de déterminer plus exactement les activités thermodynamiques, permettrait de préciser ce point et de décider si la liposolubilité est le seul facteur ou s'il faut tenir compte avec Renqvist <sup>(8)</sup> de l'adsorption ou encore des particularités stériques des molécules; 3° il faudra définir le rapport existant entre les excitations sapistes engendrées par les ions métalliques, les sucres, les alcaloïdes et les narcotiques indifférents. Il semble vraisemblable que le chlorhydrate basique de quinine et l'acétophénone, par exemple, dont les seuils respectifs d'amertume exprimés en activité thermodynamique sont de 0,00002 <sup>(9)</sup> et 0,30 n'agissent pas de la même façon; on peut penser qu'il existe ici, comme dans la narcose, une pluralité de mécanismes pouvant déclencher une même réaction cellulaire.

<sup>(6)</sup> *Bioch. Zeits.*, 120, 1921, pp. 203-211.

<sup>(7)</sup> P. DODEL et G. DASTUGUE, *Bull. Sc. Pharmacol.*, 44, 1937, pp. 145-155.

<sup>(8)</sup> *Ueber den Geschmack (Skand. Arch. f. physiol.*, 38, 1919, pp. 97-201).

<sup>(9)</sup> Calculé d'après un seuil indiqué par E. VON SKRAMLIK, *Physiologie des Geschmackssinnes (Handbuch der normalen und pathologischen Physiologie*, II, Berlin 1926, p. 382).

(Extrait des *Comptes rendus des séances de l'Académie des Sciences*, t. 224, pp. 1525-1527, séance du 28 mai 1947.)

CALCUL DES PROBABILITÉS. — Sur certains paramètres caractéristiques des systèmes d'événements compatibles et dépendants et leur application au calcul des cumulants de la répétition. Note (1) de M. MARCEL-PAUL SCHÜTZENBERGER, transmise par M. Gaston Julia.

Étant donné l'événement composé :  $A = (a_1, a_2, \dots, a_n)$ , réalisé quand et seulement quand les  $n$  événements simples  $a_1, a_2, \dots, a_n$  le sont simultanément, on appellera *déviante d'indépendance* de  $A$ , l'expression

$$\gamma(A) = \gamma(a_1, a_2, \dots, a_n) = 2^n \sum (-1)^{\nu-1} (\nu-1)! \Pr(B_1) \Pr(B_2), \dots, \Pr(B_\nu),$$

où la sommation est étendue à toutes les partitions de l'ensemble  $a, a_1, \dots, a_n$ , en  $\nu$  ( $1 \leq \nu \leq n$ ) sous-ensembles auxquels sont attachés les événements  $B_i$ .

Par exemple :

$$\begin{aligned} \gamma(ab) &= 4(\Pr(ab) - \Pr(a)\Pr(b)), \\ \gamma(abc) &= 8\{\Pr(abc) - \Pr(a)\Pr(bc) - \Pr(b)\Pr(ca) - \Pr(c)\Pr(ab) + 2\Pr(a)\Pr(b)\Pr(c)\}. \end{aligned}$$

On pose

$$\gamma(a) = 2\left(\Pr(a) - \frac{1}{2}\right).$$

Si l'on suppose que  $n_1$  événements sont identiques à  $a_1$ ,  $n_2$  à  $a_2$ , ...,  $n_i$  à  $a_i$ , on obtient une *déviante contractée* par

$$\begin{aligned} \gamma(a^2 b) &= 8[\Pr(ab) - 2\Pr(a)\Pr(ab) - \Pr(a)\Pr(b) + 2(\Pr(a))^2 \Pr(b)], \\ \gamma(a^3) &= 8[\Pr(a) - 3(\Pr(a))^2 + 2(\Pr(a))^3]. \end{aligned}$$

On démontre directement : si l'ensemble  $a_1, a_2, \dots, a_n$  peut être décomposé en deux sous-ensembles tels que tout événement composé à partir des événements simples du premier soit indépendant de tout événement composé à partir du second, alors  $\gamma(a_1, a_2, \dots, a_n) = 0$ . Notamment  $\gamma(A) = 0$  si l'un des  $a_i$  ou son négat est presque certain.

De cette constatation découle immédiatement : les  $\gamma$  sont des fonctions alternées des  $a_i$ , c'est-à-dire que si  $A'$  ne diffère de  $A$  que par la substitution de  $m$  négats d'événements simples à ces événements eux-mêmes,

$$\gamma(A') = (-1)^m \gamma(A).$$

(1) Séance du 28 juillet 1947.

( 2 )

Un calcul direct à partir des formules classiques montre alors que le  $i^{\text{ème}}$  cumulant (semi-invariant)  $x_i$  de la répétition  $r$  dans un système de  $n$  événements peut s'écrire symboliquement

$$x_i = ((a_1 + a_2 + \dots + a_n))^i,$$

où chaque terme  $a_1^{n_1} a_2^{n_2} \dots a_v^{n_v}$  du développement doit être remplacé par la déviation correspondante  $\gamma(a_1^{n_1} a_2^{n_2} \dots a_v^{n_v})$ .

On obtient ainsi pour les premiers cumulants

$$\begin{aligned} x_1 &= 2\bar{r} - n \\ x_2 &= 2 \sum_{(a,b)} (\Pr(ab) - \Pr(a)\Pr(b)) + \sum_{(a)} \Pr(a)(1 - \Pr(a)) \\ x_3 &= 6 \sum_{(a,b,c)} (\Pr(abc) - \Pr(a)\Pr(bc) - \Pr(b)\Pr(ac) - \Pr(c)\Pr(ab) + 2\Pr(a)\Pr(b)\Pr(c)) \\ &\quad + 6 \sum_{(a,b)} (\Pr(ab) - \Pr(a)\Pr(b))(1 - \Pr(a) - \Pr(b)) \\ &\quad + \sum_{(a)} (\Pr(a))(1 - \Pr(a))(1 - 2\Pr(a)). \end{aligned}$$

A partir de ces formules, on peut aisément retrouver ou établir un certain nombre de cas de tendances vers la normalité. On citera : si les  $a_i$  en nombre  $n$  sont tels que tous les  $\gamma(a_i, a_j)$  soient positifs, finis et différents de zéro, la distribution de  $\rho = (r - \bar{r})/\sigma$  tend vers une distribution normale quand  $n \rightarrow \infty$ , car les cumulants d'ordre  $i$  de  $\rho$  sont au plus d'ordre  $n^{1-(i/2)}$ .

(Extrait des *Comptes rendus des séances de l'Académie des Sciences*,  
t. 225, pp. 277-278, séance du 4 août 1947.)

1947-6. Axiomatisation de la géométrie dans un complexe linéaire... Année 1947

---

---

**« LES EXTRAITS DE LA REVUE SCIENTIFIQUE »**

---

---

**AXIOMATISATION DE LA GÉOMÉTRIE  
DANS UN COMPLEXE LINÉAIRE  
DE DROITES**

PAR

MARCO SCHÜTZENBERGER



M. CM. XL. VII.

**EXTRAIT DU N° 3278, 1<sup>er</sup> AOÛT 1947,  
FASCICULE 13 DE LA 85<sup>e</sup> ANNÉE DE LA REVUE SCIENTIFIQUE, PAGES 782 A 784**

---

---

**PUBLIÉS PAR « LES ÉDITIONS DE LA REVUE SCIENTIFIQUE »**

---

---

## AXIOMATISATION DE LA GÉOMÉTRIE DANS UN COMPLEXE LINÉAIRE DE DROITES

PAR

MARCO SCHÜTZENBERGER

**Sommaire.** — De même que la géométrie projective linéaire peut être définie de façon purement formelle, indépendamment de la réalisation des êtres qui la constituent, de même il est possible de donner à la géométrie, dans un complexe linéaire, une axiomatisation exclusivement combinatoire. Pour cela, le langage de la théorie des treillis (*Lattices*, *Verbände*, structures) est particulièrement bien adapté si l'on ajoute aux deux opérations  $\cap$  (intersection) et  $\cup$  (réunion complétée) une troisième opération  $/$ , introduisant par sa définition même les propriétés caractéristiques de la structure de : complexe linéaire de droites dans un espace projectif à trois dimensions et à coordonnées dans un corps commutatif.

**S**OIT donc un treillis  $\mathfrak{B}'$  d'êtres génériques  $\alpha, \beta, \gamma, \dots$ , constitué par :

- 1° un être minimal  $\Omega$  (élément vide);
- 2° une classe d'êtres  $a, b, c, \dots, x, \dots$  (que nous nommerons *points*) immédiatement supérieurs à  $\Omega$  ( $\Omega \subset a$ );
- 3° une classes d'êtres  $A, B, C, \dots, X, \dots$  (que nous nommerons *droites*) immédiatement inférieurs à l'être suivant : ( $A \subset \bar{\Omega}$ );
- 4° un être maximal  $\bar{\Omega}$  (espace complet).

Plus brièvement nous définirons  $\mathfrak{B}'$  comme un treillis où toutes les chaînes ont la longueur 3, les êtres  $a, b, c, \dots$  étant de hauteur 1 ( $|a| = 1$ ), et les  $A, B, C, \dots$  de hauteur 2 ( $|A| = 2$ ) (condition I).

Nous pouvons déjà donner certaines précisions sur la valeur de  $\cap$  et  $\cup$  :

$$\begin{aligned} \Omega \cap \Omega &= a \cap \Omega = A \cap \Omega = \bar{\Omega} \cap \Omega = \Omega; \\ \Omega \cup \bar{\Omega} &= a \cup \bar{\Omega} = A \cup \bar{\Omega} = \bar{\Omega} \cup \bar{\Omega} = \bar{\Omega}; \\ a \cap a &= a \cup a = a \cap \bar{\Omega} = a \cup \Omega = a; \\ A \cap A &= A \cup A = A \cap \bar{\Omega} = A \cup \Omega = A; \\ a \cap b &= \Omega; \quad A \cup B = \bar{\Omega}. \end{aligned}$$

$$a \cap A \begin{cases} = a & \text{si } a \subset A, \\ = \Omega & \text{si } a \not\subset A, \end{cases} \quad \left| \quad a \cup A \begin{cases} = A & \text{si } a \subset A, \\ = \bar{\Omega} & \text{si } a \not\subset A. \end{cases}$$

Enfin, pour certains couples d'êtres  $a$  et  $b$  dits *liés*,

$$(|a \cup b| = 2) : a \cup b = X.$$

Pour tous les autres couples,

$$a \cup b = \bar{\Omega} \quad (|a \cup b| = 3).$$

Pour certains couples d'êtres  $A$  et  $B$  dits *liés*,

$$(|A \cap B| = 1) : A \cap B = x.$$

Pour tous les autres couples

$$A \cap B = \Omega \quad (|A \cap B| = 0).$$

Mais, pour définir complètement  $\mathfrak{B}'$ , il faut encore introduire  $/$ ; nous le ferons à partir de l'axiome suivant (axiome C) :

Il existe une application univoque unique  $\varphi$  de l'ensemble  $H$  des couples  $(x, X)$  sur le sous-ensemble  $H'$  de  $H$  des couples tels que  $x \subset X$ , laissant  $H'$  invariant et telle que :  $\varphi(x, X) = (y, Y)$  entraîne  $x \cup y = Y$  et  $X \cap Y = y$ , et nous définirons  $/$  par :  $x/X = y$ .

Géométriquement, en désignant par  $\hat{x}$  l'ensemble des êtres  $z$  liés à  $x$  (c'est-à-dire tels que  $|x \cup z| = 2$ ),  $y$  est l'intersection (au sens de la Théorie des ensembles) de  $\hat{x}$  et de  $X$ , ou, si l'on préfère, de  $X$  et du plan polaire  $\hat{x}$  de  $x$ , que l'axiome C postule unique.

Désormais ce seront exclusivement des treillis satisfaisant à C que nous désignerons par  $\mathfrak{B}$ . Ces treillis jouissent de propriétés spéciales :

1° Tout  $A$  peut être représenté comme réunion d'au moins deux  $a$  et, de même, tout  $a$  est l'intersection d'au moins deux  $A$ ; cela justifie l'équivalence de nos deux définitions des  $\mathfrak{B}'$ .

2° Si

$$|x \cup y| = 2 \quad \text{et} \quad |x \cup z| = 2, \quad \text{alors} \quad |y \cup z| = 3.$$

Autrement dit, il n'existe pas de triangles dont les trois côtés appartiennent à un même complexe linéaire.

3° Dans le cas fini, toutes les droites ont le même nombre de points  $m$ , et tous les points sont sur le même nombre de droites  $n$ . Réciproquement, s'il existe  $N$

AXIOMATISATION DE LA GÉOMÉTRIE DANS UN COMPLEXE LINÉAIRE DE DROITES 783

points et M droites dans un treillis  $\mathcal{B}'$  sans triangles, et si (condition II)

$$M \cdot n \cdot (m - 1) \cdot (n - 1) = N \cdot (M - n) = M \cdot (N - m),$$

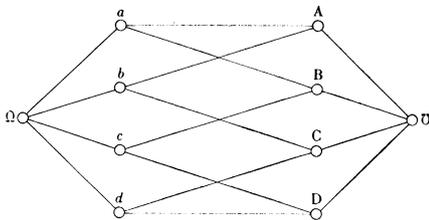
alors  $\mathcal{B}'$  satisfait à l'axiome C.

4° Quels que soient les êtres génériques  $\alpha, \beta, \gamma, \dots$  et de manière duale, en permutant  $\cap$  et  $\cup$ , l'on a les identités suivantes :

$$(I) \quad \alpha \cap (\beta \cup \gamma \cup \delta) = [\alpha \cap (\beta \cup \gamma)] \cup [\alpha \cap (\gamma \cup \delta)] \cup [\alpha \cap (\delta \cup \beta)],$$

$$(II) \quad [\beta \cup \gamma] \cap [\alpha \cup (\beta \cap \gamma)] = \{ \alpha \cap [\beta \cup \gamma] \} \cup \{ \beta \cap [\alpha \cup (\beta \cap \gamma)] \} \cup \{ \gamma \cap [\alpha \cup (\beta \cap \gamma)] \},$$

$$(III) \quad [\alpha \cup \beta] \cap [\alpha \cup \gamma] = \alpha \cup [\beta \cap (\alpha \cup \gamma)] \cup [\gamma \cap (\alpha \cup \beta)].$$



Le plus simple des  $\mathcal{B}$  est alors le treillis de la figure 1 (quadrilatère gauche), dont nous donnons la double représentation.

Mais toutes ces conditions sont insuffisantes pour caractériser les complexes habituels de la géométrie élémentaire ; en effet elles sont vérifiées si l'on identifie aux êtres  $a, b, c, \dots$  et  $A, B, C, \dots$  respectivement, les droites et les points d'un complexe linéaire. Or nous verrons tout à l'heure que cette structure (complexe dual) n'est qu'exceptionnellement isomorphe à celle que l'on obtient en identifiant les points aux  $a, b, c, \dots$  et les droites aux  $A, B, C, \dots$ .

Étant donné cinq points  $a_i$  (les indices sont pris modulo 5) distincts, nous nous dirons qu'ils forment un pentagone si les cinq réunions  $a_i \cup a_{i+1}$  sont des droites de  $\mathcal{B}$  (c'est-à-dire si  $|a_i \cup a_{i+1}| = 2$ ), et nous appellerons P-constructible tout être de  $\mathcal{B}$  pouvant s'obtenir à partir des  $a_i$  au moyen des opérations  $\cap$ ,  $\cup$  et  $/$ .

Nous introduirons alors l'axiome suivant, qui est le plus simple possible dans le système d'opérations adopté. En posant

$$b_i = a_i / a_{i+2} \cup a_{i-2},$$

l'on a l'axiome J :

$$c_i = b_{i+1} / a_i \cup b_i = b_{i-1} / a_i \cup b_i,$$

pour tout pentagone P de  $\mathcal{B}$ .

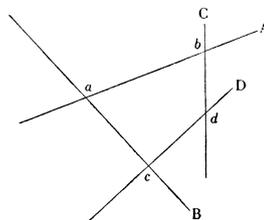
Il est assez facile de vérifier cette identité pour un complexe au sens ordinaire, en prenant, après

transformation linéaire convenable, pour équation  $xY - yX + zT - tZ = 0$  et, pour pentagone,

$$\begin{matrix} a_0 : 1, & 0, & 0, & 0, \\ a_1 : 0, & 0, & 0, & 1, \\ a_2 : 0, & 1, & 0, & 1, \\ a_3 : 1, & 1, & 1, & 1, \\ a_4 : 0, & 0, & 1, & 1. \end{matrix}$$

Nous n'exposerons pas la démonstration complète, assez lourde, du théorème suivant qui établit la validité de l'axiome J :

« Si J est vrai pour un  $\mathcal{B}$  entièrement P-constructible,  $\mathcal{B}$  est isomorphe à un complexe linéaire à coor-



$$\begin{matrix} a/C = b \\ a/D = c \\ b/B = a \\ b/D = d \\ c/A = a \\ c/C = d \\ d/A = b \\ d/B = c \end{matrix}$$

Fig. 1.

données dans un corps minimal de caractéristique quelconque. »

Nous donnerons cependant deux Lemmes qui servent de pivot à cette démonstration :

LEMME 1. — J entraîne  $c_{i+2}/a_i \cup b_i = c_{i-2}/a_i \cup b_i$ ;

LEMME 2. — J entraîne  $d_i = a_i / b_{i+1} \cup c_{i+2}$ ;

où les signes peuvent être combinés de manière quelconque.

Pour établir le théorème, il est commode d'introduire une relation ternaire d'alignement entre points de  $\mathcal{B}$  :  $(a, b, c)_{x,y}$  exprimant le fait que  $a, b$  et  $c$  sont contenus simultanément dans  $\hat{x}$  et dans  $\hat{y}$  ou, si l'on préfère, que

$$\begin{matrix} |a \cup x| = |a \cup y| = 2; \\ |b \cup x| = |b \cup y| = 2; \\ |c \cup x| = |c \cup y| = 2. \end{matrix}$$

L'on prouve alors par récurrence, pour tous les points P-constructibles :

- 1°  $(a, b, c)_{x,y}$  et  $(a, b, d)_{x,y}$  entraînent  $(a, c, d)_{x,y}$ ,
- 2°  $(a, b, c)_{x,y}$  et  $(a, b, c)_{x,z}$  entraînent  $(a, b, c)_{y,z}$ .

Cela permet de montrer que les ensembles intersections des  $\hat{x}$  jouissent des propriétés des droites de l'espace projectif.

On observera toutefois que, comme pour les géométries projectives, il n'est pas démontré que J assure la possibilité de réalisation de  $\mathcal{B}$  dans un corps quand l'ensemble des êtres P-constructibles n'épuise pas  $\mathcal{B}$ .

Quelques cas particuliers sont dignes de remarque.

Lorsque le corps des coordonnées est un corps de caractéristique finie  $p$ , il est toujours possible de trouver une égalité en  $\cap$ ,  $\cup$  et  $/$  caractérisant ce module. Par exemple, l'on a

$$c_i =: d_i \quad (\text{Loi } M_2),$$

pour les corps modulo 2 ;

$$e_i =: b_{i \pm 2} / b_{i \pm 1} \cup c_{i \pm 1} \quad (\text{Loi } M_2),$$

pour les corps modulo 3, avec les deux combinaisons de signes  $+$ ,  $-$ ,  $-$ , et  $-$ ,  $+$ ,  $+$ , etc.

Dans le cas des corps modulo 2, il existe un isomorphisme entre la structure  $\mathfrak{B}$  et la structure duale obtenue en permutant  $\cap$  et  $\cup$ .

L'on se rend facilement compte que la formule duale de J n'est pas vérifiée dans les autres cas et que par conséquent cette dualité est le privilège des  $\mathfrak{B}$  modulo 2 et du treillis donné en exemple.

Enfin, trois théorèmes importants peuvent être établis pour tous les  $\mathfrak{B}$  P-constructibles :

**THÉORÈME 1.** —  $\mathfrak{B}$  est P-constructible à partir de tout autre pentagone  $P'$  et ne l'est pour aucune autre configuration de cinq ou de moins de cinq points.

**THÉORÈME 2.** — L'ensemble des «  $a$  » a même puissance que l'ensemble des «  $A$  » ( $=: p^3 + p^2 + p + 1$ ) quand le corps des coordonnées est modulo  $p$ .

**THÉORÈME 3.** — Pour tout triple  $x, y, z$ , tel que

$$|x \cup y| = |y \cup z| = |z \cup x| = 3,$$

il existe un et un seul «  $t$  » satisfaisant à

$$|x \cup t| = |y \cup t| = |z \cup t| = 2.$$

Comme exemple de treillis  $\mathfrak{B}$  non représentable sous forme de complexe linéaire, il faut citer la configuration classique  $\Gamma$  formée par les 27 droites de la surface cubique générale.

Il est facile de voir que tout sous-treillis P-constructible de  $\Gamma$  est un complexe linéaire dans un espace modulo 2, et que les 12 autres droites n'appartenant pas à cette configuration ne peuvent être choisies dans cet espace de manière à satisfaire aux relations d'incidences.

Ce type de structure, intermédiaire entre le complexe linéaire *complet* dans un surcorps et le complexe dans un corps minimal, se généralise sans difficulté à tous les surcorps, en donnant de nouvelles familles de treillis  $\mathfrak{B}$  fermés pour  $\cap$ ,  $\cup$  et  $/$ , et où le nombre des droites est différent du nombre des points.

Pour conclure, signalons qu'il est possible, comme pour la géométrie projective plane, de remplacer la relation d'incidence traduite par l'axiome J, par une propriété du groupe d'automorphisme de  $\mathfrak{B}$ .

Partons d'un quadrilatère vrai Q de  $\mathfrak{B} : a, b, c, d$ ,

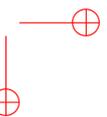
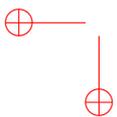
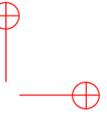
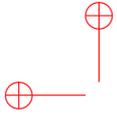
$$(|a \cup b| = |a \cup d| = |b \cup c| = |c \cup d| =: 2),$$

et soit  $m$  lié à  $a$ . Nous dirons que  $\mathfrak{B}$  est Q-homogène s'il existe un sous-groupe du groupe d'automorphisme de  $\mathfrak{B}$  laissant Q invariant et transitif en  $m$ .

L'on peut alors démontrer que  $\mathfrak{B}$  vérifie J si, pour tous les Q,  $\mathfrak{B}$  est Q-homogène avec, en outre, les propriétés suivantes :

Le sous-groupe qui conserve Q et  $M =: m \cup a$  laisse invariantes toutes les droites issues de  $a, b, c, d$ , et des points de  $\hat{a} \cap \hat{c}$  et  $\hat{b} \cap \hat{d}$  (intersection au sens de la théorie des ensembles !).

(manuscrit reçu le 3 juillet 1947)



# Année 1948

## Bibliographie

- [1948-1] Pierre Gavaudan, Hélène Poussel et Marcel-Paul Schützenberger. Le mécanisme physico-chimique de l'excitation sapide et la notion d'excitant indifférent. *C. R. Acad. Sci. Paris*, 226 :751–752, 1948. Séance du 1er avril 1948.
- [1948-2] Pierre Gavaudan, Hélène Poussel, Georges Brebion et Marcel-Paul Schützenberger. L'étude des conditions thermodynamiques de l'excitation olfactive et les théories de l'olfaction. *C. R. Acad. Sci. Paris*, 226 :1395–1396, 1948. Séance du 26 avril 1948.
- [1948-3] Marcel-Paul Schützenberger et Raymond Turpin. Recherches statistiques sur la distribution du sexe à la naissance. *C. R. Acad. Sci. Paris*, 226 :1845–1846, 1948. Séance du 31 mai 1948.
- [1948-4] Marcel-Paul Schützenberger. Valeurs caractéristiques du coefficient de corrélation par rang de Kendall dans le cas général. *C. R. Acad. Sci. Paris*, 226 :2122–2123, 1948. Séance du 28 juin 1948.
- [1948-5] Marcel-Paul Schützenberger. Fidélité des tests. *Travail et Méthode*, pages 14–16, Novembre 1948.
- [1948-6] Marcel-Paul Schützenberger. Étude statistique d'un problème de sociométrie. *Gallica Biologica Acta*, 1(Fasc. 1) :96–104, janvier - mars 1948.
- [1948-7] Pierre Gavaudan, Georges Brebion, Hélène Poussel et Marcel-Paul Schützenberger. Étude pharmacodynamique des chimiorécepteurs de l'olfaction. (1.) Principes et méthodes d'une étude thermodynamique de l'olfaction. *Gallica Biologica Acta*, 1(Fasc. 2) :147–166, avril-juin 1948.
- [1948-8] Pierre Gavaudan, Hélène Poussel et Marcel-Paul Schützenberger. Étude pharmacodynamique des chimiorécepteurs de l'olfaction. (2.) L'étude thermodynamique des séries homologues et les théories de l'olfaction. *Gallica Biologica Acta*, 1(Fasc. 2) :167–187, avril-juin 1948.

Bibliographie

Année 1948

- [1948-9] Marcel-Paul Schützenberger. Remarques sur les relations aléatoires d'ordre et leur application à la psychologie. *Gallica Biologica Acta*, 1(Fasc. 2) :191–204, avril - juin 1948.
- [1948-10] Marcel-Paul Schützenberger. An abac for the simple range. *Psychometrika*, 13(2) :95–97, juin 1948.
- [1948-11] Marcel-Paul Schützenberger. Sur certaines applications remarquables des treillis dans eux-mêmes. *C. R. Acad. Sci. Paris*, 227 :1008–1010, 1948. Séance du 3 novembre 1948.
- [1948-12] Marcel-Paul Schützenberger. *Contribution à l'étude statistique du sexe à la naissance*. Thèse en médecine, Faculté de Médecine de Paris, 1948. 52 pages, soutenue le 2 juillet 1948.

PHYSIOLOGIE. — *Le mécanisme physico-chimique de l'excitation sapide et la notion d'excitant indifférent.* Note de M. **PIERRE GAVAUDAN**, M<sup>lle</sup> **HÉLÈNE POUSSEL** et M. **MARCEL PAUL SCHUTZENBERGER**, transmise par M. Robert Courrier.

Selon les hypothèses de Cohn (<sup>1</sup>) et d'Oertly et Myers (<sup>2</sup>), les molécules devraient leur sapidité à la présence de groupements sapophores actifs. Nos précédentes recherches sur l'application de la notion d'activité thermodynamique au mécanisme de la narcose et de la gustation (<sup>3</sup>) nous ont laissé entrevoir que ces hypothèses posaient mal le problème, parce qu'elles négligeaient de définir les conditions énergétiques du système solution sapide-récepteur cellulaire.

Une étude (<sup>4</sup>) portant sur 11 sucres, 6 substances organiques sucrées et 15 substances organiques amères, nous permet de tirer les conclusions suivantes :

1° *Substances liposolubles amères ou sucrées.* — Des substances de constitutions très différentes provoquent les sensations d'amer (exemple: tétrachloréthane, éther éthylique, acétophénone, anisol) ou de sucré (exemple: tétrachloréthane, nitrobenzène, chloroforme, *p*-éthoxyphénylurée) pour des valeurs d'activité thermodynamique très voisines dans les deux cas, variant entre un et quelques centièmes. Ces substances s'accumulent dans les lipoïdes en vertu de leur coefficient de partage et de même que la narcose se produit à une valeur d'activité constante, l'excitation par ces narcotiques indifférents se produit également quand une activité et une fraction moléculaire critiques sont atteintes dans les lipoïdes cellulaires. Mais la valeur de la constante d'excitation doit

---

(<sup>1</sup>) *Die organischen Geschmackstoffe*, 1924.

(<sup>2</sup>) *Journ. Amer. Chem. Soc.*, 41, 1919, p. 855.

(<sup>3</sup>) P. GAVAUDAN, M. DODE et H. POUSSEL, *Mémor. des Services chim. de l'État*, 31, 1944, p. 384-423 et 32, 1946, p. 388-409; P. GAVAUDAN, M. P. SCHUTZENBERGER et H. POUSSEL, *Comptes rendus*, 224, 1947, p. 1525.

(<sup>4</sup>) Étude portant sur des données dues à divers auteurs et sur des essais personnels effectués isolément sur la zone antérieure de la langue pour le sucré et sur le V. lingual pour l'amer.

752

ACADÉMIE DES SCIENCES.

être nécessairement inférieure à celle de la constante de narcose (0,06 molg/l). Comme l'activité est représentée par  $N/N_0$  où  $N$  et  $N_0$  sont les fractions moléculaires actuelle et saturante, les valeurs d'activité dans la narcose et l'excitation pourront être très voisines ou égales; il suffira que  $N_0$  varie selon le récepteur cellulaire considéré.

2° *Substances hydrosolubles amères ou sucrées.* — Les valeurs d'activité et les concentrations en molg/l dans l'eau sont assez constantes pour des substances amères telles que l'urée, l'antipyrine, l'éthyluréthane, etc. L'activité varie beaucoup et inversement à la solubilité, pour des substances comme les sucres, mais les concentrations dans l'eau sont relativement constantes (0,02 à 0,12 molg/l). Tout se passe comme si l'excitation se déclenchait quand une certaine fraction moléculaire du corps dissous est atteinte dans un solvant cellulaire. L'excitation sapide ou amère se produirait quand une certaine fraction moléculaire (nombre de molécules dissoutes sur total des molécules du mélange) est présente dans un solvant cellulaire aqueux ou lipoïdique dont l'activité est modifiée d'une façon constante. Pour chacune des deux sensations il existerait *deux voies d'excitation indifférente*, ce qui s'accorde avec les idées sur la structure mixte hydro-lipophile des couches limitantes cellulaires. L'existence de substances très différentes provoquant exclusivement l'une des deux sensations ou l'existence de substances bisapides (tétrachloréthane) agissant toutes à des valeurs d'activité très voisines ôtent beaucoup de valeur à la théorie des groupes sapophores. Ce sont des forces de van der Waals et non des réactions chimiques, qui, au moins dans un premier stade du phénomène, provoquent l'excitation. Mais l'action localisée à un type de récepteur pour certaines substances conduit à admettre que, les conditions énergétiques étant respectées, les configurations moléculaires règlent la perméabilité ou la solubilité dans les solvants cellulaires et concourent dans ce sens à déterminer la qualité sapide.

Enfin des substances comme l'acide picrique, par exemple, s'écartent de la règle précédente et sont amères à une valeur de l'ordre de 0,0001. Dans de semblables cas, il faut donc tenir compte de la réactivité chimique, de considérations cinétiques, ou encore de phénomènes de *pénétration* de films superficiels analogues à ceux décrits par Schulman (<sup>5</sup>). Ces phénomènes évoquent un ou des mécanismes d'action plus complexes que ceux mis en jeu par les excitants indifférents.

En résumé, la notion d'activité permet de distinguer deux groupes d'excitants sapides : les indifférents et les spécifiques.

---

(<sup>5</sup>) *Trans. Farad. Soc.*, 39, 1943, p. 412-417.

SÉANCE DU 26 AVRIL 1948.

1395

PHYSIOLOGIE. — *L'étude des conditions thermodynamiques de l'excitation olfactive et les théories de l'olfaction.* Note de M. PIERRE GAVAUDAN, M<sup>lle</sup> HÉLÈNE POUSSEL, M. GEORGES BREBION et M. MARCEL-PAUL SCHUTZENBERGER, transmise par M. Robert Courrier.

Jusqu'à présent, dans les recherches sur le mécanisme de l'olfaction, on a seulement tenu compte des concentrations-seuils des substances odorantes exprimées en poids ou en molécule-gramme par litre. On a parfois aussi souligné le rôle important de la vitesse du courant odorant. Mais ces simples données expérimentales ne définissent pas exactement les conditions du système olfactif constitué par la substance odorante à une pression de vapeur partielle donnée dans l'atmosphère et par la phase qui est le lieu de l'excitation dans la cellule.

Nous nous sommes proposé de réaliser pour l'olfaction une étude analogue à celle déjà effectuée dans le cas de la narcose et de l'excitation sapide par les narcotiques indifférents (<sup>1</sup>). Beaucoup de substances odorantes sont aussi narcotiques liposolubles et déclenchent des sensations sâpides bien définies. Il s'agit de déterminer si les conditions énergétiques du déclenchement des sensations sâpides et odorantes sont semblables ou différentes.

Nous avons déterminé des seuils olfactifs exprimés directement en activités thermodynamiques (rapport  $p/p_0$  de la pression partielle à la pression saturante) grâce à des techniques spéciales qui seront décrites ultérieurement.

Nous avons d'abord étudié les séries homologues aliphatiques car les odeurs y sont suffisamment différentes et les propriétés physiques et pharmacodynamiques bien connues. Dans la série des alcools le seuil d'activité thermodynamique provoquant une sensation spécifique passe de l'ordre de grandeur de  $1000 \times 10^{-6}$  pour l'alcool méthylique à celui de  $6 \times 10^{-6}$  pour l'alcool butylique et à celui de  $30$  à  $60 \times 10^{-6}$  pour l'alcool octylique. Contrairement à ce que l'on observe dans la narcose et dans l'excitation sapide, où l'activité-seuil, de l'ordre des centièmes, reste à peu près constante, on observe dans l'olfaction une diminution considérable, du terme en  $C_1$  au terme en  $C_4$ , puis une légère remontée du terme en  $C_4$  au terme en  $C_8$ . Il semble qu'une pareille oscillation s'observe aussi avec les acides d'après les résultats de J. Passy (<sup>2</sup>) en extrapolant la valeur de la tension de vapeur saturante de l'acide en  $C_{10}$ , non trouvée dans les tables pour  $T = 20^\circ$ . En  $C_4$  l'activité descend à  $0,2 \times 10^{-6}$ .

La valeur parfois très faible de l'activité thermodynamique et l'allure très

(<sup>1</sup>) P. GAVAUDAN, M. DODÉ et H. POUSSEL, *Mémorial des Services chimiques de l'État*, **31**, 1944, p. 384-423 et **32**, 1946, p. 388-409; P. GAVAUDAN, H. POUSSEL et M. P. SCHUTZENBERGER, *Comptes rendus*, **226**, 1948, p. 751.

(<sup>2</sup>) *Comptes rendus*, **116**, 1893, p. 1007.

1396

ACADÉMIE DES SCIENCES.

spéciale de sa variation dans les séries homologues suggèrent que dans l'olfaction des mécanismes assez différents de ceux de la narcose et de l'excitation sapide sont mis en jeu. Il n'existe pas de règle simple analogue aux règles de Meyer et Hemmi ou de Traube permettant de prévoir l'activité-seuil et encore moins la qualité de l'odeur. La narcose et l'excitation sapide dépendent de propriétés physiques communes à toutes les molécules et assez banales : elles sont relatives à la solubilité et à l'adsorption et leurs grandeurs sont directement reliées au nombre d'atomes de carbone; toute molécule, quelle que soit sa grandeur, produit le phénomène de façon *indifférente*. Dans le cas de l'olfaction il s'agit au contraire d'une action plus spécifique comme l'indiquent les variations du seuil et de la qualité sensorielle. Mais il n'est pas exclu que des règles reliant cette spécificité à certaines propriétés physiques puissent être établies.

En effet, les incréments de diverses constantes physiques telles que longueur de la chaîne, point de fusion, coefficient d'activité, constante capillaire, varient en oscillant périodiquement dans les séries homologues, traduisant les modifications de la *forme* et de la *symétrie* moléculaires. On pourrait s'attendre à trouver une relation entre ces derniers caractères physiques, les variations des seuils d'activité thermodynamique et la qualité de la sensation. Il faut d'ailleurs noter que certains phénomènes de toxicité [H. Hurst <sup>(3)</sup>] présentent aussi une variation oscillante dans les séries homologues. Il n'est donc pas nécessaire de recourir d'emblée à une théorie vibratoire du type de celle de Dyson <sup>(4)</sup> pour expliquer l'excitation olfactive. Il paraît encore possible de rechercher des hypothèses de travail dans le cadre des propriétés physico-chimiques habituellement invoquées pour l'explication des phénomènes pharmacodynamiques et toxiques.

---

<sup>(3)</sup> *Trans. Farad. Soc.*, **39**, 1943, p. 390-411.

<sup>(4)</sup> *Perf. Essent. Oil Rec.*, **28**, 1937, p. 13 et *Chemistry and Indust.*, **57**, 1938, p. 647-651.

BIOLOGIE. — *Recherche statistique sur la distribution du sexe à la naissance.*  
 Note de MM. **RAYMOND TURPIN** et **MARCEL-PAUL SCHÜTZENBERGER**, présentée  
 par M. Léon Binet.

Nous avons étudié la distribution du sexe à la naissance dans 14230 familles contemporaines françaises d'au moins 5 enfants. Ce total se décompose ainsi :

5128 familles de 5 enfants	634 familles de 10 enfants	37 familles de 15 enfants
3255 » 6 »	384 » 11 »	22 » 16 »
1877 » 7 »	218 » 12 »	11 » 17 »
1619 » 8 »	108 » 13 »	2 » 18 »
872 » 9 »	59 » 14 »	4 » 19 »

Les documents de base proviennent des sources suivantes :

a. Pour la région parisienne : familles inscrites en 1947 aux Associations de Familles (2912 familles);

b. Pour la région du Nord : familles inscrites sur les Annuaires des Grandes Familles (1276 familles).

c. Pour une certaine part des familles de plus de 10 enfants : dossiers de la Fondation Cognacq-Jay (147 familles).

d. Enfin, et surtout, un échantillonnage de 9895 familles extrait des dossiers du Service de la Médaille des Familles au Ministère de la Santé publique et de la Population.

En raison du but poursuivi, nous n'avons tenu compte que des familles sans naissances gémellaires.

Calculé sur l'ensemble des 95875 naissances, le rapport des sexes a été trouvé égal à  $0,5066 \pm 0,0016$ . Nous n'avons pas constaté de variation significative de ce chiffre en fonction du rang de la naissance, mais seulement une indication en faveur des thèses selon lesquelles le premier et les derniers nés sont plus fréquemment des garçons.

Mais les deux faits importants que nous avons établis sont les suivants :

1° Il existe une corrélation faible mais indiscutable entre les sexes des naissances immédiatement consécutives.

2° Cette corrélation décroît très vite lorsqu'augmente l'intervalle entre les naissances, comme le montrent les chiffres suivants :

$$r_0 = 0,029 \pm 0,0035 : \text{corr. entre naissances immédiatement consécutives};$$

$$r_1 = 0,007 \pm 0,0039 : \text{corr. entre naissances séparées par une autre naissance};$$

$$r_2 = 0,004 \pm 0,0043 : \text{corr. entre naissances séparées par deux autres naissances}.$$

1846

ACADÉMIE DES SCIENCES.

Si  $r_0$  et  $r_1$  sont significativement différentes de zéro, il n'est cependant pas possible de prouver que  $r_1$  est supérieure à sa valeur théorique qui résulterait d'un tirage en chaîne simple de Markoff de corrélation  $r_0$ .

Ces faits peuvent s'interpréter, soit en supposant l'existence de *périodes gynophiles* ou *androphiles* portant sur au moins deux gestations, soit en supposant une *influence masculinisante* ou *féminisante* (éventuellement par *augmentation* de la *léthalité*) d'une grossesse masculine ou féminine sur la grossesse immédiatement consécutive. Il n'est pas possible, sur la base de notre information actuelle, de prouver ou d'écarter l'existence de périodes gynophiles ou androphiles de longue durée, non plus que la possibilité d'une prédisposition (peut-être héréditaire) de certains couples à engendrer préférentiellement des garçons ou des filles.

STATISTIQUE. — *Valeurs caractéristiques du coefficient de corrélation par rang de Kendall dans le cas général.* Note de M. MARCEL-PAUL SCHÜTZENBERGER, présentée par M. Gaston Julia.

Kendall <sup>(1)</sup> a défini un *coefficient de corrélation* par rang

$$\tau = \frac{4}{n(n-1)} \sum_{(i,j)} X_{ij} - 1,$$

où  $X_{ij}$  est une indicatrice égale à 1 ou à 0 suivant que la paire  $(i, j)$  est rangée ou non dans le même ordre dans les deux séries S et S' de  $n$  objets à comparer.

Bien que cette expression ne soit pas à proprement parler un véritable *coefficient de corrélation*, son usage se recommande pourtant dans de nombreux cas comme plus commode et plus *naturel* que le coefficient  $\rho$  de Spearman.

Pour étudier la répartition de  $\tau$  quand sa valeur moyenne *n'est pas nulle*, on considérera le modèle probabiliste suivant :

Les classements dans S et S' sont ceux des valeurs hypothétiques  $a_1, a_2, \dots, a_n$  et  $a'_1, a'_2, \dots, a'_n$  prises dans les deux séries de tirages indépendants par les variables aléatoires  $A_i$ , réparties sur un même axe, selon des fonctions  $F_i$  sur lesquelles diverses hypothèses peuvent être faites.

Dans le cas général, les moments d'ordre élevé de la répartition de  $\tau$  ont une expression compliquée, cependant au moyen des coefficients

$$\begin{aligned} \delta_{ij} &= \int_{-\infty}^{+\infty} F_i(t) dF_j(t) - F_j(t) dF_i(t), \\ \gamma_{ijk} &= \int_{-\infty}^{+\infty} \int_0^{+\infty} [F_i(t+t') - F_i(t)][F_k(t+t') - F_k(t)] dF_j(t') dF_j(t), \end{aligned}$$

[qui ne sont autres que les *déviances d'indépendance* définies dans une Note antérieure <sup>(2)</sup>], on peut calculer simplement les valeurs de la moyenne  $\bar{\tau}_{11}$  et de la variance  $\sigma_{11}^2$  de  $\tau_{11}$  entre deux classements aléatoires :

$$\begin{aligned} \bar{\tau}_{11} &= \frac{2}{n(n-1)} \sum_{(i,j)} \delta_{ij}^2, \\ \sigma_{11}^2 &= \frac{8}{n^2(n-1)^2} \sum_{(i,j,k)} S \gamma_{ijk} (\gamma_{ijk} - 2 \delta_{ij} \delta_{jk}) + \frac{4}{n^2(n-1)^2} \sum_{(i,j)} (1 - \delta_{ij}^2) \\ &\quad (S, \text{ sommation sur les 3 permutations : } ijk, jki, kij). \end{aligned}$$

<sup>(1)</sup> *A new measure of rank correlation-Biometrika*, 20, 1938, p. 81.

<sup>(2)</sup> M. P. SCHÜTZENBERGER, *Comptes rendus*, 225, 1947, p. 227.

( 2 )

Remarquons que si, pour  $m$  objets  $a, b, \dots, m$ , la probabilité qu'ils soient observés dans cet ordre est désignée par  $\text{Prob}[a, b, c, \dots, m]$ , on a

$$\begin{aligned}\delta_{ij} &= \text{Prob}[a_i a_j] - \text{Prob}[a_j a_i] \\ \gamma_{ijk} &= \text{Prob}[a_i a_j] \times \text{Prob}[a_j a_k] - \text{Prob}[a_i a_j a_k].\end{aligned}$$

La théorie des *formes polaires* permet d'en déduire les expressions correspondantes  $\bar{\tau}_{12}$  et  $\sigma_{12}^2$  relatives au cas où les deux séries de tirage aléatoire  $S$  et  $S'$  sont effectués dans deux *familles différentes* des fonctions  $F_i$  et  $F'_i$  et l'on vérifie l'inégalité de Schwartz  $\bar{\tau}_{12}^2 \leq \bar{\tau}_{11} \times \bar{\tau}_{22}$ .

Notamment la corrélation  $\tau_{01}$  entre un classement aléatoire et le classement *a priori* où l'on pose  $A_1 < A_2 < A_3 < \dots < A_n$  a pour valeurs caractéristiques

$$\begin{aligned}\bar{\tau}_{01} &= \frac{2}{n(n-1)} \sum_{(i,j)} \delta_{ij} \\ \sigma_{01}^2 &= \frac{8}{n^2(n-1)^2} \sum_{(i,j,k)} (\gamma_{jki} + \gamma_{kij} - \gamma_{ijk}) + \frac{4}{n^2(n-1)^2} \sum_{(i,j)} (1 - \delta_{ij}^2).\end{aligned}$$

Dans tous les cas si  $d$  est une *borne inférieure* des  $|\delta_{ij}|$ ,  $\sigma^2$  est majorée par

$$\frac{2}{3} \times \frac{2n-1}{n(n-1)} \times (1-d^2).$$

Cette valeur est en général *plus grande* que celle donnée par Kendall

$$\frac{2}{9} \times \frac{2n+5}{n(n-1)},$$

mais est *toujours valable* et non pas seulement dans l'hypothèse restrictive où tous les classements de tous les triples d'objets ont la même probabilité  $1/6$ .

(Extrait des *Comptes rendus des séances de l'Académie des Sciences*,  
t. 226, p. 2122-2123, séance du 28 juin 1948.)

Année 1948

1948-5. Fidélité des tests

# TRAVAIL ET MÉTHODES

REVUE TECHNIQUE  
DE LA DIRECTION ET  
DE L'ORGANISATION  
DES ENTREPRISES

*Etablissement d'une succursale en province par M. J. Damancé*  
*La fidélité des tests, par M. M.-P. Schützenberger*  
*Calcul rapide du prix de revient dans un garage station-  
service, par M. H. Grun*  
*Faits et chiffres de la publicité britannique, par M. J. W. Hobson*  
*Utilisation professionnelle des malades et mutilés*  
*Vitesse de réaction et nature des stimulants*  
*Manœuvres d'atterrissage et consignes de piste*  
*La mathématique, objet de culture et outil de travail*  
*Un règlement intérieur de direction*  
*Le tableau de contrôle américain ProducTrol*  
*Pensez à votre hall de réception*

NOVEMBRE  
1948

ÉDITIONS SODERCA 16, AVENUE DE LA LIBÉRATION, PARIS 13<sup>e</sup>

## La fidélité des tests

par M. P. SCHÜTZENBERGER

Il n'est pas besoin de revenir une fois de plus, dans ces colonnes, sur l'avantage que présente dans les problèmes de sélection l'emploi des méthodes rationnelles. Outre le fait qu'elles sont le plus souvent seules utilisables dans le cas où un nombreux personnel doit être sélectionné pour un poste d'embauche, elles présentent en outre un caractère de rigueur qui permet leur emploi et leur comparaison à divers moments et en divers endroits.

Mais les méthodes rationnelles reposent essentiellement sur une mesure de certaines caractéristiques : traits ou aptitudes ou connaissances de l'individu examiné, et il est bon de connaître la précision avec laquelle cette mesure est faite par tel ou tel dispositif psychométrique donné.

Nous n'étudierons ici que le cas le plus simple où la mesure est un *score*, c'est-à-dire une note conventionnelle obtenue en attribuant au sujet un certain nombre de points, déterminé une fois pour toutes, pour chacune des réponses aux différentes questions (ou plus fréquemment « items ») du test.

Le score se trouve donc ainsi être une véritable conciliation dialectique entre le *jugement qualitatif* porté sur un individu au moyen de chaque item (« capable de répondre à cette question ou non ») et une *mesure quantitative* déduite de la totalité de l'épreuve.

Ce processus de mesure, qui n'est d'ailleurs pas spécial et de loin, à la psychologie, présente cependant dans ce domaine une particularité extrêmement importante, celle de comporter, du fait même de son objet, des erreurs ou fluctuations très larges et presque toujours du même ordre de grandeur que les différences inter-individuelles.

Si, par exemple, l'on opère des pesées au moyen d'une balance de type courant, il sera possible de peser des poids allant de 10 grammes à 5 kilos avec une précision d'environ 10 à 15 grammes. Par conséquent l'étendue de l'instrument de mesure est un très grand nombre de fois plus grand que son erreur propre.

Bien au contraire, en psychologie, si la mesure d'une aptitude s'effectue sur une échelle telle que la quasi totalité des sujets voit son score compris entre +50 et -50 il ne sera pas rare, hélas, que chacun de ces scores ne soit connu qu'avec une précision d'environ 10 points.

La nécessité où se sont trouvés les psychologues spécialisés dans ce domaine, la psychométrie, de n'utiliser que de semblables bases numériques pour leur travail, a conduit tout naturellement à un approfondissement et un renouvellement de la théorie de la mesure.

Des outils statistiques nouveaux ont donc dû être forgés permettant de traiter des données peu précises.

Par suite, l'étude des tests en a reçu une nouvelle impulsion et ces méthodes permettent maintenant de construire des épreuves dont la précision dépasse de beaucoup celle que l'on aurait pu espérer il y a seulement une quinzaine d'années.

Il n'est pas inutile de souligner que cette étude de la précision ou *fidélité des tests* n'est pas purement académique. Comme nous le verrons ultérieurement, et comme d'ailleurs le simple bon sens l'exige, la valeur pratique d'une épreuve pour le pronostic de la réussite professionnelle dépend en grande partie de la précision de cette même épreuve et ne peut dépasser une certaine limite, fonction de l'importance des fluctuations.

### Comment mesure-t-on la fidélité d'un test ?

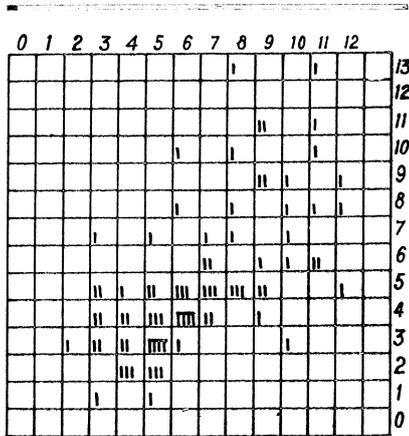
Il convient d'abord de préciser ce que l'on entend par fidélité d'un dispositif quelconque de mesure. Soit par exemple une centaine d'objets de poids différents que l'on pèse avec une balance peu sensible, d'une fidélité de 25 grammes pour fixer les idées, les objets à comparer extrêmes ne différant que d'environ 50 grammes. Une autre estimation de la fidélité de la balance serait la corrélation existant entre les poids indiqués par une première série de mesures et les poids indiqués par une deuxième série ; pour chaque objet en portant en « ordonnée » le poids obtenu à la première mesure et en « abscisse » le poids à la seconde, on pourrait représenter graphiquement l'ensemble de ces deux séries d'opérations par le tableau ci-joint (v. tabl. 1).

L'on appellera alors corrélation un coefficient pouvant varier de  $-1$  à  $+1$  et décrivant la manière dont ces points se répartissent autour de la diagonale principale sur laquelle ils seraient presque tous exactement situés si la fidélité de la balance était très grande (de l'ordre de 1 gramme). On dirait alors que la corrélation est 1. Dans le cas présent, elle serait de  $r = 0,63$ .

Enfin la corrélation entre les deux séries de mesure serait 0 si les différents objets ne différaient entre eux que par une quantité inférieure, par exemple à la fidélité de la balance ; les points se répartissant alors en un nuage où n'apparaîtrait aucune direction principale.

De même pour un test quelconque, en supposant que seules sont en jeu les erreurs de mesure dues à son imperfection propre et non à des modifications des aptitudes de chaque sujet d'une passation à l'autre, on définirait la fidélité par la corrélation entre les mesures dans deux séances consécutives.

Le coefficient ainsi obtenu s'appelle coefficient : « test-retest » et donne certainement, quand il est applicable, l'estimation la meilleure de la fidélité du test.



Dans chaque case : nombre d'objets observés la première fois avec le poids  $p_1$  (ordonnée) et la seconde fois avec le poids  $p_2$  (abscisse).  
Les chiffres indiqués désignent conventionnellement des unités de 5 grammes.  
Ici  $r = 0,63$

Cependant, son emploi n'est en général pas possible. En effet, pour la plupart des tests, l'entraînement dû à une première passation de l'épreuve modifie complètement les résultats de la seconde.

Si l'on imaginait de remplacer le test de la deuxième séance par une épreuve très voisine de la première, on obtiendrait certes un résultat meilleur. Mais encore il convient de voir que l'on a mesuré ensemble deux variations bien distinctes :

1°) La fluctuation propre due au manque de fidélité de l'appareil de mesure.

2°) La fluctuation des sujets d'une expérience à l'autre. Cette deuxième partie peut être fort importante, comme le montre l'exemple ci-dessous où l'on a corrélé entre eux et de toutes les manières possibles les notes obtenues par un groupe important de sujets dans 6 tests pratiquement identiques passés à des intervalles de temps différents : les tests 1 et 2 dans la même séance, un jour, les tests 3 et 4 dans une même séance un autre jour, et les tests 5 et 6 dans une même séance un troisième jour. La diminution des corrélations est parlante par elle-même.

SPLIT HALF RELIABILITY. — Pour obvier aux inconvénients présentés par la méthode précédente, l'on a imaginé de diviser au hasard le test en deux moitiés équivalentes et de comparer les scores qui seraient attribués au sujet en ne considérant alternativement que l'une ou l'autre des moitiés.

Il est clair que l'on élimine ainsi, *ipso facto*, la fluctuation propre des sujets d'une séance à l'autre.

Cependant de nouvelles difficultés apparaissent. En effet, si l'on peut jusqu'à un certain point, lorsque le test est long, se prémunir contre le risque de voir l'une des moitiés mesu-

	1.	2.	3.	4.	5.	6.
1	—	.9457	.9086	.8987	.8643	.8932
2	.9457	—	.9081	.8953	.8609	.8969
3	.9086	.9081	—	.9393	.8806	.8993
4	.8987	.8953	.9393	—	.8527	.8764
5	.8643	.8609	.8806	.8527	—	.9278
6	.8932	.8969	.8993	.8764	.9278	—

Le coefficient à l'intersection de la colonne x et de la ligne y indique la corrélation entre les notes x et y obtenues au test par le sujet.

On remarque la rapide décroissance de cette corrélation à mesure que l'intervalle de temps augmente.

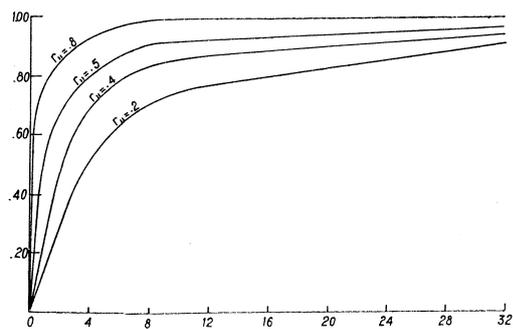
On remarque aussi que les coefficients dans les cases 1,2 — 3,4 et 5,6 correspondant à deux tests passés dans la même séance sont nettement plus élevés que les autres.

rer une aptitude différente de celle qui influe sur la réussite dans la première moitié, il n'en est pas de même lorsque le nombre des items est restreint (une dizaine par exemple). Dans la pratique et pour éviter en outre que l'une des moitiés ne comprenne systématiquement des items situés soit à la fin soit au début des tests pour lesquels on pourrait penser que l'attention ou l'entraînement du sujet influe de manière sensible, on a l'habitude de comparer les scores du sujet :

- 1) dans les questions paires ;
- 2) dans les questions impaires.

Il est évident, en outre que la fidélité de la note totale d'un test est fonction de la longueur de celui-ci.

Chaque item représentant une sorte de mesure élémentaire, il est assez intuitif que les erreurs sur chacune de ces opérations simples peuvent se compenser en moyenne quand leur nombre augmente et que la note qui résulte de 50 items, par exemple, est plus fidèle et précise que celle qui résulterait de 10 items seulement.



Cette figure donne la fidélité d'un test que l'on allonge N... fois.

En « abscisses » est indiqué le nombre de fois dont on allonge le test ;

En « ordonnée » la fidélité du test allongé. Les courbes tracées correspondent à la fidélité du test initial.

En particulier, la corrélation  $r$  que nous avons considérée plus haut entre deux passations théoriques consécutives permet d'en déduire la corrélation  $r_n$  analogue pour un test semblable  $n$  fois plus long.

C'est la fameuse formule de Spearman-Brown :

$$r_n = \frac{n \cdot r}{1 + (n - 1) r}$$

En particulier lorsque nous avons considéré les notes dans les deux moitiés du test, nous n'avons estimé que la fidélité d'un test de longueur deux fois plus faible que le test réel et la fidélité réelle est donnée par :

$$r = \frac{2r'}{1 + r'}$$

Mais cette méthode est encore insuffisante car elle a toujours présupposé que chacun des items du test mesurait bien la même aptitude (ou trait, ou connaissance) des sujets. On a donc été amené à imaginer des méthodes plus élaborées, sur le détail desquelles nous ne pouvons entrer ici, et qui consistent essentiellement à comparer le classement que donne chacun des items avec le classement total du test : si le test est homogène, c'est-à-dire mesure bien un trait ou une aptitude unique, ces classements doivent être semblables en gros.

Dans la pratique cependant, il semble bien qu'il n'y ait que rarement avantage à employer cette méthode d'analyse complexe qui nécessite des calculs très longs, et que l'on ait avantage à se concentrer le plus souvent de la split half reliability de Spearman-Brown.

D'autres aspects influent encore sur la fidélité que nous ne pouvons traiter entièrement ici : un des plus importants est ce que l'on pourrait appeler le degré de fermeture du test. On désigne par ce terme le nombre de réponses possibles qui sont offertes au sujet pour chaque question. On distingue ainsi les questions auxquelles il ne peut être répondu que par oui ou non, les questions où 2, 3, 4, 5 ou même 6 alternatives sont possibles.

Il est assez immédiat (et le calcul le confirme) que la fidélité du test est augmentée avec le nombre des alternatives offertes au sujet, car la probabilité, pour qu'un sujet trouve la bonne réponse simplement en répondant au hasard, est diminuée dans des proportions considérables. Pour corriger cette influence du hasard, le score  $s$  d'une feuille de réponse où il a été bien répondu à  $B$  questions, mal à  $M$  et où  $N$  questions sont restées sans réponse est donné par convention par la formule :

$$s = (B - M) N/n$$

où  $n$  est le nombre des alternatives.

D'autre part, des expériences nombreuses ont montré que la fidélité d'une épreuve était sensiblement accrue lorsque les sujets possédaient un entraînement suffisant à résoudre ce type de problème. C'est pourquoi dans la plupart des tests bien construits, 5 à 10 minutes du temps avant la passation sont laissées au sujet pour s'entraîner sur des questions analogues à celles qui lui seront posées ultérieurement et dont on ne tient aucun compte dans la correction définitive.

Ces diverses remarques nous permettent de donner quelques brèves indications sur la manière dont peut s'étudier

rapidement une épreuve nouvelle lorsque les exigences d'un poste de travail nécessitent l'élaboration d'un test spécial.

Il va sans dire que toutes ces indications sont données au sujet des tests « papier-crayon » que l'on tend à employer de plus en plus fréquemment mais qu'elles s'étendent, mutatis mutandis, aux autres types d'épreuves (tests d'appareils, etc.) :

1) Construire un nombre d'items environ deux fois supérieur à celui que l'on a l'intention d'inclure dans la forme définitive des tests ; utiliser de préférence des questions où les diverses alternatives possibles ont, *a priori*, la même chance d'être désignées par le sujet ; éviter les questions dont la solution est une véritable devinette et qui, de ce fait, font appel à des aptitudes différentes de celles que doit mesurer le test ; prévoir un nombre suffisant de questions pour exercer le sujet.

2) Administrer à un groupe homogène d'une centaine de sujets de même niveau que ceux auxquels est destinée ultérieurement l'épreuve, la première forme ainsi préparée.

3) Pour chaque item dénombrer le nombre de fois pour lequel la bonne réponse a été indiquée ; éliminer de la correction celles des questions auxquelles moins de 20 % ou plus de 80 % ont correctement répondu.

4) Attribuer à chaque sujet sur la base des items restant un score égal au nombre de ses bonnes réponses et vérifier que la courbe de répartition de ces scores a une forme suffisamment voisine de la normale.

5) Calculer la Split half reliability.

6) Si celle-ci est inférieure à la borne que l'on s'est fixée, il convient alors d'éliminer les items pour lesquels une forte proportion de sujets, considérés comme les meilleurs d'après le score total, a mal répondu (une limite commode est plus de 75 % des sujets classés dans les 25 % les meilleurs d'après le score total) ; rajouter de nouvelles questions et refaire passer sur une nouvelle population.

Cette opération doit souvent être répétée 2 ou 3 fois jusqu'à l'obtention d'une fidélité suffisante pour le but que l'on se propose.

Il nous faut maintenant examiner ce que signifient concrètement les coefficients de fidélité que nous venons d'étudier et quels sont les rapports de cette notion avec la validité ou efficacité des tests.

Une interprétation fort simple du coefficient de fidélité est la suivante :

Supposons que nous ignorions tout de la valeur de l'aptitude d'un sujet à un test dont les notes s'échelonnent environ de -50 à +50. *A priori* il nous est donc impossible de dire à quel point de cet échelle la note du sujet se trouvera.

Si nous connaissons déjà la note du sujet dans une épreuve très voisine de l'épreuve en question et dont la corrélation avec cette dernière se trouve être égale à  $r$  nous pouvons savoir que cette étendue se trouve diminuée dans la proportion de  $1 - r$ .

D'après ce qui précède il est donc clair qu'à moins que la fidélité n'atteigne une valeur extrêmement élevée, telle que 0,97 ou 0,98, l'amélioration de la connaissance de la valeur vraie du sujet dans un test n'est que très faible.

Ce point est important car les épreuves sont rares dont la fidélité dépasse 0,80 et notamment dans le domaine caractériel.

L'on a ainsi une idée plus précise des services que peut rendre la psychométrie susceptible d'apporter des renseignements précieux lorsqu'il s'agit de comparer des moyennes de notes relatives à un groupe de sujets, mais souvent encore trop imprécise pour pouvoir être utilisée dans la prédiction des possibilités d'un sujet donné, comme dans l'orientation professionnelle.

Sociométrie.

**Étude statistique d'un problème  
de sociométrie**

par M. M. P. SCHÜTZENBERGER.

*L'article est consacré à la formulation du traitement statistique de certains aspects d'une enquête sociométrique « par choix ». Un exemple d'application est donné, qui conduit à rejeter l'hypothèse nulle.*

Le travail suivant est issu d'une enquête du groupe de Psychométrie Pédagogique (G.P.P.) sous la direction de M. P. H. MAUCORPS au cours de l'année 1945-1946 dans quatre classes de préparation de l'Ecole Navale. L'une des séries de questions posées aux élèves était la suivante :

« Parmi les élèves qui appartiennent à la même classe de préparation que vous, désignez :

- A. - Votre meilleur ami ;
- B. - Vos quatre meilleurs amis (y compris le précédent) ;
- C. - Les trois élèves qui vous sont les plus antipathiques ».

Parmi les multiples problèmes qui peuvent se poser au cours du dépouillement de semblables matériaux, on a choisi de traiter ici la question suivante :

Dans quelle mesure peut-on affirmer que le nombre de

## PROBLÈME DE SOCIOMÉTRIE

97

réciprocités — c'est-à-dire de désignations (choix) simultanés du sujet  $i$  par le sujet  $j$  et du sujet  $j$  par le sujet  $i$  est supérieur à celui qui résulterait d'une répartition aléatoire de ces mêmes choix.

C'est ce qui constitue « l'hypothèse nulle » ( $H_0$ ) en fonction de laquelle seront calculées toutes les statistiques suivantes.

On établira des tests permettant de savoir quelle est la probabilité pour que dans « l'hypothèse nulle » les valeurs observées diffèrent d'une certaine mesure des valeurs théoriques. Quand cette probabilité tombera au-dessous d'un certain niveau (en général 5 %, 1 % ou 1 °/00) on sera amené à écarter l'hypothèse nulle. Ceci constitue le mécanisme habituel du raisonnement statistique.

Les formules approchées précédemment calculées par MORENO et JENNINGS (Statistics of Social configuration - Sociometry 1938 pp. 342-374), ne s'appliquant que difficilement au cas actuel où il n'a pas été possible de recueillir toutes les réponses, on a été conduit à formuler le problème d'une manière rigoureuse et générale permettant un test plus sûr.

**Notations.**

Soient  $1, 2, \dots, n$  les  $n$  sujets dont on a recueilli les choix (sujets « actifs ») et  $n+1, \dots, N$  les  $N - n$  sujets du groupe qui ont pu être désignés mais dont les choix sont inconnus (sujets « passifs », — soit  $k$  le nombre fixe des choix émis par chaque sujet (ici :  $k_A = 1, k_B = 4, k_C = 3$ ).

On aura avantage à représenter l'ensemble de l'enquête par une matrice  $U$  d'ordre  $n \times U$  où les éléments  $a_{ij}^i$  (à l'intersection de la colonne  $i$  et de la ligne  $j$ ) sont égaux à 1 ou nuls suivant que le  $j$  ème sujet a ou non choisi le  $i$  ème.

On appellera réciprocity  $C_{ij}$  l'événement

$$a_{ij}^j \times a_{ji}^i = 1$$

On a évidemment :

$$a_{ii}^i = 0 \text{ (aucun sujet ne peut se choisir lui-même)}$$

$$\sum_j a_{ij}^i = k$$

$$0 \leq i < N + 1$$

7

La probabilité a priori de  $a_{ij} = 1$  ( $i \neq j$ ) sera  $p_{ij}$  et l'hypothèse nulle  $H_0$  sera :

$$1^\circ p_{ij} = \frac{k}{N-1} = p$$

$$2^\circ \text{ Prob. } C_{ij} = p_{ij} p_{ji}$$

$$\text{On posera en outre : } q = 1 - p = \frac{N - k - 1}{N - 1}$$

$$d = n/N$$

Par abréviation on appellera :

1° « cas carré » le cas où  $d = 1$   
(Tous les sujets sont « actifs »).

2° « cas de Poisson » le cas infini où,  $k$  restant fixe,  $d$  a une limite non nulle  $\delta_1$

#### Formules préliminaires.

Soit  $f_x$  le nombre de sujets ayant reçu  $x$  choix ; il convient d'abord de tester l'accord de la répartition des  $f_x$  avec l'hypothèse nulle  $H_0$ .

Les valeurs théoriques de  $f_x$  sont :

$$f_x = \binom{n}{x} p^x q^{n-x} (n-x+q^n)^*$$

Ce qui conduit à une répartition dont les valeurs typiques sont

$$m = p d (N-1) \quad ; \quad \sigma^2 = (N-1) p d (1-pd)$$

Un autre moyen de tester  $H_0$  serait de comparer le nombre de sujets n'ayant jamais été choisi à sa valeur a priori  $f_0$   
En posant :

$$\Theta_r = \frac{(N-p)! (N-k-1)!}{(N-1)! (N-k-r)!}$$

On emploiera la notation  $\begin{bmatrix} a \\ b \end{bmatrix}$  pour les coefficients binomiaux.

$$* \quad \begin{bmatrix} A \\ a \end{bmatrix} = \frac{A!}{(A-a)! a!} = C_A^a$$

PROBLÈME DE SOCIOMÉTRIE

99

on trouve les moments factoriels de la répartition de fo

$$B_r = (\theta_{r..})^n \sum_{f=0}^{n+1} \binom{n}{f} \binom{N-n}{r-f} \left( \frac{N-r}{N-r-k} \right)^f$$

$$= (\theta_{r..})^n \sum_{f=0}^{r+1} \binom{n}{f} \binom{N-f}{r-f} \left( \frac{k}{N-r-k} \right)^f$$

Ces formules ne deviennent maniables que dans le cas carré où :

$$B_r = \binom{N}{r} (\theta_{r..})^{n-r} (\theta_r)^r$$

Cependant on a toujours pour valeur moyenne de fo :

$$B_1 = \bar{f}_c = N q^{n-1} (q + dp)$$

Et dans le cas de Poisson  $\bar{f}_0$  tend vers  $e^{-k\delta}$

Réciprocités.

On formulera d'abord un théorème général relatif à la répartition *a priori* de la répétition d'une configuration contenue dans une sous-matrice diagonale U' de U.

