

Apposant que l'immigration ait continué à son rythme d'avant . On sait ce qui est arrivé à beaucoup de ceux qui souhaitaient r leur pays, mais ne savaient où aller. La destruction emprunte nt des chemins détournés, mais les idées peuvent en être des , plus sûrs que les fusils et les bombes.

## C H A P I T R E V

## La véritable erreur de Cyril Burt

## L'ANALYSE FACTORIELLE ET LA RÉIFICATION DE L'INTELLIGENCE

Ce fut le mérite insigne de l'école anglaise de psychologie, depuis Sir Francis Galton, d'avoir transformé les tests mentaux : jadis manigances de charlatan, ils sont devenus, grâce à cette méthode d'analyse mathématique, un instrument scientifique de précision, reconnu de tous. (Cyril Burt, 1921, p. 130)

*L'affaire Cyril Burt*

Si j'avais eu le désir de vivre une vie nonchalante, à l'abri de tout souci maternel, j'aurais voulu être un jumeau vrai, séparé de mon frère à la naissance et élevé dans un milieu social différent. Nous pourrions tous deux louer nos services à prix d'or à une foultitude de chercheurs sociaux. Car nous serions les représentants extrêmement rares de la seule expérience naturelle dans laquelle, chez l'homme, la génétique est dissociée des effets du milieu : nous serions des individus génétiquement identiques élevés dans des milieux distincts.

Les études sur les jumeaux vrais élevés séparément devraient donc occuper une place de choix dans la littérature sur l'hérédité du QI. Ainsi en devrait-il aller s'il n'y avait une difficulté, l'extrême rareté de l'animal lui-même. Peu de chercheurs ont réussi à rassembler plus de vingt couples. Cependant, au milieu de cette pénurie, une étude semblait émerger : celle de Sir Cyril Burt (1883-1971). Sir Cyril, le doyen des pionniers des tests mentaux, avait mené deux carrières successives

qui lui permirent d'acquiescer une position prééminence en dirigeant tout à la fois la théorie et la pratique dans son domaine, la psychologie de l'éducation. Pendant vingt ans, il fut le psychologue officiel du London County Council, responsable de l'application et de l'interprétation des tests mentaux dans les écoles de Londres. Il succéda ensuite à Charles Spearman à la chaire de psychologie la plus prestigieuse de Grande-Bretagne, celle de l'University College de Londres (1932-1950). Au cours de sa longue retraite, Sir Cyril publia plusieurs articles où il fit état d'une très haute corrélation entre les QI de jumeaux vrais élevés séparément venant ainsi appuyer la thèse héréditariste. Les études de Burt se détachaient de toutes les autres, car il était parvenu à trouver cinquante-trois couples, ce qui faisait plus que doubler le total atteint par n'importe lequel de ses prédécesseurs. Il est peu surprenant qu'Arthur Jensen ait utilisé le chiffre de Sir Cyril comme son argument le plus décisif dans son célèbre article (1969) sur les prétendues différences d'intelligence, héréditaires et irréversibles, entre les Blancs et les Noirs des Etats-Unis.

L'histoire de la chute de Burt a été maintes fois contée. Un psychologue de Princeton, Leon Kamin, s'aperçut pour la première fois qu'au moment où l'échantillon de Burt était passé, en une série de publications, d'une petite vingtaine de couples de jumeaux à plus de cinquante, la corrélation moyenne du QI entre les couples était restée inchangée à la troisième décimale près, situation qui, dans le domaine statistique, est si improbable qu'elle correspond à la définition du mot impossible. Puis, en 1976, Oliver Gillie, correspondant médical du *Sunday Times* de Londres, fit passer les charges pesant sur Cyril Burt de la négligence inexcusable au truquage conscient. Gillie découvrit, entre autres choses, que les deux « collaboratrices » de Burt, Margaret Howard et J. Conway, qui avaient prétendument recueilli et analysé ses données, soit n'avaient pas pu être en contact avec Burt lorsqu'il rédigea les articles portant leur signature, soit, pire encore, n'avaient jamais existé. Ces accusations conduisirent à reconsidérer les « preuves » avancées par Burt à l'appui de sa rigide position héréditariste. Et il s'avéra que d'autres études essentielles étaient tout aussi frauduleuses, notamment ses corrélations de QI entre parents proches (suspectes, car trop belles pour être vraies et apparemment élaborées à partir de distributions statistiques idéales, plutôt que réellement mesurées — Dorfman, 1978) et ses données sur les niveaux décroissants d'intelligence en Grande-Bretagne.

Les partisans de Burt considérèrent tout d'abord ces accusations comme un complot à peine voilé, ourdi par la gauche pour discréditer la position héréditariste. H.J. Eysenck écrivit à la sœur de Burt : « Je pense que l'affaire a été montée de toutes pièces par quelques environnementalistes d'extrême gauche bien décidés à jouer un jeu politique avec des faits scientifiques. Je suis sûr, sans le moindre doute, que l'avenir réhabilitera Sir Cyril dans son honneur et sa probité. » Arthur

Jensen, qui avait dit de Burt qu'il était un « noble né » et « l'un des plus grands psychologues du monde », dut admettre qu'on ne pouvait se fier aux données sur les jumeaux vrais, bien qu'il n'attribuât leur inexactitude qu'à la seule négligence.

