

J.-P. BENZÉCRI

Validité des échelles d'évaluation en psychologie et en psychiatrie et corrélations psychosociales

Les cahiers de l'analyse des données, tome 17, n° 1 (1992), p. 55-86

http://www.numdam.org/item?id=CAD_1992__17_1_55_0

© Les cahiers de l'analyse des données, Dunod, 1992, tous droits réservés.

L'accès aux archives de la revue « Les cahiers de l'analyse des données » 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/>

VALIDITÉ DES ÉCHELLES D'ÉVALUATION EN PSYCHOLOGIE ET EN PSYCHIATRIE ET CORRÉLATIONS PSYCHOSOCIALES

[VALID. PSY.]

J.-P. BENZÉCRI

1 Psychométrie et analyse des données

Dès le milieu du XIX^{ème} siècle, la biométrie, ou étude des dimensions physiques des êtres vivants, a considéré de nombreux systèmes de grandeurs naturellement corrélées entre elles: ainsi se fortifia la statistique multidimensionnelle. Mais l'analyse des données telle que nous l'entendons, c'est-à-dire la recherche inductive de dimensions cachées définies par combinaison de mesures primaires (ou d'autres données adéquatement codées), ne se développa qu'avec la psychométrie.

Car de ce que, depuis des millénaires, les hommes ont appelé intelligence, mémoire, imagination, patience... rien ne s'exprime directement par aucune grandeur physique; d'où la nécessité d'une construction statistique, si jamais les chiffres peuvent *mesurer l'âme*. (Tel est, en effet, ce qu'invite à faire le nom même de *psychométrie*).

Il est commun de regarder une telle mesure comme l'introduction de la *quantité* dans un domaine qui avait été exclusivement celui de la *qualité*. À la vérité, les données à partir desquelles on élabore une grandeur numérique assimilée à l'intelligence, à l'anxiété, à la douleur... peuvent être elles-mêmes des nombres: nombre de questions résolues dans un problème, nombre de fois qu'un certain mot a été employé dans un discours, temps pendant lequel on a enduré un stimulus...

Cependant, la psychométrie n'en opère pas moins toujours une transformation du qualitatif en quantitatif, parce que le terme de *quantité* s'oppose à celui de *qualité* non comme l'expression numérique s'oppose à l'expression verbale, mais par une différence de valeur logique.

Suivant ARISTOTE, "le caractère propre de la quantité est qu'on peut lui attribuer l'égal et l'inégal", tandis que "semblable ou dissemblable se dit uniquement des qualités". De ce point de vue, une description multidimensionnelle est

toujours qualitative, même si elle ne comporte que des variables numériques précises, parce que la multiplicité des descriptions possibles est telle qu'on n'en rencontrera jamais d'égales, mais seulement de semblables.

L'analyse des données ne peut opérer que sur des descriptions de forme numérique. Ces descriptions constituent généralement un système de relations, exprimé si besoin est, à partir des données primaires, par un codage approprié. Par l'analyse factorielle, on réduit le nombre des dimensions; par la classification automatique, on ramène l'infini à un nombre fini de types. Ainsi on va du qualitatif au quantitatif; de ce qui ne peut être perçu que comme semblable, à ce qui est susceptible d'égalité.

Nous nous sommes appliqué ailleurs à suivre, (cf. *Histoire et Préhistoire de l'Analyse des Données*, 1975-1982), par quelles voies la statistique est parvenue au point où nous la voyons, traitant, grâce à l'ordinateur, de grands ensembles de données; puis, sans prétendre faire une histoire philosophique exhaustive des rapports entre *qualité* et *quantité*, ou entre *discontinu* et *continu*, nous avons relevé, (cf. "Qualité et quantité dans la tradition des philosophes et en Analyse des Données"; *CAD*, Vol XII, n°1, 1988), des textes de divers penseurs, qui apportent des conceptions contradictoires, toutes présentes, confusément, dans l'usage des statisticiens d'aujourd'hui.

Pour préciser le jeu réciproque entre qualité et quantité, il faudrait considérer, dans le détail, les principes du codage des données et de la construction des tableaux analysés. Une telle recherche offre matière à plusieurs mémoires. Nous tenterons seulement ici d'ordonner les réflexions qu'évoquent deux thèmes, proposés par le Docteur Marc BOSC:

- A) Évolutivité et Validité des échelles d'évaluation en psychiatrie:
Échelle qualitative $\neq$ Échelle quantitative;
- B) Problème de Méthodologie et d'Analyse Statistique d'un
Outil d'évaluation psychosocial.

Ces thèmes requièrent, sans doute, des recherches originales: mais l'expérience acquise dans la pratique du codage aide à en frayer la voie. Cependant, en développant des pensées abstraites formées à partir d'exemples dont les images flottent encore devant nos yeux, nous risquerions de n'être pas compris, si le lecteur ne connaissait, lui aussi, quelque chose de ces exemples.

Le programme du présent exposé sera donc de reprendre tout ce que les "Indices Systématiques" des années 1988-91 de *CAD* offrent sous les mots de **psychologie, échelle et codage**; en précisant le format des données traitées, les raisons du codage adopté et ce qui, dans les résultats, se rapporte aux thèmes proposés; sans nous astreindre à justifier toutes les conclusions, mais en laissant aux idées la liberté qui est permise dans un projet.

2 Codage disjonctif complet, codage flou et recadrage des notes suivant l'équation personnelle

2.1 Codage disjonctif complet

Partons du cas typique d'un questionnaire, ou ensemble Q de questions auxquelles répond un ensemble I de sujets. Le cas d'un ensemble I d'individus ou d'objets décrits suivant un ensemble Q de variables n'en diffère pas formellement; car, pour le traitement statistique, il importe peu qu'un sujet se décrive lui-même par les réponses qu'il donne; ou qu'un observateur fasse la description, en usant, s'il y a lieu, d'appareils de mesure.

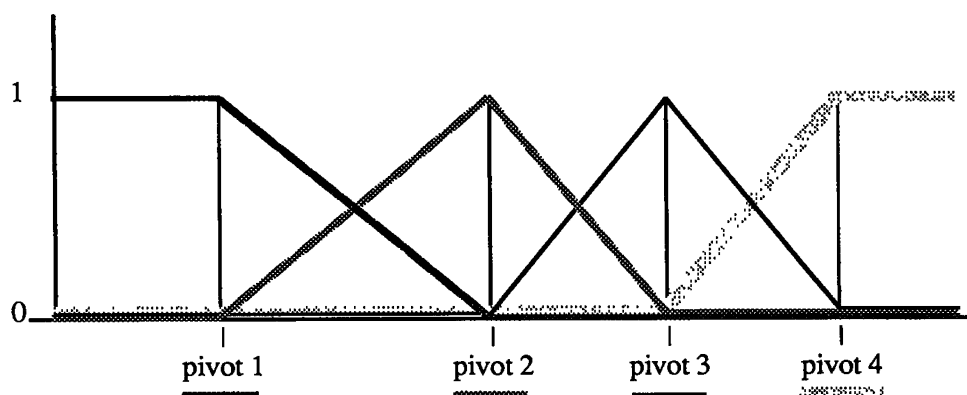
Le format de question le plus simple est la question fermée: l'enquêteur fixe lui-même un ensemble de réponses hors desquelles il n'est pas permis au sujet interrogé de rien dire. Ainsi, dans un signalement, l'ensemble des professions ou des niveaux d'études sera réduit à une nomenclature sans nuance. Ou, si l'on s'enquiert de ce que ferait le sujet dans des circonstances déterminées, on ne lui laisse choisir qu'entre un petit nombre d'issues.

En vue de l'analyse des données, les réponses à un questionnaire fermé sont toujours mises sous *forme disjonctive complète*. À chaque question q il correspond un bloc J_q de colonnes, avec une colonne par modalité: sur la ligne afférente au sujet i , on lit, dans la colonne j , le nombre $k(i, j)$ qui vaut 1 si le sujet a adopté la modalité j , et zéro sinon.

Mais souvent, l'ensemble des modalités de réponses définit seulement une échelle d'intensité, plus ou moins bien jalonnée de locutions adverbiales; allant, par exemple, de 'tout à fait opposé' à 'tout à fait d'accord'; ou de 'presque jamais' à 'très souvent'. Alors, (si l'on met à part la modalité d'abstention communément prévue), la réponse n'a d'autre signification que celle d'un nombre, allant, e.g., de 1 à 5, si l'on a cinq modalités; et cette suite discontinue représente une variable continue, dont on peut supposer que, pour un choix approprié d'unité, elle varie de 0 à 1.

C'est pourquoi, il n'est pas rare qu'on présente franchement des échelles: une question est accompagnée d'une ligne horizontale où le sujet doit marquer sa réponse; d'autant plus à droite que son approbation, son adhésion,... est plus nette; à moins qu'on ne mette à chaque extrémité de la ligne une phrase ou une locution caractérisant brièvement l'une ou l'autre de deux attitudes contraires.

Ainsi, quant au format, la question fermée rejoint la variable continue: car, d'une part, une question fermée, ne propose parfois rien d'autre que les degrés successifs d'une variable continue sous-jacente; et, d'autre part, l'intervalle de variation d'une quantité continue peut être divisé en segments consécutifs, la mesure étant traduite en une réponse discrète: le numéro du segment où elle se trouve; d'où, finalement, pour la quantité continue elle-même, un codage en (0, 1) sous forme disjonctive complète.



Recodage barycentrique d'une variable en 4 modalités autour de 4 valeurs pivot : on a figuré avec des traits différents les valeurs prises par chacune des modalités.

2.2 Codage linéaire par morceaux

Mais on peut craindre que le codage en $(0, 1)$ n'offre de la quantité continue qu'une expression trop grossière. Si, par exemple, la capacité d'exercice d'un sujet, mesurée, en secondes, sur une bicyclette ergométrique, est exprimée suivant les trois modalités 'faible' (<400), 'moyenne' (de 400 à 700) et 'forte' (>700), le cas d'un sujet atteignant 720 sera mieux décrit en mettant 0 dans la colonne 'faible' et $1/2$ dans chacune des deux colonnes 'moyenne' et 'forte' qu'en mettant 1 dans la colonne 'moyenne' et 0 dans les deux autres colonnes.

On appelle *codage flou* un codage qui s'écarte du codage disjonctif complet en ce qu'il répartit sur deux colonnes consécutives le poids de la réponse d'un sujet. Sans reprendre toutes les variantes de codage flou, bien étudiées par J. GALLEG0 (in *CAD*, Vol. VII, n°4, 1982), l'article [CODAGE LIN.] (in *CAD*, Vol. XIV, n°2, 1989), propose du *codage linéaire par morceaux*, ou codage barycentrique, une description que nous reproduisons ici afin de préciser les traitements auxquels on soumet les réponses exprimées sur des échelles.

Ainsi qu'il apparaît sur le graphique, il s'agit de recoder une variable dont on suppose que les valeurs v sont des nombres qu'on peut assimiler à une abscisse sur l'axe horizontal. (Dans les applications, il se pourra que v varie continûment, ou prenne seulement un nombre fini de valeurs entières, etc...). Les modalités sont créées autour de valeurs repères appelées ici pivots.

Sur la figure, on a choisi d'illustrer le cas où il y a quatre modalités: {mod1, mod2, mod3, mod4}. Si la variable v est inférieure ou égale à pivot 1, mod1 vaut 1; et les trois autres modalités sont nulles; entre pivot 1 et pivot 2, mod1 décroît linéairement de 1 à 0; et elle reste nulle ensuite. La modalité mod2

est nulle jusqu'à pivot 1; elle croît linéairement de 0 à 1 entre pivot 1 et pivot 2; décroît linéairement de 1 à 0 entre pivot 2 et pivot 3; et reste nulle ensuite. De même, la modalité mod3 ne diffère de 0 que si v est comprise entre pivot 1 et pivot 3; le maximum, 1, étant atteint sur le pivot 2. Enfin la dernière modalité mod4, nulle jusqu'à pivot 3 croît linéairement jusqu'à 1 entre pivot 3 et pivot 4; et garde cette valeur ensuite.

