

REVUE DE STATISTIQUE APPLIQUÉE

ODILE CADET

Un exemple de plan d'expériences

Revue de statistique appliquée, tome 29, n° 3 (1981), p. 43-53

http://www.numdam.org/item?id=RSA_1981__29_3_43_0

© Société française de statistique, 1981, tous droits réservés.

L'accès aux archives de la revue « *Revue de statistique appliquée* » (<http://www.sfds.asso.fr/publicat/rsa.htm>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

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

UN EXEMPLE DE PLAN D'EXPERIENCES

Mme Odile CADET (1)

1. INTRODUCTION

Le type de plan d'expériences décrit et analysé ci-dessous présente la particularité d'expérimenter successivement sur une même unité expérimentale tous les niveaux de certains facteurs étudiés. De façon plus précise les facteurs étudiés se décomposent en deux groupes. A niveaux fixés des facteurs du premier groupe, on considère un certain nombre d'unités expérimentales, chacune recevant successivement tous les niveaux du facteur du second groupe lors d'une expérience qui se déroule en plusieurs étapes (autant qu'il y a de niveaux des facteurs du second groupe).

Si on considère qu'il peut y avoir un effet de l'étape de l'expérience, sans interaction avec tout autre facteur, l'ordre dans lequel sont expérimentés les traitements ne doit pas être identique pour chaque unité expérimentale ; en conséquence on organise l'expérience de telle façon que tous les ordres possibles de ces traitements soient expérimentés ; à niveaux fixés des facteurs du premier groupe, le nombre d'unités expérimentales choisi est alors égal au nombre de permutations possibles des traitements.

Un cas particulier d'un tel plan a été utilisé dans une expérience d'apprentissage de la lecture et fait l'objet d'exemple d'application dans cet exposé.

II. PLAN D'EXPERIENCE

On considère deux facteurs F_1 et F_3 ayant respectivement ν_1 et ν_3 niveaux et dont on veut étudier les effets principaux et les interactions. Pour cela on dispose pour chaque niveau de F_1 d'unités expérimentales qui ne constituent pas une population a priori homogène. Pour contrôler cette hétérogénéité on applique à chaque unité expérimentale successivement chacun des ν_3 traitements définissant les niveaux de F_3 , ces unités expérimentales jouant alors un rôle de "bloc". Afin de ne pas négliger l'effet éventuel de l'ordre dans lequel sont appliqués à une

(1) Laboratoire de Statistique et Probabilités Université Paul Sabatier 118 Route de Narbonne 31 062 Toulouse Cédex.

même unité expérimentale les ν_3 traitements, on considère suffisamment d'unités expérimentales pour que chacune des ν_3 ! permutations des niveaux de F_3 soit appliquée à une unité expérimentale ce qui nécessite au total un nombre d'unités expérimentales égal à $\nu_1 \cdot \nu_3$!.

Prenons comme exemple, le plan dont nous avons été amenés à analyser les résultats ; l'expérience se déroule de la façon suivante : Deux classes (CM1 – CE1) de garçons (respectivement de filles) sont soumises à l'expérience décrite ci-dessous :

– on présente à chaque enfant quatre groupes composés de trois mots chacun, où un groupe de mots est établi en fonction du caractère imagé et de la fréquence d'apparition de ces mots dans le langage courant. Ces quatre groupes sont définis par les caractéristiques :

- “mots imagés et fréquents”
- “mots non imagés et fréquents”
- “mots imagés et non fréquents”
- “mots non imagés et non fréquents”.

Dans chaque groupe, les mots sont de nature syllabique différente. Sur chaque enfant on mesure le seuil perceptif et le temps de réaction verbale. Les niveaux de F_1 sont alors définis par le croisement de deux facteurs (sexe et âge), quant au facteur F_3 , d'une part il résulte également du croisement de deux facteurs (image et fréquence), d'autre part la mise en œuvre de l'un quelconque de ces ν_3 traitements consiste en plusieurs “sous-traitements” qui sont en nombre a si bien qu'il intervient en facteur F_4 (dans l'exemple : nature syllabique du mot) hiérarchisé dans F_3 .

Les permutations considérées plus haut n'ont porté que sur les niveaux de F_3 c'est dire dans l'exemple sur les groupes de mots ; à niveau fixé de F_3 , l'ordre de présentation des niveaux de F_4 est considéré comme sans influence.