Je crois que la magnifique biographie « officielle » de Burt qu'a récemment publiée L.S. Hearnshaw (1979) a apporté une solution à ce problème autant que les données le permettent (Hearnshaw avait été chargé de la rédaction de ce livre par la sœur de Burt, avant que la moindre accusation n'ait été portée). Hearnshaw qui, au début, était un admirateur inconditionnel de Burt dont il tend à partager les attitudes intellectuelles, finit par conclure que toutes les allégations étaient vraies, et pire encore. Malgré tout, il m'a convaincu que l'énormité même et la bizarrerie de la supercherie de Burt nous forcent à la considérer non comme le programme « rationnel » d'un personnage sournois tentant, par tous les moyens, de sauvegarder son dogme héréditariste, alors qu'il savait que la partie était perdue (ce qui fut ma première interprétation, je l'avoue), mais comme les actes d'un homme malade et tourmenté. (Tout ceci, bien entendu, ne résout en rien une autre question, celle de savoir pourquoi des données aussi manifestement contrefaites, restèrent si longtemps indiscutées et ce que cette persistance implique sur le fondement de nos présuppositions héréditaristes.)

Hearnshaw pense que Burt commença ses falsifications au début des années 1940 et que son travail précédent était honnête, bien que gâché par des *a priori* rigides, et souvent inexcusablement bâclé et superficiel, même selon les normes de son temps. Le monde de Burt commença à s'effondrer durant la guerre, en partie de son propre fait, à n'en pas douter. Les données de ses recherches furent détruites dans le blitz de Londres ; son mariage échoua ; il fut exclu de sa propre chaire lorsque, parvenu à l'âge légal, il refusa de prendre sa retraite de son propre gré et essaya de conserver les rênes du pouvoir ; il perdit son poste de rédacteur dans la revue qu'il avait fondée, là aussi pour ne pas avoir voulu céder la place au moment que lui-même avait fixé pour la cessation de ses activités ; son dogme héréditariste ne s'accrochait plus à la mentalité d'une époque qui venait d'être témoin de l'Holocauste. En outre, Burt souffrait apparemment de la maladie de Ménière, qui consiste en des troubles des organes de l'équilibre, aux conséquences souvent négatives sur la personnalité.

Hearnshaw fait état de quatre cas de fraude dans la dernière partie de la carrière de Burt. J'en ai déjà mentionné trois : l'invention de données sur les jumeaux vrais, de corrélations de QI chez des parents proches et les niveaux décroissants de l'intelligence en Grande-Bretagne. Le quatrième est, à beaucoup d'égards, le plus étrange de tous, car l'affirmation de Burt était si absurde et ses actes si manifestes qu'il lui était impossible de ne pas être démasqué un jour ou l'autre. Ce comportement ne pouvait pas être le fait d'un homme en parfaite santé

mentale. Burt tenta de commettre un véritable patricide intellectuel en prétendant être le père d'une technique appelée « analyse factorielle », alors que ce titre revient de droit à son prédécesseur et mentor, Charles Spearman. Ce dernier avait présenté l'essentiel de sa découverte dans un article fameux datant de 1904. Burt ne mit jamais en doute cette priorité — il l'affirmait en fait constamment — tant que Spearman dé tint la chaire que Burt occuperait plus tard à l'University College. Qui plus est, dans son célèbre ouvrage sur l'analyse factorielle (1940), Burt déclara que « la prééminence de Spearman est reconnue par tous les analystes » (1940, p. X).

Burt tenta une première fois de réécrire l'histoire du vivant même de Spearman et cela lui valut une réponse cinglante du détenteur honoraire de la chaire de Burt. Celui-ci se rétracta immédiatement et envoya à Spearman une lettre qui n'a pas son pareil quant à la défiance et l'obséquiosité : « Il est évident que votre priorité ne fait aucun doute. [...] Je me suis demandé à quels endroits précis j'avais pu m'égarer. Ne serait-il pas plus simple que je numérote mes déclarations, puis, comme mon maître décole le faisait jadis, vous pourriez mettre une croix en face de chaque point où votre élève a commis une bévue et cocher ceux où votre thèse a été correctement interprétée. »

Mais, après la mort de Spearman, Burt déclencha une campagne qui, tout au long du reste de sa vie « devint de plus en plus effrénée, obsessionnelle et outrancière » (Hearnshaw, 1979). Hearnshaw écrit (1979, pp. 286-287) : « Les murmures contre Spearman qui étaient à peine audibles à la fin des années 1930 s'enflèrent en une campagne stridente de dénigrement qui grossit jusqu'à ce que Burt se soit arrogé la totalité de la renommée de Spearman. Il apparut que Burt semblait de plus en plus obsédé par les questions de priorité et de plus en plus susceptible et égocentrique. » La falsification de Burt se résume à peu de chose : c'est en 1901, trois ans avant la publication de l'article de Spearman, que Karl Pearson avait inventé l'analyse factorielle (ou une technique qui s'en approchait fort). Mais Pearson ne l'avait pas appliquée à des problèmes psychologiques. Burt se rendit compte des applications possibles et adapta la technique aux études sur les tests mentaux, en lui apportant au passage quelques modifications et améliorations décisives. La filiation s'établit donc de Pearson à Burt. L'article de Spearman de 1904 n'était simplement qu'une diversion.