*Théorème.* - Pour un U' formé de  $l_1$  lignes à 1 choix.  
 $l_2$  lignes à 2 choix.  
 $l_3$  lignes à 3 choix

et dont les lignes sont elles-mêmes fixées, la probabilité a priori est :

$$p_{U'} = \binom{N-1}{k}^{n-1} \prod_{i=0}^{k+1} \binom{N-i-1}{k-i}^{l_i} = \prod \left\{ \frac{k!}{(k-i)!} \times \frac{(N-i-1)!}{(N-1)!} \right\}^{l_i}$$

Par conséquent si g est l'ordre du groupe de substitution de U' la valeur moyenne du nombre de répétition de U<sub>i</sub> dans l'hypothèse H<sub>0</sub> est

$$\bar{r}_{U'} = p_{U'} \times \frac{n!}{(n-\sum l_i)!} \times g^{-1}$$

100

GALLICA BIOLOGICA ACTA

**Calculs des moments factoriels.**

On considère successivement la matrice  $U'$  d'ordre  $2 \times 2$  que constitue une réciprocité puis les matrices  $U'$  constituant une double réciprocité et l'on obtient les deux premiers moments factoriels d'où les valeurs typiques

011	0100	$\bar{p} = p^2 \frac{n(n-1)}{2} ; \quad \sigma^2 = \bar{p} q \left( q + 2 \frac{N-n}{N-2} p \right)$
100	1000	
10G	0001	
	0010	

d'où l'on déduit que  $\sigma_r^2$  est *toujours* compris entre  $\bar{r}$  et  $\bar{r}q^2$ .

Le calcul des moments du 3<sup>e</sup> ordre entraînerait l'application du théorème général aux matrices suivantes :

0111	0110	00010	010000	011
1000	1001	00001	100000	101
1000	1000	10000	000100	110
1000	0100	10000	001001	
		01100	000001	

d'où la valeur suivante du moment factoriel de troisième ordre :

$$\begin{aligned}
 B_3 = & \bar{p} p^4 \frac{(n-2)(n-3)(n-4)(n-5)}{24} + \bar{p} p^3 \frac{(n-2)(n-3)(n-4)}{2} \frac{k-1}{N-2} \\
 & + \bar{p} p^3 (n-2)(n-3) \left( \frac{k-1}{N-2} \right)^2 + \bar{p} p^2 \frac{(n-2)(n-3)}{3} \frac{(k-1)(k-2)}{(N-2)(N-3)} \\
 & + \bar{p} p \frac{(n-2)}{3} \left( \frac{k-1}{N-2} \right)^3
 \end{aligned}$$

L'expression du moment du 4<sup>e</sup> ordre serait encore plus compliquée (11 termes). Cependant, pour  $k = 1$ , on a les formules simples :

$$B_r = \frac{n!}{(n-2r)!r!} 2^{-r} (N-1)^{-2r}$$

## PROBLÈME DE SOCIOMÉTRIE

101

car une seule matrice  $U'$  est seulement à considérer et l'emploi des inégalités de Loeve ou de Bonferati est extrêmement facile.

**Cas de Poisson.**

Il en est ici de même car la probabilité des  $U'$  ayant deux choix dans une ligne est infiniment petite par rapport aux  $U'$  n'ayant qu'un choix par ligne.

Un calcul classique montre la répétition des réciprocités, suit une loi de Poisson de paramètre  $k^2 \delta^2/2$ .

Il est vraisemblable que dans la pratique et dès que  $k/N-1$  est suffisamment petit, cette approximation est applicable au cas fini.

**Données incomplètes.**

Ce cas est celui où les sujets n'ont pas tous effectué le même nombre de  $k$  de choix ; il est encore possible, si l'on connaît pour chaque sujet actif les nombres  $k_i$  et  $h_i$  de choix *émis* et reçus, d'obtenir une *valeur rigoureuse* de  $\bar{r}$  :

$$\bar{r} = \frac{1}{2} \sum_{i=0}^{n+1} \frac{h_i k_i}{N-1} .$$

que l'on pourra utiliser en 1<sup>re</sup> approximation comme paramètre d'une loi de Poisson (si  $N$  est petit).

Il faut signaler que dans le cas carré et quand tous les  $k_i$  sont égaux à  $k$  l'on retrouve la valeur précédemment calculée dans l'hypothèse  $H_0$ .

Ainsi donc, dans ce cas, la connaissance complète de la distribution marginale des choix ne *change en rien* la valeur moyenne a priori du nombre des réciprocités.

**Extension.**

Le théorème général permet également d'étudier des configurations plus complexes : par exemple les chaînes d'ordre  $\lambda$  c'est-à-dire des événements  $C(i, j, k, \dots)$  constitués par le choix simultané de  $i$  par  $j$ , de  $j$  par  $k$ , de  $k$  par  $l$  etc... ( $\lambda$  sujets) : on trouve encore pour valeur moyenne du nombre des répétitions de chaîne d'ordre  $\lambda$  :

102

GALLICA BIOLOGICA ACTA

$$\bar{r}_\lambda = \frac{n!}{(n-\lambda)!} \frac{(N-1)^{-\lambda}}{\lambda} K^\lambda$$

et dans le cas de Poisson une distribution de Poisson de paramètre :

$$K^\lambda e^{-K}$$

On peut également (toujours par application du théorème général) chercher la valeur moyenne du nombre de « blocs compacts d'ordre  $\mu$  » c'est-à-dire d'ensemble de  $\mu$  ( $\mu \leq k + 1$ ) sujets se choisissant tous entre eux; les matrices  $U'$  sont alors :

$\mu = 2$	$\mu = 3$	$\mu = 4$
01	011	0111
10	101	1011
(réciprocité)	110	1101 etc...
	(double chaîne d'ordre 3)	1110

on a sans peine :

$$\bar{r}_\mu = \binom{n}{\mu} K \left\{ \frac{K!}{(k-\mu+1)!} \times \frac{(N-\mu)!}{(N-1)!} \right\}^\mu$$

$\bar{r}_\mu$  tend vers zéro pour  $\mu \geq 3$  lorsque,  $k$  étant fixe, tend vers l'infini.

#### Application à l'enquête G.P.P.

On utilisera les résultats précédents pour tester l'hypothèse nulle relativement à chacun des trois types de choix indiqués.

Le tableau I donne pour chacune des quatre classes le nombre  $N$  d'élèves dans la classe, le nombre de sujets ayant répondu au questionnaire et les nombres AA, BB et CC de réciprocity pour les choix A. B. et C.

Le tableau II, donnant les valeurs typiques  $\bar{r}$  et  $\sigma^2$  pour chacun de ces  $4 \times 3 = 12$  cas, permet de voir qu'il serait fort légitime d'utiliser une répartition de Poisson pour tester  $H_0$ . Cependant dans le cas AA, pour la 1<sup>re</sup> et la 2<sup>e</sup> classe,  $k$  étant égal à 1, on a appliqué les formules classiques :

PROBLÈME DE SOCIOMÉTRIE

103

$$P_r = B_r - [1]B_{r,1} + [2]B_{r,2} - \dots$$

TABLEAU I.

Classe	N	n	AA k=1	BB k=4	CC k=3
1	44	36	3	18	5
2	36	33	6	22	10
3	39	27	3	17	2
4	37	20	7	17	2

Données de l'enquête :

N = nombre de sujets dans la classe.

n = nombre de sujets dont les réponses ont été recueillies.

AA = nombre de réciprocités pour le 1<sup>er</sup> choix.

BB = nombre de réciprocités pour le 2<sup>e</sup> choix.

CC = nombre de réciprocités pour le 3<sup>e</sup> choix.

TABLEAU II.

Valeur théorique dans l'hypothèse nulle.

Classe	AA Réciprocités	Réciprocités BB	Réciprocités CC
1	$\bar{r} = 0,341$ $\sigma^2 = 0,328$ $\sigma = 0,573$	$= 5,45$ $= 4,66$ $= 2,16$	$= 3,07$ $= 2,73$ $= 1,65$
2	$\bar{r} = 0,431$ $\sigma^2 = 0,404$ $\sigma = 0,639$	$= 6,90$ $= 5,54$ $= 2,35$	$= 3,88$ $= 2,62$ $= 1,62$
3	$\bar{r} = 0,243$ $\sigma^2 = 0,234$ $\sigma = 0,436$	$= 3,89$ $= 3,35$ $= 1,83$	$= 2,19$ $= 1,96$ $= 1,40$
4	$\bar{r} = 0,147$ $\sigma^2 = 0,142$ $\sigma = 0,377$	$= 2,34$ $= 2,07$ $= 1,44$	$= 1,32$ $= 1,21$ $= 1,10$

$\bar{r}$  = valeur moyenne du nombre de réciprocités.

$\sigma^2$  = écart type > > >

qui permettent d'affirmer rigoureusement que l'hypothèse nulle (évidemment rejetée dans la 2<sup>e</sup> et la 4<sup>e</sup> classe a beaucoup plus de 1 0/00) doit également l'être dans la 1<sup>re</sup> et la 3<sup>e</sup> classe, avec des probabilités de 13 o/oo et de 14 o/oo respectivement.

La signification des écarts entre les valeurs théoriques et les valeurs observées pour BB est hors de contestation (plus d'une chance sur mille), soit que l'on emploie une approximation de Poisson (classes 3 et 4) soit que l'on utilise la règle des  $3\sigma$  (classes 1 et 2).

En ce qui concerne CC il semble difficile de rejeter  $H_0$  pour les classes 1, 3 et 4 où la valeur empirique est très voisine de la moyenne. L'étude de la classe 2 est relativement facile : en effet le théorème général donne l'expression de chacun des termes qui contribuent à former les moments factoriels et il est aisé de voir que si une matrice  $U'$  (pour l'ordre  $\rho$ ) est une matrice  $(2\rho - x) \times (2\rho - x)$  l'ordre en  $n$  et  $N$  du terme correspondant est  $-x$  ; l'on est ainsi amené à ne prendre en considération que les  $U'$  d'ordre  $2\rho$ ,  $2\rho - 1$  et  $2\rho - 2$ , les autres  $U'$  n'introduisent que des différences de l'ordre de  $1^{\circ}/10000^{\circ}$  chacune. On trouve aussi que la probabilité d'atteindre par le seul jeu des fluctuations 10 réciprociétés est inférieure à 4 %.

### RESUME ET CONCLUSIONS

Diverses statistiques ont été établies pour analyser une « enquête par choix » dans quatre classes de préparation aux grandes écoles.

La valeur exacte de la moyenne et de l'écart type de la répétition du nombre de choix réciproques ont été calculées dans l'hypothèse nulle de l'indépendance et de l'équiprobabilité des choix.

Une application en a été faite d'où il semble ressortir que les antipathies ne manifestent le phénomène de réciprociétés qu'à un degré beaucoup moindre que les sympathies pour lesquelles l'hypothèse de la répartition indépendante des choix est absolument à rejeter.

# GALLICA BIOLOGICA ACTA

---

VOLUME 1  
FASCICULE 2  
AVRIL-JUIN 1948

---

LIBRAIRIE LE FRANÇOIS  
91, Boulevard Saint-Germain - PARIS

# GALLICA BIOLOGICA ACTA

publiés sous la direction de P. Gavaudan  
*Directeur du Service de Biologie  
des Services Chimiques de l'Etat*

VOL. 1 FASC. 2 JUIN 1948

## S O M M A I R E

### Articles originaux.

- S. Bazin** : Etude de la pénétration de diverses substances chimiques dans le système vacuolaire de « Chara » .... p. 107
- P. Gavaudan et H. Poussel** : Remarques sur les théories de la cancérisation chimique ..... p. 111
- H. Poussel** : Sur les mécanismes des actions toxiques de l'hexachlorocyclohexane  $\gamma$  ..... p. 114
- P. Gavaudan, H. Vogel et G. Debraux** : Etudes des propriétés rhizogènes de la  $\beta$  lactone de l'acide  $\beta$  isopropylmalonique . . . . . p. 120
- G. Brebion et P. Gavaudan** : Etude de la photosynthèse des plantes aériennes par la méthode de Warburg ..... p. 124
- G. Debraux** : Recherches sur les feuilles isolées en survie. .... p. 131
- G. Debraux et P. Gavaudan** : Recherches sur la prophylaxie du mildiou de la vigne par les composés organiques ..... p. 138
- P. Gavaudan, G. Brebion, H. Poussel et M. P. Schützenberger** : Etude pharmacodynamique des chimiorécepteurs de l'olfaction; (1) Principes et méthodes d'une étude thermodynamique de l'olfaction . . . . . p. 147
- P. Gavaudan, H. Poussel et M. P. Schützenberger** : Etude pharmacodynamique des chimiorécepteurs de l'olfaction; (2) L'étude thermodynamique des séries homologues et les théories de l'olfaction ..... p. 167
- N. Kobozieff et N. A. Pomriaskinsky-Kobozieff** : Possibilité de sélection sur l'anomalie de l'oreille externe chez la souris. p. 184
- M. Schützenberger** : Remarques sur les relations aléatoires d'ordre ..... p. 187

### Revue générale.

- P. Gavaudan** : Echanges de matériaux figurés entre noyau et cytoplasme..... p. 205

LIBRAIRIE LE FRANÇOIS  
91, Boulevard Saint-Germain - PARIS

Physiologie des sensations.

**Étude pharmacodynamique des  
chimiorécepteurs de l'olfaction.**

(I) *Principes et méthodes d'une étude thermodynamique  
de l'olfaction*

par P. GAUUDAN, G. BREBION, H. POUSSEL  
et M. P. SCHÜTZENBERGER.

*Dans ce travail, après un exposé critique concernant les méthodes utilisées jusqu'à présent pour étudier le mécanisme de l'olfaction et exprimer des seuils, on propose de donner une meilleure définition physique des conditions de déclenchement de la sensation olfactive en utilisant la notion d'activité thermodynamique. On propose d'exprimer les seuils de sensation en fonction de l'activité thermodynamique des vapeurs odorantes et on décrit des méthodes propres à déterminer directement les seuils en fonction de l'activité.*

**I. Examen des méthodes olfactométriques anciennes basées sur l'expression de la concentration des substances odorantes en poids ou sur des unités olfactives arbitraires.**

Nous avons précédemment étudié les mécanismes de la narcose (1 et 2) et de l'excitation sapide en appliquant la notion d'activité thermodynamique (3 et 4). Nous avons considéré le problème de l'excitation gustative comme un problème pharmacodynamique normal susceptible d'être étudié par les méthodes classiques. Nous essayerons d'appliquer les mêmes principes à l'étude du mécanisme de l'olfaction. Ce mécanisme est encore à l'heure actuelle complètement obscur les hypothèses proposées sont nombreuses mais elles doivent faire l'objet d'un contrôle sévère, car les données sur les-

quelles elles reposent ont été établies dans des conditions expérimentales qui ne nous renseignent pas exactement sur l'état physique du système étudié.

Depuis ZWAARDEMAKER (5) et PASSY (6), les pionniers des études sur l'olfaction, on a l'habitude en physiologie des sensations d'exprimer les seuils olfactifs en poids de substance ou en M.g./L.

Nous ne reviendrons que très rapidement sur les diverses méthodes olfactométriques qui ont été utilisées jusqu'à présent.

Elles ont en général utilisé les principes suivants :

1° *Méthode de PASSY* (6). Dilution d'une quantité connue des substances odorantes dans l'alcool permettant d'obtenir par évaporation une concentration connue dans un volume d'air donné.

PASSY utilisait une bouteille d'un litre ; dans diverses variantes de cette méthode les auteurs ont employé des caissons de grande capacité permettant au sujet de respirer plus normalement.

2° *Méthode de ZWAARDEMAKER*. Cette méthode est elle-même inspirée de celle imaginée par REUTER (7). En principe on fait passer l'air inspiré sur une surface variable recouverte de substance odorante. Pratiquement on utilise deux tubes coulissant l'un dans l'autre et la mesure olfactométrique s'effectue par appréciation de la longueur sur laquelle l'air inspiré a lèché la surface odorante. Se basant sur les résultats obtenus avec des olfactomètres basés sur le principe, ZWAARDEMAKER a établi une notion d'« olfactie » qui n'a qu'une valeur purement relative et empirique et paraît n'avoir ni signification physiologique ni sens physique.

3° Divers osmoscopes à dilution ont été décrits ; notamment celui de BACH (8) qui comprend une série de flacons reliés par des tubulures ; l'appareil permet de diluer la vapeur contenue dans un premier récipient. Il faut citer également l'appareil de HOFFMAN et KOHLRAUSCH (9), basé sur le principe suivant. L'appareil comprend un réservoir de vapeur saturante et des burettes à mercure permettant de prélever une quantité donnée de cette vapeur de la diluer et de la faire respirer au sujet.

Connaissant la température, la pression de vapeur satu-

rante et le poids moléculaire de la substance, les auteurs calculaient les concentrations seuils en gr/L. L'appareil est certainement un des plus précis qui ait été proposé mais les résultats sont exprimés en concentrations en gr./L. Cet appareil a été perfectionné par MORIMURA (10) par addition de burettes supplémentaires et dispositifs thermostatiques, mais l'utilisation est la même.

## II. Olfactométrie basée sur la notion d'activité thermodynamique de la vapeur odorante.

Dans toutes ces recherches les seuils olfactifs ont été exprimés en poids de substance ou en fraction de molécule-gramme par litre. L'olfactométrie, qui est à la base des recherches sur l'olfaction est donc elle-même fondée, comme on va le voir sur un principe inexact, qui ne peut servir à établir des comparaisons valables entre les seuils et la constitution chimique des substances. C'est là peut-être, une des causes qui ont empêché de pénétrer le mécanisme de l'olfaction jusqu'à présent.

En effet, lorsque l'on compare les seuils de deux alcools homologues de la série aliphatique, l'alcool méthylique et l'alcool octylique, par exemple, on indique simplement le poids de substance dans l'atmosphère ; mais ces alcools ont des tensions de vapeur très différentes, celle de l'alcool méthylique est de 95 mm. de mercure et celle de l'alcool octylique de quelques dixièmes de millimètre. Il en résulte que dans les deux cas l'état physique pourra être tout à fait différent comme nous le montrerons plus loin. On se trouve avec les deux alcools, aux concentrations données (en poids) à des distances très différentes de la tension de vapeur saturante. Cela revient

à dire que les rapports  $\frac{p}{p^0}$ , où  $p$  est la pression de vapeur partielle et  $p^0$  la tension de vapeur saturante sont très différents.

Ces rapports  $\frac{p}{p^0}$  ne sont autre chose que l'activité thermodynamique du système, que l'on représente par :

$$a = \frac{p}{p^0}$$

La valeur de ce rapport correspond à la tendance que

manifeste la substance à sortir de la phase gazeuse où elle se trouve. L'activité est une valeur relative. Celle de la pression de vapeur saturante est égale à l'unité. Nous ne reviendrons pas ici sur les notions de thermodynamique qui conduisent à cette valeur et qui ont fait l'objet d'exposés antérieurs (1 et 2) destinés aux biologistes.

Nous avons complètement renoncé dans cette étude à exprimer les seuils de sensation olfactive en fonction de concentrations en poids ou de fractions de molécule-gramme par litre d'atmosphère inspirée. Les seuils ont été exprimés en activités thermodynamiques. C'est la seule façon de comparer l'activité physiologique des substances odorantes en se plaçant dans les mêmes conditions physiques.

Nous avons étudié et mis en œuvre plusieurs méthodes :

**1° Méthode activométrique basée sur le principe de l'olfactomètre de Hoffman et Kohlrausch.**

On a vu que HOFFMAN et KOHLRAUSCH se sont servis d'un appareil leur permettant de diluer une quantité connue de vapeur saturante et de calculer à partir de cette dilution une concentration en poids.

Il est évident que cet appareil indique, en principe, directement l'activité de l'atmosphère conditionnée de cette façon. En vertu de la loi de Mariotte un volume donné de vapeur saturante amené à un second volume connu fournit une activité proportionnelle au rapport des volumes choisis.

Nous avons d'abord entrepris des déterminations de seuils avec un olfactomètre de HOFFMAN et KOHLRAUSCH dont le récipient de dilution était constitué par une burette de 100 cm<sup>3</sup> et dont le récipient de mesure de vapeur saturante était une burette de 5 cm graduée au 1/100<sup>e</sup>. Un volume de vapeur saturante, par exemple 0,1 cm<sup>3</sup>, était mesuré et introduit dans le récipient dilueur par la manœuvre des ampoules à mercure. Ces dernières, ainsi que les robinets de communication étaient manœuvrés de telle façon que la vapeur saturante était admise sous un léger vide réalisé dans le récipient de dilution ; on obtenait ainsi une légère détente créant une turbulence permettant le mélange. En effet, un des inconvénients de cet appareil est la fréquente médiocrité du mélange de la vapeur saturante avec l'air. Le second inconvénient réside dans la très faible quantité d'air à la disposition du sujet (100 cm<sup>3</sup>).

Pour remédier à ce dernier inconvénient nous avons construit un second appareil équipé avec une burette de dilution de 1.000 cm<sup>3</sup> de capacité construite en verre épais. Pour éviter la manipulation d'une quinzaine de kilogs de mercure contenu dans un gros récipient destiné à être alternativement abaissé et remonté, le déplacement du métal était assuré par l'artifice suivant : l'extrémité inférieure de la burette de dilution était en relation avec un récipient pouvant être mis sous pression au moyen d'une bouteille à air comprimé. La montée du mercure dans la burette s'effectuait sous l'action de la pression de l'air, la descente se réalisait par détente au moyen d'un robinet. En outre, la burette était munie vers son extrémité inférieure et supérieure de deux canalisations reliées à une petite pompe en verre aspirante et foulante munie d'une soupape en verre ; ce dispositif était destiné à obtenir un mélange homogène de la dilution avant le refoulement vers les narines du sujet. Le sujet avait ainsi à sa disposition un litre de dilution de vapeur odorante ; ce volume pouvait être refoulé en un temps variable car on pouvait régler à volonté la vitesse d'ascension du mercure en agissant sur un robinet de commande reliant le pot de mercure sous pression à la burette de dilution. Malgré tous les perfectionnements qu'il comportait par rapport à l'appareil de HOFFMAN et KOHLRAUSCH cet appareil a été abandonné, car l'inconvénient commun des appareils de ce type est l'impossibilité d'étudier les phénomènes de fatigue ou de compensation des odeurs, d'une part, et la difficulté technique de la parfaite désodorisation lorsque l'on change de substance, d'autre part. L'appareil comportait une circulation d'air chaud permettant un lavage des canalisations, mais il était nécessaire de nettoyer périodiquement le mercure. En outre, l'adsorption par le mercure et par les parois en verre dans les conditions d'emploi constituait une cause d'erreurs dans la mesure des seuils.

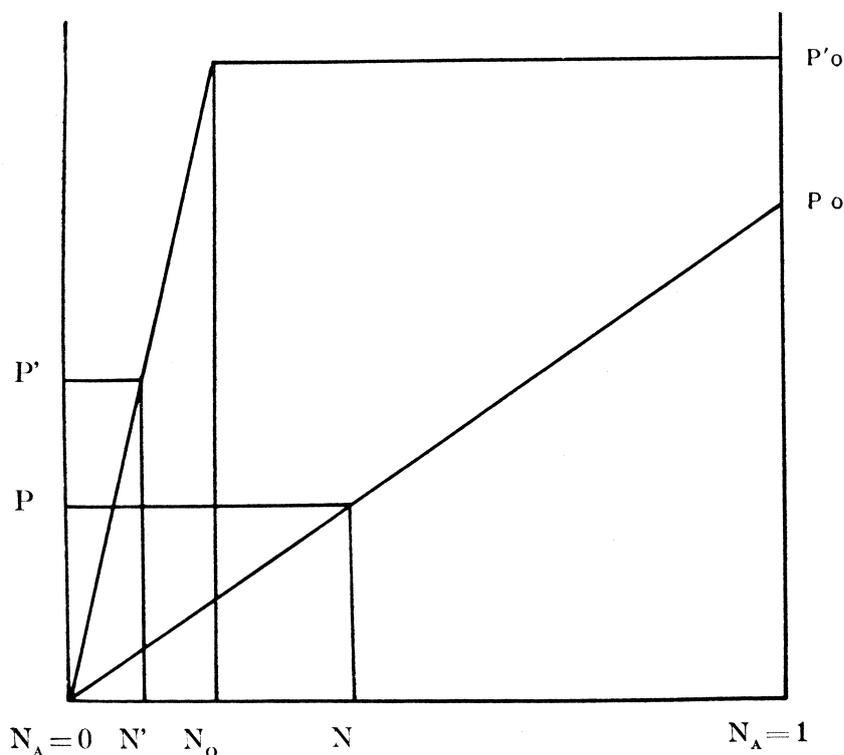
2° *Méthode activométrique basée sur l'obtention d'une valeur d'activité thermodynamique connue dans l'atmosphère en fonction d'une valeur d'activité thermodynamique connue en solution.*

a) *Principe de la méthode.*

Les résultats obtenus avec le précédent appareil n'étaient pas proportionnels aux difficultés considérables rencontrées

dans sa manipulation, nous nous sommes tournés vers une autre méthode beaucoup plus simple, dont la mise en œuvre est très facile et qui fournit des résultats rapides permettant d'orienter les recherches.

Le principe de la méthode est le suivant. Lorsqu'un liquide volatil pur est en équilibre avec sa vapeur dans une enceinte fermée l'activité dans les deux phases est égale à l'unité. C'est la définition de l'activité thermodynamique. Dans les deux phases le corps est à l'état pur, sous deux états différents, mais ses tendances à sortir de chaque phase pour pénétrer dans l'autre sont les mêmes. Si l'on désigne par  $p$  et  $p_0$  les tensions de vapeur partielle et saturante, comme  $p = p_0$  on aura nécessairement  $p/p_0 = 1$ . Si l'on considère le liquide pur, non mélangé à un solvant quelconque, sa concentration moléculaire désignée par  $N$  est égale à 1. L'activité est dans ce cas  $a = N = 1$ . Considérons maintenant le même liquide A miscible en toutes proportions avec un solvant B, le mélange cons-



tituant une solution idéale. La valeur de la tension de vapeur en fonction de la concentration moléculaire obéit à la loi des solutions idéales :

$$p = N \times p^0$$

La pression partielle est égale au produit de la concentration moléculaire par la pression de vapeur saturante. Lorsqu'il s'agit d'une substance non miscible en toutes proportions la loi des solutions idéales ne s'applique plus.

La droite représentant la variation de pression en fonction de la concentration possède une pente très différente parce qu'on atteint la pression  $p^0$  pour une concentration moléculaire  $N' = N_0$  plus petite que  $N = 1$ . La valeur de  $N_0$  correspond à la solubilité du corps dans le solvant. La valeur  $p'$  de la tension de vapeur partielle n'est plus donnée par la précédente égalité ; la pression de vapeur augmente relativement plus vite, si bien qu'on a :

$$p' = k \times N' p^0$$

Il est également possible de représenter tout ceci d'une autre façon en fonction de l'activité thermodynamique.

Quand les substances sont miscibles en toutes proportions on a :

$$a = \frac{p}{p^0} = N$$

L'activité est égale à la concentration moléculaire.

Quand les substances ne sont pas miscibles en toutes proportions on a :

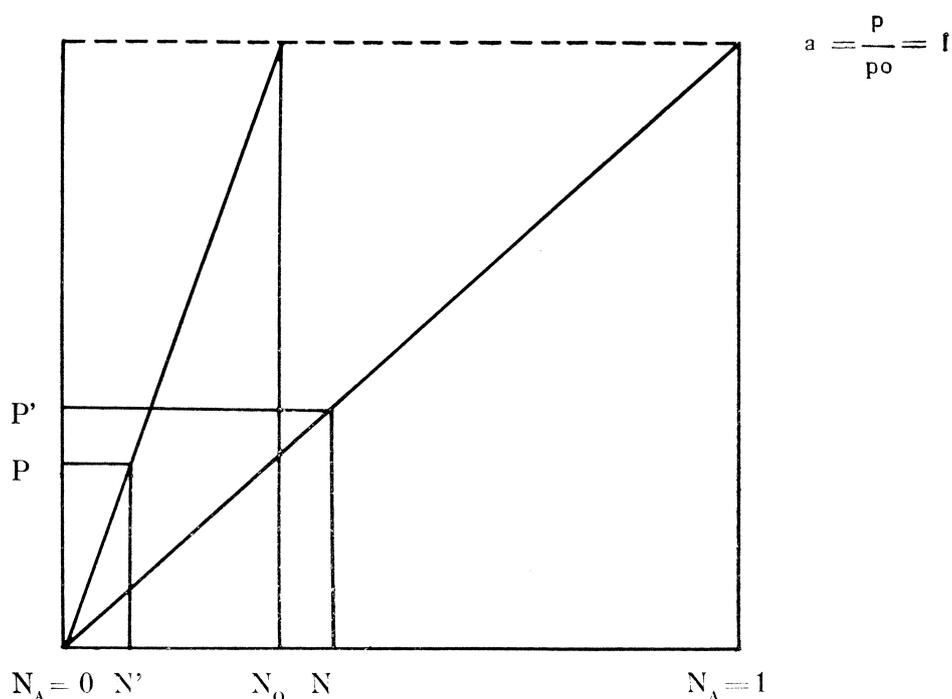
$$\frac{p'}{p^0} = N' \times k$$

Ici l'activité n'est plus égale à la concentration moléculaire mais à cette dernière multipliée par un coefficient que l'on représente par  $\gamma$  et qui est le coefficient d'activité. Nous admettrons, bien que sans nul doute ceci ne soit qu'une approximation, que la valeur de  $\gamma$  correspond à l'inverse de la solubilité  $N_0$  de la substance ou que

l'activité de la solution peut être représentée par  $a = \frac{N'}{N_0}$ . En

réalité il faudrait connaître la courbe réelle des tensions de vapeur en fonction des concentrations moléculaires. Mais pré-

cisément pour les substances que nous allons étudier, comme pour celles que nous avons étudiées dans le cas de la narcose ou de l'excitation-sapide, ces valeurs de pression partielle ne sont pas connues et souvent nous ne possédons pas les valeurs des tensions de vapeur saturante. Nous supposons donc qu'aux concentrations très faibles auxquelles nous avons affaire les solutions se comportent d'une façon suffisamment idéale. Nous n'avons d'ailleurs pratiquement pas l'embaras



du choix et se refuser à cette hypothèse c'est, pour le moment, renoncer à toute possibilité d'étude.

Dans quelques rares cas il sera possible de connaître commodément avec certitude la tension de vapeur de diverses solutions lorsque la valeur du coefficient d'activité sera connue et encore, à condition de se trouver dans le domaine des solutions diluées ; ce sera le cas des alcools, par exemple, dont les coefficients d'activité ont été mesurés expérimentalement. En principe, en vertu de ce qui précède nous admet-

trons qu'il nous suffira de disposer d'une solution d'activité thermodynamique connue dans une enceinte close pour connaître l'activité de la vapeur au-dessus de cette solution au moment où l'équilibre est atteint. Pratiquement on préparera des solutions qui seront des dilutions de titre connu par rapport à la solution saturante que l'on laissera se mettre en équilibre avec l'atmosphère dans un récipient clos. Le sujet respirera la vapeur ainsi obtenue et l'on mesurera ainsi des seuils olfactifs exprimés en activité thermodynamique à partir des solutions aqueuses des substances.

Mais auparavant il importe de se prémunir d'une cause d'erreur qui peut porter sur le détitrage de la solution dans certaines conditions. Tout d'abord, le volume de l'atmosphère surplombant la solution doit évidemment ne pas être trop considérable par rapport à la masse de substance contenue dans un volume donné de solution. On peut facilement apprécier l'erreur commise dans des conditions déterminées de rapports entre les volumes des phases au moment de l'établissement de l'équilibre. Si l'on place dans un récipient clos de deux litres, par exemple, un volume de un litre de solution à une activité de 0,01, l'activité de la vapeur et de la solution ne seront plus de 0,01 mais de  $0,01 - x$  après établissement de l'équilibre. Cette diminution d'activité est proportionnelle à la quantité de substance susceptible de passer dans la phase gazeuse quand l'équilibre s'établit. Elle est proportionnelle aux volumes en présence et au coefficient de solubilité de la vapeur. On pourra calculer ainsi l'erreur que l'on commet lorsqu'on pense avoir dans l'atmosphère une activité de 0,01, par exemple, en partant d'une solution à cette activité pour des volumes égaux de vapeur et de solution.

Soit une solution d'alcool d'un volume d'un litre surmontée par un volume d'atmosphère de un litre. Pour que la vapeur soit à l'activité 0,01 il faut que la pression partielle  $p$  soit :  $p_0 \times 0,01$ .

On calculera le poids de substance contenue dans un litre de vapeur à cette pression,  $N$  étant la concentration en molécule gramme.

$$p(20^\circ) = \frac{22.400 \times 293 \times 760 \times N}{273 \times 1.000} = \frac{4.988.032 \times N}{273}$$

en arrondissant le numérateur on aura :

$$N = \frac{p \times 273}{5.000.000}$$

et la quantité  $Q_a$  en poids par litre d'air, le poids moléculaire de la substance étant  $M$  sera :

$$Q_a = \frac{p \times 273}{5.000.000} \times M = 0,000055 \times p \times M$$

D'autre part l'activité 0,01 de l'alcool dans l'eau sera atteinte quand la solution contiendra un poids donné par le raisonnement suivant :

l'activité est :

$$a = N \gamma$$

où  $N$  est égal au quotient du nombre de molécules  $m$  d'alcool par le nombre total de molécules d'alcool et de molécules d'eau  $m + m'$  par litre de mélange. On peut négliger  $m$  au dénominateur et la valeur de  $m'$  est environ 55.

on aura donc :

$$a = 0,01 = \frac{m \times \gamma}{55}$$

$$\text{D'où } \frac{0,55}{\gamma} = m$$

le poids moléculaire étant  $M$  la quantité  $Q_e$  dans l'eau sera :

$$Q_e = \frac{0,55 \times M}{\gamma}$$

On connaît ainsi les quantités  $Q_a$  et  $Q_e$  de substance nécessaires pour produire une même activité 0,01 dans les phases vapeur et liquide. Le coefficient de détitrage lorsque l'on mettra en équilibre des volumes égaux d'air et de solution sera :

$$D = 1 / \frac{Q_e}{Q_a}$$

Nous avons calculé les indices de détitrage pour quelques alcools aliphatiques utilisés dans nos déterminations de seuils. Il est nécessaire de faire une remarque au sujet de l'alcool octylique. N'ayant pas trouvé dans la bibliographie la valeur de la tension de vapeur de ce corps à 20° nous avons dû extrapoler en admettant qu'en gros les tensions de vapeur des

alcools de la série aliphatique obéissent à la règle de TRAUBE. Dans ce cas le coefficient moyen  $p_o n / p_o n + 1$  est égal à 3. De façon à être certains de ne pas commettre d'erreur grave par excès dans cette approximation nous avons calculé la pression de vapeur de l'alcool octylique à partir de celle de l'alcool amylique en prenant un coefficient de 2 seulement, ce qui est certainement un minimum. Même dans ces conditions l'erreur de détitrage commise sur l'alcool octylique est négligeable.

Alcools	M	$p_o$	$\gamma$	$Q_e / l L$	$Q_a / l L$	$\frac{Q_e}{Q_a}$
Méthyl .....	32	95	1,5	11,73	0,0016	1/7330
Ethyl .....	46	43	3,6	7,03	0,0009	1/7800
Propyl .....	60	14	14,4	2,29	0,0004	1/4872
Butyl .....	74	4,4	53	0,76	0,0001	1/4293
Amyl .....	88	2,8	214	0,22	0,000135	1/1692
Octyl .....	130	0,3	12300	0,005	0,00002	1/250

*Indice de détitrage dans la série des alcools pour des phases liquide et vapeur de volumes égaux*

Comme on le voit, l'établissement d'une phase vapeur aux dépens d'une phase liquide dans un système où le rapport des phases est égal à l'unité abaisse le titre de la phase liquide, et par conséquent l'activité du système entier, d'une valeur oscillant entre quelques centièmes et quelques millièmes. Étant donné que pratiquement on se contente de connaître l'activité dans les meilleures conditions à un dixième près de sa valeur l'erreur commise par le détitrage sera complètement négligée.

Pratiquement nous avons appliqué les principes précédents au moyen des deux méthodes suivantes :

b) *Méthode statique de l'éprouvette.*

Cette méthode est d'une très grande simplicité car de simples éprouvettes d'un litre de capacité et des solutions saturées de substances odorantes suffisent. On réalise dans des séries d'éprouvettes d'un litre de capacité une série de solutions dont l'activité est connue par simple dilution de la solution saturante à essayer. On utilise des éprouvettes non graduées jusqu'au sommet laissant entre le sommet et la graduation un volume de 100 à 150 cm<sup>3</sup> occupé par l'atmosphère.

Dans ces conditions le détitrage est encore plus faible bien entendu que celui qui a été calculé précédemment pour des volumes égaux des deux phases. On ferme avec la paume de la main ou mieux avec un rodage (éprouvettes spéciales) l'éprouvette correspondant à l'activité thermodynamique qui doit faire l'objet de l'essai et on agite très vigoureusement de façon à émulsionner l'atmosphère de l'éprouvette avec le liquide et à réaliser l'équilibre entre les deux phases. On débouche ensuite brusquement l'éprouvette sous le nez du sujet qui perçoit une bouffée très fugace de l'odeur essayée. Le sujet a auparavant senti dans les mêmes conditions une éprouvette témoin ne renfermant que de l'eau. En général après un court apprentissage avec un parfum caractéristique (salicylate de méthyle, alcool octylique, etc.), on peut procéder à des déterminations de seuil. Le sujet caractérise souvent l'odeur de l'éprouvette contenant de l'eau avec les vocables d'odeur de mouillé, d'eau battue, de vase, et la considère comme « odeur nulle » vis-à-vis de la première valeur d'activité pour laquelle il perçoit une sensation nouvelle correspondant au parfum introduit dans la solution à essayer.

L'inconvénient de la méthode est de ne permettre qu'une sensation très fugace et de nécessiter une nouvelle manipulation si l'on veut faire sentir à nouveau l'odeur au sujet. Nous avons mis en œuvre une méthode semi-dynamique basée sur le même principe.

c) *Méthode semi-dynamique des deux ampoules à décantation.*

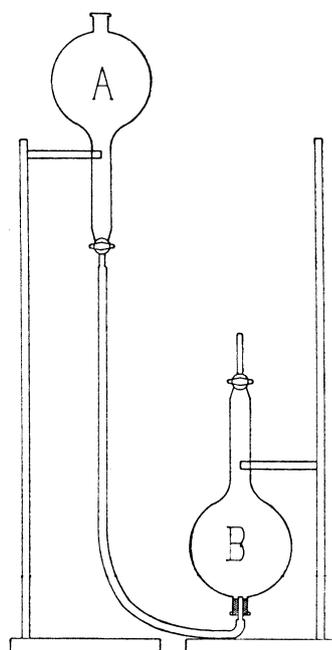
Alors que dans la précédente méthode la sensation ne peut être perçue que pendant un court instant (de l'ordre de la seconde), la seconde méthode permet d'insuffler vers les narines du sujet la vapeur odorante pendant une quinzaine de secondes de façon continue ou de pratiquer facilement plusieurs essais espacés de plus courte durée.

L'appareil comprend deux ampoules à décantation de 2 litres de capacité reliées entre elles par un tube de caoutchouc et remplies de solution odorante à la même activité thermodynamique. Les solutions sont obtenues, comme précédemment par dilution de solutions saturées ou de solutions titrées à une activité donnée à partir du coefficient d'activité, s'il

GAVAUDAN, ETC... : OLFACTION

159

s'agit des alcools. Les ampoules sont fixées chacune à un support stable. L'ampoule B possède une orientation inverse de l'ampoule A, de telle façon que le sujet reçoive la vapeur odorante par le robinet de B. Cette disposition a de plus l'avantage d'éviter un freinage trop considérable du liquide venant de A, au niveau d'un second robinet et de conserver un déplacement de liquide assez rapide permettant une vi-



tesse d'insufflation relativement grande vers les narines du sujet. On amène le niveau du liquide à une graduation correspondant à un volume de 1 litre dans l'ampoule B qui sert d'échangeur entre l'atmosphère et le liquide, puis cette ampoule détachée de son support est agitée violemment, de façon à émulsionner l'atmosphère et le liquide et à réaliser rapidement l'équilibre des phases, comme dans la méthode de l'éprouvette. Puis l'ampoule B étant remise en place sur son support et son robinet ouvert, le sujet se met en position, une des narines placées au-dessus de ce robinet. Le robinet A est alors ouvert, ce qui permet le déplacement de la vapeur par le liquide venant de A par gravité. Cette méthode donne de bons résultats.

3° *Activomètre dynamique basé sur le principe du couplage de débitmètres assurant des débits de vapeur (saturante ou à une activité déjà connue) et d'air dans des rapports définis et variables à volonté.* (Photographie, schéma et légende pp. 161, 162, 163).

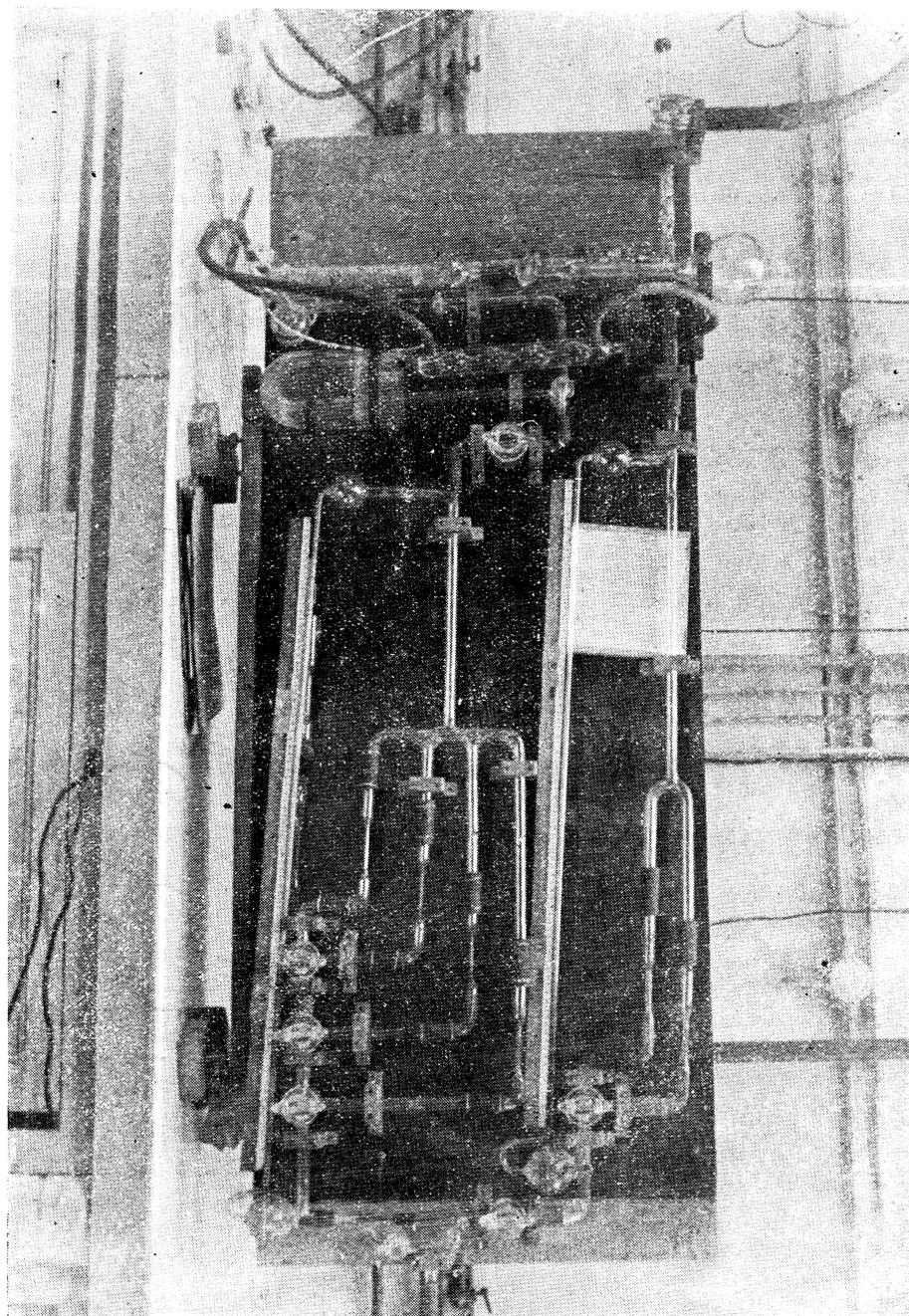
Nous avons réalisé un appareil permettant d'obtenir une activité connue, comme dans l'appareil statique dérivé de celui de HOFFMAN et KOHLRAUSCH, mais sous un débit constant. Le principe consiste à mélanger de façon continue des proportions connues de vapeur à la pression saturante ou à une pression partielle donnée et d'air pur.

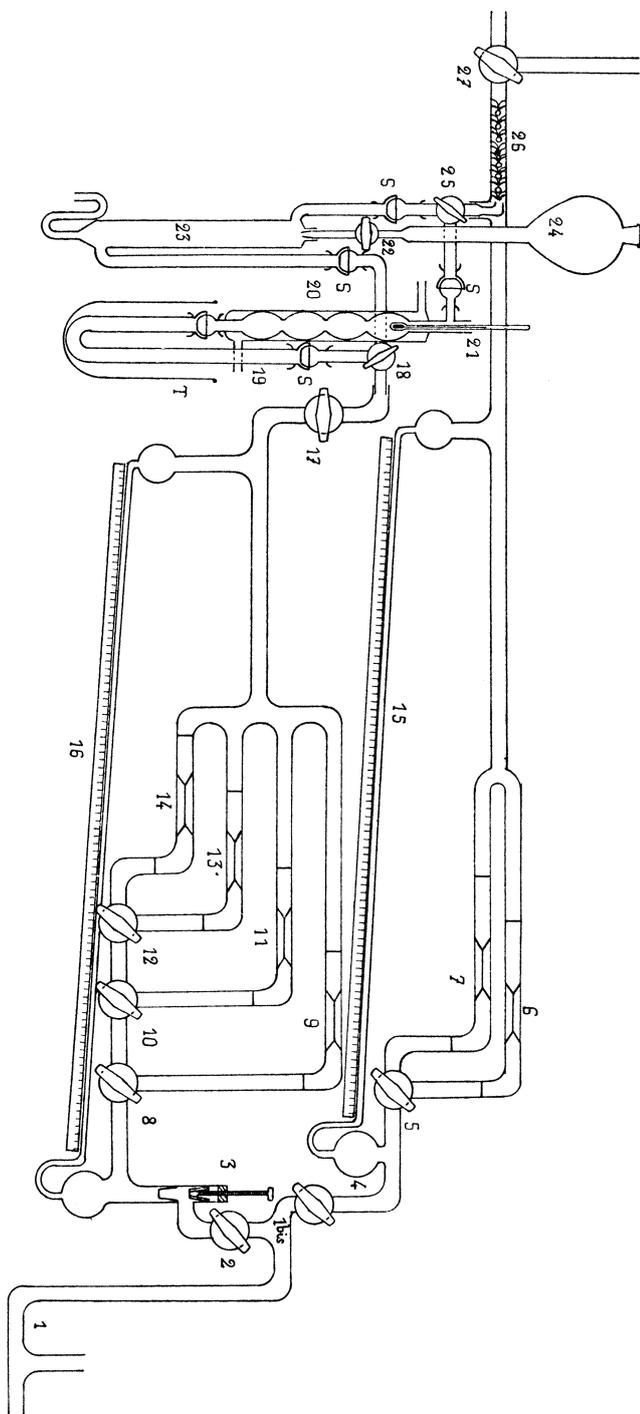
L'appareil est constitué par une association de débitmètres à niveau de liquide permettant de contrôler à volonté dans certaines limites les débits d'air et de vapeurs odorantes. La source d'air est constituée par une turbine à air entraînée par un moteur électrique. Le courant d'air est partagé en 1 bis en deux parties. La plus grande partie du volume d'air total se dirige vers un débitmètre à grand débit pourvu de deux capillaires et d'un robinet à deux voies. Les débits peuvent varier dans de larges limites en utilisant les deux capillaires et la graduation du débitmètre. Une petite partie du volume d'air total peut être dirigée, grâce à un jeu de robinets vers une série de capillaires calibrés (9, 11, 13 et 14) en relation avec un second débitmètre ; un robinet et un pointeau (3) permettent de régler avec précision le débit dans cette partie de l'appareil. Les deux débitmètres ont été étalonnés en charge sur les deux séries de capillaires à grand et petit débit. Cet étalonnage a été effectué au moyen d'un gazomètre et d'une bouteille à air pour les capillaires à grand débit et pour les capillaires à faible débit de  $8/10^{\circ}$  de mm. et  $5/10^{\circ}$  de mm. Les capillaires à très faible débit de  $1/10^{\circ}$  de mm. et  $5/100^{\circ}$  de mm. ont été étalonnés par déplacement d'un index mobile dans un tube gradué. Une abaque a été ensuite construite pour les diverses combinaisons de capillaires à grand et petit débit en portant en coordonnées les graduations des débitmètres et les volumes débités. En pratique, dans ces essais on n'a utilisé que le capillaire à grand débit de 0,5 à 6 l. minute fonctionnant constamment sous un débit de  $5.000 \text{ cm}^3$  par minute. On a établi au moyen de l'abaque un tableau permettant d'obtenir directement en fonction de la graduation du débitmètre à petits débits l'activité pour les différents capillaires. Les divers ca-

1948-7. Étude pharmacodynamique des chimiorécepteurs de...

Année 1948

*Réalisation du débitmètre-actinomètre.*





*Schéma du débitmètre-actinomètre.*  
(Légende page 163)

## GAVAUDAN, ETC... : OLFACION

163

1. Arrivée d'air.
  - 1 bis. Bifurcation vers 6 et 7 ou vers 9, 11, 13, 14.
  2. Dérivation d'air destinée à être soit saturée de vapeur, soit portée à une activité thermodynamique initiale donnée.
  3. Robinet pointeau de réglage de débit de la dérivation d'air pur.
  4. Robinet de réglage de débit d'air pur (grand débit).
  5. Robinet permettant le passage de l'air par les capillaires 6 et 7.
  6. Capillaire pour débits de 0 l. 500 à 6 litres minute.
  7. Capillaire pour débits de 2 litres à 15 litres minute.
  8. Robinet permettant le passage de l'air vers le tube capillaire 9 ou vers 10.
  9. Capillaire de  $8/10^{\circ}$  de  $\frac{m}{m}$  pour débits de 50 à 700 c.c. min.
  10. Robinet permettant le passage de l'air vers le tube capillaire 11 ou vers 12.
  11. Capillaire de  $5/10^{\circ}$  de  $\frac{m}{m}$  pour débits de 10 à 150 c. c. min.
  12. Robinet permettant le passage de l'air vers les capillaires 13 ou 14.
  13. Capillaire de  $1/10^{\circ}$  de  $\frac{m}{m}$  pour débits de 0,5 à 12 c. c. min.
  14. Capillaire de  $5/100^{\circ}$  de  $\frac{m}{m}$  pour débits de 0,1 à 1,5 c. c. min.
  15. Réglette graduée indiquant les débits des capillaires 6 ou 7.
  16. Réglette graduée indiquant les débits des capillaires 9 - 11 - 13 et 14.
  17. Robinet commandant le passage ou l'arrêt de l'air destiné à être saturé ou porté à une activité thermodynamique initiale connue après réglage au débit voulu par 3 et lecture sur 16.
  18. Robinet commandant le passage de l'air à travers 19 ou 23.
  19. Tube saturateur en U contenant le corps pur étudié.
  20. Réfrigérant.
  21. Thermomètre.
  22. Robinet permettant l'écoulement d'une solution d'activité thermodynamique connue dans l'échangeur 23.
  23. Echangeur rempli aux  $3/4$  de pierre ponce et contenant la solution d'activité thermodynamique connue.
  24. Réserve de solution d'activité thermodynamique connue.
  25. Robinet d'arrivée de la vapeur saturante ou de la vapeur à une activité thermodynamique connue.
  26. Mélangeur de vapeurs et d'air.
  27. Robinet permettant de diriger l'air chargé de vapeur odorante à une activité thermodynamique donnée soit vers l'extérieur de la pièce, soit vers le sujet en expérience.
- S. Rodages sphériques.  
T. Thermostat.

pillaires permettent ainsi de débiter par minute un nombre de centimètres cubes représentant une fraction connue du débit total de 5.000 cm<sup>3</sup> d'air. Cet air ainsi débité par les petits capillaires peut passer par le jeu d'un robinet à trois voies (18) dans deux parties différentes de l'appareil destinées à obtenir soit un débit de la vapeur saturante (19) soit un débit d'une dilution déjà connue de vapeur saturante (13) ; les volumes débités dans l'un et l'autre cas sont destinés à être mélangés à l'air du débitmètre à grand débit grâce à un robinet à deux voies et à un gicleur capillaire commun débouchant dans le courant gazeux principal. On s'est assuré, au moyen d'hydrogène sulfuré et de papier à l'acétate de plomb, que le courant d'air principal n'entraîne pas, par effet de succion, les vapeurs se trouvant dans le circuit en amont du capillaire. Primitivement le débitmètre-activomètre était destiné à obtenir une activité donnée à partir de la vapeur saturante. Cette dernière est obtenue dans un tube en U (19) contenant un peu de ponce imbibée de liquide odorant et dont les parois sont recouvertes d'un dépôt, dans le cas d'un corps odorant solide. L'air arrive par la branche de droite à raison de quelques centimètres cubes par minute. De façon à faciliter la saturation le tube en U est contenu dans un récipient thermostatique à 50° et est surmonté d'un réfrigérant ramenant la température des vapeurs à 20° (20). Un thermomètre permet de contrôler cette température. Cette méthode est convenable pour obtenir des débits correspondant à des activités variant entre quelques dix millièmes et quelques centièmes. Cet appareil peut être utilisé avec succès dans divers travaux de pharmacodynamie et de toxicologie chaque fois que l'on doit maintenir des êtres vivants dans une atmosphère renouvelée d'une substance active à une concentration ou une activité connues. Nous reviendrons d'ailleurs sur ce point. Cet appareil est plus pratique que les appareils basés sur le principe de la distribution continue d'une substance par atomisation mécanique ou par vaporisation thermique directe de quantités mesurées pondéralement ou volumétriquement.

Mais dès nos premières études sur l'olfaction nous nous sommes rendus compte que l'appareil ainsi conçu ne pourrait nous rendre les services que nous en attendions, beaucoup de seuils olfactifs se trouvant dans le domaine inférieur aux dix-millièmes. Or au-dessous des dix-millièmes l'appareil devient

inutilisable. Le débit de capillaires de diamètre inférieur à  $1/10^{\text{e}}$  de mm. ne pouvait demeurer constant ni être connu avec une précision suffisante ; de faibles variations thermiques suffisent à provoquer de grosses erreurs. Il était possible de reprendre des dilutions de vapeurs saturantes correspondant à des activités de  $1/1.000^{\text{e}}$  et de les diluer à nouveau dans la marge de  $1/10.000^{\text{e}}$  à  $1/1.000.000^{\text{e}}$  par le jeu d'un second appareil analogue à celui que nous venons de décrire. Mais nous avons toutefois renoncé à cette complication et nous nous sommes bornés à ajouter à notre appareil un dispositif permettant de diluer les vapeurs provenant d'une dilution aqueuse de la substance à une activité thermodynamique connue. Nous avons vu précédemment que l'activité d'une vapeur au-dessus d'un liquide représentant une certaine dilution de la solution saturée était proportionnelle en gros à la valeur de cette dilution. On fait donc passer dans un tube échangeur (23) contenant de la ponce imbibée de solution aqueuse odorante à une activité connue, l'air provenant du circuit du débitmètre à faibles débits. On obtient ainsi, à volonté, des dilutions du  $1/10^{\text{e}}$ ,  $1/100^{\text{e}}$  ou  $1/1.000^{\text{e}}$  de l'activité établie dans le tube à ponce. Ce tube est pourvu à sa partie supérieure d'un réservoir (24) de solution que l'on peut faire ruisseler très lentement, dans les expériences de longue durée. Ce tube est porteur à sa partie inférieure d'un siphon permettant l'évacuation du trop plein de liquide sans modifier le régime gazeux. On peut ainsi obtenir des valeurs d'activité de  $1/1.000.000^{\text{e}}$  par exemple et au-dessous. Nous avons vérifié que les résultats ainsi obtenus correspondaient avec ceux obtenus par la simple méthode de l'éprouvette.

Notons toutefois que l'activomètre dynamique permettra l'étude de la variation du seuil d'activité thermodynamique en fonction de la vitesse ainsi que celle d'autres problèmes inabordables par la méthode de l'éprouvette ou par celle des deux ampoules à décantation. Enfin l'activomètre dynamique peut être utilisé dans diverses recherches pharmacodynamiques où il est indispensable de soumettre des êtres vivants à l'action d'atmosphères renouvelées contenant des teneurs définies de substances volatiles. Nous reviendrons ultérieurement sur cette application particulière, notamment en ce qui concerne l'étude de certains insecticides.

166

GALLICA BIOLOGICA ACTA

VOL. I 1948

On étudiera dans un travail suivant les premiers résultats de l'étude thermodynamique du déclenchement de la sensation olfactive.

*Laboratoire de Pharmacodynamie du Ministère  
de la Santé Publique au Centre d'Etudes du Bouchet.*

Mai 1948.

#### BIBLIOGRAPHIE

- P. GAVAUDAN, M. DODE et H. POUSSEL : La toxicologie générale et la notion d'activité thermodynamique. *Memor. des Serv. Chim. de l'Etat*, 13, 1944, 384-423.
- M. DODE, P. GAVAUDAN et H. POUSSEL : Contribution à l'étude de la narcose par les narcotiques indifférents. *Memor. des Serv. Chim. de l'Etat*, 32, 1946, 388-409.
- P. GAVAUDAN, M. P. SCHUTZENBERGER et H. POUSSEL : L'excitation des chimiorécepteurs de la langue par les substances du groupe des narcotiques indifférents et la règle thermodynamique de la narcose. *C. R. de l'Ac. des Sc.*, 224, 1947, 1525-1527.
- P. GAVAUDAN, H. POUSSEL et M. P. SCHUTZENBERGER : Le mécanisme physico-chimique de l'excitation sapide et la notion d'excitant indifférent. *C. R. de l'Ac. des Sc.*, 226, 1948, 751-752.
- H. ZWAARDEMAKER : *L'odorat*, Paris, 1925, 302 p.
- J. PASSY : *C. R. de la Soc. de Biol.*, 44, 1892, 84-88 et 239-243.
- C. REUTER : Demonstration eines speciall zu klinischen zwecken bestimmten Riechmessers. *Verhandl. d. Gesell. deut. Naturf. u. Aerzte*, 70, pt. 2., 311-313, Leipzig.
- BACH : *Gesundheits, Ing.*, 60, 1937, 222-225.
- HOFFMAN et KOHLRAUSCH : *Bioch. Zeit.*, 156, 1925, 287.
- S. MORIMURA : Untersuchungen über den Geruchsinn. *The Tohoku Jour. of exper. Med.*, 22, 1933-1934, 417-448.

# GALLICA BIOLOGICA ACTA

publiés sous la direction de P. Gavaudan  
*Directeur du Service de Biologie  
des Services Chimiques de l'Etat*

VOL. 1 FASC. 2 JUIN 1948

## S O M M A I R E

### Articles originaux.

- S. Bazin** : Etude de la pénétration de diverses substances chimiques dans le système vacuolaire de « Chara » .... p. 107
- P. Gavaudan et H. Poussel** : Remarques sur les théories de la cancérisation chimique ..... p. 111
- H. Poussel** : Sur les mécanismes des actions toxiques de l'hexachlorocyclohexane  $\gamma$  ..... p. 114
- P. Gavaudan, H. Vogel et G. Debraux** : Etudes des propriétés rhizogènes de la  $\beta$  lactone de l'acide  $\beta$  isopropylmalonique . . . . . p. 120
- G. Brebion et P. Gavaudan** : Etude de la photosynthèse des plantes aériennes par la méthode de Warburg ..... p. 124
- G. Debraux** : Recherches sur les feuilles isolées en survie. .... p. 131
- G. Debraux et P. Gavaudan** : Recherches sur la prophylaxie du mildiou de la vigne par les composés organiques ..... p. 138
- P. Gavaudan, G. Brebion, H. Poussel et M. P. Schützenberger** : Etude pharmacodynamique des chimiorécepteurs de l'olfaction; (1) Principes et méthodes d'une étude thermodynamique de l'olfaction . . . . . p. 147
- P. Gavaudan, H. Poussel et M. P. Schützenberger** : Etude pharmacodynamique des chimiorécepteurs de l'olfaction; (2) L'étude thermodynamique des séries homologues et les théories de l'olfaction ..... p. 167
- N. Kobozieff et N. A. Pomriaskinsky-Kobozieff** : Possibilité de sélection sur l'anomalie de l'oreille externe chez la souris. p. 184
- M. Schützenberger** : Remarques sur les relations aléatoires d'ordre ..... p. 187

### Revue générale.

- P. Gavaudan** : Echanges de matériaux figurés entre noyau et cytoplasme..... p. 205

LIBRAIRIE LE FRANÇOIS  
91, Boulevard Saint-Germain - PARIS

Physiologie des sensations.

**Étude pharmacodynamique des  
chimiorécepteurs de l'olfaction.**

(II) *L'étude thermodynamique des séries homologues  
et les théories de l'olfaction*

par P. GAVAUDAN, H. POUSSEL et M. P. SCHÜTZENBERGER.

*On étudie dans ce travail les variations du seuil sensoriel dans les séries homologues en essayant de tirer les conséquences que comportent ces constatations vis-à-vis des théories de l'olfaction. Le problème de l'olfaction est considéré comme un problème pharmacodynamique ordinaire.*

Nous examinerons dans ce travail les premiers résultats que nous avons obtenus dans nos recherches sur l'olfaction en appliquant la notion d'activité thermodynamique à l'expression des seuils selon les principes et les méthodes décrites dans le précédent travail (\*).

Nous avons choisi les séries homologues comme premier objet d'étude parce que les variations des propriétés physiques et physiologiques sont également bien connues dans les composés de cette catégorie. On peut ainsi espérer mettre plus facilement en évidence, si elles existent, des relations entre les propriétés physiques des molécules et les conditions du déclenchement de la sensation odorante. Comme les odeurs des divers termes des séries sont assez différentes et que le problème du seuil doit être plus ou moins lié à celui de la qualité sensorielle, dans ce dernier domaine également la

(\*) Voir : ce journal p. 147 et bibliographie n° (34).

connaissance de la variation des propriétés physiques pourra être utile.

Les variations du pouvoir odorant dans les séries homologues ont déjà retenu l'attention de divers auteurs et certains d'entre eux ont tenté de trouver l'explication de ce fait.

PASSY (1) avait signalé une oscillation du pouvoir odorant dans la série des acides gras. BACKMAN (18) a insisté sur la variation du pouvoir odorant dans la série des alcools. Les alcools méthylique et éthylique sont bien solubles dans l'eau, mais peu dans les lipoïdes, inversement, l'alcool cétylique, soluble dans les lipoïdes, est pratiquement insoluble dans l'eau et est inodore. Les termes intermédiaires, tels que l'alcool butylique, ont des odeurs puissantes et sont solubles à la fois dans l'eau et les lipoïdes. D'après BACKMAN la solubilité dans l'eau et les lipoïdes sont toutes deux nécessaires au conditionnement du pouvoir odorant. Notons que le terme de solubilité dans les lipoïdes est assez vague, car la solubilité des alcools présente de grandes différences selon le type de lipoïdes. Si l'alcool méthylique est très peu soluble dans l'huile d'olive, il est au contraire miscible en toutes proportions à l'alcool oléique (solvant choisi par MEYER et HEMMI (4) dans leurs travaux sur la théorie lipoïdique de la narcose). Mais il est fort possible que la remarque de BACKMAN soit exacte et présente un intérêt. On peut en effet concevoir une spécificité des lipoïdes solvants dans lesquels les différents alcools n'auraient qu'une solubilité limitée. Des considérations analogues peuvent s'appliquer aux acides gras. Une variation s'observe également dans la série des carbures saturés où l'odeur apparaît vers le 4<sup>e</sup> ou 5<sup>e</sup> terme, devient très puissante au 8<sup>e</sup> terme (odeur de pétrole) et s'atténue pour les paraffines supérieures.

Mais nous n'anticiperons pas sur les explications possibles des variations du pouvoir odorant dans les séries homologues.

Nous avons étudié 3 alcools normaux et nous avons trouvé les seuils suivants (exprimés en activités thermodynamiques) :

C <sub>1</sub>	$1.000 \times 10^{-6}$
C <sub>4</sub>	$6 \times 10^{-6}$
C <sub>6</sub>	$30 \times 10^{-6}$

Comme on le voit, l'activité thermodynamique seuil diminue considérablement de l'alcool méthylique à l'alcool bu-

GAVAUDAN-POUSSEL-M.P. SCHUTZENBERGER : OLFACTION 169

tylique puis augmente légèrement (environ 5 à 10 fois plus forte) pour l'alcool octylique.

Nous rapprocherons tout d'abord nos résultats de ceux obtenus par PASSY (1) dans la série des alcools. Mais malheureusement notre comparaison ne peut porter que sur la portion de la série homologue qui va de  $C_1$  à  $C_5$ , car PASSY n'a pas étudié l'alcool octylique normal primaire. Néanmoins cette simple confrontation est déjà intéressante car elle révèle une concordance des résultats. Nous avons transcrit ici les résultats de PASSY en calculant les activités thermodynamiques d'après les concentrations en poids par litre d'air indiquées par l'auteur et la tension de vapeur des différents alcools.

	C en gr/L	M C	M. gr./L	p	$p_0$	$a = \frac{p}{p_0}$
méthyl . . . . .	0,001	32	0,00003	0,54	95	0,005
éthyl . . . . .	0,00025	46	0,0000054	0,09	44	0,002
propyl . . . . .	0,000010	60	0,00000016	0,002	14	0,00014
butyl . . . . .	0,000001	74	0,000000013	0,0002	4,4	0,000045
amyl . . . . .	0,000001	88	0,000000011	0,0002	2,8	0,00007

Bien que la valeur de l'activité seuil de l'alcool butylique soit 10 fois plus élevée que celle que nous avons admise, on constate que les valeurs seuils d'activité calculables d'après nos résultats de PASSY varient dans le même sens que dans nos propres essais. Dans les deux séries d'expérience l'activité thermodynamique diminue considérablement du terme  $C_1$  au terme  $C_5$ . D'après les résultats de PASSY cette diminution doit cesser en passant du terme  $C_5$  au terme  $C_6$ . Ce point sera d'ailleurs vérifié plus tard par nous.

PASSY (2), nous l'avons déjà dit, avait noté la forme oscillante de la courbe du pouvoir odorant dans la série des acides gras ; mais les résultats de cet auteur ont été exprimés en concentrations en poids par litre d'air. Il nous a paru intéressant de calculer les activités thermodynamiques correspondantes. Nous n'avons pas trouvé dans les tables de constantes la tension de vapeur de l'acide décylrique à 20°. Mais la tension de vapeur des acides gras varie pour les premiers termes d'un

$$\text{facteur environ égal à : } \frac{P_{on}}{P_{on+1}} = 3$$

selon une règle analogue à la règle de TRAUBE.

Supposant qu'il en était de même pour les acides à nombre d'atomes de carbone plus élevé, nous avons extrapolé la valeur de la tension de vapeur de l'acide décylique, en admettant pour nous mettre à l'abri d'une erreur par excès dans l'évaluation de la constante, qu'à partir du terme en  $C_7$  le quotient de la tension de vapeur d'un terme par celle du suivant est seulement égal à 2, ce qui est un minimum.