On peut donc considérer la situation générale définie par :

- un premier groupe de k_1 facteurs $F_{1,1} \dots F_{1,k_1}$ (dont le croisement définit le facteur F_1) ayant respectivement $\nu_{1,1} \dots \nu_{1,k_1}$ niveaux avec

$$\nu_1 = \prod_{i=1}^{k_1} \nu_{1,i}$$

- un second groupe de k_3 facteurs $F_{3,1} \dots F_{3,k_3}$ (dont le croisement définit le facteur F_3) et ayant respectivement $\nu_{3,1} \dots \nu_{3,k_3}$ niveaux avec

$$\nu_3 = \prod_{i=1}^{k_3} \nu_{3,i}$$

et un facteur F_4 hiérarchisé dans le croisement de ces k_3 facteurs et ayant au total $\nu_4 = \nu_3 \cdot a$ niveaux dont a sont “emboîtés” dans chacun des niveaux de F_3 et constituent les a opérations dont se compose le traitement défini par ce niveau de F_3 .

- un facteur F_2 (dans l'exemple : facteur enfant) ayant ν_2 niveaux qui constitue le facteur unité expérimentale et qui est emboîté dans le facteur F_1 : chaque unité expérimentale est soumise à chacun des niveaux du facteur F_4 , une unité expérimentale donnée recevant au cours de ν_3 étapes consécutives les ν_3

niveaux du facteur F_3 suivant l'une des permutations de ces ν_3 niveaux. On peut donc considérer l'ensemble E des cellules e définies par :

$$E = \{e = (e_{1,1}, \dots, e_{1,k_1}, e_2, e_{3,1} \dots e_{3,k_3}, e_4, e_5)\}$$

où

- $e_{1,1} \dots e_{1,k_1}$ sont les niveaux respectivement de $F_{1,1} \dots F_{1,k_1}$
- e_2 est la coordonnée de l'unité expérimentale $1 \leq e_2 \leq \nu_2$; une unité expérimentale donnée est affectée à des niveaux donnés des facteurs du premier groupe si bien qu'il existe une application surjective γ_2 de $\bar{\nu}_2$ sur $\prod_{i=1}^{k_1} \bar{\nu}_{1,i}$ qui

à toute unité expérimentale donnée e_2 associe les niveaux des facteurs du premier groupe auxquels elle est affectée. (de façon générale $\bar{\nu}$ désigne l'ensemble $\{1, 2, \dots, \nu\}$ où ν est un entier naturel quelconque). γ_2 est telle que $\text{card}(\gamma_2^{-1}(e_1)) = \nu_3 !$.

- $(e_{3,1}, \dots, e_{3,k_3})$ sont les niveaux définissant F_3 et e_4 est le niveau F_4 avec $1 \leq e_4 \leq \nu_4$ ($\cong a\nu_3$). Il existe une application surjective γ_4 de $\bar{\nu}_4$ sur $\prod_{i=1}^{k_3} \bar{\nu}_{3,i}$ qui précise les niveaux de F_4 hiérarchisés dans chacun des niveaux de F_3 ; $\text{card} \{\gamma_4^{-1}(e_3)\} = a$.
- e_5 désigne l'étape de l'expérience et varie de 1 à ν_3 .

Le plan d'expérience utilisé est la partie E^* de l'ensemble ci-dessus qui contient les cellules telles que e_5 est une fonction de e_2 et de $e_3 = (e_{3,1} \dots e_{3,k_3})$ puisque une unité expérimentale donnée e_2 reçoit tous les traitements dans un ordre déterminé, si bien qu'un traitement donné lui est appliqué à une étape donnée de l'expérience.

De façon plus précise, à toute unité expérimentale e_2 , on peut associer la permutation π_{e_2} de $\bar{\nu}_3$ sur lui-même qui précise l'ordre dans lesquels sont appliqués à e_2 les traitements e_3 ; on a alors $e_5 = \pi_{e_2}(e_3)$ et par suite :

$$E^* = \{e = (e_{1,1} \dots e_{1,k_1}, e_2, e_{3,1} \dots e_{3,k_3}, e_4, \pi_{e_2}(e_3))\}$$

Puisque pour e_1 fixé il y a autant de valeurs de e_2 que de permutations de $\bar{\nu}_3$ sur lui-même, on peut identifier chaque unité expérimentale e_2 à la permutation π_{e_2} et c'est pourquoi on peut éventuellement écrire E^* sous la forme :

$$E^* = \{e = (e_{1,1} \dots e_{1,k_1}, \pi, e_{3,1} \dots e_{3,k_3}, e_4, \pi(e_3))\}$$

où π appartient au groupe symétrique S_{ν_3} de degré ν_3 des permutations de $\bar{\nu}_3$ sur lui-même.