On notera qu'il n'y a jamais plus de deux modalités différentes de 0; que les valeurs des modalités sont toutes positives ou nulles; et que le total en est 1, quelle que soit la valeur v .

On donne généralement pour principal mérite au codage flou qu'il conserve l'information. Dans les résultats de l'analyse des données, cet avantage est surtout sensible s'il n'y a qu'un petit nombre d'individus d'après lesquels apprécier les corrélations entre modalités de questions différentes; ou si, le questionnaire n'étant pas redondant, on ne peut espérer reconnaître qu'un sujet est (d'un certain point de vue) *entre moyen et fort* par le seul fait que, dans un certain bloc de questions apparentées, ses réponses se partagent également entre 'faible' et 'fort'.

Un autre avantage du codage flou est la commodité du découpage des variables. Afin de définir les bornes des classes (ou modalités) consécutives d'une variable, ou de choisir les pivots d'un codage barycentrique, l'utilisateur considère ordinairement divers histogrammes de la variable à coder. Changeant plusieurs fois l'étalement, le nombre des créneaux, il tente de reconnaître des intervalles de faible densité séparant des modes; suppute la longueur de la queue de distribution qu'il convient d'enfermer dans la dernière classe ou de laisser au delà du dernier pivot...

Or la décision est d'autant plus difficile à prendre qu'on redoute qu'en déplaçant de peu une borne on ne modifie nettement le codage d'un groupe d'individus pour lesquels le '1' sautera d'une colonne à la suivante...; rien de tel n'est à craindre avec un pivot, parce que le résultat du codage en dépend continûment.

2.3 Recadrage des notes suivant l'équation personnelle

Mais plus décisif encore en faveur du codage flou nous paraît être le fait qu'il permet d'interpréter les réponses des sujets, de recadrer les notes, d'après ce qu'on a appelé l'*équation personnelle* (cf. [ÉQ. PERS.], in CAD, Vol. XIV, n°3, 1989). Cette locution nous vient de l'astronomie; où, selon Littré, elle désigne le "temps qui s'écoule entre voir et enregistrer, qui varie avec les différents observateurs, et qui entre comme élément dans la correction d'une observation". Présentement, il s'agit d'un décalage, qui varie avec les sujets, entre l'intensité de ce qu'on éprouve, et la note (ou la marque sur une échelle) par laquelle on exprime cette intensité.

Comme il n'y a pas, à ce qu'on éprouve, de corrélat matériel tangible (et que d'ailleurs une opinion ne saurait objectivement être assimilée à une note) la correction ne pourra être qu'approximative et conjecturale. La pratique de l'analyse des données a cependant établi qu'on gagne à considérer comme le zéro propre de l'échelle d'un sujet la moyenne des notes qu'il a attribuées; on utilise ensuite ce zéro pour recadrer les notes entre (-1) et $(+1)$.

En bref, on recadre entre -1 et $+1$ l'ensemble des notes attribuées par un même sujet i , en en calculant la moyenne, le Max et le min. Les notes sont d'abord centrées en en retranchant la moyenne moy. Puis toute note >0 est divisée par $(\text{Max} - \text{moy})$; toute note <0 est divisée par $(\text{moy} - \text{min})$; ainsi les notes données par le sujet i varient de -1 (qui représente min) à $+1$ (qui représente Max). Le tableau des notes recadrées de tous les sujets est alors codé en prenant, pour chaque variable v , les valeurs pivot $\{-1, 0, +1\}$; d'où trois modalités qu'on notera $\{v-, v=, v+\}$.

Ce choix a toutefois l'inconvénient de produire, pour certaines variables v , une modalité extrême, $v-$ ou $v+$, dont le poids est très faible; ce qui peut perturber l'analyse. Supposons, par exemple, qu'une entité v (phrase ou personne) soit bien notée par la majorité des sujets: la note recadrée, $k(i, v)$, sera souvent proche de 1 , la modalité $v+$ sera très lourde; au contraire $k(i, j)$ descendra rarement au dessous de 0 , la modalité $v-$ sera très légère (voire de masse nulle).

On peut équilibrer les modalités en choisissant des pivots adaptés à chaque variable v : qui seront, respectivement, $\{mnv, myv, Mxv\}$, le minimum, la moyenne et le maximum de l'ensemble des valeurs prises par la variable recadrée v , pour tous les sujets. Il est clair que, selon que v est généralement forte ou faible, le pivot central, myv , sera positif ou négatif. Les modalités ainsi définies, par ce qu'on peut appeler un double recadrage (des lignes d'abord; puis des colonnes), sont communément notées: $\{v<, v\approx, v>\}$.

Si nous avons choisi, en définitive, trois modalités par variable, plutôt que deux $\{v\leq, v\geq\}$, c'est afin d'éviter que le partage des attitudes moyennes entre les deux modalités extrêmes ne crée, entre celles-ci, une corrélation préjudiciable à la clarté des résultats. Un tel effet était apparu dans une analyse des "Scrutins de 1967 à l'Assemblée des Nations Unies" (cf. A. HAMROUNI, [O.N.U. 67], in *CAD*, Vol. I, n°3, 1976), en codant les abstentions comme une attitude partagée entre le 'Oui' et le 'Non'.

Nous savons bien que plus les sujets interrogés ont de liberté dans leur réponse, plus celle-ci est difficile à interpréter. D'un autre côté, un format trop rigide (tel que: 'pour' ou 'contre', deux modalités) met le sujet si mal à l'aise qu'on ne peut se fier à ce qu'il a dit. Le statisticien doit s'accommoder des données recueillies, tout en s'appliquant à en aviver, en quelque sorte, la couleur, par le codage.

3 Applications du recadrage suivant l'équation personnelle

3.1 Analyse des réponses de 2000 électeurs à un thermomètre de sympathie vis-à-vis de personnalités politiques grecques, [POLIT. GREC.], par T. Béhrakis et I. Nicolacopoulos, in *CAD*, Vol. XIII, n°2, 1988.

La culture politique de l'Europe du Sud a fait l'objet d'une étude comparative réalisée simultanément (été 1985) en Grèce, en Italie, en Espagne et au Portugal sur la base d'un questionnaire quasi identique pour les quatre pays. L'enquête grecque tombait à la veille des élections générales du 2.6.1985; élections qui, venant à la fin de la cohabitation entre le gouvernement socialiste et le Président de la République C. CARAMANLIS (conservateur), ont abouti à la deuxième victoire du parti socialiste (PASOK).

Parmi les différentes questions, il y avait un thermomètre de sympathie vis-à-vis de dix personnalités politiques, qu'on demandait de noter sur une échelle allant de 1 à 10. Outre ce thermomètre, qui ne diffère pas, quant au format, des échelles utilisées en psychométrie, on a considéré, afin de confirmer l'interprétation, la question "Pour qui voteriez-vous"; dont les six modalités de réponse désignent les cinq partis principaux, en réservant la place du 'non exprimé'.

Nous savons que, laissés libres de choisir entre de nombreux degrés (ici 10), les sujets ont un comportement qui n'est que partiellement déterminé par leur opinion sur le thème de la question. Indépendamment de ce thème, certains tendent à donner des réponses extrêmes, soit favorables (10 ou 9) soit défavorables (1 ou 2); d'autres se refusent à quitter le centre de l'échelle.

D'après une enquête du Pr. A. De BOECK sur les "Attitudes des étudiants du Zaïre vis-à-vis des actes attentatoires aux personnes, aux biens et aux mœurs", le regretté F. MUTUMBO a, dès 1969 (cf. *L'Analyse des Données*, TII C n°9, §3.2), analysé la correspondance entre sujets interrogés et notes répondues (de 0 à 10), indépendamment du contenu précis de la question proposée ($k(i, n)$ étant le nombre de fois que le sujet i a adopté la note n). Dans le plan (1, 2), est apparu le cycle fermé des notes: la conjonction des extrêmes donne à la note zéro une place non seulement voisine de 1, mais intermédiaire entre 1 et 10; mais le 5 tend à rejoindre le centre du cycle, du fait que cette note ronde est systématiquement utilisée par certains sujets concurremment avec 0 et 10, à l'exclusion de toute autre note.

Les auteurs ont analysé le tableau croisant l'ensemble I des individus interrogés, avec l'ensemble J des 30 modalités de réponse {-, =, +} utilisées pour coder les valeurs accordées aux 10 hommes politiques. L'axe 1 apparaît comme un facteur général, opposant la droite à la gauche. En ce sens, le thermomètre fournit une échelle d'évaluation des convictions politiques. Mais, comme de règle, l'analyse révèle d'intéressantes nuances.

Le plan (1,2) montre clairement un effet Guttman; l'ensemble J dessine un croissant parabolique allant des opinions franches (+,-) propres à la gauche (approbation des personnalités de Gauche; rejet de celles de Droite), aux opinions de la droite, en passant par les attitudes mitigées (=).

Parmi les hommes politiques de la droite, Caramanlis, ex-Président, conservateur, apparaît comme le plus généralement accepté: aussi l'opposition à Caramanlis (-KAR) est-elle très excentrique; tandis que l'attitude mitigée (=KAR) est à la jonction des attitudes de la gauche et du sous-nuage des {=}; et que l'approbation de Caramanlis (+KAR) est à la jonction des attitudes de la droite et des {=}; et plus proches de celles-ci.

Le rejet de Kyrkos (eurocommuniste) semble réaliser une certaine conjonction des extrêmes, car -KYR est à l'intérieur de la parabole. La classification ascendante hiérarchique permet de préciser cette intuition. Ayant retenu, de l'ensemble I des sujets interrogés, une partition en 8 classes, on a construit un tableau qui, d'une part, donne les poids relatifs, dans chaque classe, des modalités d'attitude vis-à-vis des hommes politiques et dénombre, d'autre part, les intensions de votes (déclarées dans la question complémentaire). Ce tableau montre que le rejet de Kyrkos, quasi absolu à l'extrême droite (dans la classe 1), se trouve aussi à l'extrême gauche (dans les classes 7 et 8) avec un taux quelque peu supérieur à la valeur moyenne.

3.2 Analyse des questionnaires en ergonomie: l'appréciation des réglages d'un poste de travail, [ERGO. RÉGLAGES], par P. Loslever, T.M. Guerra et D.Roger, in *CAD*, Vol. XIII, n°2, 1988.

Une étude préliminaire ayant montré que les réglages relatifs à la hauteur de l'assise, à celle de la table et à la position du dossier d'un poste de travail bureautique déterminaient le plus la posture adoptée par l'opérateur; la tâche de dactylographie étant, en outre, la plus contraignante, par sa rigidité, sur le plan postural et visuel; les auteurs ont construit un protocole expérimental, autour d'une tâche de dactylographie, afin d'étudier l'influence des trois réglages d'assise, table et dossier. Nous considérerons ici les données subjectives recueillies par questionnaire.

En fonction des données anthropométriques des sujets, et compte tenu des contraintes imposées à la vue par la dactylographie, on a établi pour chaque réglage trois niveaux possibles soit:

hauteur d'assise : {Haute, Moyenne, Basse } ;
 éloignement de la table à l'assise : {Éloignée, Moyenne, Proche } ;
 éloignement du dossier au clavier : {Éloigné, Moyen , Proche } .

par exemple, le niveau moyen de la hauteur d'assise a été fixé, pour chaque sujet, à la hauteur {poplité, sol+talon}; et un écart de 2,5 cm par rapport à ce niveau a été pris pour les niveaux extrêmes.