Burt débâta son histoire maintes et maintes fois. Il la raconta même par le truchement d'un de ses nombreux pseudonymes dans une lettre qu'il écrivit à sa propre revue et qu'il signa du nom d'un psychologue français inconnu, Jacques Lafitte. À l'exception de Voltaire et de Binet, M. Lafitte ne mentionnait que des auteurs anglais et déclarait : « Il est certain que la première formulation correcte et formelle fut la démonstration de la méthode des axes principaux que Karl Pearson exposa en 1901. » Cependant quiconque aurait pu montrer, après une heure de recherche, que cette version était pure invention, car Burt

n'avait jamais cité l'article de Pearson dans aucun de ses écrits antérieurs à 1947, alors que toutes ses études précédentes sur l'analyse factorielle accordent la paternité de la technique à Spearman et mentionnent à l'évidence que les méthodes de Burt en dérivent.

L'analyse factorielle devait bien revêtir une très grande importance pour que Burt, dans sa recherche fébrile de la postérité, en ait fait le centre de ses préoccupations et se soit efforcé avec tant d'acharnement de s'en faire passer pour l'inventeur. Cependant, en dépit de tous les textes de vulgarisation sur le QI et l'histoire des tests mentaux, pratiquement rien n'a été écrit (en dehors des cercles professionnels) sur l'influence et la signification de l'analyse factorielle. Je soupçonne que la principale raison de ce désintérêt réside dans le caractère abstrait, mathématique de cette technique. Le QI, cette échelle linéaire qui fut conçue tout d'abord comme une mesure grossière et empirique, est facile à comprendre. L'analyse factorielle, qui tire son origine d'une théorie statistique abstraite et dont le principe se fonde sur la recherche de structures « sous-jacentes » au sein de grands tableaux de données numériques, est, pour dire les choses comme elles sont, d'abord malaisée. Mais négliger l'analyse factorielle est une grave omission pour quiconque désire comprendre l'histoire des tests mentaux et leur raison d'être aujourd'hui. Car, comme Burt le faisait remarquer à juste titre (1914, p. 36), l'histoire des tests mentaux se compose de deux tendances principales et connexes : les méthodes des échelles d'âge (les tests de Binet) et les méthodes basées sur les corrélations (l'analyse factorielle). En outre, comme Spearman l'a sans cesse souligné tout au long de sa carrière, la justification théorique de l'usage d'une échelle unilinéaire de QI repose sur l'analyse factorielle elle-même. Burt a pu faire preuve de perversité dans sa campagne, mais il avait raison quant à la stratégie adoptée : une place permanente et glorieuse était réservée dans le panthéon de la psychologie au promoteur de l'analyse factorielle.

J'ai débuté dans ma carrière de biologiste en employant l'analyse factorielle pour étudier l'évolution d'un groupe de reptiles fossiles. On m'avait appris cette technique comme si elle découlait de principes premiers utilisant la pure logique. En fait, pratiquement tous ses procédés sont nés pour permettre la justification de certaines théories spécifiques de l'intelligence. L'analyse factorielle, en dépit de son statut de pure technique mathématique déductive, fut inventée pour des raisons et dans un contexte social bien précis. Et, bien que son fondement mathématique soit inattaquable, son utilisation persistante comme outil de la connaissance de la structure physique de l'intellect s'est embourbée à sa naissance dans de profondes erreurs conceptuelles. La principale, en fait, participe d'un des thèmes majeurs de ce livre, la réification, en l'occurrence la notion selon laquelle un concept nébuleux, socialement défini, comme l'intelligence, pourrait s'identifier à une « chose » possédant une localisation précise dans le cerveau

et un degré donné d'hérédabilité — et selon laquelle on pourrait mesurer cette chose et la réduire à un chiffre unique permettant de classer les individus en fonction de la quantité qu'ils en possèdent. En assimilant un axe factoriel mathématique au concept d'« intelligence générale », Spearman et Burt apportèrent une justification théorique à l'échelle linéaire que Binet avait proposée comme un simple guide empirique.

Les discussions animées autour de l'œuvre de Cyril Burt se sont concentrées sur les supercheries de la fin de sa carrière. Cette perspective a masqué la profonde influence que Sir Cyril a exercée en tant que spécialiste des tests mentaux : car il fut le plus puissant de tous ceux qui se sont évertués à tirer de l'analyse factorielle un modèle présentant l'intelligence comme une « chose » réelle et unique. Les convictions de Burt s'enracinaient dans cette erreur qu'est la réification. Les manipulations frauduleuses ultérieures furent la réaction tardive d'un homme vaincu ; mais son erreur précédente, « honnête », s'est répétée sur tout notre siècle et a eu des conséquences sur des millions de vies humaines.