Pour e_1 et $\pi(e_3)$ fixés, lorsque π décrit S_{ν_3} on rencontre alors chaque valeur de e_3 un même nombre de fois qui est $(\nu_3 - 1) !$.

III. MODELE

On suppose que pour chaque élément e de E^* on dispose d'une observation x_e d'une variable aléatoire: $X_e: N[m(e), \sigma^2]$ où $m(e)$ et σ^2 sont inconnus et où les $n(= \nu_1 \nu_3 ! \nu_3 a)$ variables aléatoires X_e sont supposées indépendantes entre elles. (La présentation et les notations spécifiées sont celles utilisées par J.R. BARRA [1]). m est une application de E^* dans \mathbf{R} et on note Ω l'espace vectoriel des applications de E^* dans \mathbf{R} , muni du produit scalaire :

$$\langle m, m' \rangle = \sum_{e \in E^*} m(e) m'(e) \quad \forall m, m' \in \Omega.$$

Puisque le plan E^* est incomplet on est amené à supposer dans le modèle que certaines interactions sont nulles. On suppose donc que m peut s'écrire de la façon suivante :

$$\begin{aligned} m = m_\phi + m_1(e_1) + m_{1,2(e_1)}(e_1, \pi) + m_3(e_3) + m_{3,4(e_3)}(e_3, e_4) \\ + m_{\pi(3)}(\pi(e_3)) + m_{1,3}(e_1, e_3) + m_{1,2(e_1),3}(e_1, \pi, e_3) \\ + m_{1,3,4(e_3)}(e_1, e_3, e_4) \end{aligned} \quad (1)$$

nous supposons de plus que les relations suivantes sont satisfaites :

$$\begin{aligned} \sum_{e_1 \in \bar{\nu}_1} m_1(e_1) = \sum_{e_3 \in \bar{\nu}_3} m_3(e_3) = \sum_{\pi(e_3) \in \bar{\nu}_3} m_{\pi(3)}(\pi(e_3)) = 0 \\ \sum_{\pi \in \gamma_2^{-1}(e_1)} m_{1,2(e_1)}(e_1, \pi) = 0 \quad \forall e_1 \in \bar{\nu}_1 \\ \sum_{e_4 \in \gamma_4^{-1}(e_3)} m_{3,4(e_3)}(e_3, e_4) = 0 \quad \forall e_3 \in \bar{\nu}_3 \\ \sum_{e_1 \in \bar{\nu}_1} m_{1,3}(e_1, e_3) = \sum_{e_3 \in \bar{\nu}_3} m_{1,3}(e_1, e_3) = 0 \\ \sum_{\pi \in \gamma_2^{-1}(e_1)} m_{1,2(e_1),3}(e_1, \pi, e_3) = \sum_{e_3 \in \bar{\nu}_3} m_{1,2(e_1),3}(e_1, \pi, e_3) = 0 \quad \forall e_1 \in \bar{\nu}_1 \\ \sum_{e_1 \in \bar{\nu}_1} m_{1,3,4(e_3)}(e_1, e_3, e_4) = \sum_{e_4 \in \gamma_4^{-1}(e_3)} m_{1,3,4(e_3)}(e_1, e_3, e_4) = 0 \\ \forall e_3 \in \bar{\nu}_3 \end{aligned}$$

la signification de ces différents paramètres est alors :

- m_ϕ : effet moyen général
- $m_i(e_i)$: effet différentiel du niveau e_i du facteur F_i : pour $i \in \{1, 3, \pi(3)\}$
- $m_{i,j(e_i)}(e_i, e_j)$: effet différentiel du niveau e_j de F_j (emboîté dans le facteur F_j) à niveau e_i du facteur F_i fixé : pour $(i, j(e_i)) \in \{(1, 2(e_1)), (3, 4(e_3))\}$
- $m_{i,j}(e_i, e_j)$: effet d'interaction entre le niveau e_i du facteur F_i et le niveau e_j du facteur F_j : pour $(i, j) = (1, 3)$.

$m_{i,j(e_i),k}(e_i, e_j, e_k)$: à niveau e_i du facteur F_i fixé, effet d'interaction entre le niveau e_j du facteur F_j (emboîté dans le facteur F_i) et le niveau e_k du facteur F_k .

pour $(i, j(e_i), k) \in \{(1, 2(e_1), 3), (3, 4(e_3), 1)\}$

Remarque

$m_{3,4(e_3),1}(e_3, e_4, e_1)$ a même signification que $m_{1,3,4}(e_3)(e_1, e_3, e_4)$

Implicitement, compte tenu de la structure du plan d'expérience, nous supposons nulles les interactions entre le facteur F_5 et tout autre facteur, ainsi que l'interaction entre le facteur F_2 et le facteur F_4 ceci à niveaux e_1 du facteur F_1 et e_3 du facteur F_3 fixés.