Des 27 combinaisons possibles des niveaux des 3 réglages, on en a retenu 9, comme étant, selon la littérature ergonomique disponible, les plus acceptables, tout en respectant les principes d'un plan d'expérience en carré latin. (En effet, l'analyse de la variance est très couramment, voire même exclusivement, utilisée en ergonomie, pour évaluer l'influence des facteurs sur une variable donnée; l'expérience a cependant montré, selon les auteurs, que cette méthode ne convenait pas ici, parce que le modèle gaussien est inacceptable.)

L'expérience a porté sur 12 sujets, chacun effectuant (sous chacun des 9 réglages) quatre périodes de travail consécutives, après chacune desquelles il répondait au même questionnaire Q, comprenant 17 questions posées en forme

	beaucoup	pas du tout
1. Avez-vous envie de bouger.....		

d'échelle; dont les 12 premières concernent ce que le sujet éprouve: {envie de bouger, mal à la tête, fatigue, mal aux jambes, aux mains, aux bras, aux yeux, à la nuque, au dos, difficulté de concentration, intérêt pour le texte}; et les 5

	beaucoup trop loin	beaucoup trop près
17. Comment trouvez-vous le dossier		

dernières formulent des appréciations sur les conditions ergonomiques {son, lumière, hauteur d'assise, table, dossier}.

On a d'abord analysé, l'utilisation de la ligne d'échelles par les sujets; la ligne étant divisée en 7 segments égaux consécutifs, on a dénombré les marques mises par les 12 sujets dans ces segments: e.g. $k(s, 2)$ = nombre des marques mises par s dans le 2-ème segment, à partir de la gauche. Huit sujets sur douze utilisent beaucoup la position la plus à droite du segment (7) et la position médiane (4), la modalité 6 n'étant alors pratiquement pas utilisée; mais deux sujets affectionnent les modalités (5, 6); un seul recourt véritablement à la totalité de l'échelle.

Pour chaque sujet, on a recadré séparément, suivant l'équation personnelle, les réponses qu'il a fournies (dans l'ensemble des questionnaires remplis par lui) aux questions {1...12} (ce qu'on éprouve) et {13...17} (appréciations ergonomiques). Les 17 questions ont ainsi été codées suivant 51 modalités. Avec 12 sujets, répondant sous chacune des 9 conditions de réglage, à 4 présentations du questionnaire, on a $12 \times 9 \times 4 = 432$ lignes. En cumulant, pour chaque couple (réglage, rang de présentation), les lignes afférentes aux 12 sujets, on a un tableau 36×51 qui résume l'évolution intra et inter-expérimentale des réponses de l'ensemble des sujets.

L'axe 1 oppose les modalités '+', ($F1 > 0$), aux modalités '-', ($F1 < 0$), des variables {envie de bouger, fatigue, douleur dans le dos}: il s'agit donc d'un facteur de confort général. Comme on le présumait, les plaintes s'intensifient de

la première à la quatrième présentation; mais, selon les réglages, les plaintes sont plus ou moins vives, et s'intensifient plus ou moins. Le meilleur réglage semble être HMP {assise Haute, table Moyenne et dossier Proche}; ceux qui suscitent le plus de plaintes étant BMÉ et BÉM.

Il est plus surprenant, mais aussi plus instructif, de découvrir sur l'axe 2 une opposition entre {table trop basse, dossier trop proche}, ($F2 > 0$), et {table trop haute, dossier trop éloigné} ($F2 < 0$), associée à des oppositions entre périodes de travail. Avec plusieurs réglages, en effet, les appréciations ergonomiques se modifient de la 1-ère à la 4-ème d'une série de 4 périodes de travail consécutives: il semble notamment qu'une table proche de l'assise (i.e. basse), bien appréciée à la 1-ère période, paraisse trop basse à la 4-ème période.

3.3 Analyse des réponses des étudiants à un questionnaire relatif au mémoire de recherche de la maîtrise en administration des affaires, [QUEST. MÉM. RECH.], par B. McGibbon Taylor, P. Leduc, J.S. Tibeiro, in *CAD*, Vol. XIV, n°3, 1989.

Ensemble des 25 phrases retenues pour l'analyse, avec leurs sigles

- v4 Quand j'ai compris ce que représentait le mémoire de recherche, j'ai regretté ma décision de m'inscrire au MBA de Concordia.
- v5 Les objectifs du mémoire de recherche ne m'ont pas été clairement expliqués.
- V6 J'estime que les objectifs assignés au mémoire ont été atteints dans mon cas.
- V10 J'ai attendu la date limite pour m'engager avec un directeur de mémoire.
- ?11 J'avais en tête un sujet bien défini, quand je cherchais un directeur de mémoire.

On extrait d'un questionnaire un bloc de 25 questions, de même format, dont chacune propose une phrase vis-à-vis de laquelle l'enquêté manifeste son attitude par une note chiffrée variant de 1 à 5. Ces notes sont analysées après un recadrage tenant compte de l'équation personnelle, propre à chaque sujet. Simple et double recadrage ayant fourni des résultats concordants, on ne rendra compte ici que des résultats obtenus par simple recadrage.

Les phrases proposées aux sujets constituent des variables, notées Vx. Toutefois, pour aider à la lecture des graphiques, il a paru bon de remplacer éventuellement le 'V' majuscule par un autre caractère, le plus souvent un 'v' minuscule. En bref, les phrases favorables à l'institution du mémoire, ou en décrivant la bonne réalisation, conservent le 'V'; tandis que la minuscule 'v' symbolise un aspect négatif. Pour certaines phrases, compte tenu non seulement du sens obvie mais des résultats d'analyse, on a mis '?', voire '!'.

Il faut ici prendre garde que, dans le calcul de l'équation personnelle, l'estimation du zéro est biaisée si les phrases proposées vont toutes dans le même sens; la moyenne des notes attribuées par un sujet étant alors élevée ou basse selon que celui-ci approuve ou rejette la tendance générale; sans qu'il soit possible de savoir à quelle note, pour le sujet, s'arrête le rejet et commence l'approbation.

De ce point de vue, le présent questionnaire nous semble prêter à critique en ce que les phrases favorables au mémoire y semblent les plus nombreuses: on trouve, en effet, sur le tableau complet des phrases, 13 fois 'V', 8 fois 'v', 3 fois '?' et un '!'. De plus, de longues suites de phrases de même tendance (on a 9 'V' consécutifs...) peuvent assourdir l'attention du sujet et susciter des réponses stéréotypées.

Comme dans l'analyse du thermomètre grec, le nuage des modalités se présente, dans le plan (1,2), sous la forme en croissant caractéristique de l'effet Guttman. Dans le quadrant ($F1 > 0$, $F2 > 0$) sont les modalités d'opposition au mémoire: rejet de phrases favorables (sigles V-, selon nos notations) ou acceptation de phrases opposées (sigles v+). Dans le demi-plan ($F2 < 0$), avec des valeurs absolues modérées de F1, on trouve groupées les modalités '=' d'appréciation de toutes les phrases. Enfin, en achevant de suivre le croissant jusque dans le quadrant ($F1 < 0$, $F2 > 0$), on rencontre les modalités favorables à l'institution du mémoire ou à sa réalisation dans le cas du sujet interrogé: sigles V+ ou v-.

efficacité satisfaction
louange du directeur
heureux d'avoir appris
non scolaire, pratique

je n'ai rien appris!

directeur incompetent et inaccessible
déploire l'existence du mémoire
un gâchis de temps et d'argent
ne savait où aller et a été mené
ni contacts ni pratique

Ci-dessus, classes de modalités favorables;
à droite, classes de modalités défavorables.

On ne s'étonnera pas de trouver dans le plan (3,4) des aspects plus nuancés, plus difficiles à exprimer en un seul mot que le facteur de niveau général qui occupe le plan (1,2): mais on remarque des classes de modalités, dont la CAH permet de proposer une interprétation générale. Renvoyant à l'article original pour les détails des phrases, nous reproduisons l'interprétation des subdivisions retenues, d'une part, pour la classe des modalités favorables; et d'autre part, pour les modalités défavorables.

Quant aux nuances, il apparaît, e.g., que l'approbation de la phrase '?11':

J'avais en tête un sujet bien défini, quand je cherchais un directeur de mémoire,

s'écarte peu de l'origine sur l'axe 1 mais se place cependant dans une classe de modalités favorables qu'on a étiquetée: non scolaire, pratique.

3.4 Essai d'analyse des notes attribuées par un ensemble de sujets aux mots d'une liste, [NOTES MOTS], par J.-P. Benzécri, in CAD, Vol. XIV, n°1, 1989.

Monsieur Émeric Deutsch, directeur de SOFRES Communication, a entrepris, selon le projet de Jean-François Steiner, de fonder les analyses d'attitudes et opinions qui lui sont demandées, sur une typologie sociométrique globale établie d'après les notes données par des sujets aux mots d'une liste.

Quand, dans la même liste, on trouve Dieu, neige, séduire et venger, il peut sembler sacrilège de demander au sujet de se prêter à la consigne de noter sans réfléchir, de -3 à +3, les symboles de tant et de telles réalités. Le fait est que 90% des sujets le font d'assez bonne grâce; et que l'analyse nous convainc *a posteriori* que leur comportement n'est pas dépourvu de cohérence; qu'il manifeste, au contraire, des attitudes qui nous sont familières.

L'article est fondé sur un tableau de notes croisant un ensemble S de 1140 sujets avec un ensemble M de 286 mots. Comme les données appartiennent à la SÔFRES, et que la méthode, et la liste même des mots, sont la propriété de la Société SÉMIOMÉTRICA, on ne pouvait songer à une publication sans bride. On se bornera ici à évoquer l'enchaînement d'analyses et de codages où l'équation personnelle joue son rôle.

Le premier tableau analysé croise l'ensemble M des mots avec l'ensemble N des 7 notes (de -3 à +3), avec $k(m, n)$ = nombre des sujets de S ayant attribué au mot m la note n . L'analyse de ce tableau $M \times N$ fournit un résultat familier. Dans le plan (1,2), les notes s'ordonnent de -3 à +3, suggérant un arc de parabole, suivant le classique effet Guttman. Les sigles des mots dessinent un croissant, beaucoup plus dense du côté positif que du côté négatif. Aux deux extrémités se lisent clairement les deux sigles d'une antithèse: au pôle négatif GRRR, au pôle positif, PAIX.

Pour naturel qu'il soit, cet étalement des mots, (depuis ceux généralement mal notés jusqu'à ceux le mieux notés), est cause d'embarras. Car d'une part, il apparaît que deux mots, dont la valeur moyenne diffère sur le marché des opinions, ne seront pas l'objet de comportements directement comparables; d'autre part, il est certain *a priori* qu'une corrélation peut exister entre des mots qui, tout en étant distants sur l'échelle moyenne, ont en commun d'être relativement bien notés ou mal notés, simultanément, par des sujets appartenant à la même tendance.

Dans le croissant des mots, la classification automatique permet seulement de découper une ou deux classes d'environ cent mots, au sein de chacune desquelles les profils sur l'ensemble des notes soient assez voisins pour qu'on puisse comparer directement les perceptions des sujets, sans avoir à compenser, par un calcul plus ou moins arbitraire, la distance entre leurs niveaux moyens. C'est ainsi qu'on a retenu une classe M' de 93 mots.

Avant tout essai de codage, montrons le parti qu'on peut tirer de l'analyse de données univoques: nous dénombrerons, pour chaque paire de mots, les sujets ayant donné aux deux mots la même note. Nous dirons qu'il y a coïncidence de note entre deux mots, m et m' , quand un sujet s donne à m et m' la même note, quelle que soit celle-ci. D'où un tableau de correspondance $M \times M$, 286×286 , où $k(m, m')$ est le nombre des sujets ayant donné à m et m' des notes qui coïncident.