### *Corrélation, cause et analyse factorielle*

#### CORRÉLATION ET CAUSE

L'esprit de Platon a la vie dure. Nous sommes incapables d'échapper à cette tradition philosophique qui veut que ce que nous voyons et mesurons dans le monde ne soit que la représentation superficielle et imparfaite d'une réalité cachée. L'essentiel de la fascination des statistiques tient dans ce sentiment viscéral — méfiez-vous tous jours des sentiments viscéraux — que les mesures abstraites résumant de grands tableaux de données doivent exprimer quelque chose de plus réel et de plus fondamental que les données elles-mêmes. (Pour acquérir un bon niveau professionnel, les statisticiens doivent faire un effort conscient pour contrebalancer cette tendance naturelle.) La technique de *corrélation* a donné lieu à des abus, car elle semble fournir une voie privilégiée aux déductions sur la causalité (et c'est bien ce qu'elle fait parfois, mais seulement parfois).

La corrélation évalue la tendance qu'à une mesure de varier de concert avec une autre. Au cours de la croissance d'un enfant, par exemple, ses bras et ses jambes s'allongent ; cette tendance commune à changer dans la même direction s'appelle une *corrélation positive*. Tous les organes du corps ne présentent pas de corrélations positives semblables pendant la croissance. Les dents, par exemple, ne grossissent pas après leur percée. Le rapport entre la taille de la première

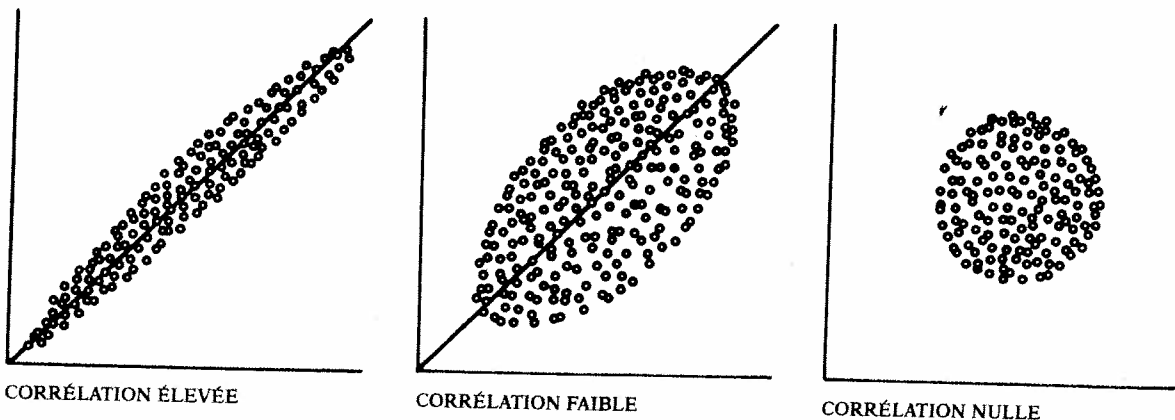
incisive et la longueur des jambes depuis l'âge de, mettons, dix ans jusqu'à l'âge adulte représente une *corrélation nulle* : les jambes s'allongent alors que les dents ne changent pas du tout. D'autres corrélations peuvent être négatives — un phénomène s'accroît tandis que l'autre décroît. Nous commençons à perdre des neurones à un âge désespérément précoce, et ils ne sont jamais remplacés. Ainsi, le rapport entre la longueur des jambes et le nombre de neurones après la mi-enfance représente une *corrélation négative* — la longueur des jambes s'accroît alors que le nombre des neurones décroît. Remarquez bien que je n'ai pas parlé de causalité. Nous ne savons pas pourquoi ces corrélations existent ou n'existent pas, nous savons seulement si elles sont là ou non.

La mesure classique de la corrélation est appelée le coefficient de corrélation de Pearson, ou, plus brièvement symbolisé,  $r$ . Le coefficient de corrélation va de  $+1$  pour une corrélation positive parfaite, à  $0$  pour une corrélation nulle, puis à  $-1$  pour une corrélation négative parfaite\*.

En gros, le  $r$  mesure la forme d'une ellipse de points tracés sur un diagramme (voir fig. 5.1). Des ellipses très efflanquées représentent de hautes corrélations — la plus mince de toutes, la ligne droite correspond à un  $r$  de 1.0. Des ellipses rebondies sont le signe de corrélations plus faibles, et la plus obèse de toutes, le cercle, traduit une corrélation nulle (l'augmentation d'une mesure ne permet absolument pas de prévoir si l'autre va augmenter, diminuer ou rester la même).