D'autre part $m_i(e_i)$ est une application appartenant à Ω , constante par rapport à tout e_ρ $\rho \neq i$, de même $m_{i,j}(e_i, e_j)$ est une application de Ω , constante par rapport à tout e_ρ $\rho \neq i$ et $\rho \neq j$, $m_{i,j(e_i)}(e_i, e_j)$ est une application de Ω constante par rapport à tout e_ρ $\rho \neq i$ et $\rho \neq j(e_i)$ enfin $m_{i,j(e_i),k}(e_i, e_j, e_k)$ est une application de Ω , constante par rapport à tout e_ρ $\rho \neq i$, $\rho \neq j(e_i)$ et $\rho \neq k$.

Si on désigne par :

- Ω_ϕ le sous ensemble de Ω des applications constantes
- Ω_i le sous ensemble de Ω des applications constantes par rapport à tout e_ρ $\rho \neq i$ pour $i \in \{1, 3, \pi(3)\}$
- $\Omega_{i,j(e_i)}$ le sous ensemble de Ω des applications constantes par rapport à tout e_ρ $\rho \neq i$ et $\rho \neq j(e_i)$ pour $(i, j(e_i)) \in \{(1, 2(e_1)), (3, 4(e_3))\}$
- $\Omega_{i,j}$ le sous ensemble de Ω des applications constantes par rapport à tout e_ρ $\rho \neq i$ et $\rho \neq j$ pour $(i, j) = (1, 3)$
- $\Omega_{i,j(e_i),k}$ le sous ensemble de Ω des applications constantes par rapport à tout e_ρ $\rho \neq i$, $\rho \neq j(e_i)$ et $\rho \neq k$, pour $(i, j(e_i), k) \in \{(1, 2(e_1), 3), (3, 4(e_3), 1)\}$.

Tous ces sous ensembles de Ω , sont orthogonaux deux à deux, relativement au produit scalaire défini sur Ω .

D'autre part Ω est la somme directe des sous espaces Ω_I ($I \in \mathcal{J}$) et du sous-espace orthogonal à $\bigoplus_{I \in \mathcal{J}} \Omega_I$, noté $\Omega_{\mathcal{J}}^\perp$ ($\Omega_{\mathcal{J}}^\perp$ est en fait la somme directe des sous-espace de Ω définissant les interactions négligées dans le modèle ci-dessus). On peut écrire : $\Omega = \bigoplus_{I \in \mathcal{J}} \Omega_I \oplus \Omega_{\mathcal{J}}^\perp$

Chaque m_I est laprojection de m sur le sous espace Ω_I ($I \in \mathcal{J}$), la projection de m sur $\Omega_{\mathcal{J}}^\perp$ étant supposée nulle $m = \sum_{I \in \mathcal{J}} m_I(e)$.

Remarque

Cette décomposition ainsi obtenue est une décomposition "condensée". En effet, citons un exemple ; $m_1(e_1)$ est en réalité $m_1(e_{1,1}, \dots, e_{1,k_1})$ et peut se décomposer en fonction de l'effet principal de chacun des facteurs $F_{1,1}, \dots, F_{1,k_1}$, de leurs interactions d'ordre 1, 2, ..., dont les diverses estimations font appel à des procédures classiques d'analyse de la variance.

IV. ANALYSE

Le modèle est le suivant :

$$X(e) = m(e) + U(e)$$

- où – $U(e)$ est une variable aléatoire $N(0, \sigma^2)$ ($\sigma^2 \in \mathbb{R}^{+*}$) ; toutes les variables aléatoires $U(e)$ étant supposées indépendantes entre elles.
- $m(e)$ résume l'influence de l'ensemble des facteurs $F_1 \dots F_5$ et a pour décomposition $m(e) = \sum_{I \in \mathcal{J}} m_I(e)$.

Par la méthode d'estimation moindre carré nous pouvons déterminer les $\widehat{m_I(e)}$ de $m_I(e)$ ($I \in \mathcal{J}$).

$\widehat{m_I(e)} = x_I(e)$ où $x_I(e)$ a pour expression celle de $m_I(e)$ dans laquelle $m(e)$ est remplacé par son estimation $x(e)$.