	M	$P_0$	Concentration pour 1.000	p	P/ $P_0$
formique .....	46	33	0,000025	0,0099	0,0003
acétique .....	60	11,7	0,000005	0,00152	0,00012
butyrique .....	88	0,75	0,00000001	0,000000208	$0,27 \times 10^{-7}$
décylique .....	172	0,01	0,00000005	0,000005	0,0005

Comme dans le cas des alcools l'activité thermodynamique diminue du terme en  $C_7$  au terme en  $C_6$  puis remonte puisqu'elle se trouve, vers le terme en  $C_6$ , du même ordre de grandeur que pour le terme en  $C_7$ .

On doit faire deux remarques relatives aux valeurs d'activité thermodynamique seuils déclenchant la sensation olfactive :

1° Ces valeurs peuvent être très basses : dans le cas de l'alcool butylique l'activité narcotique est d'environ 50.000 fois plus forte que l'activité seuil déclenchant la sensation olfactive (0,038 contre 0,000006) ;

2° Alors que les valeurs sont constantes dans la narcose et dans l'excitation sapide (activités calculées pour les alcools d'après les valeurs de RENQVIST (3) (pour l'excitation sapide) elles varient considérablement dans l'excitation olfactive.

On peut donc penser que les mécanismes mis en jeu dans la narcose d'une part, et dans l'excitation olfactive, d'autre part, sont assez différents. Il est évident dès l'abord qu'il n'existe pas dans le cas de l'olfaction de règles simples analogues à la règle de TRAUBE, à celle de MEYER et HEMMI (4) ou à celle de la constance de l'activité (FERGUSON (5), GAVAUDAN, DODE et POUSSSEL (6), et permettant de prévoir le seuil d'excitation et encore moins la qualité sensorielle. Mais des valeurs d'activités thermodynamiques seuils de l'ordre de  $10^{-2}$  pour l'excitation sapide et de  $6 \times 10^{-5}$ , pour l'excitation olfactive, pour si différentes qu'elles soient, doivent-elles nous entraîner

GAVAUDAN-POUSSEL-M.P. SCHUTZENBERGER : OLFACTION 171

à conclure qu'il faut faire une distinction radicale entre les forces physiques qui sont à l'origine des deux types d'excitation ?

En particulier, ces résultats doivent-ils entraîner nécessairement l'adoption d'une des théories vibratoires ? Ces théories, d'ailleurs en général très imprécises, visent à établir une relation entre les vibrations intramoléculaires des substances odorantes et la qualité sensorielle ; elles font également état de la présence de pigment dans la tache olfactive. Dès 1874 TYNDALL (7) observait que les matières odorantes absorbent l'infra-rouge. Depuis cette époque de nombreux auteurs ont tenté de relier le pouvoir odorant des substances à l'adsorption de diverses radiations. HEYNINX (8) a tenté d'établir une relation entre l'odeur et l'absorption dans l'ultra-violet, et à côté de considérations assez bizarres (il rapporte, par exemple, les propriétés antiseptiques du camphre et des terpènes aux « vibrations propres moléculaires ultra-violettes » de ces composés), il exprime toutefois nettement l'idée que l'olfaction est en relation avec la structure électronique (1919).

La théorie vibratoire de l'olfaction a reçu son expression la plus moderne et la plus scientifique dans le travail de DYSON (9) qui a tenté d'établir une correspondance entre les types d'odeur, par exemple imines, aldéhydes-cétones, acétyléniques, mercaptans, etc., et le spectre RAMAN des molécules correspondantes. La fréquence des vibrations moléculaires provoquerait la sensation olfactive comme certaines longueurs d'onde caractérisent certaines sensations colorées.

En résumé, il existerait des récepteurs cellulaires sensibles à des vibrations intra-moléculaires d'une certaine fréquence. En supposant que cette relation entre spectre RAMAN et odeur soit strictement valable et qu'il n'y ait pas des exceptions et des chevauchements (probables) entre spectres et types d'odeur, il reste à l'interpréter. Il n'est pas certain que des phénomènes pharmacodynamiques ou physiologiques puissent être provoqués directement par des vibrations internes de la molécule. Jusqu'à présent les actions pharmacodynamiques ou physiologiques connues sont dues soit à la réactivité chimique de certaines fonctions de la molécule active, soit à la molécule dans son ensemble, dont les propriétés physiques viennent modifier le milieu thermodynamique du solvant.

Des phénomènes de mésomérie sont également possibles

dans l'adsorption par les substrats biologiques. Il faudrait rechercher s'il existe des relations entre mésométrie, pouvoir odorant et spectre RAMAN.

La théorie vibratoire de l'olfaction ne paraît pas encore très claire et il est nécessaire d'arriver à définir si les molécules agissent sur les récepteurs par leurs vibrations elles-mêmes ou si le déclenchement de la sensation est dû à d'autres propriétés physico-chimiques corrélatives de ces vibrations intramoléculaires. La première tâche des partisans d'une théorie vibratoire pure devra être de montrer comment les récepteurs cellulaires sont influencés par les vibrations intramoléculaires. Dans le cas d'une démonstration positive la pharmacodynamie se trouverait à un tournant important.

Mais les théories vibratoires sont encore très incertaines et ne nous fournissent pas pour le moment d'hypothèses de travail très nettes. Au contraire, il semble qu'il soit encore possible de rechercher une explication du mécanisme de l'olfaction dans le cadre des phénomènes physico-chimiques qui sont habituellement invoqués en pharmacodynamie et en physiologie.

Il est évident que dans la série homologue des alcools toute molécule, du terme en  $C_1$  au terme en  $C_5$ , par exemple, (pourvu que l'on soit avant le terme où se manifeste le « cut off » d'activité biologique) exerce la même action physiologique, narcose ou déclenchement d'une excitation sapide, quelle que soit sa longueur pourvu que l'on se trouve dans des conditions de concentration correspondant à la règle d'activité thermodynamique. Il suffit qu'un certain nombre de molécules, considéré comme pratiquement constant soit présent dans le récepteur cellulaire pour que le phénomène pharmacodynamique prenne naissance. Pratiquement il n'existe aucune spécificité d'action. On peut considérer que la sensation d'amaïertume déclenchée au niveau des papilles circumvallées ou la narcose du têtard sont réalisées par n'importe quelle molécule de la série homologue. Il s'agit d'un mode d'excitation ou de narcose indifférentes (14). L'excitation dépend seulement de la présence d'un nombre déterminé de molécules quelconques dans un certain solvant cellulaire. Ce sont des propriétés des molécules en somme banales et communes, relatives à la dissolution ou à l'adsorption dans un solvant ou sur une surface cellulaires, qui sont ici en cause.

## GAVAUDAN-POUSSEL-M.P. SCHUTZENBERGER : OLFACION 173

Il n'en n'est pas de même dans l'olfaction où nous voyons non seulement la valeur des seuils d'excitation varier considérablement mais également la nature de l'excitation se modifier d'un terme à l'autre. Un phénomène de spécificité physiologique n'existant pas précédemment est introduit ici : l'odeur des alcools inférieurs est peu marquée et très différente de l'odeur de pomme bien caractéristique de l'alcool butylique, laquelle est également très différente de l'odeur florale de l'alcool octylique. On doit supposer, ou bien que les mêmes récepteurs ne sont pas influencés de la même façon ou bien que des récepteurs différents ou des combinaisons différentes de récepteurs sont excités. La longueur de la chaîne est un facteur qui influe à la fois sur la valeur du seuil thermodynamique et sur la qualité de la sensation. ZWAARDEMAKER avait remarqué il y a déjà longtemps que dans les séries homologues « c'est l'allongement de la chaîne qui détermine les transitions » (15 p. 242). D'après le même savant on constate une transition au 5<sup>e</sup> terme, quand l'odeur alcoolique prend un caractère fétide et au 7<sup>e</sup> terme, quand cette odeur devient aromatique. Il est évident que l'indication des « transitions » présente un caractère un peu arbitraire et personnel. On pourrait également faire remarquer que l'odeur des alcools ne devient vraiment caractéristique qu'au 3<sup>e</sup> terme. Mais il est incontestable que l'odeur prend une nuance fétide au 5<sup>e</sup> terme, aromatique au 7<sup>e</sup> puis un caractère « aigret » (odeur de savon) au 12<sup>e</sup> terme.

ZWAARDEMAKER donnait le tableau suivant de la répartition des poids des groupements atomiques dans les molécules des alcools de la série homologue :

1		1						31
2		15						31
3		15		14				31
4		15	14			14		31
5		15	14		14		14	31
6		15	14		14	14		14 31
7		15	14	14		14	14	14 31
8		15	14	14		14	14	14 31

Ce simple tableau ne peut évidemment, comme le disait ZWAARDEMAKER lui-même, nous faire pénétrer jusqu'au fond

de la compréhension du mécanisme de la stimulation, mais il met en évidence le déplacement du centre de gravité de la molécule et les modifications de sa symétrie. On voit que dans la série, pour la première fois le centre de gravité (par rapport aux poids des groupements atomiques) se trouve entre deux atomes de carbone avec l'alcool propylique. C'est précisément à ce terme que se manifeste, à notre avis, la première transition dans la série des alcools. L'alcool méthylique et éthylique purs n'ont pratiquement pas d'odeur caractéristique, l'alcool propylique, au contraire, possède une odeur caractéristique. Le centre de gravité se déplace par la suite dans la série en occupant alternativement pour les termes pairs et impairs une position située au niveau d'un atome de carbone ou entre deux atomes de carbone.

Ces considérations très simples montrent que les molécules des termes successifs d'une série homologue ne se ressemblent pas entièrement ; il n'y a entre elles qu'une homologie générale : elles appartiennent à la même famille, mais il ne faut pas oublier que chaque type possède ses particularités. Les particularités moléculaires sont particulièrement bien dévoilées par les valeurs des points de fusion des cristaux. La fusion est due à l'augmentation d'agitation thermique des molécules qui conduit à la rupture du réseau cristallin. Le point de fusion dépend des forces intermoléculaires, de la densité de l'entassement des groupes atomiques ou des atomes, enfin du réseau cristallin. Tous ces facteurs sont interdépendants, c'est ce qui explique que les particularités structurales des molécules influent beaucoup sur le point de fusion des cristaux. Dans beaucoup de séries homologues, les acides gras, par exemple, on observe des points singuliers dans la courbe exprimant la valeur du point de fusion en fonction du nombre d'atomes de carbone.

Les acides gras montrent un point de fusion minimum pour le terme en  $C_5$ . De plus, la variation du point de fusion à partir du terme en  $C_5$  présente une allure oscillante et la courbe représentative est une ligne brisée, le terme pair ayant un point de fusion plus élevé que le terme impair suivant. En outre, si l'on calcule la valeur de la variation du point de fusion des termes successifs on voit qu'elle n'est pas constante. L'addition d'un atome de carbone n'a pas toujours le même effet : lorsqu'on passe d'un terme pair à un terme

GAVAUDAN-POUSSEL-M.P. SCHUTZENBERGER : OLFACTION 175

impair l'incrément du point de fusion diminue régulièrement, surtout à partir du passage de  $C_7$  en  $C_8$ . Il n'en n'est pas de même lorsqu'on passe d'un terme impair à un terme pair. L'incrément varie d'une façon irrégulière et oscillante. L'addition d'un atome de carbone n'a pas du tout le même effet sur le point de fusion s'il s'agit du passage de l'acide formique à l'acide acétique ou du passage de l'acide en  $C_7$  à l'acide en  $C_8$ . L'oscillation de la valeur du point de fusion en fonction de la parité ou de l'imparité a été attribuée au fait que les chaînes hydrocarbonées ont dans les cristaux des formes en zigzag, ce qui a été confirmé par les études aux rayons X ; les termes les plus symétriques ont un incrément positif, les moins symétriques un incrément négatif. Les points singuliers dans les courbes des points de fusion, par exemple le minimum des acides gras en  $C_8$  ont été expliqués par un arrangement de la chaîne en spirale ou en hélice.

Comme on le voit, la forme et la symétrie des molécules homologues peuvent être différentes dans les cristaux. Il est possible que certaines de ces particularités soient conservées dans les solutions où elles sont plus difficiles à déceler, par suite de l'agitation des molécules du solvant et du corps dissous. Bien que concernant l'état solide de la matière, l'oscillation des points de fusion n'est peut-être pas absolument sans lien avec l'oscillation d'activité biologique. On sait, en effet, que la chaleur et la température de fusion interviennent dans une relation théorique (CLAPEYRON) permettant de calculer la solubilité idéale des substances. On pourrait donc s'attendre *a priori* à l'existence d'une certaine relation entre les deux types d'oscillations biologique et physique. Il est évident que les deux types d'oscillations ne peuvent être parallèles, l'oscillation biologique dépendant des conditions spatiales nouvelles intervenant en milieu liquide aux interfaces, de la nature du solvant, de l'affinité du solvant et du corps dissous, etc...

Nous admettrons donc que la forme et la symétrie des molécules varient dans les séries homologues et que ces variations sont décelées par celles de certaines de leurs constantes physiques ou des incréments de ces dernières.

L'idéal serait de pouvoir établir une relation entre les variations de certaines propriétés physiques et la sensation odorante, mais le nombre des facteurs physiques qui interviennent et l'incertitude même de certaines de ces données

physiques concernant la forme des molécules dans les solutions et les couches paucimoléculaires ne permettent pas encore de réaliser une telle analyse. Aux transitions de ZWAARDEMAKER dans les séries homologues il faudra dès que possible substituer la notion plus précise de modifications de la forme et de la symétrie moléculaires et partir de l'idée que ces deux facteurs peuvent conditionner la valeur de l'activité thermodynamique seuil et la qualité sensorielle.

La forme, la symétrie des molécules, la présence de certains groupes (COOH ou OH par exemple), détermineraient les quantités de molécules dissoutes ou adsorbées par les solvants spécifiques des cellules réceptrices ; ces molécules pourraient dans certains cas former des complexes (pour une certaine longueur de chaîne), au sens de SCHULMAN (17) et RIDEAL.

D'autre part, la forme et la symétrie des molécules interviennent également pour déterminer les mouvements de rotation et de translation de ces dernières ; or il est possible que la vitesse de pénétration des molécules joue un rôle dans le mécanisme de l'excitation. On peut concevoir que la rotation et le passage de certaines molécules dans les couches limitantes puissent provoquer des perturbations ioniques ou électriques spécifiques. RENQVIST (3) dans sa théorie du goût avait admis que la vitesse de passage des molécules constituait l'excitant sapide. Une pareille hypothèse n'est évidemment pas interdite dans le cas des chimiorécepteurs de l'olfaction, puisque l'on ignore à l'heure actuelle si le déclenchement de la sensation est dû au fait de l'arrivée même des molécules sur des récepteurs spécifiques ou à la libération d'un médiateur chimique.

En résumé, nous constatons d'une part l'existence de variations considérables de l'activité thermodynamique seuil et de la qualité sensorielle dans les séries homologues et d'autre part nous rapportons ces variations à des oscillations de propriétés physiques. Il faut noter que l'existence d'activités pharmacodynamiques d'intensité oscillante dans les séries homologues est un fait assez fréquent.

J. H. SCHULMAN et RIDEAL ont rapporté le maximum d'activité des séries homologues au passage de la formation de complexes à des phénomènes de dispersion dans les couches minces. C'est ainsi qu'ils expliquent l'optimum d'activité oes-

## GAVAUDAN-POUSSEL-M.P. SCHUTZENBERGER : OLFACTION 177

trogène des p-p' dihydroxystilbène, lorsque les deux groupes alkyls substituants sont des radicaux éthyl. L'activité biologique optima s'observe pour le composé donnant lieu à l'adsorption la plus considérable sur une interface protéique, l'activité diminuant immédiatement pour les composés qui exercent une action dispersante sur les molécules de l'interface protéique. L'activité hémolytique optima dans une série homologe coïncide avec le maximum d'action dispersante sur les molécules de cholesterol ou de protéine d'une interface.

L'optimum de pouvoir antiseptique de divers alkyls phénols, crésols et resorcinols (voir ING) (19), est observé pour un certain terme des séries. Des phénomènes d'oscillation de l'activité biologique ont été aussi signalés dans le cas des heteroauxines : d'après GRACE (20) dans les séries des acides  $\alpha$  naphthalène acétique et  $\beta$  indole-acétique les composés à nombre pair d'atomes de carbone dans la chaîne latérale ont une plus grande activité que ceux à nombre pair (test racine). THIMANN et BONNER (21) ont signalé une oscillation analogue pour les acides indolacétique, propionique, butyrique avec un test différent (*Avena*).

VELDSTRA (10) a observé que l'action turgescence des acides gras sur le coacervat d'oléate et la courbure dans le test *Pisum* présentent un maximum  $C_{11}$ . La courbe est asymptotique aux abscisses vers le terme en  $C_6$  et après un minimum très bas en  $C_{16}$  remonte légèrement en  $C_{18}$ .

D'après DILLS et MENUSAN (27) le pouvoir insecticide des acides gras pour *Aphis rumicis* présente un maximum d'activité pour le terme en  $C_{10}$ .

HURST (23) a observé sur les larves de *Phormia terraevovae* que la toxicité des alcools purs augmentait de  $C_1$  en  $C_5$  puis diminuait ; en  $C_6$  la toxicité est redevenue à peu près égale à ce qu'elle était en  $C_1$ . Il faut noter que les oscillations de toxicité (à l'exception de HURST) que nous avons relatées ont été exprimées par tous les auteurs en fonction de la *concentration* de substance active ; il faudrait donc reprendre ces résultats et examiner ce que donnerait leur traduction en activité thermodynamique. Nous avons pu facilement réaliser cette transformation dans le cas des résultats obtenus par HURST, car cet auteur a fait agir des corps purs en y immergeant les larves. L'auteur a donc toujours réalisé ses essais à une activité thermodynamique égale à l'unité (le corps pur

ayant par définition une activité égale à un). La position du maximum d'activité toxique en  $C_5$  paraît donc bien avoir dans ce cas une signification absolue. L'allure de la courbe de HURST<sup>7</sup> concernant les essais d'intoxication par injection dans la cavité générale est très différente. C'est une courbe exponentielle qui obéit simplement à la règle de TRAUBE. Ainsi que l'a fait remarquer HURST la toxicité est déterminée dans les expériences par immersion par un facteur limitant qui est le transport à travers la couche de chitine réglé par une spécificité des couches de passage et par la forme des molécules. C'est ainsi qu'à une même activité la toxicité résultante peut être très différente avec divers alcools. Nous voyons apparaître ici, dans un phénomène de toxicité assez banal, qui est en définitive une simple narcose, un facteur intermédiaire de perméabilité qui peut créer dans l'intoxication des conditions de spécificité.

Nous avons cité tous ces exemples pour indiquer que la variation du seuil d'activité thermodynamique provoquant une sensation olfactive *quelconque* dans une série homologue n'est pas un phénomène isolé. Il est possible qu'il existe une relation entre cette variation et les variations de diverses actions pharmacodynamiques, c'est-à-dire qu'elle soit explicable par des considérations analogues à celles développées par SCHULMAN et RIDEAL (op. cit.). Mais il est possible aussi que la variation soit due à d'autres causes.

De façon générale, il n'est pas impossible de concevoir l'existence d'une théorie de l'olfaction (substances liposolubles du type des narcotiques indifférents) basée sur l'influence perturbatrice exercée sur des couches limitantes par la présence de molécules dont les formes, la symétrie, l'arrangement des groupes polaires et apolaires sont adaptés à la nature des récepteurs. Les propriétés physiques mises en jeu dans un tel mécanisme d'action seraient relatives aux molécules considérées comme un tout et non à certaines de leurs parties ou à des vibrations de ces parties. Il est bien entendu impossible d'écarter complètement « a priori » la possibilité de théorie basées sur une adsorption spécifique avec phénomène de mésomérie. La limite entre l'excitation déclenchée par la formation de complexes et celle provoquée par des phénomènes de mésomérie sera d'ailleurs peut-être difficile à trancher dans certains cas. On ne peut non plus refuser complètement d'ac-

GAVAUDAN-POUSSEL-M.P. SCHUTZENBERGER : OLFACTION 179

cordier une part d'attention à des théories du type de celle de DYSON dont nous avons déjà parlé plus haut. Il est évident que les vibrations intramoléculaires sont étroitement liées à la structure, à la forme et à la symétrie des molécules. Enfin on n'oubliera pas qu'il peut exister *une pluralité de mécanismes déclenchant l'excitation olfactive*. Il est possible que des substances comme H<sub>2</sub>S, les amines, l'oxyde de cacodyle, les halogènes, etc., agissent par des mécanismes chimiques.

Nous avons voulu simplement montrer qu'il est encore possible de rechercher une explication au mécanisme de l'excitation des chimiorécepteurs de l'olfaction en ayant recours aux méthodes classiques de la pharmacodynamie.

Nous avons étudié ici le problème de l'olfaction comme un problème de pharmacodynamie cellulaire en essayant de calquer cette étude sur celle que nous avons précédemment effectuée sur le déclenchement de la sensation sapide. Cette méthode est une des voies par lesquelles le problème de l'olfaction peut être abordé. Il est évident qu'il en existe d'autres. Ainsi l'on pourra étudier avec beaucoup d'intérêt et de fruit, comme l'on fait divers auteurs, dont LE MAGNEN (24), soit l'influence de la vitesse du courant odorant [voir ELSBERG] (25), soit la fatigue des récepteurs par combinaisons de substances odorantes [voir : ZWAARDEMAKER (15)]. Nous n'insisterons pas ici sur l'importance des conclusions que l'on pourra tirer de ces recherches et qui sont relatives aux variations du seuil sensoriel avec la vitesse du courant gazeux et à la discrimination des divers types de récepteurs.

On accordera aussi une très grande importance aux recherches sur les anosmies spécifiques. MONCRIEFF (26) (p. 43), a souligné que les anosmies partielles étaient probablement très communes ; une des plus connues est celle de l'acide cyanhydrique. Ainsi les recherches de GULLOT (27) offrent un grand intérêt. D'après cet auteur, à chaque type d'anosmie correspondrait un type de perception indépendant, c'est-à-dire une odeur fondamentale. Par exemple, la seule odeur de muse intéresserait au moins 3 fondamentales, ce qui conduit l'auteur à supposer l'existence chez l'homme de plusieurs centaines de fondamentales.

Il est certain que la confrontation de tous les résultats sur l'influence de la vitesse, sur la fatigue par les diverses odeurs, sur la détermination des fondamentales par la méthode des

anosmies, avec les données de l'étude thermodynamique des seuils ne manquera pas de faire progresser la question.

Il faut noter que l'introduction de la notion d'activité thermodynamique dans l'étude de tous les problèmes relatifs au mécanisme de l'olfaction paraît souhaitable toutes les fois que l'on se préoccupe de donner un sens physique précis à la détermination des seuils de sensation. Les recherches de GUILLOT (27) posent en particulier un problème intéressant : les muscs des trois catégories, nitrés, macrocycliques et stéroïdes sont de constitutions chimiques très différentes. Il serait important de connaître à quelle activité thermodynamique ils agissent.

Il est certes possible du point de vue pratique de continuer à avoir recours aux anciennes méthodes olfactométriques, parfois très simples, qui sont utilisées par les spécialistes de la parfumerie dont l'expérience et les observations sont d'un si grand secours, mais il est indispensable de donner une signification physique précise aux conditions expérimentales du déclenchement de la sensation olfactive. La notion d'activité thermodynamique le permet. Le problème de la sensation olfactive est un problème pharmacodynamique et doit être résolu selon des règles générales auxquelles n'ont échappé ni la narcoïse ni la sensation sapide.

#### Considérations sur l'importance des études sur le mécanisme de l'olfaction en biologie cellulaire

Nous terminerons la seconde partie de ce travail par quelques considérations générales qui, bien qu'en marge d'une étude sur l'olfaction ne manquent pas d'intérêt.

Nous nous sommes efforcés de considérer le problème de l'olfaction comme un problème ordinaire de pharmacodynamie cellulaire. Ce faisant, nous avons été conduits à trouver certaines analogies entre des problèmes communs au déclenchement des sensations odorantes et à la production de diverses actions toxiques ou pharmacodynamiques : nous ne reviendrons pas sur la question de l'existence fréquente d'un optimum d'activité biologique pour une certaine longueur de la chaîne carbonée.

Le fait le plus intéressant dans le déclenchement de la sensation olfactive est la spécificité. Cette spécificité est une

GAVAUDAN-POUSSEL-M.P. SCHUTZENBERGER : OLFACTION 181

véritable électivité pharmacodynamique qui ressemble à celle que l'on rencontre dans les actions des hormones sexuelles, des hormones de croissance des végétaux, des stéroïdes agissant sur le cœur, des cancérigènes. Le déclenchement de la sensation olfactive paraît comme dû à une série d'actions pharmacodynamiques parfois très spécifiques et localisées. Il ne semble pas que leur mécanisme doive être profondément différent de celui des diverses actions envisagées et nous pensons avec VELDSTRA (10) que dans ces différents cas le champ extérieur de la molécule joue un grand rôle.

Nous pensons que l'hypothèse de travail des « ergones » groupant les heteroauxines, les hormones sexuelles, les glucosides cardiaques, les substances odorantes, dans un groupe de substances agissant par l'adhésion avec le substrat mérite une étude très attentive. Il est possible que dans le cas de l'odorat, malgré l'homogénéité apparente du matériel cellulaire, uniquement constitué de cellules nerveuses, la question soit très complexe et que chaque odeur pose un problème pharmacodynamique spécial. Cette complexité même doit être riche d'enseignements.

Il apparaît d'ailleurs déjà qu'il existe certaines relations entre l'excitation des chimiorécepteurs par les hormones sexuelles et les propriétés sexualisantes de ces substances.

SANO (28), a indiqué que le musc naturel provoquait une action de croissance par badigeonnage de la crête du chapon.

PRELOG, RUZICKA, MEISTER et WIELAND (29) ont noté chez certains androsténols une variation concomitante des pouvoirs odorants et hormonaux. Ces faits sont extrêmement intéressants. On peut être en effet tenté de rechercher à la base de certains phénomènes relatifs à l'action sexualisante des hormones des phénomènes d'excitation plus ou moins comparables à ceux dont les chimiorécepteurs de l'olfaction sont le siège.

Il serait évidemment important de reconnaître sur quel type de cellule, nerveuse ou non, s'exerce une telle excitation.

D'autre part, ELSBERG (30) et LE MAGNEN (31) ont mis en évidence l'existence de variations de sensibilité olfactive liées au sexe et sans nul doute au cycle hormonal. Il est possible que les hormones sexuelles interviennent donc pour modifier la sensibilité des chimiorécepteurs de l'olfaction ou la conductibilité des étages nerveux supérieurs responsables de la sen-

sation olfactive. Il a été montré par SELYE (32) que les hormones sexuelles possédaient des propriétés narcotiques. Ce fait est peut-être en relation avec les phénomènes observés par les auteurs précédents ; nous savons en effet que les narcotiques peuvent selon les doses exciter ou déprimer parfois très électivement les fonctions de certaines cellules nerveuses.

Il existe donc des relations étroites et peut-être physiologiquement importantes entre des phénomènes en apparence aussi disparates que la sensation olfactive et les actions hormonales. On entrevoit là un champ de travail immense et l'on comprend la nécessité qu'il y a de traiter l'olfaction comme un problème général de pharmacodynamie cellulaire.

*Laboratoire de pharmacodynamie du Ministère de  
la Santé Publique au Centre d'Etudes du Bouchet.  
Mai 1948.*

#### BIBLIOGRAPHIE

- (1) J. PASSY : C. R. de l'Ac. des Sc. : les propriétés odorantes des alcools de la série grasse, *114*, 1892, 1141.
- (2) J. PASSY : Forme périodique du pouvoir odorant dans la série grasse, C. R. de l'Ac. des Sc., *116*, 1893, 1007.
- (3) Y. RENQVIST : Ueber den Geschmack, *Skand Arch. f. Physiol.*, *38-39*, 1919, 97-201.
- (4) K. H. MEYER et H. HEMMI : Beiträge zur Theorie der Narkose, III, *Bioch. Zeit.*, *277*, 1935, 39-71.
- (5) J. FERGUSON : The use of chemical potentials as indices of toxicity. *Proc. Roy. Soc. ser. B.*, *127*, 1939, 387-404.
- (6) P. GAVAUDAN, M. DODE et H. POUSSEL : La toxicologie générale et la notion d'activité thermodynamique, *Mémor. des Serv. Chim. de l'Etat*, *31*, 1944, 384-423.
- (7) TYNDALL : absorption de la chaleur rayonnante par les odeurs, traduction française abbé Moigno, 1874.
- (8) A. HEYNINX : Essai d'olfactique physiologique, Thèse Médecine Univ. libre Bruxelles, 1919, 289 p.
- (9) DYSON : *Chemistry and Industry*, *57*, 1938, 647-651.  
DYSON : *Perf. Essent. Oil Rec.*, *28*, 1937, 13.
- (10) H. VELDSTRA : Actualités biochimiques, *10* ; Aspects actuels de l'Enzymologie ; considérations sur l'interaction des ergones et de leurs substrats : 97-115.
- (11) J. LE MAGNEN : Analyse d'odeurs complexes et homologues par la fatigue. C. R. de l'Ac. des Sc., *226*, 1948, 753.

## GAVAUDAN-POUSSEL-M.P. SCHUTZENBERGER : OLFACTION 183

- (12) A. PULLMANN : *Ann. de Chimie*, 2, 1947, 5.
- (13) R. DAUDEL, *C. R. de la Soc. de Biol.*, 142, 1948, 5.
- (14) P. GAVAUDAN, H. POUSSEL et M. P. SCHUTZENBERGER : Le mécanisme physico-chimique de l'excitation sapide et la notion d'excitant indifférent. *C. R. de l'Ac. des Sc.* 226, 1948, 751-752.
- (15) H. ZWAARDEMAKER : *L'odorat*, Paris, 1925, 302 p.
- (16) H. FREUNDLICH : *Kapillarchemie*, t. I, Leipzig, 1930.
- (17) J. H. SCHULMAN : Some physical chemical properties of biologically active molecules : *Trans. Farad. Soc.* 39, 1943, 412-417.
- (18) BACKMAN : *Physiol. abstracts*, 2, 1917, 479.  
Id. : *Journ. Physiol. Pathol. Gener.* 17, 1917, 1-4.
- (19) H. R. ING : Chemical constitution and pharmacological action, *Trans. Farad. Soc.* 39, 1943, 372-380.
- (20) N. H. GRACE : *Journ. of Canad. Res.*, 17, 1939, 247, 373.
- (21) K. V. FRIEMANN et J. BONNER : *Physiol. rev.* 18, 1938, 545.
- (22) DILLS et MENUSAN : *Contr. Boyce Thomson Inst.*, 7, 1935, 63-82.
- (23) H. HURST : Principles of insecticidal action as guide to drug reactivity-phase distribution relationships, *Trans. Farad. Soc.* 39, 1943, 390-411.
- (24) J. LE MAGNEN : *C. R. de l'Ac. des Sc.*, 225, 1947, 1378.
- (24) A. ELSBERG, E. D. BREWER et I. LEVY : The relative importance of volume and pressure of the impulse for the sensation of Smell and the olfactory process. *Bull. of the neurol. Instit. New-York*, 4, 1935, 264-269.
- (26) R. W. MONCRIEFF : *The chemical senses*, London, 1944, 424 p.
- (27) M. GUILLOT : Anosmies partielles et odeurs fondamentales. *C. R. de l'Ac. des Sc.*, 226, 1948, 1307.
- (28) SANO : *Journ. Pharm. Soc. Japan.*, 56, 913-914 et *Chemical abstracts*, 32, 1938, 8519.
- (29) V. PRELOG, L. RUZICKA, P. MEISTER et P. WIELAND, *Helv. chim. Acta*, 28, 1945, 618.
- (30) ELSBERG : *Arch. neurol. psychiatr.*, Chicago, 37, 1937, 223-236.
- (31) J. LE MAGNEN : *C. R. de l'Ac. des Sc.* 226, 1948, 694.
- (32) H. SELYE : The pharmacology of steroid hormones and their derivatives, *Rev. Can. de biol.*, I, N° 1, 1942, 577-632.
- (33) G. EHRENSVARD : Ueber die primarvorgänge bei Chemozeptoren-beeinflussung, *Acta physiol. Scand.* 3, suppl. IX, 1942, 148 p.
- (34) P. GAVAUDAN, H. POUSSEL, G. BREBION, M. P. SCHÜTZENBERGER : l'Étude des conditions thermodynamiques de l'excitation olfactive et les théories de l'olfaction : *C. R. Ac. Sc. Paris*, 226, 1948, p. 1395.
- (35) A.F.H. WARD : *Trans. Farad. Soc.*, 42, 1946, 399.

## Appendice

Au moment de l'impression de ce travail nous avons pris connaissance d'un mémoire de G. EHRENSVARD (33). Cet auteur a étudié le problème de l'excitation des chimiorécepteurs de l'olfaction. EHRENSVARD s'est livré à une étude théorique détaillée et très intéressante du rôle joué par la structure et la forme des molécules. Il a accordé une grande importance à l'adsorption des molécules odorantes au niveau d'une surface limite entre deux phases et a réalisé une étude de ce phénomène au moyen de modèles physiques. Ces modèles sont constitués par des « chaînes » de liquides non miscibles de densités différentes. On mesure la différence de potentiel qui prend naissance lorsque l'on crée une dissymétrie au niveau d'une des interfaces d'une telle chaîne en ajoutant des substances organiques variées dans une des phases non aqueuses. Les phases aqueuses sont constituées par des solutions de KCl et les phases « lipoïdiques » sont constituées par des substances telles que la salicylaldéhyde, le nitrobenzène ou l'alcool benzylique. EHRENSVARD ajoute à une des phases lipoïdiques des quantités déterminées d'alcools de la série homologue (alcools normaux primaires) et trace les courbes des différences de potentiel (en mV) en fonction des quantités d'alcool (en vol %). Les courbes des alcools en C<sub>1</sub>, C<sub>2</sub>, C<sub>3</sub>, C<sub>4</sub>, C<sub>5</sub>, C<sub>6</sub>, C<sub>8</sub>, C<sub>10</sub>, C<sub>11</sub>, C<sub>12</sub>, se disposent dans un domaine commun assez réduit ; au contraire les courbes des alcools en C<sub>7</sub> et C<sub>9</sub> s'écartent beaucoup de ce domaine.

Le même résultat apparaît si l'on porte en abscisses le nombre d'atomes de carbone et en ordonnées la valeur des différences de potentiel pour une même concentration (4 %) des différents alcools. EHRENSVARD a également étudié par la même méthode les alcools secondaires. L'allure de la courbe de la valeur des potentiels en fonction du nombre d'atomes de carbone pour une même concentration en alcool est dans ce cas plus simple et il est curieux de noter le déplacement du maximum pour les termes en C<sub>5</sub> et C<sub>6</sub>. D'après EHRENSVARD

## OLFACTION : APPENDICE

185

ces phénomènes d'action spécifique pour un certain nombre d'atomes de carbone sont dus à la disposition spatiale des chaînes carbonées.

EHRENSVARD a rapproché les oscillations qu'il a observées de l'oscillation des seuils olfactifs remarquée par divers auteurs. Il est tout à fait intéressant de noter non seulement l'existence d'une activité maxima pour un certain nombre de maillons carbonés ( $C_7$  et  $C_8$ ) dans le phénomène d'EHRENSVARD mais même d'une oscillation pair - impair (tout au moins pour les systèmes salicyladéhyde-alcools n-primaires). EHRENSVARD attribue le déclenchement de l'excitation olfactive à une modification de potentiel prenant naissance au niveau d'une interface appartenant aux récepteurs.

Selon cet auteur le phénomène d'oscillation du pouvoir odorant dans les séries homologues ne possède pas de correspondant parmi les phénomènes physiques, à l'exception des modifications alternantes qu'il a observées avec des modèles électriques (1).

Les remarques qui suivent ne sont nullement destinées à diminuer l'importance des remarquables phénomènes observés par EHRENSVARD, mais il faut observer qu'il n'est pas exact de dire que l'oscillation des seuils odorant est un phénomène biologique qui n'a pas son correspondant dans le domaine des phénomènes physiques. L'oscillation des points de fusion est un phénomène qui n'est pas sans analogie avec l'oscillation des potentiels observés par EHRENSVARD. On a vu que nous sommes parvenus à partir de cette simple constatation à des conclusions voisines de celles d'EHRENSVARD puisque nous admettons que c'est à des oscillations de propriétés physiques actuellement plus ou moins bien connues, que sont dues les oscillations du seuil et la qualité sensorielle. Ces propriétés physiques particulières de chaque type de molécule de la série homologue, que nous connaissons d'ailleurs très mal, sont évidemment liées à la structure et à la forme des molécules et ce sont elles qui sont responsables du déclenchement des mécanismes d'excitation olfactive et des

(1) Das Auftreten von Alternationserscheinungen und sprunghaften Intensitäts Schwankungen in homologen Reihen gewisser Riechstoffe ist z.B. ein Vorgang für den kein physikalisch chemisches Analogon gegeben werden kann, mit Ausnahme von entsprechenden Phänomenen in gewissen Grenzflächensystemen (*Loc. cit.*, p. 133).

variations de potentiel observées par EHRENSVARD.. On pourrait tenter de rechercher s'il existe dans les propriétés physiques des corps en solution quelque chose qui rappelle l'oscillation du seuil olfactif.

Il se peut que l'on puisse obtenir quelques indications analogues à celles que donne l'étude des variations de points de fusions en considérant les variations de certaines constantes, telles que la constante capillaire et le coefficient d'activité.

On constate que les incréments de ces constantes présentent des fluctuations en fonction du nombre d'atomes de carbone, mais il est difficile de connaître la part imputable aux erreurs d'expérience et c'est aux physiciens qu'il appartient de nous renseigner à ce sujet qui est essentiel pour le biologiste. En effet, la théorie de la règle de TRAUBE pour les solutions capillairement actives ne nous fournit que des renseignements trop sommaires sur les paramètres décrivant la configuration stérique des molécules. La théorie (WARD) (35) ne fournit d'indications que sur la dimension principale du volume enveloppant des molécules dont les formes sont supposées varier simplement du cylindre au sphéroïde en fonction du nombre des atomes de carbone. Dépassant cette discrimination sommaire de formes, les dispositifs olfactifs semblent permettre une reconnaissance plus fine de l'architecture moléculaire (parité et imparité ainsi que nombres critiques d'atomes de carbone, etc.).

Quelles que soient les possibilités actuelles de l'étude physique d'une éventuelle variation des constantes auxquelles il a été fait allusion et des causes profondes du phénomène d'EHRENSVARD, il n'en demeure pas moins exact que le principe d'une relation entre l'oscillation du seuil biologique et une oscillation des propriétés physiques doit être exact.

Remarquons en terminant que pratiquement la règle biologique de TRAUBE n'obéit elle-même, qu'en *moyenne* au coefficient 3. Des expérimentateurs aussi consciencieux que FÜHNER et WARBURG ont respectivement trouvé des quotients methyl/ethyl = 1,8 et heptyl/octyl = 3,4 dans l'inhibition du développement de l'œuf d'oursin d'une part, des quotients methyl/ethyl = 1,5 et ethyl/propyl = 7,8, dans l'inhibition de la réaction de BLACKMANN chez *Chlorella*. Des irrégularités de ce type ont été constatées et soulignées par divers auteurs

Année 1948      1948-8. Étude pharmacodynamique des chimiorécepteurs de...

OLFACTION : APPENDICE

187

dans leurs critiques de la théorie capillaire de la narcose. La cause de ces irrégularités mériterait d'être étudiée. On suppose naturellement que les auteurs ont toujours utilisé des alcools rigoureusement purs.

# GALLICA BIOLOGICA ACTA

publiés sous la direction de P. Gavaudan  
*Directeur du Service de Biologie  
des Services Chimiques de l'Etat*

VOL. 1 FASC. 2 JUIN 1948

## S O M M A I R E

### Articles originaux.

- S. Bazin** : Etude de la pénétration de diverses substances chimiques dans le système vacuolaire de « Chara » ..... p. 107
- P. Gavaudan et H. Poussel** : Remarques sur les théories de la cancérisation chimique ..... p. 111
- H. Poussel** : Sur les mécanismes des actions toxiques de l'hexachlorocyclohexane  $\gamma$  ..... p. 114
- P. Gavaudan, H. Vogel et G. Debraux** : Etudes des propriétés rhizogènes de la  $\beta$  lactone de l'acide  $\beta$  isopropylmalonique . . . . . p. 120
- G. Brebion et P. Gavaudan** : Etude de la photosynthèse des plantes aériennes par la méthode de Warburg ..... p. 124
- G. Debraux** : Recherches sur les feuilles isolées en survie. .... p. 131
- G. Debraux et P. Gavaudan** : Recherches sur la prophylaxie du mildiou de la vigne par les composés organiques ..... p. 138
- P. Gavaudan, G. Brebion, H. Poussel et M. P. Schützenberger** : Etude pharmacodynamique des chimiorécepteurs de l'olfaction; (1) Principes et méthodes d'une étude thermodynamique de l'olfaction . . . . . p. 147
- P. Gavaudan, H. Poussel et M. P. Schützenberger** : Etude pharmacodynamique des chimiorécepteurs de l'olfaction; (2) L'étude thermodynamique des séries homologues et les théories de l'olfaction ..... p. 167
- N. Kobozieff et N. A. Pomriaskinsky-Kobozieff** : Possibilité de sélection sur l'anomalie de l'oreille externe chez la souris. p. 184
- M. Schützenberger** : Remarques sur les relations aléatoires d'ordre ..... p. 187

### Revue générale.

- P. Gavaudan** : Echanges de matériaux figurés entre noyau et cytoplasme..... p. 205

LIBRAIRIE LE FRANÇOIS  
91, Boulevard Saint-Germain - PARIS

Statistique appliquée à la psychologie**Remarques sur les relations  
aléatoires d'ordre  
et leur application à la psychologie**

par Marco P. SCHÜTZENBERGER.

Dans de nombreux domaines de la psychologie et de la psycho-sociologie intervient la notion de *relation aléatoire d'ordre* ; ayant à comparer deux ou plusieurs objets, un sujet ou un groupe de sujets les classe par ordre de grandeur, de préférence, d'utilité..., etc... En fait, ce classement n'est pas rigide mais se traduit par des expressions telles que « A est en général préféré à B » .. « Les R sont dans l'ensemble supérieurs aux S »... « X a été élu contre Y » ou avec plus de rigueur « A a été préféré à B dans p pour cent des cas »... « q pour cent des X sont supérieurs aux Y » « X a obtenu une majorité de n voix... ».

On rappellera schématiquement ici les hypothèses probabilistes plus ou moins classiques qui servent de base à l'interprétation de ces résultats et on s'efforcera de montrer certaines de leurs implications relatives à la construction d'une échelle absolue sur laquelle pourrait se disposer tout ensemble de tels objets.

Enfin on indiquera des possibilités d'expérimentation destinées à contrôler la validité des théories exposées.

*L'échelle psychologique.*

Le sujet, les yeux fermés, compare, par couple, en les soulevant trois poids A, B, et C en indiquant à chaque comparaison lequel des deux objets est le plus léger à l'exclusion de toute autre réponse alors que dans la méthode des trois choix

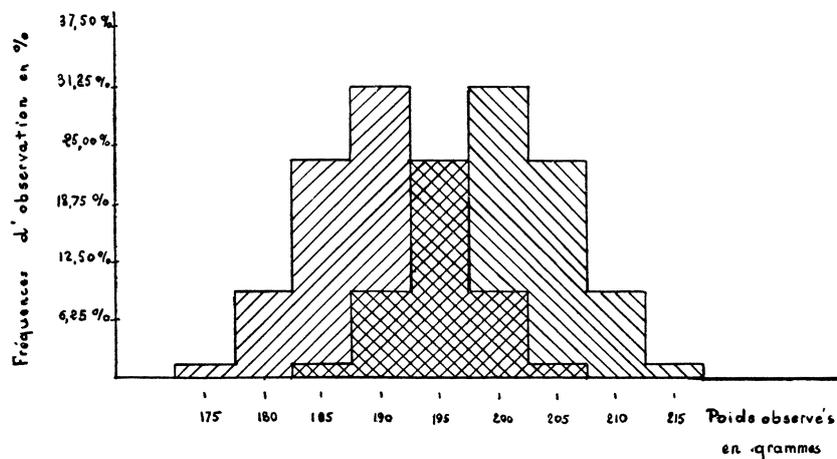
(Dreihauptfälle de WUNDT) une réponse « douteux » est admise. Cette dernière méthode introduisant des éléments étrangers au point discuté ici n'esera d'ailleurs pas abordé dans le cours du présent article.

Ce processus de comparaison est répété une centaine de fois et les résultats suivants ont été obtenus.

A plus léger que B	80 %
B » A	20 %
A » C	95 %
C » A	5 %
B » C	70 %
C » B	30 %

FIG. 1. — On a porté en abscisses les fréquences pour lesquelles l'objet A (hachures de haut en bas et de droite à gauche) et l'objet B (hachures de haut en bas et de gauche à droite) sont observés dans les différentes marges de poids.

Ainsi dans 1,5625 % des cas A est observé entre 172,5 gm et 177,5 gm dans 9,375 % entre 177,5 gm et 182,5 gm, etc...



Ces chiffres correspondent à peu près à A = 190 gm  
B = 200 gm C = 205 gm.

C'est l'exemple le plus simple d'une *relations aléatoire d'ordre* : ici la relation perçue « plus léger ».

Pour interpréter cette expérience, on peut la rapprocher d'une autre, analogue, mais réalisable par des dispositifs exclusivement mécaniques ; on se borne à la comparaison de deux poids. A et B qui sont pesés successivement au moyen

## SCHUTZENBERGER : RELATIONS ALÉATOIRES D'ORDRE 193

d'une balance d'une sensibilité de l'ordre de 5 gm ; à la  $i^{\text{ème}}$  mesure les poids obtenus sont  $a_i$  et  $b_i$  ; l'observateur est censé ne pas connaître ces chiffres mais seulement, savoir à chaque fois si  $a_i$  est plus petit que  $b_i$  ou inversement (fig. 1).

On se trouve ainsi dans la même situation que dans l'expérience de psychophysiologie précédente où n'étaient connus que les résultats des comparaisons entre A et B.

Dans la seconde expérience, la théorie des erreurs montre que les poids successifs observés  $a_i$  se distribuent selon une certaine loi de fréquence autour de la *vraie valeur*  $\bar{a}$  de A ; c'est-à-dire que chaque  $a_i$  peut être décomposé en deux parties  $a_i = \bar{a} + \alpha_i$  de même  $b_i = \bar{b} + \beta_i$  ou  $\alpha_i$  et  $\beta_i$  fluctuations (ou erreurs) sont les variables aléatoires *indépendantes* l'une de l'autre à chaque expérience et d'une expérience à l'autre.

Par analogie, on est donc amené à supposer qu'à la  $i^{\text{ème}}$  comparaison par paire de A et de B, le sujet perçoit le poids des deux objets aux niveaux  $a_i$  et  $b_i$  sur une *échelle psychologique*, cette opération comportant des *fluctuations* réparties selon certaines lois de fréquence, qui, à chaque expérience successives font percevoir plus ou moins légers les deux poids A et B (fig. 2 et fig. 3).

De même l'intensité du « sentiment » (ou emploie ici à dessein ce mot peu précis) que le sujet éprouve en examinant, par exemple, un objet X, ou son adhésion à une thèse Y, est censé fluctuer d'une expérience à l'autre et selon une loi définie, autour d'une valeur centrale, sur une certaine échelle psychologique continue (continuum affectif).

Ainsi si les stimuli A, B, C, .. etc... étaient simultanément proposés au sujet, à la  $i^{\text{ème}}$  de ces expériences de comparaison, l'ordre qu'il indiquerait serait celui des intensités  $a_i, b_i, c_i$  sur cette échelle psychologique.

Plus ou moins implicitement l'on a donc fait les hypothèses suivantes :

1° *Hypothèse d'isomorphie* (ou d'ordre total).

L'échelle psychologique E a une structure analogue à l'échelle physique ce qui implique

1-1 Etant donnés deux objets A et B comparés dans la  $i^{\text{ème}}$  expérience, c'est-à-dire pris aux niveaux  $a_i$  et  $b_i$ , on a toujours

$$a_i \leq b_i \text{ ou } b_i \leq a_i, \\ \text{et si } a_i \leq b_i \text{ et } b_i \leq a_i \\ \text{alors } a_i = b_i. —$$

1-2 Etant donnés trois objets quelconques A, B et C dans la  $i^{\text{ème}}$  expérience :

Si  $a_i < b_i$  et  $b_i < c_i$  alors  $a_i < c_i$

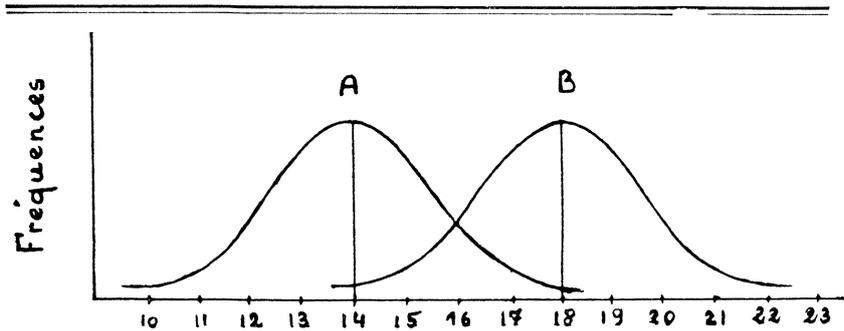


FIG. 2. — Voici par exemple les courbes de répartition relatives à deux objets A et B sur une certaine échelle psychologique.

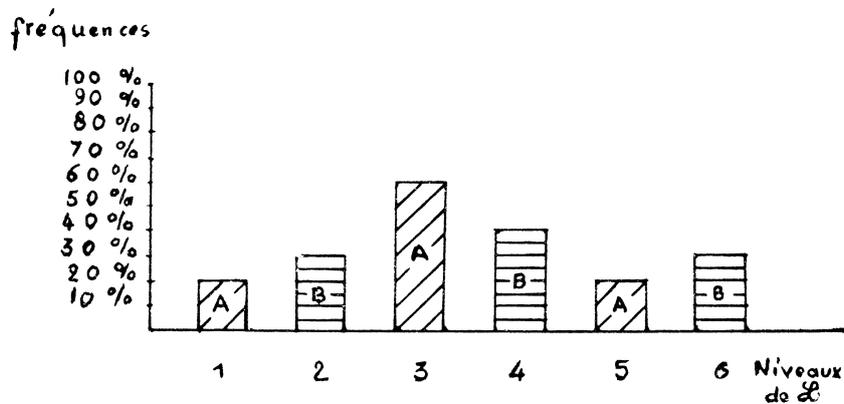
Dans les exemples suivants, on a pris des répartitions discontinues conventionnelles au lieu des répartitions continues qui se rencontrent dans la pratique. Ainsi ce schéma signifie que A peut être observé aux niveaux 1, 3 et 5 avec les probabilités respectives 20 %, 60 %, 20 %, B aux niveaux 2, 4 et 6 avec les probabilités respectives 30 %, 40 %, 30 %.

Pour calculer Prob (ab) on cherche tous les cas possibles où  $a < b$

A B	Prob A	Prob B	Prob (A) × Prob (B)
1 2	10/100	30/100	600/10.000
1 4	20/100	40/100	800/10.000
1 6	20/100	30/100	600
3 4	60/100	40/100	2.400
3 6	60/100	30/100	1.800
5 6	20/100	30/100	600

Total : Prob (ab) = 6.000/10.000 = 0,60

FIG. 3.



2° *Hypothèse d'indépendance.*

2-1 *Indépendance globale* : la fonction de répartition des fluctuations ou erreur de A est *indépendante de l'objet B* auquel il est comparé (tout au moins pour certains ensembles d'objets B).

2-2 *Indépendance locale* : la valeur de la fluctuation  $\alpha_i$  de A dans la  $i^{\text{ème}}$  expérience est *indépendante de la valeur de la fluctuation*  $\beta_i$  de B dans cette même expérience.

2-3 *Indépendance temporelle* : les valeurs  $\alpha_i$ ,  $\beta_i$ , etc., sont *indépendantes des valeurs antérieures* auxquelles ont été prises les objets A, B et C dans les comparaisons précédentes.

3° *Hypothèse de normalité.*

3. Il est possible de représenter la distribution des fluctuations de A sur l'échelle psychologique E par une *loi de répartition normale* (courbe en cloche de GAUSS).

$$y = (2 \pi \sigma)^{-1/2} \exp -\frac{1}{2} \left( \frac{x - a}{\sigma} \right)^2$$

3' Les dispersions (déviations standards) de ces courbes sont égales.

L'hypothèse 3 est issue directement de la théorie physique des erreurs et de même que l'hypothèse 3' introduite par THURSTONE dans un but de simplification des calculs, elle présente des avantages considérables pour le traitement mathématique de la théorie. Sa validité est peut-être contestable dans de nombreux cas ; malgré son utilité pratique et sa justification théorique dans certaines limites, elle n'a qu'un rôle accessoire dans l'étude des relations aléatoires d'ordre que l'on tente ici.

*Généralisation.*

A partir de ce modèle probabiliste, deux extensions successives se proposent d'elles-mêmes :

Tout d'abord, au lieu d'un seul sujet effectuant un grand nombre de fois la comparaison entre A et B, on la fait répéter par un grand nombre de sujets.

Si l'on peut supposer que l'échelle E et les fonctions de distribution des fluctuations sont les mêmes pour tous les sujets (population homogène), il est clair que l'on se trouve formellement dans la même situation que lorsqu'il s'agit d'un sujet unique.

Si non, c'est-à-dire, si les fonctions de répartition varient

d'un sujet à l'autre, il est nécessaire de faire jouer un quatrième hypothèse d'indépendance de ces variations pour se ramener au cas initial, en confondant les fluctuations inter-individuelles et intra-individuelles.

Mais la généralisation la plus intéressante est celle relative aux cas où n'existe pas préalablement d'échelle physique *mesurable* des objets, comme celle des poids, des dimensions, des températures... etc... Ainsi en est-il, par exemple, pour les jugements éthiques ou esthétiques.

Il est alors tentant de chercher à établir une échelle objective, en utilisant les méthodes valables dans les cas envisagés jusqu'ici, échelle sur laquelle on pourrait repérer tout nouvel objet par sa position relative aux objets déjà étudiés.

On a pu ainsi établir des échelles pour la valeur publicitaire de « slogans », pour le mérite relatif de dissertations littéraires, pour la qualité de dessins d'enfants, etc...

*Construction de l'échelle psychologique.*

On ne discutera pas ici les critiques formulées contre la validité des hypothèses de base, notamment contre 1-1 et 1-2 le caractère multidimensionnel des jugements qualitatifs, contre 2-1 et 2-2 le déplacement du point de vue selon le couple d'objets à comparer, l'effet de contraste et enfin, contre 2-3 l'émoussement de la sensation. On laissera de côté les hypothèses de normalité, d'ailleurs plus spécialement techniques, pour se placer dans le cas général et pour supposer que les répartitions sont quelconques.

Ainsi sur l'échelle psychologique E les fluctuations (ou erreurs) relatives à l'objet A sont réparties selon une certaine loi  $F_A$  et, indépendamment de celles relatives à B, qui sont elles mêmes réparties selon  $F_B$  ... etc...

La probabilité pour que A soit observé plus petit que B se calcule sans difficulté théorique

$$\text{Prob (ab)} = \int_{-\infty}^{+\infty} F_A(t) d F_B(t)$$

Pour trois objets A, B et C la probabilité pour qu'ils soient observés dans l'ordre A, B, C est donnée par

$$\text{Prob (abc)} = \int_{-\infty}^{+\infty} F_A(t) (1 - F_C(t)) d F_B(t)$$

On calculerait de même la probabilité Prob (Sn)

SCHUTZENBERGER : RELATIONS ALÉATOIRES D'ORDRE 197

que les  $n$  objets  $A, B, \dots$  soient classés dans l'ordre défini par la séquence  $S_n$  au cours d'un processus de comparaison simultanée. Evidemment il s'agit là d'un calcul purement formel, pouvant d'ailleurs conduire dans la pratique à des difficultés considérables, dès que le nombre d'objets est supérieur à trois ou à quatre.

D'autre part, les diverses probabilités entretiennent entre elles des rapports d'égalité, comme :

$$\text{Prob}(ab) + \text{Prob}(ba) = 1 \text{ ou comme :}$$

$\text{Prob}(abc) + \text{Prob}(acb) + \text{Prob}(cab) = \text{Prob}(ab)$  ou d'inégalité comme :

$1 \leq \text{Prob}(ab) + \text{Prob}(bc) + \text{Prob}(ca) \leq 2$  (1) et ces relations peuvent d'ailleurs se calculer systématiquement au moyen d'une algèbre spéciale assez maniable.

Réciproquement, si les lois  $F$  sont *partiellement spécifiées*, c'est-à-dire si elles ne dépendent que d'un nombre fini de paramètres inconnus, il est clair que l'on peut, sous certaines conditions, remonter aux paramètres inconnus à partir de la connaissance empirique des  $\text{Prob}(S)$  (que l'on suppose égales aux fréquences relatives d'apparition de ces séquences).

Notamment THURSTONE a indiqué une méthode permanente de calcul, valable quand les  $F$  sont des courbes normales.

Si les fonctions de répartition sont entièrement inconnues le problème se pose de la construction de  $F_A, F_B, F_C$  donnant lieu aux mêmes valeurs que les  $\text{Prob}(S_n)$  que l'on connaît, ce qui généralise le problème des moments de TCHEBYCHEFF. Il n'est pas résolu sauf dans un cas particulier et où deux objets distincts sont seuls en cause ce qui est d'ailleurs précisément le cas étudié par cet auteur quand on connaît les valeurs de :  $\text{Prob}(ab) \text{ Prob}(aab) \text{ Prob}(aa\dots ab) \dots$

*Relation statistique d'ordre.*

Puisqu'il n'est pas encore possible de déterminer les  $F$  à partir des valeurs observées des  $\text{Prob}(S_n)$ , dans le cas fréquent dans la pratique, où l'hypothèse de normalité ne peut être légitimement adoptée, il s'impose donc d'essayer de caractériser simplement les rapports qu'entretiennent entre eux les

(1) Cette inégalité nous a été indiquée par M. le Pr. FRECHET, que nous tenons à remercier ici de ses conseils qui nous ont été précieux.

objets A, B, C, ...etc... de manière à pouvoir appréhender l'essentiel de la structure de leur ensemble.

Dans l'expérience psycho-physiologique exposée au début de cet article la manière la plus immédiate de décrire les rapports entre A, B, et C est évidemment « A est plus léger que B et B que C » ; l'on essaiera donc, toujours par analogie, de définir une relation de ce type, applicable aux objets qui se rencontrent dans les diverses généralisations de cette expérience.

La première idée, serait de prendre les indices de tendance centrale, par exemple les valeurs moyennes de A et de B sur l'échelle E, pour poser : A *est inférieur* à B si la *valeur moyenne*  $\bar{a}$  de A *est inférieure* à la *valeur*  $\bar{b}$  de B.

Cette généralisation directe n'est cependant pas possible, en effet, les seules données observées sont certaines Prob (Sn) et il est possible de construire, par exemple, deux systèmes de  $F_A$  et de  $F_B$  ayant même Prob (ab) et tels que dans l'un  $\bar{a} < \bar{b}$  et dans l'autre  $\bar{b} < \bar{a}$ .

Il en serait de même si l'on prenait la valeur médiane pour indice de tendance centrale.

Ceci est d'ailleurs évident si l'on remarque que les Prob (Sn) étant données, E n'est défini qu'à un isomorphisme près, conservant sa structure d'ensemble totalement ordonné.

Pour prendre une image physique, plus intuitive, si l'on suppose que les  $F$  représentent des répartitions de masse sur un fil extensible les intégrales :

$$P(xy) = \int_{-\infty}^{+\infty} F_x(t) dF_y(t)$$

c'est-à-dire, les Prob (xy), ne changent pas quand les longueurs relatives des différents segments du fil sont modifiées.

En fait supposant connues la totalité des Prob (Sn) pour un ensemble donné d'objets, il est seulement possible (théoriquement) de déterminer tous les couples de répartition relatives. (Pour deux objets A et B, la répartition relative est la valeur de  $F_A(t)$  en fonction de la valeur de  $F_B(t)$  où t est une description paramétrique quelconque de l'échelle.)

On est donc logiquement conduit à définir la relation statistique d'ordre : « A statistiquement inférieur à B » à partir

SCHUTZENBERGER : RELATIONS ALÉATOIRES D'ORDRE 199

des seules données immédiatement observables, et on posera

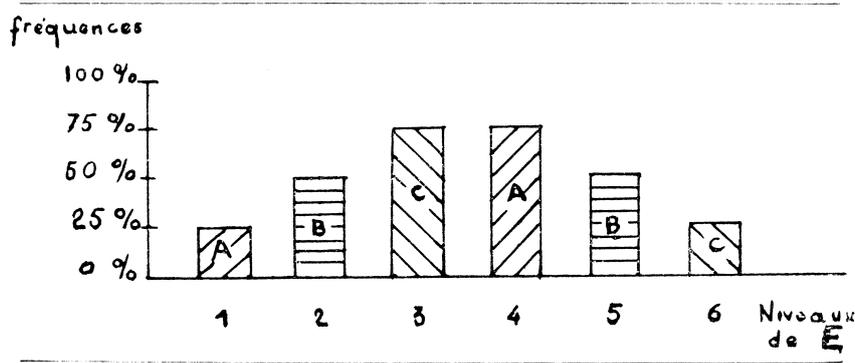
« A statistiquement inférieur à B » équivalent à :  $\text{Prob}(ab) \leq 1/2$ , équivalent à : « A est observé inférieur à B dans au moins 50 % des cas ».

Si l'on revient maintenant à l'exemple initial, on constate sans peine que cette définition est équivalente, dans ce cas précis, aux autres définitions avancées au début de ce paragraphe et qui semblent peut-être plus directes.

De même, chaque fois que les courbes  $F_X$  peuvent être considérées comme normales, la notion d'ordre statistique coïncidera avec l'ordre des indices de tendance centrale des  $F_X$  sur E ce qui pourrait se traduire par l'expression imagée suivante : si les répartitions  $F_A$  et  $F_B$  sont normales, « A en moyenne inférieur à B » est équivalent à « A en général inférieur à B ».

*Les contradictions de la relation statistique d'ordre.*

Mais si la notion ainsi définie n'entraîne aucune difficulté dans les cas simples que l'on vient de citer il n'en est pas de même dans l'exemple suivant (fig. 4) :



On a  $\text{Prob}(ab) = 0,625$   
 $\text{Prob}(bc) = 0,625$   
 $\text{Prob}(ca) = 0,5625$

Les trois probabilités sont plus grandes que 0,50 et l'on

doit donc admettre que peuvent être *simultanément vraies* les trois affirmations suivantes :

A est statistiquement inférieur à B	
B	» C
C	» A

D'ailleurs malgré son caractère artificiel cet exemple est représentatif des cas assez fréquents, en effet la relation  $\text{Prob}(ab) + \text{Prob}(bc) + \text{Prob}(ca) = 1 + \text{Prob}(abc) + \text{Prob}(cab) + \text{Prob}(bca)$  montre que si  $\text{Prob}(ab)$  et  $\text{Prob}(bc)$  sont légèrement supérieurs à  $1/2$  il en sera de même de  $\text{Prob}(ca)$  dès que les trois probabilités du second nombre seront assez grandes.

On doit donc admettre que la conception d'ordre statistique est en contradiction, dans certains cas, avec les propriétés caractéristiques habituelles d'une relation d'ordre c'est-à-dire la *transitivité* (1).

Il faut en outre remarquer que ce caractère de non transitivité résulte de la nature statistique de la définition car pour chaque expérience concrète où A, B et C sont simultanément comparés l'on a nécessairement l'un des six classements possibles :

$$a_i \leq b_i \leq c_i \quad \text{ou} \quad b_i \leq c_i \leq a_i \quad \text{ou} \quad c_i \leq a_i \leq b_i$$

$$\text{ou} \quad a_i \leq c_i \leq b_i \quad \text{ou} \quad b_i \leq a_i \leq c_i \quad \text{ou} \quad c_i \leq b_i \leq a_i$$

dans chacun desquels  $\leq$  est une relation transitive puisque portant sur les *valeurs actuelles*  $a_i, b_i, c_i$  de A, B et C (hypothèse 1-2).

On ne discutera pas ici les applications de cette contradiction à la critique des *théories économiques* dites *marginalistes* qui utilisent largement la notion de préférences statistiques dans leur théorie de la valeur. Mais, si la réfutation du marginalisme ne peut se faire valablement que sur le terrain de la réalité économique concrète des profits et salaires, il n'est cependant pas inutile de montrer la possibilité d'erreurs graves qu'il y a à étendre sans précaution des règles valables dans un cas concret (transitivité de préférence) à un niveau

(1) Une relation est dite transitive, si, pour trois objets quelconques a, b et c, ab et bc entraîne ac ; telles sont par exemple les relations « plus petit » - « plus aigu » - « héritier de » - « contenu dans » - « frère de » -

Une relation est dite non transitive si l'affirmation précédente n'est pas toujours valable ; tel est le cas par exemple de la relation « à l'est » (au sens strict, à moins de 180°) Moscou à l'est de New-York ; New-York à l'est de Shangaï mais : Moscou à l'ouest (et non pas à l'est) de Shangaï.

## SCHUTZENBERGER : RELATIONS ALÉATOIRES D'ORDRE 201

supérieur (non transitivité statistique) où leur dépassement dialectique est nécessaire.

Il faut en outre remarquer que le modèle probabiliste que l'on vient d'exposer peut être théoriquement suffisant pour expliquer les prétendus *jugements inconsistants* (A préféré à B ; B préféré à C ; C préféré à A) que l'on a rencontrés dans des enquêtes psychologiques. Il serait intéressant de voir, si, abandonnant l'hypothèse de la normalité des courbes, on ne pourrait pas interpréter plus simplement les faits observés qu'en supposant une « inconsistance » des préférences ou une impossibilité de classer simultanément les objets proposés parce que ceux-ci relèvent de plusieurs dimensions (points de vue) différents.

Mais à l'intérieur même de cette contradiction il est possible de fixer des limites à la non transitivité.

L'inégalité

$1 \leq \text{Prob}(ab) + \text{Prob}(bc) + \text{Prob}(ca) \leq 2$  montre que si  $\text{Prob}(ab)$  et  $\text{Prob}(bc)$  sont plus grands que  $3/4$ ,  $\text{Prob}(ac)$  est au plus égal à  $1/2$ .

On pourrait traduire ceci d'une manière moins précise mais plus parlante en disant que si A est *statistiquement très inférieur* à B et B *très inférieur* à C alors A est *statistiquement inférieur* à C.

Autrement dit encore, la relation d'ordre statistique est transitive lorsqu'elle n'intervient qu'entre des objets assez éloignés les uns des autres sur leur échelle commune.

Une deuxième limitation de la non transitivité se rencontre dans le cas où l'un des objets A est *toujours inférieur* à l'autre B c'est-à-dire quand  $\text{Prob}(ab) = 1$ .

On démontre alors que si B est *statistiquement inférieur* à C il en est de même de A.

Autrement dit, quand la relation « *statistiquement inférieure* » est dans son cas limite de « *toujours inférieur* », elle redevient transitive.

On doit cependant remarquer que « A *toujours inférieur* à B » et « B *statistiquement inférieur* à C » n'entraîne pas

A *toujours inférieur* à C

mais seulement

A *statistiquement inférieur* à C

comme le montre l'exemple suivant (fig. 5).