Selon cette définition, on doit inscrire dans toute case de la diagonale de ce tableau le nombre 1140, effectif de l'échantillon des sujets considérés. En effet, dans la mesure où chacun des sujets n'est interrogé qu'une fois, l'unique note donnée à tout mot m coïncide avec elle-même! Dans l'analyse d'un tel tableau où prédominent les termes diagonaux, chaque mot tend à créer par lui-même un facteur: on a donc corrigé la diagonale, en donnant pour valeur à $k(m, m)$ le maximum atteint par $k(m, m')$ pour $m' \neq m$.

Disons d'abord qu'à partir du sous-tableau $M' \times M'$ recensant les coïncidences de notes pour 93 mots (dont les niveaux d'appréciation sont voisins: cf. *supra*), on obtient des facteurs et des classes dont l'interprétation se fait aisément, non en terme de niveau, de pouvoir moyen d'attraction ou de répulsion des mots, mais selon leurs sens et les attitudes qu'ils évoquent.

Éliminant seulement quelques mots qui, tels GRRR, sont l'objet d'une attitude quasi unanime, ce qui empêche de les situer, quant au sens, par rapport au reste du vocabulaire, nous avons retenu 262 mots. Dans l'analyse (quasi globale) du tableau de correspondance 262×262 (recensant les coïncidences de notes avec la diagonale écrasée), l'influence du niveau général des mots est très forte, mais elle semble ne marquer que les deux premiers facteurs, où le classique nuage en croissant parabolique signe un net effet Guttman.

Si on affiche à l'écran divers plans croisant deux axes de rang supérieur à 2, on voit se détacher plusieurs îlots ou groupes de mots nettement interprétables quant au sens. Bien que les facteurs de rang 1 et 2 apportent plus de 90% de l'inertie, il apparaît légitime d'effectuer une classification des mots d'après les facteurs de rang supérieur à 2.

En termes géométriques, ceci revient à projeter le nuage des mots sur le supplémentaire orthogonal au plan (1,2). Une analogie simple est celle d'un objet tridimensionnel transparent allongé que l'on regarde longitudinalement; ou encore, en projection sur un plan perpendiculaire à l'axe. Ainsi, toutes les coupes transversales se superposent, et d'éventuels filets longitudinaux sont vus comme une petite tache. Ce procédé d'observation n'est toutefois pas d'une sûreté parfaite; car il se peut que l'objet allongé sinue autour de son axe.

On conçoit donc que les résultats de la classification des 262 mots, très satisfaisants dans leur ensemble, offrent matière à critique. Afin de ne conserver que des classes bien interprétables, nous avons examiné l'arbre de la CAH générale; pour retenir finalement une partition des mots en 44 classes. Au sein de chaque classe, la signification de tous les mots est analogue, en ce sens que même si le niveau d'appréciation moyen qu'il reçoivent est inégal, ce sont généralement les mêmes sujets qui les notent le mieux, ou le plus mal. Après codage des réponses des individus suivant l'équation personnelle, cette partition a fourni une typologie très suggestive.

Dans le cas présent, appliquer le codage sur trois modalités à 262 variables-mot, implique de créer un tableau à 786 colonnes dont nous ne pourrions rien faire. On crée d'abord un tableau à 262 colonnes des notes recadrées (ligne par ligne) suivant l'équation personnelle; puis, à partir de celui-ci, un tableau à 132 colonnes, donnant, pour chacune des 44 classes q de la partition retenue, les cumuls $q+$, $q=$ et $q-$ des notes $j+$, $j=$ et $j-$ afférentes aux mots compris dans cette classe.

Finalement, on a analysé un tableau 1140×132 , croisant l'ensemble S des sujets avec 132 modalités $\{q-, q=, q+\}$ qui expriment respectivement le rejet, l'indifférence ou l'approbation vis-à-vis non de mots isolés, mais de classes de mots. Il nous suffira de citer ici l'interprétation proposée pour les facteurs: à l'exception du 1-er, qui oppose les modalités $q=$, ($F1 > 0$), aux $\{q+, q-\}$, ($F1 < 0$), tous les facteurs ont une valeur morale ou idéologique.

($F1 < 0$) : certainement ;	($F1 > 0$) : peut-être ;
($F2 < 0$) : contestation ;	($F2 > 0$) : religion et morale ;
($F3 < 0$) : centre ;	($F3 > 0$) : conjonction ascétique ;
($F4 < 0$) : les damnés de la terre ;	($F4 > 0$) : le jardin d'Épicure ;
($F5 < 0$) : agressivité ;	($F5 > 0$) : refus de la lutte ;
($F6 < 0$) : idéal monarchique.	

Le signalement, par âge-sexe et attitude de vote déclarée (comme dans l'analyse du thermomètre grec), a confirmé cette interprétation; encourageant le projet d'une sociométrie globale, fondée sur les notes données à des mots.

3.5 Analyse des réponses d'un millier d'algériens à un questionnaire sur l'avenir politique de l'Algérie, [QUEST. ALGÉRIE], par M. Zerroug, in *CAD*, Vol. XVI, n°3, 1991.

Dans un pays qui depuis son indépendance et jusqu'à une date récente n'a connu que le régime du Parti Unique, il n'y a pas de tradition établie des sondages d'opinion; pas d'enquêteurs professionnels, pas d'échantillons constitués, considérés comme "représentatifs" et régulièrement interrogés sur les problèmes les plus divers. Mais, en revanche, les idées ont toujours très bien couru dans le tissu social algérien, et le proverbial "Téléphone Arabe" a une très large bande passante...

À l'approche des élections législatives, dont le 1-er tour était fixé au 27 Juin 1991, la Société TRANSMÉDIA, d'Alger, a tenté une première: proposer aux sujets, non de simples déclarations d'intension de vote, mais un ensemble de phrases vis-à-vis desquelles chacun puisse, en exprimant son attitude d'approbation ou de réprobation plus ou moins forte, composer un projet personnel pour l'Avenir de l'Algérie.

Quand fut conçu le questionnaire, les élections municipales avaient mis en mouvement toutes les familles de pensées; on annonçait fusions et scissions:

ceci interdisait toute référence précise à des initiales symboliques, atomes crochus en permanente recombinaison... C'est pourquoi, évitant les slogans agressifs, on avait voulu, en termes aussi généraux que possible, pointer dans toutes les directions que les algériens, en quête d'avenir, ont depuis longtemps explorées du regard.

Nous ne reproduirons pas ici le questionnaire en entier, ni son analyse, publiée récemment, dans l'article cité, telle qu'elle avait été rédigée avant que les élections ne fussent différées. Quant aux résultats, nous dirons seulement qu'avec deux facteurs principaux, le premier de disponibilité idéologique, le second d'intégration sociale, l'analyse suggérait "un projet de ralliement associé aux cadres supérieurs", projet qu'à l'automne de 1991, le gouvernement de l'Algérie peut sembler avoir fait sien. Mais nous insisterons sur la conception du questionnaire et sa présentation, autant que sur la méthode d'analyse elle-même.

Les 22 phrases soumises à l'appréciation des sujets sont choisies pour faire allusion à la quasi-totalité des thèmes débattus: économie, héritage des luttes pour l'indépendance, relations avec l'Occident et avec les pays arabes, enseignement, religion, socialisme, conflits de générations. Mais, afin d'éviter toute agressivité, le questionnaire a été livré aux critiques de lecteurs sensibles. L'ordre même des phrases ménage une entrée et une sortie acceptable pour tous, et jamais rien de brûlant n'est avancé sans le voisinage de quelque chose de tiède. L'expérience a montré que circonspection et courtoisie ne nuisent aucunement à l'acquisition de l'information.

les vacances doivent être passées en famille, à la maison, pour bien se reposer	tout à fait tout à fait opposé x d'accord
---	---

Le questionnaire, préparé en français et en arabe, est présenté sur trois pages; dont la première explique (sur un exemple rappelé ci-dessus), l'usage d'une échelle analogique; et les deux suivantes présentent les 22 phrases, avec une échelle en face de chacune d'elles. En laissant aux sujets plus de liberté, l'échelle offre l'occasion d'échapper aux réflexes de partis pour formuler une opinion nuancée; et le recadrage des notes suivant l'équation personnelle permet de s'affranchir des modes d'expression propres à chaque sujet.

L'expression analogique est d'autre part à la portée de la plupart des sujets dont on peut attendre qu'ils répondent à un questionnaire politique. Dans une pré-enquête, organisée par M. MAAÏTI auprès d'une centaine de sujets seulement, les échelles ont été correctement utilisées. Dans le sondage étendu, il semble que la consigne n'ait pas été soutenue par un commentaire verbal des enquêteurs, mais qu'au contraire ceux-ci aient suggéré aux sujets d'interpréter l'échelle comme une question à trois modalités: 'Oui' à droite, 'Non' à gauche, 'Ne sait pas' au milieu; ce qui amplifie les tendances individuelles à une stéréotypie des réponses: marquées, selon la personne, soit par la réserve, soit par la force de l'assentiment ou du refus.

Dans ces conditions, le recadrage des notes suivant l'équation personnelle n'a pas apporté de clarification à l'analyse: les résultats obtenus après recadrage ont même structure que ceux extraits des données codées en (0, 1), avec, pour chaque phrase 'Phr' 3 modalités {Phr<, Phr?, Phr>}. Or (suivant une tendance qui n'est pas indépendante de leur place dans la société), certains sujets ont une prédilection pour les modalités évasives; d'autres tendent à formuler toujours une préférence: les uns multiplient les approbations et les autres les rejets; alors que l'on s'est appliqué à construire un questionnaire équilibré, où la diversité des opinions était, en quelque sorte, présentée suivant ses deux faces; témoins ces deux phrases, la 12-ème et la 19-ème, relatives au socialisme:

12 Le socialisme a échoué même en URSS et il n'a rien apporté de valable à l'Algérie.

19 Pour les pays du tiers monde, l'idéal socialiste reste valable: l'économie libérale ne suffit pas aux aspirations du peuple.

Mais un nouveau système de codage a apporté des résultats satisfaisants. On a fait en sorte que, dans la ligne de chaque sujet, prennent même poids total les trois groupes de 22 modalités, {Phr<}, {Phr?} et {Phr>}. En termes imagés, nous dirons que, quelles qu'aient été les réponses exprimées par un sujet, on alloue à celui-ci un capital déterminé de rejet, de réserve et d'approbation. En termes mathématiques, le but sera atteint si l'on a, en notant Q l'ensemble des 22 phrases, et $kij(i, q_\epsilon)$ la note recodée de l'intensité de la modalité q_ϵ de la phrase q , pour le sujet i :

$$\forall i \in I, \forall \epsilon \in \{<, ?, >\} : \sum \{ kij(i, q_\epsilon) \mid q \in Q \} = \text{card}Q = 22 ;$$

Cette relation sera satisfaite si, e.g., à supposer qu'un sujet i ait approuvé 4 des 22 phrases proposées, $\{q_1, q_2, q_3, q_4\}$, on met à la valeur 22/4 les quatre notes correspondantes $kij(i, q_1>), \dots, kij(i, q_4>)$, les autres notes $k(i, q>)$ restant à zéro; et de même pour les signes $\{?, <\}$. En bref, (à d'importants détails près qu'on ne reprendra pas ici), au lieu de rester sous le format en (0, 1), on donne au nombre $kij(i, q_\epsilon)$ correspondant à une modalité q_ϵ adoptée par le sujet i , une valeur d'autant plus forte que les modalités de signe ϵ adoptées par i sont moins nombreuses.