On a donc :

$$\widehat{m_\phi(x)} = x_\phi = \frac{1}{\nu_1 \nu_3 ! \nu_3 a} \sum_{e \in E^*} x(e)$$

$$\widehat{m_1(e_1)}(x) = x_1(e_1) = \frac{1}{\nu_3 ! \nu_3 a} \sum_{e \in E_{1,1}^*(e_1)} x(e) - x_\phi \quad (= \overline{x_1(e_1)} - x_\phi)$$

$$\widehat{m_{1,2}(e_1)}(e_1, \pi)(x) = x_{1,2}(e_1)(e_1, \pi) = \frac{1}{\nu_3 a} \sum_{e \in E_{1,2}^*(e_1)(e_1, \pi)} x(e) - \overline{x_1(e_1)}$$

$$\widehat{m_3(e_3)}(x) = x_3(e_3) = \frac{1}{\nu_1 \nu_3 ! a} \sum_{e \in E_{3,3}^*(e_3)} x(e) - x_\phi \quad (= \overline{x_3(e_3)} - x_\phi)$$

$$\widehat{m_{3,4}(e_3)}(e_3, e_4)(x) = x_{3,4}(e_3)(e_3, e_4) = \frac{1}{\nu_1 \nu_3 !} \sum_{e \in E_{3,4}^*(e_3)(e_3, e_4)} x(e) - \overline{x_3(e_3)}$$

$$\widehat{m_{\pi(3)}(\pi(e_3))}(x) = x_{\pi(3)}(\pi(e_3)) = \frac{1}{\nu_1 \nu_3 ! a} \sum_{e \in E_{\pi(3)}^*(\pi(e_3))} x(e) - x_\phi$$

$$\widehat{m_{1,3}(e_1, e_3)}(x) = x_{1,3}(e_1, e_3) = \frac{1}{\nu_3 ! a} \sum_{e \in E_{1,3}^*(e_1, e_3)} x(e) - \overline{x_1(e_1)} - \overline{x_3(e_3)} + x_\phi$$

$$\begin{aligned} \widehat{m_{1,2}(e_1),3}(e_1, \pi, e_3)(x) &= x_{1,2}(e_1),3(e_1, \pi, e_3) = \frac{1}{a} \sum_{e \in E_{1,2}(e_1),3(e_1, \pi, e_3)} x(e) \\ &\quad - \frac{1}{\nu_3 a} \sum_{e \in E_{1,2}^*(e_1)(e_1, \pi)} x(e) - \frac{1}{\nu_3 ! a} \sum_{e \in E_{1,3}^*(e_1, e_3)} x(e) + \overline{x_1(e_1)} \end{aligned}$$

$$\widehat{m}_{1,3,4}(e_3)(e_1, e_3, e_4)(x) = x_{1,3,4}(e_3)(e_1, e_3, e_4) = \frac{1}{\nu_3!} \sum_{e \in E_{1,3,4}^*(e_3)(e_1, e_3, e_4)} x(e) - \frac{1}{\nu_3! a} \sum_{e \in E_{1,3}^*(e_1, e_3)} x(e) - \frac{1}{\nu_1 \nu_3!} \sum_{e \in E_{3,4}^*(e_3, e_4)} x(e) + \overline{x_3(e_3)}$$

où on désigne par :

$E_i^*(e_i)$: l'ensemble des éléments e de E^* pour lesquels le facteur F_i est fixé au niveau e_i (pour $i \in \{1, 3, \pi(3)\}$).

$E_{i,j}^*(e_i, e_j)$: l'ensemble des éléments e de E^* pour lesquels à niveau e_i fixé du facteur F_i , le facteur F_j (emboîté dans le facteur F_i) est au niveau e_j (pour $(i, j(e_i)) \in \{(1, 2(e_1)), (3, 4(e_3))\}$).

$E_{i,j}^*(e_i, e_j)$: l'ensemble des éléments e de E^* pour lesquels les facteurs F_i et F_j sont fixés respectivement aux niveaux e_i et e_j (pour $(i, j) = (1, 3)$)

$E_{i,j}^*(e_i, k)$: l'ensemble des éléments e de E^* pour lesquels à niveau e_i fixé du facteur F_i , le facteur F_j (emboîté dans le facteur F_i) est au niveau e_j et le facteur F_k au niveau e_k (pour $(i, j(e_i), k) \in \{(1, 2(e_1), 3), (3, 4(e_3), 1)\}$).