202

GALLICA BIOLOGICA ACTA

VOL. I 1948

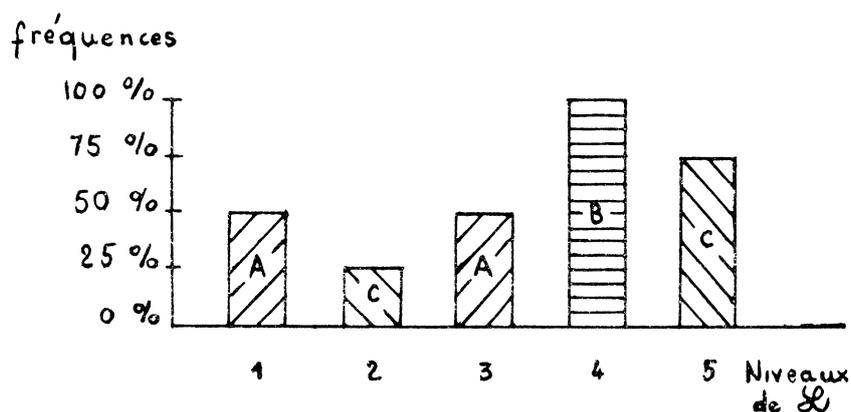


Fig. 5. — Ici on a :  
 Prob (ab) = 1,00  
 Prob (ac) = 0,775  
 Prob (bc) = 0,75

#### *Retour à l'expérimentation.*

Il serait fastidieux d'insister plus longuement sur les propriétés nouvelles (plus riches) de la relation statistique d'ordre.

On indiquera encore quelques possibilités d'expérimentation permettant de contrôler directement la théorie développée ici, expériences en cours que l'on espère pouvoir exposer et discuter dans un travail ultérieur.

#### *Test d'indépendance.*

On part des égalités du type

$\text{Prob} (abx) + \text{Prob} (axb) + \text{Prob} (xab) = \text{Prob} (ab)$   
 vraies quelque soient les objets A, B et X.

Il est possible, en faisant classer aux sujets les triples d'objets A, B, X A, B, Y A, B, Z... etc.. d'obtenir plusieurs estimations de  $\text{prob} (ab)$  qui doivent être égales, si les hypothèses d'indépendance sont valables dans le cas étudié.

#### *Test de normalité.*

Tout d'abord en se propose de rechercher systématiquement des ensembles *effectifs* d'objets donnant lieu à une non transitivité de la relation d'ordre.

SCHUTZENBERGER : RELATIONS ALÉATOIRES D'ORDRE 203

D'autre part, pour tester directement la normalité des répartitions dans le cas des expériences psycho-physiologiques il est souvent possible d'utiliser la méthode suivante :

On donne à classer à un sujet 4 objets A, A', B, B' ; A et A' d'une part, B et B' d'autre part étant égaux entre eux du point de vue de l'expérience.

Six classements sont évidemment possibles pour le triple A, A', B.

$$\begin{aligned} & A, A', B - A', A, B \\ & A, B, A' - A', B, A \\ & B, A, A' - B, A', A \end{aligned}$$

Par définition, les deux classements situés sur chacune des lignes ont la même probabilité, respectivement :

$$\text{Prob}(aab), \text{Prob}(aba), \text{Prob}(baa)$$

On a la relation :

$$2 \text{Prob}(aab) + \text{Prob}(aba) = \text{Prob}(ab)$$

De même avec le triple B, B' et A on a :

$$2 \text{Prob}(abb) + \text{Prob}(bab) = \text{Prob}(ab)$$

Sur les dix *degrés de liberté* de l'expérience (5 relatifs au triple A, A', B et 5 relatifs au triple A, B, B').

- 6 servent à comparer les deux estimations de Prob(aab), Prob(aba), Prob(abb), Prob(baa), Prob(bab),
- 1 permet de comparer l'estimation de Prob(ab) due à l'expérience avec A, A' et B avec celle due à A, B et B'.

- Les deux suivantes, *dans l'hypothèse de la normalité*, permettent d'estimer les deux rapports  $\sigma_A/d$  et  $\sigma_B/d$  où d désigne la distance des moyennes des répartitions de A et de B et  $\sigma_A$  et  $\sigma_B$  les déviations standards de ces mêmes répartitions.

Il reste donc encore un degré de liberté pour tester l'accord de la théorie avec l'expérience et cela sans que le sujet ait eu à effectuer de démarches plus complexes que le *classement simultané d'un triple d'objets*.

#### RESUME ET CONCLUSIONS

On a exposé le modèle probabiliste classique servant à expliquer le mécanisme des relations aléatoires d'ordre.

On en a déduit la définition d'une « relation statistique d'ordre » non toujours transitive dont quelques propriétés, parmi les plus immédiates, ont été soulignées.

204

GALLICA BIOLOGICA ACTA

VOL. I 1948

Enfin on a suggéré des expériences de contrôle dont les résultats seront présentés et discutés dans un travail ultérieur.

*Section de Laboratoires du Ministère de la Santé Publique  
au Centre d'études du Bouchet.*

#### BIBLIOGRAPHIE

Pour un exposé classique des théories développées ici, on renvoie le lecteur aux ouvrages suivants qui contiennent une bibliographie complète :

GUILFORD : Psychometric Methods, New-York, 1936.

LUNDBERG : Social research, New-York, 1940.

et surtout aux mémoires fondamentaux de :

THURSTONE : Psychophysical analysis, Am.. J. Psy. 1927, pp. 368-389 - A study of comparative judgement, Psy. Rev. 1927, pp. 273-286 - Theory of attitude measurement. Psy. Rev. 1929, pp. 222-241.

PSYCHOMETRIKA—VOL. 13, NO. 2  
JUNE, 1948

## AN ABAC FOR THE SAMPLE RANGE

MARCO P. SCHÜTZENBERGER  
PARIS, FRANCE

An abac is computed which gives the probability that at least  $q\%$  of the whole population will be included in the range of a random sample of given size. Applications are suggested for testing homogeneity of a sampling from a given population.

Let a sample of  $N$  values be randomly drawn from an infinite continuous distribution. The present chart gives the probability  $p$  that at least  $q$  per cent of the whole population lie between the extreme values of the sample.

For instance, if the universe is that of the speeds in a given performance test, and the sample is constituted of  $N$  subjects passing in a given day,  $q$  will be the proportion to all subjects of those who will complete the task slower than the best of the sample and faster than the poorest.

It is easy to prove that the probability  $p$ , the percentage  $q$ , and the sample size  $N$  are related by

$$p = 1 - \left(\frac{m-1}{m}\right)^{N-1} \left(\frac{N+m-1}{m}\right)^*, \quad (1)$$

where

$$m = \frac{1}{1-q}. \quad (2)$$

For large enough  $N$  and  $m$ , (1) may be closely approximated by

$$p = 1 - (r+2)e^{-r-1}, \quad (3)$$

where

$$r = \frac{N}{m} - 1. \quad (4)$$

The important fact is that (1) holds for every *continuous infinite* distribution, even leptokurtic or platykurtic, skew or multimodal.

\* Wilks, S. S. *Mathematical statistics*. Princeton, New Jersey: Princeton University Press, 1943. P. 93.

*Use of the Abac*

In ordinates are plotted, on a logarithmic scale, values of  $q$  (proportion of the whole population included in the sample range) from 50 per cent up to 998 per thousand.

In abscissas, on a logarithmic scale also, are values of the sample size  $N$  from 5 up to 1000.

Equi-probability curves are drawn for  $p = 999/1000, 99/100, 90/100, 75/100, 50/100, 25/100, 10/100, 1/100, 1/1000$ , so that interpolation can be made easily, and the corresponding values are written in the left and upper margins of the abac.

For example:

for  $N = 20$  and  $p = 99/100$ ,  $q$  is 75 per cent.

for  $N = 200$  and  $p = 999/1000$ ,  $q$  is 955 per thousand.

*Applications*

Two main applications may be suggested in the field of current psychology:

In forecasting the proportion of subjects included in the range of a given sample. For example, in a preliminary trial on 20 subjects of the speed test referred to above, extreme values of 45 and 135 seconds have been observed. Taking the equi-probability curve for  $p = 99/100$ , we may guess with 99 odds against one, that at least 75 per cent of all future subjects will score between 45 and 135 seconds. With 9 odds against one (that is to say, with  $p = 90/100$ ) and 100 subjects instead of 20, this proportion would rise up to 960 per thousand.

An alternative use of the chart may be in testing a bias in sampling when there is a hint of range discrepancies and when classical tests are too difficult or lengthy to compute. For example:

Let us suppose we got from a large population the decilage of a test and that in a sample of seven, we found a subject in the first percentile and one in the last. Here  $q$  is at least 98 per cent,  $N = 7$ , and there is just a single chance out of a hundred ( $p = 1/100$ ) that a random sample from the known population would be so scattered. Conversely, if in another sample of thirty subjects, we found none in the first nor in the 10th decile ( $q \leq 80$  per cent), the probability is  $99/100$  that this sample would be a biased one.

Other applications of this chart could readily be found in every case where the distribution function departs from usual forms or is unknown.

MARCO P. SCHÜTZENBERGER

97

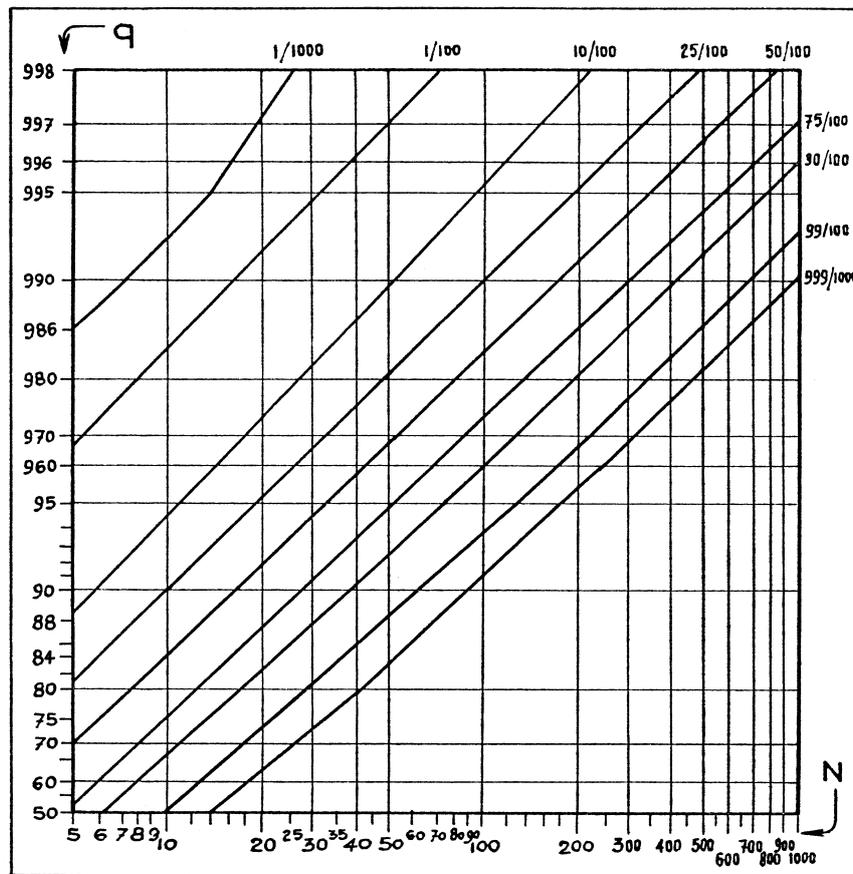


FIGURE 1

1008

ACADÉMIE DES SCIENCES.

ordonné *produit restreint* des ensembles ordonnés  $E_1$  et  $E_2$ ,  $\Omega$  désignant l'ordre produit de  $\Omega_1$  et  $\Omega_2$ .

Soit  $\mathcal{F}$  l'ensemble des parties non vides  $X$  de  $E^*$  telles que

$$X = \Omega^*(X) \quad (*)$$

$$(x, y_i) \in X \quad \text{quel que soit } i \in I \quad \text{et} \quad \bigvee_{i \in I} y_i \text{ existe} \rightarrow \left(x, \bigvee_{i \in I} y_i\right) \in X$$

$$(x_j, y) \in X \quad \text{quel que soit } j \in J \quad \text{et} \quad \bigvee_{j \in J} x_j \text{ existe} \rightarrow \left(\bigvee_{j \in J} x_j, y\right) \in X.$$

La relation d'ordre produit restreint  $\Omega^*$  peut être immergée dans la relation d'inclusion  $\mathcal{J}_{\mathcal{F}} \subset \mathcal{F} \times \mathcal{F}$ . Nous appellerons ce prolongement de  $\Omega^*$  l'*ordre produit tensoriel* (restreint) de  $\Omega_1$  et  $\Omega_2$ .

On démontre alors que, lorsqu'on suppose que  $\Omega_1$  et  $\Omega_2$  sont des relations d'ordre latticiel complet, le quotient de  $\Omega^*$  par les lois distributives redonne la définition du produit tensoriel de deux lattices complètes étendu au cas infini telle qu'elle a été énoncée dans les Notes citées en (\*).

ALGÈBRE. — *Sur certaines applications remarquables des treillis dans eux-mêmes.* Note (\*) de M. MARCEL-PAUL SCHÜTZENBERGER, présentée par M. Gaston Julia.

Étant donné un treillis quelconque complet  $\mathcal{L}$ , on appellera *S-application croissante* une application  $s$  de  $\mathcal{L}$  dans lui-même telle que :

I. Pour tout  $x$ ,  $x \leq s(x)$ .

II. Pour tout  $x$  et  $y$ , si  $x \leq y$ , alors,  $s(x) \leq s(y)$ .

III. Pour tout  $x$  et  $y$ ,  $s(x)s(y) = s(xy)$ ; de I et III on déduit sans peine :

III'. Pour tout  $x$  et  $y$ ,  $s(x) + s(y) \leq s(x + y)$ .

Manifestement, les itérées  $s^n$  de  $s$  sont aussi des S-applications. D'après III, si  $u$  et  $u'$  sont tels que  $x \leq s(u)$  et  $x \leq s(u')$ , alors,  $x \leq s(uu')$ , donc,  $\mathcal{L}$  étant complet, l'intersection  $t(x)$  de tous les  $u$  tels que  $x \leq s(u)$  définit une application  $t$  de  $\mathcal{L}$  dans lui-même telle que,  $ts(x) \leq x \leq st(x)$  et l'on montre que :

I. Pour tout  $x$ ,  $t(x) \leq x$ .

II. Pour tout  $x$  et  $y$ , si  $x \leq y$ , alors  $t(x) \leq t(y)$ , car

$$x \leq xy \leq st(x)st(y) = s[t(x)t(y)],$$

donc  $t(x) \leq t(x)t(y)$ , donc II.

(\*) En posant  $\Omega[X] = \bigcap_{x \in X} \Omega(x)$ ,  $\Omega(X) = \bigcup_{x \in X} \Omega(x)$ .

(\*) Séance du 3 novembre 1948.

SÉANCE DU 15 NOVEMBRE 1948.

1009

III. Pour tout  $x$  et  $y$ ,  $t(x) + t(y) = t(x + y)$ , car, d'après III',

$$x + y \leq st(x) + st(y) \leq s[t(x) + t(y)],$$

donc  $t(x + y) \leq t(x) + t(y)$ , mais, d'après II,  $t(x) \leq t(x + y)$  et  $t(y) \leq t(x + y)$ , donc III,  $t$  sera appelée *S-application décroissante adverse* de  $s$  et l'on montre que :

III'. Pour tout  $x$  et  $y$ ,  $t(xy) \leq t(x)t(y)$ .

IV et IV'. Pour tout  $x$ ,  $s^n t^n(x) = s^n(x)$ ,  $t^n s^n(x) = t^n(x)$ , car  $u \leq s^n t^n(u)$  et  $t^n s^n(x) \leq x$ .

Par leur composition, une S-application  $s$  de  $\mathcal{L}$  et son adverse  $t$  engendrent un *monoïde associatif*  $\mathcal{M}(\mathcal{S}, \mathcal{L})$  et, d'après IV et IV', on montre qu'il existe une représentation de tout élément de  $\mathcal{M}(\mathcal{S}, \mathcal{L})$  par  $s^{\alpha_1} t^{\beta_1} s^{\alpha_2} t^{\beta_2} \dots s^{\alpha_i} t^{\beta_i} s^{\alpha_{i+1}} \dots s^{\alpha_j} t^{\beta_j}$ , où la suite  $\alpha_1, \beta_1, \alpha_2, \beta_2, \dots, \alpha_i, \beta_i, \alpha_{i+1}, \beta_{i+1}, \dots, \alpha_j, \beta_j$  est *unimodale*, c'est-à-dire telle que, pour un certain  $i$ ,

$$\begin{array}{ccccccc} \alpha_1 & \leq & \beta_1 & \leq & \alpha_2 & \leq & \dots & \leq & \alpha_i & \leq & \beta_i \\ \beta_j & \leq & \alpha_j & \leq & \beta_{j-1} & \leq & \dots & \leq & \alpha_{i+1}. \end{array}$$

A toute S-application on peut associer une *S-relation*, écrite  $\leq$  et définie par :  $x \leq y$  équivalent à  $x \leq y \leq s(x)$ ; équivalent à  $t(y) \leq x \leq y$ .

On montre que :

5. Pour tout  $x, y$  et  $z$ , si  $x \leq y$ , alors  $x + z \leq y + z$  et  $xz \leq yz$  car, d'après II,  $s(x) \leq s(x + y)$ , donc si  $x \leq y \leq s(x)$ , alors  $x + z \leq x + y + z \leq s(x + z)$ .

Réciproquement, étant donnée une relation  $\leq$  satisfaisant à :

1. Pour tout  $x$ ,  $x \leq x$ .

2. Pour tout  $x$  et  $y$ , si  $x \leq y$ , alors  $x \leq y$ .

31. Pour tout  $x, y$  et  $z$  si  $x \leq y \leq z$  et  $x \leq z$ , alors  $x \leq y$ .

32. Pour tout  $x, y$  et  $z$  si  $x \leq y \leq z$  et  $y \leq z$ , alors  $y \leq z$ .

Cette relation définit deux applications adverses l'une de l'autre et en outre :

4. Pour tout  $x, y$  et  $z$ , si  $x \leq y$  et  $x \leq z$ , alors  $x \leq y + z$ .

4. Pour tout  $x, y$  et  $z$ , si  $y \leq x$  et  $z \leq x$ , alors  $yz \leq x$ .

Et ces applications satisfont à III et III' si l'une des deux conditions 5 est satisfaite.

Pour un treillis  $\mathcal{L}$ , l'ensemble des S-applications (ou des S-relations), ordonné par la relation :

$S \prec S'$ , équivalent à : pour tout  $x$ ,  $s(x) \leq s'(x)$ ;

Équivalent à : pour tout  $x$ ,  $t'(x) \leq t(x)$ ;

Équivalent à : pour tout  $x$  et  $y$ , si  $xy \leq$  alors  $x \leq y'$  constituent un *treillis*  $\mathfrak{S}(\mathcal{L})$  dont le plus petit élément  $S_0$  est :  $t_0(x) = s_0(x) = x$ ;  $x \leq y$  équivalent à  $x = y$ .

Et le plus grand élément  $S_1$  est :  $s_1(x) = 1$ ,  $t_1(x) = 0$ ;  $x \leq y$  équivalent à  $x \leq y$ .

On désignera par  $\sigma$  et  $\tau$  la plus petite S-application telle que pour tout  $x$  et  $y$  : si  $x \leq y$ , alors  $x \leq (x + z)\sigma(x)$ .

1948-11. Sur certaines applications remarquables des treillis dans... Année 1948

1010

ACADÉMIE DES SCIENCES.

Si  $\mathcal{L}$  est modulaire,  $\Sigma$  correspond à la relation :  $x \leq y$  équivalent : à  $x|y$  est complémenté et le treillis quotient  $S_0|\Sigma$  est un treillis distributif complémenté de puissance  $2^N$  où  $N$  est le nombre de paramètres indépendants dont on dispose pour normer  $\mathcal{L}$ . Si  $\mathcal{L}$  est non seulement modulaire mais distributif,  $\mathfrak{C}(\mathcal{L})$  est aussi distributif. Enfin, si  $\mathcal{L}$  est complémenté,  $\mathfrak{C}(\mathcal{L})$  est isomorphe à  $\mathcal{L}$ .

ALGÈBRE. — Une méthode pour la décomposition spectrale et l'inversion des matrices. Note de M. JEAN-MARIE SOURIAU, présentée par M. Jean Chazy.

Soit  $A$  une matrice carrée d'ordre  $n$ . Nous désignerons par  $T(A)$  la trace de cette matrice (c'est-à-dire la somme des éléments de la diagonale principale), et par

$$P(x) = x^n + k_1 x^{n-1} + \dots + k_n$$

le polynôme caractéristique de cette matrice.

Définissons comme suit les matrices  $B_i$  :

$$B_0 = I \text{ (matrice unité d'ordre } n),$$

$$B_1 = A - T(A),$$

$$B_2 = AB_1 - \frac{1}{2} T(AB_1),$$

$$B_3 = AB_2 - \frac{1}{3} T(AB_2),$$

.....

$$B_n = AB_{n-1} - \frac{1}{n} T(AB_{n-1}).$$

Nous avons démontré les résultats suivants :

1° Les quantités  $-T(A)$ ,  $-(1/2)T(AB_1)$ , ...,  $-(1/n)T(AB_{n-1})$  sont les coefficients  $k_1, k_2, \dots, k_n$  du polynôme caractéristique. En particulier, le déterminant de  $A$  vaut  $[( - 1)^{n-1}/n] T(AB_{n-1})$ .

2° La matrice  $B_n$  est nulle, la matrice  $(-1)^{n-1} B_{n-1}$  est la transposée de la matrice adjointe de  $A$ , et, si  $A$  possède une inverse, on a

$$A^{-1} = -\frac{1}{k_n} B_{n-1}.$$

3° En posant

$$Q(x) = x^{n-1} B_0 + x^{n-2} B_1 + \dots + B_{n-1},$$

on a

$$[x - A] Q(x) = P(x),$$

$$T(Q(x)) = P'(x),$$

et, si  $x_i$  est une racine simple de  $P(x)$ , la matrice  $Q(x_i)$ , qui n'est pas nulle, est le covariant de Frobenius attaché à la valeur propre  $x_i$  (c'est-à-dire le produit du mode à droite et du mode à gauche correspondants).

Année 1948      1948-12. Contribution à l'étude statistique du sexe à la naissance

---

FACULTE DE MEDECINE DE PARIS  
=====

THESE

pour le

DOCTORAT EN MEDECINE

(Diplome d'Etat)

par

Marcel Paul SCHUTZENBERGER

Né le 24.10.1920 à Paris

présentée et soutenue publiquement le 2 Juillet 1948

CONTRIBUTION A L'ETUDE STATISTIQUE  
\*\*\*\*\*

DU SEXE A LA NAISSANCE  
\*\*\*\*\*

Président : Monsieur Raymond TURPIN, Professeur.

A

Mon Président de Thèse

Monsieur le Professeur R. TURPIN

Professeur de thérapeutique à la Faculté de Médecine de Paris

Médecin de l'Hôpital Saint Louis

Directeur du Centre de génétique de l'Hôpital Saint Louis

avec l'expression de ma reconnaissance  
pour le grand honneur qu'il a bien voulu me  
faire en me confiant le sujet de cette thèse,  
développement de ses travaux de génétique, en  
me guidant constamment de ses conseils au cours  
de sa réalisation et en acceptant d'en présider  
la soutenance.

A M E S M A I T R E S  
= =====

DANS LES HOPITAUX

Monsieur le Professeur P.GUILLAUME-LOUIS, Directeur de l'Ecole de Médecine de Tours (I.&L.) 1938/1941  
Monsieur le Docteur D.GAUDEAU, Professeur à l'Ecole de Médecine de Tours - 1939/1940  
Monsieur le Docteur F.CHEVE, Professeur à l'Ecole de Médecine de Tours - 1940/1941  
Monsieur le Docteur A.TZANCK, Médecin Chef du Centre de transfusion sanguine replié à Blois (I.&C.) 1939/1940  
Monsieur le Docteur J.TUSQUES, Médecin Chef de service à l'Hôpital Psychiatrique de Blois -1941/1942  
Monsieur le Docteur A.CARBILLET, Médecin-Directeur de l'Hôpital Psychiatrique d'Yzeure- 1943/1945  
Monsieur le Docteur E.SIDI, Médecin Assistant à l'Hôpital Saint Louis.

AU COLLEGE DE FRANCE

Monsieur le Professeur P.LANGEVIN, Membre de l'Institut de France, Directeur de l'Ecole de physique et de chimie industrielle de la Ville de Paris (in memoriam).

A LA FACULTE DES SCIENCES DE PARIS

Monsieur le Professeur A.CHATELET, Ancien Recteur de l'Université de Lille, Professeur d'arithmétique supérieure  
Monsieur le Professeur M.FRECHET, Professeur de calcul des probabilités et de statistique

AU CENTRE D'ETUDES DU BOUCHET

Monsieur P.GAUDAUDAN, Chef du service de biologie cellulaire des services chimiques de l'Etat.

I N T R O D U C T I O N  
=====

Depuis l'antiquité la plus reculée, le problème de la détermination du sexe chez l'animal et chez l'homme a suscité un intérêt considérable, en raison de son importance économique, sociale et politique.

De nombreuses techniques et théories ont été imaginées et il suffirait de citer les aphorismes de la médecine hippocratique sur l'influence masculinisante du côté droit et féminisante du côté gauche, les traditions religieuses des diverses civilisations, les idées sur l'influence de certains régimes alimentaires, les opinions sur le rôle de la configuration astronomique au moment de la conception.

A l'époque moderne, l'utilisation de statistiques de plus en plus rigoureuses et étendues, l'emploi de la méthode expérimentale, le progrès des sciences biologiques ont permis de dégager un certain nombre de faits indiscutables et de fournir des interprétations satisfaisantes de la plupart d'entre eux.

°°°

Le présent travail, exclusivement statistique, a un objectif plus limité; se bornant aux familles humaines, on s'efforcera de prouver la réalité d'une "corrélation" entre naissances immédiatement consécutives, sans préjudice d'ailleurs de l'existence d'autres effets.

Faisant suite à la présentation par Monsieur le Doyen L.BINET, à l'Académie des Sciences d'une note (1) de Monsieur le Professeur R.TURPIN et moi-même, cette thèse apporte-ce qui n'avait pu y figurer- :

- le détail des éléments numériques,
- la discussion des sources,
- le principe des modes de calcul,
- l'analyse statistique des principaux résultats,
- la construction et le test d'un schéma interprétatif.

En attirant ainsi l'attention sur un phénomène qui semble ne pas avoir été remarqué par d'autres Auteurs, on espère apporter un élément nouveau dont toute théorie se devra de tenir compte, si des recherches ultérieures, reposant sur d'autres sources documentaires, en confirment la généralité et la permanence.

°  
° °

Nous tenons à remercier ici tous ceux qui nous ont permis de mener à bien le présent travail, en nous permettant les indispensables recherches dans les archives et le dépouillement des dossiers; en particulier Monsieur KERGALL, Directeur du service de la "Médaille d'Honneur aux Mères" au Ministère de la Santé publique et de la Population, à Messieurs les Présidents des Associations de familles nombreuses, à Messieurs les Membres du personnel administratif de la Fondation COGNACQ-JAY.

---

(1) - Recherches statistiques sur la distribution du sexe à la naissance par M.L.R.TURPIN & M.P. SCHÜTZENBERGER à l'Académie des Sciences- Séance du 1<sup>er</sup> juin 1948.

Nous remercions également pour une collaboration  
bénévole, qui nous fut précieuse, Mesdames COUIN, POUSOL et  
DUBROCHÉ du Centre de Génétique de l'Hôpital Saint Louis.

Nous exprimons enfin, toute notre gratitude à Monsieur  
le Docteur Raymond TURPIN, Professeur de thérapeutique à la Faculté  
de Médecine de PARIS, Directeur du Centre de Génétique à l'Hôpital  
Saint Louis. Il nous a accueilli dans son service hospitalier de  
l'Hôpital Saint Louis, il nous a confié ce travail et nous a aidé  
constamment dans sa réalisation; qu'il soit également remercié  
pour le très grand honneur qu'il nous a fait en acceptant de prési-  
der cette thèse.

BUT , METHODE e t S O U R C E S  
-----

BUT  
---

Avant de commencer l'étude de l'enquête statistique actuelle, il peut être bon de rappeler certains points importants de la théorie de la détermination des sexes, en ce qui concerne l'espèce humaine.

Sex ratio : Depuis VILLANI (1300 environ), GRAUNT (1661) et SÜSSMILCH (1761) on sait que la quasi-égalité numérique des sexes à la naissance est un phénomène général et constant; plus précisément il est admis que le nombre de naissances masculines pour CENT féminines est d'environ 106; ce rapport constituant une "véritable caractéristique" de l'espèce humaine comme l'écrivait HALBWACHS.

Il n'est pas besoin d'insister sur le fait que ce taux de masculinité à la naissance n'est pas identique à celui qui s'observerait à la conception, non plus qu'à celui de la population effective, en raison de l'inégale mortalité des individus des deux sexes aux différents âges de la vie (cf.3 et 7.)

D'autre part, cette constance de la sex ratio doit être complétée par la constatation de sa variation et des limites de celles-ci selon les pays, les époques et les milieux; par exemple, de manière significative, la sex ratio en Angleterre a évolué de 1910 à 1919 entre 0,510 et 0,515.

On ne discutera pas ici, car elles sortent du cadre de cette étude, les causes les plus fréquemment invoquées pour rendre

compte de ces fluctuations (influence du changement de régime alimentaire, âge moyen des parents .. etc .. etc.....)

On a aussi avancé que les primipares donnaient plus fréquemment naissance à des garçons et que cette augmentation de la sex ratio se rencontrait également chez les mères âgées; cette hypothèse est encore en suspend et des chiffres contradictoires ont été publiés, (cf.12 et 1.).

Enfin, un fait très important a été établi grâce aux matériaux recueillis par GEISSLER (cf.6.); éliminant les naissances gemellaires, la théorie classique voudrait que la répartition par nombre des garçons des familles de  $n$  enfants se fit selon une loi binomiale.

Or, si ce modèle probabiliste est très suffisant dans la pratique, il est possible de prouver en utilisant une base assez large de documentation, qu'il est théoriquement inacceptable.

Notamment FISCHER (cf.4.) a montré que les variances des histogrammes relatifs aux familles de  $n$  enfants sont significativement plus grandes que leur valeur théorique dans l'hypothèse d'une répartition binomiale; ce qui revient à dire que les familles extrêmes ( c'est à dire celles où l'on observe un fort excès de garçons ou de filles) sont proportionnellement plus nombreuses que dans le cas de cette répartition.

Confirmant cette observation et offrant une possibilité d'interprétation, il faudrait pour être complet, exposer les théories selon lesquelles, chaque couple serait prédisposé à engendrer préférentiellement des garçons ou des filles selon certaines particularités de son génotype.(cf.10).

C'est sur la base de ces différents travaux que l'on

discutera les résultats obtenus au cours de cette enquête; c'est à dire successivement:

- 1°- la sex ratio et sa variation
- 2°- la répartition du sexe en fonctions du rang des naissances.
- 3°- l'excès des familles extrêmes
- 4°- la liaison entre les sexes à la naissance dans une même famille.

A chaque stade de l'étude, on sera amené à formuler une hypothèse probabiliste, permettant de tester l'hypothèse physiologique sous jacente.

Finalement, on construira un modèle statistique, dont on évaluera les paramètres et dont on montrera l'accord avec les faits; cependant qu'on donnera des indications sommaires sur une méthode plus complexe d'analyse de semblables documents.

#### METHODE

Etant donné le but poursuivi dans ce travail, les dispositions suivantes ont été prises pour la collecte des documents.

- 1°- Il a semblé nécessaire de se borner aux familles "assez" nombreuses pour que l'on puisse penser avoir écarté au maximum l'influence possible d'une restriction volontaire à la naissance; dans ce sens, il est évident que les familles les plus nombreuses (plus de 8 enfants) apportaient les meilleures garanties; mais la décroissance rapide de leur nombre en fonction du nombre des enfants (coefficient d'environ 0,55 par enfant supplémentaire pour une population non sélectionnée) a nécessité l'abaissement de cette borne inférieure.

Finalement a prévalu une considération d'ordre pratique : la

fixation administrative du chiffre de CINQ Enfants pour la définition de la "famille nombreuse" facilitant la collecte des documents basés sur cette valeur; c'est pourquoi, en définitive, l'enquête a été limitée aux familles de cinq enfants et plus.

2°- Est-il utile de mentionner que n'ont été prises en considération que les familles pour lesquelles les progéniteurs sont censés être les mêmes pour toutes les naissances; "pater est quem nuptiae demonstrant".

3°- On a éliminé toutes les familles comprenant des grossesses multiples; tant en raison de l'impossibilité d'en fixer l'uni ou la pluri-vitellinité (précision pourtant essentielle) qu'en considération de l'alourdissement et de la complexité dans le dépouillement et le traitement mathématique.

4°- On n'a conservé que les familles pour lesquelles était connu le sexe de tous les enfants nés viables, laissant de côté celles où cette condition n'était pas remplie. Une confiance plus grande, à ce point de vue, doit être accordée au stock 1 (dossiers recueillis au Ministère de la Santé Publique et de la Population) dont les documents de base contiennent systématiquement tous les enfants vivants et morts; il n'a pas été toujours possible d'avoir ce renseignement dans les autres stocks (2: région parisienne et 3: grandes familles du Nord) et surtout, hélas, dans le stock 4 (Fondation COGNACQ-JAY).

5°- J'ai dû collecter moi-même les matériaux provenant des stocks 2, 3, 4 et 4, ainsi que la moitié de ceux issus du stock 1; l'autre moitié de celui-ci a été dépouillée par des collaborateurs bénévoles (qui ignoraient à dessein tous les buts

de cette enquête).

Ce point est important en raison de la part de l'influence personnelle de l'observateur sur le choix des documents; il les récolte très souvent (inconsciemment d'ailleurs) avec partialité en faveur de ses thèses, cette remarque faite, il faut indiquer qu'aucune différence significative entre les documents issus des deux moitiés du stock 1, non plus qu'entre les parties de la moitié du stock 1 étudiées par divers collaborateurs n'a été constatée en ce qui concerne les points principaux de cette enquête.

6°- Dans le même esprit on s'est interdit, et également afin d'éviter des duplications possibles, d'introduire dans cette étude des familles issues de source autre que systématique telles que: familles connues personnellement ou citées dans des ouvrages, travaux ..etc.. etc.

7°- Une partie importante (et difficilement déterminable) des familles étudiées sont des familles "en cours" c'est-à-dire qu'une famille de cinq enfants, par exemple, ne représente ici seulement qu'une famille ayant cinq enfants au moment de l'enquête (1<sup>er</sup> mars 1948) sans que l'on puisse préjuger si le cinquième sera le dernier enfant ou sera suivi de frères ou de soeurs.

Étant données les sources utilisées, et les conditions matérielles de réalisation il n'a pas été possible d'enregistrer les dates de naissance des enfants successifs, non plus qu'aucune particularité concernant les parents: âge, statut social et économique, habitat, état de santé .. etc ..

Ainsi les données recueillies sont le nombre de famille, de chaque type caractérisé par la succession des naissances masculines ou féminines.

Par exemple:

On a observé 162 familles de 5 enfants dont 2 garçons du type: fille-garçon-fille-fille-garçon- (f.g.f.f.g.) et deux familles de 10 enfants du type : fille-garçon-garçon-garçon-garçon-garçon-fille-garçon-fille-garçon. (f.g.g.g.g.g.f.g.f.g.).

LES SOURCES

STOCK .1: On a utilisé les dossiers constitués au Ministère de la Santé publique et de la Population en vue de l'attribution de la Médaille d'Honneur des familles (promotion 1946 et 1947).

Les irrégularités dans les fréquences relatives des familles de 5, 6, 7, 8, ..etc... enfants s'expliquent aisément du fait qu'il existe.

a- une Médaille de bronze à laquelle ont droit les mères de cinq enfants et plus.

b- une Médaille d'argent pour huit enfants et plus.

c- une Médaille d'or pour dix enfants et plus.

Des considérations diverses font que certaines mères reçoivent la Médaille dès la naissance du nombre minimum d'enfants requis alors que d'autres, ignorées jusque là par le service, sont décorées de la Médaille de bronze, par exemple, alors que leur vie obstétricale est largement achevée et qu'elles ont déjà dix enfants.

Dans l'ensemble; les dates de naissance s'étendent d'environ 1900 à 1945 et les 9.895 familles utilisées peuvent être considérées comme un échantillon représentatif de l'ensemble du territoire métropolitain français.

STOCK.2: Les familles de ce stock ont été extraites

des fichiers des associations de famille du département de la Seine; il a été possible d'enquêter dans la plupart des sièges d'arrondissement et d'étudier ainsi 2.912 familles pour lesquelles des remarques chronologiques, semblables à celles faites pour le stock 1; sont valables.

STOCK. 3: On a dépouillé les annuaires dits "Les grandes familles de Roubaix-Tourcoing" (dernière édition 1947) et "Les grandes familles de Lille-Armentières-Cambrai-Douai-Valenciennes (édition 1931).

Un soin tout particulier a été apporté pour éviter qu'une même famille ne soit enregistrée en double et l'on s'est volontairement borné à la dernière génération figurant dans l'annuaire.

STOCK.4: On a consulté les dossiers des années 1920, 1921, 1922, du Grand Prix COGNACQ-JAY (réservé aux familles de plus de dix enfants) et rassemblé ainsi 147 familles particulièrement intéressantes, en raison du très grand nombre de leurs enfants.

Malheureusement, il semble que les candidats aient souvent tenu à honneur de minimiser le nombre de leurs enfants décédés en bas âge; en raison du travail de vérification supplémentaire qui se trouvait imposé de ce fait, on a dû se limiter à ce nombre très restreint de familles.

Remarques générales sur l'échantillonnage.

Comme on vient de le montrer, l'échantillonnage utilisé ne peut pas, en toute rigueur, prétendre à être représentatif de la composition actuelle de la population française; cependant il semble suffisamment adapté au but poursuivi; l'étude pure-

15

ment physiologique des naissances dans les familles nombreuses.

D'autre part, divers tests, sur lesquels on ne s'étendra pas, ont permis de constater l'homogénéité des différentes sources et autorisent ainsi à traiter en bloc les 14.250 familles sans se livrer à des ventilations et sous-ventilations qui alourdiraient les calculs et masqueraient les résultats.

Année 1948 1948-12. Contribution à l'étude statistique du sexe à la naissance

TABLEAU N° 1  
Origine des documents

Nombre d'enfants par famille	Ministère de la Santé	Région parisienne	Grandes familles du Nord	Fondation COGNACQ- JAY	TOTAL
5	3.309	1.387	432		5.128
6	2.214	740	301		3.255
7	1.298	374	205		1.877
8	1.269	211	139		1.119
9	662	109	81		872
10	527	48	59		634
11	313	27	34	10	384
12	159	12	15	32	218
13	68	3	6	31	108
14	26	3	1	29	59
15	14		2	21	37
16	9		1	12	22
17	5			6	11
18				2	2
19				4	4
<b>Total des familles</b>	<b>9.693</b>	<b>2.914</b>	<b>1.276</b>	<b>147</b>	<b>14.230</b>

1948-12. Contribution à l'étude statistique du sexe à la naissance

Année 1948

FABRIAU N° 2  
Répartition numérique des familles  
selon le nombre total de leurs enfants et le nombre de garçons

Nombre d' enfants par famille	Nombre de garçons														Nombre total des familles	
	0	1	2	3	4	5	6	7	8	9	10	11	12	13		14
5	144	810	1544	1585	841	204										5.128
6	54	325	717	979	785	330	65									3.255
7	21	111	287	480	529	304	126	19								1.877
8	8	61	180	327	412	406	170	49	6							1.619
9		20	72	155	204	184	140	74	22	1						872
10		5	23	64	120	165	128	78	39	11	1					634
11		3	15	27	51	85	92	61	37	9	3	1				384
12		4	3	18	19	32	50	49	23	16	2	1	1			216
13			2	1	3	9	12	24	19	14	9	13	2			108
14					2	6	9	14	11	8	6	1				59
15					2	2	4	6	8	5	5	1	4			37
16					2	3	3	4	3	4	1		2			22
17						1			1	3	2	3	1			11
18													1			2
19										1	1	1		1	1	4
																<u>14.230</u>

S E X     R A T I O  
=====

Calculée sur les 95.875 naissances, la probabilité, à priori de naissances d'un garçon est de 0,5006 ± 0,0016, sensiblement plus faible que les valeurs trouvées habituellement.

Aucune autre interprétation plus satisfaisante n'ayant pu être donnée de ce fait, il paraît raisonnable de ne voir là qu'une fluctuation d'échantillonnage.

Pour comparer maintenant la sex ratio entre les familles d'importance différente, on a employé le test classique de  $\chi^2$  correspondant au coefficient de LEXIS et donné par  $\chi^2 = \sum N_n \frac{(p_n - \bar{p})^2}{\bar{p}}$  avec K-1 degré de liberté ou  $N_n$  est le nombre d'enfants dans les familles de n enfants;  $p_n$  la sex ratio correspondante et  $\bar{p}$  la valeur des moyennes des  $p_n$ ; on trouve ainsi  $\chi^2 = 26,7$  pour 14 degrés de liberté, ce qui est une valeur nettement significative (0,02 %).

Mais si l'on se borne aux familles de moins de 12 enfants le  $\chi^2$  n'est plus que de 7,7 et par conséquent n'est pas significatif.

Rapprochant ce fait de la non-augmentation de la sex ratio jusqu'à 12 enfants; il semble plausible de ne pas attribuer de signification physiologique à l'écart constaté entre ces 15 valeurs (voir tableau N° 3.).

Et il ne faudrait peut être y voir que l'expression des conditions économiques rendant plus fréquentes, parce que plus aisés, les développements de familles très nombreuses à prédo-

minance masculine dans les milieux ruraux, d'où provient précisément la majeure partie de ces familles exceptionnelles.

Année 1948 1948-12. Contribution à l'étude statistique du sexe à la naissance

TABLEAU N° 3

## Valeur caractéristique des histogrammes

Nombre d' enfants	Nombre de familles	Fréquence des naissances masculines	Variance observée	Ecart de la variance (1)
5	3.128	0,5085	1,300	+ 4,6
6	3.255	0,5057	1,595	+ 6,3
7	1.877	0,5061	1.892	+ 6,5
8	1.619	0,5011	2.069	+ 2,7
9	872	0,4980	2.636	+10,7
10	634	0,5178	2.583	+ 1,9
11	384	0,565	2.976	+ 3,8
12	218	0,5022	3.667	+ 8,0
13	108	0,5207	4,990	+ 2,9
14	59	0,5170	3.322	- 1,0
15	37	0,5441	4,695	+ 4,3
16	22	0,4682	5,110	+ 3,7
17	11	0,6112	3,072	= 0,8
18	2	0,5566	8,000	+ 3,5
19	4	0,6316	3,333	- 0,1

(1) Ecart en valeur observée et valeur théorique exprimée en écart type.

REPARTITION DES SEXES SELON  
 =====

LE RANG DE NAISSANCE.  
 ---

Deux hypothèses ont été avancées à ce sujet:

1°- La probabilité que le premier né soit un garçon est supérieur à la probabilité de naissance d'un garçon dans les grossesses ultérieures.

2°- La proportion des naissances masculines augmente chez les mères d'âge avancé.

Pour tester la première hypothèse, la méthode la plus simple et la plus efficace, est de comparer les fréquences; on trouve ainsi

Probabilité de naissance d'un garçon à la première grossesse  $7.253 / 14.230 = 5097 \pm 0,0042$ .

Probabilité de naissance d'un garçon dans les autres grossesses:  $\frac{41.242}{81.495} = 0,5062 \pm 0,0018$ .

La différence entre ces deux chiffres, étant de l'ordre de 1 fois son écart type ne peut être considérée comme significative et les documents recueillis n'apportent donc qu'une très faible confirmation de l'hypothèse.

Pour tester la deuxième hypothèse, on a employé une méthode qui semblera peut être un peu artificielle, mais qui permet une utilisation assez précise de l'information contenue dans les documents pour les familles de  $n$  enfants, on a considéré que, dans l'hypothèse nulle le nombre des garçons à chaque rang de naissances est réparti selon une loi rectangulaire discrète, autrement dit qu'à chaque rang devait se trouver sensiblement le même nombre de garçon

On introduira donc deux concepts commodes, qui sont les familles de  $n$  enfants.

1°-  $\bar{r}_n$  rang moyen des garçons.  $\bar{r}_n$  se trouvera déplacé vers l'une ou l'autre des extrémités selon que les garçons seront plus ou moins fréquents à la fin OU au début des familles.

$$\text{Sa valeur est estimée par } \hat{r}_n = \frac{\sum i f_i}{N_g}$$

$f_i$  = nombre des naissances masculines au  $i$  ième rang,  $N_g = \sum f_i$  = nombre total de garçons dans les familles de  $n$  enfants.

$\hat{r}_n$  sera considéré comme distribué normalement autour de sa valeur moyenne  $\frac{n-1}{2}$  avec une variance égale à  $1/12 (N_g - 1) / (n^2 - 1)$

2°-  $\sigma_p^2$  : variance des rangs qui se trouvera augmentée ou diminuée selon que les garçons sont ou non plus fréquents au début ET à la fin des familles.

$$\sigma^2 \text{ est estimée par } \left( \sum i f_i - N_g (\bar{r}_n) \right)^2 / (N_g - 1) - 1$$

et sera considéré comme distribué normalement autour de sa valeur moyenne :  $1/12 (n^2 - 1)$  avec une variance que pratiquement l'on prendra égale à  $1/80 (N_g)^{-1} (n^4 - 5n^2 + 4)$ .

Le tableau ci-joint N° 4 montre à simple examen qu'il n'y a que peu de preuves pour écarter l'hypothèse nulle d'une distribution rectangulaire.

Pour les  $\sigma^2$  où il pourrait y avoir doute, on a combiné les 15 tests par une méthode classique ( cf.8.).

L'on sait en effet que l'expression  $M^2 = -2 \sum \log_e p_i$  ou  $p_i$  est pour chaque test la probabilité d'observer un écart aussi grand avec la valeur théorique est distribué comme  $\chi^2$  avec  $2K$  degrés de liberté.

Ici  $M^2 = 39,7$  et  $2K = 30$ ; l'on ne peut donc écarter

20

l'hypothèse nulle (probabilité comprise entre 0,15 et 0,10.).

Par la suite, on considèrera donc comme uniforme la probabilité à priori de naissance d'un garçon à un rang quelconque.

Année 1948 1948-12. Contribution à l'étude statistique du sexe à la naissance

TABLEAU N° 4

Valeurs caractéristiques des distributions rectangulaires

Nombre d' enfants	Rangs moyens	Ecart des rangs moyens	Variance des rangs	Ecart d' varian (1)
5	2.015	+ 1,2	2.072	+ 1
6	2.602	+ 1,2	2.502	+ 1,5
7	3.045	+ 1,9	3.045	+ 1,7
8	3.534	+ 1,2	5.320	+ 1,2
9	4.004	+ 5,1	6.727	+ 0,6
10	4.416	- 1,6	8.250	0,0
11	4.966	- 0,5	10.000	0,0
12	5.445	- 0,6	11.694	- 0,1
13	6.026	+ 0,2	14.253	+ 0,5
14	6.312	- 1,0	16.356	- 0,1
15	6.993	0,0	17.975	- 0,7
16	7.550	+ 0,4	22.834	+ 0,4
17	7.467	- 0,1	26.175	+ 1,0
18	9.050	+ 0,5	23.438	-0, 8
19	8.625	- 0,5	33.771	+ 0,2

(1) Ecart entre valeur observée et valeur théorique exprime en écart type.

V A R I A N C E S     D E S     R E P A R T I T I O N S  
 =====

Soit la répartition de  $N_n$  familles de  $n$  enfants en familles de  $g$  garçons et  $n-g$  filles.

C'est une observation déjà faite par FISCHER sur les matériaux de GEISLER que les familles à forte prédominance masculine ou féminine sont sensiblement plus nombreuses que ne le voudrait une répartition binomiale.

Opérant sur un matériel, d'où les naissances gémellaires ont été éliminées, il est possible, ici, de préciser ce point en comparant la valeur observée de la dispersion de ces distributions à la valeur théorique dans le cas d'une répartition binomiale qui constitue l'hypothèse nulle.

Dans ce cas, on devrait avoir les familles de  $n$  enfants  $f_g =$  nombre de familles de  $g$  garçons  $= \binom{n}{g} p^g q^{n-g} N_n$  où  $p = \frac{\sum g f_g}{N n}$  est la probabilité a priori de naissance d'un garçon.

La variance est estimée par la formule  $\frac{1}{N n} (\sum g^2 f_g - N n p^2)$  et  $N$  étant grand, il sera légitime de considérer qu'elle est distribuée comme une variable normale de moyenne  $n p q$  et de variance  $\frac{1}{n^2} = n^{-1} \mu_4 + \frac{3-n}{n(n-1)} \mu_2^2$  où  $\mu^2$  et  $\mu^4$  sont respectivement les moments à la moyenne du 2<sup>ème</sup> et du 4<sup>ème</sup> ordre de la répartition binomiale considérée; pratiquement on a pris ici pour valeur de  $1/n^2$  l'expression  $1/N \cdot 1/8 (N-1) (n-1)^{-1} (n^2 - n + 1)$

Or, comme le montre le tableau 3 dans 12 cas sur 15, la variance observée est supérieure à sa valeur moyenne et dans 10 cas sur 15 cette différence est significative à plus de 1 %.

Année 1948      1948-12. Contribution à l'étude statistique du sexe à la naissance

23

Il n'est pas besoin d'autre test combinant les résultats des quinze valeurs observées pour pouvoir établir que, d'une manière fort appréciable, les histogrammes sont plus étalés que ne le permettrait l'hypothèse d'une répartition binomiale, qui doit donc être rejetée.

S E Q U E N C E S  
=====

Si, comme dans le travail de GEISLER, l'information dont on dispose se trouvait limitée aux histogrammes, il serait difficile d'aller plus loin et de nombreuses possibilités (entre lesquelles on ne saurait trancher) s'offriraient pour expliquer l'étalement des distributions; notamment en supposant que la probabilité a priori  $p$  de naissance d'un garçon constante dans une même famille est répartie d'une façon aléatoire dans la population, on aurait les fréquences suivantes

$$f \ g = E ( p \ \xi \ q \ n - \xi ) \binom{n}{\xi} X \ N_n$$

où  $E(x)$  = valeur moyenne de  $x$ ; et l'on pourrait estimer la variance de  $p$  dans la population par l'expression

$$\sigma_p^2 = \binom{n}{2}^{-1} \sum \binom{y-2}{2} S_y$$

toujours positive quand, et seulement quand, la variance de l'histogramme des familles de  $n$  enfants est supérieure à sa valeur théorique.

Mais, dans l'enquête présente, l'information est plus complète et comprend à l'intérieur du total des familles de même composition globale de  $g$  garçons et de  $f$  filles, la répartition entre les différents arrangements possibles.

On peut donc tester l'hypothèse nulle  $H_1$  selon laquelle tous les arrangements auraient la même fréquence relative d'apparition pour des valeurs données de  $g$  et de  $f$ .

Ayant déjà examiné dans la seconde partie la possibilité que les garçons ou les filles se rencontrent préférentiellement à des moments privilégiés de la vie obstétricale (premier enfant,

naissances tardives) et n'ayant constaté aucune déviation significative, on est amené à tester un éventuel groupement des naissances de même sexe.

Pour tester  $H_1$  contre cette nouvelle hypothèse, il serait facile d'imaginer des "coefficients d'accumulation" plus ou moins complexes, faisant appel à telle ou telle propriété des suites de tirage aléatoire ou de permutations.

Mais d'abord, de telles expressions seraient surtout valables pour des suites assez longues (soit par exemple 10 à 20 Enfants) et l'on a affaire ici à des suites beaucoup plus courtes (10% seulement des naissances ont été observées dans des familles de onze enfants et plus). Ensuite, et surtout, on risquerait de tomber dans l'erreur fréquente qui consiste à forger des concepts statistiques étrangers au contenu concret du mécanisme aléatoire qu'ils doivent servir à étudier.

On a donc utilisé pour tester  $h_1$  le nombre des "coupures" c'est-à-dire des événements définis par la succession dans une même famille d'une naissance masculine à une naissance féminine la précédant immédiatement ou inversement.

A ce concept se rattache directement le concept de "séquence masculine" (respectivement féminine) c'est-à-dire de suites de naissances masculines (respectivement féminines) immédiatement consécutives.

Pour une famille de  $g$  garçons de  $f$  filles, on a immédiatement

$$S_m = S_f + A + B - 1$$

où  $S_m$  désigne le nombre des séquences masculines,

$S_f$  désigne le nombre des séquences féminines,

et  $A$  et  $B$  sont des indicatrices telles que  $A = 1$  si le premier enfant est un garçon,  $A = 0$  si le premier enfant est une fille,

$B = 1$  si le dernier enfant est un garçon et  $B = 0$  si le dernier enfant est une fille, et

$C = S_m + S_f - 1$ ; où  $C$  désigne le nombre de coupures; par exemple dans la famille du type : g.f.f.g.g.f.f.f. (3 garçons et 5 filles) on a  $g = 3$ ;  $f = 5$ ;  $B = 0$ ;  $S_g = 2$ ;  $S_f = 2$ ;  $C = 3$ ; et dans la famille du type: f.g.f.g.g.f.f.g.f. (5 garçons et 6 filles) on a  $g = 5$ ;  $f = 6$ ;  $A = 0$ ;  $B = 0$ ;  $S_g = 4$ ;  $S_f = 5$ ;  $C = 2$ .

Il est évident que  $C$  répond assez exactement à ce que l'on peut entendre plus ou moins intuitivement par "groupage des naissances de même sexe."

L'emploi de la notion de "coupures" (ou "séquences") se recommande par son étroite liaison avec l'estimation de la probabilité conditionnelle d'une naissance masculine quand est connu le sexe de l'enfant précédent; notion elle-même interprétable d'une manière concrète sur le plan de la physiologie comme on le montrera plus loin.

Les diverses répartitions liées aux  $S_g$ ,  $S_f$ , et  $C$ , ont été étudiées théoriquement par STEVENS (c.f. 9) - (1); et on a complété son travail en calculant l'expression formelle des coefficients de corrélation entre  $S_g$  et  $S_f$ , ce qui a permis d'obtenir la variance de  $C$ .

On a ainsi pour une famille de  $g$  garçons et  $f$  filles:

$$\text{Prob} ( S_g = u ) = \frac{\binom{g-1}{u-1} \binom{f+1}{u}}{\binom{g+f}{g}} - 1$$

valeur moyenne de  $S_g = \frac{(g)(f+1)}{(g+f)}$

(1) Il conviendrait, en outre, de citer les travaux (6 et 12) dont je n'ai pu avoir directement connaissance.

$$\text{variance de } S_g = \frac{(g-1)(f+1)f_g}{n(n-1)}$$

$$\text{corrélation entre } S_g \text{ et } S_f = \sqrt{\frac{g-1}{g+1} \times \frac{f-1}{f+1}}$$

Pour C, la distribution est un peu plus compliquée; on a :

$$\text{Prob}(C = 2w) = \binom{g-1}{w} \binom{f-1}{w-1} + \binom{g-1}{w-1} \binom{f+1}{w} \binom{f+g}{g} - 1$$

$$\text{Prob}(C = 2w+1) = 2 \binom{g-1}{w} \binom{f-1}{w} \binom{f+g}{g} - 1$$

$$\text{valeur moyenne de } C = \frac{2fg}{f+g}$$

$$\text{Variance de } C = \frac{2gfg(2gfg - g-f)}{(g+f)^2(g+f-1)}$$

La distribution des S n'est qu'une forme de la distribution hypergéométrique; quand à celle de C, elle est un peu moins maniable.

On peut donner quelques exemples de ces répartitions; ainsi pour les familles de trois garçons et de trois filles on a pour distribution de C:

$$\begin{array}{l} \text{Valeur de C} \quad 1 \quad 2 \quad 3 \quad 4 \quad 5 \\ \text{Prob C} \quad : \quad 2/20 \quad - \quad 4/20 \quad - \quad 8/20 \quad - \quad 4/20 \quad - \quad 2/20 \end{array}$$

$$\text{la valeur moyenne de } C = \frac{2 \times 3 \times 3}{3+3} = 3$$

$$\text{variance de } C = \frac{2 \times 3 \times 3 \times 3 \times (2 \times 3 \times 3 - 6)}{6 \times 6 \times 5} = 4/10$$

Par conséquent, une famille telle que g.f.f.f.g.g. (3 garçons et 3 filles) ou C = 2, est "plus groupée" que la moyenne représentée elle-même, par exemple, par g.f.f.g.g.f.; une famille "moins groupée" serait g.f.g.f.f.g. ou C = 4

Pratiquement, la répartition de C étant relativement assez voisine d'une distribution binomiale, on a convenu d'assimiler la répartition du nombre total des coupures dans les N familles

ayant toutes les mêmes nombres  $g$  et  $f$  de garçons et de filles à une répartition normale dès que  $N$  atteint ou dépasse 20.

De fait, l'approximation est encore meilleure qu'on pourrait le penser, car cette valeur élimine la majeure partie des cas où  $g$  et  $f$  sont très différents et partant, où la distribution de  $C$  est fortement asymétrique.

On a ainsi calculé pour toutes les classes de famille de  $g$  garçons et de  $f$  filles, représentées par plus de 20 cas, la déviation du nombre total des "coupures" à sa valeur théorique, différence exprimée en écart type (Tableau N° 5).

On voit ainsi que sur les 48 classes, la valeur observée dans 36 cas est plus faible que la valeur théorique; combinant ces résultats par la méthode de E.S. PEARSON déjà utilisée pour les variances on trouve un  $\chi^2$  de 177,4 pour 26 degrés de liberté, ce qui correspond à une probabilité d'erreur de première espèce inférieure à 1 / 100.000 ième.

Moins rigoureusement, on pourrait employer, comme test, le nombre de cas où la valeur observée est inférieure à la valeur théorique, en supposant que dans l'hypothèse nulle cette fréquence est 1/2.

On trouve ainsi:

a- Toutes les classes de famille	
64 cas plus petit / 43 cas plus grand	Probabilité : 0,025
b- Classes représentées par au moins 10 familles	
40 cas plus petit / 17 cas plus grand	Probabilité : 0,0006
c- Classes représentées par au moins 20 familles	
36 cas plus petit / 13 cas plus grand	Probabilité : 0,0003
d- Classes représentées par au moins 100 familles	
22 cas plus petit / 6 cas plus grand	Probabilité : 0,002

e- Classes représentées par au moins 400 familles

10 cas plus petit / 1 cas plus grand Probabilité : 0,059

Il n'est pas besoin d'insister sur le caractère un peu sommaire de ce dernier test en raison de la forte asymétrie des répartitions de C pour les classes représentées par un petit nombre de familles.

Grâce au  $\chi^2$  et à la faible valeur des probabilités qui viennent d'être calculées on doit donc considérer que les diverses familles de même nombre de garçons et de filles n'ont pas la même fréquence et que sont favorisées (faiblement mais indiscutablement) celles où les enfants de même sexe se trouvent "groupés".

Ce fait est d'une extrême importance, car il prouve le caractère intrinsèque de la corrélation entre naissances successives, corrélation qui pourrait, sans cette démonstration, être attribuée, par exemple, à la variabilité de la probabilité de naissance d'un garçon entre les différentes familles.

TABLEAU N° 5

Nombre de coupures ( écart de la valeur observée à la valeur théorique en écart type) pour les familles représentées par plus de 20 Cas.

Nombre de filles	1	2	3	4	5	6	7	8
1				1,2	1,7	0,5	0,7	1,2
2			2,1	1,3	-1,1	1,9	0,9	1,4
3		2,5	2,7	1,2	-0,8	0,5	-0,4	-0,2
4	0,6	1,5	1,0	3,2	0,2	0,6	0,6	-0,1
5	-1,1	1,3	2,1	2,4	-0,5	0,3	-1,0	
6	0,8	0,8	-1,1	0,6	2,0	0,4		
7	0,9	0,1	1,1	1,4	-1,2	1,0		
8	0,8	0,6	-0,3					

C O R R E L A T I O N   E N T R E   L E S   S E X E S  
 =====  
 D A N S   U N E   M E M E   F A M I L L E  
 =====

Ayant établi par le test précédent l'existence d'une liaison entre les naissances consécutives on en évaluera l'importance afin de tester son influence dans la répartition des familles entre les différents types.

Pour ce faire on représentera provisoirement la probabilité d'apparition d'une famille de  $g + f$  enfants par l'expression:

$$p^A (1-p)^{1-A} x^R g (1-x)^{R'g} y^{Rf} (1-y)^{R'f}$$

où  $A$  comme précédemment est égal à 1 ou à 0, selon que le premier né est un garçon ou une fille;  $Rg$  : nombre des naissances masculines précédées d'une naissance masculine;  $R'g$  : nombre des naissances masculines précédées d'une naissance féminine;  $Rf$  : nombre des naissances féminines précédées d'une naissance féminine;  $R'f$  : nombre des naissances féminines précédées d'une naissance masculine;  $p$  = probabilité a priori d'un premier né masculin;  $x$  = probabilité conditionnelle d'un enfant du sexe masculin quand le précédent est du sexe masculin;  $y$  = probabilité conditionnelle d'un enfant du sexe féminin quand le précédent est féminin.

Comme il a été annoncé plus haut, les valeurs de  $Rg$  et de  $Rf$  sont liées au  $Sg$ ,  $Sf$  et  $C$  par les relations (bien évidentes)  $Rg = g - Sg$ ;  $R'g = Sg - B$ ;  $Rf = f - Sf$ ;  $R'f = Sf - (1-B)$ .

Autrement dit, on emploiera provisoirement un schéma de MARKOFF à trois paramètres.

Plus généralement, on étudiera la liaison entre naissances séparées par  $i$  autres naissances c'est-à-dire les valeurs de  $p_{gi}$  (respectivement  $p_{fi}$ ) : probabilité conditionnelle de naissance d'un garçon (respectivement d'une fille) quand un garçon (respectivement une fille) est né précédemment à  $i$  naissances d'intervalles.

Dans les paragraphes précédents on avait  $x = P_{g0}$  et  $y = P_{f0}$ ; un calcul simple, sur lequel on ne s'étendra pas, montre que l'estimation par la méthode du maximum de vraisemblance (qui est tout indiquée ici où le nombre d'observations est grand) conduit aux valeurs suivantes pour  $P_{gi}$ :  $\hat{P}_{gi} = \frac{R_{gi}}{N_g - B_{gi}}$ ; où  $N_g$  représente le nombre total des garçons dans les 14.230 familles;  $B_{gi}$  la somme du nombre des garçons aux rangs  $n-1$ ;  $n-2$ ... $n-1-i$  pour les familles de  $n$  enfants;  $R_{gi}$ : le nombre total des garçons précédés d'un garçon à  $i$  intervalles ( Pour  $i = 0$   $R_{g0} = N_g - S_g$  ); et  $B_{i0}$  = nombre des derniers nés masculins.

On a, évidemment, pour les naissances féminines des expressions correspondantes.

Les tableaux N° 6 et 9 donnent les éléments numériques résultants du dépouillement des 14.230 familles pour les 4 premières valeurs de  $i$  seules intéressantes :

1°- en raison de l'augmentation de la variance des estimations.

2°- en raison du caractère artificiel que présenterait l'assimilation du rapport entre la première et la dernière naissance dans une famille de cinq enfants à celui qui existe entre naissance de rang  $J$  et  $J + 4$  dans les familles de dix enfants, par exemple.

D'autre part, ces tableaux peuvent être considérés comme des tables de contingence où un  $\chi^2$  permet de les tester rigoureusement

En ce qui concerne les variances des estimations, un calcul immédiat montre que  $\hat{P}_{gi}$  et  $\hat{P}_{fi}$  seraient indépendants, si leur valeur vraie était  $1/2$  et qu'elles auraient alors une variance égale à  $1/4$   $(N_{gi} - B_{gi})^{-1}$ ; on admettra que les valeurs actuelles permettent d'employer cette approximation (la différence est de moins  $1\%$ ).

Enfin, on a calculé le coefficient de corrélation  $r_i$  entre les sexes des naissances séparées par  $i$  intervalles.

L'ensemble de ces résultats est résumé au tableau N° 7.

Or, si les chiffres sont significatifs pour  $i=0$  et  $i=1$ ; il n'en est pas de même pour  $i=2$ .

D'autre part, si l'on calcule les probabilités conditionnelles  $p_{g1}$  et  $p_{f1}$ , qui résulteraient de la seule existence de  $p_{g0}$  et  $p_{f0}$ ; c'est-à-dire la liaison entre événements de rang  $r$  et  $r+2$  dans une chaîne de MARKOFF simple; on trouve

$$p'_{g1} = p_0^2 p_{g0} + (1 - p_{g0})(1 - p_{f0}) = 0,5066$$

$$p'_{f1} = p_0^2 p_{f0} + (1 - p_{g0})(1 - p_{g0}) = 0,4943$$

en utilisant les estimations  $\hat{p}_0$  pour valeur de  $p_0$ .

Les différences entre les  $p'_{f1}$  et  $\hat{p}_{f1}$  étant de l'ordre de l'écart type de ces dernières, on voit qu'il n'est pas possible d'écarter l'hypothèse du schéma de MARKOFF sur la base de l'information recueillie, en raison de la faible valeur de ces expériences; et il en sera de même à fortiori pour les  $p_2$ .

Autrement dit, si une théorie propose d'interpréter les désaccords entre la répartition observée et la répartition théorique par la seule influence d'une grossesse sur le sexe à la naissance suivante, il n'existe pas de preuve -permettant de l'écarter- que l'on puisse tirer de l'étude des corrélations dans cette enquête.

Or l'exemple bien connu des "free martins" étudiés par LILLIE, dans le bétail, suggère la possibilité d'une imprégnation masculini-

sante ( ou peut être féminisante) de l'embryon pouvant entraîner une augmentation de la létalité du foetus ultérieur si celui-ci est de sexe différent.

On ne discutera pas ici les nombreux aspects biologiques de l'hypothèse que l'on avance -en faisant d'ailleurs toutes réserves quand à la possibilité d'existence d'autres actions- et on se bornera à faire ressortir que le fait établi ici "liaison intrinsèque entre naissances immédiatement consécutives" ne se trouve complété par l'assertion "suffisant à expliquer les autres corrélations" que relativement au volume de l'information dont on dispose actuellement.

D'autre part, la comparaison entre les valeurs de  $\hat{p}_{g0}$  et de  $\hat{p}_f$  et les fréquences à priori de naissances d'un garçon ou d'une fille montrent également que les différences  $0,5206 - 0,5066 = 0,0140$  et  $0,5087 - 0,4934 = 0,0153$  ne peuvent être considérées comme significativement différentes; ceci revient à dire que l'amélioration de la prédiction permise par la connaissance du sexe de l'enfant précédent peut être considérée comme égale pour les deux sexes, elle est d'environ 1, 5 %.