Le coefficient de corrélation, quoique aisément calculé, a été victime d'erreurs d'interprétation. On peut en donner des exemples. Supposez que je trace point par point les données sur la longueur des bras et celles sur la longueur des jambes pendant la croissance d'un enfant. Je vais obtenir une corrélation élevée avec deux implications intéressantes. En premier lieu, j'aurai réalisé une *simplification*. J'ai commencé avec deux dimensions (la longueur des bras et des jambes) que j'ai à présent, effectivement, réduites à une. La corrélation étant très forte, on peut dire que la ligne elle-même (une seule dimension) représente quasiment toutes les informations qui ont été originellement fournies sous une forme bidimensionnelle. En second lieu, je peux, dans ce cas, faire une déduction raisonnable sur la *cause* de cette réduction à une dimension. La longueur des bras et celle des jambes

\* Le  $r$  de Pearson n'est pas une mesure convenant pour toutes les sortes de corrélations, car il n'évalue que ce que les statisticiens appellent l'intensité de la liaison linéaire entre deux mesures, c'est-à-dire la tendance qu'ont tous les points à s'aligner sur une droite. D'autres liaisons de stricte dépendance ne présenteront pas une valeur de 1.0 pour le  $r$ . Si, par exemple, chaque augmentation de deux unités dans une variable correspondait à une augmentation de 2<sup>2</sup> unités dans une autre variable, le  $r$  serait inférieur à 1.0, même si les deux variables pouvaient être parfaitement « corrélées » au sens vulgaire du terme. Leur représentation graphique serait une parabole et non une droite, et le  $r$  de Pearson mesure l'intensité de la ressemblance linéaire.



5.1 La force d'une corrélation en fonction de la forme d'un nuage de points. Plus l'ellipse est allongée, plus la corrélation est élevée.

sont étroitement corrélées car ce sont toutes les deux des mesures partielles d'un seul même phénomène sous-jacent, à savoir la croissance.

Mais de peur qu'on ne s'imagine trop facilement que la corrélation représente une méthode magique pour déceler de façon non équivoque la cause des phénomènes liés entre eux, regardons la liaison entre mon âge et le prix de l'essence pendant ces dix dernières années. La corrélation est presque parfaite, mais personne ne s'aviserait d'y voir une cause commune. Le fait de la corrélation n'implique rien sur la cause. Il n'est même pas vrai que des corrélations intenses soient plus susceptibles que les faibles de représenter un lien causal, car la corrélation entre mon âge et le prix de l'essence est proche de 1.0. J'ai parlé de cause dans le cas de la longueur des bras et des jambes non pas parce que la corrélation était élevée, mais à cause de ce que je savais sur la biologie de la situation. La déduction sur la cause doit venir d'ailleurs, non du simple fait de la corrélation — bien qu'une corrélation inattendue puisse nous amener à chercher une cause commune tant que nous gardons en mémoire le fait qu'il est possible que nous ne la trouvions pas. La très grande majorité des corrélations dans notre monde sont, sans aucun doute, non causales. Tout ce qui a décrit régulièrement ces dernières années sera fortement corrélié à la distance séparant la terre de la comète de Halley (qui diminue aussi depuis quelque temps), mais même l'astrologue le plus acharné ne distinguerait aucune causalité dans la plupart de ces liaisons. L'hypothèse selon laquelle corrélation équivaut à cause est probablement l'une des deux ou trois erreurs les plus répandues et les plus graves du raisonnement humain.

Peu de personnes se laisseraient prendre par une démonstration absurde telle que la corrélation âge-essence. Mais prenons un cas intermédiaire. On me donne un tableau de données sur les longueurs des lancers de balle effectués par vingt enfants. J'établis la courbe de ces données et calcule un  $r$  élevé. La plupart des gens, je pense, partageraient mon intuition, à savoir qu'il ne s'agit pas là d'une corrélation sans signification ; cependant, en l'absence de tout autre renseignement, la corrélation en elle-même ne m'apprend rien sur les causes sous-jacentes. Car je peux proposer au moins trois interprétations causales différentes et plausibles de cette corrélation (la vraie raison se trouvant probablement dans un mélange des trois) :

1. Les enfants sont simplement d'âge différent, les plus âgés lançant la balle le plus loin.
2. Les différences représentent des niveaux variés de pratique et d'entraînement.
3. Les différences proviennent de disparités dans les capacités naturelles qui ne peuvent être effacées même par un entraînement intensif. (La situation serait encore plus complexe si l'échantillon renfermait des garçons et des filles d'éducation conventionnelle. La corrélation pourrait alors être attribuée prioritairement à une qua-

trême cause, les différences sexuelles ; et nous pourrions nous soucier, en plus, de la cause de cette différence sexuelle : instruction, constitution propre, ou quelque combinaison entre l'un et l'autre.)

En résumé, la plupart des corrélations sont non-causales ; lorsqu'elles sont causales, le fait et la force de la corrélation spécifient rarement la nature de la cause.

#### CORRÉLATION À PLUS DE DEUX DIMENSIONS

Ces exemples bidimensionnels sont faciles à saisir (aussi difficiles soient-ils à interpréter). Mais qu'en est-il des corrélations entre plus de deux mesures ? Un corps est formé de nombreux organes et pas seulement de bras et de jambes. Que faire si l'on désire savoir combien de mesures sont interdépendantes pendant la croissance ? Supposons, pour simplifier, que l'on ajoute une seule mesure supplémentaire, la longueur de la tête, pour réaliser un système à trois dimensions. On peut à présent décrire la structure des corrélations entre ces trois mesures de deux façons :