L'estimateur $\hat{\sigma}^2(x)$ de la variance résiduelle σ^2 est défini par :

$$\hat{\sigma}^2(x) = \frac{S^2(x)}{n - r}$$

où $S^2(x) = \sum_{e \in E^*} (x(e) - \sum_{I \in \mathcal{I}} x_I(e))^2$, représente l'estimation du carré de la norme de la projection de m sur $\Omega_{\mathcal{J}}^\perp$.

n désigne le nombre total d'observations c'est-à-dire $\nu_1 \nu_3! \nu_3 a$
 r , nombre de paramètres à estimer a pour valeur

$$\sum_{I \in \mathcal{J}} \dim \Omega_I = \nu_1 \nu_3 (\nu_3! + a - 1) + \nu_3 - 1$$

Tests d'hypothèses

On peut construire à partir des résultats précédents de nombreux tests d'hypothèses relatives aux effets des facteurs $F_1 \dots F_5$; ces hypothèses sont de la forme : $\mathcal{H}_0 : "m_I(e) = 0 \quad I \in \mathcal{J}_0 \subset \mathcal{J}"$ (\mathcal{J}_0 non vide)

$$\Leftrightarrow m = \sum_{I \in \mathcal{J} - \mathcal{J}_0} m_I(e)$$

Le test de l'hypothèse \mathcal{H}_0 est alors défini par la statistique suivante :

$$F_{\mathcal{H}_0}(x) = \frac{\sum_{I \in \mathcal{J}_0} \langle x_I, x_I \rangle}{S^2(x)} \quad \frac{n - r}{\sum_{I \in \mathcal{J}_0} \dim \Omega_I}$$

où $S^2(x)$, n , r sont définis dans le paragraphe précédent.

Sous l'hypothèse \mathcal{H}_0 la statistique $F\mathcal{H}_0(x)$ a pour loi une loi de Fisher ayant pour degrés de liberté $\sum_{I \in \mathcal{J}_0} \dim \Omega_I$ et $n - r$.

Le seuil α étant fixé, le test de l'hypothèse \mathcal{H}_0 est alors immédiat.

Remarque :

Dans toute l'analyse nous avons supposé qu'il y avait une seule observation par unité expérimentale. Dans le cas de plusieurs observations, il serait alors possible d'introduire un terme d'interaction supplémentaire dans le modèle, on sera d'autre part conduit à diverses modifications en particulier dans la détermination de l'estimateur de la variance résiduelle du modèle.

V. APPLICATION

L'exemple de plan d'expériences décrit dans le paragraphe II a pour but d'analyser l'influence de différents facteurs respectivement sur le seuil perceptif et sur le temps de réaction verbale de l'enfant.

En dehors des facteurs image et fréquence dont l'étude est l'objet même de l'expérience, l'organisation de l'expérience a nécessité la recherche des facteurs susceptibles d'influencer les résultats lorsqu'ils auront été introduits dans le dispositif expérimental (ex. : sexe, âge). Le matériel expérimental non homogène a été stratifié en sous groupes aussi homogènes que possible (sexe et âge fixés). A l'intérieur de chacun des sous groupes, chaque unité expérimentale (enfant), jouant ici le rôle de "bloc", a été choisie au hasard et soumise à la série d'expériences. Il est naturel de penser qu'un effet d'accoutumance peut intervenir. Considérant ainsi que l'ordre de présentation des groupes peut avoir une influence, les quatre groupes de mots sont présentés à chaque enfant dans un ordre différent. L'ordre précis dans lequel sont présentées à chaque unité expérimentale les différentes expériences (différents groupes) est apparu un moyen d'atténuer voir d'éviter des liaisons entre expériences successives et d'éviter de favoriser certaines expériences au détriment des autres. Le nombre de permutations possibles des quatre groupes détermine le nombre d'enfants choisis au hasard dans chaque classe. Bien que dans l'expérience, l'ordre de présentation des mots à l'intérieur d'un groupe soit permuté, nous faisons l'hypothèse que l'ordre des mots pour un groupe donné est sans influence ; peut être est-ce là l'hypothèse la plus contestable...