Légende du tableau N° 6 ci-contre

les chiffres figurant dans les neuf cases du tableau N° 6 doivent être interprétés comme suit pour les différentes valeurs de  $i$   
 $i = 0 ; i = 1 ; i = 2.$

	Nombre de garçons précédés à $i$ intervalles	Nombre de filles précédées à $i$ intervalles	Valeurs marginales
Nombre de garçons suivis à $i$ intervalles	d'un enfant du même sexe ( $g_i$ ) = $R_{gi}$	d'un enfant d'un autre sexe ( $g_f$ ) = $N_g - B_{gi} - R_{gi}$	Nombre total des garçons aux $n-1$ premiers rangs ( $g^0$ ) = $N_g - B_{gi}$
Nombre de filles suivies à $i$ intervalles	d'un enfant d'un autre sexe ( $f_g$ ) = $N_f - B_{fi} - R_{fi}$	d'un enfant du même sexe ( $f_f$ ) = $R_{fi}$	Nombre total des filles aux $n-1$ premiers rangs ( $f^0$ ) = $N_f - R_{fi}$
Valeurs marginales	Nombre total des garçons aux $n-i$ derniers rangs ( $0g$ ) = $N_g - A_{gi}$	Nombre total des filles aux $n-i$ derniers rangs ( $0f$ ) = $N_f - A_{fi}$	Nombre total des couples étudiés ( $00$ ) = $N_g + N_f - A_{gi} - A_{fi}$ = $N_g + N_f - B_{gi} - B_{fi}$

TABLEAU N° 6

## Liaisons entre naissances successives

	Nombre de gar- çons précédés à $i$ intervalles	Nombre de filles précédés à $i-1$ intervalles	Valeurs marginales
	2 1 3 9 8	1 9 7 0 3	4 1 1 0 1
Four $i = 0$	1 9 8 4 4	2 0 5 5 0	4 0 3 9 4
	4 1 2 4 2	4 0 2 5 3	8 1 4 9 5
	1 7 2 9 3	1 6 5 2 5	3 3 8 1 8
Four $i = 1$	1 6 8 0 1	1 6 5 4 0	3 3 3 4 7
	3 4 0 9 4	3 3 0 7 1	6 7 1 6 5
	1 3 6 1 7	1 3 1 4 1	2 6 7 5 8
Four $i = 2$	1 3 6 0 6	1 2 9 7 1	2 6 1 7 7
	3 6 8 2 3	2 6 1 1 2	5 2 9 3 5

Tableau N° 7

Probabilités conditionnelles et corrélations

	chi carré = 70,25	Probabilité : plus petite que 1/1.000
	$\hat{p}_{y0}$	= 0,5206 ± 0,0025
i = 0	$\hat{P}_{f0}$	= 0,5087 ± 0,0025
	$P_0$	= 0,0294 ± 0,0035
		°
		° °
	chi carré = 3,80	Probabilité = 5%
	$\hat{P}_{g1}$	= 0,5114 ± 0,0028
i = 1	$\hat{P}_{f1}$	= 0,4962 ± 0,0028
	$p_1$	= 0,0075 ± 0,0039
		°
		° °
	chi carré = 1,04	Probabilité = 30%
	$\hat{p}_{g2}$	= 0,5089 ± 0,0031
	$\hat{P}_{f2}$	= 0,4955 ± 0,0031
	$p_2$	= 0,0044 ± 0,0043
		°
		° °

MODELE      PROBABILISTE  
=====

Il convient maintenant de faire le point des différentes hypothèses testées successivement; afin de pouvoir construire un modèle: probabiliste de la théorie biologique adoptée; ce modèle devra être le plus simple possible et permettre, après estimation de ses paramètres, de tester l'accord entre la nouvelle théorie et les faits qu'apporterait un retour à l'observation de nouveaux documents.

L'on sait déjà

1° qu'il est possible (malgré les cas exceptionnels des familles de plus de quinze enfants) de considérer comme constante la probabilité à priori de naissance d'un garçon dans les différentes familles et aux différents rangs.

2°- que la seule liaison significative est entre naissances immédiatement consécutives.

Enfin, et surtout, la théorie biologique adoptée implique une influence d'une grossesse sur la grossesse ultérieure.

On est ainsi conduit, tout naturellement, à un schéma de MARKOFF, c'est-à-dire à un tirage aléatoire où la composition de l'urne au ième coup dépend linéairement du résultat du i -1 ième.

Pour éviter de multiplier le nombre des paramètres (qu'on se souvienne de la boutade de SMIRNOV..) on introduira seulement

- a- la probabilité à priori de naissance d'un garçon
- b- un incrément constant quand les deux naissances sont de même sexe.

Plus précisément -et pour des raisons techniques que l'on exposera plus loin- on posera:

Prob. de naissance d'un enfant du sexe  $S_2$  (quand l'enfant précédent est du sexe  $S_1$ ) =  $1/2 (1 + x U + y V)$ .

où  $x$  est une indicatrice égale à  $+1$  ou  $-1$  selon que  $S_2$  est masculin ou féminin et  $y$  une indicatrice égale à

+ 1 si  $S_1 = S_2$  ( les deux enfants sont du même sexe

0 si  $S_1$  n'existe pas (l'enfant est le premier né de la famille).

- 1 si  $S_1 \neq S_2$  ( les deux enfants sont de sexes différents).

C'est là, évidemment, le cas le plus simple d'un tirage en chaîne.

Grâce à l'expression analytique employée, il est possible de considérer comme nulles les puissances de  $U$  et de  $V$  supérieures à la deuxième puisque l'on sait que  $U$  est de l'ordre de  $2\%$  (donc  $U^2$  de l'ordre de  $4/10.000$ ) et  $V$  de l'ordre de  $3\%$  (donc  $V^2$  de l'ordre de  $1/1000$ )/

Pour cette même raison, il est possible avec une bonne approximation de réduire l'équation au maximum de vraisemblance donnant les estimations  $\hat{U}$  et  $\hat{V}$  à des équations linéaires (1)

$$U = \frac{N_g - N_f}{N_g + N_f} = \frac{1265}{95.725} = 0,0132$$

$N_g$  étant le nombre total des naissances masculines et  $N_f$  le nombre total des naissances féminines.

(1) A dessein, on a laissé de côté l'exposé du traitement mathématique qui n'offre d'ailleurs aucune difficulté spéciale.

$$V = \frac{N + M - 2S}{N - M} = \frac{2401}{81.495} = 0,0295.$$

$N$  étant le nombre total des naissances soit  $N = N_g + N_f$  ;  $M$ : le nombre total des familles,  $S$  : le nombre total des séquences (ces chiffres se déduisent d'ailleurs immédiatement des calculs des chapitres précédents).

2°- de considérer les estimations  $\hat{U}$  et  $\hat{V}$  comme quadratiquement indépendantes.

3°- de prendre comme expression de leur écart type les suivantes

$$\text{écart type de } \hat{U} = \frac{1}{\sqrt{N}} = 0,0032$$

$$\text{écart type de } \hat{V} = \frac{1}{\sqrt{N - M}} = 0,0035$$

On est alors en mesure de représenter, comme on l'a déjà fait au début du chapitre précédent la fréquence des familles de chaque type par l'expression:

$$\left(\frac{1 + AU}{2}\right) \left(\frac{1 + U + V}{2}\right)^{R_g} \left(\frac{1 + U - V}{2}\right)^{R'_g} \left(\frac{1 - U + V}{2}\right)^{R_f} \left(\frac{1 - U - V}{2}\right)^{R'_f}$$

avec les notations  $A$ ,  $R_g$  et  $R_f$  du chapitre précédent.

Toujours, conformément au même principe d'approximation, on pourra calculer les fréquences relatives dans l'ensemble des familles de  $n = g + f$  enfants, de la classe des familles de  $g$  garçons et  $f$  filles comme

$$N_{gf} = \begin{bmatrix} g + f \\ g \end{bmatrix} \times 2^{-(g+f)} \times \left[ 1 + (g - f)U + (g+f-1-2C_{gf})V \right]$$

où  $C_{gf}$  désigne la valeur moyenne du nombre des "coupures".

Pour les premiers histogrammes (  $g + f = 5, 6, 7, 8$  ) on a comparé les valeurs observées dans les différentes classes aux valeurs théoriques.

a- dans l'hypothèse adoptée d'un tirage en chaîne simple de MARKOFF.

b- dans l'hypothèse d'une répartition binomiale.

Les paramètres étant estimés sur la base de toutes les familles il semble logique de prendre pour nombre des degrés de liberté pour les  $\chi^2$  relatifs aux familles de  $n$  enfants la valeur  $n$  et non pas  $n - 1$  et  $n - 2$  comme on devrait le faire si  $\hat{U}$  et  $\hat{V}$  étaient estimés pour chaque valeur de  $g + f$ .

On constate (cf. tableau N° 8 ) qu' l'hypothèse de MARKOFF rend compte des observations d'une manière beaucoup plus précise sauf en ce qui concerne  $n = 8$  ) .

Les probabilités des  $\chi_1^2$  (schéma binomial) et  $\chi_2^2$  (schéma de MARKOFF ) sont d'ailleurs dans les intervalles suivants:

	$n = 5$	$n = 6$	$n = 7$	$n = 8$
$\chi_1^2 =$	$0,05 \leq 0,10$	$0,10 \leq 0,20$	$0,02 \leq 0,005$	$0,10 \leq 0,20$
$\chi_2^2 =$	$0,30 \leq 0,50$	$0,30 \leq 0,90$	$0,50 \leq 0,70$	$0,05 \leq 0,10$

montrant un bon accord entre l'hypothèse et la réalité.

Il est intéressant de remarquer qu'une cause de désaccord, entre les deux hypothèses et l'observation, pour  $n = 5$  est le faible nombre de familles composées exclusivement de filles; peut être faut-il ne voir là que l'expression de causes économiques et sociales qui pousseraient particulièrement à de nouvelles grossesses, les parents n'ayant eu que des enfants du sexe féminin ( ?.??. )

Les chiffres sont trop faibles d'ailleurs pour pouvoir

constater l'augmentation du nombre des familles de 6 enfants du type f,f,f,f,f, g ou f,f,f,f,f,f, qui devrait en résulter.

On pourrait calculer également la valeur théorique de la variance de ces répartitions qui devraient être (pour n enfants)

$$\frac{n}{4} - \frac{n^2}{4} U^2 + \frac{1}{n^{2n}} \times h(n) \hat{V}$$

où h(n) est une expression complexe égale à

$$h(n) \sum_{J=0}^{n+1} \binom{n}{J} \left[ 4 J^4 - 4 n J^3 + (n^2 - n) J^2 \right]$$

Enfin, on trouverait pour valeur théorique du nombre moyen des "coupures" dans les familles de g garçons et f filles:

$$C_{gf} = \frac{1 + (g-f) U + (g+f-1-2D_{gf}) V}{1 + (g-f) U + (g+f-1-2C_{gf}) V}$$

où  $C_{gf}$ , comme précédemment est le nombre moyen des "coupures" dans l'hypothèse nulle =  $\frac{2 g f}{g+f}$  et  $D_{gf}$  une constante liée à la variance

des "coupures" et égale à  $\frac{2 g f - 1}{g + f - 1}$

Quelques exemples numériques, comme celui des familles de cinq garçons et cinq filles; où la valeur moyenne du nombre des "coupures" devient 4,98 au lieu de 5,00, suffiraient à convaincre combien est faible, pratiquement, l'influence de l'incrément.

Année 1948 1948-12. Contribution à l'étude statistique du sexe à la naissance

Tableau N° 8

Comparaisons entre les histogrammes observés et les schémas probabilistes

Nombre de garçons	Nombres observés	Nombres théoriques dans l'hypothèse binomiale	Nombres théoriques dans l'hypothèse de MARKOFF	
FAMILLES DE CINQ ENFANTS				
5	204	171	190	
4	841	833	852	chi carré = 9,73
3	1.585	1.624	1.386	
2	1.544	1.581	1.543	chi carré = 5,88
1	810	769	788	
0	144	150	169	
FAMILLES DE SIX ENFANTS				
6	65	55	63	
5	330	321	336	
4	785	783	776	chi carré = 9,75
3	979	1.017	987	
2	717	743	733	chi carré = 2,24
1	325	289	304	
0	54	47	54	
FAMILLES DE SEPT ENFANTS				
7	19	16	19	
6	126	109	117	
5	304	320	323	CHI carré = 15,00
4	529	520	507	
3	480	507	493	chi carré = 8,16
2	287	296	206	
1	111	96	104	
0	21	13	10	
FAMILLES DE HUIT ENFANTS				
8	6	7	5	
7	49	55	60	
6	170	185	190	chi carré = 11,34
5	406	356	351	
4	412	444	432	chi carré = 14,44
3	327	348	343	
2	180	171	176	
1	61	47	52	
0	8	6	9	

DISCUSSION      SOMMAIRE      DES  
 =====

METHODES      EMPLOYEES  
 =====

On a volontairement employé une démarche progressive, cherchant à serrer de plus en plus près la réalité par le jeu dialectique des hypothèses biologiques successives, traduites statistiquement en hypothèses nulles et modèles probabilistes et confrontées à l'observation.

Une autre approche eut été possible, partant directement du schéma de MARKOFF et l'éprouvant par une double méthode d'estimation que l'on peut brièvement esquisser.

Il est possible, en effet, d'estimer de deux manières l'incrément  $V$  lié à la succession immédiate de deux naissances de même sexe = a-: par la seule connaissance de l'histogramme des répartitions b-: pour chaque classe de familles de  $g$  garçons et de  $f$  filles par la connaissance de la valeur observée du nombre des "coupures".

a- En employant la première méthode et en se contentant de l'approximation déjà constamment utilisée, on trouve pour valeur de  $\hat{V}$  une expression simple, absolument correcte (unbiaissée), asymptotiquement équivalente à la solution de l'équation au maximum de vraisemblance à laquelle elle est tangente quand  $V$  est un infiniment petit de premier ordre:

$$\hat{V} = \sum_j \frac{N_g H_g}{K^2 M}$$

la sommation étant étendue, pour l'ensemble des  $M$  familles de  $(g + f)$  = (constante) enfants, à toutes les valeurs de  $g$  : les  $H_g$  sont liés simplement au nombre moyen des "coupures" et à la probabilité

$$\frac{1 + V}{2} \quad \text{et} \quad K = \frac{2(g + f - 1)}{g + f}$$

La varxiance de cette estimation est alors sensiblement égale à  $(KM) - 1$ .

b- En considérant comme seules connues les familles de  $g$  garçons et  $f$  filles et la fréquence relative  $M_C$  de celles qui présentent  $C$  "coupures", on trouve une deuxième estimation, également tangente à l'équation du maximum de vraisemblance, mais affectée d'une erreur systématique, facile à éliminer.

Il serait donc possible de comparer ces deux estimations à l'écart type de leur différence et de tester l'hypothèse nulle de la non existence d'une corrélation résiduelle entre les sexes des naissances dans une même famille.

On ne discutera pas, ici, ces calculs; mais on montrera seulement la perte d'information sur  $\hat{V}$ , qui résulterait de la seule connaissance des histogrammes pour  $M$  familles de  $g + f$  enfants; en effet dans les documents employés, chaque famille apportait une information de  $g + f - 1$ ; alors que l'on aurait ici seulement

$$\frac{2(g + f - 1)}{g + f}$$

La perte d'information est donc de  $\frac{g+f}{2}$ .

Dans le cas présent, on peut évaluer que pour obtenir la même précision, il eut fallu recueillir 3,1 fois plus de familles,

soit environ 43.000.

Il n'est pas inintéressant, non plus, à la lumière des chiffres calculés de voir rapidement ce que peut apporter comme information un certain ensemble de documents et tout particulièrement le nombre minimum de familles qu'il faudrait rassembler pour pouvoir l'opposer à la valeur de l'incrément  $V$ , obtenue ici.

Comme il ne s'agit là que d'indications très sommaires, on prendra deux cas extrêmes pour les familles de trois garçons et de trois filles:

1°- Nombre  $M'$  minimum de famille requis pour établir que  $V$  est significativement plus faible que 0,0295 si la valeur de  $\hat{V}$  dans l'échantillon est zéro; on a sans peine

$$0,0295 \geq 3 \sqrt{\frac{1}{81.495} + \frac{1}{5 M'}}$$

soit  $M'$  plus grand que 2.388.

2°- Valeur moyenne maxima de  $V'$  dans un échantillon de cinquante familles; pour établir que  $V$  est significativement plus petit que 0,0295, on a :

$$0,0295 - V' \geq 3 \sqrt{\frac{1}{81.495} + \frac{1 - V'^2}{250}}$$

soit  $V'$  plus petit que - 0,56; soit encore  $S$  plus grand que 245; ce qui implique notamment que moins de quinze familles doivent avoir un nombre de "coupures" inférieur à la valeur moyenne 3.

Ces quelques remarques sont suffisantes pour montrer, on l'espère, combien ces effets de "groupage" sont difficiles à étudier sur des faibles populations et combien pèse peu dans la discussion un échantillon limité à quelques dizaines de cas même s'il

Année 1948      1948-12. Contribution à l'étude statistique du sexe à la naissance

47

**s'agit de familles très nombreuses: une famille de dix sept enfants par exemple, n'apporte que deux fois plus d'information sur ce point qu'une famille de neuf enfants et quatre fois plus qu'une famille de cinq enfants.**

**C O N C L U S I O N S**  
=====

L'étude a porté sur 14.230 familles de [ 5 à 19 enfants ], soit au total 95.725 naissances; étant donné le but poursuivi, une sélection préalable des documents a permis de ne retenir que les familles françaises sans multiparité; divers aspects concrets, tels que le milieu économique et social, l'âge et la santé des parents, l'influence des causes économiques et psychologiques-pouvant favoriser tel ou tel type de famille- ont du être négligés au profit de phénomènes statistiques plus généraux.

. ° .

1°- Les résultats suivant peuvent être considérés comme établis:

L'hypothèse d'une répartition des sexes à la naissance représentable par un mécanisme de tirages successifs indépendants dans une urne à composition constante doit être rejetée.

- en raison de l'étalement trop grand des histogrammes, c'est à dire l'excès des familles à forte prédominance masculine ( ou féminine) sur les valeurs théoriques qui résulteraient d'une répartition binomiale (hypothèse a ) issue du schéma précédent;

- en raison de la diminution du nombre moyen des séquences de même sexe, traduisant une liaison entre naissances consécutives ( hypothèse b ) indépendantes du fait précédent.

++++

II°- Au degré d'approximation, que permet le montant total de l'information recueilli, un modèle statistique de tirage aléatoire en chaîne (simple) de MARKOFF a été construit, qui rend compte de l'ensemble des chiffres obtenus d'une manière satisfaisante.

++++

III°- Par contre quatre hypothèses ( c, d, e, f .) n'ont pu être définitivement confirmées ou infirmées:

- L'amélioration de la prévision due à la connaissance du sexe de l'enfant précédent est différent selon les sexes (hypothèse c ).

- La probabilité de naissance d'un garçon ou d'une fille varie selon le rang de celle-ci (hypothèse d.)

- Se surajoutent à la corrélation entre naissances immédiatement consécutives-

- il existe une corrélation résiduelle non nulle entre naissances non immédiatement consécutives (hypothèse e).

- il existe une variation, entre familles, de la probabilité de naissance d'un garçon (hypothèse f).

°  
° °

Mais si l'on a tenu à présenter ainsi abstraitement les résultats de cette enquête, afin d'en rendre plus clair l'énoncé, il convient évidemment d'avoir toujours présent à l'esprit le contenu concret, biologique des hypothèses et modèles probabilistes qui seul légitime l'emploi de ces derniers; simultanément, il est nécessaire de voir les limitations internes qu'imposait à ce

travail la nature même des documents rassemblés.

Le rejet de l'hypothèse nulle (hypothèse a) de la répartition binomiale traduit le fait déjà constaté par d'autres auteurs que l'hypothèse d'une détermination du sexe par un mécanisme ( inégale vitalité de l'embryon ou fécondation élective .. etc.. etc..) constant d'une grossesse à l'autre et caractérisé par les mêmes paramètres pour tous les couples ne peut être conservée.

L'existence d'une liaison entre naissances immédiatement consécutives (hypothèse b) rendant compte des faits observés, peut traduire deux hypothèses biologiques bien distinctes qui sont:

A- L'hypothèse adoptée pour donner un modèle probabiliste : il existe une influence masculinisante (ou féminisante) d'une grossesse sur une grossesse immédiatement consécutive, égale pour les deux sexes (hypothèse d).

B- Il existe des périodes plus ou moins androphiles(ou gynophiles), qui, pour le volume d'informations recueilli ne se manifestent de manière significative que par leurs actions sur les naissances contiguës; mais qui pourraient (et peut être simultanément).

varier de manière systématique, constante d'une famille à l'autre et en fonction de l'âge et du passé obstétrical de la mère (hypothèse e).

avoir une plus ou moins grande durée (3, 4 gestations au plus) étant dues au milieu ou au genre de vie des parents (hypothèse c).

être caractéristique d'un couple donné (et peut être lié à des particularités du génotype des parents) - hypothèse f).

Ce n'est que l'élargissement ( enrichissement) de la nature des documents de base, l'observation et la comparaison avec d'autres espèces, où l'expérimentation serait possible; qui pourraient permettre de préciser ces faits; mais surtout d'élucider les mécanismes (letalité ou fécondation élective ou détermination épigamique) par lesquels influent les causes dont on a pu mettre en évidence un des effets les plus marquants.

Le Président de Thèse

R. T U R P I N

VU

Le Doyen

L. B I N K E T

VU et PERMIS D'IMPRIMER

Le Recteur:

SARRAIL

## REFERENCES

=====

La liste suivante n'a évidemment aucune prétention à être même l'ébauche d'une bibliographie sur un sujet qui a tellement fait couler d'encre et dont les ramifications sont si nombreuses.

On a cité quelques travaux particulièrement représentatifs des thèses étudiées ici statistiquement. En outre, pour certains points de technique statistique, qui risquaient d'être peu familiers aux lecteurs, on a tenu à indiquer les mémoires originaux.

°  
° °

- 1- CIOCCO.A.1938. Variations in the sex ratio at birth in the United States- Human Bio.10. pp 36-44
- 2- CREW.F.A.E.1937. The sex ratio. Amer.Nat.71. pp 529-559. British Assoc.Report.1937
- 3- FISCHER?R.A.1944. Statistical Methods for research workers p 129
- 4- FRIEDA.S.SWED and EISENHART.C. 1943. Tables for testing randomness of grouping in a sequence of alternatives. An.of Math.Stat Vol 14.
- 5- GEIBLER.A.1889. Beiträge der K.Sächsischen Stat.Bureau 35 pp 1-2
- 6- KENDALL.M.G. and B.BINGTON SMITH B. 1938. Distribution of digits obtained by an observer in using a Randomising machine supp J.R.Stat.Soc. pp 6-21
- 7- PARKES.A.S. 1938. The factors governing the mammalian sex ratio Sci.Progress 18 pp 426-435
- 8- PEARSON.E.S. 1938. The probability integral transformation for testing goodness of fit and combining independent test of significance. Biom.10. pp 134
- 9- STEVENS.W.L. 1939.Distribution of groups in a sequence of alternatives. Ann.of Eug. Vol IX.
- 10- TSCHERMAK SEYSENEG.G.A. 1939. Versuch einer Mendelistischen Erklärung des Geschlechterverhältniss des menschlichen Neugeborenen. Bio.Generalis 14 pp 542(551)
- 11- WALD.A. and WOLFOVITZ.J.1940.On a test of wheter two samples are from the same population. Ann.of Math.Stat. Vol XI.
- 12- WICKSELL.S.D. 1936. Sex proportion and parental age. Kungl Fysiogra. Sälls Handl. 37 . N° 6 .

=====

# Année 1949

## Bibliographie

- [1949-1] Marcel-Paul Schützenberger et Raymond Turpin. Sur une asymétrie latérale statistique, attribut du phénotype sexuel humain. *C. R. Acad. Sci. Paris*, 228 :431–432, 1949. Séance du 31 janvier 1949.
- [1949-2] Marcel-Paul Schützenberger et Raymond Turpin. Sur la détermination du sexe chez l’homme. *Semaine des Hôpitaux de Paris*, 25(60) :2544–2545, 14 août 1949.
- [1949-3] Raymond Turpin, Marcel-Paul Schützenberger et M. Tisserand. Remarque sur la statistique des becs-de-lièvre. *Semaine des Hôpitaux de Paris*, 25(60) :2545–2546, 14 août 1949.
- [1949-4] Marcel-Paul Schützenberger et Raymond Turpin. L’étude des dermatoglyphes. *Semaine des Hôpitaux de Paris*, 25(60) :2553–2562, 14 août 1949.
- [1949-5] Marcel-Paul Schützenberger. Une application de l’analyse séquentielle. *Semaine des Hôpitaux de Paris*, 25(60) :2562–2564, 14 août 1949.
- [1949-6] Marcel-Paul Schützenberger. Résultats d’une enquête sur la distribution du sexe dans les familles nombreuses. *Semaine des Hôpitaux de Paris*, 25(61) :2579–2582, 18 août 1949.
- [1949-7] Marcel-Paul Schützenberger. Fondements de la statistique appliquée à la psychologie, I. Principes des tests statistiques. *Bull. Groupe Étude Psychol. Univ. Paris*, 2e année(Fasc. 8–9) :73–74, 11 mai 1949.
- [1949-8] Marcel-Paul Schützenberger. Fondements de la statistique appliquée à la psychologie, II. Estimation d’un paramètre inconnu. *Bull. Groupe Étude Psychol. Univ. Paris*, 2e année(Fasc. 10–11–12) :111–112, 23 mai 1949.
- [1949-9] Marcel-Paul Schützenberger. À propos de la “cybernétique” (mathématiques et psychologie). *Évolution Psychiatrique*, pages 585–607, 1949.

Bibliographie

Année 1949

- [1949-10] Marcel-Paul Schützenberger. Sur l'extension des théorèmes de dualité aux treillis distributifs non complémentés. *C. R. Acad. Sci. Paris*, 228 :33–35, 1949. Séance du 15 novembre 1948.
- [1949-11] Marcel-Paul Schützenberger. A non-existence theorem for an infinite family of symmetrical block designs. *Ann. Eugenics*, 14 :286–287, 1949.
- [1949-12] Henri Duchêne et Marcel-Paul Schützenberger. Quelques réserves sur la méthodologie de L. Szondi. *Inform. Psychiat.*, 25e année, 4e série(6) :210–216, juillet 1949.
- [1949-13] Anne Ancelin, Henri Duchêne et Marcel-Paul Schützenberger. Investigaciones críticas sobre la teoría y el test de L. Szondi. *Revista de Psicología General y Aplicada*, IV(11) :437–449, juillet–septembre 1949.
- [1949-14] Nicolas Koboziëff, Mme. Pomriaskinsky-Koboziëff et Marcel-Paul Schützenberger. Sur une différence pondérale à la naissance entre deux lignées de souris : *Mus musculus*. *C. R. Acad. Sci. Paris*, 229 :1267–1268, 1949. Séance du 28 novembre 1949.

GÉNÉTIQUE. — *Sur une asymétrie latérale statistique, attribut du phénotype sexuel humain.* Note de MM. **RAYMOND TURPIN** et **MARCEL-PAUL SCHÜTZENBERGER**, présentée par M. Pierre-P. Grassé.

Morphologiquement, les dermatoglyphes digitaux se classent en : *arches*, *boucles* (à ouverture *radiale* ou *cubitale*), et *tourbillons*. Chez les sujets normaux, la distribution statistique de ces types varie selon le doigt, la main droite ou gauche, le sexe et la population envisagée. Apparemment plus soucieux des différences qui opposent entre eux les groupes ethniques, la plupart des auteurs ont négligé les caractères communs à l'espèce humaine et qui dépendent moins, sans doute, de la fréquence relative des gènes en cause qui diffère suivant l'aire géographique. Cependant, malgré l'extrême variabilité des proportions des différents types de figures, le rapport de leurs fréquences (ou, plus grossièrement, l'ordre de celles-ci), dans les quatre catégories :

Main gauche chez l'homme (MG),	
»	la femme (FG):
main droite chez l'homme (MD),	
»	la femme (FD):

est, de façon remarquable, relativement stable et caractéristique de chaque doigt.

Utilisant tous les documents que nous avons pu réunir et qui constituent un échantillonnage assez varié de la population du globe, nous avons établi le fait suivant sur les six groupes qui comprennent au moins 400 individus de chaque sexe (1) :

*Au médius et à l'index, de façon constante, les boucles radiales prédominent à droite chez l'homme et à gauche chez la femme.* Cette opposition est d'autant plus remarquable qu'elle ne s'accompagne que de différences peu sensibles sur les fréquences cumulées. Notamment, au médius, l'ordre de fréquence des boucles radiales est en général :  $FD < MG < MD < FG$ , alors qu'à l'index cet

(1) 101 511 danois (J. N. Bugge); 66 000 hispano-indiens (K. O. Henckel); 22 000 grand russes (P. S. Semenovski); 2 000 javanais (J. Dankmeijer); 800 belges (H. T. Piebenga); et enfin, 1 000 français dont nous avons rassemblé les empreintes grâce à M. le Professeur Sannié, Directeur de l'Identité Judiciaire.

( 2 )

ordre moyen est  $FD < FG < MG < MD$ . Un phénomène semblable s'observe, pour les tourbillons, au médius, avec le même ordre  $FD < FG < MG < MD$ .

En ce qui concerne les autres doigts et les autres types de figures, les interactions *côté*  $\times$  *sexe* sont souvent significatives <sup>(2)</sup>, mais n'ont presque jamais ce caractère de véritable *renversement de symétrie* (surtout en ce qui concerne les *arches*, d'une part, les *quatrième et cinquième* doigts d'autre part). En l'absence d'une théorie suffisamment développée, il est difficile d'en apprécier la signification biologique, car leur mesure dépend trop de l'outil mathématique employé.

Bien que nous ne puissions encore interpréter ce phénomène d'asymétrie latérale dont la réalité est indiscutable, il mérite à notre avis d'attirer l'attention des chercheurs, car il constitue un type nouveau d'attribut sexuel.

---

<sup>(2)</sup> L'existence d'interactions *côté*  $\times$  *sexe* pour les arches et pour les tourbillons a été souvent mentionnée mais sans précision, les auteurs s'étant bornés le plus souvent à comparer les fréquences totales. Il semble que seul Semenovski (*Russkij antropologičeskij žurnal*, 16, 1927, p. 7-21) ait aperçu l'asymétrie qui fait l'objet de la présente Note, mais sans disposer, à cette époque, de documents assez nombreux et assez variés pour en pouvoir prouver la généralité.

(Extrait des *Comptes rendus des séances de l'Académie des Sciences*,  
t. 228, p. 431-432, séance du 31 janvier 1949.)

## SUR LA DÉTERMINATION DU SEXE CHEZ L'HOMME

par R. TURPIN et M. P. SCHÜTZENBERGER

La constance de la *sex ratio* dans l'espèce humaine pose en même temps le problème de ses variations en fonction du milieu (1) et du développement biologique propre des progéniteurs (2) et divers travaux ont attiré l'attention sur l'excès des garçons parmi les premiers nés, parmi les naissances « après les guerres », etc., sans que les chiffres publiés concordent jamais parfaitement entre eux.

Pour éviter les inconvénients propres aux statistiques de masse nous avons jugé intéressant de recueillir l'ordre des sexes à la naissance dans un nombre suffisamment important de familles, espérant pouvoir mettre ainsi en évidence des déviations plus nettes par rapport à la loi binomiale qui régirait le déterminisme du sexe si aucun autre facteur extérieur au schéma chromosomique n'intervenait dans son mécanisme.

## LE MATÉRIEL STATISTIQUE.

Afin de diminuer l'influence de restrictions volontaires ou non à la natalité, nous avons borné notre enquête aux « familles nombreuses » définies plus ou moins arbitrairement par la limite de 5 enfants. En outre, les naissances gemellaires introduisant des complications techniques considérables, nous avons seulement recueilli les familles qui n'en comportaient pas. Pour cette collecte nous nous sommes adressés successivement : aux associations de familles nombreuses de la Seine, à des annuaires donnant le relevé des « grandes familles » des départements du Nord de la France, aux dossiers de la fondation Cognacq-Jay, enfin et surtout aux dossiers que déposent au Ministère de la Santé Publique et de la Population les postulantes à la Médaille des Familles.

Grâce à l'appoint (très inégal d'ailleurs) de ces diverses sources, nous disposons donc de 14.230 familles françaises contemporaines, d'au moins 5 enfants, ce total se décomposant de la manière suivante (tableau I) :

TABLEAU I

512 familles de 5 enfants	108 familles de 13 enfants
3255 » 6 »	59 » 14 »
1877 » 7 »	37 » 15 »
1619 » 8 »	22 » 16 »
872 » 9 »	11 » 17 »
634 » 10 »	2 » 18 »
384 » 11 »	4 » 19 »
218 » 12 »	

Il convient de faire observer que l'importance relative de ces divers types de familles n'a nullement la prétention de refléter la composition de la population : d'une part la manière même dont nos documents ont été rassemblés favorisait les familles très nombreuses (et ceci à dessein), d'autre part, nombre de ces familles étaient « en cours », c'est-à-dire qu'elles ont été fixées ici à un moment où la vie obstétricale de la mère n'était probablement pas terminée. Signalons enfin que dans toutes ces familles le père était censé être le même pour tous les enfants, les cas où la mère s'était remariée ayant été éliminés.

## VARIATIONS DE LA « SEX RATIO ».

Calculé sur l'ensemble des 95.875 naissances le rapport des sexes à la naissance a été trouvé égal à 0,5066 + 0,0016, légèrement plus faible que les valeurs habituelles. Aucune variation systématique certaine de ce chiffre n'a pu être établie en fonction du nombre des enfants dans la famille

non plus qu'en fonction du rang de naissance. Le léger excès de garçons parmi les premiers nés pourrait aussi bien être interprété comme une fluctuation d'échantillonnage. Par contre, lorsque l'on compare le nombre des familles présentant un fort excès de l'un des sexes à la valeur que devrait avoir ce nombre dans le cas d'une répartition du sexe avec une probabilité constante (de fait on compare la valeur observée de la variance des répartitions pour  $n$  enfants à sa valeur théorique dans un schéma binomial) on trouve une nette prédominance de ce type de familles. Ce résultat déjà obtenu par R. A. Fisher travaillant sur les matériaux de Geissler (3) (familles de 8 enfants) peut s'interpréter de plusieurs manières :

1° D'abord, en supposant que certaines familles ont une prédisposition à procréer des garçons ou des filles différente de la valeur moyenne générale. Cette théorie se trouve étayée par diverses observations montrant des familles où se transmet de manière héréditaire l'excès des individus d'un sexe ou d'un autre.

2° En supposant qu'il existe une liaison entre les sexes des naissances successives, ce qui se traduit analytiquement par le remplacement du schéma binomial par un schéma de Markoff plus ou moins complexe.

3° En supposant que la prédisposition à engendrer des garçons ou des filles varie sensiblement au cours de l'existence d'un certain nombre de femmes. Il faut alors admettre que ces périodes gynophiles ou androphiles se distribuent plus ou moins au hasard ; ou tout au moins sans préférence nette pour le début ou la fin de la vie obstétricale. En outre, leur durée doit excéder assez fréquemment la durée normale de l'intervalle entre deux gestations.

## TEST DES HYPOTHÈSES.

La structure même de nos documents permet sans peine, sinon de trancher entre ces diverses hypothèses, tout au moins d'en évaluer l'influence relative à condition que celle-ci soit suffisamment importante.

Tout d'abord, le calcul des corrélations entre le sexe pour les naissances immédiatement consécutives ( $r_0$ ), séparées par une autre naissance ( $r_1$ ), par deux ( $r_2$ ), etc., donne les résultats suivants :

$$\begin{aligned} r_0 &= 0,029 \pm 0,0035 \\ r_1 &= 0,007 \pm 0,0039 \\ r_2 &= 0,004 \pm 0,0043 \end{aligned}$$

(Les autres corrélations doivent être considérées comme pratiquement nulles.)

Un calcul simple montre alors que si l'on élimine l'effet de la liaison entre naissances immédiatement consécutives, l'influence des causes invoquées dans l'hypothèse 1 ne peut être que très faible et, en tous cas, impossible à mettre en évidence sur nos documents actuels. De même, le seul fait que ces coefficients soient à la limite de la signification ne permet pas d'espérer que l'on puisse trancher entre les hypothèses 2 et 3 puisque la corrélation apparente  $r_1$  peut parfaitement apparaître comme résultant d'un schéma de Markoff en chaîne simple comme d'un modèle statistique plus complexe.

Par contre, pour tester rigoureusement l'existence d'une liaison entre sexes indépendante d'une variation entre les familles de la probabilité de naissance d'un garçon, on dispose d'une méthode parfaitement rigoureuse :

## ETUDE DES SÉQUENCES.

On a employé la méthode déjà utilisée (4) pour tester l'indépendance d'une série d'événements (dans le contrôle statistique des fabrications industrielles) en l'adaptant au cas actuel où les séries sont relativement très courtes. Les familles étant classées d'après le nombre d'enfants et la proportion des garçons, la répartition des rangs de ceux-ci peut être étudiée de plus près : par exemple pour les familles de 3 garçons et 2 filles, 10 types de familles sont possibles *a priori* : gggff, ggfgf, ggffg, gfgfg, gffgg, fgggf, fggfg, fgfgg, fgggg. Dans certains de ces types (par exemple gggff ou fgggg) les enfants de même sexe sont groupés ensemble, dans d'autres types ils alternent plus ou moins régulièrement (par exemple gfgfg) et pour chaque type on peut définir un indice (nombre de successions d'un enfant à un enfant du sexe opposé) traduisant numériquement cette différence.

Dans le cas où seuls joueraient des facteurs constants tout le long de la vie de la mère (hypothèse 1), tous ces types auraient la même probabilité *a priori*. Inversement dans les autres hypothèses les types tels que gggff ou fgggg devraient être relativement plus fréquents que les types tels que gfgfg. C'est effectivement ce que l'on constate pour chaque classe de familles de  $x$  garçons et  $y$  filles, la valeur moyenne de l'indice que l'on vient de définir est calculée et comparée à sa valeur *a priori* dans le schéma binomial. Pour les 47 classes représentées par plus de 10 familles, la valeur observée est plus grande que la valeur théorique dans 40 cas. De manière moins grossière, en combinant tous ces tests par la méthode de E. S. Pearson (par transformation des vraisemblances en chi-carrés) on obtient sur l'ensemble un résultat significatif à plus de 1/100.000 ce qui établit sans contestation la réalité du phénomène, tout au moins pour nos documents.

## CONCLUSION

Utilisant le modèle probabiliste le plus simple pour re-

présenter ces faits, nous avons été conduits à admettre qu'après la naissance d'un garçon la probabilité pour qu'il en naisse un second est augmentée d'environ 1 %, de même la probabilité de naissance d'une fille après une autre fille est accrue de la même valeur. Nos documents ne permettent pas de préciser plus le phénomène observé et le schéma statistique rend compte d'une manière satisfaisante de l'ensemble des faits.

Il est encore difficile d'interpréter ce phénomène et nous espérons que d'autres chercheurs pourront en reprendre l'étude et en confirmer la généralité et la permanence. Nous poursuivons d'ailleurs ces recherches en utilisant cette fois la notion de durée de l'intervalle entre les naissances dont nous n'avions pu tenir compte dans cette première enquête (en raison du volume des documents à rassembler). Elles permettront d'approfondir le mécanisme de la détermination du sexe chez l'homme. Bien qu'il soit encore trop tôt pour porter un jugement définitif sur leurs résultats, il semble qu'elles confirment déjà nos conclusions antérieures : l'intervalle qui sépare deux naissances consécutives de même sexe est en moyenne sensiblement plus court que l'intervalle entre deux naissances de sexe différent.

## REFERENCES

- (1) CREW F. A. E., 1937. — The sex ratio. *Amer. Nat.*, 71, 529-559.
- (2) WICKSELLE S. D., 1926. — Sex proportion and parental age. *Kung. Fysiogr. Sälls Handling*, 37, n<sup>o</sup> 6.
- (3) GEISSLER A., 1889. — *Beiträge der K. Sächsischen Stat. Bureau*, 35.
- (4) WALD A. and WOLFOWITZ J., 1940. — On a test of whether two samples are from the same population. *Ann. of Math. Stat.* (XI).
- (5) TURPIN R. et SCHÜTZENBERGER M. P., 1948. — Recherche statistique sur la distribution du sexe à la naissance. *Compte Rendus Acad. Sci.*, 226, pp. 1845-1846.

(Travail subventionné par l'Institut National d'Hygiène et le Fonds d'études de la Société Médicale des Hôpitaux de Paris).

REMARQUE SUR LA STATISTIQUE DES BECS-DE-LIÈVRE

par R. TURPIN, M. TISSERAND et M. P. SCHÜTZENBERGER

La plupart des auteurs ont signalé la variabilité de la fréquence d'apparition des becs-de-lièvre selon le sexe, le côté, l'intensité de la forme, etc. Nous avons pensé qu'il serait intéressant de comparer sous cet angle les différents chiffres publiés afin de contrôler l'existence et l'étendue des interactions. Malheureusement, soit que les documents aient été hors de notre portée (et notamment en ce qui concerne les mémoires de G. Kloeppe, R. R. W. Hinrichsen, L. G. Grace), soit que les chiffres publiés aient été peu propres à cette nouvelle exploitation statistique (par exemple : J. Sanders, W. Lange, K. Eickler), nous avons dû limiter notre étude aux deux derniers travaux de Veau et de P. F. Andersen, analysés d'un autre point de vue dans l'article précédent.

PRINCIPES DE L'ANALYSE STATISTIQUE.

Dans chaque échantillon, les chiffres sont classés :

- selon le sexe;
- selon le côté (gauche, droit, bilatéral) ;
- selon l'intensité (simple, complet).

Le dispositif statistique est donc du type  $2 \times 2 \times 3 \times 2$  puisque deux échantillons sont utilisés.

La nature même du problème suggère la définition des interactions : on dira qu'il y a interactions entre plusieurs principes de classification quand la fréquence correspondante sera différente du produit des probabilités marginales (prises égales aux fréquences marginales observées). Ainsi une interaction (ES) (Echantillon  $\times$  Sexe) signifie-t-elle que la proportion totale des sujets masculins (par exemple) n'est pas la même dans les deux échantillons. Une interaction (SCI) (Sexe  $\times$  Côté  $\times$  Intensité) signifierait que dans les deux sexes la répartition des formes simples et complètes n'est pas la même selon le côté. Mettant provisoirement à part toute préoccupation probabilitaire, il est bien évident que si l'on fait préalablement subir aux fréquences observées une transformation logarithmique, on se trouve formellement dans la situation de l'analyse de variance. Cependant, si cette transformation possède un contenu algébrique satisfaisant, elle ne remplit pas les conditions statistiques nécessaires et en particulier elle ne stabilise pas la variance (qui ici est proportionnelle à la fréquence puisqu'il s'agit d'une loi de Poisson). En conséquence on a préféré utiliser la méthode de chi-carré qui se prête assez bien à ce type de recherches. Toutefois on a également donné ici les résultats de l'analyse de variance dont on voit sans peine qu'elle aurait plutôt conduit à sous-estimer les significances.

	d. l.	carré moyen
. . . . .	23	
e . . . . .	1	30,858
. s . . . . .	1	12,117
. . c . . . . .	2	10,682
. . . i . . . . .	1	28,185
e s . . . . .	1	0,480
e . c . . . . .	2	0,401
e . . i . . . . .	1	0,012
. s c . . . . .	2	0,673
. . s i . . . . .	1	0,171
. . . c i . . . . .	2	1,712
c s c . . . . .	2	0,221
e s . i . . . . .	1	0,454
e . c i . . . . .	2	0,032
. s c i . . . . .	2	0,094
e s c i . . . . .	2	0,642

TABLEAU I

d. l. : degré de liberté.  
e : échantillon.  
s : sexe.  
c : côté.  
i : intensité.

Les effets principaux.

Au total on compte une proportion de 0,6476 sujets masculins. Les formes bilatérales représentent une fraction de 0,2602 du total, cependant que parmi les formes unilatérales la proportion des cas à gauche est : 0,6927. Enfin la fréquence relative des formes complètes est 0,7166. Tous ces chiffres sont significativement différents de ceux qu'indiqueraient les diverses hypothèses nulles simples que l'on pourrait former a priori. On les utilisera donc sous leur forme brute pour la suite des calculs concurremment avec la fréquence relative des sujets de l'échantillon français (0,7183).

Les interactions du premier ordre.

On dispose de six comparaisons indépendantes entre elles qui conduisent aux valeurs suivantes :

Interaction	Degrés de liberté	Valeur du chi-carré
Echantillon $\times$ sexe	1	7,34
» $\times$ côté	2	1,57
» $\times$ intensité	1	0,13
Sexe $\times$ côté	2	8,03
» $\times$ intensité	1	0,30
Côté $\times$ intensité	2	32,97

Seule la dernière de ces interactions est hautement significative ; la première est significative — ( $P < 1/0,00$ ) ; l'interaction « sexe  $\times$  côté » est presque significative.

Pour analyser plus profondément l'influence du côté, on pourrait utiliser une décomposition orthogonale puisqu'il y a deux degrés de liberté à chaque fois. Il semble plus logique de tester d'abord le rapport des formes bilatérales aux formes unilatérales, puis à l'intérieur de celles-ci le rapport de la gauche à la droite.

Bien que ces deux tests ne soient pas indépendants les résultats sont ici tellement nets que cette méthode semble suffisante.

On trouve :

Interaction	Valeur du chi-carré	
Sexe $\times$ unilatérale/bilatérale..	6,28	( $P = 1,2\%$ )
Sexe $\times$ gauche/droit .....	2,93	( $P = 8,5\%$ )
Intensité $\times$ unilatérale/bilatér	32,44	
Intensité $\times$ gauche/droit ....	0,44	

Interactions du 2<sup>e</sup> ordre.

L'existence des interactions précédentes rend illusoire le calcul des interactions du 2<sup>e</sup> ordre par la méthode des décompositions orthogonales.

Mais renonçant cette fois encore à l'indépendance, on peut employer la méthode de chi-carré pour vérifier l'égalité des birapports.

L'analyse de variance indique que seule l'interaction : Echantillon  $\times$  Sexe  $\times$  Intensité mérite d'être testée. Le chi-carré correspondant est trouvé égal à 1,97.

CONCLUSIONS

A. Les deux échantillons diffèrent probablement par l'inégalité de la proportion des sexes : 62,8 % d'hommes dans l'échantillon de Veau contre 69,8 % dans celui de P. F. Andersen.

Peut-être pourrait-on aussi les soupçonner de différer quant à la gravité des formes selon le sexe (excès relatif des formes complètes chez les femmes en France et corrélativement des formes simples chez les hommes au Danemark) bien que les totaux des formes simples et complètes concordent. Une différence de localisation latérale entre les deux échantillons paraît encore moins défendable.

SEM. HÔP. PARIS  
14 AOUT 1949

STATISTIQUE DES BECS-DE-LIEVRE

— 2546 —

B. Dans l'ensemble les sexes ne se différencient que par la localisation et ceci surtout en raison de l'excès significatif des formes bilatérales chez l'homme : 27,9 % de tous les cas ; chez la femme 22,5 %.

La diminution des formes droites est ici peu nette chez l'homme : à droite 29,5 % des formes unilatérales ; chez la femme 32,9 %.

C. Enfin il existe un contraste indiscutable entre les formes complètes beaucoup plus souvent bilatérales (29,8 % contre 16,6 % pour les formes simples) qu'unilatérales sans qu'aucune préférence pour le côté gauche ou le côté droit ne puisse être mise en valeur.

(Travail subventionné par l'Institut National d'Hygiène et le Fonds d'Etudes de la Société Médicale des Hôpitaux de Paris).

## L'ETUDE DES DERMATOGLYPHES

par R. TURPIN et M. P. SCHÜTZENBERGER

L'intérêt de l'étude des dermatoglyphes humains tient à la contribution qu'elle peut apporter à la solution de nombreux problèmes de pathologie et de biologie générale.

Du point de vue médical, des travaux récents ont mis en valeur les particularités des crêtes papillaires dans certaines maladies : oligophrénie mongolienne, neurogliomatose, etc. (1, 2). Nous voyons dans ces associations les exemples d'états morbides qui méritent mieux que tous autres, peut-être, le nom de neuro-ectodermoses.

Du point de vue biologique les études dermatoglyphiques sont nées d'inspirations diverses : les plus pratiques sont l'obligation de porter un diagnostic de gémellité uni ou multi ovulaire (3), de faire la preuve d'une paternité. Ce sont là, en quelque sorte, les applications courantes de travaux poursuivis parallèlement sur le plan théorique. De nombreux chercheurs se sont efforcés, sans obtenir entière satisfaction, de mesurer la part respective de l'ambiance et de l'hérédité à l'origine de ces dessins. S'ils pouvaient, par leur hérédité génique nous apporter quelques nouveaux repères chromosomiques, nous permettre d'isoler quelques nouvelles liaisons factorielles, ils contribueraient grandement aux progrès de l'eugénique rationnelle et, par voie de conséquence, ils nous fourniraient ces tests de l'identité gémellaire et ces tests de paternité que nous recherchons.

Mais, au delà de la biologie humaine, les dermatoglyphes peuvent servir à l'étude de problèmes encore plus généraux : En premier lieu le problème de l'évolution : la seule comparaison des dermatoglyphes de l'homme et des autres primates reconstitue en partie et complète les arbres généalogiques déjà établis par d'autres méthodes d'étude de la filiation des espèces (4).

En second lieu, les problèmes de la composition des forces héréditaires et des forces ambiantes. Les dessins épidermiques sont un matériel de choix pour l'analyse des

conflits successifs dans lesquels se réalise la morphogénèse par le jeu contradictoire des facteurs individuels et de milieu plus ou moins contingents, des facteurs mendéliens dont la répartition statistique varie d'un stock de population à un autre, et, enfin des facteurs héréditaires d'une nature plus complexe et plus riche, sans doute, déterminant le plan de l'espèce et sa réalisation selon le sexe. Pour l'étude de ces interactions, les dermatoglyphes offrent des possibilités remarquables : par la constance et la précocité du moment de leur formation, par la richesse de leur expression, au doigt, au côté, au membre, où interviennent ces différentes causes et où peut les saisir l'analyse des corrélations. Enfin par leur nature même qui s'exprime dans un quantitatif qui devient toujours qualitatif et qui doit s'étudier de ce double point de vue.

C'est à l'examen de diverses questions relatives au problème de la détermination des dermatoglyphes que sont consacrés les travaux suivants. Destinés à préciser certains points préalables, ils ne constituent nullement les aboutissements de recherches définitives, mais des travaux préliminaires à d'autres recherches.

*Références.* — Etant donnée la richesse de la bibliographie, nous nous sommes bornés à citer quelques travaux parmi les plus importants parus depuis le livre fondamental de CUMMINS et MIDLO (CUMMINS H. and MIDLO C. — *Finger prints, Palms and Soles*, Philadelphia, 1943) qui contient une bibliographie relativement très complète pour les travaux antérieurs à 1940.

(1) TURPIN R. et CASPAR-FONMARTY (Mme). — La dactyloscopie des mongoliens. *Sem. Hôp.*, 13, 1945, p. 341.

(2) PENROSE L. S. — Familial studies on palmar patterns in relation to mongolism. Proc. 8<sup>e</sup> Congrès International de Génétique, Stockholm, 1948.

(3) TURPIN R. et CASPAR-FONMARTY (Mme). — Intérêt diagnostique chez les triplés des dessins des plis palmaires. *Sem. Hôp.*, 7, 1944, p. 156.

(4) MIDLO C. and CUMMINS H. — *Palmar and plantar dermatoglyphics in primates* (Philadelphia, 1942).

SÉM. HÔP. PARIS  
14 AOUT 1949

L'ETUDE DES DERMATOGLYPHES

— 2554 —

I. — LA CORRÉLATION ENTRE LES DOIGTS

La nécessité de posséder un échantillon de contrôle assez vaste nous a permis de reprendre en le complétant un travail déjà ancien de H. Waite par d'autres méthodes et sur un matériel différent (1). (500 français et 500 français au lieu de 2.000 sujets anglais masculins).

Grâce à l'obligeance de M. le Professeur Sannié nous avons pu avoir accès à un jeu d'empreintes digitales relevées par les services de l'Identité judiciaire. Bien que certains travaux semblent indiquer que la population pénitentiaire n'est pas absolument représentative de la population dite normale du point de vue des dermatoglyphes, nous avons pensé qu'en ce qui concerne l'objet de notre étude la déviation ne saurait être que négligeable. En effet, si les théories pathologiques de la criminalité s'appliquent à de nombreux cas concrets, elles ne sont pas suffisantes pour expliquer l'ensemble ni même la majorité des délits et, très vraisemblablement seule une fraction extrêmement faible de notre matériel appartient à un stock constitutionnellement différent de la population totale.

Par exemple, nous avons trouvé dans les fiches que nous avons utilisées un certain nombre de délits tels que « usage de faux papiers d'identité » qui relèvent plus de causes sociales, politiques ou simplement éthiques que de la pathologie mentale.

Ainsi les 1.000 « formules digitales » (500 hommes et 500 femmes) que nous analyserons ici semblent pouvoir constituer un assez bon échantillon des dessins épidermiques de la population contemporaine française. D'ailleurs, le service de l'Identité Judiciaire de Paris centralisant toutes les fiches de la province, nous avons éliminé, pour plus d'homogénéité, les sujets nés en dehors du territoire métropolitain français.

Pour chaque individu nous disposons donc d'un nombre de dix chiffres codant l'appartenance de chacun des dix doigts à l'une des cinq classes suivantes (2) :

- arches,
- boucles à ouverture radiale,
- boucles à ouverture cubitale (cu « u'naire »),
- tourbillons,
- dessins plus complexes (figures « composites ») comprenant aussi bien les poches latérales et centrales que les figures en S (*twisted* de Bonnevie), etc.

TABLEAU I

	G					D					
	19	305	1	105	70	19	250	0	176	55	
♂	39	284	5	108	64	22	301	2	121	54	I
♀	80	162	95	140	23	80	127	107	174	12	
♂	111	170	67	133	19	117	150	72	140	21	III
♀	48	341	6	93	12	36	344	15	97	8	
♂	91	321	10	66	12	7	351	4	65	16	V
♀	20	286	1	189	4	12	211	8	265	4	
♂	42	282	4	166	6	29	265	7	198	1	
♀	13	431	0	53	3	10	380	2	106	2	
♂	29	426	1	42	2	15	430	3	50	2	
	Arches	Boucles cubitales	Boucles radiales	Tourbillons	Figures composées	Arches	Boucles cubitales	Boucles radiales	Tourbillons	Figures composées	

Distribution des figures pour les 1.000 sujets de l'échantillon.

LA TECHNIQUE STATISTIQUE

Pour calculer les corrélations entre doigts nous avons réduit les cinq classes précédentes à trois seulement correspondant aux types de Galton : les *arches*, les *boucles* et les *tourbillons*. Sur la base de cette classification, on peut alors attribuer à chaque figure un indice « d'intensité » (Cummins et Steggerda) (3) égal à -1 pour les arches, 0 pour les boucles et +1 pour les tourbillons. Ces valeurs, pour arbitraires qu'elles soient, reflètent cependant de manière assez satisfaisante les relations qui existent entre les différents dessins et, tout au moins, l'ordre de leur complexité : c'est un phénomène constant que dans les oppositions de main à main, de sexe à sexe, de groupe ethnique à groupe ethnique, tout accroissement de la proportion des arches se traduise avant tout par une diminution de celle des tourbillons et réciproquement. Séparément,

$$\text{pour chaque sexe, nous avons alors calculé les } \frac{10 \times 9}{2} = 45$$

corrélations entre tous les doigts par la méthode de Brauvas-Pearson. Signalons que cette méthode diffère considérablement de celle de Waite qui avait utilisé : 1° une division des figures en cinq classes (arches, petites boucles, grandes boucles, tourbillons et « composites », correspondant à notre 5° classe) ; 2° le coefficient de contingence. Le parallélisme entre les résultats, que nous montrerons plus loin n'en est que plus significatif.

LES CORRÉLATIONS

Les résultats de ces calculs sont présentés dans le tableau II où à l'intersection de chaque ligne et de chaque colonne figure la corrélation obtenue, pour les hommes dans le champ supérieur droit, pour les femmes dans le champ inférieur gauche.

En raison de l'arbitraire de la double division trichotomique sur laquelle repose le calcul de ces coefficients, et du fait que la liaison entre les types de figures ne pourrait que très difficilement se rattacher à une distribution normale à *n* variables (les cumulants d'ordre élevé sont très significativement importants), nous n'avons pas calculé systématiquement la variance des corrélations obtenues. Cependant nous avons des raisons de penser que les fluctuations se traduisent par des dérivations standard de l'ordre de 0,05.

TABLEAU II

		G					D				
		V	IV	III	II	I	I	II	III	IV	V
G	V						39	36	35	36	52
	IV	45					30	40	48	54	43
	III	40	56				29	44	52	43	33
	II	37	55	60			35	55	47	37	36
	I	31	30	27	28		59	30	30	27	21
D	I	25	31	37	38	60		37	38	30	24
	II	30	50	54	73	30	36		49	34	34
	III	38	48	57	35	20	26	51		42	37
	IV	41	62	53	54	30	33	52	55		43
	V	63	39	39	32	26	26	26	36	38	

Corrélations. — ♂ : sup. droit ; ♀ : inf. gauche

« Les corrélations ont été multipliées par 100 pour des raisons de commodité typographique. — Ainsi : 60 doit être lu : 0,60 (corrélations entre les deux touches chez la femme) ».

Tout d'abord il est manifeste que toutes ces valeurs étant positives et même parfois assez élevées, la liaison entre les doigts est significative, à la fois dans son ensemble, et pour chaque paire de doigts quelconques, et qu'elle témoigne de l'existence de causes importantes agissant dans le même sens et de manière très générale pour déterminer un ressemblance entre les types d'empreintes d'un même sujet (1). Les valeurs que nous avons obtenues sont très sensiblement inférieures aux valeurs correspondantes de Waite. Sans aucun doute, cette différence tient à la technique statistique employée comme nous avons pu le vérifier en appliquant notre technique à certain nombre de tableaux originaux de Waite. Pour aller plus loin on peut faire l'histogramme des diverses corrélations. Bien que leurs erreurs probables soient difficiles à évaluer, et que, d'autre part, la valeur de ces corrélations dépende vraisemblablement de la répartition statistique marginale, différente d'un doigt à un autre, des indications peuvent être apportées qui concordent pour les deux sexes :

1° Les corrélations les plus élevées sont, dans l'ensemble, les corrélations entre doigts homologues : Chez l'homme les valeurs se rangent dans l'ordre suivant : III < V < IV < II < I, et chez la femme : III < I < IV < V < II. Sur son matériel, Waite avait obtenu : II < I < III < IV < V et il est difficile de saisir exactement la signification de ces différences. Par contre, ces corrélations semblent nettement plus élevées chez la femme que chez l'homme, phénomène d'autant plus remarquable qu'il ne s'accompagne pas, semble-t-il, d'un accroissement des autres corrélations, du moins aussi net ; la comparaison est d'ailleurs, pour ce dernier point discutable en l'absence de théorie précise sur les mécanismes causaux. Ce fait est certainement en rapport cependant avec la diminution de la différence entre les deux mains dans le sexe féminin, telle qu'elle se traduit par l'emploi de chi-carré sur la répartition de chaque doigt (tableau III).

TABLEAU III

	I	II	III	IV	V
♂	19,84	23,75	1,92	2,59	13,37
♀	5,16	4,69	3,64	0,87	5,10

Chi-carré des différences entre main droite et main gauche

2° A l'exception des corrélations précédentes les valeurs les plus hautes s'observent entre doigts immédiatement voisins, de la même main, puis du membre contralatéral.

3° Si dans chaque sexe on compare pour chaque paire de doigts la corrélation pour la paire gauche à la corrélation pour la paire droite, cette dernière est la plus faible (sauf pour le couple I/II dans nos deux échantillons mais non dans celui de Waite).

4° Les corrélations du POUCE avec les autres doigts sont dans l'ensemble systématiquement plus faibles que les autres ce qui semble assez naturel, a priori, si l'on considère la situation anatomique spéciale de ce doigt.

En conclusion, reprenant un par un les divers points du travail de Waite nous avons pu confirmer ses conclusions et les étendre aux deux sexes ; il est à signaler qu'outre une différence dans la technique statistique le matériel employé par Waite différait significativement du nôtre du point de vue de la répartition des différents types de figure

comme nous a permis de le mettre en évidence un test par chi-carré sur le nombre total des trois classes dans chaque main.

LIAISON ENTRE LES FIGURES « COMPOSITES »

La nature des documents que nous possédons nous permet de contrôler la liaison entre les figures de la dernière classe qui avaient été mises avec les tourbillons dans les calculs précédents : En raison de leur fréquence relativement faible, nous avons pu calculer seulement la corrélation entre le nombre total à droite et le nombre total à gauche (tableau IV). Les valeurs obtenues dans chaque sexe sont significatives de l'existence de causes favorisant simultanément l'apparition de ce type aux deux mains (r = 0,33 chez l'homme ; r = 0,30 chez la femme).

TABLEAU IV

	0	1	2	3	main gauche	0	1	2	3	main gauche
♂	358	60	5	0		371	50	6	0	
1	46	23	3	2		32	25	2	2	
2	0	0	0	2		6	3	1	0	
3	0	1	0	0		1	0	1	0	
main droite						main droite				
	Hommes					Femmes				

Tables de contingence pour les figures composées

LIAISON ENTRE LES BOUCLES RADIALES

Ce type de figure encore plus rare que le précédent a une distribution tout à fait remarquable puisqu'il est relativement très fréquent à l'index et extrêmement rare aux autres doigts. Il a semblé intéressant d'abord de voir si ces apparitions aux doigts I, III, IV et V étaient purement accidentelles ou bien relevaient d'une prédisposition générale, ensuite d'étudier la liaison de ces phénomènes avec la présence de ces mêmes boucles radiales à l'index.

Pour l'étude du premier problème l'hypothèse nulle est que l'apparition d'une boucle radiale à l'un quelconque des doigts sauf l'index est un phénomène indépendant, de probabilité variable selon le doigt. Les formules classiques permettent alors, en supposant que les fréquences totales observées pour chaque doigt sont égales à ces probabilités, de calculer la probabilité P<sub>0</sub>, P<sub>1</sub>, P<sub>2</sub>, P<sub>3</sub>, etc. pour qu'un individu ait zéro, une, deux, ou trois boucles radiales. Les chiffres calculés montrent que cette hypothèse doit être rejetée au profit d'une liaison entre boucles radiales aux doigts I, III, IV et V. (Tableau V).

TABLEAU V

	Valeurs calculées		Valeurs observés	
	♂	♀	♂	♀
P <sub>0</sub>	0,93550	0,93011	472	470
P <sub>1</sub>	0,06301	0,06780	23	25
P <sub>2</sub>	0,00147	0,00205	5	4
P <sub>3</sub>	0,00001	0,00003	0	1
P <sub>4</sub>	0,00000	0,00000	0	0

Test d'indépendance pour les boucles radiales

En ce qui concerne la liaison entre boucles à l'index et boucles à un autre doigt, les chiffres sont trop faibles pour

SEM. HÔP. PARIS  
14 AOUT 1949

L'ETUDE DES DERMATOGLYPHES

— 2556 —

qu'aucune corrélation significative ait pu être établie. Enfin en considérant pour chaque sexe la table  $2 \times 2$  (boucles radiales/ulnaires  $\times$  index droit/gauche), on obtient des chi-carrés de 15,14 et 17,16 respectivement qui établissent que, même compte tenu de la prédisposition générale aux boucles, il existe entre le caractère radial ou cubital de celles-ci une association certaine aux index.

## CONCLUSION

Nous avons repris un travail de Waite sur un nouveau matériel et avec d'autres méthodes. Dans l'ensemble, nos conclusions confirment celles de cet auteur. En outre la comparaison entre les deux sexes qui n'avait pas encore été effectuée de ce point de vue, nous a montré un ac-

croissement chez la femme de la liaison entre doigts homologues. Enfin, nous avons établi l'existence d'une association entre l'apparition aux différents doigts de certains type de figures plus rares.

(Travail subventionné par l'Institut National d'Hygiène et le Fonds d'Etudes de la Société Médicale des Hôpitaux de Paris).

(1) WAITE H. — Association of finger-prints. *Biometrika*, vol. 10 (1915), pp. 421-478.

(2) SANNIÉ G. — Eléments de police scientifique, pp. 45-46. Paris (1939).

(3) CUMMINS H. and STEGGERDA M. — Finger Prints in a dutch family series. *Am. J. Phys. Anthr.*, vol. 20 (1935), pp. 19-41.

## II. — FACTEURS LATÉRAUX ET FACTEURS SEXUELS DANS LA MORPHOLOGIE DES EMPREINTES DIGITALES

De nombreux travaux ont montré que la proportion des différents types d'empreintes digitales variait de manière considérable d'un groupe ethnique à un autre. Il n'y a là rien de surprenant si l'on admet que leur détermination est sous le contrôle, tout au moins partiel, d'une série de gènes dont la distribution doit varier d'un stock de population à un autre. Mais la simple constatation de cette diversité ne permet pas d'aller bien loin et même, au contraire, risque de conduire à des « hiérarchisations » abusives entre les isolats plus ou moins grande que l'on peut considérer. Par contre, l'étude comparative de ces groupes possédant des répartitions statistiques différentes de gènes, même si ceux-ci sont encore très imparfaitement connus est un moyen puissant d'analyse. Car, au delà des divergences on retrouve les constantes anatomiques plus cachées de l'espèce humaine et les associations entre particularités morphologiques que ne permettrait pas de mettre en évidence l'emploi d'un matériel trop homogène. Nous avons essayé ici de donner quelques indications sur les rapports qui existent entre les facteurs latéraux et le phénotype sexuel, sans nous dissimuler le caractère provisoire de ce travail.

## LES SOURCES DOCUMENTAIRES.

Nous avons recherché dans la littérature les répartitions d'empreintes aux différents doigts pour le plus grand nombre possible de populations. Pour diverses raisons (qui seront exposées plus bas), nous avons dû renoncer à utiliser un certain nombre de travaux, ainsi que ceux qui ne réunissaient qu'un nombre trop faible de sujets. Enfin deux mémoires importants (Maasland sur les Tobabatacs et Kubo sur les peuples de la Corée, de la Chine et du Japon) n'ont encore pu être consultés. En définitive, nous avons pu ainsi rassembler des chiffres complets sur six échantillons comprenant au moins 400 individus de chaque sexe : les danois (1), les américano-indiens (2), les grand-russes (3), les javanais (4), les belges et les hollandais (5) auxquels nous avons adjoint un échantillon de 500 français et 500 françaises recueilli par nous-mêmes grâce à l'obligeance de M. le Professeur Sannié, Directeur de l'Identité judiciaire à Paris qui a bien voulu nous ouvrir certains de ses fichiers. Bien que l'ensemble de ces échantillons ne puisse évidemment prétendre à être représentatif de quoi que ce soit, il nous semble cependant assez étendu pour qu'une cer-

taine valeur de généralité puisse être attachée aux conclusions que l'on en pourra tirer.

## LA MÉTHODE STATISTIQUE.

Ainsi que l'ont signalé tous les auteurs (6), les répartitions globales varient de manière très significative d'une population à une autre, et ceci d'autant plus que certaines différences de classification viennent encore s'ajouter aux autres causes.

D'autre part, à l'intérieur d'une même population, et ceci pour chaque type d'empreinte, la répartition est extrêmement inégale, selon le doigt, le sexe, le côté. Ainsi donc, en l'absence d'une théorie qui donnant une signification à certaines moyennes de pourcentages nous autorise à les effectuer, l'usage de celles-ci est extrêmement discutable dès que des recherches un peu fines doivent être entreprises et l'on doit regretter que de nombreuses publications ne contiennent que des totaux, qui, même s'ils ne portent que sur un petit nombre de valeurs initiales observées, diminuent considérablement l'intérêt des résultats. Par exemple : le rapport du nombre des arches à l'index au nombre correspondant à l'auriculaire peut varier entre 4 et 20 dans une même population selon le sexe considéré, et on voit sans peine que la signification du « nombre total d'arches dans un certain sexe » est au moins problématique.