Une telle transformation supprimerait un facteur tel que le premier issu de l'analyse des notes attribuées par un ensemble de sujets à des mots (cf. §3.4, *in fine*):

(F1<0) : certainement ;

(F1>0) : peut-être ;

dans la CAH, elle favorise l'agrégation de modalités moyennes avec d'autres modalités; et sans doute faudra-t-il, dans l'avenir en expérimenter l'usage, même là où le recadrage suivant l'équation personnelle suffit à révéler des résultats interprétables, au travers des stéréotypes de réponse des sujets. (Le recadrage, en effet, ne suffit pas à compenser la tendance de certains sujets à affectionner

les notes moyennes, tout en utilisant exceptionnellement les extrémités de l'échelle; or c'est précisément cette tendance qui, même après recadrage, favorise l'agrégation des modalités moyennes en une seule classe.)

4 Les échelles dans les études cliniques et thérapeutiques

Dans ce §4, nous abordons le thème proposé par le Dr. BOSC. Nous rendrons compte d'abord, en détail, (au §4.1), d'une étude, à laquelle le Dr. BOSC lui-même a participé, et qui, fondée sur un vaste corpus de données, offre des exemples de traitements variés, sans toutefois recourir au recadrage. Au §4.2, nous parcourons diverses études, généralement beaucoup moins étendues que la première, mais qui suggèrent des réflexions que nous croyons utiles pour les études diachroniques. Nous terminons en considérant, au §4.3, une analyse qui, certes fondée sur un petit nombre de protocoles, offre un exemple de ce que le recadrage peut apporter à la critique de la validité d'une échelle.

4.1 Échelles d'appréciation de l'anxiété et évaluation des plaintes somatiques: analyse de protocoles recueillis dans des essais thérapeutiques, [ÉCHELLES], par J.-P. Benzécri, M. Bosc, G. D. Maïti, in CAD, Vol. XIV, n°3, 1989.

4.1.1 Nature et origine des données analysées

Nous décrivons le format des données, et les essais d'où elles proviennent.

Afin de prendre en compte les divers aspects de l'anxiété, a été conçue l'échelle de COVI, qui évalue l'anxiété par le total de trois notes (de 0 à 4) :

Discours du sujet: nerveux, n'est pas dans son assiette, agité, effrayé sans raison, etc...

Comportement: semble effrayé, mal à l'aise, agité, angoissé;

Plaintes somatiques: sudation, tremblements, sensation de striction cardiaque, tachycardies, oppression respiratoire, sommeil agité, etc...

Un problème connexe, mais plus complexe en ce que l'anxiété n'en est qu'une composante, est celui de la pathologie dépressive. Dans ce domaine, le diagramme H.A.R.D., de Rufin et Ferreri (1984), permet de caractériser l'état d'un patient par un système de 12 notes (de 0 à 6) rangées sous 4 titres:

Humeur: Tristesse, Désintérêt, perte d'Appétit.

Angoisse: Anxiété psychique, Anxiété somatique, Insomnie.

Ralentissement: Lassitude, difficulté de Concentration, Attitude.

Danger: Dévalorisation, idées Suicidaires, idées Délirantes.

L'intérêt propre à ce système est qu'il ne sert pas seulement à calculer une note globale, mais offre, de l'état dépressif, un profil dont les auteurs s'entendent à apprécier l'évolution en cours de traitement.

Les plaintes somatiques ont été évoquée avec l'échelle de COVI. On peut ranger sous ce titre des symptômes pathologiques nets, des troubles psychosomatiques difficiles à cerner et des plaintes dont l'objectivité est douteuse. Pour éviter d'étiqueter l'ensemble des plaintes d'un patient selon un jugement hypothétique peut-être hâtivement formé, on utilise communément un inventaire systématique des plaintes suivant une trentaine d'items, au nombre desquels nous citerons:

Diminution d'appétit; Gastralgies; Constipation, Diarrhées, Nausées, Vomissements, Troubles de la Vue, Sécheresse de la bouche, Transpiration, Fourmillements des membres, Sensations de vertige, etc...

On considère deux essais, relatifs l'un à la dépression de l'adulte et l'autre aux troubles du sommeil, et entrepris afin de préciser l'efficacité d'un antidépresseur sédatif.

Le premier essai concerne 87 patients des deux sexes, âgés de 18 à 80 ans, ayant, initialement, dans l'échelle HARD (présentée ci-dessus) une note totale supérieure à 18; mais dont l'état n'apparaît pas requérir une hospitalisation. L'état des patients est évalué au cours d'un premier examen (jour J0); puis, les patients étant traités par le produit, sont examinés après 1, 2 et 4 semaines (J7, J14 et J28). Chacun des 4 examens (J0, J7, J14, J28) comporte notamment une évaluation du diagramme HARD. Au cours de l'examen initial, on fait de plus un inventaire systématique des plaintes somatiques suivant 32 items, pour chacun desquels est donnée une note de 0 à 4.

Le deuxième essai concerne 263 patients adultes des deux sexes se plaignant de troubles du sommeil dans un contexte d'anxiété ou au cours d'un état dépressif. Il y a, en principe, 3 examens par patient, à J0, J14 et J28. L'examen initial comporte un inventaire des plaintes somatiques suivant le même protocole que dans le premier essai. Il y a, à chaque examen (J0, J14 et J28), évaluation du diagramme de HARD ainsi que des 3 notes d'anxiété de l'échelle de COVI.

Aux données manquantes près, nous avons donc 350 inventaires de plaintes somatiques provenant de chacun des patients de l'un ou l'autre essai; et 1137 diagrammes HARD (rangés par groupes de 3 ou de 4 afférents à un même patient); de ces diagrammes HARD, la plupart (issus du deuxième essai relatif aux troubles du sommeil) sont associés à une évaluation de l'échelle de COVI.

4.1.2 Codages des notes et structure des échelles

Partons de l'exemple du diagramme HARD. Ses 12 notes, prises telles quelles, définissent un profil de dépression, plus ou moins chargé sur l'une ou l'autre des composantes. En soumettant à l'analyse des correspondances le tableau brut 1137×12 , on découvrira entre les notes des proximités, plus ou moins conformes à la quadridimensionnalité postulée par les auteurs; et l'on

pourra considérer, indépendamment de son intensité globale, l'orientation de la dépression des divers sujets (e.g. plutôt vers le danger ou plutôt vers le ralentissement), avec son évolution au cours du traitement. Mais, pour reconnaître une dimension générale d'intensité, on usera d'un codage.

Une note de l'échelle de COVI, du diagramme HARD, de l'inventaire des plaintes somatiques, apparaît comme une question fermée (cf. §2.1) dont le nombre de modalités est respectivement 5 ou 7 selon que la note peut varier de 0 à 4 ou de 0 à 6. On a donc pu utiliser le codage disjonctif complet (en 0, 1).

Mais ce codage présente ici deux inconvénients: d'une part, surtout pour le diagramme de HARD, il multiplie les modalités; d'autre part, le niveau moyen de gravité différant beaucoup d'un trait à un autre, on a, pour diverses notes, certaines modalités extrêmes, soit fortes soit faibles, qui se rencontrent très rarement et perturbent l'analyse. C'est pourquoi on a eu recours au codage linéaire par morceaux (cf. §2.2), en créant pour chaque variable trois modalités {-, =, +}.

Qu'on utilise le codage disjonctif complet ou le codage sur 3 modalités, l'analyse du diagramme HARD fournit d'abord un premier facteur de niveau général, étroitement corrélé au total des notes, calculé d'ordinaire pour évaluer l'état dépressif; et le plan (1, 2) montre un net effet Guttman (cf. *supra*, §§3.1 & 3.4). Mais seul le codage suivant 3 modalités aboutit, par la CAH, à une typologie claire, où l'on reconnaît, (dans les subdivisions de la classe des modalités '+'), les quatre titres du HARD, à deux importantes corrections près: la *Dévalorisation*, trait de gravité modérée, rejoint la *Tristesse* et le *Désintérêt*, sous le titre 'Humeur' dont se sépare la *diminution d'Appétit*; ce dernier trait est isolé, parce qu'il n'est pas constant dans la dépression, qui peut engendrer la boulimie.

L'analyse des notes brutes a offert l'occasion de considérer ensemble, dans leur aspect multidimensionnel, l'échelle de COVI et le diagramme HARD. On considère 683 examens effectués aux jours J0, J14 ou J28, dans l'essai relatif à la qualité du sommeil. Le tableau analysé comprend 15 colonnes principales qui sont d'une part les 12 notes du diagramme HARD et d'autre part les 3 composantes de l'échelle de COVI; avec, en colonnes supplémentaires, les totaux Htot et Ctot des notes du diagramme HARD et de l'échelle de COVI.

Tandis que l'axe 2 est créé par la singularité, déjà notée de la note d'Appétit, on a dans le plan (1,3) un schéma tripolaire assez satisfaisant: {Danger (F1<0; F2>0), Angoisse (F1>0; F2>0), Ralentissement (F2<0)}. Toutes les composantes de l'échelle de COVI sont du côté positif de l'axe 1; et c'est aussi suivant cet axe que s'opposent, globalement, les totaux Htot et Ctot: car le diagramme HARD prend en compte un aspect de Danger (F1<0) que l'échelle de COVI ne distingue pas.

La classification ascendante hiérarchique précise les rapports entre HARD et COVI. On ne s'étonnera pas de voir les *Plaintes somatiques* de COVI s'agréger avec l'*anxiété somatique* de HARD). Les deux autres notes de l'échelle de COVI, rendant compte du *Discours du sujet* et de son *Comportement*, sont très proches entre elles et s'agrègent aux items de *Lassitude* et d'*Anxiété Psychique* du diagramme HARD.

Pour l'analyse des plaintes somatiques, comme pour le diagramme HARD, le codage suivant 3 modalités est préférable. On a un axe 1 de niveau général; et le plan (1, 2) présente un effet Guttman, moins net toutefois qu'avec le diagramme HARD. Une typologie des plaintes somatiques peut être obtenue soit d'après l'analyse du tableau codé, soit, directement, par l'analyse du tableau brut des notes. En montrant des liaisons entre symptômes indépendantes de l'organe source, cette typologie invite à chercher une réponse thérapeutique plus nuancée que celle possible aujourd'hui.

4.1.3 Typologie des sujets et variation de la pathologie dépressive au cours des essais thérapeutiques

Compte tenu de ce que les deux essais, relatifs à la dépression de l'adulte et aux troubles du sommeil, comportent respectivement des examens à {J0, J7, J14, J28} et {J0, J14, J28}, les diagrammes de HARD sont répartis en 7 groupes désignées par les sigles {DEPa, DEPb, DEPc, DEPe, somA, somC, somE}. Il sera commode d'appeler en bref ces groupes des "étapes des essais".

À tout tableau croisant l'ensemble des 1137 diagrammes avec un ensemble de modalités, on peut adjoindre 7 lignes, calculées par cumul des lignes rentrant dans chacune des "étapes des essais". Dans l'analyse, on peut mettre en principal soit les 1137 diagrammes, soit les 7 étapes. D'une manière ou d'une autre, il apparaît manifeste que la dépression régresse au cours des essais.

Une autre voie a été explorée: d'après une analyse factorielle (des notes codées ou brutes), effectuer une classification des 1137 diagrammes, de laquelle on retient une partition, e.g. en 20 classes; puis construire le tableau de contingence ventilant chacune de ces classes suivant les 7 étapes des essais. Dans le cas des notes codées, l'analyse de ce tableau 20×7 donne un premier axe de gravité, sur lequel les étapes des deux essais se rangent, dans leur ordre naturel, du début (gravité maxima) à la fin (gravité minima). Mais dans le plan (2, 3), les étapes des deux essais se séparent, parce qu'il s'agit de pathologies différentes.