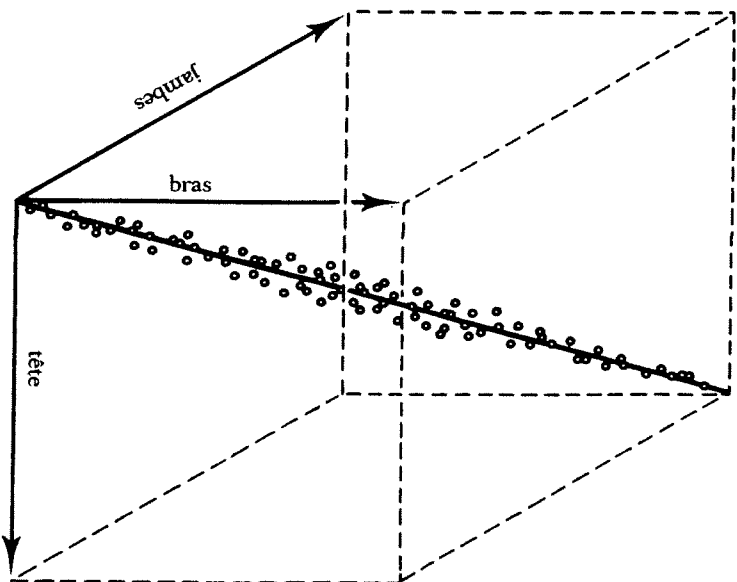
1. Nous pouvons rassembler tous les coefficients de corrélation entre les couples de mesures en un seul tableau ou *matrice* de coefficients de corrélation (fig. 5.2). La droite qui va du coin supérieur gauche au coin inférieur droit suit la ligne de corrélation, nécessairement parfaite, de chaque variable avec elle-même. On l'appelle la diagonale principale ; toutes les corrélations y sont de 1.0. La matrice est symétrique de part et d'autre de la diagonale principale, puisque la corrélation entre la mesure 1 et la mesure 2 est la même que la corrélation entre 2 et 1. Ainsi, les trois valeurs, qu'elles soient au-dessus ou

	bras	jambes	tête
bras	1.0	91	72
jambes	91	1.0	63
tête	72	63	1.0

5.2 Matrice de corrélations pour trois mesures.

au-dessous de la diagonale sont les corrélations que nous cherchons : bras et jambes, bras et tête, jambes et tête.

2. Nous pouvons placer les points concernant tous les individus sur un diagramme tridimensionnel (fig. 5.3). Comme les corrélations sont toutes positives, le nuage de points est orienté comme un ellipsoïde (un ballon de rugby). En deux dimensions, il s'agissait d'une ellipse. Une droite passant par le grand axe de ce ballon de rugby exprime les fortes corrélations positives entre les mesures.



5.3 Diagramme tridimensionnel montrant les corrélations pour trois mesures.

Il nous est possible de saisir ce cas tridimensionnel, à la fois mentalement et visuellement. Mais comment s'imaginer un espace à vingt, à cent dimensions ? Si nous mesurons cent organes d'un corps en croissance, notre matrice de corrélations contiendrait 10 000 nombres. Pour mettre ces informations sous forme graphique, il nous faudrait travailler avec un espace à cent dimensions, avec cent axes mutuellement perpen-

diculaires représentant les mesures originales. Bien que ces cent axes ne constituent pas un problème mathématique (ils forment, en termes techniques, un hyperespace), il nous est impossible de les porter sur un diagramme dans notre monde euclidien à trois dimensions.

Ces cent mesures d'un corps en croissance ne représentent vraisemblablement pas cent phénomènes biologiques différents. Exactement comme la plupart des informations de notre exemple à trois dimensions pouvaient se résumer en une seule dimension (le grand axe du ballon de rugby), de la même manière nos cent mesures pourraient être simplifiées en un petit nombre de dimensions. Nous perdions certaines informations dans le processus, c'est certain, comme nous l'avons fait lorsque nous sommes passés de cet ellipsoïde très allongé, qui restait encore une structure tridimensionnelle, à une seule droite représentant son grand axe. Mais nous sommes sans doute prêts à accepter cette perte en échange de la simplification qu'elle entraîne et de l'éventualité qu'elle pourrait ouvrir vers une interprétation, en termes biologiques, des dimensions retenues.

#### L'ANALYSE FACTORIELLE ET SES BUTS

Avec cet exemple, nous arrivons au cœur de ce que l'analyse factorielle essaie de réaliser. L'analyse factorielle est une technique mathématique permettant de réduire un système complexe de corrélations en un plus petit nombre de dimensions. Elle agit, littéralement, en mettant en facteurs un tableau, ordinairement une matrice de coefficients de corrélation. Rappelez-vous cet exercice d'algèbre appelé « mise en facteurs » consistant à simplifier d'horribles expressions en regroupant les multiplicateurs communs. Géométriquement, le processus de la mise en facteurs revient à placer des axes dans un ballon amassés suffisamment d'informations sur une seule droite, le long de ce grand axe de cet hyperballon de rugby, droite appelée *première composante principale*. Nous aurons besoin d'axes supplémentaires. Par convention, nous représenterons la deuxième dimension par une droite *perpendiculaire* à la première composante principale. Ce deuxième axe, ou *deuxième composante principale*, se définit comme la droite qui « explique » (le mot n'a pas ici de signification causale) une quantité de variables restantes plus grande qu'aucune autre droite qui pourrait être tracée perpendiculairement à la première composante principale. Si, par exemple, l'hyperballon de rugby était aplati comme une limande, la première composante principale passerait par le centre, de la tête à la queue, et la deuxième également par le centre perpendiculaire aux axes précédents et expliquerait un nombre régulièrement décroissant de variables prises parmi celles qui reste-