L'ensemble des facteurs retenus est alors :

$$\begin{array}{l}
 F_{1,1} : \text{facteur "sexe"} \quad \nu_{1,1} = 2 \\
 F_{1,2} : \text{facteur "âge"} \quad \nu_{1,2} = 2 \\
 F_2 : \text{facteur "enfant"} \quad \nu_2 = 4 \times 4! \text{ facteur emboîté dans les deux précédents.} \\
 \\
 F_{3,1} : \text{facteur "image"} \quad \nu_{3,1} = 2 \\
 F_{3,2} : \text{facteur "fréquence"} \quad \nu_{3,2} = 2 \\
 F_4 : \text{facteur "nature du mot"} \quad \nu_4 = 4 \times 3 \text{ (a = 3)} \\
 F_5 : \text{facteur "ordre de présentation des groupes"} \quad \nu_5 = 4 \text{ (= } \nu_3 \text{)}
 \end{array}
 \left. \vphantom{\begin{array}{l} F_{1,1} \\ F_{1,2} \\ F_2 \\ F_{3,1} \\ F_{3,2} \\ F_4 \\ F_5 \end{array}} \right\} \begin{array}{l} F_1 : \text{facteur "catégorie d'enfant"} \\ \nu_1 = 4 \\ \\ F_3 : \text{facteur "groupe de mots"} \\ \nu_3 = 4 \end{array}$$

Cette organisation de l'expérience nous conduit à supposer que chaque observation résulte de l'additivité d'un certain nombre de paramètres (partie fixe)

et d'un terme "erreur" (partie aléatoire). Ce dernier terme est supposé de distribution normale et indépendant. Bien que ce type de modèle se rencontre dans la littérature (ex. : [4]) il est certain que cette dernière hypothèse est contestable. Il serait souhaitable de pouvoir auparavant procéder, dans les mêmes conditions, à des expériences (identiques à celle-ci) permettant de vérifier en particulier l'hypothèse de normalité et d'indépendance des erreurs.

La partie fixe se décompose de la façon suivante :

$$\begin{aligned}
 m(e) = & m_{\phi} + m_1(e_1) + m_{1,2}(e_1)(e_1, \pi) + m_3(e_3) \\
 & + m_{3,4}(e_3)(e_3, e_4) + m_{\pi(3)}(\pi(e_3)) + m_{1,3}(e_1, e_3) \\
 & + m_{1,3,4}(e_3)(e_1, e_3, e_4)
 \end{aligned}$$

où $m_1(e_1)$ se décompose en fonction de l'effet principal des facteurs $F_{1,1}$ (sexe) et $F_{1,2}$ (âge) et de leur interaction

$m_3(e_3)$ se décompose en fonction de l'effet principal des facteurs $F_{3,1}$ (image) et $F_{3,2}$ (fréquence) et de leur interaction.

$m_{1,3}(e_1, e_3)$ se décompose :

en fonction de l'effet d'interaction d'ordre 2 entre les facteurs $F_{1,1}$ (sexe) et $F_{3,1}$ (image) (respectivement $F_{1,1}$ et $F_{3,2}$; $F_{1,2}$ et $F_{3,1}$; $F_{1,2}$ et $F_{3,2}$) puis en fonction de l'effet d'interaction d'ordre 3 entre les facteurs $F_{1,1}$, $F_{1,2}$ et $F_{3,1}$ (respectivement $F_{1,1}$, $F_{1,2}$ et $F_{3,2}$; $F_{1,1}$, $F_{3,1}$ et $F_{3,2}$; $F_{1,2}$, $F_{3,1}$ et $F_{3,2}$).

enfin en fonction de l'effet d'interaction d'ordre 4 entre les facteurs $F_{1,1}$, $F_{1,2}$, $F_{3,1}$ et $F_{3,2}$.

$m_{1,3,4}(e_3)(e_1, e_3, e_4)$ se décompose :

en fonction de l'effet d'interaction d'ordre 2 entre les facteurs $F_{1,1}$ et F_4 (nature du mot) (respectivement $F_{1,2}$ et F_4), puis en fonction de l'effet d'interaction d'ordre 3 entre les facteurs $F_{1,1}$, $F_{1,2}$ et F_4 , ceci à niveau fixé du facteur F_3 (c'est-à-dire pour un groupe de mots donné).

Dans ce modèle nous supposons qu'à niveau fixé du facteur F_1 , c'est-à-dire pour un groupe d'enfants déterminé, il n'y a pas d'interaction entre les facteurs F_2 (enfant) et F_3 (groupe de mots).

Les résultats concernant le seuil perceptif et le temps de réaction verbale sont résumés dans les tableaux I et II.

L'expérience s'est déroulée sur 96 sujets et on dispose pour chacune des deux variables étudiées de $n = 1\ 152$ observations ; sur l'ensemble de la population étudiée, le temps de réaction verbale moyen est de 1 183.58 ms. et le seuil perceptif moyen est de 55.97 ms.