Voici enfin comment on a procédé pour éprouver la stabilité du profil d'anxiété dépressive au cours du traitement. D'après l'analyse simultanée des notes brutes de l'échelle de COVI et du diagramme HARD (cf. *supra* §4.1.2), on a réparti les 683 examens en 13 classes, homogènes quant au type de dépression mais hétérogènes quant au niveau. Puis on a construit un tableau de cooccurrence 13×13 , donnant, en bref, pour chaque couple de classes (c, c') le

nombre $k(c, c')$ de sujets ayant présenté, à des dates différentes, un examen dans la classe c et un dans c' ; (il faut seulement veiller à compter judicieusement les sujets présentant plusieurs examens dans l'une des deux classes...).

L'analyse de la matrice de cooccurrence retrouve, approximativement, la disposition relative qu'ont les classes dans la hiérarchie issue de l'analyse du tableau des données initiales; d'où l'on conclut que les profils pathologiques des sujets sont relativement stables au cours de nos essais, où pourtant le niveau de gravité diminue fortement dans la plupart des cas; et que les changements de classes se font de façon cohérente, c'est-à-dire principalement entre classes voisines.

4.2 Étude diachronique et données psychométriques

4.2.1 Comparaison entre les états de sujets après l'effort sous divers traitements, [COMPAR. EFFORT], par A. Youbi, in *CAD*, Vol. XIV, n°4, 1989.

Sans pouvoir divulguer les résultats qui lui ont été confiés, l'auteur de la présente étude montre, sur un exemple, que l'analyse de tableaux cumulés croisant traitements et modalités descriptives est une méthode d'analyse logique à la fois simple et efficace. Cette étude nous intéresse parce qu'elle analyse des différences entre valeurs successives de variables psychométriques.

Huit sujets sains réalisent chacun, successivement, trois épreuves d'effort, E0, E1, E2. Les épreuves sont les mêmes pour tous. L'épreuve E0 est réalisée dans des conditions usuelles et hors de tout traitement. Les épreuves E1 et E2 sont toutes deux réalisées dans des conditions éprouvantes et sous traitement. Il y a deux traitements possibles: A et B; chaque sujet est soumis à l'un des traitements, A ou B, lors de l'épreuve E1; et à l'autre traitement, lors de l'épreuve E2. Après chaque épreuve, un sujet est soumis à des mesures psychométriques et physiologiques.

Les mesures psychométriques sont résumées par les quatre notes d'un diagramme HASE:

{HUMeur, ANXiété, EFFort, SOMmeil};

chacune des quatre notes peut varier de 0 à 15; et elle est d'autant plus élevée que l'état du sujet est plus perturbé dans le domaine concerné.

Quant aux variables physiologiques, nous nous bornerons à les désigner par les sigles SSS et VVV, en précisant que l'état d'un sujet est d'autant meilleur que ces variables, cotée chacune, par le laboratoire, de 0 à 5, sont plus basses.

Dans l'esprit des expérimentateurs, l'épreuve E0, réalisée hors traitement, constitue une référence à laquelle comparer les deux autres épreuves E1 et E2, réalisées sous traitement. Les valeurs brutes des six variables mesurées après

une épreuve E1 ou E2, ne sont pas à interpréter en elles-mêmes mais relativement aux valeurs mesurées après E0 chez le même sujet. L'auteur s'est borné à opérer sur des différences de valeurs, sans exclure la possibilité que, dans d'autres études, il soit meilleur de calculer des rapports.

L'individu statistique est un couple (sujet, épreuve sous traitement). On attribue à chaque couple un sigle de trois caractères, tel que Ab1: le chiffre qui termine ce sigle est le numéro du sujet, lequel peut varier de 1 à 8. Les deux lettres désignent les traitements auquel ce sujet a été soumis: soit, dans le cas du sujet 1, A pour l'épreuve E1 et B pour l'épreuve E2. Si l'on considère l'épreuve E1, on met en capitale la première lettre et en minuscule la seconde (comme dans Ab1); et on fait le contraire pour la deuxième épreuve (aB1).

Ceci posé, on construit un tableau ayant 16 lignes, les 16 épreuves sous traitement (2 par sujet) et 6 colonnes, les 6 variables (psychométriques ou physiologiques). À l'intersection de la ligne Ab1 et de la colonne EFF, on place la valeur -2, obtenue en retranchant la valeur de la note EFF, attribuée au sujet 1 après l'épreuve E0, de la valeur de EFF pour 1 après E1. De même, à l'intersection de la ligne bA5 et de la colonne HUM, on place la valeur 4, obtenue en retranchant la valeur de la note HUM, attribuée au sujet 5 après l'épreuve E0, de la valeur de HUM pour 5 après E2. Etc...

Les 6 variables sont découpées chacune en trois modalités; soit, au total 18 modalités, qu'il est commode d'affecter de l'un des signes {+, =, -} selon que le comportement relatif dans l'épreuve sous traitement a été bon, moyen ou mauvais. Il importe de noter que "relativement bon" ne veut pas dire "meilleur qu'à l'état basal", mais seulement que la variation, même s'il s'agit d'une détérioration (comme c'est généralement le cas pour l'HUMeur), est parmi les plus favorables enregistrées dans les 16 épreuves sous traitement. À ces 18 modalités, on en adjoint 4, {Ab, Ba, aB, bA} qui spécifient le rang de l'épreuve et le traitement. D'où, finalement, un tableau 16×22 , sous forme disjonctive complète, en (0, 1) (cf. §2.1).

D'après le tableau 16×22 , on construit un tableau de cooccurrence des modalités, ou tableau de BURT, 22×22 ; dont on analyse deux sous-tableaux: 18×18 (croisant les modalités des 6 mesures avec elles-mêmes); et 18×4 , (croisant les modalités des 6 mesures avec {Ab, Ba, aB, bA}); les épreuves individuelles étant adjointes, en supplément, à l'une et l'autre analyse. L'axe 1 est, chaque fois, un axe de niveau général, avec dans le plan (1, 2), effet Guttman.

Sur l'axe 1, les points Ab et Ba, figurant l'épreuve E1 sous les deux traitements possibles, sont proches de l'origine; et l'axe est créé par l'opposition entre aB et bA, issues de l'épreuve E2 sous les deux traitements. Si, comme le postulaient les expérimentateurs, l'influence du traitement sous lequel a été effectuée l'épreuve E1 ne s'étend pas à l'épreuve E2, il faut conclure que ce sont

les sujets eux-mêmes qui ne deviennent sensibles à la différence entre traitements qu'après avoir effectué E1. Associé aux modalités favorables (de signe '+'), le point aB atteste la supériorité du traitement B.

Fondée sur des principes de codage désormais classiques, la construction et l'analyse des tableaux cumulés, croisant traitements et modalités descriptives des effets, apparaît ici comme une méthode d'analyse logique simple et efficace, apte à traiter conjointement des données psychométriques et physiologiques.

4.2.2 Comparaison entre les effets d'un produit et d'un placebo chez les insuffisants cardiaques évolués résistants au traitement conventionnel, [COMPAR. CARD.], par J.-P. Benzécri, G.D. Maïti, A. Tarral, O. Dubourg, in CAD, Vol. XIV, n°4, 1989.

Le but est double: d'une part, apprécier le bénéfice que de les patients peuvent retirer d'un produit; et d'autre part, déterminer la fiabilité des méthodes non invasives, seules pratiquées ici, pour évaluer la performance circulatoire. L'effet favorable du produit sur la performance circulatoire est établi ainsi que sa supériorité sur le placebo pour assurer la survie des patients; et la cohérence des résultats atteste la fiabilité des mesures. Mais seule nous concerne ici l'étude de la qualité de vie des patients.

Après qu'on a fait l'inventaire des données manquantes, il reste 86 protocoles de réponses complets, avec 5 notes (de 0 à 3) relatives à la qualité de vie: {activité Physique; Sommeil; Appétit; Énergie; Respiration}; (cette dernière question étant destinée à sonder l'état circulatoire des patients, selon la manière dont elle est ordinairement interprétée par ceux-ci). Le tableau de base 86×5 est codé sous forme disjonctive complète.

L'analyse fournit deux premiers facteurs nettement séparés des suivants, et liés entre eux par le classique effet Guttman: dans le plan (1, 2) les modalités s'ordonnent de faible (mauvais) à fort (bon) en dessinant un croissant parabolique. L'interprétation est donc claire; et il ne reste qu'à observer la place des protocoles successifs des individus relativement à ce croissant.

Dans l'ensemble, les cas d'amélioration subjective sont rares. En général, sous produit ou sous placebo, les patients, quel que soit leur état circulatoire ou leur pronostic vital, donnent, au fil du traitement, des réponses de moins en moins bonnes pour la qualité de vie. Mais le lien entre plainte subjective et qualité de vie n'est pas univoque: il évolue au cours du traitement, ce qui pose un difficile problème de psychométrie et interdit toute conclusion hâtive.

4.2.3 Appréciation comparée des variations, sous deux traitements, de la gêne éprouvée par des patients, [COMP. VAR.], par J. Hayek, A. Loemba, in CAD, Vol. XVI, n°1, 1991.

Il s'agit de plusieurs dizaines de patients, dont on ne précisera pas le mal.

absence
de gène ————— gène
maxima

Entre autres données, les observations médicales comportent une appréciation de la gène (dont le patient est chroniquement affecté), fournie quotidiennement du jour J00, d'entrée dans le traitement, au jour J14. Cette appréciation est portée sur une ligne horizontale, suivant le format usuel des échelles. L'échelle ayant une longueur de 100 mm, l'appréciation est notée sur le protocole d'examen comme un nombre variant de 0 à 100.

Comme il est bien connu, les ordres de grandeur des abscisses sur l'axe diffèrent grandement d'un sujet à l'autre; sans qu'on puisse d'après de telles différences apprécier le niveau de la gène éprouvée par ces sujets. En revanche, il semble légitime d'apprécier la suite des notes afférentes à chacun des sujets comme un profil de gène. On a donc analysé un tableau (49×15) croisant l'ensemble I des 49 sujets, pour lesquels on dispose de notes complètes, avec l'ensemble J des 15 notes de gène désignées par les sigles de Gn00 à Gn14.

L'axe 1, nettement détaché des suivants quant au taux d'inertie, s'interprète clairement: la suite des modalités de gène est en effet, presque sans permutation, dans l'ordre des jours de Gn00 à Gn14. Un sujet situé du côté ($F1 < 0$), associé aux modalités des premiers jours de traitement, a un profil de gène dont la décroissance au cours du temps est plus rapide que la moyenne; au contraire, on trouve avec ($F1 > 0$) les sujets pour lesquels le poids de la gène en fin de traitement est relativement élevé: même s'il y a eu une décroissance (ce qui est général), elle est inférieure à la moyenne.

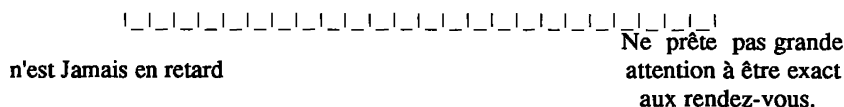
Une épreuve de rang par simulation suggère qu'il n'y a pas de différence significative entre les profils de gène des sujets soumis aux deux traitements.

On a considéré les profils de gène: la présentation verbale de l'échelle, autant que son orientation naturelle de la gauche vers la droite suggère, en effet, que l'attention des sujets portait sur le mal, à l'exclusion de son complémentaire: le bien-être! On a toutefois considéré un tableau à 30 colonnes, créé en adjoignant à chaque colonne Gn_j une colonne Bn_j donnant son complément à 100; en d'autres termes, la distance entre le point marqué, par le sujet, sur l'échelle et l'extrémité droite de celle-ci.