raient. On peut très bien considérer que cinq composantes principales expliquent presque toutes les variables de notre hyperballon de rugby, c'est-à-dire que cet hyperballon dessiné en cinq dimensions ressemble suffisamment à l'original pour nous convenir, tout comme une pizza ou une limande dessinée en deux dimensions peut exprimer toute l'information dont nous avons besoin, même si les deux objets renferment originellement trois dimensions. Si nous choisissons de nous arrêter à cinq dimensions, nous pouvons réaliser une simplification considérable au prix acceptable d'une perte minimale d'information. Nous pouvons saisir conceptuellement les cinq dimensions ; nous pouvons même les interpréter biologiquement.

Puisque la mise en facteurs est faite sur une matrice de corrélations, j'utiliserai une représentation géométrique des coefficients de corrélation eux-mêmes afin de mieux expliquer comment la technique fonctionne\*. Les mesures originales peuvent se représenter sous forme de vecteurs de la longueur d'une unité, irradiant à partir d'un point commun. Si deux mesures ont une corrélation élevée, leurs vecteurs sont situés l'un près de l'autre. Le cosinus de l'angle que forment deux vecteurs représente le coefficient de corrélation entre eux. Si deux vecteurs se chevauchent, leur corrélation est parfaite, c'est-à-dire de 1,0, le cosinus de 0° étant 1. Si deux vecteurs sont à angle droit, ils sont totalement indépendants, c'est-à-dire avec une corrélation de zéro, le cosinus de 90° étant de zéro. Si deux vecteurs se dirigent dans des

\* (Note réservée au mordu : les autres peuvent poursuivre sans dommage.) Je vais présenter ici un procédé appelé techniquement l'« analyse en composantes principales », qui est légèrement différent de l'analyse factorielle. Dans l'analyse en composantes principales, on conserve toutes les informations des mesures originales et on leur adapte de nouveaux axes selon le même critère utilisé dans l'analyse factorielle pour l'orientation des composantes principales, c'est-à-dire que le premier axe rend compte (explique) d'un plus grand nombre de données que tout autre axe et que les axes suivants sont situés à angle droit par rapport à tous les autres axes et regroupent des quantités d'informations régulièrement décroissantes. Dans l'analyse factorielle véritable, on décide au préalable (par divers moyens) de ne pas inclure toutes les informations sur les axes factoriels. Mais les deux techniques — la véritable analyse factorielle avec son orientation des composantes principales et l'analyse en composantes principales — jouent le même rôle conceptuel et ne diffèrent que dans le mode de calcul. Dans les deux analyses, le premier axe (le *g* de Spearman pour les tests d'intelligence) est la dimension « la mieux adaptée » qui explique plus d'informations dans un ensemble de vecteurs que tout autre axe.

Durant les dix dernières années, une confusion sémantique s'est répandue dans les cercles de statisticiens ; elle s'est exprimée dans la tendance à restreindre le terme d'« analyse factorielle » aux seules rotations d'axes habituellement réalisées après le calcul des composantes principales et à étendre le terme d'« analyse en composantes principales » à la fois à la véritable analyse en composantes principales (où toutes les informations sont retenues) et à l'analyse factorielle réalisée par l'orientation des composantes principales (nombre de dimensions réduit et perte d'informations). Ce changement de définition est en complet désaccord avec l'histoire du sujet et des termes. Spearman, Burt et quantité d'autres psychométriciens ont travaillé pendant des dizaines d'années avant que Thurstone et d'autres inventent les rotations axiales.