Aussi, tenant compte de ce fait et également de ce que la diversité du volume de nos échantillons aurait risqué de masquer les phénomènes les plus intéressants, nous avons préféré employer une approche peut-être plus sommaire mais plus exacte en définitive pour mettre en évidence l'influence des facteurs latéraux et de sexe.

Pour cela, à l'intérieur de chaque population, pour chaque doigt et pour chaque type de figure, nous avons rangé par ordre de grandeur les quatre fréquences GF (à gauche chez les femmes), GM (à gauche chez les hommes), DF (à droite chez les femmes), et DM (à droite chez les hommes) et effectué la somme des rangs sur les six populations. Grâce aux tables calculées par M. G. Kendall et B. Babington Smith (7), il est alors possible de tester l'accord des classements contre l'hypothèse nulle qui impliquerait une distribution « au hasard » de ces valeurs. On obtient ainsi un premier résultat qui nous semble remarquable à savoir que malgré la variabilité des proportions des différents types de figure il existe pour la distribution à chaque doigt

de chacun d'eux un accord presque complet entre toutes les populations (cf tableau I).

des tourbillons dont c'est l'augmentation à droite qui est la plus nette.

TABLEAU I

	Arches	Boucles radiales	Boucles cubitales	Tourbillons
I	163,50 H	114,00 S	162,00 H	127,50 H
II	64,5 N	170,00 H	91,50 P	123,50 S
III	146,00 H	145,50 H	146,00 H	123,50 S
IV	170,00 H	163,50 H	170,00 H	180,00 H
V	170,00 H	89 P	126,50 H	180,00

Dans chaque case la somme des carrés des déviations et leur probabilité codée par :

H :	prob.	<	1/1000
S :	prob.	<	1/100
P :	prob.	<	5/100
N :	prob.	>	5/100

Les effets principaux.

Mais la méthode précédente nous permet d'aller plus loin. Désignons par (df), (gf), (dm) et (gm) les rangs des fréquences pour un certain doigt et un certain type de figure. Les expressions

$$(D) = (dm) + (df) - (gm) - (gf) \text{ et } (M) = (dm) + (gm) - (df) - (gf)$$

sont manifestement des mesures de l'influence des facteurs « côté droit » et « sexe masculin » respectivement. Dans l'hypothèse nulle il est facile de calculer la probabilité pour qu'une de ces expressions (D), (M) ou (I) = (dm) + (gf) - (gm) - (df) (que nous utiliserons plus loin) excède une valeur x. On a :

- x plus petit que 13 : p plus grande que 5 % ;
- x entre 13 et 17 : p entre 5 % et 1 % ;
- x entre 17 et 21 : p entre 1 % et 0,1 % ;
- x plus grand que 21 : p plus petite que 0,1 %.

(D), (M) et (I) ne sont pas rigoureusement indépendantes mais, en pratique on pourra les considérer comme telles en première approximation puisque les covariances et tous les cumulants du troisième ordre de leur distribution simultanée sont identiquement nuls. Les tableaux II et III mettent en évidence les résultats suivants :

Les facteurs latéraux agissent surtout sur les tourbillons (prédominance à droite sauf au médius). Ils agissent également sur les arches (prédominance constante à gauche). En ce qui concerne les boucles cubitales la prédominance est aussi à gauche sauf pour le médius dont l'excès à droite est particulièrement remarquable. Les boucles radiales semblent prédominer à droite pour les trois doigts du groupe radial et à gauche pour les deux doigts du groupe cubital.

Les facteurs de sexe ont une action très certaine sur les arches et les tourbillons, favorisés respectivement chez les femmes et chez les hommes. Les boucles cubitales sont elles aussi plus fréquentes chez les femmes. En ce qui concerne les boucles radiales la localisation au pouce (plus fréquente chez les femmes) s'oppose de façon singulière à la localisation à l'index (plus fréquente chez les hommes). On doit remarquer enfin que les arches semblent être surtout sensibles aux facteurs de sexe (excès chez les femmes) au contraire

TABLEAU II

	Arches	Boucles radiales	Boucles cubitales	Tourbillons
I	22	16	0	-7
II	13	-24	7	-7
III	16	-2	2	-18
IV	22	-1	10	-12
V	22	+4	9	-18

ACTION DU FACTEUR F

	Arches	Boucles radiales	Boucles cubitales	Tourbillons
I	13	10	18	-19
II	5	-2	14	-21
III	18	7	-18	7
IV	14	-22	24	-24
V	14	-18	19	-18

ACTION DU FACTEUR G

	Arches	Boucles radiales	Boucles cubitales	Tourbillons
I	-1	-10	18	-10
II	-8	-10	11	-2
III	-2	-23	16	-11
IV	0	-13	2	0
V	0	-4	8	0

Les interactions.

Si l'on se bornait à utiliser les fréquences brutes, l'étude des interactions serait elle aussi très difficile. A l'annulaire par exemple l'ordre de fréquence relatif des tourbillons est à peu près constamment le suivant :

GF (gauche femme) < GM (gauche homme) < DF (droite femme) < DM (droite homme). En poussant plus loin l'analyse on constate que ces différences sont significatives dans chaque échantillon dès que celui-ci est assez large. En outre, considérant les quatre fréquences comme les quatre valeurs d'un tableau 2 x 2, un test par chi-carré révèle l'hétérogénéité des deux sexes dans les échantillons les plus importants.

Autrement dit, l'on ne peut pas admettre que la probabilité pour qu'apparaisse un tourbillon à l'annulaire soit la même à droite qu'à gauche, ni qu'elle soit la même chez les hommes et chez les femmes ; on ne peut pas admettre non plus que le rapport d'un côté à l'autre soit le même dans les deux sexes. Mais, précisément, en ce qui concerne ce dernier test par chi-carré de l'interaction, il s'introduit de par son seul emploi l'hypothèse que les effets latéraux et de sexes se combinent comme un produit de probabilité. Or, c'est ce seul fait que nous ayons testé et il est difficile de conclure de ce non agrément avec un mode spécial de composition (ou mieux d'additivité) à une différence d'asymétrie entre les sexes. En effet, l'emploi d'une autre méthode analytique pourrait dans certains cas et spécialement ici où les différences sont faibles (quelques pour cent) de même que leurs erreurs conduire à des résultats entièrement opposés : par exemple si l'on avait utilisé la transformation préalable  $\text{Arcsin } \sqrt{x}$  ou  $\sqrt{x+0,5}$  pour effectuer une analyse de variance de l'ensemble des pourcentages. Mais il existe des cas qui ne se prêtent pas à cette objection : ce

SEM. HÔP. PARIS  
14 AOUT 1949

L'ETUDE DES DERMATOGLYPHES

— 2558 —

sont ceux où l'asymétrie est telle qu'elle subsiste pour n'importe quelle transformation de variable, ou du moins pour la classe de celles qui respectent les relations d'ordre initiales entre pourcentages (c'est-à-dire qui sont monotones croissantes ou décroissantes).

Cette remarque a été faite par J. B. S. Haldane (8) et nous suivrons ici sa méthode qui lui a servi à discuter les effets relatifs de l'hérédité et du milieu.

## INTERACTIONS

D'un point de vue abstrait l'ordre des fréquences dans le tableau carré que nous évoquons plus haut est indifférent parce qu'il n'y a que 24 classements possibles des quatre fréquences et parce que ce nombre est précisément l'ordre du groupe d'automorphisme du groupe abélien principal d'ordre quatre qui schématise cette double classification simultanée. Mais, concrètement, on doit distinguer les principes classificateurs : d'une part il y a des facteurs de sexe et latéraux, d'autre part, leur interaction. Celle-ci correspond au fait que dans certains cas les fréquences GM et DF par exemple sont plus grandes que les fréquences GF et DM, or si les effets principaux jouaient seuls et indépendamment (par exemple Droit et Masculin) la fréquence correspondant à GF devrait être la plus faible.

On aura donc trois types principaux de tableaux :

ceux où les cases d'une même diagonale ont respectivement les rangs 1 et 4 auquel aucune interaction ne peut être établie par cette méthode ;

ceux où les cases d'une même diagonale ont respectivement les rangs 1 et 3 : il y a une interaction qui se traduit soit par un renversement de l'asymétrie d'un sexe d'une main à l'autre ;

enfin ceux où les deux plus fortes fréquences sont dans la même diagonale, et ici on pourrait dire que l'effet le plus important est l'interaction.

Pour appliquer cette méthode il est donc parfaitement justifié d'utiliser la troisième composante orthogonale, sur les totaux des six populations.

$$I = (dm) + (gf) - (df) - (gm).$$

Des phénomènes remarquables apparaissent :

Sauf pour les boucles cubitales où l'interaction se traduit par une augmentation des fréquences (dm) et (gf), ce sont les fréquences de la diagonale (gm), (df) qui sont plus élevées qu'elles ne devraient l'être si les facteurs de sexe et de main agissaient indépendamment et ceci est surtout vrai pour les boucles radiales. D'autre part, il semble bien que les doigts du groupe externe soient l'effet d'interactions plus fortes que ceux du groupe interne. Enfin, le phénomène que nous avons déjà signalé (9) (renversement complet d'asymétrie selon le sexe pour les boucles radiales du médium) est mis en valeur de manière particulièrement nette.

## III. — RECHERCHES SUR LA VALEUR QUANTITATIVE

## La théorie de l'hérédité de la valeur quantitative de Bonnevie

On appelle « valeur quantitative » (V. Q.) d'une empreinte le nombre de crêtes séparant deux points caractéristiques de celles-ci : le centre de figure et un triradius.

D'après Bonnevie (1), trois facteurs héréditaires, mendéliens indépendants à dominance incomplète concourraient à sa détermination :

— un facteur principal (V, v.), les homozygotes récessifs

Il est possible de compléter cette analyse par l'examen des cas où les méthodes précédentes ne donnent que des résultats incertains avec l'aide des procédés classiques de comparaison des rapports de fréquences. Dans l'ensemble, on retrouve les résultats précédents mais certains autres apparaissent qui mériteraient une étude plus poussée : les interactions aux deux doigts du groupe cubital, pour les tourbillons et en partie pour les arches semblent être plutôt dans le sens (DM), (GF) ainsi d'ailleurs que l'interaction au pouce pour les arches.

## CONCLUSION

Nous avons montré que la répartition statistique des dermatoglyphes varie selon le sexe et le côté, mais de telle sorte que l'ordre des fréquences dans les quatre cas qui résultent de cette double dichotomie reste constant et constitue une caractéristique et du doigt et du type d'empreinte envisagé. En même temps, nous avons mis en évidence l'existence d'interactions entre ces facteurs. La permanence de ces effets est d'autant plus remarquable que d'une population à une autre la répartition brute des différents types d'empreintes varie d'une façon extrêmement importante. Malheureusement, nos conclusions ne reposent que sur un échantillonnage très faible de la population du globe et où de nombreux groupes ethniques ne sont pas représentés. Nous espérons que des séries de dactylogrammes recueillis dans des conditions systématiques et couvrant des domaines plus vastes permettront d'approfondir ce phénomène et, par suite, d'apporter quelque lumière sur les processus de développement neuro-ectodermique et leur vitesse relative.

## REFERENCES

- BUGGE J. N. (1932). — Sur la statistique des empreintes digitales (en danois). *Meddel. om Danmarks Antrop.*, vol. 3, pp. 387-392.
- HENCKEL K. O. (1933). — Beiträge zur Anthropologie Chilies. *Ztschr. für Morph. und Anthrop.*, vol. 34, pp. 290-309.
- SEMEŃOVSKI P. S. (1927). — Sur la distribution des différents types, etc... (en russe). *Russkij antropologičeskij žurnal*, vol. 16, pp. 7-26.
- DANKMEIJER J. (1938). — Some anthropological data on finger prints. *Am. J. Phys. Anthrop.*, vol. 23, pp. 377-388.
- PIEBENGA H. T. (1938). — Systematische und erbbiologische Untersuchungen. *Ztschr. für Morph. und Anthrop.*, vol. 37, pp. 140-165.
- Voir à ce sujet une bibliographie très complète dans la thèse : DANKMEIJER J. : La signification des empreintes digitales en anthropologie (en néerlandais). Utrecht (1934).
- KENDALL M. G. et BABINGTON-SMITH B. (1939). — The problem of m-rankings. *Ann. of Math. Stat.*, vol. 10, p. 275.
- HALDANE J. B. S. (1946). — The interaction of nature and nurture. *Ann. of Eugenics*, vol. 13, pp. 197-206.
- TURPIN R. et SCHÜTZENBERGER M. P. (1949). — Sur une asymétrie latérale statistique, attribut du phénotype sexuel humain. *C. R. Acad. Sci.*

(Travail subventionné par l'Institut National d'Hygiène et le Fonds d'études de la Société Médicale des Hôpitaux de Paris).

se reconnaissent à une valeur quantitative élevée pour le doigt où elle est maximum (dépassant 22).

— deux facteurs secondaires, les homozygotes dominants se reconnaissent à un abaissement important de la V. Q. à l'un ou plusieurs des trois doigts radiaux (couple R, r) ou des deux doigts cubitales (ou ulnaires) (couple U, u).

Nous avons repris trente deux paires de jumeaux identi-

ques étudiées par Bonnevie plus une paire donnée par Enenbach (2) et nous avons cherché si par l'emploi de l'analyse de variance, nous ne pourrions pas préciser la portée de cette théorie.

#### LA MÉTHODE STATISTIQUE

A priori, les facteurs au sens le plus large du terme qui déterminent la V. Q. peuvent être classés en :

1° facteurs généraux valables pour tous les sujets, véritables facteurs caractéristiques de l'espèce humaine.

2° facteurs familiaux héréditaires ou congénitaux variant d'une paire à une autre paire de jumeaux.

3° facteurs individuels responsables des différences entre individus d'une même paire gémele.

A l'intérieur de chacune de ces classes, il convient de distinguer : les facteurs de mains valables pour les cinq doigts d'une main et opposant celle-ci au membre controlatéral.

Les facteurs de paires de doigts influençant simultanément deux doigts homologues (par exemple les pouces) et les différenciant d'une autre paire d'homologues (par exemple les index).

Les facteurs de doigt unique dont l'action se limite à un seul doigt (par exemple médus gauche ou auriculaire droit).

La liste des facteurs admissibles a priori, dont notre documentation permet l'analyse, est donc :

##### 1° Facteurs généraux :

m : facteur de main droite ou gauche (m = 1 ou 2).  
d : facteur de paire de doigts (d = I, II, III, IV ou V).  
md : facteur de doigts.  
(1 1, 1 II, 1 III, 1 IV, 1 V)  
(2 1, 2 II, 2 III, 2 IV, 2 V)

2° Facteurs familiaux (congénitaux ou héréditaires, nos documents ne permettent pas de trancher ce point) :

f : facteur s'étendant à tous les vingt doigts du couple gémele.  
fm : facteur familial de mains s'étendant aux deux mains homolatérales d'une même paire gémele.  
fd : facteur familial de paires de doigts.  
fmd : facteur familial de doigt.

##### 3° Facteurs individuels :

fj : facteur individuel (s'étendant sur dix doigts d'un même individu).  
fjm : facteur individuel de main.  
fjd : facteur individuel de doigts.  
fjmd : facteur individuel de doigt.

Du point de vue de la technique statistique, mentionnons enfin que les catégories fj, fjm, fjd et fjmd doivent être entendues comme étant les sommes fj + j, fjm + jm, fjd + jd et fjmd + jmd respectivement puisque aucune signification ne peut être attachée à l'ordre arbitraire dans lequel ont été numérotés les individus à l'intérieur de chaque paire gémele.

Il n'est peut être pas inutile d'expliquer la signification biologique de ces diverses causes.

Les facteurs généraux représentent la « physionomie propre » de chaque doigt dans l'espèce humaine ou plutôt dans la population envisagée.

Le facteur familial « f » correspond au facteur principal de Bonnevie (ou à un système plus complexe de gènes de même fonction).

Le facteur familial « fd » traduit le fait que l'abaissement (ou l'élévation) de la V. Q. à telle ou telle paire de doigts a une cause héréditaire ou congénitale.

Le facteur familial « fm » s'interpréterait par l'hérédité

d'un abaissement (ou élévation) systématique de la V. Q. à droite ou à gauche.

La signification de « fmd » est assez claire d'après ce qui précède.

L'importance relative des facteurs individuels fj, fjm, fjd, est une indication du rôle relatif des mécanismes centraux et périphériques.

Signalons que « fm » devrait ressortir nettement si des « jumeaux en miroir » faisaient partie de notre matériel puisqu'il correspond à une *asymétrie opposée* des deux individus d'une même paire gémele.

En ce qui concerne « fjmd », enfin, il est clair que ce dernier facteur est en quelque sorte la mesure de la fluctuation irréductible de la V. Q. Il totalise toutes les influences incontrôlables et insystematisables (dans les limites que fixe la structure même des documents que nous utilisons) qui ont pu s'exercer pour dévier la V. Q. d'un certain doigt, d'une certaine main, d'un certain individu, de la valeur que leur assignaient les facteurs d'ordre hiérarchique supérieur. C'est donc par rapport à son effet moyen que se contrôlera l'étendue de l'action de ces derniers.

L'analyse de variance se propose de fixer la part que joue chacun de ces facteurs éventuels.

Pour cela, une certaine grandeur (la *variance*) mesurant l'hétérogénéité totale des documents est calculée puis répartie entre les différents facteurs qu'indique l'analyse logique préalable qui vient d'être effectuée.

On peut ainsi *tester* la signification de la différence des moyennes des V. Q. des empreintes appartenant à chacune des catégories correspondant aux divers facteurs. C'est-à-dire que l'on cherchera à savoir si l'imputation de ces écarts à des fluctuations aléatoires indépendantes ne conduit pas à admettre qu'il s'est produit un événement de probabilité extrêmement faible, auquel cas, l'on soupçonnerait l'existence d'une cause systématique (l'un des facteurs prévus a priori) altérant le mécanisme probabilitaire initial.

L'aspect analytique de cette technique statistique est très simple et l'on se bornera pour un exposé moins sommaire à renvoyer aux chapitres qui sont consacrés à ce problème classique dans tous les traités modernes.

L'hétérogénéité totale est mesurée par la somme des carrés des écarts de la V. Q. à la moyenne générale. Et, de même, que l'inertie d'un solide autour d'un point fixe peut être fractionnée en ses composantes le long de chaque axe, cette somme de carrés est répartie entre diverses formes quadratiques indépendantes correspondant aux différences de moyenne que suppose l'existence des facteurs précités. On compare ensuite ces expressions par le test F de Snedecor (ou z de Fischer), après avoir introduit le nombre de comparaisons indépendantes sur lesquelles reposent ces estimations (nombre de *degrés de liberté*).

Pour que ces derniers calculs soient rigoureux, un certain nombre de conditions doivent être remplies traduisant le caractère « gaussien » des fluctuations résiduelles. Nous n'avons ici que bien peu d'arguments à apporter en faveur de ces hypothèses. Cependant de nombreux travaux théoriques ont montré que les résultats obtenus en traitant les données brutes comme si elles satisfaisaient aux conditions de normalité avaient une stabilité assez grande par rapport aux changements de distribution statistique.

En outre, les hypothèses de base de l'analyse de variance comprennent l'égalité de la variance des fluctuations.

Pour l'ensemble des documents cette condition n'est pas remplie et il serait vraisemblablement nécessaire d'introduire une transformation de la V. Q. pour s'y ramener. Cependant, étant donné le faible nombre de paires de jumeaux

SFM. HÔP. PARIS  
14 AOUT 1949

L'ETUDE DES DERMATOGLYPHES

— 2560 —

TABLEAU I

FACTEURS	GÉNOTYPES										Toutes les paires : N = 33
	VVRRUU N = 2	VvRRUU N = 9	VvRrUu N = 3	VvRvUu N = 2	VvRRUU N = 3	vVRUUU N = 2	vVRrUu N = 2	vVRvUu N = 4	vVRRUU N = 4	vVRRUU N = 4	
	d.l.	d.l.	d.l.	d.l.	d.l.	d.l.	d.l.	d.l.	d.l.	d.l.	d.l.
<b>Généraux</b> $\left. \begin{matrix} m \\ d \\ md \end{matrix} \right\}$	1 0,60 4 8,75 4 6,35	1 1,80 4 350,70 4 2,58	1 14,01 4 70,44 4 4,71	1 10,00 4 29,38 4 3,25	1 58,01 4 308,30 4 24,73	1 0,60 4 112,21 4 21,81	1 5,61 4 35,41 4 6,66	1 12,01 4 67,91 4 3,20	1 12,01 4 67,91 4 3,20	1 12,01 4 67,91 4 3,20	1 54,21 4 774,59 4 11,02
<b>Familiaux</b> $\left. \begin{matrix} f \\ fm \\ fd \\ fmd \end{matrix} \right\}$	1 0,60 1 1,20 4 6,38 4 1,20	8 203,30 8 8,98 32 572,25 32 10,91	2 646,87 8 6,47 8 3,30	1 176,40 1 10,10 4 18,53 4 5,48	2 101,51 2 8,31 8 24,89 8 16,45	1 7,23 1 0,03 4 16,41 4 3,26	1 148,23 1 0,60 4 3,41 4 4,06	3 32,28 3 25,25 12 3,30 12 9,27	3 32,28 3 25,25 12 3,30 12 9,27	3 32,28 3 25,25 12 3,30 12 9,27	32 586,37 32 7,85 128 42,40 128 8,19
<b>Individuels</b> $\left. \begin{matrix} fj + j \\ fim + jm \\ fj + fd + jmd \end{matrix} \right\}$	2 0,93 9 5,02 36 0,98	9 17,00 9 5,02 36 0,98	3 6,85 3 1,21 12 8,35	2 0,45 2 3,25 8 5,68	3 15,41 3 3,29 12 11,38	2 4,23 2 1,13 8 2,16	2 0,23 2 13,23 8 2,91	4 0,74 4 2,59 16 1,80	4 0,74 4 2,59 16 1,80	4 0,74 4 2,59 16 1,80	33 7,75 33 3,73 132 0,26
<b>Incontrôlables</b> $\left. \begin{matrix} fj + jmd \\ fj + fd + jim \\ fj + fd + jmd \end{matrix} \right\}$	8 0,56 20 0,78	36 10,02 90 10,07	12 5,13 30 6,22	8 1,50 20 3,20	12 9,83 30 10,03	8 3,31 20 2,73	8 4,66 20 4,38	16 2,84 40 2,18	16 2,84 40 2,18	132 14,31 330 6,98	

d. l. : Nombre de degrés de liberté.  
N : Nombre de paires de jumeaux.  
m. s. : Déviation standard.

X significativement plus grand que « i »  
X X » au niveau 5 %  
? » 1 %  
? petit » 5 %  
? » » 1 %

envisagés dans cette étude préliminaire, on s'est contenté d'opérer sur la V. Q. brute. Enfin, la classification préalable adoptée semble atténuer cette source d'erreurs ; en effet, pour étudier les documents, nous les avons classés d'après les génotypes des individus *déjà établis* par Bonnevie et effectué l'analyse pour chacun des génotypes représentés par plus de deux paires de jumeaux. Si donc, la théorie génétique qu'elle a conçue se trouvait vraie, il ne devrait rester de significatifs que les facteurs généraux et, éventuellement, les facteurs individuels, les facteurs familiaux disparaissent de par la classification elle-même.

DISCUSSION DES RÉSULTATS

On trouve les résultats suivants (tableau I).

La comparaison des facteurs individuels au terme résiduel (fluctuation irréductible) fjmd, montre que leur effet ne peut en être distingué tout au moins sur le petit nombre de cas que contient notre échantillon. Conformément à l'habitude, il convient donc de regrouper les termes fj, fjm, fjd et fjmd en une seule catégorie : i, qui contient tous les facteurs individuels et qui est connue avec une précision supérieure à celle de fjmd.

On peut alors avec plus de sûreté comparer les autres facteurs à i c'est-à-dire établir le rapport entre chaque cause générale et familiale et l'influence moyenne de tous les facteurs purement individuels.

Manifestement le facteur « d » a une action différentielle ; certainement aussi les facteurs « f » et « fd ».

En effet, soit qu'on la compare à la variance entre paires, soit à la fluctuation irréductible, la variance imputable à cette catégorie est significativement grande pour les génotypes VVRRUU, VvRRUU, VvRRUu, VvRrUU, vvRRUU, vvRrUU, c'est-à-dire dans presque tous les cas.

Compte tenu de la remarque que nous faisons plus haut, il est difficile de s'avancer plus, mais ceci est déjà ce que nous voulions prouver.

*Aux delà de la théorie de Bonnevie, de deux facteurs commandant la V. Q. des doigts des groupes radiaux et cubitiaux respectivement, il semble bien qu'il faille envisager l'existence d'une série plus complexe de facteurs familiaux agissant sur la V. Q. de chacune des paires de doigts.*

De même la V. Q. moyenne est sous la dépendance d'un facteur faisant apparaître des différences importantes à l'intérieur des groupes de même génotype Bonnevien.

En ce qui concerne cette dernière source de variation, il est peut-être difficile d'y voir nécessairement une origine entièrement héréditaire (et non due à l'ambiance commune intra-utérine) alors que pour « fd » l'extension élective à quatre doigts (2 droits et 2 gauches à l'intérieur d'une paire gémellaire) plaide en faveur d'une cause plus « centrale » que ne le pourrait être une influence de milieu.

Pour aller plus loin nous avons rassemblé ces matériaux en une analyse unique portant sur toutes les paires à notre disposition (32 de Bonnevie plus 1 d'Ennenbach ; tableau II).

TABLEAU II

	RR Rr rr		RR Rr rr		RR Rr rr
UU	2 . .	UU	9 . .	UU	3 2 .
Uu	. . .	Uu	3 2 .	Uu	1 2 .
uu	. 1 .	uu	1 1 1	uu	1 4 .
	VV (3 paires)		Vv (17 paires)		vv (13 paires)

Répartition par génotype de Bonnevie des 33 paires gémellaires

Evidemment « f », « d », « fd » sont significatifs ce qui ne nous apporte rien de nouveau, mais aussi « m » ce qui montre l'existence d'une asymétrie entre les deux mains d'importance à peu près égale pour tous les individus puisque « fm » et « fjm » ont une influence très faible.

La faible action de ces facteurs potentiels ainsi que de « fmd » nous semble être en rapport avec un fait connu : celui de la non hérédité de l'expressivité latérale ; ici comme ailleurs, l'asymétrie existe, mais sans que s'héritent des particularités à gauche comme opposées à des particularités à droite (à l'exception de la sénestralité).

Enfin « md » n'est pas significatif pour l'ensemble des diverses analyses de variance et il semble que l'on puisse voir là le fait qu'en ce qui concerne la V. Q. (et dans la limite du petit nombre de cas étudiés), les effets des facteurs latéraux (de main) et des facteurs de paire d'homologues s'additionnent assez exactement.

Traduits en différences de moyennes, les résultats conduisent à fixer à  $+ 0,58 \pm 0,14$  l'excès du nombre de crêtes de chaque doigt de la main droite sur son homologue gauche ; pour chaque paire les déviations sont ici :

I :	.....	+	2,78	± 0,77
II :	.....	-	2,28	»
III :	.....	-	1,86	»
IV :	.....	+	2,43	»
V :	.....	-	1,07	»

Si l'on cherche par un calcul (dont la signification est surtout formelle) à chiffrer la part relative des trois classes de facteurs dans l'hérédité des V. Q. on trouve :

Facteurs généraux (ou d'espèce) .....	10 %
Facteurs familiaux .....	82 %
Facteurs individuels .....	8 %

Ceci revient à dire qu'alors que la déviation standard de la V. Q. d'une empreinte quelconque prise dans les 660 empreintes que contiennent nos documents était de 6,86 ; elle n'est plus en moyenne que de 6,54 lorsqu'on la considère dans une sous-population de 66 doigts homologues et de même côté (c'est-à-dire après élimination des facteurs généraux). Enfin cette déviation tombe à 2,64 lorsque sont éliminés à leur tour les facteurs familiaux. Et cette valeur rejoint assez bien, en la précisant, la limite de 4 que Bonnevie attribuait à l'influence des causes incontrôlables dans son analyse du même matériel.

Il resterait à expliquer la valeur extraordinairement faible qu'ont les chiffres correspondant à « fjm » et surtout « fjd ».

Biologiquement on serait tenté de faire appel à un mécanisme tel que la particularité génétique « abaissant de la V. Q. d'une paire de doigts », se manifeste de façon aléatoire à droite ou à gauche et plus rarement aux deux mains simultanément.

Il semble cependant plus prudent étant donné qu'à une exception près (VvRRUU) le phénomène n'apparaît que dans l'analyse simultanée des 33 paires, de rechercher une raison statistique. On est alors conduit à suspecter cette « hyponormalité » de provenir de corrélations entre les diverses influences et surtout de l'inégalité de la variance résiduelle dans les divers génotypes, ce qui expliquerait également assez bien le cas de VvRRUU où les neuf paires constituent un stock fortement hétérogène.

CONCLUSIONS

Reprenant le matériel classique de Mme K. Bonnevie, nous avons cherché à l'exploiter plus complètement par

SEM. HÔP. PARIS  
14 AOÛT 1949

L'ANALYSE SEQUENTIELLE

— 2562 —

l'emploi de techniques statistiques modernes. Cette analyse nous a permis de retrouver les faits suivants :

- a) une différence constante entre les moyennes de la valeur quantitative aux deux mains;
- b) une différence constante entre les moyennes chaque paire de doigts homologues, et de faire apparaître à côté et au-delà des trois facteurs mendéliens *V*, *R* et *U* de Bonnevie;
- c) il existe des facteurs plus complexes héréditaires ou ambiants intra-utérins, ou mixtes responsables de l'abaissement de la *V. Q.* sur l'ensemble des dix doigts et sur chacune des paires de doigts homologues.

REFERENCES

- (1) BONNEVIE K. — Was lehrt die Embryologie de Papillarmuster über ihre Bedeutung als Rassen- und Familiencharacter ? III Zur Genetic des quantativen Wertes der Papillarmuster. *Ztschr. für ind. Abst. und Vererbungslehre*. Vol. 59 (1931), pp. 1-60.
- (2) ENNEBACH S. — Fingerabdrücke bei ein- und zweieigen Zwillingen. *Ztschr. für mensch. Vererb. und Konstitutionslehre*. Vol. 23 (1939), pp. 555-586.

(Travail subventionné par l'Institut National d'Hygiène et le Fonds d'études de la Société Médicale des Hôpitaux de Paris).

## UNE APPLICATION DE L'ANALYSE SEQUENTIELLE

par M. P. SCHÜTZENBERGER

La nécessité d'améliorer et surtout d'accélérer le contrôle des fabrications industrielles au stade de la production en grande série est à l'origine de la découverte de l'analyse séquentielle par Abraham Wald.

Mais en dehors de ce domaine déjà vaste, il existe de nombreux autres problèmes où cette méthode pourrait être appliquée avec profit. C'est pour attirer l'attention sur sa facilité d'emploi et sur ses divers avantages que l'on a essayé d'en décrire ici une utilisation particulièrement typique.

## LE PROBLÈME STATISTIQUE

Une série de substances doit être testée afin de retenir les plus actives pour des recherches ultérieures plus approfondies. Aucune de ces drogues n'a un effet absolument constant et chacune d'elles produit seulement un pourcentage plus ou moins élevé de lésions constatables après sacrifices sur les animaux auquel elles ont été administrées. C'est précisément en fonction de la proportion de sujets ayant réagi positivement à chacune de ces substances que l'on veut les séparer en deux groupes : les substances actives et les substances inactives. D'autre part, des travaux antérieurs ont montré qu'il n'était intéressant de ranger dans la première catégorie que les drogues donnant un minimum de  $2/3$ , soit  $66,6\%$  de réactions, et que celles pour lesquelles ce pourcentage était inférieur à  $1/3$ , soit  $33,3\%$  devaient être classées comme inactives. Pour les substances donnant des pourcentages intermédiaires, on désire seulement les classer au mieux d'après leur voisinage de ces deux limites. Abstraitement, le problème est bien défini : en expérimentant sur un nombre suffisamment grand d'animaux (quelques centaines par exemple) on peut déterminer avec telle précision que l'on se fixe à l'avance le pourcentage des cas où agit chacune des substances et prendre ensuite celles où sa valeur atteint la valeur choisie. Mais, dans la pratique, il n'en est pas absolument ainsi : si l'on veut éviter de poursuivre les études sur les drogues peu actives c'est parce que ces travaux sont coûteux, mais de même c'est une dépense de temps et d'argent, peut-être inutile, que d'employer un nombre d'animaux trop grand, s'il est possible d'obtenir un classement pratiquement assez exact avec un nombre d'essais très inférieur. Aussi, compte tenu de ces facteurs extérieurs au problème purement statistique, doit-on se fixer deux autres limites : la limite supérieure de la probabilité pour

qu'une substance active ne soit pas retenue parce que des fluctuations d'échantillonnage l'auront fait ranger dans l'autre catégorie, et la limite supérieure de la probabilité de l'erreur inverse consistant à garder comme active une substance qui ne l'est pas. Ici, ces limites ont été prises égales à  $2\%$ , c'est-à-dire que, pour chaque substance dont le pourcentage est en dehors de l'intervalle  $33,3\%$ ,  $66,6\%$ , il n'y a qu'une chance sur cinquante d'être mal classée.

## LA MÉTHODE SÉQUENTIELLE

La grande nouveauté de l'analyse séquentielle est de procéder essai par essai en s'arrêtant aussitôt que le niveau de certitude fixé à l'avance est atteint. Pour cela, deux séries de chiffres sont calculées à l'avance :  $am$  et  $rm$ . Pour chaque substance, on examine les animaux un par un jusqu'à ce que, au  $m^{\text{ème}}$  animal examiné, le nombre de ceux qui ont réagi positivement atteigne  $am$  ou soit rejoint par  $rm$  auquel cas on accepte ou rejette la substance et on passe à l'étude d'une autre drogue. Dans le cas présent, aucune conclusion ne peut être tirée avant qu'un minimum de 6 animaux ait été examiné, et en fonction des chiffres choisis plus haut les valeurs de  $am$  et de  $rm$  sont les suivantes :

Pour la première substance, les essais se sont déroulés ainsi : quatre animaux avaient réagi dans le lot initial de six, de même le septième et le huitième, pas de réaction chez le neuvième, mais une réaction chez le dixième, le onzième et le douzième ; à ce stade, sur douze animaux examinés, neuf ont réagi : la limite  $a_{12} = 9$  est atteinte et l'on range la première substance dans la classe des substances actives.

Pour la deuxième substance, deux animaux seulement du lot initial avaient réagi, puis successivement le huitième, le dixième et le treizième. C'est au seizième animal seulement que la limite  $r_{16} = 5$  est atteinte ce qui fait classer cette drogue parmi les inactives, etc.

Une objection s'offre d'elle-même : est-il sûr que l'on ne sera pas parfois obligé de sacrifier un nombre extrêmement élevé d'animaux avant d'arriver à pouvoir prendre une décision ? Outre le fait (que l'on démontre) que c'est là un événement de probabilité infiniment petite, on convient en général d'arrêter les essais dès qu'un certain nombre d'animaux a été examiné et de classer alors la drogue d'après le pourcentage de réactions observé.

Cette nouvelle limite est prise en général comme deux

TABLEAU I

m : nombre d'animaux examinés	r m : limite inférieure	a m : limite supérieure	RÉSULTAT DES ESSAIS :					
			NOMBRE D'ANIMAUX AYANT RÉAGI :					
			1 <sup>re</sup> sub.	2 <sup>e</sup> sub.	3 <sup>e</sup> sub.	4 <sup>e</sup> sub.	5 <sup>e</sup> sub.	6 <sup>e</sup> sub.
6	0	6	4	2	1	3	3	5
7	0	7	5	2	1	4	3	6
8	1	7	6	3	2	4	3	7
9	1	8	6	3	2	5	3	
10	2	8	7	4	2	5	4	active
11	2	9	8	4	4	5	4	
12	3	9	9	4	4	5	4	
13	3	10	9	5	inactive	5	4	
14	4	10	active	5		6	5	
15	4	11		5		6	5	
16	5	11		5		6	6	
17	5	12		inactive		7	7	
18	6	12				7	7	
19	6	13				8	7	
20	7	13				9	8	
21	7	14				10	8	
22	8	14				11	8	
23	8	15				12		
24	9	15				12	inactive	
25	9	16				13		
26	10	16				14		
27	10	17				14		
28	11	17				15		
29	11	18				16		
30	12	18				17		
31	12	19				17		
32	13	19				17		
33	13	20				18		
34	14	20				19		
35	14	21				19		
36	15	21				20 > 18		active

fois la valeur moyenne du nombre des essais (que l'on peut calculer *a priori*, ce qui est un avantage précieux). Ici, ce nombre moyen étant 18 on s'arrêtera au 36<sup>e</sup> animal si aucune limite n'a encore été atteinte et on rangera comme actives les substances ayant occasionné moins de 18 réactions (exemple : subst. 4). On peut comparer ces chiffres (17 et 35) au nombre d'animaux par lots que les méthodes classiques (non séquentielles) exigeraient pour le même niveau de précision et de sécurité (il faudrait une quarantaine d'animaux) et l'on mesure ainsi l'économie réalisée.

EXTENSION DU PROBLÈME

Le choix actuel de chiffres simples pour les quatre valeurs initiales ne doit pas laisser croire à une limitation imposée par la méthode.

Il est en évident que dans un autre cas, les proportions fixant les caractéristiques *a priori* des substances actives et inactives auraient pu être tout autres : dans une autre série d'essais analogues les produits étaient réputés actifs s'ils conduisaient à un pourcentage supérieur à 95 % et inférieur en dessous de 80 %. De même les limites des probabilités d'erreur ne sont pas nécessairement égales : par exemple si les recherches ultérieures avaient été très coûteuses et le nombre des substances actives considérable, il aurait surtout été important de se prémunir contre le risque de conserver un produit relativement inefficace, d'où élévation d'une des limites à un pour mille par exemple et abaissement de l'autre à cinq pour cent pour évi-

ter d'allonger trop les essais. Dans un autre cas on aurait pu surtout craindre de ne pas reconnaître comme telle une substance active, d'où renversement de ces proportions, etc. (Les chiffres critiques correspondants sont donnés en exemple dans le Tableau II).

Pour chaque problème de ce type, non seulement il est possible de calculer simplement les valeurs *am* et *rm* mais encore l'on peut savoir à l'avance pour une substance dont le pourcentage d'action est donné quelle chance elle a d'être rangée dans l'une ou l'autre des catégories et cela, en moyenne, au bout de combien d'essais.

LA GÉNÉRALITÉ DE L'ANALYSE SÉQUENTIELLE

Bien plus, le type de problème que l'on a traité ici, n'est qu'un exemple des extraordinaires possibilités d'application de l'analyse séquentielle : elle peut être employée pour contrôler si un certain pourcentage ne dépasse pas une valeur donnée, pour déterminer de deux procédés celui qui donne le plus souvent le meilleur résultat, pour tester l'égalité d'une moyenne avec une valeur prévue à l'avance, etc., etc.

Mais peut-être, à côté de l'économie d'effort qu'elle permet de réaliser cette méthode présente-t-elle un autre avantage moins apparent : celui d'obliger l'expérimentateur à prendre clairement conscience de ce qu'il recherche et surtout de la tolérance qu'il est disposé à admettre puisque, aussi bien, celle-ci est toujours bornée chaque fois qu'une connaissance plus concrète et plus profonde des mécanismes causaux ne permet pas d'éviter l'usage des méthodes statistiques.

SEM. HÔP. PARIS  
14 AOUT 1949

QUESTIONS DES PRATICIENS

— 2564 —

TABLEAU II

$m$	$r^m$	$a^m$	$r^m$	$a^m$
5	2	.	0	.
10	6	.	4	.
15	11	.	8	.
20	15	.	13	20
25	20	.	17	25
30	24	.	22	29
35	29	.	26	33
40	33	40	31	38
45	38	45	35	42
50	42	49	40	47
55	46	54	44	51
60	51	58	48	56
65	55	63	53	60
70	60	67	57	65
75	64	72	62	69
80	69	76	66	73
85	73	80	71	78
90	78	85	75	82
95	82	89	79	87
100	86	94		
105	91	98	87	
110	95	103		
115	100	107		
120	104	112		
125	109	116		
130	111	120		
135	117	125		
140	122	129		
145	126	134		
150		135		

Nous avons donné ici les valeurs critiques d'acceptation et de rejet correspondants au cas où les substances « actives » sont définies par un pourcentage d'action supérieur à 95 % et où les substances « inactives » sont celles qui agissent dans moins de 80 % des cas.

Ces valeurs étant peu différentes il y a avantage à procéder par lots de 5 animaux et non pas animal par animal.

La 1<sup>re</sup> colonne correspond au nombre total «  $m$  » d'animaux examinés.

Les 2<sup>e</sup> et 3<sup>e</sup> colonnes donnent les chiffres critiques  $r^m$  et  $a^m$  pour une première expérience où il convient surtout de se prémunir contre la conservation d'une substance inactive sans que le rejet d'une substance active soit très grave. (Probabilités respectives de ces deux erreurs : 1/1.000 et 5/100).

Les 4<sup>e</sup> et 5<sup>e</sup> colonnes donnent les chiffres critiques  $r^m$  et  $a^m$  pour une deuxième expérience où les probabilités sont inverses des précédentes (5/100, 1/1.000), il s'agit ici essentiellement de ne pas laisser échapper une substance active.

Dans la 1<sup>re</sup> expérience aucune substance ne peut être acceptée avant le 40<sup>e</sup> animal (20<sup>e</sup> dans la seconde). Les nombres moyens d'animaux à examiner sont :

	1 <sup>re</sup> expérience	2 <sup>e</sup> expérience
Acceptation d'une substance active .....	68	32
Rejet d'une substance inactive .....	21	46

REFERENCES

Il est impossible de donner ici une bibliographie de ce domaine qui se développe très vite. Cependant on peut citer les ouvrages suivants :

WALD A. — Sequential Analysis (N. Y., 1947); il convient également d'associer à cette découverte le nom de G. A. BARNARD, qui, indépendamment de WALD, a décrit une méthode analogue pour les doubles dichotomies dans : Economy in Sampling with référence..., etc... *British Ministry of Supply. Advisory service of Stat.* série R, n° Q6/R7.

Pour de nombreux travaux on trouvera de nombreuses tables et graphiques dans : FREEMAN H. A.; FRIEDMAN M.; MOSTELLER F. and WALLIS W. A. : Sampling inspection (N. Y., 1948).

(Travail subventionné par l'Institut National d'Hygiène et le Fonds d'études de la Société Médicale des Hôpitaux de Paris).

RÉSULTATS D'UNE ENQUÊTE SUR LA DISTRIBUTION DU SEXE  
DANS LES FAMILLES NOMBREUSES

par M. P. SCHÜTZENBERGER.

Attaché de Recherche à l'Institut National d'Hygiène

L'étude des processus qui déterminent le sexe de l'enfant à naître nécessite concurremment à toute recherche biologique une connaissance statistique précise qui lui indique des perspectives, contrôle ses déductions, fixe les limites *a priori* de sa validité.

Sous la direction de M. le Professeur R. Turpin, et grâce à l'appui du Fonds d'Etudes de la Société Médicale des Hôpitaux de Paris, nous avons effectué une semblable enquête. Elle nous a permis de mettre en évidence un fait nouveau que nous discuterons ici.

## LES DOCUMENTS

Une attention particulière doit être apportée dans les enquêtes démographiques à ce que les facteurs économiques et sociaux aux sens le plus large, ne viennent masquer les phénomènes proprement biologiques dont on se propose l'étude. Ainsi, peut-on craindre qu'une restriction volontaire de la natalité n'intervienne dans une famille quand est né, enfin, un héritier mâle; ou qu'au contraire une longue succession de filles ne décourage certains parents... A ces influences, qui viennent le plus immédiatement à l'esprit (et dont nous ignorons complètement la portée effective), s'en ajoutent certainement d'autres plus subtiles, plus variables et plus importantes. La composition actuelle d'une famille du point de vue de sa répartition en garçons et en filles n'est sûrement pas sans influer sur l'attitude des progéniteurs dans un très grand nombre de cas. Aussi, pour éviter au maximum les distortions par rapport aux processus physiologiques qu'occasionnent ces influences, avons-nous limité notre enquête aux familles nombreuses et plus précisément aux familles ayant eu plus de 5 enfants. Peut-être eut-il mieux valu prendre une limite plus élevée; malheureusement, la rapidité de la décroissance du nombre des familles en fonction de leur importance, nous eut exposé à un autre danger: celui de ne disposer que d'un échantillon trop restreint. En effet, dans la population, on constate qu'avec une régularité assez remarquable il y a à peu près 1,6 fois moins de familles de  $n + 1$  enfant que de familles de  $n$ , pour  $n$  plus grand que 5 tout au moins.

TABLEAU N° I  
ORIGINE DES DOCUMENTS

Nombre d'enfants par famille	Ministère de la Santé	Région parisienne	Grandes familles du Nord	Fondation Cognac-Jay	Total
5	3.309	1.387	432		5.128
6	2.214	740	301		3.255
7	1.298	374	205		1.877
8	1.269	211	139		1.119
9	682	109	81		872
10	527	48	59		634
11	313	27	34	10	384
12	159	12	15	32	218
13	68	3	6	31	108
14	26	3	1	29	59
15	14		2	21	37
16	9		1	12	22
17	5			6	11
18				2	2
19				4	4
Total des familles.	9.893	2.914	1.270	147	14.230

larité assez remarquable il y a à peu près 1,6 fois moins de familles de  $n + 1$  enfant que de familles de  $n$ , pour  $n$  plus grand que 5 tout au moins.

D'autre part, la considération des naissances gemellaires et de leur répartition introduisant un élément étranger au but que nous nous proposons (et qui n'eut pu être éliminé que par des artifices mathématiques compliqués), nous n'avons recueilli que des familles n'en comportant pas. Ainsi notre échantillon comprend au total 14.230 familles françaises contemporaines sans jumeaux, soit 95.725 naissances. Le détail des sources est indiqué au tableau I :

— Service de la Médaille des Familles au Ministère de la Santé Publique et de la Population (9.893 familles).

TABLEAU N° II  
RÉPARTITION NUMÉRIQUE DES FAMILLES  
SELON LE NOMBRE DE LEURS ENFANTS ET LE NOMBRE DE GARÇONS

Enfants	GARÇONS															
	0	1	2	3	4	5	6	7	8	9	10	11	12	13		14
5	5.144	810	1.544	1.585	841	24										8.128
6	54	325	717	979	785	350	65									3.255
7	21	111	287	480	529	304	125	19								1.877
8	8	61	180	327	412	406	170	49	6							1.619
9		20	72	155	204	184	140	74	22	1						872
10		5	23	64	120	165	128	78	39	11	1					634
11		3	15	27	51	85	92	61	37	9	3	1				384
12		4	3	18	19	32	50	49	23	16	2	1	1			218
13		2	1	3	9	12	24	19	14	9	13	2				108
14				2	3	6	9	14	11	8	6	1				59
15					2	2	4	6	8	5	5	1	4			37
16						2	3	3	4	3	4	1	2			22
17						1			1	3	2	3	1			11
18													1			2
19									1					1	1	4

## Année 1949 1949-6. Résultats d'une enquête sur la distribution du sexe dans...

SEM. HÔP. PARIS  
18 AOUT 1949

DISTRIBUTION DU SEXE DANS LES FAMILLES NOMBREUSES

— 2580 —

— Fichiers des Associations de Familles Nombreuses de la Région Parisienne (2.914 familles).

— Annaire dit « Les grandes familles du Nord de la France » (édition 1948), (1.276 familles).

— Dossiers de la Fondation Cognacq-Jay (années 1920-1924 : 145 familles). Outre les personnes appartenant à ces divers services et qui ont bien voulu faciliter notre accès à leurs documents, nous tenons à remercier Mmes Godin, Deroche et Pouzol qui nous ont aidé à recueillir près de la moitié des familles au Ministère de la Santé Publique et de la Population. En plus de son aspect matériel, leur aide nous a été précieuse puisqu'elle nous a permis de contrôler l'objectivité de notre échantillonnage : les calculs limités aux seuls documents que nous n'avons pas recueillis personnellement conduisent dans l'ensemble aux mêmes conclusions que ceux que nous avons collectés nous-mêmes. Ce point est plus important qu'il ne paraît puisqu'au aussi bien une certaine part d'arbitraire était inévitable dans l'enregistrement d'un aussi grand nombre de familles et qu'on eut pu suspecter l'influence d'une tendance inconsciente en faveur de nos thèses dans les cas douteux (prénoms ambigus tels que Claude, Dominique, etc., élimination de documents de base suspects, ou même erreur pure et simple dans l'enregistrement).

LA VARIATION DE LA SEX RATIO  
SELON LE RANG DE NAISSANCE

Préalablement à toute autre recherche, il était nécessaire de contrôler l'hypothèse fréquemment admise d'une variation systématique de la fréquence des garçons selon le rang de naissance (les premiers-nés seraient plus fréquemment du sexe masculin) et en fonction de l'âge de la mère (les femmes âgées ayant proportionnellement moins de filles). En ce qui concerne la première de ces théories, nos documents donnent une réponse très nette : si cette déviation existe, elle ne saurait être que très faible, puisque la proportion des garçons parmi les premiers-nés n'est que de 0,5097 contre 0,5062 aux autres rangs. Le léger écart constaté pourrait très légitimement être imputé à une simple fluctuation d'échantillonnage car l'écart type de la première fréquence est de 0,0042 et celui de la seconde de 0,0018. En ce qui concerne la deuxième hypothèse, nos documents se prêtent moins bien à une analyse rigoureuse. Cependant, il est clair que si elle était vraie on devrait constater une augmentation du nombre des garçons parmi les naissances de rang le plus élevé. On en déduit un test asymptotiquement équivalent à celui de J. B. S. Haldane et Cedric B. Smith (pour tester l'influence du rang de naissance sur l'incidence d'une maladie héréditaire) en calculant le rang moyen de naissance des garçons. Aucune déviation de la valeur observée à la valeur théorique que prévoit l'hypothèse de la constance de la probabilité n'a pu être établie de façon statistiquement significative. En présence de ce résultat négatif, qui retrouve une recherche antérieure de Wicksell sur les naissances en Suède (corrélation entre sex-ratio et âge maternel), nous supposons dorénavant que la probabilité de naissance d'un garçon ne varie pas systématiquement avec le rang de la naissance.

LA VARIANCE DE LA SEX RATIO  
D'UNE FAMILLE A L'AUTRE

L'affirmation de la constance de la sex ratio dans l'espèce humaine, même si on tient compte de possibilités de variations systématiques telle que l'augmentation si souvent invoquée du nombre des garçons après les guerres, re-

couvre en réalité toute une série d'hypothèses entre les deux théories limites suivantes :

— à chaque naissance, la probabilité *a priori* de naissance d'un garçon a une certaine valeur, constante d'une famille à une autre ;

— au contraire, d'une famille à une autre, cette probabilité varie de façon importante et c'est seulement sa valeur moyenne qui est constante dans la population.

Pour étudier cette question, la statistique théorique montre qu'il suffit de posséder pour un échantillon assez nombreux non pas seulement la sex ratio totale, mais la composition de chacune des familles en garçons et en filles. En effet, si la première hypothèse était vraie, l'onde  $x$  devrait observer à peu près  $N_p \binom{g+f}{g}$  familles de  $g$  garçons et  $f$  filles si le nombre total des familles de  $g+f$  enfants est  $N$ , et si la sex ratio générale est  $p$ . Au contraire, si la deuxième hypothèse se trouvait vraie, les valeurs probables correspondant aux familles où l'un des sexes prédomine nettement sur l'autre devraient être plus fortes que ne l'indiquent les formules précédentes. Sans entrer dans le détail ni des calculs ni des méthodes employées, nous nous bornerons à affirmer que, dans notre échantillon, les chiffres observés obligent à rejeter sans appel la première hypothèse : pour qu'une déviation aussi forte entre les valeurs qu'elle prévoit et les faits existant il eut fallut que se produisit un événement de probabilité inférieure à 1/100.000. Confirmant un résultat déjà obtenu par R. A. Fisher sur les matériaux réunis par Geissler, nous avons donc montré que *lorsque sont en causes des collections très importantes de familles, la théorie selon laquelle la probabilité a priori de naissance d'un garçon est une constante absolue ne peut être observée*. Pour donner une idée de l'importance de ce phénomène, indiquons que 65 familles de 6 enfants sans une seule fille ont été recueillies, alors que 55 seulement auraient dû être observées ; que nous avons trouvé 21 familles de 7 filles sans un seul garçon au lieu des 13 prévisibles dans l'hypothèse d'une probabilité identique pour toutes les familles, etc. Il n'est pas inutile d'insister sur le caractère, somme toute minime, de ces divergences et qui seraient passées inaperçues dans une enquête moins étendue (quelques milliers de familles seulement).

TABLEAU N° III  
VALEURS CARACTÉRISTIQUES DES HISTOGRAMMES

Nombre d'enfants	Nombre de familles	Fréquence des naissances masculines	Variance observée	Ecart de la variance (1)
5	5.128	0.5085	1.300	+ 4,6
6	3.255	0.5057	1.595	+ 6,3
7	1.877	0.5061	1.892	+ 6,5
8	1.619	0.5011	2.069	+ 2,7
9	872	0.4989	2.535	+ 10,7
10	634	0.5178	2.583	+ 1,9
11	384	0.5066	2.976	+ 3,8
12	218	0.5022	3.667	+ 8,0
13	118	0.5207	4.990	+ 2,9
14	59	0.5170	3.322	— 1,0
15	37	0.5441	4.095	+ 4,3
16	22	0.4682	5.110	+ 3,7
17	11	0.6112	3.072	— 0,8
18	2	0.5566	8.000	+ 3,5
19	4	0.6316	3.333	— 0,1

(1) Ecart entre valeur observée et valeur théorique exprimé en écart-type.

18 AOUT 1949

TABLEAU N° VI

COMPARAISONS ENTRE LES HISTOGRAMMES OBSERVÉS ET LES SCHEMAS PROBABILISTES			
Nombre de garçons	Nombre observés	Nombres théoriques dans l'hypothèse binominale	Nombres théoriques dans l'hypothèse de Markoff.
FAMILLES DE CINQ ENFANTS			
6	204	171	190
4	841	833	852
3	1,585	1,624	1,386
2	1,544	1,581	1,543
1	810	769	788
0	144	150	169
FAMILLES DE SIX ENFANTS			
6	65	55	63
5	330	321	336
4	785	783	776
3	979	1,017	987
2	717	743	733
1	325	289	304
0	54	47	54
FAMILLES DE SEPT ENFANTS			
7	19	16	19
6	126	109	117
5	304	320	323
4	529	520	507
3	460	507	493
2	327	296	298
1	111	96	104
0	21	13	10
FAMILLES DE HUIT ENFANTS			
8	6	7	5
7	49	55	60
6	170	185	190
5	406	356	351
4	412	444	432
3	327	348	343
2	180	171	176
1	61	47	52
0	8	6	9

## DISCUSSION DU SCHEMA DE MARKOFF

La première interprétation de ces résultats, celle qui correspond le plus étroitement au formalisme analytique, est l'hypothèse d'une influence réelle du sexe d'un enfant sur le sexe de l'enfant à la grossesse suivante : pour étayer biologiquement cette théorie, on pourrait faire appel à l'exemple du *free martinisme* où, en effet, une semblable action (mais contemporaine !!!) est indiscutable. Il faudrait alors admettre une sorte d'imprégnation de l'organisme maternel, persistant quelques mois ou plutôt quelques années et pouvant influencer le sexe de l'embryon suivant. Plus vraisemblablement, pourrait-on penser que cette imprégnation augmente électivement le nombre de ces avortements précoces qu'a étudiés Parker quand le sexe du nouveau fœtus diffère de celui de l'enfant précédent. Mais il est une autre interprétation de ces phénomènes qui nous semble plus aisée à admettre dans l'état actuel de nos connaissances. Supposons en effet que, sous l'influence de causes extérieures variées, l'organisme maternel présente des périodes de durée et d'intensité aléatoires pendant lesquelles est favorisée la conception d'enfants de tel ou tel sexe, ou bien est augmentée la létalité des embryons mâles ou femelles (on pourrait songer par exemple à des modifications du ph utérin ?). Si ces périodes sont assez longues pour chevaucher de manière relativement fréquente sur

plusieurs grossesses, leur action se manifesterait en première approximation par des naissances réparties selon un schéma de Markoff. A vrai dire, si ces influences étaient très importantes, ou bien si leur fréquence ou leur durée compensaient leur faible action, il serait possible d'en tester rigoureusement l'existence. De fait, l'ordre de grandeur de l'incrément de probabilité auquel nous avons abouti ne permet guère de pousser plus à fond la discussion, à moins de pouvoir doubler ou tripler le volume des documents de base, ce qui semble pratiquement irréalisable en France actuellement. Par conséquent, il nous faut faire des hypothèses qui peuvent sembler parfaitement arbitraires pour évaluer la portée réelle de l'action de ces périodes gynophiles et androphiles. Supposons que les périodes aient une durée moyenne de deux gestations avec une probabilité de  $p+x$  et de  $p-x$ . Pour obtenir un incrément de liaison de 0,015, il suffit que  $x=0,146$ . Autrement dit, si l'on veut bien admettre que la probabilité de naissance d'un garçon fluctue lentement et de manière aléatoire entre 0,35 et 0,65 environ, mais de telle sorte que les deux tiers des naissances en moyenne se trouvent dans la même phase, cela suffit pour rendre compte de la corrélation que nous avons observée. Si ces périodes étaient en moyenne deux fois plus longues, il suffirait que la probabilité oscillât entre 0,40 et 0,60. Pour discuter l'hypothèse qui nous semble la plus vraisemblable : seules certaines femmes présenteraient de semblables fluctuations de la probabilité de naissance d'un garçon, il faudrait faire encore de nouvelles hypothèses qui perdraient alors tout sens. Cette dernière théorie a cependant pour elle et d'expliquer mieux la constance relative de la sex ratio générale et le fait que la liaison résiduelle entre naissance aux rangs  $n$  et  $n+2$  est très faible, sinon nulle, résultat peu compatible avec l'existence relativement fréquente de périodes de longue durée.

## CONCLUSIONS

Arrivé au terme de cet exposé, il convient de faire le point des résultats obtenus et des perspectives qu'ils peuvent ouvrir à de nouvelles recherches.

- 1° Il ne semble pas que la fréquence des garçons varie de façon importante ni systématique selon le rang de naissance ;
- 2° L'hypothèse d'une distribution des sexes selon un schéma binomial (c'est-à-dire spécifiant la constance de la probabilité et l'indépendance des naissances) ne peut être conservée ;
- 3° L'hypothèse d'une liaison simple entre les sexes dans les naissances consécutives rend compte convenablement des faits observés et s'interprète de façon satisfaisante par l'existence de périodes aléatoires gynophiles et androphiles ;
- 4° Il ne semble pas possible en se limitant à une formation de même nature que celle utilisée ici (succession des sexes dans les familles), d'obtenir autre chose qu'une confirmation ou une infirmation de nos thèses.

Par contre, en faisant usage d'autres éléments, et en particulier de la durée de l'intervalle qui sépare les naissances, il serait sans nul doute possible de préciser les résultats déjà obtenus et de serrer de plus près les hypothèses entre lesquelles il est encore difficile de choisir. Sous la direction de M. le Professeur R. Turpin, ces travaux sont actuellement en cours au Centre de Génétique de l'hôpital Saint-Louis et laissent espérer de rapides progrès dans la connaissance des mécanismes qui déterminent le sexe dans l'espèce humaine.

## REFERENCES

- R. TURPIN et M. P. SCHÜTZENBERGER. — Recherche statistique sur la distribution du sexe à la naissance. *Comptes rendus à l'Académie des Sciences*, 226. 31-5-1948.

Année 1949 1949-7. Fondements de la statistique appliquée à la psychologie, I...

# BULLETIN

## DU GROUPE D'ÉTUDES DE PSYCHOLOGIE DE L'UNIVERSITÉ DE PARIS

Les 2 numéros : 200 francs (80 pages)

17, rue de la Sorbonne, Paris V<sup>e</sup>

### SOMMAIRE

	Pages
Congrès de l'Union Nationale des Etudiants, G. Milhaud	2
<b>PSYCHOLOGIE GENERALE</b>	
La Conduite Humaine, M. Lagache, VII, XI, XII, XIII, XIV, XV, XVI, XVII.....	3 à 28
Etude sur le Langage, M. Besson, I, II.....	28 à 31
<b>PSYCHOLOGIE PHYSIOLOGIQUE</b>	
Agencement des superstructures cérébro-spinales et végétatives (La vision), M. Tournay.....	32 à 34
L'organisation de la physiologie des sociétés animales. M. Grassé, I, II.....	34 à 40
La Connaissance de l'Espace, M. Piéron, VIII, IX....	40 à 45
Notion d'histoire des méthodes, M. Fraisse, VI et suite	45 à 54
<b>PSYCHOLOGIE DE LA VIE SOCIALE</b>	
Problèmes et techniques de Psychologie Sociale, M. Lagache .....	54 à 61
<b>PSYCHOLOGIE PATHOLOGIQUE</b>	
Leçons de psychanalyse théorique, M. Lagache, V, VI, VII .....	62 à 68
Conférences du stage de psychiatrie de Saint-Anne, M. Baruk .....	75
<b>PSYCHOLOGIE APPLIQUEE</b>	
La méthode des tests et son champ d'application, (Fidélité et Validité), M. Reuchlin.....	28
<b>PSYCHOLOGIE PEDAGOGIQUE</b>	
Les Syndromes caractériels, M. Wallon.....	28
Problèmes et méthodes de Psychologie de l'Enfant, M. Zazzo .....	77
<b>LA VIE DU GROUPE DE PSYCHOLOGIE</b>	
Dates des examens de l'Institut de Psychologie....	74
Examen de statistique.....	74
Année propédeutique et licence de psychologie.....	61
Annonces de démonstrations et de conférences.....	53
Questions d'oral de psychologie .....	78
Questions d'oral de psychologie de la vie sociale....	79
<b>CONFERENCES</b>	
Principes des tests statistiques, D <sup>r</sup> M.-P. Schützenberger .....	73

MERCREDI 11 MAI 1949

2<sup>e</sup> Année - N<sup>o</sup> 8-9

# LA VIE DU GROUPE DE PSYCHOLOGIE

## CONFÉRENCES

Résumé des conférences faites au Groupe

D<sup>r</sup> M.-P. SCHÜTZENBERGER

### PRINCIPE DES TESTS STATISTIQUES

La statistique du XIX siècle n'était guère qu'une technique de condensation et de description de ces phénomènes qui nécessitent le traitement de larges masses de documents.

Sous la poussée des nécessités techniques (contrôle de fabrication, expérimentation agricole, etc...), ce point de vue a complètement changé, grâce à l'esprit nouveau que lui a donné le calcul des probabilités. Aujourd'hui, *tout problème concret de statistique se ramène au schéma suivant* :

A chaque expérience, une et une seule des éventualités, A, B, C, se trouve réalisée. Nous ignorons laquelle, mais nous savons que, si c'était A, un certain phénomène observé avait des chances, P<sub>A</sub>, P'<sub>A</sub>, P''<sub>A</sub>... de se produire, que si c'était B ces chances étaient P<sub>B</sub>, P'<sub>B</sub>, P''<sub>B</sub>, etc...

A l'aide de la seule connaissance des phénomènes en question, nous devons dire quelle était l'éventualité qui était réalisée et prendre, en conséquence, une décision qui sera plus ou moins heureuse, selon l'écart entre la réalité et notre supposition. Pour cela, nous demandons au statisticien de nous donner une « règle de conduite » fixant de manière optimale la décision à prendre selon les observations.

Manifestement, c'est là le problème-type du diagnostic en psychologie : les éventualités sont l'appartenance du sujet à l'une des catégories A, B... (par exemple, catégories nosologiques). Si les tests sont faillibles — ce qui semble être parfois le cas, dans le domaine caractériel au moins — le problème devient un problème de statistique, à partir du moment où nous connaissons :

1° les probabilités pour qu'un « sujet A » apparaisse (à tort) à l'examen comme un « sujet B », comme un « sujet C », etc... ;

2° La gravité relative de nos erreurs possibles :

Comme il s'agit surtout ici de démontrer le mécanisme logique du raisonnement statistique, nous adopterons un exemple très simple avec deux catégories seulement.

Un expérimentateur me présente soit une boîte A, soit une boîte B, ayant la composition suivante :

A : trois boules rouges et une boule noire ;

B : trois boules noires et une boule rouge.

Je tire au hasard une boule de la boîte qui m'est présentée, et je dois dire si cette boîte était une boîte A ou une boîte B.

La traduction psychologique de ce schéma serait le dépistage des malades au moyen d'une épreuve telle que :

A : si le sujet est normal, il a une chance sur quatre de réagir positivement ;

B : si le sujet est malade, il a trois chances sur quatre de réagir positivement.

Examinons les quatre règles de conduites possibles :

- (1) Quelle que soit la couleur de la boule que j'ai tirée, je répons A ;
- (2) Quelle que soit la couleur de la boule que j'ai tirée, je répons B ;
- (3) Si la boule est rouge, je répons A, et je répons B si elle est noire ;
- (4) Si la boule est rouge, je répons B, et je répons A si elle est noire.

Il est facile de tracer le tableau suivant, qui donne, selon la règle de conduite adoptée, le pourcentage de réponses justes :

	(1)	(2)	(3)	(4)
A.....	100 %	0 %	75 %	25 %
B.....	0 %	100 %	75 %	25 %

Si nous nous limitons à des règles de conduite rigides, ce sont là les seules possibilités. Clairement, la règle de conduite (3) est meilleure que (4), puisqu'elle donne un pourcentage de réponses justes supérieur dans les deux cas. Le choix entre (1), (2) et (3), par contre, n'est pas aussi facile et l'on doit à von Neumann d'avoir formulé clairement le principe suivant qui est généralement adopté :

« Choisir la règle de conduite qui rende maximum le nombre minimum de réponses justes. »

Ceci revient, en quelque sorte, à s'assurer contre les pires échecs :

Si j'adoptais (1), l'expérimentateur pourrait, en ne me donnant jamais que des boîtes B, me faire commettre 100 % d'erreurs, et de même *mutatis mutandis*, pour (2) ; par contre, pour (3), l'expérimentateur ne pourra pas faire descendre le pourcentage des réponses exactes en dessous de 75 %.

C'est donc cette règle de conduite qu'il conviendrait d'employer dans l'exemple du diagnostic psychologique évoqué plus haut.

Cette manière de raisonner nous permet d'éviter l'usage d'assertions telles que : « Tel sujet ayant répondu positivement dans l'épreuve a trois chances sur quatre d'être un malade », qui n'ont rigoureusement aucun sens dans la statistique moderne (elles n'auraient un sens que si les sujets eux-mêmes étaient tirés au hasard).

En effet, selon que la population que nous examinons est un quartier d'asile d'aliénés ou, au contraire, un atelier d'usine, la « probabilité » pour qu'un sujet, au hasard, soit un malade n'est pas la même.

Par contre, décider que nous classerons comme malades ceux qui réagissent positivement, et normaux les autres, nous garantit que, quelle que soit la population que nous examinons, nous nous tromperons dans 25 % des cas au plus.

Il faudrait, pour être complet, discuter le cas où, précisément, est connue à l'avance la composition de la population, mais, en réalité, le domaine d'application de cette théorie est assez limitée dans la pratique.