L'analyse du tableau à 30 colonnes montre d'abord, sur l'axe 1, une opposition entre l'ensemble des Gn_j ($F1 < 0$) et celui des Bn_j ($F1 > 0$): du côté ($F1 < 0$) sont donc les sujets qui ont exprimé leur gène par des valeurs élevées, ce qui, selon nous, ne signifie aucunement que celle-ci ait été particulièrement forte. L'axe 2, lui, montre le succès du traitement: ($F2 > 0$), conjonction entre gène au début et bien à la fin; ou l'échec: ($F2 < 0$), conjonction entre gène à la fin et bien-être au début (relativement du moins). Cette dernière analyse éclaire l'usage que les sujets font des notes.

4.3.1 Le questionnaire d'auto-évaluation de BORTNER

Conçu d'après les travaux de ROSENMAN et FRIEDMAN, le questionnaire de BORTNER vise à apprécier la tendance plus ou moins forte des sujets à présenter un comportement de type 'A', caractérisé par la hâte, l'impatience, l'agressivité. De façon précise, il s'agit d'un questionnaire d'auto-évaluation comprenant 14 items. Chaque item est défini par un couple de phrases antithétiques, disposées aux deux extrémités d'un axe gradué sur lequel le sujet doit placer une croix exprimant sa propre attitude entre les deux pôles opposés.



Dans l'item n° 1, donné ci-dessus à titre d'exemple, la phrase écrite à gauche est censée représenter le pôle 'A'; tandis qu'à droite on a l'autre pôle, appelé 'B'. La réponse d'un sujet est traduite en une note, ou nombre entier pouvant aller de 1 à 24; cette dernière valeur correspondant au pôle 'A'.

Comme il est bien connu que les réponses placées sur une échelle sont choisies par chaque sujet en fonction d'une tendance personnelle à aller plus ou moins loin vers la droite, indépendamment du sens de la question, les pôles 'A' et 'B' des 14 items sont disposés diversement, à l'une ou l'autre des extrémités de l'échelle: le pôle 'A' est à droite pour les questions {2, 5, 7, 9, 11, 12, 14}; il est à gauche pour les sept autres questions; (pour 9, cf. *infra*, §4.3.2).

D'ordinaire, l'on calcule d'après l'ensemble du questionnaire une note 'A' globale, qui est la moyenne des 14 notes afférentes aux diverses questions; ces dernières étant comptées soit à partir de l'extrémité droite (comme pour la question 1), soit à partir de la gauche, selon la polarité de la présentation.

Du point de vue de l'analyse multidimensionnelle, on peut faire à cette pratique plusieurs objections. D'abord, la diversité des phrases (dont le libellé précis est donné dans la suite), laisse attendre une structure des réponses plus complexe que celle du schéma bipolaire {A, B} adopté par les auteurs. Ensuite, l'alternance des polarités ne compense pas complètement le biais dû à l'équation personnelle des sujets: en effet, ceux-ci ne se distinguent pas seulement entre eux par une tendance plus ou moins forte à se placer vers la droite, mais aussi par l'attraction ou la répulsion qu'exercent sur eux les extrémités de l'échelle.

Enfin, même si l'on suppose que l'auto-évaluation fournit une description fidèle du comportement du sujet (par exemple, pour la question 1, de sa ponctualité), ce comportement ne résulte pas seulement d'une "tendance" ('A'

ou 'B') mais du rôle dévolu au sujet dans la société; (la ponctualité n'étant pas pour tous également facile à tenir, ni également cruciale dans les rapports avec les autres...).

De certaines recherches, on a conclu que la note globale 'A' était associée à une majoration du risque de cardiopathie ischémique; d'où l'intérêt porté par les cardiologues au questionnaire de BORTNER. Les données dont nous disposons concernent seulement 97 patients des deux sexes, âgés de 35 à 77 ans. Mais disposant pour chacun des sujets d'un signalement (âge, sexe, profession, ...) avec quelques variables proprement médicales (antécédents cardiologiques, pathologie, hypertension) nous pouvons explorer la complexité des réponses au questionnaire de BORTNER, et mettre ces réponses en rapport avec le reste des données.

4.3.2 Analyse des réponses au questionnaire de BORTNER

À l'ensemble des 14 items du questionnaire de BORTNER, on a appliqué le codage sur 3 modalités suivant l'équation personnelle avec double recadrage des notes, très utile ici pour équilibrer les modalités d'items qui attirent inégalement les sujets (cf. §§2.3 & 3). Plus précisément, les modalités des questions {1, 3, 4, 6, 8, 10, 13}, pour lesquelles le pôle 'A' est présenté à gauche, sont notées {Az<, Az=, Az>} (z étant le numéro); pour les autres questions, dont c'est le pôle 'B' qui est mis à gauche, on note {Bz<, Bz=, Bz>}. Selon l'hypothèse des auteurs du questionnaire, les modalités {Az>, Bz<} caractérisent un comportement de type 'A', agressif; et les {Az<, Bz>} un comportement 'B'.

L'analyse du tableau 97 × 42 produit un facteur 1 qu'on peut, approximativement, interpréter comme un indice du comportement agressif 'A'. En effet, pour 12 des 14 questions, on a, du côté (F1<0) les modalités A< (ou, ce qui est équivalent: B>); les modalités {A>, B<} allant avec (F1>0). Mais la question 9 est franchement inversée, car B09>, "veut être reconnu", est du côté (F1>0), avec les modalités A>; alors que son antithèse, B09<, "Uniquement soucieux d'être content de lui" va avec (A<0), contrairement à ce que postule la version du questionnaire qu'a utilisée D. DAROUKH (version peut-être erronée: cf. CAD, ce n°, [NOTE MÉD.]). Pour la question 12 ("cacher ou Démontrer ses sentiments") il y a une inversion (d'ailleurs peu marquée, les COR1 étant très faibles), dont seule la CAH révèle la complexité.

Comme l'analyse factorielle, la classification des 42 modalités des 14 items s'accorde approximativement avec la notion d'une échelle de BORTNER. Au sommet de la hiérarchie, on distingue 3 classes: {j80, j69, j81}, qui correspondent au schéma {A>, A=, A<}, à quelques exceptions près, dont celle déjà signalée pour la question 12; les deux modalités {B12> et B12<} étant dans j81. En interprétant les subdivisions de j81, on découvre notamment comment les sujets ont compris la question 12. Il n'y a pas d'agressivité dans j81; la subdivision qui contient B12<, 'Démonstratif', suggère un caractère aimable et

jovial; tandis que celle qui contient B12>, 'cache ses sentiments') dessine une physionomie taciturne, où la tension n'est pas absente. C'est aussi avec B12> qu'on trouve la modalité B09< ("Uniquement soucieux d'être content de soi, quoi que les autres puissent penser") à laquelle l'analyse attribue une polarité opposée à celle (de type A>) prévue par les auteurs.

4.3.3 Le signalement: agressivité et rôle dans la société

Le tableau en (0,1) des modalités du signalement a été adjoint en supplémentaire à l'analyse du tableau de BORTNER, 97 × 42. Sur le demi-axe (F1>0) se détache la modalité 'profession de cadre supérieur', suivie de 'âge<55' (les sujets les plus jeunes de notre échantillon); avec, dans l'histoire de la maladie, angor, plutôt qu'infarctus ou syndrome de menace. Or la classe de modalités des items de BORTNER situées à l'extrémité (F1>0) de l'axe 1, manifeste, plutôt que l'agressivité et l'ambition, la hâte et la tension; suggérant un comportement qui peut être celui de cadres encore relativement jeunes.

Du côté (F1<0) on a Pouv (ouvrier), A>65 ans, avec des classes de modalités des items où on a cru reconnaître soit un caractère aimable et jovial, soit lenteur, voire apathie. Ces variables évoquent des sujets qui, en partie du fait de leur condition d'ouvrier ou de retraité (le tableau de données attribue une profession à des sujets même très âgés), n'ont pas une agressivité ou une tension qui majore le risque cardiaque.

S'inscrit sur l'axe 2 une distinction très nette entre les deux sexes. Ainsi 'Sexe féminin', se trouve, du côté (F2<0), opposé à la classe des modalités du questionnaire de BORTNER qui dessinent une personne agressive et ambitieuse dans la compétition. Dans notre échantillon, constitué pour un quart de femmes, généralement plus âgées que les hommes et dont une sur trois n'a pas exercé de profession rémunérée, aucune n'ayant atteint le niveau de cadre supérieur, cette opposition semble due autant à l'amabilité des dames qu'à la place qui leur a été faite dans la société.

L'axe 3 nous offre, de même, une association qu'on expliquera par la situation sociale des sujets: avec la modalité 'Sans profession' vont les classes de réponses 'non ponctuel, n'écoute pas' et aussi 'insouciant, tranquille'.

Le fait que, dans son étude multicentrique des sujets de langue française, de la Belgique à Marseille, le *Groupe coopératif de la fondation P. NEUMANN*, a trouvé une progression des notes globales du Nord au Sud, confirme que le questionnaire de BORTNER ne concerne pas seulement la tension ou l'agressivité, mais aussi d'autres facteurs qui ne jouent pas dans le même sens sur le risque cardiaque. On soulignera ici, d'une part, le cas des questions 12 et 9; et, d'autre part, le lien entre les réponses des sujets et certaines modalités du signalement qui ne résultent qu'en partie d'un comportement volontaire ou d'un type caractérologique.

5 Conception et utilisation des échelles en psychiatrie

5.1 Acquisition de l'information

Nous avons dit (cf. §2.1) que, pour le traitement statistique, il importe peu qu'un sujet se décrive lui-même par les réponses qu'il donne; ou qu'un observateur fasse la description, en usant, s'il y a lieu, d'appareils de mesure. Quant à la validité des informations recueillies, au contraire, la différence est grande.

D'une part, on peut présumer que l'observateur saura, dans un entretien libre avec le sujet, acquérir l'information pertinente; et qu'il déduira de celle-ci les réponses appropriées aux items du questionnaire; tandis que le sujet lui-même, au contraire, pourra ne pas comprendre un terme technique, être embarrassé pour fixer le degré d'intensité de ce qu'il éprouve, hésiter à écrire ce que son attitude révèle pourtant sans ambiguïté.

Mais, d'autre part, l'observateur, qui a une conception théorique de la pathologie, sera tenté de broser un tableau où tous les items se groupent en syndromes classiques; alors que le sujet répond aux questions successives avec d'autant plus d'ingénuité que celles-ci ne sont pas groupées en chapitres, mais, au contraire, disposées pour se tempérer mutuellement (cf. [QUEST. ALGÉRIE], §3.5); et nos critiques à propos de [QUEST. MÉM. RECH.], §3.3).

On a pu obtenir des réponses chargées de sens, en donnant aux sujets pour consigne de noter, de -3 à +3, sans réfléchir, les mots d'une longue liste (cf. [NOTES MOTS], §3.4). Sans prôner une telle consigne, nous croyons qu'il faut tenter de recueillir les réponses mêmes des sujets; et que, pour cela, l'énoncé des questions et leur enchaînement doivent être mûrement pesés. Le questionnaire de BORTNER, (cf. §4.3), même s'il ne mesure pas une dimension unique d'agressivité, offre un exemple appréciable d'auto-évaluation.

5.2 Format des questions

Sans doute, certains admettent-ils encore qu'il y a une différence majeure entre les données numériques continues (voire entières), dites quantitatives, et les réponses discrètes, dites qualitatives, prises dans un ensemble fixé de locutions ou de mots. Nous avons dit (cf. §1) qu'une description multidimensionnelle est toujours de l'ordre de la qualité; la quantité n'étant véritablement atteinte qu'après analyse. Le médecin sait d'expérience qu'un chiffre de pression artérielle ou de glycémie peut ne prendre son sens que dans un contexte; et que le diagnostic, (tranché: donc susceptible d'égalité et non seulement de ressemblance; qualitatif donc, pour ARISTOTE), est au terme d'une synthèse.