Les tests significatifs obtenus dans le cas des deux analyses sont indiqués dans les tableaux I et II ; à titre d'exemple analysons un des résultats obtenus pour le seuil perceptif, à savoir, l'existence d'interaction significative entre les facteurs : âge - sexe et fréquence.

Pour cela considérons les performances moyennes résumées dans les deux tableaux suivants :

CM1			CE1		
	garçons	filles		garçons	filles
fréquence forte	33.45	50.60	fréquence forte	36.64	24.31
fréquence faible	71.74	120.10	fréquence faible	62.53	48.40

Considérons ensuite les écarts entre les performances des filles et des garçons :

	CM1	CE1
fréquence forte	17.15	- 12.33
fréquence faible	48.36	- 14.33

la différence de supériorité fille, garçon entre les petits et les grands n'est pas la même pour la fréquence forte et la fréquence faible. Ce phénomène se traduit par une interaction significative entre les facteurs âge - sexe et fréquence. Une interprétation plus approfondie de tous ces résultats a été faite par le praticien, compte tenu de ses connaissances et de son expérience [2].

TABLEAU I
Seuil de perception

Source de variation	Somme des carrés	Degrés de liberté	F. de Fisher
Sexe	27 426.77	1	5.718*
âge	194 766.01	1	40.609*
Int. sexe × âge	152 237.02	1	31.742*
enfant	3 676 459.72	92	8.332*
Ordre de présentation	47 203.11	3	3.281
fréquence	448 049.44	1	93.419*
image	62 172.19	1	12.963*
Int. fréquence × image	3 672.96	1	0.766
nature du mot	168 674.39	8	4.396*
Int. sexe × fréquence	15 568.77	1	3,246
Int. sexe × image	1 708.69	1	0,356
Int. âge × fréquence	60 161.13	1	12.544*
Int. âge × image	29 050.54	1	6.057*
Int. sexe × fréquence × image	5 899.69	1	1.230
Int. âge × fréquence × image	18 456.01	1	3.848*
Int. sexe × âge × fréquence	19 610.25	1	4.089*
Int. sexe × âge × image	781.77	1	0.163
Int. sexe × âge × fréquence × image	6 908.02	1	1.440
Int. sexe × nature du mot	47 435.58	8	1.236
Int. âge × nature du mot	29 470.74	8	0.768
Int. sexe × âge × nature du mot	67 478.32	8	1.759
Résidus	4 839 298.80	1 009	

* signifie test significatif au seuil de 5 %

TABLEAU II
Temps de réponse verbale

Source de variation	Somme des carrés	Degrés de liberté	F. de Fisher
Sexe	4 167 625.09	1	22.265*
Age	55 967 318.34	1	299.003*
Int. sexe × âge	6 861 968.84	1	36.660*
Enfant	108 167 481.90	92	6.281*
Ordre de présentation	576 049.84	3	1.026
Fréquence	2 986 364.34	1	15.954*
Image	2 227 632.78	1	11.901*
Int. fréquence × image	74 144.59	1	0.396*
Nature du mot	5 912 503.94	8	3.948*
Int. sexe × fréquence	26 699.75	1	0.143
Int. sexe × image	640 995.03	1	3.424
Int. âge × fréquence	505 934.17	1	2.703
Int. âge × image	5859.03	1	0.031
Int. sexe × fréquence × image	87 327.17	1	0.466
Int. âge × fréquence × image	36 607.92	1	0.196
Int. sexe × âge × fréquence	160 319.53	1	0.856
Int. sexe × âge × image	247 866.67	1	1.324
Int. sexe × âge × fréquence × image	120 991.00	1	0.646
Int. sexe × nature du mot	1 060 641.81	8	0.708
Int. âge × nature du mot	937 874.15	8	0.626
Int. âge × sexe × nature du mot	2 131 420.10	8	1.423
Résidus	188 864 452.30	1 009	

* signifie "test significatif au seuil de 5 %"

REFERENCES

- [1] J.R. BARRA. — *Notions fondamentales de Statistique mathématique*, Dunod, 1980.
- [2] M.T. BREU-FACCA. — *L'image mentale dans l'apprentissage de la lecture*, Thèse 3^e Cycle, Toulouse, 1980.
- [3] PHAM DINH TUAN. — *Contribution à l'analyse de la variance et aux plans d'expérience*. Thèse 3^e Cycle, Grenoble, 1970.
- [4] MAXWELL. — *Experimental Design in Psychology and the medical sciences*, Methuen, London, 1958.