Grâce à l'exemple précédent, il nous est maintenant possible d'expliquer comment se justifient les « tests statistiques » si fréquemment utilisés :

Prenons le cas le plus simple du test de la différence de deux moyennes :

On a fait passer une épreuve à deux populations,  $x$  et  $y$ , et l'on veut savoir si l'hypothèse  $H_0$ , selon laquelle les deux populations ont la même moyenne peut être conservée ou si l'on doit, au contraire, préférer l'hypothèse  $H_d$ , selon laquelle la moyenne de  $x$  est plus grande de «  $d$  » unités que celle de  $y$ . (On suppose égales et connues les variances.)

La logique de l'argumentation est la suivante : si  $H_0$  était vraie, une différence de moyenne aussi forte que celle qui est observée aurait une probabilité  $P_0$  ; si, au contraire,  $H_d$  était vraie, cette probabilité aurait pu atteindre  $P_d$ .

En rejetant  $H_0$  chaque fois que  $P_0/P_d$  est inférieur à une certaine valeur, et en la conservant dans le cas contraire, nous agissons, en définitive, de la même manière que dans la règle de conduite (3) de notre exemple :

Si  $H_0$  est vraie, nous tomberons juste dans  $a$  % des cas.

Si  $H_d$  est vraie, nous tomberons juste dans  $b$  % des cas ;  $a$  et  $b$  peuvent se calculer a priori et sont choisis de telle sorte que, quelle que soit la véritable valeur de la différence, notre pourcentage d'erreurs ne puisse s'élever au-dessus d'une certaine limite.

Ici encore, répétons-le, il ne s'agit nullement de « probabilité que  $H_0$  soit vraie ou fausse », mais de règles de conduite conduisant, en moyenne, à des résultats optimaux (au sens de von Neumann).

Pour conclure, il convient de remarquer, une fois de plus, que les tests statistiques ne peuvent se définir que lorsque les deux conditions suivantes sont réalisées :

- 1° Lorsqu'un processus aléatoire connu intervient à un stade ou à un autre ;
- 2° Lorsqu'est spécifiée la nature des hypothèses en faveur desquelles peut être rejetée l'hypothèse nulle. (Une affirmation n'est bien posée qu'en même temps que les négations correspondantes.)

Si ces conditions sont remplies, la méthode statistique permet d'éviter que le nombre des erreurs excède certaines valeurs — que l'on pourrait rendre arbitrairement petites dans certains cas — mais au prix d'une limitation supérieure corrélative du nombre des réussites. Et cette limitation essentielle de ce type de raisonnement ne doit pas être oubliée.

Dr M.-P. SCHÜTZENBERGER.

# BULLETIN

## DU GROUPE D'ÉTUDES DE PSYCHOLOGIE DE L'UNIVERSITÉ DE PARIS

Les 3 numéros : 300 francs (112 pages)

17, rue de la Sorbonne, Paris V<sup>e</sup>

### SOMMAIRE

	Pages
L'Équipe du Bulletin vous parle.....	2
<b>PSYCHOLOGIE GÉNÉRALE</b>	
La conduite humaine (VIII, IX, X), M. Lagache.....	3
Études du langage (III, IV), M. Bresson.....	10
<b>PSYCHOLOGIE DE LA VIE SOCIALE</b>	
Conscience et Société. — Les faux problèmes de la sociologie du XX <sup>e</sup> siècle, M. Gurvitch.....	14
Psychologie sociale aux U.S.A. (I, II, III, IV, V, VI), M. Klineberg.....	26
<b>PSYCHOLOGIE PHYSIOLOGIQUE</b>	
Agencement des superstructures cérébrospinales et végétatives (goût, odorat, audition, le cervelet, le système végétatif), M. Tournay.....	38
Le système nerveux élémentaire (V à XII), M. Fessard	51
Notion d'histoire des systèmes et des méthodes (fin), M. Fraisse .....	56
La connaissance de l'espace. . . Vision (suite), M. Piéron	69
L'organisation de la physiologie des sociétés animales (III, IV, V), M. Grasse.....	77
<b>PSYCHOLOGIE DE L'ENFANT ET PÉDAGOGIE</b>	
Les syndromes caractériels (fin), M. Wallon.....	86
Travaux pratiques. — Test de la figure complexe, M. Mialaret .....	87
<b>PSYCHOLOGIE APPLIQUÉE</b>	
Psychologie différentielle, M. Oleron.....	89
<b>PSYCHOLOGIE PATHOLOGIQUE</b>	
Leçons de psychanalyse théorique (VIII et IX), M. Lagache .....	94
Conférences de psychiatrie :	
M. Delay .....	100
M. Baruk .....	109
M. Pichot .....	106
D <sup>r</sup> Marty .....	107
<b>CONFÉRENCES</b>	
Fondements de la statistique appliquée à la psychologie, II (Estimation d'un paramètre inconnu), D <sup>r</sup> M.-P. Schützenberger .....	112

LUNDI 23 MAI 1949

2<sup>e</sup> Année - N<sup>os</sup> 10-11-12

# CONFÉRENCES

Résumé des conférences faites au Groupe

## LES FONDEMENTS DE LA STATISTIQUE APPLIQUÉE A LA PSYCHOLOGIE

D<sup>r</sup> M.-P. SCHUTZENBERGER

### II

#### ESTIMATION D'UN PARAMÈTRE INCONNU

Dans la dernière conférence nous avons examiné le problème central de la statistique moderne. Aujourd'hui nous étudierons un problème particulièrement important pour les psychologues : celui de l'estimation d'une grandeur inconnue.

En quoi consiste-t-il exactement :

L'on sait à l'avance que les observations dépendent de façon bien déterminée de la valeur d'un certain paramètre inconnu. On fait un certain nombre d'observations et l'on essaye d'en déduire une valeur « optimale » du paramètre.

Par exemple le nombre des mâles dans les portées de quatre petits d'une certaine espèce est une fonction binomiale d'un paramètre inconnu « la sexratio »  $p = 1 - q$ , on devrait avoir, si nombre observé des portées était très grand :

Nombre de portées avec :

4 mâles	$A_4 p^4$ ,
3	$A_3 4 p^3 (1-p)$ ,
2	$A_2 6 p^2 (1-p)^2$ ,
1	$A_1 4 p (1-p)^3$ ,
0	$A_0 (1-p)^4$ .

Du fait dans une étude donnée, les valeurs observées  $A_1$  à  $A_4$  ne satisfont pas en général les diverses relations qu'impliquent ces équations. Il faut donc trouver un principe qui nous donne le moyen de savoir si telle valeur de  $p$  représente mieux les chiffres observés que telle ou telle autre. Encore une fois, la notion d'utilité « opérationnelle » joue un rôle prédominant : l'estimation que nous effectuons à tel ou tel but pratique — et l'on peut chiffrer le prix de tel ou tel type d'erreur. Le point de vue peut commander l'expérimentation elle-même et il est parfois moins cher de réunir de larges masses d'observations, sur lesquels on effectue des calculs simples que d'effectuer des calculs pénibles sur des observations peu nombreuses. Ceci vaut surtout dans la pratique industrielle ; dans la recherche on se préoccupe en général de tirer des observations toute la précision qu'elles peuvent donner. Pour cela il existe deux méthodes générales :

La méthode des moindres carrés (on rend minimum la somme du carré des écarts, entre observations et valeurs théoriques déduites de la valeur estimée du paramètre).

La méthode du maximum de vraisemblance

Abonnement annuel : 1.000 Fr.

(on prend la valeur du paramètre qui aurait rendu maximum la probabilité a priori de l'ensemble des observations).

Les deux méthodes sont asymptotiquement équivalentes dans des cas très généraux et c'est une discussion sans cesse renouvelée entre les élèves de Egon S. Pearson et de R.-A. Fisher que de savoir quelle est la meilleure (la plus simple et la plus efficace) dans chaque cas particulier.

Pour montrer plus clairement le caractère optimal de ces estimations, prenons le cas courant de l'estimation de la « vraie valeur » de l'aptitude d'un sujet dans un certain test et sa réussite lors d'une passation.

Moyennant les hypothèses usuelles, le problème est ramené au suivant :

Chaque sujet  $i$  est caractérisé par une « vraie valeur  $x_i$  » de son aptitude.

Les  $x_i$  sont distribués normalement, et avec une moyenne nulle et une variance égale à un. La note  $y_i$  obtenue par le sujet  $i$  lors de la passation de test est une variable aléatoire distribuée avec une variance égale à  $(1-r^2)$  autour d'une moyenne  $r x_i$ . Nous expliquerons plus bas la signification de  $r$ .

Pour un sujet  $y_i$  est donnée. Que peut-on dire sur  $x_i$ ?  $X_i$  n'est pas une variable aléatoire en général —  $x_i$  est seulement inconnu. La « meilleure » décision sera de le prendre égal à  $r y_i$ , comme on le montre facilement par l'une ou l'autre des méthodes invoquées plus haut. Meilleur signifie ici que si nous recommençons un assez grand nombre de fois l'expérience. La somme des carrés des écarts des valeurs  $x_i$  estimées à leurs vraies valeurs serait un minimum.

Quelle est la nature de l'information que nous a apporté le test ?

Avant, nous savions seulement que  $x_i$  appartenait à une population normale de variance 1. Nous savons par exemple que 3 0/00 des sujets avaient une « vraie valeur », plus grande en valeur absolue que 3, etc. Une fois que le sujet a passé le test, la connaissance que nous avons de  $x_i$  est plus précise :

D'une part, la valeur moyenne des « vraies valeurs » conduisant à un score  $y_i$  n'est pas zéro, mais  $r y_i$ ; d'autre part, la variance de cette répartition est  $1-r^2$  « 1, et par exemple si  $r = 0,8$ , nous n'aurions eu que

(Suite page 111.)

C.C.P. Lombes Paris 614.662.

Directeur-Gérant : M. Kirschen

I.C.C. (J. LONDON, Impr.), 13, rue de la Grange-Batelière, Paris (9<sup>e</sup>)

**STATISTIQUES**

3 0/00 d'avoir  $y_i$ , si  $x_i$  avait différé de  $r y_i$  de 1,8 (et non pas 3 comme précédemment). De la comparaison de l'information initiale à l'information après observation se déduit une quantité, « la quantité d'information » qui partage de nombreuses propriétés formelles avec une grandeur bien connue des physiciens : l'entropie (cf. Wiener and Doob).

Revenons à la grandeur  $r$  des formules précédentes : c'est une fonction de la reliabilité de test ; on peut la calculer (théoriquement) par divers artifices souvent contestables. De fait, quelques exemples montrent à quel point la précision est faible quand  $r$  tombe en-dessous de 0,8 ou 0,9. Ainsi par exemple pour certains tests caractériels

Suite de la page 112.

(sinon tous)  $r$  plus petit ou égale 0,50 ; la réduction de variance est donc de 25,25 %. C'est-à-dire que la variance des  $x_i$  pour  $y_i$  fixe est de 0,75 (au lieu de un). Cependant que pour chaque unité dans le score observé l'estimation ne progresse que d'une demi-unité. Pour pouvoir (au niveau de 1 % rejeter l'hypothèse qu'un sujet a une vraie valeur en-dessous de la moyenne, il faut donc que son score observé soit au moins de l'ordre de +4,3.

On voit le danger qu'il y a à conclure sans précaution du score observé à la vraie valeur dans le cas de tests peu reliables. Notes revues par le D<sup>r</sup> M.-P. Schützenberger.

**M. P. SCHÜTZENBERGER**

Année 1949 1949-9. À propos de la "cybernétique" (mathématiques et psychologie)

*L'ÉVOLUTION PSYCHIATRIQUE*

*N° IV - Année 1949*

*Pages 585 à 607*

## A PROPOS DE LA "CYBERNÉTIQUE" (MATHÉMATIQUES ET PSYCHOLOGIE)

par *M. P. SCHÜTZENBERGER*

*Attaché à l'Institut National d'Hygiène*

La publication de la « Cybernétique » de Norbert WIENER (1) aura fait pousser de hauts cris à certains : introduire dans la psychologie, domaine de l'« Einfühlung » et du « Verstehen », les concepts et l'esprit de la physique mathématique, établir un parallèle même lointain entre le fonctionnement des machines électroniques et certains aspects de la vie mentale, c'était jeter un défi à ceux qu'effraie encore l'ombre de la statue de CONDILLAC. Pour d'autres le scandale était aussi grand qui pensaient avoir déjà fait la part du feu avec la concession de la psychométrie à ceux qu'ils se contentaient de flétrir du nom de statisticiens. Et pour ceux-là même, la cybernétique était une atteinte à leurs privilèges puisqu'elle employait des groupes continus de transformation et des algèbres de BOOLE encore suspects d'esprit révolutionnaire dans l'enseignement classique élémentaire au lieu des respectables coefficients de corrélation.

Et pourtant, depuis quelques décades, préparant l'apparition de cette discipline nouvelle, des travaux nombreux bien que dispersés avaient montré la possibilité — sinon le besoin — d'une mathématisation de certains chapitres de la psychologie.

Un théoricien de la Gestalt Psychologie (2) employait les mots les plus modernes de la topologie pour décrire les « champs sociologiques » et ses élèves, dans le même but, recouraient au calcul matriciel (3). D'un autre côté, le grand algébriste BIRKHOFF (4) s'essayait à une approche expérimentale des fondements de l'esthétique

Depuis la physiologie des sensations, dès l'origine physico-mathéma-

(1) Norbert WIENER : *Cybernetics or control and communication in the animal and the machine* (1948).

(2) — Kurt LEWIN — *Principles of topological Psychology* (1936).

(3) — L. FESTINGER. *The analysis of sociograms*. (1949). *Hum. Rel.* (2) Pp 153-159.

(4) — G. D. BIRKHOFF, *Aesthetic measure* (1933).

tique de par son objet même, jusqu'à l'exposé : *more geometrico* par HULL (1) de la théorie des réflexes conditionnés, on assistait à un assaut général pour utiliser dans le domaine des sciences de l'homme l'outil qui avait assuré, par l'intermédiaire de la physique, de tels succès dans la connaissance et la maîtrise de la nature.

Il est encore trop tôt pour établir un bilan, fût-il provisoire, de la cybernétique et prématuré même de fixer des limites à ses recherches mais il n'est peut-être pas inutile de la replacer dans ce courant de pensée qu'elle s'efforce d'intégrer et sur lequel elle a attiré l'attention. Plus spécialement, laissant à d'autres plus qualifiés que nous la charge de discuter les aspects proprement techniques de ces travaux, nous essayerons de les situer en fonction du rôle méthodologique qu'y jouent les mathématiques. Aussi bien, nous pourrions ainsi mieux les référer au cadre scientifique général dans lequel ils se développent et qu'ils sollicitent.

Et comme il faut bien classer malgré l'arbitraire et le schématisme, nous distinguerons trois niveaux de jonction :

— Celui de l'expérimentation et de l'observation : c'est la *statistique mathématique appliquée à la psychologie* ;

— Celui du maniement de concepts et de lois fournis eux-mêmes, à chaque pas par la psychologie. Certaines conséquences seulement peuvent alors, moyennant des hypothèses d'approximation, être soumises au traitement analytique, puis confrontées à la réalité. C'est l'utilisation la plus simple de l'outil mathématique, comme prolongement numérique du raisonnement direct. C'est la *méthode des modèles*.

— Enfin, au troisième stade, les concepts initiaux sont par nature, physico-mathématiques — leur développement s'effectue sur le plan abstrait et dans la structure de certains des êtres ainsi formés; le psychologue est invité à retrouver les phénomènes psychiques qui lui sont familiers. C'est la *méthode axiomatique* qui a déjà conquis la physique mathématique moderne. Dans cette revue forcément limitée de théories récentes et parfois mal connues en France, nous avons dû opérer une sélection des moments les plus caractéristiques, laissant parfois dans l'ombre des domaines entiers. Un autre écueil, difficile à éviter, résidait dans l'appareil mathématique, généralement peu classique, nécessaire à une exposition complète. Nous espérons y avoir remédié partiellement en développant un peu longuement quelques exemples concrets, et, dans l'impossibilité de citer tous les mémoires originaux importants, en renvoyant à une série d'ouvrages de base indispensables à une étude plus approfondie.

\*  
\*\*

(1) Clark L. HULL — Principles of Behaviour (1949).

Année 1949 1949-9. À propos de la "cybernétique" (mathématiques et psychologie)

M. P. SCHUTZENBERGER

587

## STATISTIQUE ET PSYCHOLOGIE

C'est là pour beaucoup de psychologues l'unique possibilité d'application des mathématiques à leurs recherches et c'est d'ailleurs la seule qui ait déjà conduit à des résultats pratiques sur une large échelle.

Sous la nécessité de l'action et avec la complicité de l'idéologie dominante en psychologie différentielle (ne s'agissait-il pas de marquer chaque homme d'une aptitude lui conférant une place irrémédiable dans une société irrémédiable aussi ?), la psychométrie a enregistré des succès notables. Aussi n'est-il pas étonnant qu'on identifie — à tort — *statistiques, statistique mathématique* et même *méthode mathématique* quand la solution des problèmes apparemment les plus importants nécessite la coopération de tout un bureau de calcul largement pourvu de machines à cartes perforées.

Nous laisserons de côté l'emploi des méthodes statistiques désormais usuelles et qui sont communes à tout un ensemble de disciplines, comme la météorologie ou la génétique où l'on ne saurait se passer de « chi carrés », non plus que de « t de Student ». Mais, au centre de la psychométrie, il convient de discuter la signification de L'ANALYSE FACTORIELLE qui, elle, est directement issue des recherches psychologiques.

L'analyse factorielle est le visage théorique de la méthode de régressions linéaires multiples, elle-même fondement géométrique à l'*ingenueering problem* de la sélection et de l'orientation de masse du personnel.

Depuis les premiers travaux de SPEARMAN (1), au début de ce siècle l'analyse factorielle est devenue majeure. Ses méthodes se sont affinées en se diversifiant sous l'influence de la rivalité qui oppose les tenants de l'école anglaise aux élèves de THURSTONE (2). Ses domaines eux aussi se sont élargis : d'abord l'intelligence et les aptitudes sensori-motrices, puis le caractère et la personnalité (3), et même par une sorte de retournement la typologie psychologique (4) (5). Jusqu'à déborder le cadre de la psychologie comme l'attestent des applications — les plus valables peut-être — aux problèmes de la biométrie (G. TESSIER) (6). Certains veulent y voir le triomphe des méthodes analytiques en psychologie et en parlent même comme d'une « psychologie mathématique ». D'autres lui nient presque toute valeur — qui peut-être la connaissent moins bien.

La théorie factorielle nous servira ici d'exemple type du rôle des sta-

(1) Ch. SPEARMAN : General intelligence objectively determined and measured (1904) *Am. J. Psy.* (15) Pp. 201-293.

*The Abilities of man*, 1932.

(2) L. L. THURSTONE — Multiple factor analysis (1947).

(3) EYSENK : Dimensions of personality.

(4) W. STEPHENSON : Correlating persons instead of tests (1935). *Char. and Personal.* (4)

(5) C. BURT : The factors of the mind (1940).

(6) G. TESSIER. Un essai d'analyse factorielle... (1938). *Biotypologie* Pp. 73-96.

tistiques dans un domaine comme celui des facultés supérieures ou de la personnalité où l'expérimentation est pauvre et où les « gros phénomènes » sont encore trop rares et trop isolés.

A la base on postule que les performances d'un individu, (plus généralement ses comportements), sont la résultante d'un certain nombre de causes hypothétiques, (les facteurs) *indépendantes, et totalement intérieures à chaque individu*. Réciproquement, celui-ci est parfaitement défini par la possession de chacun de ces facteurs à un degré plus ou moins élevé. C'est donc un atomisme des fonctions mais qui se sépare radicalement des théories antérieures par son approche vers la définition de ses éléments. Refusant (provisoirement) les catégories à priori telles que l'attention, la mémoire, la rapidité, etc., on n'admettra comme facteurs que ceux qui seront livrés par l'observation et le calcul au moyen de la technique de factorisation.

Mais, pour pouvoir soumettre ces faits observés à l'algorithme choisi, une série d'hypothèses (mathématiques celles-ci) est nécessaire :

- *normalité* : (caractère « gaussien ») de la distribution des facteurs dans la population ;
- *proportionalité* entre la performance (ou le comportement) du sujet et le degré auquel il possède le facteur en cause ;
- *additivité* simple de l'action spécifique de chaque facteur.

Mais, il faut bien le noter, ces hypothèses sont bien plus imposées par les limites actuelles de nos possibilités techniques de calcul qu'elles ne sont l'expression d'une structure supposée du fonctionnement mental. Schématiquement, elles reviennent à étendre à la totalité de l'échelle de variations des causes l'approximation linéaire usuelle dans chaque voisinage restreint. Elles postulent, implicitement, que les instruments d'observation sont tels qu'aucune différence qualitative ne bouleverse cette échelle quand les facteurs varient quantitativement d'un extrême à l'autre. Enfin la légitimation de ces hypothèses est proprement psychologique, le calcul des probabilités n'étant appelé que pour constater qu'aucun désaccord flagrant n'interdit l'application de la méthode.

A partir de ces prémisses, le statisticien est requis d'extraire les facteurs du matériel expérimental qu'on lui soumet. Il ne saurait s'y refuser : toute technique factorielle appliquée à tout ensemble de chiffres fournit des facteurs selon les hypothèses dont on l'a munie au départ et qui la différencient des autres techniques (1). De manière plus ou moins ouverte ces hypothèses complémentaires contiennent déjà le postulat d'existence d'une certaine structure factorielle. Dans deux batteries de tests pratiquement équivalentes une méthode multifactorielle met en valeur les « *primary mental*

(1) On trouvera une discussion complète des diverses méthodes dans : G. Thomson (1939 et nouvelle édition en 1948). *The factorial analysis of human abilities*.

*abilities* » (verbales, numériques, de mémoire brute, de visualisation, et une méthode uni-factorielle retrouvera d'abord un *facteur général prédominant*. Même si ces désaccords sont finalement moins graves qu'ils n'apparaissent, car des médiations existent (*facteur du second ordre, facteurs de groupe*), qui réconcilient les deux théories et préservent la valeur des résultats acquis, ils posent avec urgence le problème de la signification des facteurs. Mais, sur le terrain de l'analyse factorielle, cette question est extra-statistique, extra-mathématique : aucune esquisse de modèle n'ayant été conçue, le mathématicien peut au plus indiquer si certains effets sont imputables au seul hasard, c'est typiquement la position de Lawley (1). Et le psychologue, (qui avait peut-être espéré plus de son bureau de calcul), est de nouveau face à face avec le problème qu'il se posait au départ et pour la résolution duquel il n'a qu'une élaboration plus riche et plus claire de ses chiffres expérimentaux.

De cette contradiction découle une double conséquence : plus modestes ou plus réalistes certains factorialistes ne prétendent plus qu'à voir résumées en quelques formules leurs volumineuses tables d'intercorrélation ou même leurs chiffres initiaux. A ceci les invitera par exemple la méthode de HOTELLING (2). La statistique semble alors retourner aussi à ses taches originelles : elle ne teste pas une théorie probabilitaire, elle n'estime pas des paramètres au mieux d'observations entachées d'erreurs selon une loi connue ; elle fonctionne comme une méthode d'enregistrement et de condensation des données brutes.

Mais ce rôle elle l'assume en le dépassant : avec la théorie du « design of experiments » de R. A. FISHER qui s'enseigne maintenant en Angleterre dans des chaires spéciales, la statistique prend un sens nouveau : elle cherche à se placer avant toute expérimentation pour donner à celle-ci le maximum d'efficacité par un agencement rationnel du dispositif d'observation. A cet égard, il est symptomatique de voir l'importance croissante que prennent en psychologie même des techniques comme l'analyse discriminative (3) ou l'analyse de variance (4). Dans l'une comme dans l'autre, les axes, les « di-

(1) D. N. LAWLEY : « The estimation of factors loadings by the method of maximum likelihood » (1940). Proc. Roy. Soc. Edim. (60) — Pp. 64-82.

(2) H. HOTELLING : « Analysis of a complex of statistical variables into principal components (1933), Jour. Ed. Psy. (24) Pp. 417-441, 498-520.

(3) Pour un exposé technique voir : C. R. Rao — The utilisation of multiple measurements in problems of biological classification (1948).

Et aussi : C. A. B. Smith. Some examples of discrimination (1947) Ann. of Eug. (13) 272-274. Et spécialement en psychologie :

M. S. Bartlett. Internal and external factor analysis (1948). Brit. J. Psy. Stat. suppl. (1). Pp 73-81.

(4) H. E. Garret et and J. Zubin, The analysis of variance in psychological research (1943). Psy Bull. (40). 233-267.

D. A. Grant. On « The analysis of variance ». (1944). Psy. Bull (41) 158-166.

mensions » sont inscrits à l'avance dans le schéma opératoire. Dans l'analyse discriminative, l'accent est mis sur le pouvoir d'un système de mesures, de retrouver à posteriori une classification posée à priori par le choix de deux ou plusieurs groupes sélectionnés. Dans l'analyse de variance on cherchera plutôt à inventorier les effets et les interactions d'un système de causes en utilisant un matériel strictement choisi à l'avance en fonction de celles-ci.

Cependant quelque brillant que soit ce rôle de la statistique comme guide de l'expérimentateur, quelque avantage qu'en puissent retirer sans doute d'autres domaines de la psychologie, il faut reconnaître qu'il est purement incident.

L'intervention du calcul des probabilités ne s'impose que du fait des contingences multiples qui troublent l'observation et que l'on répute aléatoires pour pouvoir les éliminer par l'opération « *moyenne de :* » Le déterminisme est strict et il ne s'agit que de le retrouver. Ceci n'est d'ailleurs évidemment pas une critique (au contraire) mais la marque d'une limite d'utilité.

Idéalement on pourrait se dispenser de ce formalisme et de ces longs calculs si les expériences pouvaient être faites à une échelle suffisamment large. *La statistique mathématique* n'est nécessaire que parce qu'il est impossible de recueillir assez de *statistiques*, — et seulement pour cela, même si l'impossibilité est essentielle puisque l'on ne peut s'adresser qu'à une partie de la population présente pour établir une loi sur la population présente et future. Si la statistique est un moyen puissant de lever cette contradiction (sur laquelle a buté plus d'un philosophe du cercle de Vienne), elle n'a cependant pas encore un rôle déterminant dans la psychologie des aptitudes.

Une technique factorielle (celle des « *Cluster* » de CATTELL) (1). montre d'ailleurs la possibilité d'une démarche non mathématique. Elle doit être interprétée comme la recherche plus ou moins naïve (analytiquement s'entend !!) de groupes de traits de caractères hautement liés entre eux, — c'est dans ce domaine le chaînon entre la statistique et les statistiques.

Mais, d'ailleurs, ces statistiques qui parleraient d'elles-mêmes ne sont nécessaires qu'en raison des distorsions qu'imposent aux phénomènes mille circonstances concrètes, et combien subtiles. On pourrait concevoir, à la limite qu'elles seraient rendues négligeables par une stricte élimination de tout sujet ne répondant pas aux conditions prescrites par un protocole minutieux. Evidemment, c'est là un dessein utopique que dément bien vite toute tentative de recherche selon ces seuls canons. La statistique permet en partie, d'en limiter les exigences et d'en corriger les imperfections. Selon le domaine on jugera son rôle primordial ou modeste.

(1) R. B. CATTELL : « Cluster search methods (1944) — *Psychometrika* (9), Pp. 169-184  
« Description and measurement of personality » (1946).

## II — LA BIOPHYSIQUE MATHÉMATIQUE DU SYSTÈME NERVEUX

C'est peut-être parce qu'ils pensaient que les mathématiques pouvaient leur apporter bien autre chose que l'affinement de leurs méthodes d'observations que divers psychologues se sont efforcés de lui faire une place plus grande dans leurs recherches. Déjà l'interprétation proposée par G. THOMSON (1) des théories factorielles de SPEARMAN représentait un pas décisif vers la construction de modèles mathématiques : on suppose que chaque performance implique la mise en jeu d'un certain nombre « d'unités » anatomo-physiologiques (les « *bonds* ») regroupées en une série d'ensembles se chevauchant mutuellement. La corrélation entre les scores de deux tests est la résultante de cette communauté de mécanismes élémentaires intervenant simultanément dans les deux tâches. Des hypothèses très générales de distribution statistique permettent alors de retrouver comme conséquence la loi d'annulation des Tétrades que l'on peut considérer comme un fait expérimental assez bien établi (sous certaines conditions). Il n'est plus besoin de postuler un facteur général ou plutôt celui-ci n'est qu'une interprétation de la réalité, une commodité de langage.

Ici, au contraire de ce que nous avons déjà rencontré, les hypothèses de base ne portent plus sur les modalités d'apparition du phénomène mais sur son mécanisme, sur sa structure même : il n'y a pas seulement description mathématique mais explication.

De semblables tentatives, extrêmement nombreuses (et de valeur fort inégale d'ailleurs) ont eu lieu dans bien d'autres domaines de la psychologie. Sans essayer d'en donner une revue même sommaire nous nous contenterons d'étudier plus en détail la « *bio-physique mathématique du système nerveux* », qui présente en outre l'avantage d'être comme un avant-propos à la cybernétique.

Quel est l'objet de cette théorie ?

Essentiellement, elle se propose de construire des modèles schématiques de neurones et de systèmes de neurones satisfaisant à une triple exigence :

- ne pas être en contradiction avec les données actuelles de l'anatomie et de la physiologie nerveuse ;
- être assez simples pour se prêter convenablement à un traitement mathématique relativement approfondi ;
- représenter sans trop de désaccord les faits d'observation chaque fois qu'il est possible de conduire la théorie assez avant pour qu'elle les rencontre.

On sera peut-être surpris de la teneur et de la hiérarchie de ces condi-

(1) G. THOMSON : « On complete families of conclusion coefficients... » (1935), *But. Jour. Psy.* (26), pp. 63-92.

tions ; on trouvera que si la première va de soi, la seconde n'a qu'une importance minime eu égard à la nécessité pour une semblable construction de mener à des conséquences vérifiables. De fait, on doit voir là le rôle considérable de la personnalité et des conceptions de M. RACHEVSKY (de l'Université de Chicago), qui n'ont pas été sans influencer dans ce domaine la plupart des chercheurs : dans son livre monumental, (1) RACHEVSKY étudie toute une série de questions (de la division cellulaire au jugement esthétique en passant par les rapports entre le poids et l'appareil locomoteur des êtres vivants), non en vue d'en constituer d'abord une théorie valable selon les critères de la méthode hypothético-déductive mais avec la volonté « provocante » (« *stimulating* ») de démontrer par des exemples multipliés la possibilité d'une approche mathématique aux phénomènes de la vie.

Pour cela il lui faut naturellement, et de manière fort consciente, choisir les problèmes qu'il peut géométriser plutôt que géométriser des problèmes choisis.

Selon ces principes, deux schémas principaux de fonctionnement neuronal ont été élaborés. Dans l'un (RACHEVSKY, puis LANDAHL, HOUSEHOLDER, RAPPOPORT) (2) on discute les circuits du point de vue dynamique du développement dans le temps de leurs états d'excitation. Dans l'autre, (Mc CULLOCH et PITTS) (3), le neurone encore simplifié n'est plus qu'une unité inanalysable, — c'est un véritable relai électrique dont seul l'agencement des interconnexions peut être objet d'études (4).

Comme exemple de problème étudié grâce au premier schéma, nous choisirons le plus simple : celui de la variation du temps de réaction en fonction de l'intensité du stimulus.

Une grandeur «  $e$  » caractérise l'état d'excitation d'un neurone et, par définition son action sur tout neurone efférent. — En l'absence de stimulation extérieure cette grandeur décroît proportionnellement à elle-même, — c'est-à-dire selon une loi exponentielle.

Inversement elle peut croître (proportionnellement aussi) en fonction de la fréquence  $f$  des impulsions afférentes ; d'où ( $a$  étant un paramètre), la première équation :

$$\frac{de}{dt} = -a(e-f)$$

Toujours en première approximation on posera :

- (1) N. RACHEVSKY : *Mathematical Biophysics* (1938).  
 (2) — A. S. HOUSEHOLDER and H. D. LANDAHL, *Mathematical biophysics of the central nervous system* (1945).  
 (3) — W. S. Mc CULLOCH and W. PITTS. • *A logical calculus of the ideas immanent in the nervous activity*. *Bull. Math. Bioph.* (1943) Pp 115-133.  
 (4) TELSON WEY. *On matrices of neural nets* (1948).  
*Bull. Math. Bioph.* (10). 63-67.

$f = b \log. (S/h)$  pour  $S$  plus grand que  $h$ ,  
 $= 0$  pour  $S$  plus petit que  $h$ ,  
 avec  $b$  : paramètre caractéristique du récepteur,  
 $S$  : intensité du stimulus.  
 $h$  : valeur liminaire.

Le circuit nerveux étant enfin supposé réduit à sa forme la plus simple possible (deux neurones, l'un récepteur, l'autre effecteur), les méthodes classiques du calcul différentiel et intégral conduisent à représenter le temps de réaction  $t_r$  par :

$$t_r = t_0 - 1/a \log (1 - h/b \log (S/h))$$

Les paramètres inconnus ( $t_0$ ,  $a$ ,  $b$ ,  $h$ ) sont estimés dans chaque cas par les procédés habituels de la statistique. Avec les observations (notamment celles de H. PIERON sur les stimuli auditifs et gustatifs) on constate un accord très satisfaisant.

A vrai dire il est assez difficile d'en apprécier la vertu probante en raison même du grand nombre des constantes qui peuvent être ajustées. Dans d'autres cas, où la courbe expérimentale est de forme moins simple, la concordance est plus significative. Mais dès que l'on essaye d'appliquer la méthode à des questions complexes (seuils différentiels, vision stéréoscopique, apprentissage) les calculs deviennent effroyablement pénibles et ce n'est que par le recours à des hypothèses simplificatives d'approximation, souvent assez troublantes pour le mathématicien, qu'ils peuvent être achevés. De plus, ces fonctions que l'on a introduites à l'origine contiennent déjà une large part d'arbitraire : caractéristique de la méthode des modèles, on ne leur demande que de représenter qualitativement l'allure générale, par exception seulement la loi, des phénomènes. Dans leur choix même se retrouve alors ce critère idéaliste de commodité dont POINCARÉ s'était fait l'avocat.

Dans le modèle de Mc CULLOCH et PITTS, qui nous achemine vers le point de vue axiomatique, toutes les approximations sont au contraire ramassées dans les postulats initiaux :

- l'unité de temps convenablement choisie, on suppose que toutes les décharges neuroniques s'effectuent à ce rythme ;
- $N_i(t)$  symbolise la décharge du temps  $t$  du neurone ( $i$ ) ; celle-ci provoque sous certaines conditions la décharge (ou l'inhibition)  $N_j(t+1)$  d'un neurone efférent ( $j$ ) au temps  $t+1$ .

Tout fonctionnement d'un circuit nerveux peut être alors représenté par une proposition logique au moyen des signes de la logique symbolique (1) :

$\& (et)$  :  $N_i(t_i) \& N_j(t_j)$  se lit :  
 « décharge de ( $i$ ) à ( $t_i$ ) ET de ( $j$ ) à  $t_j$  »  
 $V$  : (*vel*) :  $N_i(t_i) V N_j(t_j)$  se lit :  
 « décharge de ( $i$ ) à ( $t_i$ ) OU de ( $j$ ) à  $t_j$  »

(1) Voir par exemple W. V. Quine, *Mathematical logic* (1940).



cette forme mécanique, rigide, dont nous venons de donner un aperçu et il n'est pas impossible, au contraire, que la situation change avec l'introduction récente des méthodes probabilistes.

Sur le plan épistémologique, on voit sans peine l'influence des théories physiques du milieu du XIX<sup>e</sup> siècle et le désir de ramener les phénomènes à ce jeu de leviers et de poulies que réclamait Lord KELVIN. Que les organes élémentaires soient plutôt des solénoïdes que des roues d'engrenage ne traduit que le progrès technique, tout comme le remplacement des équations aux dérivées partielles par les anneaux de Boole (c'est le nom général des systèmes algébriques munis d'opérations telles que  $\&$ ,  $\vee$  et  $\neg$ ) ; le but est le même : construire un modèle tangible tel que le calcul permette d'en déduire des conséquences numériques une fois attribuées aux constantes des valeurs convenables ; les concepts et les lois sont ici physiologiques et déduits à chaque pas de l'expérience. Il ne s'agit que de les coordonner et d'en projeter plus loin certaines conséquences : on pourrait parler de « mathématisation locale » et c'est un stade déjà supérieur à celui que nous avons trouvé dans la psychométrie. Si fructueuses que puissent être ces recherches, dans l'état actuel de nos connaissances, il est cependant intéressant de voir s'il n'est pas possible d'aller plus loin.

### III — UNE TENTATIVE D'AXIOMATISATION

Le troisième niveau de jonction entre mathématique et psychologie est certainement représenté le plus parfaitement par cette « *théorie des jeux et au comportement économique* » (1) due initialement à VON NEUMANN et qu'il a développée en collaboration avec l'économiste MORGENSTERN. D'autres exemples, beaucoup moins purs se rencontrent d'ailleurs dans l'œuvre de ZIPF (2) où est tentée une réduction des processus sociologiques à un principe maupertuisien de moindre effort ou dans la théorie des guerres de Lewis F. RICHARDSON (3) ou encore dans le livre récent de RACHEVSKY (4).

Dans la théorie de VON NEUMANN il n'existe pas de référence à un substrat physique dont les équations serviraient à symboliser le mécanisme. Il n'intervient non plus que le minimum d'observations psychologiques ou sociologiques dont des schémas analytiques plus ou moins fragmentaires chercheraient l'approximation. Au contraire, dès l'origine, des postulats, les plus généraux possibles, choisis à coup sûr en fonction de considéra-

(1) J. von NEUMANN and OSKAR MORGENSTERN. • *Theory of games and economic behaviour*. (1946).

(2) Kingsley ZIPF: *Human behaviour* (1949).

(3) L. F. RICHARDSON. • *War Moods* • *Psychometrika* (1948) Pp. 147-173 et 197-232.

On remarquera la communauté de propos de ces travaux dont l'objet final est l'explication la plus abstraite possible des crises et des guerres et on ne sera pas surpris qu'ils recourent tous à des méthodes axiomatiques et formelles impliquant d'elles-mêmes l'impuissance de l'homme vis-à-vis des fléaux sociaux.

(4) N. RACHEVSKY : *Mathematical Theory of human relations* (1947).

tions extra-mathématiques, qui les légitiment à priori, mais accessibles de façon directe au traitement mathématique. Ce sont des hypothèses géométriques à contenu psychologique — non des constats psychologiques géométrisables.

Prenons l'exemple de la théorie du « jeu » entre deux adversaires : les règles de la partie définissent de façon précise et dans chaque cas concret une fonction  $g = f(T_1, T_2)$  des stratégies  $T_1$  et  $T_2$  de chacun des deux joueurs.  $g = f(T_1, T_2)$  est le gain moyen du joueur quand il adopte la stratégie  $T_1$  et que son adversaire adopte  $T_2$ . De même que la dépendance de  $f$  de  $T_1$  et de  $T_2$  est fixée par la règle d'un jeu, l'univers des stratégies possibles est aussi défini.

On cherche alors la valeur  $t_1$  de  $T_1$  qui rend  $g$  maximum; quand  $t_2$  est choisi au contraire de sorte à diminuer le plus possible cette grandeur. Ce couple de valeur  $t_1$  et  $t_2$  est appelé « les stratégies optimales ». Analytiquement, il réalise  $\text{Max. } T_1 (\text{Min. } T_2 (f(T_1, T_2)))$ .

Une analogie matérielle fera mieux saisir le principe : soit une chaîne de montagne orientée E. O. que traverse un col N. S. Il s'agit pour le premier joueur de fixer un point le plus haut possible, mais il ne dispose que du choix du parallèle. Son adversaire, dont les desseins sont opposés détermine finalement le point en optant pour tel ou tel méridien. Ici le gain du premier joueur est l'altitude du point ; l'ensemble de ses stratégies possibles est l'ensemble des parallèles (respectivement, des méridiens pour le deuxième joueur) la stratégie optimale du premier joueur est évidemment le choix du parallèle passant par le sommet du col puisque c'est là le point le plus bas auquel pourra, au pire, le fixer son adversaire, tout autre choix laissant à ce dernier une possibilité de réduire encore plus l'altitude finale. Après avoir démontré que toute situation de jeux à deux compétiteurs se ramenait précisément à ce schéma, on prouve que  $\text{Max. } t_1 \text{ Min. } t_2 = \text{Min. } t_2 \text{ Max. } t_1$  ; c'est-à-dire que la théorie conduit aux mêmes résultats quand elle est développée du point de vue du second joueur (ce qui est intuitif dans notre exemple) : Il est alors possible de déterminer numériquement ces stratégies optimales et de montrer qu'elles jouissent de la propriété de rendre minimum la perte maximum. Ce qu'elles offrent est donc avant tout une protection contre le risque et la seule hypothèse psychologique — qu'on peut ou non refuser — est que tel est bien le mobile profond de la conduite effective des individus en certaines circonstances définies (1).

(1) Que cet axiome soit trop pauvre pour rendre compte du comportement économique strictosensu des individus ou des sociétés capitalistes est assez évident. — Il semble cependant s'appliquer assez bien dans les situations de jeu. — Entre adultes tout au moins car nous tenons d'une communication personnelle de Mr. Léon Festinger que d'après des expériences (non publiées) de Mosteller il s'appliquerait mal au jeu chez les enfants, — la théorie de von Neumann ne permettant pas l'introduction de l'élément émotionnel qui est alors prédominant.

La portée de ces considérations dépasse largement celle d'une mise en forme pure et simple : on y trouve d'abord la solution de cette récurrence infinie, véritable paradoxe éleate, où serait conduit le joueur d'Edgard POE s'il rencontrait un adversaire à sa taille. (1)

Il ne s'agit pas seulement de construire une martingale contre un système fixe qui ne réagirait pas, — mais de fonder une théorie de certains rapports interindividuels. Auteur d'une des œuvres les plus importantes sur les fondements de la physique quantique, VON NEUMANN a fort bien expliqué lui-même la nouveauté du problème : dans toute question de physique, il ne s'agit en définitive que de l'étude des extremum d'une certaine fonction. C'est aussi la caractéristique des relations d'un sujet à l'objet. Par contre, ici, ce sont les relations entre deux sujets antithétiques, en permanente réaction mutuelle, qu'il s'agit d'approfondir. Aussi l'outil mathématique lui-même doit se renouveler et il n'est pas surprenant qu'intervienne la théorie des corps convexes, relativement si peu utilisée ailleurs.

Et enfin cette position originale pour le mathématicien lui permet d'éclairer dans d'autres domaines des solutions auxquelles il était arrivé par d'autres voies : assimilons à un joueur l'observateur en face d'une expérimentation statistique, (son adversaire) qui ne se révèle partiellement à lui qu'à travers un mécanisme, probabilitaire peut-être, mais d'abord, et surtout, inconnu. La stratégie optimale est alors la technique qui permet d'extraire le maximum d'information à moindre frais pour un risque dont la limite est donnée à l'avance. Destinée initialement au contrôle des fabrications industrielles et forme particulièrement hardie de cette statistique comme méthode expérimentale que nous évoquions plus haut, — l'analyse séquentielle d'Abraham WALD (1) est un autre aspect à peine différent du jeu à deux partenaires tel que le pose VON NEUMANN.

A l'intérieur même de la situation interindividuelle, la théorie nous offre d'autres particularités intéressantes; construisant une variété de poker (très simplifiée pour les nécessités du calcul), VON NEUMANN découvre que la

(1) (« Stolen letter »). Le jeu se schématise ainsi : A et B choisissent indépendamment et à l'insu l'un de l'autre pile ou face. Si les deux choix sont reconnus identiques A gagne sinon B l'emporte. La stratégie optimale, — qui assure à chaque joueur le gain minimum de 1 partie sur 2, — consiste pour chacun à choisir au hasard tantôt pile, tantôt face avec une probabilité de 0,50. La règle du jeu peut être modifiée, la coïncidence des deux pièces apportant plus ou moins à A selon qu'elle s'est effectuée sur pile ou sur face ; et de même les deux discordances possibles étant pénalisées différemment. La théorie indique alors qu'elles doivent être les probabilités avec lesquelles chaque joueur a avantage à tirer au hasard le côté qu'il choisit.

On remarquera que cette situation est aussi (entre autres) celle du pointeur de D. C. A. et du pilote de l'avion : il y a ou non coïncidence de deux trajectoires. Le pilote peut choisir sa feinte (à gauche ou à droite) — son adversaire au sol devant la deviner.

Depuis la publication du livre de VON NEUMANN aucun travail important nouveau n'a plus paru au sujet de cette théorie...

(1) Abraham WALD (1947) — *Sequential Analysis*.

stratégie optimale comprend en général l'obligation de « relancer » avec une main faible : c'est-à-dire de « bluffer ». Et lorsque l'on enrichit le modèle on voit apparaître des nuances plus fines : bluff agressif chez le premier joueur (relance élevée avec un jeu faible), défensif chez le second (« voir » de temps à autre avec une main moyenne). On peut même préciser de quelle manière doivent être infléchies ces formules pour tirer avantage au maximum de la faute de l'adversaire, de sa non observance des règles de la stratégie optimale.

Dans ce domaine où l'explication « intuitive » semblerait devoir prédominer, un procédé mathématique permanent apporte, non seulement des précisions numériques, mais directement des processus que l'on peut bien qualifier de psychologiques. Ceci est d'autant plus remarquable qu'il n'y a pas là le produit d'un modèle construit *ad hoc*, que tout arbitraire semble avoir été éliminé dans le choix des fonctions puisque celles-ci sont fournies, de l'extérieur, par les règles du jeu.

C'est le miracle de la méthode mathématique qui non seulement précise les conditions et les modalités du phénomène mais encore l'explique seule pleinement. Et cela en même temps qu'elle abolit l'explication que nous cherchions sur un plan trop proche pour être efficace. En physique, c'est le remplacement des lois de KEPLER par la loi unifiante de l'attraction universelle.

Nous n'avons pas la place ici de discuter comment VON NEUMANN passe du jeu à deux adversaires au cas le plus général de  $n$  joueurs, — et comment le psychologue retrouve au détour des calculs des concepts comme celui de domination, de coalition qu'il croyait lui appartenir en propre et qu'il doit replacer comme cas particuliers d'être *abstrait plus généraux*. Dans cette géométrie de la concurrence universelle, disparaissent encore des indéterminées puisque seuls entrent en ligne de compte les gains possibles et non la manière dont ils sont obtenus.

Enfin, dans ce but (à l'opposé de RACHEVSKY et de son école), VON NEUMANN est contraint de forger un outil analytique entièrement nouveau, de faire appel à la théorie des groupes et à celle des relations binaires : c'est encore une caractéristique de la méthode axiomatique que d'être en même temps à chaque fois un nouveau chapitre des mathématiques.

Quelle que soit la valeur de toute la construction du point de vue de son application à la réalité économique (et il faut se rappeler la critique pertinente et radicale qu'en a faite entre autres N. WIENER), la théorie de VON NEUMANN mérite une place importante dans l'histoire des sciences comme le premier essai conscient et cohérent d'étude axiomatique d'une situation psychologique.

## IV. LA CYBERNETIQUE

Quelle est maintenant la place de la cybernétique au milieu de ces théories, en quoi s'en différencie-t-elle ? D'abord en cela qu'elle est avant tout axée sur les problèmes techniques concrets, elle est la science des servo mécanismes et des appareils automatiques. En seconde intention seulement elle se retourne vers la psychologie pour chercher des perspectives et proposer des solutions. Cette dépendance étroite des nécessités de la praxis l'empêche de tomber dans cette arbitraire et cette gratuité que l'on peut reprocher à certaines des recherches que nous avons évoquées plus haut. Pour saisir la raison de cette liaison qui est la substance même de la cybernétique, il n'est que de voir, même très sommairement, sa situation historique: La physique du XIX<sup>e</sup> siècle a été comme centrée sur le problème de l'énergie, de sa conservation, de ses transformations : c'était la clé du développement de toute industrie, et cette préoccupation dominante s'est réfléchiée dans les autres disciplines ; la physiologie a été en grande partie celle de l'organisme vivant envisagé comme machine à vapeur. Avec la solution, même provisoire, des aspects les plus urgents de ces problèmes d'autres nécessités sont apparues progressivement : de plus en plus ce n'est pas seulement le côté mécanique du travail de l'ouvrier, sa force physique, qu'il s'est agi de remplacer à meilleur compte, mais aussi les aspects les plus complexes de son habileté technique, voire même certaines de ses facultés intellectuelles.

La cybernétique est née de l'effort scientifique nécessaire à cette mécanisation sur une large échelle de secteurs nouveaux de l'activité humaine. Dans ce domaine qui est le sien propre elle a enregistré des succès prodigieux avant même d'avoir pris conscience de son existence autonome en tant que système cohérent de doctrines.

Sans que l'on puisse dater avec quelque précision le moment d'un changement radical il y a loin du métier à tisser de JACQUART au pilote automatique. Et elle justifie un changement de point de vue la différence qui sépare l'arithmomètre de THOMAS grâce auquel (en 1851) « la racine carrée d'un nombre de 12 chiffres peut être extraite en 1 minute 20 secondes » de l'électro-intégrateur de GUTENMAGER et KOROLKOV dont l'emploi permet aux ingénieurs soviétiques de calculer en quelques heures les caractéristiques optimales d'un nouveau modèle d'hélice, performance qui demanderait des mois de travail même avec les machines à calculer de type classique les plus rapides.

Enhardie par ses résultats, la cybernétique offre à la physiologie et à la psychologie d'étudier l'organisme vivant comme servo-mécanisme et de tirer le maximum d'enseignements de cette confrontation.

Mais n'oublions pas que ceci est un projet, non un bilan malgré certains résultats récents dans l'électro-encéphalographie ou l'étude des phénomènes élémentaires de motricité.

Certains n'ont voulu voir dans la cybernétique qu'une « psychopathologie humaine et électronique » forcément assez dérisoire. Ils ont été abusés par ce qui n'était que suggestions (probablement fructueuses) et non théorie constituée. Le langage anthropomorphique de certains chapitres les incitaient à cette vision sommaire, mais il est commode d'utiliser les termes de « mémoire » de « but », de « conduite » pour nommer et décrire certains « fonctionnements » et n'est-ce pas le cas de bien d'autres concepts de la physique dont l'objectivation ne s'est opérée que progressivement ?

Et d'ailleurs, on ne peut nier qu'elle soit féconde cette tendance de chaque ère technique à se croire à la veille de fabriquer l'homonculus faustien. Aussi, la cybernétique comme géométrisation de la psychologie ne rentre-t-elle qu'à l'état de programme dans le cadre que nous avons cherché à définir. C'est à ce titre que nous en résumerons les principaux chapitres dans ce qui ne sera essentiellement qu'une très brève paraphrase du livre si riche de Norbert WIENER, qui en est à la fois l'acte de naissance et la charte provisoire.

#### L'INFORMATION

De même que le concept d'énergie et ses dérivés, force, potentiel, etc... ont été l'objet d'un travail d'élaboration constant autour duquel se sont orientés de larges groupes de recherches il semble que ce soit celui d'*information* qui constitue à l'heure actuelle le noyau de la cybernétique. Concept extrêmement abstrait dont la diffusion n'est encore probablement pas achevée et qui se trouve au carrefour de bien des préoccupations contemporaines, — avec l'importance qu'elle conduit à attacher à la structuration fine et à l'organisation des parties dans le tout, — avec le primat accordé à la signification sur le signifiant et le signifié. Des sociologues comme LEVY STRAUSS s'en rapprocheront avec l'explication de l'exogamie systématique comme moyen de communication intergroupe parallèle au langage. Pendant ce temps des linguistes appliquent des critères statistiques à la caractérisation des langues artificielles (1). De fait, sous sa forme mathématique, la notion d'information provient presque exclusivement des recherches techniques dans le domaine des télécommunications (2) et c'est à partir de là qu'elle s'est étendue et a reçu un statut précis avec SHANNON (3).

Axiomatiquement l'information est définie comme l'opposé de l'entropie. C'est donc une mesure de l'indifférenciation structurale que notre ignorance ne peut que supposer à l'objet. Ainsi elle est en rapport nécessaire avec le calcul des probabilités puisque toute connaissance ne peut s'exprimer que

(1) — J. C. DODD : « On measuring languages » (1944). J. Am. Stat. Ass. (44) 77-88.

(2) — R. V. L. HARTLEY : « The transmission of information » (1928). Bell. Syst. Tech. J. (3) — 535-546.

(3) — C. E. SHANNON : « A mathematical theory of communication » (1948) Bell. Syst. Tech. J. (27) 379-423 et 623-656.

relativement à un cadre préalable d'hypothèses plus ou moins plausibles à priori : toute assertion est réponse à une implicite question : ou bien... ? ou bien, moins vraisemblablement... ?

La mesure de l'information est définie par une fonction assez simple de ces probabilités à priori. Elle s'accroît obligatoirement à chaque expérience nouvelle qui, réduisant le nombre des possibles initiaux, organise plus étroitement notre connaissance de l'objet. Au contraire elle se dissipe lorsque cette organisation s'affaiblit: c'est ce qui se produit par exemple quand, par l'emploi d'une notation plus grossière nous confondons sous une seule rubrique des données différenciées auparavant. Ou encore quand s'introduit une confusion inséparable d'un facteur inconnu à un autre qui ne l'était que partiellement : c'est le « bruit de fond », les « parasites », qui appauvrissent la structure voulue du message musical en diminuant l'intensité de signification. Incidemment, nous retrouvons là le « design of experiment » et c'est le même processus mathématique qui conduira à sélectionner cette organisation des observations successives ou simultanées grâce auquel est extrait d'une certaine situation le maximum de connaissance, relativement à un certain objectif.

Mais ce n'est pas tout d'en avoir trouvé l'aune, cette information il va falloir la communiquer, la transformer, l'emmagasiner avec le moins de pertes possibles.

Un théorème prouve que pour cela la meilleure forme que puisse avoir un message est celle d'une séquence d'alternatives :

Pour un chiffre, une indication numérique quelconque, c'est sa représentation dans le système binaire « 10 » signifiant 2, « 11 » : 3, « 100 » : 4, « 101 » : 5, etc... Pour une description complexe c'est la méthode des « clefs dichotomiques ». Plus simplement, l'exemple de la photographie montre la possibilité de toute représentation approchée par un réseau discontinu de points noirs et blancs. En cybernétique tout message optimum est un Pissaro. Dans tous les cas le signe élémentaire sera un « oui » ou un « non » et cette utilisation de systèmes « tout » ou « rien » nous ramène au modèle neuronique de Mc. CULLOCH et PITTS. Du point de vue mathématique cette conception toute pythagoricienne, cette prédominance de discontinu entraîne tout naturellement l'emploi d'anneaux booléens, instruments idéaux pour la représentation et l'étude de semblables ensembles. C'est ce que proposaient d'ailleurs dès 1936 COUFFIGNAL (1) (en France) et TURING (2) (en Angleterre) pour réaliser ces « engins à penser » dépassant de loin le piano logique originel de Jevons. La logique formelle se voit ainsi réhabilitée en partie et fixer ses limites comme « vraie métaphysique » des machines à calculer.

De ce système qui constitue la base théorique des grands appareils

(1) L. COUFFIGNAL (1936) — Thèse de doctorat ès-Sciences.

(2) T. TURING (1936) : J. Lond. Math. Soc. (42) — 230-265.

modernes comme l'ENIAC ou l'EDVIAC, découlent pour le physiologiste un certain nombre de modèles, ou plutôt de suggestions pour l'élaboration de modèles :

les opérations élémentaires qui combinent ces « oui » et « non » sont extrêmement simples. Théoriquement même on pourrait les réduire toutes à une seule: la « barre de Sheffer » exprimant la disjonction de deux propositions  $a$  et  $b$  :

$a/b$  : si  $a$  vrai alors  $b$  faux et réciproquement (par exemple «  $a = b$  » s'écrit (se construit) :

$(a/b) / (a/b)$  et «  $a$  ou  $b$  »  $(a/a) / (b/b)$ . De fait il semble préférable de partir des opérations légèrement plus complexes  $\&$ ,  $\vee$  et  $\neg$  dont est susceptible le neurone axiomatique de Mc. CULLOCH et PITTS.

Et même, certaines séquences d'opérations intervenant fréquemment (l'addition de deux chiffres par exemple, ou encore leur comparaison) auront avantage à être regroupées en une seule unité fonctionnelle. Mais ici intervient un second principe: celui de l'économie de matériel. Au lieu d'immobiliser une de ces unités dans l'attente de l'acte unique qu'elle aura à effectuer au cours d'un calcul, il est plus logique de constituer une réserve de tels systèmes élémentaires interchangeable, parmi lesquels un préselecteur, véritable « bureau de planning » industriel viendra chercher l'élément disponible pour l'intégrer à chaque fois dans un système provisoire plus complexe qui sera dissous aussitôt la tâche terminée. Ainsi se balancent avec souplesse l'indifférenciation des organes élémentaires et la structuration du tout.

Mais la multiplicité des « réflexions » élémentaires est la rançon de leur simplicité et la possibilité d'une erreur ne peut être tenue pour négligeable. D'où la nécessité du contrôle.

Celui-ci, naturellement, doit être aussi rapide que le calcul proprement dit. Par la méthode du « contrôle en chaîne » on en fait une partie intégrante de tout le dispositif : le message, l'information, est traitée parallèlement dans une série de séquences fonctionnelles semblables (1). A la fin de chaque opération les résultats sont confrontés dans un comparateur et le message ne poursuit son chemin que s'il y a accord. L'hypothèse de l'impossibilité de deux erreurs menant à des résultats identiques qui justifie cette procédure permet même de réduire à trois le nombre des séries parallèles : au cas où l'une d'elles diffère des deux autres, seul le message de la majorité poursuit son chemin et annule l'autre résultat.

Vient maintenant la mémorisation de l'information que nous sommes d'ailleurs nous-mêmes habitués à réaliser sous forme dichotomique sur les cartes perforées des machines statistiques. Ici on propose un schéma plus

(1) Il est intéressant de voir en rapport avec ceci :  
N. C. WOOSLEY : « Patterns of sensory representation in the cerebral cortex (1947). Fed. Proc. (6) — 437-441.

Année 1949      1949-9. À propos de la "cybernétique" (mathématiques et psychologie)

---

M. P. SCHUTZENBERGER

603

---

dynamique : la série de « tops » qui constitue un message parcourt un circuit fermé. Un organe la lit, la « redresse » si elle a subi une distortion ; la relance pour la reprendre ensuite ; et le même cycle se répète jusqu'à ce que ce souvenir soit « utilisé » ou effacé irrémédiablement.

### LES GESTALT

Ici se greffe, en marge de ces mécanismes déjà largement utilisés ou tout au moins à l'étude, l'une des hypothèses les plus hardies de la cybernétique : celle de la représentation des Gestalt et des « concepts ».

L'information que fournissent les organes des sens n'est pas tellement importante sous la forme brute que possède le message initial envoyé par les récepteurs.

Ce qui par contre est essentiel c'est sa valeur gnosique, de signe (positif ou négatif) de telle ou telle réalité. C'est-à-dire, en définitive son appartenance à une certaine classe de stimuli, tous caractéristiques du même objet mais perçu sous des angles différents. Mais ces diverses modalités appartenant à une même classe ne sont pas une collection quelconque : elles se relient les unes aux autres selon des règles prescrites à priori par la conformation du champ sensoriel en cause. On suppose alors qu'au lieu d'un répertoire de tous ces « synonymes » auxquels serait confronté à chaque fois le message, existe un organe qui compare celui-ci à un seul prototype de la classe envisagée en lui faisant subir successivement toutes les modifications possibles. Notion duale de la notion de classe d'équivalence à la signification de laquelle elle donne seule une plénitude complète apparaît alors le concept de *groupe de transformation* qui s'est montré si fécond dans toute la pensée mathématique moderne. En effet, un seul groupe, donc un seul organe, servira pour toute une série de classes élémentaires comme le montre, dans le domaine de la vision, l'exemple de ces mêmes altérations qu'infligent à tout contour les lois de la perspective. Ainsi est réalisée une double économie fonctionnelle : un seul circuit nerveux représente le « carré », un autre le « triangle », un autre le « cercle », etc... et un seul organe réalisant le groupe des projections linéaires vient seul comparer les formes successives d'un message émis par la rétine. De plus c'est par l'intermédiaire de paramètres intimement liés à la structure de ce groupe que peut s'effectuer de façon systématique cette confrontation. Le *scanning* », (le « balayage ») de la collection des gestalts déjà enregistrés s'effectue selon un rythme synchronisé, pense-t-on, par certaines des ondes électriques que décèlent nos électroencéphalogrammes. Le modèle de circuit conditionable proposé par RACHEVSKY vient à point nommé pour fournir une explication de la mémorisation de nouveaux concepts que leur répétition inscrit dans un enlacement de neurone qu'elle structure. Perdant leur immatériabilité platonicienne les gestalts ne sont plus, sous la forme d'un rythme d'influx nerveux, que les invariants d'un groupe d'opérateurs.

Elles ne préexistent pas à la perception, mais sont susceptibles, une fois constituées de s'interconnecter et de former sous la nécessité de l'action des unités plus riches qui dépassent le cadre perceptif d'un seul organe sensoriel.

Toujours dans le domaine des hypothèses on ne peut passer sous silence l'interprétation que donne WIENER de l'émotion comme message particulier à la transmission duquel concourent les voies hormonales. L'émotion infléchit dans un certain sens les autres messages, d'une structure plus fine comme les messages sensoriels, modifie comme en les survoltant les paramètres caractéristiques des effecteurs. Ceci nous rapproche de thèses plus classiques comme celles de CANNON, de LURIA et des psychologues soviétiques sur le rôle mobilisateur de l'émotion.

Enfin, à cette même partie de la cybernétique se rattachent certaines possibilités d'explications pathogénétiques qui elles aussi sont encore seulement des hypothèses. Citons :

— l'encombrement de certains trajets nerveux, dont l'insuffisance relative de développement provoque des troubles du trafic. WIENER pense en particulier aux rapports entre bégayement et latéralité dus au faible « débit » des commissures cérébrales.

— les idées obsessionnelles envisagées comme un message dont la ronde sans fin à l'intérieur d'un circuit neuronique fermé peut aussi bien être traitée par l'électro-choc (et plus radicalement la lobotomie) que par une psychothérapie qui faisant jouer de nouvelles interconnexions parvient à liquider le trouble.

— l'épilepsie provoquée par la mise en résonance d'un ensemble de circuits sous l'influence d'impulsions extérieures convenablement rythmées (1).

#### LES « FEED BACK » (2)

A l'autre extrémité de l'échelle, la multiplicité des discontinuités, la complexité de leurs lois impose de recourir à un changement de point de vue: c'est le cas par exemple d'un message téléphonique perturbé brownienement par des parasites et qu'il convient de traiter par d'autres méthodes que celles qui étaient efficaces pour une simple série de « tops ». Dans ce but, l'appareil mathématique idéal est l'analyse harmonique avec les transformations de LAPLACE et de FOURIER. C'est d'ailleurs le domaine qu'a cultivé particulièrement N. WIENER.

(1) — V. J. WALTER and W. G. WALTER : « The central effect of rhythmic sensory stimulation » (1949).

E. E. G. Clin. Neur. (1) — 57-86.

(2) Voir du point de vue de la technique: Mc Coll Servomechanisms (1946) H. M. JAMES. N. B. NICHOLS, R. S. PHILIPS : Theory of servomechanisms (1947) H. W. BODE : Network analysis and feedback amplifiers design (1945).

Année 1949 1949-9. À propos de la "cybernétique" (mathématiques et psychologie)

M. P. SCHÜTZENBERGER

605

Ici l'organe élémentaire est directement issu de la technique radio-électrique : c'est le filtre. Le courant d'entrée,  $f(t)$ , variable avec le temps est transformé selon une loi fixe caractéristique de l'appareil en une autre fonction,  $g(t)$ , reliée à  $f(t)$  par une équation différentielle que matérialise un système de résistance, de selfs, et de condensateurs.

Analytiquement ceci revient (sous certaines conditions) à remplacer  $f(t)$  par une fonction linéaire de ses valeurs dans le passé

$$f(t - t'), f(t - 2t'), f(t - nt')$$

On voit le lien avec la partie la plus importante de la théorie de l'information (celle des « times-séries ») où il s'agit précisément de prédire le « futur » de  $f(t)$  en utilisant à chaque instant la connaissance que l'on a de ses valeurs antérieures (1).

Le feed-back, lui est un ensemble plus complexe dont le filtre n'est qu'un élément et qui déborde même le domaine des réalisations électriques : Un exemple de feed-back purement mécanique est le régulateur à boule de Watt, contrôlant la distribution de la vapeur en fonction de la charge de la machine.

Le but élémentaire d'un feed-back est de modifier l'énergie entrant dans un mécanisme de manière à maintenir constante une certaine grandeur qui dépend du fonctionnement de ce mécanisme. C'est le cas d'un thermostat par exemple.

Pour cela, le feed-back réintroduit en quelque sorte dans le circuit l'information qu'il possède à chaque instant sur le résultat de l'organe qu'il contrôle. Dans des systèmes plus complexes, ce n'est pas seulement l'écart avec le but à atteindre qui doit intervenir mais aussi la vitesse avec laquelle cet écart diminue afin de tenir compte de l'inertie de l'engin. Des exigences très précises et très variées ont ainsi amené à réaliser des extraordinaires pointeurs de D. C. A., dans lesquels se combinent les caractéristiques propres à la pièce, les données météorologiques et les renseignements sur la position et la vitesse de l'avion que fournit un radar. Enfin, dans un modèle encore plus perfectionné, le feed back non seulement utilisera ces informations extérieures mais encore dirigera le mécanisme afin de les acquérir. Le but principal d'un système de feed back est donc la réalisation d'un acte conçu dans sa totalité. Le moyen en est une sorte de remise à zéro autour de l'écart actuellement constaté et ceci nous introduit dans la théorie des systèmes de l'équilibre homeostatique. Elargissement de l'équilibre statique où les forces s'annulent mutuellement, de l'équilibre dynamique où des mouvements de vitesse opposés se compensent à chaque instant, l'équilibre homeostatique s'établit dans le temps, toute action extérieure déclenchant une réaction qui

(1) Cf. N. WIENER : The extrapolation, interpolation and smoothing of stationary time series... (1949).

la corrige. C'est le phénomène type des organismes vivants avec ses mécanismes de défense et, réciproquement ASHBY (1) se propose de construire un « cerveau » par un système de feed-back réalisant un homéostase complexe.

Sans prétendre discuter ces questions que nous ne pouvons qu'effleurer brièvement, notons cependant le lien que présente cette théorie avec la statistique moderne. Les performances de nombre de ces appareils doivent être évaluées du point de vue du calcul des probabilités. Dans la « R.M.S. » de WIENER qui sert à définir un critère d'optimalité on retrouve la « variance » de la statistique courante.

Fondamental aussi est le rôle des oscillations par lesquelles se réalise le but à atteindre. Les méthodes purement analytiques (diagramme de NYQUIST) permettent de distinguer a priori les domaines de variation dans lesquels l'amortissement est possible de ceux dans lesquels la machine va à la catastrophe après une série d'oscillations d'amplitude croissante.