Pour l'analyse mathématique, le format initial des données importe assez peu. On peut coder de la même façon, par codage barycentrique, en tenant

compte éventuellement de l'équation personnelle, des abscisses relevées sur une échelle (cf. [ERGO. RÉGLAGES], §3.2; etc...), des notes variant de 1 à 10 (cf. [POLIT. GREC.], §3.1), ou seulement de 0 à 6 (diagramme HARD, cf. §4.1) voire de 0 à 4 (échelle de COVI, *ibid.*). Si l'ensemble des modalités de réponse n'est pas ordonné, (e.g. de 0 à 4, ou de 'très faible' à 'très fort'), seul convient le codage en (0,1), sous forme disjonctive complète; mais ce codage peut aussi servir si l'on découpe en modalités successives l'intervalle de variation d'une grandeur continue.

Le format intéresse donc, principalement, l'acquisition de l'information. Il faut, d'une part, que le sujet, (ou l'observateur qui remplit le questionnaire), ne soit pas gêné par un format trop rigide (tel que: 'pour' ou 'contre', deux modalités); et, d'autre part, qu'il comprenne la tâche qui lui est assignée. Nous soulignerons ici que les échelles continues, qui ont notre faveur, ne sont pas, quoi qu'on dise, d'un usage délicat. Dans le cas d'un questionnaire politique, (cf. [QUEST. ALGÉRIE], §3.5), on a vu, d'abord, les échelles bien utilisées, sur toute leur longueur, par 100 sujets rassemblés par des enquêteurs improvisés; puis des sociologues de profession donner pour consigne de ne marquer de réponse qu'au centre ou à l'une des extrémités!

Le statisticien, muni d'un outil de calcul, peut traiter des données de tout format; la difficulté majeure est d'obtenir une réponse pertinente. On trouve, dans notre exposé, plusieurs exemples d'échelles, dessinées telles qu'elles ont été soumises à des sujets: cf. §§3.2, 3.5, 4.2.3 & 4.3.1. Selon les cas, il y a, aux deux extrémités de l'échelle, des énoncés ou des expressions nuancées, qui précisent les deux pôles d'une opposition; ou seulement des expressions antithétiques passe-partout (telles que 'tout-à-fait d'accord', 'tout-à-fait opposé'). L'important est que, sans être agressif, le texte proposé soit clair; et que, dans leur quasi-totalité, les sujets soient touchés par la question.

On sait que les questionnaires à tiroirs qui, (en ouvrant des blocs de questions différents selon les réponses précédemment données), offrent à chaque sujet l'occasion de se poser en cas d'espèce, se prêtent mal à la synthèse statistique; le cas des questions qui n'ont de sens véritable que pour une fraction des sujets n'est, au fond, pas meilleur. La rédaction des questionnaires est un genre littéraire particulier, qui requiert le talent de proposer, en termes familiers, des thèses ou des situations qui, bien que générales, ne suscitent pas l'unanimité.

5.3 Des échelles de réponse à l'échelle de mesure

Il est commun de cumuler l'expression numérique des réponses fournies à un questionnaire en un total qu'on interprète comme mesurant une grandeur psychique. Prenons pour exemple l'*Échelle de Dépression* de HAMILTON: les modalités de réponse à chacune des 21 questions sont numérotées à partir de zéro, jusqu'à 3, 4 ou 5 selon les questions; ainsi, pour l'*Hypochondrie*, on a 5

degrés, de 0 = Absence à 4 = Idées délirantes hypochondriaques, le milieu étant 2 = Préoccupations sur la santé. On additionne toutes ces notes pour calculer un "total score".

De ce total, on entend faire l'analogie d'une mesure physique; en sorte que le score de HAMILTON guide le diagnostic et le traitement de la dépression, comme la pression artérielle pour l'hypertension. La confiance du psychiatre dans une échelle reposant, en bref, croyons-nous, sur une forte corrélation entre les items et sur la reproductibilité du total.

Cette prétention au rang de science quantitative culmine dans la formidable construction du DSM III, qui s'applique à embrasser, dans des règles strictes, une nosographie complète. Ne taxons pas trop vite de naïveté les auteurs du DSM III: ceux-ci savent mieux que nous ce qu'il y a d'arbitraire dans leur œuvre, (d'ailleurs fréquemment remise à jour: d'où le sigle DSM III (R)!); et c'est par un choix délibéré qu'ils entendent fonder sur un consensus verbal un système de communication qu'ils veulent universel.

Nous jugeons cependant que cette universalité est illusoire. Par exemple, l'Hypochondrie, évoquée ci-dessus, peut signer une dépression; mais des Préoccupations sur la santé peuvent être légitimes chez un malade chronique ou un sujet âgé qui n'est aucunement déprimé. Au sein d'une population de mal-portants, le score de HAMILTON s'en trouvera indûment majoré.

Pour nous, les notes globales ont, au mieux, valeur de suggestion; il convient de reprendre sans cesse, sur des populations particulières relativement homogènes, rassemblées pour une étude déterminée, des analyses telles que celles dont nous avons multiplié les exemples; le cas du BORTNER étant sans doute le plus complet quant à la méthode, même si, avec 97 sujets, la base en est étroite.

Certes l'analyse des correspondances obtient presque toujours, comme premier facteur, l'équivalent d'un score global; mais il nous semble imprudent d'accepter sans réserve ce facteur comme on accepterait une mesure physique. Les opérations mathématiques les plus audacieuses, (par exemple des calculs de différences ou de rapports), ont notre faveur, à condition toutefois qu'on ne perde jamais de vue l'ensemble des résultats de l'analyse factorielle et des classifications, qui seuls permettent de critiquer ce qu'on découvre: c'est ce qu'on s'est appliqué à faire dans les exemples.

Même dans le domaine cardiovasculaire, où les mesures physiques ou biochimiques abondent, la recherche de l'universalité est décevante. Les grands essais thérapeutiques n'aboutissent pas à des conclusions stables parce qu'ils sont fondés sur des populations hétérogènes dont les diverses composantes se mêlent en des proportions non maîtrisées. (Sous le titre de "Révolte d'Esculape", in *CAD*, Vol. XI, n°2 & n°3, 1986, nous avons rassemblé maints

témoignages du mécontentement des médecins devant des constructions prétentieuses et fragiles.)

Dans notre quête de la quantité, la définition des grandeurs est en cause avant la précision de la mesure. Les savants qui, il y a quatre siècles, ont entrepris d'étudier le chaud et le froid, avaient à découvrir, avec la température, la quantité de chaleur, la capacité calorifique, la conductibilité thermique, les chaleurs latentes de changements d'états... : l'ensemble de ces notions intervient en effet dans ce qu'avant toute science, on appelle chaud et froid. Or en psychométrie, la complexité ne peut être moindre.

5.4 Variations des réponses et variations des sujets au cours du temps

Dans plusieurs des études rapportées aux §§3 & 4, les mêmes sujets répondent, plusieurs fois de suite, au même questionnaire. Dans [QUEST. ERGO.], §3.2, on découvre que, de la première à la quatrième séance, l'appréciation des réglages d'un poste de travail varie: une table qui paraissait d'abord bien placée finit par paraître trop basse. Dans [ÉCHELLES], §4.1.3, on a pu, à la fois, suivre la régression de l'intensité de la pathologie dépressive et la stabilité des profils individuels. Dans [COMPAR. EFFORT], §4.2.1, l'efficacité des traitements est appréciée en analysant les différences entre des valeurs basales et des valeurs relevées dans des conditions éprouvantes et sous traitement. Dans [COMP. VAR.], §4.2.3, le profil de 15 notes de gêne est présenté globalement comme un point, dont la place s'interprète en terme de croissance ou de décroissance. Dans [COMPAR. CARD.], on s'étonne de ce que la qualité de vie déclarée par les sujets régresse, en général, sans relation avec une amélioration clinique éventuelle.

L'outil multidimensionnel nous paraît au point; reste, ici comme ailleurs, à réussir dans la critique des résultats. Ainsi une régression de la qualité de vie déclarée ne doit pas être indûment assimilée à une détérioration objective: le sujet peut, au cours du traitement, perdre patience. Plus embarrassant est l'apprentissage: nous avons dit notre confiance en l'ingénuité du sujet (cf. §5.1): l'ingénuité ne survit pas à des passations répétées.

Cet obstacle se rencontre même dans des mesures physiques. Une étude de la "Variation diachronique de la capacité d'exercice des insuffisants cardiaques", (cf. [EXERC. INSUFF.], in CAD, Vol. XV, n°2), a d'abord établi qu'indépendamment du traitement reçu, la majorité des patients voit sa *capacité* d'exercice ergométrique croître au cours du temps. Mais il s'agit d'un phénomène complexe, qui semble présenter un fléchissement au bout de 3 mois (à la 12-ème semaine); et la *capacité* serait corrélée négativement à la Fraction d'éjection. D'où cette conclusion de l'étude: "Pour faire de l'ergométrie une méthode d'investigation non invasive précise, il faudra cerner de plus près cette sorte

d'apprentissage (ou d'accoutumance); notamment en mesurant la *Fraction d'éjection*, avant chaque exercice ergométrique, pour des séries de patients”.

On s'affranchit des effets de l'accoutumance en considérant non les variations propres à un sujet, mais les différences entre profils de variation, de sujet à sujet (cf. §4.2.3); il reste que l'accoutumance, sans doute commune à tous, ne revêt pas, chez tous, la même forme.

5.5 Extension et confrontation des questionnaires

Dans [ÉCHELLES], §4.1.2, on a pu insérer les 4 items de l'échelle de COVI parmi les 12 du diagramme HARD. L'analyse de [COMPAR. EFFORT], §4.2.1, ne souffre aucunement de ce qu'aux quatre notes l'échelle HASE soient adjointes deux variables physiologiques. Les analyses globales aident à reconnaître plusieurs syndromes cohérents là où l'on ne percevait d'abord qu'une seule dimension. De plus, en insérant graduellement de nouveaux items immédiatement confrontés à ceux déjà utilisés, on renouvellera un questionnaire sans jamais s'aventurer dans l'inconnu. Des modalités extérieures au questionnaire psychométrique peuvent également être projetées en éléments supplémentaires, comme on l'a fait pour le signalement, avec l'échelle de BORTNER.

On peut donc, de plusieurs manières, comme l'entreprend le Dr. BOSC, traiter conjointement, les items de l'Échelle de Dépression de HAMILTON et ceux d'une nouvelle échelle d'auto-évaluation de l'Adaptation Sociale. Et si, comme il nous paraît, on a, dans cette échelle, sous le titre d'Adaptation Sociale, plusieurs fois inventorié des aspects de la Dépression, l'analyse multidimensionnelle aidera à reconnaître ces contaminations.

Il faut toutefois rappeler que, comme l'agressivité sondée par BORTNER, la pathologie de la dépression se manifeste dans un contexte social. De ce point de vue, la note de dépression devrait être une note relative. Sur les axes issus d'une analyse croisant items de dépression et signalement, chaque sujet a deux projections déterminées chacune par l'un des deux questionnaires qu'il a remplis: le décalage entre les deux projections serait l'expression de la tension psychosociale.

Nous espérons avoir convaincu le psychiatre de la puissance et de la souplesse des méthodes multidimensionnelles. Le plus grand mérite de ces méthodes est, sans doute, qu'elles offrent les résultats numériques dans un contexte qui y apporte des nuances. Car au calcul, il ne convient pas de lâcher la bride. Riche en suggestions, prompt à réussir les constructions les plus audacieuses, le calcul n'est fécond que s'il se fait, avant tout, l'instrument de la critique.