De tous ces travaux résulte, comme pour la théorie de l'information, une cascade de suggestions: le terme de « feed back » peut aussi bien désigner l'ensemble des circuits nerveux commandant un effecteur, même si nous n'avons encore que l'espoir de prouver une réelle identité de structure entre feed back électrique et feed back vivant. En tous cas il ouvre des possibilités immédiates de recherches comme le montre les travaux expérimentaux de WIENER et ROSENBLUETH sur les fibrillations cardiaques et mouvements cloniques (2). Pour cela on tentera d'expliquer les particularités qualitatives d'un fonctionnement au moyen de la combinaison la plus simple de « filtres » et de « compensateurs » que permette d'envisager nos connaissances anatomiques. Mais cet agencement de par les propriétés même de ses éléments, est soumis à des règles strictes qui entraînent des relations entre variables quantitatives. Parfois même encore des comportements nouveaux sont révélés d'abord par la théorie du modèle. On a ici l'exemple déjà ancien des travaux de VAN DER POOL sur l'étude des oscillations de relaxation, qui avaient conduit à découvrir des types pathologiques inédits d'électrocardiogrammes avant même que la clinique les eut isolés.

Enfin, il est impossible de ne pas citer la suggestion de WIENER de rechercher dans l'ataxie tabétique et le tremblement intentionnel ou le tremble-

(1) — W. R. ASHBY : The electronic Brain (1946). Electronic Engineering (XII) Pp. 372.

(2) — N. WIENER et A. ROSENBLUETH : « Conduction of impulse in cardiac muscles (1941). Arc. Inst. Card.

voir aussi:

— D. G. ELLSON : « The application of operational analysis to human motor behaviour (1949). Psy. Rev. (56) Pp. 9-17.

— F. V. TAYLOR, R. Y. WALKER, A. S. HOUSEROLDER : « Some aspects of eye-hand coordination in a simplified tracking situation (1946).

Abstract: Am. Psy. (1) — 282 — 283.

— H. GASTAUT : « Sur l'auto-régulation comparée des machines et du cerveau » (1949) Sem. Hôp Pp. 2710-2716.

Année 1949 1949-9. À propos de la "cybernétique" (mathématiques et psychologie)

M. P. SCHUTZENBERGER

607

ment parkinsonien des troubles spécifiques de certains feed-back. Dans tout ceci c'est évidemment la méthode des modèles, mais celui-ci n'est pas construit abstraitement, avec la seule préoccupation de représenter tel ou tel processus. La base technique de l'axiomatisation lui impose des délimitations extrêmement précises et c'est sur celles-ci qu'il est possible de s'appuyer pour passer de la structure à la fonction ou réciproquement pour remonter des observations aux hypothèses anatomiques. La recherche de la simplicité des postulats a déplacé son point d'application : cette exigence ne porte pas sur leur formulation mathématique mais sur la construction physique des êtres dont ils symbolisent le fonctionnement. C'est un moyen puissant de lever la contradiction entre la généralité voulue et la finesse indispensable de toute démarche axiomatique.

\*  
\*\*

Les nécessités de l'exposition nous ont amené à opposer entre elles plusieurs modalités de l'approche mathématique à la psychologie. Ici, comme ailleurs, nous avons pu distinguer la description, la représentation, l'explication. Mais à vouloir séparer on risque de perdre de vue l'unité du mouvement commun qui entraîne toutes ces recherches. Il n'est pas de description, pas de règle expérimentale qui ne présuppose un modèle logique si rudimentaire soit-il ; chaque fait, chaque observation même implique pour exister une théorie. Et, par principe, celle-ci est toujours explicative à un niveau ou à un autre. Ainsi s'il nous est apparu que l'analyse factorielle était la forme la plus immédiate d'un processus de géométrisation des phénomènes psychiques, nous avons pu, par paliers, accéder à la théorie parfaitement abstraite de VON NĒUMANN ; cette opposition seule fixant la définition. En psychologie comme en physique une théorie est axiomatique par rapport à celles qui la précèdent et descriptive en regard de celles qui la suivent. L'ensemble des recherches sociométriques modernes est éloquent à cet égard qui va des sociogrammes de MORENO — simples cartes d'une configuration sociale finie — aux travaux de RAPPOPORT où les acteurs du drame ne sont plus que des matrices stochastiques. Donnant une signification physiologique à une nouvelle constante qu'on tenait encore pour arbitraire, intégrant dans une formule plus compréhensive des relations déjà établies dans des domaines restreints, résolvant les anciennes questions et ouvrant de nouveaux problèmes, chaque système relègue dans l'histoire des sciences son prédécesseur quand il arrive à permettre une prédiction plus riche et une action plus efficace. Si la cybernétique n'en est encore là pour aucun chapitre de la psychologie, la faute en est peut-être à notre impatience et nous ne pouvons qu'espérer ses prochains développements qui nous amèneront à approfondir le problème central de la conscience comme « mode d'existence de la matière la plus hautement organisée ».

ALGÈBRE. — *Sur l'extension des théorèmes de dualité aux treillis distributifs non complémentés.* Note (\*) de M. MARCEL-PAUL SCHÜTZENBERGER, présentée par M. Gaston Julia.

Deux S-applications adverses  $x^t = \sigma(x) =$  réunion des  $y$  tel que  $x|y$  soit complémenté et  $x_1 = \tau(x) =$  intersection des  $y$  tels que  $y|x$  soit complémenté ont été définies dans une Note antérieure (1) pour les treillis modulaires quelconques.

On se restreindra ici aux *treillis distributifs* tels que pour tout  $x < y$ , il existe au moins un  $z$  et un  $t$  tels que  $x < z \leq y$  et  $x \leq t < y$  et que  $x|z$  et  $t|y$  soient complémentés. Cette condition, vérifiée notamment pour les treillis finis, équivaut à

$$(x + y)x^t = x, \text{ équivaut à } y \leq x \quad \text{et} \quad xy + x_1 = x, \text{ équivaut à } x \leq y.$$

Pour tout  $x$  on appellera *complémentaire relatif inférieur* et l'on désignera par  $\underline{x}$ , tout élément  $y$  tels que  $xy = x_1$  et  $x + y = y^t$  et *complémentaire relatif supérieur*  $\bar{x}$ , tout  $y$  tel que  $x + y = x^t$  et  $xy = y_1$ .

Si le treillis est lui-même complémenté,  $x^t = 1$  et  $x_1 = 0$  pour tout  $x$ , et  $\underline{x}$  et  $\bar{x}$  se confondent avec le complémentaire absolu de  $x$  au sens habituel.

THÉOREME 1. — *Si  $y$  et  $z$  sont des  $\underline{x}$ , alors,  $y = z$  car*

$$(y + z)y^t = y + zy^t = y + z(x + y) = y + x_1 + zy = y,$$

d'où, d'après (1),  $z \leq y$  et, de même,  $y \leq z$ , donc  $y = z$ .

THÉOREME 2. — *Pour tout  $x$  et tout  $z$ ,  $xz \leq \underline{x}$  est équivalent à  $z \leq \underline{x}$  (même raisonnement).*

THÉOREME 3. — *La réunion  $y$  des  $z$  tels que  $zx \leq x_1$  est un  $\underline{x}$ .*

Évidemment,  $yx = x_1$ ; d'autre part, si  $y + x \neq y^t$ , il existe un complément absolu  $u$  de  $y + x$  dans le treillis complémenté  $y|y^t$ ; c'est-à-dire que  $y^t = y + x + u$ , avec  $(y + x)u = y$ ; d'où  $yu + xu = y$ ; donc  $xu \leq y$ , donc  $xu \leq yx = x_1$ , donc, par hypothèse,  $u \leq y$ , soit  $x + y = y^t$ .

(\*) Séance du 15 novembre 1948.

(1) *Comptes rendus*, 227, 1948, p. 1008.

Année 1949 1949-10. Sur l'extension des théorèmes de dualité aux treillis...

( 2 )

Les applications  $x \rightarrow \underline{x}$  et  $x \rightarrow \overline{x}$  sont donc des *applications partout définies* et l'on a

$$\overline{\underline{x}} = x, \quad \underline{\overline{x}} = x, \quad \text{mais en général } \overline{\overline{x}} \neq \underline{\underline{x}} \neq x.$$

**THÉORÈME 4.** — Si  $u + v = x$ , alors  $\underline{u}\underline{v} \leq x$ . En effet  $u_1 + v_1 = x$ , d'où

$$\underline{u}\underline{v} + \underline{v}\underline{u} = \underline{x}\underline{u}\underline{v} \quad \text{et} \quad \underline{u}\underline{v} + \underline{v}\underline{u} = \underline{x}\underline{x}\underline{u}\underline{v}$$

d'où  $\underline{x}\underline{u}\underline{v} = \underline{x}\underline{x}\underline{u}\underline{v}$ , d'où (d'après le théorème 2)  $\underline{u}\underline{v} \leq x$ ; de même, si  $uv = x$ , alors  $\overline{x} \leq \overline{u} + \overline{v}$ .

**THÉORÈME 5.** — Si  $x$  est un élément + irréductible  $\overline{g}_i$ ,  $\underline{x}$  est un élément  $\times$  irréductible  $\underline{g}_i$  et réciproquement. Car  $x_1 | x$  est isomorphe à  $\underline{x} | \underline{x}'$ .

**THÉORÈME 6.** — Si  $x$  est un élément + irréductible  $\overline{g}_i$ ,  $x$  et  $\underline{x}$  clivent strictement le treillis (c'est-à-dire que pour tout  $y$ , ou bien  $x \leq y$ , ou bien  $y \leq x$  (immédiat d'après le théorème 2)).

**THÉORÈME 7.** — Si  $x = \Sigma \overline{g}_i$  est une représentation minimale de  $x$  comme réunion d'éléments + irréductibles, alors  $\underline{x} = \Pi \underline{g}_i$  et réciproquement (extension des formules de Morgan) (immédiat d'après le théorème précédent).

**THÉORÈME 8.** — Si  $x < y$  sont des éléments + irréductibles, alors  $\underline{x} < \underline{y}$ .

Car d'après le théorème 6,  $y \leq \underline{x}$ , ou  $\underline{x} \leq y$ , mais,  $y \leq \underline{x}$  entraînerait  $x \leq \underline{x}$ , ce qui est impossible pour  $x \neq 0$ .

De ce résultat, on déduit sans peine :

**THÉORÈME 9.** — Tout treillis distributif fini  $\mathcal{L}$  est défini biunivoquement par la donnée de l'ensemble partiellement ordonné  $\mathcal{J}^+(\mathcal{L})$  obtenu en restreignant la structure aux éléments + irréductibles et  $\mathcal{J}^+(\mathcal{L})$  est isomorphe à  $\mathcal{J}^\times(\mathcal{L})$  défini de la même manière sur les éléments  $\times$  irréductibles.

**THÉORÈME 10.** — Pour tout  $n$  entier, le nombre  $N_n^+$  des éléments de  $\mathcal{L}$  précédés immédiatement par  $n$  éléments est égal au nombre  $N_n^\times$  des éléments de  $\mathcal{L}$  suivis immédiatement de  $n$  éléments (car  $x_1 | x$  est isomorphe à  $\underline{x} | \underline{x}'$ ).

(Extrait des *Comptes rendus des séances de l'Académie des Sciences*,  
t. 228, p. 33-35, séance du 3 janvier 1949.)

REPRINTED FROM *ANNALS OF EUGENICS*, VOL. 14, PART 4, pp. 286-287, 1949  
 [All rights reserved] PRINTED IN GREAT BRITAIN

A NON-EXISTENCE THEOREM FOR AN INFINITE FAMILY OF  
 SYMMETRICAL BLOCK DESIGNS

BY M. P. SCHÜTZENBERGER\*

*Centre de Génétique de l'Hôpital Saint Louis (Paris)*

An incomplete block design consists of a set of  $b$  subsets or 'blocks', each block containing  $k$  different 'varieties' chosen out of  $v$  varieties in all, in such a way that each of these varieties occurs the same number  $r$  of times or 'replications' in the whole design, and each pair of varieties occurs together in a fixed number  $\lambda$  of blocks. The design is called 'symmetrical', if  $v = b$  and  $r = k$ : then  $v$ ,  $r$  and  $\lambda$  are related by

$$v = 1 + \frac{r(r-1)}{\lambda}. \tag{1}$$

We show in this paper that, if  $v$  is even, then such a symmetrical design can exist only if  $r - \lambda$  is a perfect square.

The design may be schematized by a square  $v \times v$  matrix  $A = \| a_{ij} \|$  in which columns correspond to blocks and rows to varieties, and  $a_{ij} = \alpha$  if the  $i$ th variety is a member of the  $j$ th block and  $a_{ij} = \beta$  if not. Without loss of generality we may suppose that  $\alpha = 1$  and  $\beta = 0$ .

Further,  $A$  may be interpreted in a Euclidean space  $E_v$  as a set of  $v$  row vectors  $V_i$ , the coordinates of which are the  $a_{ij}$ 's; and the  $V_i$ 's satisfy the following conditions:

I. The  $V_i$ 's have the same length  $l = \sqrt{r}$ , for

$$l^2 = r\alpha^2 + (v-r)\beta^2 = r. \tag{2}$$

II. The angle  $\phi_{ii'}$  between any two  $V_i$ 's, has the same value  $\cos^{-1} \lambda/r$ , for

$$V_i V_{i'} = l^2 \cos \phi_{ii'} = \lambda\alpha^2 + 2(r-\lambda)\alpha\beta + (v-2r+\lambda)\beta^2 = \lambda. \tag{3}$$

From these relations,  $\det |A|$ , the measure of the multiple vector  $A$  may be computed by reducing  $A$  to a triangular matrix  $B$  by rotations in  $E_v$ . Thus  $B$  is defined by:

$$\text{If } i \leq j + 1, \text{ then } b_{ij} = 0, \tag{4}$$

$$\text{for all } i: \sum_{j=1}^n (b_{ij})^2 = l^2 = r, \tag{5}$$

$$\text{for all } i \text{ and } i': \sum_{j=1}^n b_{ij} b_{i'j} = l^2 \cos \phi = \lambda. \tag{6}$$

From (4) and (6) it follows immediately that:

$$\text{For all } i \leq i' \text{ and } j \leq i - 1: b_{ij} = b_{i'j} = c_j, \tag{7}$$

$$\text{and from (5) and (6)} \quad b_{ii}^2 = d_i^2 = r - \sum_{j=1}^{i-1} c_j^2, \tag{8}$$

$$d_i c_i = \lambda - \sum_{j=1}^{i-1} c_j^2 = \lambda - r + d_i^2. \tag{9}$$

$$\text{Hence} \quad d_{i+1}^2 = d_i^2 - c_i^2 = d_i^2 - (d_i^2 + \lambda - r)^2 d_i^{-2}. \tag{10}$$

\* We wish to express our grateful thanks to Prof. Penrose and to Dr Cedric A. B. Smith for their advice in the preparation of this paper.

287

M. P. SCHÜTZENBERGER

Hence, writing  $P_\mu = \prod_{i=1}^\mu (d_i)^2$ , we get the linear recurrence equation

$$\left. \begin{aligned} \frac{P_{\mu+1}}{P_\mu} &= \frac{P_\mu}{P_{\mu-1}} - \frac{(P_\mu - (r-\lambda)P_{\mu-1})^2}{P_\mu P_{\mu-1}}, \\ P_{\mu+1} &= 2(r-\lambda)P_\mu - (r-\lambda)^2 P_{\mu-1}. \end{aligned} \right\} \quad (11)$$

For the initial conditions  $p_1 = r, P_2 = r^2 - \lambda^2$ , the solution of (11) is

$$P_\mu = (r-\lambda)^{\mu-1}(r+(\mu-1)\lambda), \quad (12)$$

and, finally, we get, by putting  $\mu = v = 1 + \frac{r(r-1)}{\lambda}$ ,

$$\det |A| = \det |B| = P_v^{\frac{1}{2}} = r(r-\lambda)^{\frac{1}{2}(r-1)}. \quad (13)$$

Hence we have the following theorem: if  $v$  is even, a symmetrical incomplete block design may exist only if  $r-\lambda$  is the square of an integer, for  $\det |A|$  may be computed directly by rational operations.

It is obvious that the theorem applies only if  $\lambda$  admits 2 as a divisor and that, for any such  $\lambda$  it works for an infinite number of values of  $r$ . Obviously too, the theorem may be used indifferently for a design or its complement, for

$$r-\lambda = (V-r) - (v-2r+\lambda). \quad (14)$$

We have listed below all the cases of application of the theorem for  $v \leq 100$  or  $r \leq 20$  ( $r \leq v/2$ , and  $r \neq v$  or  $v-1$ ):

	$v$	$r$	$\lambda$		$v$	$r$	$\lambda$		$v$	$r$	$\lambda$
(1)	16	6	2		*34	12	4		70	24	8
(2)	*22	7	2	(3)	40	13	4		*76	25	8
	*46	10	2		96	20	4		66	26	10
	56	11	2		36	15	6		*88	30	10
	*92	14	2		*52	18	6		*94	31	10
	*106	15	2		*58	19	6		64	28	12
	*172	19	2		78	22	6		*86	35	14
									100	45	20

\* Impossible by application of the theorem.

(1) Kummer's configuration.

(2) Already discarded by Q. M. Hussain. (Cf. R. A. Fisher & F. Yates, *Statistical Tables...*, 3rd ed., 1948, pp. 17 and 65.)

(3) Finite geometry modulo 3.



## Editorial

*La publication d'articles doctrinaux à position tranchée provoque naturellement des réactions. Nous vous offrons celle pleine de verve de DUCHÊNE, qui, avec le fils de SCHÜTZENBERGER, pourfend le test de Szondi prôné en février par REQUET.*

*Plusieurs collègues ont insisté pour que l'Information contienne des comptes rendus d'expériences: BELEY, FOLLIN et moi-même vous apportons un exemple de ce que nous avons fait dans des établissements fort différents en réunissant les infirmiers de certains pavillons. Dans un prochain numéro nous publierons un article de P. ROYER sur le laboratoire de l'Hôpital Psychiatrique. Je pense que ces publications établiront avec les spéculations théoriques qui nous ont occupé un équilibre harmonieux.*

*Si nos finances nous le permettent, si vous nous en fournissez la matière nous aimerions que tous nos numéros aient la même physionomie.*

*La parution de mon rapport sur la vie du Syndicat me libère pour ce numéro de l'obligation d'une mise au point sur les problèmes de l'heure.*

G. DAUMÉZON.

## Quelques réserves sur la méthodologie du Docteur Szondi

par H. DUCHÊNE et M.-P. SCHÜTZENBERGER

La vogue inquiétante des théories et du test de SZONDI dans les milieux de psychologues et de psychiatres, attestée par l'article du Docteur REQUET sur le destin schizophrène que *l'Information Psychiatrique* a publié dans son numéro de février 1949, nous paraît nécessiter l'expression de critiques qui se présentent en foule dès que l'on aborde ces travaux et ne cessent de croître lorsqu'on examine les applications qui en découlent.

Nous avons pris connaissance des théories de SZONDI par la lecture du volume consacré par cet auteur à la méthodologie de son test (1), et des articles très documentés de Mme R. BEJARANO (2) et du Docteur H. ELLENBERGER (3).

(1) « SZONDI-TEST », Experimentelle Triebdiagnostik — Text Band Verlag Hans Hüber — Bern — 1947.

(2) R. BEJARANO : « Le test du destin de Szondi » *Psyché*. — N° 23-24, septembre-octobre 1948.

(3) H. ELLENBERGER : « A propos de l'analyse du destin de Szondi ». *Evolution Psychiatrique* 1948, fascicule IV.

## — 211 —

Nous ne reprendrons pas ici un exposé systématique de cette œuvre extraordinairement touffue, désirant nous borner aux critiques essentielles. Nous renvoyons donc ceux qui voudraient l'étudier aux textes cités ci-dessus.



L'œuvre de SZONDI comporte deux parties distinctes: d'une part un ensemble de conceptions théoriques, d'autre part un test (choix parmi six séries de huit photographies des deux plus sympathiques et des deux plus antipathiques).

Il importe de souligner d'emblée que les conceptions théoriques n'ont pas été déduites de l'application du test mais que celui-ci a été conçu en fonction de ces conceptions théoriques qu'il serait destiné à confirmer.

Ceci légitime une critique de la théorie et des bases du test indépendante des résultats de celui-ci. Sans doute pourra-t-on ensuite poser la question de la signification des choix au cours du test. Celle-ci devrait être discutée qu'à partir d'une large expérimentation.

Notons seulement au point de vue de la valeur du test en tant qu'instrument de mesure, deux problèmes qui ne semblent avoir nullement inquiété les différents auteurs qui ont étudié cette épreuve:

— La possibilité d'une interaction entre le sujet et l'opérateur bien difficile à retrouver après coup mais fort plausible dans de nombreux cas (nécessité de hâter le travail du sujet — de le limiter à deux choix et à deux rejets par planche, etc...);

— Le caractère presque exorbitant de l'hypothèse de base qui postule l'équivalence des choix dans les séries successives: du point de vue de l'exploitation il importe nullement qu'une figure « hy » (par exemple) soit choisie à la nième plutôt qu'à la mième série.

On reste perplexe devant la naïveté de l'auteur qui, sans le plus faible contrôle expérimental a implicitement postulé cette équivalence.

Nous nous contenterons donc de marquer ici le caractère éminemment contestable des hypothèses qui ont servi à échafauder cet édifice impressionnant



Le point de départ des recherches de SZONDI a été la constatation d'une coïncidence entre l'état obsessionnel d'une jeune femme et celui de sa belle-mère. N'y aurait-il point des tendances profondément inscrites dans l'être humain qui guideraient le choix du conjoint, des amis, de la profession, etc... Pour vérifier cette hypothèse, SZONDI rassemble de 1937 à 1944, 4.000 observations de sujets relevant les « coïncidences », non seulement dans leur ascendance, dans le choix de leurs conjoints, de leurs amis, mais également dans l'ascendance desdits conjoints ou amis.

Mais quelles coïncidences retiendra-t-il?..

C'est ici qu'intervient une construction qui représente un véritable pot-pourri des acquisitions psychiatriques, psychanalytiques, biotypologiques, etc... Les travaux de FREUD, de JUNG, de BLEULER, de KRETSCHMER, de RORSCHACH, de Mme MINKOWSKA, etc... sont harmonisés dans une synthèse qui pose en principe que « correspondant aux quatre groupes de maladies de tendances héréditaires, nous ne devons admettre ni plus, ni moins que quatre tendances fondamentales ». Ces quatre tendances fondamentales seraient:

- la tendance sexuelle (?) ou Vecteur S;
- la tendance paroxysmale ou de surprise ou éthique (?) ou Vecteur P;

## — 212 —

- la tendance du moi ou schizophrène ou Vecteur Sch;
- la tendance de contact ou circulaire ou Vecteur C.

Ces vecteurs se subdivisent en 8 (facteurs pulsionnels) correspondant « à une famille de maladies héréditaires ».

- |                    |                                      |
|--------------------|--------------------------------------|
| I. — Vecteur S     | 1 — facteur h (homosexualité)        |
|                    | 2 — facteur s (sadisme)              |
| II. — Vecteur P    | 3 — facteur e (épilepsie)            |
|                    | 4 — facteur hy (hystérie)            |
| III. — Vecteur Sch | 5 — facteur r (catatonie)            |
|                    | 6 — facteur p (psychoses paranoïdes) |
| IV. — Vecteur C    | 7 — facteur d (états dépressifs)     |
|                    | 8 — facteur n (états maniaques).     |

On nous fera grâce des 16 tendances pulsionnelles résultant de la subdivision des huit facteurs pulsionnels, bien que nous regrettions de ne pouvoir citer certains exemples des sous-groupes tel que le sous-groupe 13 Cd — dont SZONDI écrit: « Ce sont des hommes du type: « acrobate-Ohhh ». — Une note explique que ce titre se réfère à une attraction dans laquelle un clown s'exclamait « acrobate Ohhh » devant les performances de certains acrobates avant de les imiter lui-même (1).

Il est, somme toute, satisfaisant pour l'esprit de voir s'ordonner ainsi les conceptions psychiatriques dans un schéma qui représente assez bien l'idée que l'on se faisait des « constitutions ». Certains auraient sans doute discuté l'hérédité de telle ou telle catégorie; auraient préféré voir le masochisme s'opposer au sadisme ou auraient émis d'autres critiques de détail. Ne nous arrêtons pas aux objections de ces esprits chagrins: voyons plutôt l'apport original de SZONDI.

Il consiste en une combinaison du schéma précédent avec celui de la génétique. La « tendance pulsionnelle » est déterminée par un seul gène. Le « facteur pulsionnel » résulte de l'association de deux gènes, c'est-à-dire l'accouplement de deux tendances. Le « Vecteur pulsionnel » provient de deux « facteurs pulsionnels », soit deux paires de gènes, formées de quatre tendances pulsionnelles isolées dans chacun des quatre gènes.

De ces gènes, certains sont dominants et vont donc se manifester dans le phénotype. D'autres sont récessifs. Ils peuvent néanmoins se manifester également, affirme SZONDI, soit dans le phénotype, soit, et c'est le point capital de sa théorie, dans les tropismes qui dirigeront la vie pulsionnelle.

C'est de ces gènes récessifs latents que dépendront donc le libidotropisme (choix de l'être aimé) l'idéalotropisme (choix des amis, des idéaux), l'opérotropisme (choix d'une profession), le morbotropisme (tendances morbides) et le thanatotropisme (choix du genre de mort - sic).

Les exemples d'opérotropisme cités par Mme BEJARANO sont trop savoureux pour que nous résistions au plaisir de les citer: « Le criminaliste peut ainsi satisfaire ses gènes « criminels » récessifs dans le contact avec des criminels manifestes; le pompier satisfera dans son activité professionnelle ses gènes pyromaniaques récessifs — sans que pour cela tous les criminalistes ou pompiers entrent dans ce cas »...

(1) Szondi-Test, p. 77. Il faudrait également pouvoir citer tout le passage relatif à l'aptitude au bonheur des diverses catégories (p. 198).

## — 213 —

Nous ne nous étonnerons plus guère d'apprendre que le métier de psychiatre correspond à la manifestation de tendances pulsionnelles paranoïdes...

On s'inquiétera peut-être de savoir comment un sujet choisit un conjoint ou un ami pour des tendances pulsionnelles présentes chez ce conjoint ou cet ami, à l'état de gènes récessifs latents et donc ne se manifestant pas dans la personnalité?... L'explication de SZONDI est fort simple: il doit s'agir d'une radiation émise par les gènes et il souhaite fort à propos que ces radiations trouvent un jour leur Mme Curie.

Une légère difficulté toutefois: comment le sujet soumis au test des photos choisit-il justement celles de personnes vers lesquelles le guide son libidotropisme ou son idéalotropisme?... Il est évident (pour SZONDI) qu'il ne s'agit pas là d'un mécanisme génobiologique ou génotropique. Il suffit de recourir à la théorie des réflexes conditionnés et à celle des engrammes pour comprendre que la vue d'une photographie suffit à réveiller le tropisme associé à cet engramme (*Szondi-Test*, page 32).



On reste confondu devant cette théorie qui évoque irrésistiblement la formule bien connue: tout ce qui est valable en elle, n'est pas nouveau, mais tout ce qui est nouveau est des plus contestables.

La référence à la génétique représente un des meilleurs exemples d'utilisation purement verbale de celle-ci. Pourquoi un gène par tendance pulsionnelle et non par facteur ou par vecteur?... Pourquoi la vie pulsionnelle serait-elle le domaine de gènes récessifs hétérozygotes et donc latents?... Pourquoi cette explication plutôt qu'une variation de pénétrance et d'expressivité, ou une épistasie?...

Nous tenons à la disposition de ceux de nos confrères qui ne seraient pas entièrement satisfaits par l'interprétation de SZONDI, deux ou trois explications génétiques différentes tirées des notions ci-dessus et susceptibles d'être appliquées à toute combinaison de tendances qu'ils estiment préférable.

Pour ceux qui seraient à court de système pulsionnel, nous pouvons également fournir une théorie très étudiée représentant une simplification et une modernisation fort appréciables de celle SZONDI:

— à la base, une seule pulsion V ou élan vital (BERGSON, textes hindous ,etc...);

— deux aspects de cette pulsion: attraction et rétraction (propriétés fondamentales du protoplasma);

— un système convenablement calculé de gènes conditionnels permettant de faire prédominer l'une ou l'autre tendance, et des effets épistatiques qui donnent toute facilité de multiplication des sous-groupes;

— un principe de manifestation de ces tendances sur divers plans hiérarchisés (somatique, métabolique, psychologique) suivant une conception jacksonnienne (important aspect de la pensée contemporaine négligé dans l'œuvre de SZONDI);

— un ingénieux système de «feed-back» emprunté à la cybernétique permet le contrôle réciproque de ces manifestations et l'explication de la stabilité des comportements malgré les variations du milieu;

— enfin, les choix du conjoint, des amis, etc... dépendent des caractéristiques ondulatoires du regard (dont l'importance est bien mise en évidence dans le chapitre que SARTRE lui consacre dans « l'être et le néant »). Un sujet ne choisit que les individus dont le regard est en phase avec le sien. S'il y a décalage on trouve le phénomène bien connu des « batte-

— 214 —

ments », d'où sans peine une théorie mathématique et pataphysique des rixes, duels, guerres et révolutions renouvelées des meilleurs auteurs comiques.



Sans doute le lecteur pensera-t-il qu'une telle présentation de la théorie de SZONDI ne lui rend pas justice en négligeant les 4.000 observations sur lesquelles elle se fonde. C'est évidemment l'objection que nous nous sommes adressés en premier. Mais là encore, on est stupéfait de la méthode qui a présidé à la collection de ces observations. Il s'agit en effet de relever dans la famille, les amis, les alliés d'un sujet ainsi que parmi les parents et alliés de ces amis ou conjoints un certain nombre de caractéristiques *sans que jamais aucune analyse statistique vienne donner à ces constatations une autre valeur que celle d'une simple coïncidence.*

Voici un exemple donné par SZONDI lui-même des « vérifications » de son test par une observation. Il s'agit d'un prêtre catholique de 40 ans dont les « formules » établies par le test étaient :

Proportions de latence $Pe+ = Ss = Schp = C$ <hr style="width: 80%; margin: 0 auto;"/> 4 : 3 : 2 : 0	Catégorie pulsionnelle $Hy$ <hr style="width: 80%; margin: 0 auto;"/> h, K, d, m <hr style="width: 80%; margin: 0 auto;"/> e p s
------------------------------------------------------------------------------------------------------------	-------------------------------------------------------------------------------------------------------------------------------------------

Que ces formules correspondent bien au « destin » du sujet est démontré, pour SZONDI, par les particularités suivantes :

A. — Manifestation du destin dans le domaine épileptiques Pe.

a) Maladies (morbotropisme)

1. *Migraineux* sont le père, la grand'mère maternelle et le grand-père du sujet.
2. *Asthmatique* est la grand'mère maternelle.  
*Rhume des foins* atteint le probant, et une tante maternelle
3. *Enurétique* est le frère du probant.
4. *Gaucher* est le probant.

b) Occupations (operotropisme)

*Homo Sacer*: le probant (il est prêtre).

c) Amitiés (idéaltropisme)

1. *Epileptique* est le père d'un ami qui fut même interné pour une psychose épileptique.
2. *Absences* présentées par un deuxième ami du probant.
3. *Bègue* est un troisième ami du probant.
4. *Asthmatique* est un quatrième ami du probant.

B. — Manifestations du destin dans le domaine Schizoïde Schp :

1. *Démence paranoïde* fut le diagnostic porté sur son arrière grand-mère.
2. Sous une forme socialisée cette tendance p s'est manifestée chez le probant par son attirance pour la philosophie et son désir de prêcher. Il a écrit des œuvres philosophiques et fut un champion du néo-catholicisme.

## C. — Manifestations du destin dans le domaine sadique Ss

1. *Boucher* était un arrière grand-père paternel.
2. *Chirurgien* est un oncle maternel.

Signalons cependant quelques autres preuves omises (par modestie sans doute) par le Docteur SZONDI :

1° un parent, ou allié ou ami du proposant, ou un parent ou allié ou ami de l'un de ceux-ci a lu un livre de HÖLDERLIN (domaine schizoïde) ;

2° un parent ou allié, etc... est né le jour de Sainte Barbe (patronne des mineurs: domaine épileptoïde) ;

3° un parent ou allié, etc... a été examiné du point de vue mental par le Docteur SZONDI (domaine paranoïde).

Si d'ailleurs, on éprouvait quelque difficulté, il suffit de savoir que SZONDI a constaté lui-même que certains individus peuvent passer d'une classe à l'autre (*Szondi-Test*, page 211) pour être rassuré à ce sujet.

En définitive, on ne comprend pas très bien pourquoi SZONDI se laisse aller au découragement en affirmant que les observations ne sont pas toujours si simples et qu'il y a encore beaucoup de cas où il faut véritablement se casser la tête (*Wirklich den Kopf Zerbrechen*) pour arriver, à l'aide des trois méthodes, la méthode de l'analyse libre, de l'analyse du moi et de l'analyse de l'évolution de la pulsion, à trouver enfin le bon diagnostic » (*Szondi-Test*, page 249).

Il est inutile, croyons-nous, de poursuivre une critique de ces observations.

Un dernier fait donnera pourtant la mesure du crédit que l'on peut leur accorder.

SZONDI apporte comme une preuve de sa conception génétique l'examen des jumeaux monozygotes qui ont présenté un test « à peu près semblable ». Sans chicaner sur l'à peu près, contentons-nous de relever qu'il s'agit d'une *paire* de jumeau. Une seule!!!

\*\*

Ce n'est pas seulement le test de SZONDI, malheureusement mais aussi sa méthode de travail qui tend à se répandre. Un article récent du Docteur REQUET sur le « destin des déments » (1) en est l'exemple le plus frappant. Il vise à prouver que les déments (D, séniles, paralysie générale, etc...) ont une vie caractérisée par l'échec. Il pense le démontrer en accumulant des cas dans lesquels le malade était célibataire, veuf, divorcé, remarié, ou sans enfants: toutes ces situations étant considérées comme des échecs de la vie conjugale.

Dans d'autres cas, ce sont des échecs professionnels, des accidents: « enfin, si tout semblait avoir été en ordre dans l'épanouissement de la vie, nous avons relevé avec une insistance particulière l'hérédité psychopathique, souvent démentielle, dans la lignée du sujet lui-même ou de son conjoint. »

A aucun moment, la fréquence de ces « échecs » déjà si contestables dans leur définition n'est comparée à ce qu'elle pourrait être dans une autre population. Pourtant le Docteur Requet écrit: « On pourrait nous reprocher de n'avoir pas introduit d'approximations statistiques à l'appui de nos exemples. Mais telle n'était pas notre intention. Nous avons voulu, pour le moment, qualifier seulement le destin démentiel, quitte à laisser

(1) A. REQUET: « Le destin des déments ». — *Psyché*, n° 23-24, septembre-octobre 1948, page 1042.

— 216 —

pour plus tard à d'autres organismes le soin d'en évaluer l'importance quantitative et la probabilité statistique.

« Qu'il nous suffise de dire pour l'instant que toutes les démences ont un destin démentiel, là il n'y a pas d'erreurs possibles, ce n'est pas une probabilité, c'est une certitude. Mais cela ne veut pas dire, naturellement, que tous les destins démentiels ont une démence; il y en aurait vraiment trop. »

✱

La contradiction impliquée dans les dernières lignes nous paraît trop flagrante pour que nous ajoutions un seul mot de critique.

Ce que nous voudrions par contre souligner, c'est qu'une telle conception du « destin des déments » est la conséquence nécessaire de l'œuvre de SZONDI, en raison des erreurs méthodologiques sur lesquelles celle-ci est basée.

Ces erreurs nous semblent, à l'un et à l'autre, trop néfastes pour ne pas être dénoncées.

En outre, l'un de nous (M.P. SCHÜTZENBERGER) pense que la théorie et la méthode de SZONDI sont représentatives d'un type de théories et méthodes psychologiques dont les exemples se multiplient. Le succès, la prolifération de ces théories simplistes combinant hors de tout contact avec la réalité, le mécanisme le plus grossier avec l'idéalisme le plus stérile, traduit une crise grave d'une certaine idéologie: sentant la nécessité de constituer une science de l'homme, elle s'y épuise par refus d'en admettre les seuls fondements valables mais dont la reconnaissance serait ipso facto sa ruine.



1949-13. Investigaciones críticas sobre la teoría y el test de L. Szondi Année 1949

SRTA. A. E. ANCELIN y DRES. H. DUCHENE  
y M. SCHÜTZENBERGER

*Investigaciones críticas sobre  
la teoría y el test de L. Szondi*



PUBLICADO EN LA  
**REVISTA DE PSICOLOGIA  
GENERAL Y APLICADA**  
M A D R I D  
1 9 4 9

Année 1949      1949-13. Investigaciones críticas sobre la teoría y el test de L. Szondi

*REVISTA DE PSICOLOGIA  
GENERAL Y APLICADA*

•

Vol. IV. - Número 11.

*JULIO - SEPTIEMBRE*

1      9      4      9

*DIRECTOR Dr. JOSE GERMAIN  
PUBLICACION TRIMESTRAL  
— 800 PAGINAS AL AÑO —  
SUSCRIPCION PARA ESPAÑA  
Y AMERICA 80 PESETAS - PARA  
EL EXTRANJERO 100 PESETAS  
NUMERO SUELTO 25 PESETAS*

## INVESTIGACIONES CRITICAS SOBRE LA TEORIA Y EL TEST DE L. SZONDI

POR LA

SRTA. A. E. ANCELIN Y DRES. H. DUCHENE Y M. SCHÜTZENBERGER  
PARÍS

**E**NTRE las teorías que pretenden reducir a un esquema único (más que unificador) la totalidad compleja de la personalidad normal o patológica hay pocas que presenten tan francamente esta característica como la teoría de L. Szondi. A los psicólogos puros ofrece la garantía de apoyarse en la genética mendeliana, y a los biólogos la que aparentemente le confiere el empleo de palabras tomadas de diversas concepciones psiquiátricas, contradictorias por supuesto. Finalmente, la gran multitud de utilizadores de tests y de semimédicos es seducida por la elegante simplicidad de un test cuyo examen revela el pasado, el presente y el futuro, tan bien si no mejor que la bola de cristal de sus predecesores.

### LA TEORÍA DE L. SZONDI.

Cuatro pares de genes, cuatro “loci”, cada uno provisto de dos alelos, dan cuenta de toda la herencia de la personalidad.

Cada “locus” dirige una “pulsión” (vector) que se descompone en dos necesidades “pulsionales” (factores) correspondientes cada una a una enfermedad mental.

Vector S (Sexualidad)	factor h factor s	(homosexualidad) (sadismo)
Vector P (Ética)	factor e factor h	(epilepsia) (histeria)
Vector Sch (pulsión del yo)	factor k factor p	(catatonia) (paranoia)
Vector C (contacto)	factor d factor m	(depresión) (mania)

Y de estos cuatro pares de genes van a resultar :

*El libidotropismo:* elección (?) de consorte y de amigos, gracias a “radiaciones” emitidas por los genes que atraen a los portadores del mismo sistema genético.

*El operatropismo:* elección (?) de oficio.

*El morbotropismo:* elección (?) de enfermedad.

*El tanatotropismo:* elección (?) del género de muerte.

En resumen, todo el destino (“Schicksal”) del individuo se encuentra inexorablemente reglamentado por este esquema de desconcertante simplicidad.

Naturalmente, la “teoría” de L. Szondi no se ha usado en absoluto para explicar los árboles genealógicos ya publicados en la inmensa literatura consagrada a la herencia de enfermedades mentales o caracterológicas.

Los únicos ejemplos aducidos como pruebas por L. Szondi han sido recogidos por él mismo, y la ausencia total del más pequeño tratamiento estadístico hace aún más incierto el valor de sus datos. Por otra parte, la “prueba” del carácter epileptoide, por ejemplo, del sujeto y la validez de la interpretación genética parece consistir a veces tan sólo en el hecho de que el “padre de un amigo del sujeto es epiléptico” (prueba por idealotropismo) razón que parece débil aun cuando se añada la

SRTA. A. E. ANCELIN Y DRES. H. DUCHENE Y M. SCHÜTZENBERGER 439

circunstancia de que “un tío materno es bombero” (prueba de operotropismo).

No obstante la admiración que pueda merecer el trabajo realizado en la reunión de un número tan elevado de árboles genealógicos de anormales es, sin embargo, cierto que, por una parte, su número sería insuficiente para hacer válida la teoría incluso si se hubiera empleado un método riguroso (cosa que no se ha hecho) y que, por otra, el desdén completo con que L. Szondi olvida las investigaciones de los demás en este campo no puede sino suscitar graves sospechas con respecto a la parte llamada genética de su teoría.

#### EL TEST DE L. SZONDI.

Es para L. Szondi un juego explicar esta atracción entre portadores de un mismo gene recurriendo a la teoría de los engramas y reflejos condicionados: el sujeto encontrará simpáticas las fotografías de los enfermos que tienen en el estado más puro las necesidades que su “Yo” conoce y acepta, y antipáticas las fotografías correspondientes a necesidades actualmente reprimidas.

En la práctica se presentan al sujeto seis series de ocho fotografías de enfermos mentales bajo la influencia de las ocho enfermedades “pulsionales” enumeradas más arriba. Aquél indica cada vez las dos más simpáticas y las dos más antipáticas. La administración del test se resume por entero en un “score” para cada factor obtenido de la manera siguiente:

O (necesidad descargada, satisfecha) menos de dos fotografías simpáticas o antipáticas en la serie de ocho fotografías dependientes de este factor.

± (actitud ambivalente) al menos dos fotografías simpáticas y dos fotografías antipáticas.  
 — (necesidad reprimida) al menos dos fotografías antipáticas, lo más una simpática.  
 + (necesidad admitida por el “Yo”) al menos dos fotografías simpáticas, lo más una antipática.

La interpretación se hace en dos etapas:

1.ª Después de una sola administración del test, la combinación de los “scores” en cada uno de los dos factores dependientes de un mismo vector permite, gracias al “Psychodiagnostiktafel”, trazar un retrato psicológico del sujeto. Por ejemplo:

e + y h —: significa que el sujeto es “der reine Abel”.  
 e — y h +: significa que el sujeto es “der reine Kain”.  
 e ± y h ±: significa que “Abels Kampf mit Kain”

o bien

h + s: sexualidad normal.  
 h O s O: sexualidad senil (o infantil con agresividad).  
 h O s +: “El verdugo de corazón de tórtola” (*sic*).  
 h O s O: abstinencia.

Los  $4 \times 4 = 16$  casos posibles que corresponden a cada uno de los cuatro vectores, están previstos y descritos. *Una gran parte de la obra de L. Szondi está consagrada, por lo demás, a la discusión de casos individuales sobre la única base de una sola administración del test.*

2.ª Se repite el test diez veces en veinticuatro horas. Los resultados sucesivos se van integrando, teniendo en cuenta tan sólo los “scores” O, +, — ó ± para cada factor. Una serie de artificios bastante complicados permite hacer entonces el análisis completo “del destino”.

Según esto, parece que el significado de este proceso es obtener muestras de la personalidad del sujeto por medio de una investigación repetida; se espera así que las necesidades “pulsionales” más intensas aparezcan libres de las contingencias accidentales que podrían viciar una

SRTA. A. E. ANCELIN Y DRES. H. DUCHENE Y M. SCHÜTZENBERGER 441

sola administración del test. Hay que advertir que esto no impide que ciertos discípulos (por otra parte muy autorizados) administren el test diez veces seguidas a un mismo sujeto en el espacio de una tarde, lo que no deja de ser, a pesar de todo, una selección de muestras muy curiosa.

#### CRÍTICA DEL TEST.

El simple enunciado del método suscita ya algunas cuestiones bastante perturbadoras: del solo hecho de que  $6 \times 2 = 12$  elecciones de cada tipo han de ser efectuadas necesariamente por el sujeto, éste, cualquiera que sea su estado psicológico, se ve obligado a atribuir como mínimo un “score” + y un “score” —; dicho de otra forma, la teoría postula implícitamente que en cada instante existe como mínimo una necesidad pulsional que es admitida y otra que es rechazada por el Yo.

Esta aserción —deducida de la propia estructura del test— no guarda ninguna relación con la teoría de L. Szondi, ni se deduce de ella la imposibilidad —impuesta por el test— de que un sujeto manifieste a la vez seis necesidades “pulsionales” o más (puesto que son necesarias dos elecciones “simpáticas” para obtener un “score” positivo, y no hay en total más que doce elecciones).

A estas críticas *a priori* se añaden las avanzadas por algunos de los que han usado el test: se le reprocha la mala calidad de las fotografías, se pone en duda el diagnóstico inicial de los enfermos mentales escogidos, se deplora el carácter “extranjero” de ciertas fotografías..., el que sean todas demasiado antipáticas...

En nuestra opinión, por muy justas que sean estas críticas, no atañen a lo esencial, que es la teoría psicoló-

gica subyacente al test, única fuente de significación del mismo. En efecto, si no se admite la teoría de L. Szondi de la atracción de los epileptoides hacia los epileptoides, la de los maniáticos hacia los maniáticos ... ninguna experiencia, ninguna observación de la psicología corriente justifica la interpretación que se da a la respuesta del sujeto. Y esto es tanto más si se considera que la técnica de examen de L. Szondi es sumamente arbitraria y carece en absoluto de sustentación experimental.

Por ejemplo, en una muestra de cien varones adultos normales de la región parisina encontramos las proporciones siguientes para dos factores típicos:

	<u>homosexualidad (h)</u>	<u>Catatonía (C)</u>
"score" 0	11 %	10 %
"score" +	77 %	8 %
"score" -	3 %	78 %
"score" ±	9 %	4 %

lo que demuestra el carácter perfectamente *a priori* del "scoring" elemental. Ninguna escala previa justifica esta proporción, cuando menos curiosa, de 77 por 100 de homosexualidad consciente y aceptada, y de 78 por 100 de catatonía rechazada (reprimida). Esto es, pues, una definición más que una descripción: se consideran como de tal o cual tipo los sujetos que responden al test de tal o cual manera en virtud de un decreto perfectamente *a priori*.

Así, pues, nos parece imposible utilizar el test, a menos de admitir los elementos de la teoría de L. Szondi relativos a su prueba: existencia de ocho sistemas de causas combinándose de forma sencilla y determinando la simpatía o antipatía para cada uno de los grupos de fotografías. Pero entonces, esta simple exigencia impone al test un cierto número de condiciones que debería cumplir exactamente si la teoría fuese válida; estas condiciones

SRTA. A. E. ANCELIN Y DRES. H. DUCHENE Y M. SCHÜTZENBERGER 443

son: homogeneidad estadística de las diferentes fotografías, homogeneidad interna de las ocho series, un mínimo de constancia en el tiempo.

I. *Homogeneidad de las fotografías:*

Hemos utilizado 50 estudiantes (hombres), 50 estudiantes (mujeres), 50 adultos normales y consiguientemente, fotografía por fotografía, el porcentaje de las elecciones recogidas; 12 de las 40 fotografías reciben la mitad de las elecciones simpáticas, y entre ellas, seis reciben cerca de  $1/3$  del total; lo mismo ocurre para la elección de las antipáticas. Pero esta irregularidad tiene consecuencias aun más graves; por ejemplo, dadas las frecuencias de elección de las fotografías de enfermos histéricos, la mitad de los sujetos que han recibido un "score" negativo en este vector han rechazado la fotografía de la 3.<sup>a</sup> serie y tendrían un "score" cero si no hubieran designado esta fotografía. Así, puesto que el 60 por 100 de la población total es "h", el 30 por 100 de estos diagnósticos dependen, en definitiva, del solo hecho de que L. Szondi haya escogido de la familia de los histéricos esta fotografía tan generalmente antipática, al contrario de las otras que representan la misma enfermedad. Del mismo modo, el 98 por 100 de los sujetos escogen como simpático el homosexual de la 5.<sup>a</sup> serie: la significación de esta elección es totalmente incompatible con la teoría szondiana. En efecto, la sola interpretación de esta elección tan general sería que esta fotografía permite una detección muy sensible de las más débiles pulsaciones homosexuales conscientes. Así, pues, la minoría que considera esta fotografía como antipática debería rechazar con fuerza las otras fotografías de la serie "h". Un cálculo rápido muestra,

por el contrario, que esta minoría no difiere sensiblemente del conjunto de la población en lo que concierne a su actitud frente a otras fotografías de homosexuales.

## II. *Homogeneidad de las series:*

Pero hay algo más grave. Si realmente en una aplicación dada existía en los sujetos la tendencia a designar como “simpáticas” o “antipáticas” las fotografías de los enfermos representativos de una misma enfermedad pulsacional, esta tendencia debería explicarse más o menos claramente en todas las series del test, puesto que nada *a priori* obliga a que sea precisamente seis el número de estas series. Es lógico, en efecto, suponer que un test reducido a tres series daría, más groseramente sin duda, pero de manera estadísticamente significativa, un diagnóstico análogo al sugerido por otras tres series. Aquí está el argumento de la coherencia interna, valedero y utilizado en favor o en contra de todos los otros tests mentales y, por tanto, sin apelación, puesto que la no-homogeneidad de las partes probaría la no-existencia de la facultad (aquí la necesidad pulsacional) que él pretende medir. Hemos, pues, hallado las correlaciones en 50 sujetos adultos, varones y normales, entre el número de elecciones recibidas en cada factor para las tres primeras y las tres últimas series del test en cuatro aplicaciones sucesivas: ninguna de las ocho correlaciones es significativamente “no nula” al nivel del 1 por 100: no existe homogeneidad apreciable entre fotos de una misma enfermedad.

Esta prueba, realmente crucial para el test, ha sido efectuada, además, a la vez en el Maudsley's Hospital de Londres por el Sr. A. Lubin en el servicio del doctor

SRTA. A. E. ANCELIN Y DRES. H. DUCHENE Y M. SCHÜTZENBERGER 445

Eysenk con el mismo resultado. Ello nos permite concluir que:

- 1.º No hay concordancia entre la teoría de L. Szondi y los resultados experimentales.
- 2.º El número total de elecciones por categoría nosológica del test carece de significación, incluso con respecto a una interpretación "ingenua".

Al contrario, el estudio, quizá no del todo minucioso, de las elecciones revela la presencia de factores de simpatía que resultan de categorías completamente ajenas a la nosología de L. Szondi. Demos algunos ejemplos:

Nuestro grupo de 50 estudiantes (hembras) difiere estadísticamente del grupo de varones correspondiente en su actitud frente a las fotografías de mujeres; estas 17 fotografías reciben, en total, 216 antipatías en las dos categorías, pero 184 simpatías en las estudiantes, contra sólo 140 en los estudiantes ( $\chi^2 = 8,026$  para dos grados de libertad, luego la diferencia es significativa, y no puede ser atribuída al azar). Esta diferencia se debe, sobre todo, a la catatónica de la 5.ª serie y a la histórica de la 6.ª serie. Por tanto, hemos buscado en un tercer grupo (50 adultos varones normales) la correlación entre estas dos fotografías; es de 0,398, es decir, muy significativa. Esto quiere decir que las dos fotografías tienen "algo" de común que lleva a los sujetos a preferirlas o despreciarlas simultáneamente (se entiende de una manera estadística) en mucho mayor grado que el atribuible al puro azar. No hay necesidad de insistir sobre lo que este hecho supone, tanto contra la teoría de L. Szondi (las dos fotografías pertenecen a dos grupos nosológicos diferentes) como contra la manera de puntuar, ya que estas dos fotografías actúan en cierto modo a la par.

Otro ejemplo: Hay dos perfectos "barbudos" ( $p$  en

la 2.<sup>a</sup> serie y *e* en la 4.<sup>a</sup>). En el mismo colectivo hemos encontrado una correlación de 0,430. Esto equivale a decir que al 70 por 100 de los sujetos a los que les “gusta” la primera de estas fotos, les gusta también la segunda. Por tanto, que sólo el 30 por 100 de los que rechazan la primera escogen la segunda.

Otro ejemplo: parece —aunque nuestras muestras sean un poco débiles— que los estudiantes y los aprendices difieren en su actitud frente al histérico de la 1.<sup>a</sup> serie (que de un modo casi exclusivo no despierta simpatías más que en los estudiantes) y del sádico de la 2.<sup>a</sup> serie (que suscita mucho menos antipatía entre los aprendices). En otros 100 sujetos adultos varones hemos encontrado una correlación de —0,195, significativa al 5 por 100 y negativa, como cabía esperar *a priori*, dado el poder discriminativo de estas fotografías en los dos grupos anteriores.

Aquí, también, como en los ejemplos precedentes, el examen de las fotografías del test muestra que los factores que motivan estas elecciones son factores de la más ingenua psicología cotidiana y no tienen nada que ver con genes misteriosos ni engramas inquietantes: a unos les gustan los barbudos; a otros, no; para los unos resulta simpático el rostro de la mujer preocupada (*soucieuse*); otros prefieren la vulgaridad risueña del sádico de la 2.<sup>a</sup> serie, desdeñando el tono (*allure*) vagamente ascético y meditativo del histérico de la 1.<sup>a</sup> serie. Un análisis factorial de las elecciones de los sujetos revelaría, con seguridad, factores interesantes como fundamento de las simpatías y las antipatías.

En todo caso, según creemos haberlo probado, estos factores reales no tienen nada que ver con los que L. Szondi ha fijado *a priori*, y nada parece sugerir que estos fac-

SRTA. A. E. ANCELIN Y DRES. H. DUCHENE Y M. SCHÜTZENBERGER 447

tores puedan ser relacionados de manera útil con los caracteres psiquiátricos.

Los ejemplos que hemos dado no solamente demuestran la inanidad de la teoría de L. Szondi, sino que también impiden, al parecer, conceder el menor valor al test, en espera de un largo estudio experimental que queda enteramente por hacer.

#### LA FIDELIDAD DEL TEST.

Parecería, pues, superfluo poner de relieve nuevas contradicciones entre la teoría de L. Szondi y los resultados experimentales; solamente deseamos hacer aquí algunas observaciones relativas a la fidelidad del test, que constituirán una nueva prueba del carácter artificial de todo el edificio.

En un grupo de 100 varones adultos normales que pasaron dos veces la prueba en un intervalo de veinticuatro horas, los psicodiagnósticos vectoriales, para los cuales L. Szondi da tablas, experimentaron los cambios siguientes:

3 sujetos: 4 psicodiagnósticos concordantes.  
 7 sujetos: 3 psicodiagnósticos concordantes.  
 35 sujetos: 2 psicodiagnósticos concordantes.  
 38 sujetos: 1 psicodiagnóstico concordante.  
 17 sujetos: 0 psicodiagnósticos concordantes.

vemos, pues, que en el espacio de una noche, sólo tres sujetos se mantienen constantes y 17 cambian todos los psicodiagnósticos vectoriales.

Incluso si se admite la posibilidad de “carga” y de “descarga” de la necesidad “pulsional”, hemos de reconocer en los resultados una labilidad inquietante:

De los 12 “rein Abel” que teníamos un día, sólo dos lo eran todavía al siguiente día.

- 1 se había convertido en: anfallsartig auftretende Angst mit Aggression.
- 2 se habían convertido en: Paniktodstill reflex.
- 3 se habían convertido en: hysteriform Beklemmung.
- 1 se habían convertido en: Angst von bewussten Elementen.
- 3 se habían convertido en: paranoid Angst (\*).

De los tres "rein Cain" iniciales:

- 1 se había convertido, durante la noche, en "rein Abel",
- 1 se habían convertido, durante la noche, en "Angst von bewussten Elementen",
- 1 se había convertido, durante la noche, en "Beklemmung mit Jammern".

Sería cruel insistir más.

Señalamos, sin embargo, para los que quieran estudiar de nuevo los tests que, con veinticuatro horas de intervalo, sólo el 50 por 100, aproximadamente, de las elecciones permanecen constantes, si bien en las simpáticas suele haber más constancia que en las antipáticas. Esta fidelidad mediocre, propia del material utilizado, deja pocas esperanzas, incluso para una revisión total de la prueba.

#### CONCLUSIONES.

Las investigaciones experimentales que hemos efectuado en más de 500 sujetos —de los cuales no hemos dado aquí más que algunos de los resultados más salientes— levantarán ciertamente protestas: ¿cómo pueden explicarse, se nos dirá, los notables resultados obtenidos por L. Szondi y algunos de sus alumnos? ¿Por qué no se ha aplicado el test a la población de enajenados a la cual está destinado? A esta última cuestión la respuesta es fácil: ante todo, porque, saliendo del terreno de la psiquiatría, el test y la teoría de L. Szondi se lanzan constantemente a la conquista de dominios nuevos y hacen grandes estragos en la orientación y la selección profesionales: se los quiere aplicar a la eugenesia y a la pedagogía.

(\*) En alemán en el texto original.

SRTA. A. E. ANCELIN Y DRES. H. DUCHENE Y M. SCHÜTZENBERGER 449

a la criminología, etc. Era, pues, necesario probar el test en el medio normal.

Pero hay un segundo punto: el carácter eminentemente personal de la compilación nosológica, que constituye la base misma de la prueba y que deja la puerta abierta a todas las discusiones sobre la validez de los diagnósticos que hubiéramos podido utilizar para valorar el test.

Puesto que, según L. Szondi y nuestros tanteos estadísticos, el 70 por 100 de los varones adultos normales no están comprendidos en la “sexualidad normal”, parece que para L. Szondi este término tiene un sentido especial. ¿Qué contenido operacional puede tener el “Drill Ich” que caracteriza a 1/3 de nuestros sujetos?

Hagamos notar, asimismo, que L. Szondi y algunos de sus mejores discípulos, emplean, juntamente con el test propiamente dicho, la técnica del análisis libre, donde el enfermo, inducido a comentar oralmente cada una de sus elecciones, sufre una verdadera prueba proyectiva que sirve como soporte a la conversación clínica. Para esta técnica, infinitamente más rica y más verdadera, L. Szondi no ofrece más que algunas raras indicaciones, y con razón, ya que toda la ciencia clínica y psiquiátrica no se puede reducir a esquema.

Esto basta, a nuestro parecer, para explicar el éxito de L. Szondi y de ciertos psiquiatras utilizadores del test; para los otros, la vaguedad y la originalidad de las categorías nosológicas sirven como soporte verbal al servicio de su “flair” eventual que no se somete a la vigilancia de ningún criterio operacional de validez.

Precisamente es éste, a nuestro parecer, el problema psicológico más apasionante que ofrece el test de Szondi: el éxito obtenido con él, a pesar de su falta absoluta de valor.

29

BIOMÉTRIE. — *Sur une différence pondérale à la naissance entre deux lignées de souris Mus musculus*. Note (\*) de M. **NICOLAS KOBOZIEFF**, M<sup>me</sup> **N. POMRIASKINSKY-KOBOZIEFF** et M. **MARCEL-PAUL SCHÜTZENBERGER**, présentée par M. Maurice Caullery.

La comparaison des poids des souris à la naissance a déjà fait l'objet de deux travaux antérieurs [C. C. Green (<sup>1</sup>) entre *Mus musculus* et *Mus bactrianus* et T. Vetulani (<sup>2</sup>) entre les lignées *P.* et *Lilliac* de *Mus musculus*].

En prenant la précaution d'effectuer les pesées avant que l'âge du souriceau n'atteigne quinze heures, nous avons comparé deux lignées consanguines depuis 1937 : *P* et *L*, cette dernière provenant du stock mutant (troncature de l'oreille, malformations oculaires, nanisme, hydrocéphalie, etc.), mais ne présentant plus d'anomalies.

Aucune différence de fécondité entre les deux lignées n'a pu être mise en évidence, puisque ni la différence du nombre moyen de petits par portée, ni le rapport des variances ne sont significatifs (lignée *P* : 5,21 petits, lignée *L* : 4,91 petits en moyenne; *t* de Student = 0,987 pour 134 degrés de liberté). Cette constatation nous autorise à comparer les poids moyens dans les deux lignées sans tenir compte du nombre des petits par portée.

On trouve : lignée *P* : 1<sup>er</sup>,76; lignée *L* : 1<sup>er</sup>,30, soit un *t* de Student égal à 23,94, pour 686 degrés de liberté. On constate également que la distribution des poids dans la lignée *P* est moins dispersée que dans la lignée *L*, et ceci de manière significative, puisque le rapport des deux variances atteint 2,399.

La conclusion est donc certaine : les souris de la lignée *P* sont plus lourdes à la naissance que celles de la lignée *L* et les deux lignées sont génotypiquement différentes.

Pour obtenir un résultat plus précis, on est amené à comparer entre elles les souris appartenant à des portées de même nombre de petits (*cf.* tableau) et l'on vérifie que la différence des poids est constante.

(\*) Séance du 28 novembre 1949.

(<sup>1</sup>) *J. exp. Zool.*, 59, 1931, p. 213-245.

(<sup>2</sup>) *Mém. Inst. Polon. Econ. rurale Pulawy*, 11, 1930, p. 456-476.

( 2 )

		Nombre de petits dans la portée.								
		1.	2.	3.	4.	5.	6.	7.	8.	9.
Nombre de portées.	$P...$	1	3	5	14	11	15	9	6	1
	$L...$	1	4	9	17	15	12	6	3	3
Poids moyen d'un petit.	$P...$	2,400	1,675	2,123	1,752	1,960	1,738	1,680	1,620	1,783
	$L...$	1,400	1,563	1,454	1,336	1,367	1,218	1,204	1,192	1,209

Nous avons cherché à exprimer cette régression du poids à la naissance  $y_n$  en fonction du nombre  $n$  de petits dans la portée : on obtient ainsi les deux équations :

$$\begin{aligned} \text{Lignée } P..... & y_n = 2,1754 - 0,0871n + 0,024n^2 \\ \text{Lignée } L..... & y_n = 1,8468 - 0,1590n + 0,098n^2 \end{aligned}$$

Par contre, la régression du poids moyen  $Y_n$  de toute la portée sur le nombre de petits qu'elle comporte se laisse représenter par des équations qui sont seulement linéaires :

$$\begin{aligned} \text{Lignée } P..... & Y_n - 9,154 = 1,503x(n - 5,212) \\ \text{Lignée } L..... & Y_n - 6,373 = 1,072x(n - 4,914) \end{aligned}$$

Ces équations permettent une comparaison plus précise de l'ensemble des résultats.

Dans une prochaine Note, nous montrerons comment s'hérite ce caractère selon le type du croisement effectué.

(Extrait des *Comptes rendus des séances de l'Académie des Sciences*,  
t. 229, p. 1267-1268, séance du 5 décembre 1949.)

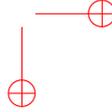
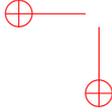
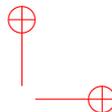
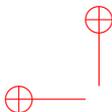
# Table des matières

## Tome I

<b>Préface</b>	<b>iii</b>
<b>Introduction</b>	<b>v</b>
<b>1943–1946</b>	<b>1</b>
1943-1 Sur la théorie des structures de Dedekind . . . . .	2
1944-1 Sur les structures de Dedekind . . . . .	4
1945-1 Sur certains axiomes de la théorie des structures . . . . .	6
1946-1 Méthodes typologiques en sélection professionnelle . . . . .	9
<b>1947</b>	<b>13</b>
1947-1 Remarques sur la notion de clivage dans les structures algébriques . . . . .	14
1947-2 Sur certains treillis gauches . . . . .	17
1947-3 Remarques sur des relations d'ordre entre variables aléatoires indépendantes . . . . .	20
1947-4 L'excitation des chimiorécepteurs de la langue par les substances du groupe des narcotiques indifférents et la règle thermodynamique de la narcose . . . . .	22
1947-5 Sur certains paramètres caractéristiques des systèmes d'événements compatibles et dépendants et leur application au calcul des cumulants de la répétition . . . . .	24
1947-6 Axiomatisation de la géométrie dans un complexe linéaire de droites . . . . .	26
<b>1948</b>	<b>31</b>
1948-1 Le mécanisme physico-chimique de l'excitation sapide et la notion d'excitant indifférent . . . . .	33
1948-2 L'étude des conditions thermodynamiques de l'excitation olfactive et les théories de l'olfaction . . . . .	35

## Table des matières

1948-3	Recherches statistiques sur la distribution du sexe à la naissance . . . . .	37
1948-4	Valeurs caractéristiques du coefficient de corrélation par rang de Kendall dans le cas général . . . . .	39
1948-5	Fidélité des tests . . . . .	41
1948-6	Étude statistique d'un problème de sociométrie . . . . .	45
1948-7	Étude pharmacodynamique des chimiorécepteurs de l'olfaction. (1.) Principes et méthodes d'une étude thermodynamique de l'olfaction . . . . .	54
1948-8	Étude pharmacodynamique des chimiorécepteurs de l'olfaction. (2.) L'étude thermodynamique des séries homologues et les théories de l'olfaction . . . . .	76
1948-9	Remarques sur les relations aléatoires d'ordre et leur application à la psychologie . . . . .	98
1948-10	An abac for the simple range . . . . .	113
1948-11	Sur certaines applications remarquables des treillis dans eux-mêmes . . . . .	116
1948-12	Contribution à l'étude statistique du sexe à la naissance . . . . .	119
<b>1949</b>		<b>173</b>
1949-1	Sur une asymétrie latérale statistique, attribut du phénotype sexuel humain . . . . .	175
1949-2	Sur la détermination du sexe chez l'homme . . . . .	177
1949-3	Remarque sur la statistique des becs-de-lièvre . . . . .	179
1949-4	L'étude des dermatoglyphes . . . . .	181
1949-5	Une application de l'analyse séquentielle . . . . .	191
1949-6	Résultats d'une enquête sur la distribution du sexe dans les familles nombreuses . . . . .	194
1949-7	Fondements de la statistique appliquée à la psychologie, I. Principes des tests statistiques . . . . .	197
1949-8	Fondements de la statistique appliquée à la psychologie, II. Estimation d'un paramètre inconnu . . . . .	200
1949-9	À propos de la "cybernétique" (mathématiques et psychologie) . . . . .	203
1949-10	Sur l'extension des théorèmes de dualité aux treillis distributifs non complémentés . . . . .	226
1949-11	A non-existence theorem for an infinite family of symmetrical block designs . . . . .	228
1949-12	Quelques réserves sur la méthodologie de L. Szondi . . . . .	230
1949-13	Investigaciones críticas sobre la teoría y el test de L. Szondi . . . . .	238
1949-14	Sur une différence pondérale à la naissance entre deux lignées de souris: <i>Mus musculus</i> . . . . .	253



## Marcel-Paul Schützenberger

### ŒUVRES COMPLÈTES

éditées par Jean Berstel, Alain Lascoux et Dominique Perrin

Les treize tomes de cette édition contiennent l'ensemble des œuvres de Marcel-Paul Schützenberger qui ont fait l'objet d'une publication dans une revue scientifique ou un livre. Ses travaux couvrent une période de plus de 50 ans, depuis sa première note aux Comptes Rendus en 1943 jusqu'à son dernier article, paru en 1997.

Les publications sont présentées dans l'ordre chronologique. Chaque tome est précédé d'une courte introduction qui essaie d'éclairer certains des travaux, tant pour leur intérêt scientifique intrinsèque que pour l'écho qu'ils ont rencontré et les développements qu'ils ont suscités.

---

#### Tome 1 : 1943 – 1949

*Ce premier tome des travaux de Schützenberger contient ses écrits de 1943 jusqu'en 1949. Il a publié à cette époque une série de notes aux Comptes-Rendus sur la théorie des treillis, qu'il a étudiés sous l'impulsion d'Albert Châtelet.*

*On trouve également dans ce tome un premier ensemble de publications de statistique appliquée à la médecine, ou de statistique pure. Son expérience du calcul statistique lui permet, selon ses propres dires, « de remplacer, dans la mesure du possible, les longs calculs par des méthodes expéditives souvent aussi efficaces ». Durant cette période, il a fait des exposés introductifs de statistique mathématique, appliquée à la biologie ou à la psychologie.*

*Il soutient une thèse en médecine, intitulée « Contribution à l'étude statistique du sexe à la naissance » à Paris le 2 juillet 1948. L'étude a porté sur 14 230 familles de cinq enfants ou plus, soit un total de 95 875 naissances. Ce travail a été couronné du prix du Baron Larrey par l'Académie Nationale de Médecine.*