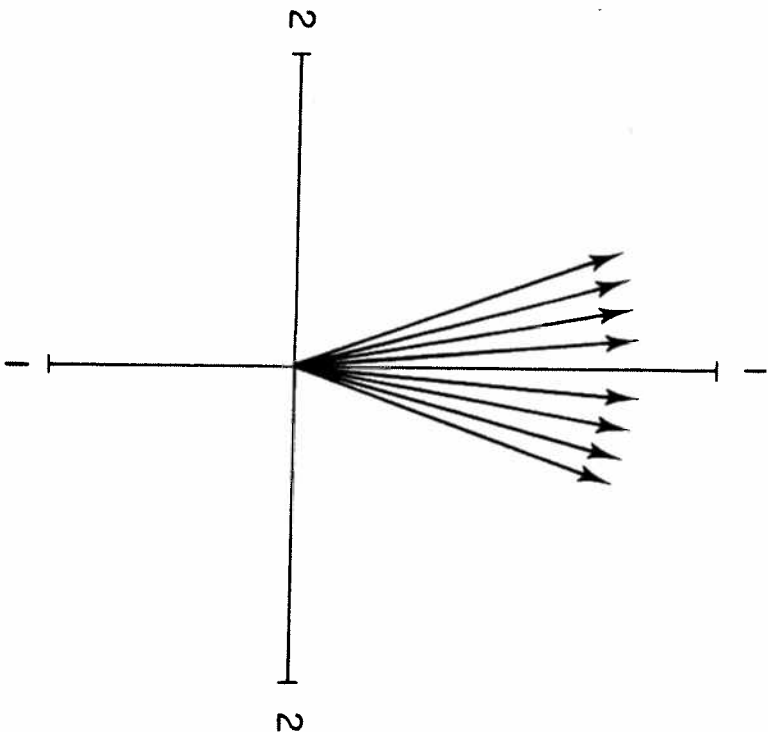
directions opposées, leur corrélation est parfaitement négative, c'est-à-dire de  $-1,0$ , le cosinus de  $180^\circ$  étant  $-1$ . Une matrice présentant des coefficients de corrélation positifs élevés sera représentée par un faisceau de vecteurs séparés entre eux par des angles aigus très petits (fig. 5.4). Lorsque l'on factorise ce faisceau en un nombre plus restreint de dimensions en calculant les composantes principales, on choisit comme première composante l'axe permettant d'expliquer la plus grande quantité d'informations, et qui forme une sorte de moyenne projetant chaque vecteur sur l'axe. On atteint ce but en traçant une droite allant de l'extrémité du vecteur à l'axe, perpendiculairement à celle du vecteur. Le rapport entre la longueur projetée sur l'axe à la longueur réelle du vecteur lui-même donne une mesure du pourcentage des informations d'un vecteur expliquées par l'axe. (Ce point est difficile à exposer, mais je pense que la figure 5.5 permet de dissiper tout malentendu.) Si un vecteur est situé près de l'axe, il est expliqué en très grande partie et l'axe englobe la plupart de ses informations. Au fur et à mesure que le vecteur s'éloigne de l'axe jusqu'à la séparation tante des informations du vecteur.

Nous plaçons la première composante principale (ou axe) de manière à ce que, parmi tous les vecteurs, elle explique plus d'informations que tout autre axe ne le pourrait. Pour notre matrice de hauts coefficients de corrélation positive représentés par un ensemble de vecteurs en faisceau serré, la première composante principale traverse le centre de l'ensemble (fig. 5.4). La deuxième composante principale est située perpendiculairement à la première et explique une quantité maximale des informations restantes. Mais si la première composante a déjà expliqué la plupart des informations dans tous les vecteurs, le deuxième axe principal et tous ceux qui suivent ne peuvent traiter que la petite quantité d'informations restantes (fig. 5.4).

On trouve fréquemment dans la nature ce type de système aux corrélations positives élevées. Dans la première étude, par exemple, où

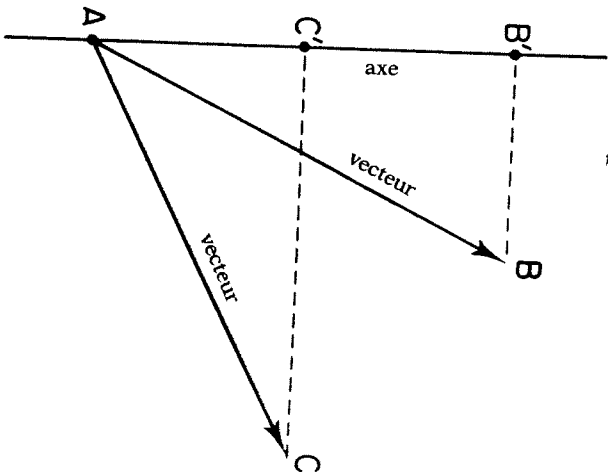
Ils réalisaient tous leurs calculs grâce à l'orientation des composantes principales et s'appelaient eux-mêmes des *factor analysts*. Je continue donc à utiliser le terme « analyse factorielle » dans son sens originel et j'y inclue toutes les orientations d'axes — que ce soit celles des composantes principales, qu'elles aient subi une rotation ou qu'elles soient orthogonales ou obliques.

J'emploierai aussi un raccourci commun, même s'il est peu rigoureux, lorsque j'abor-derai la fonction des axes factoriels. Techniquement, les axes factoriels expliquent la variance des mesures originales. Je dirai d'eux, comme on le fait souvent, qu'ils « expliquent » les informations — comme ils le font dans le sens courant (mais non technique) du mot information. C'est-à-dire que, quand le vecteur d'une variable originale se projette fortement sur un ensemble d'axes factoriels, seule une faible partie de sa variance reste inexpiquée dans des dimensions supérieures, à l'extérieur du système des axes factoriels.



5.4 Représentation géométrique des corrélations entre huit tests où tous les coefficients de corrélation sont élevés et positifs. La première composante principale, marquée 1, est proche de tous les vecteurs, alors que la seconde composante principale, marquée 2, est située à angle droit par rapport à la première et n'explique qu'une faible quantité d'informations dans les vecteurs.





5.5 Calcul de la quantité d'informations contenues dans un vecteur expliqué par un axe. Tracer un trait de l'extrémité du vecteur à l'axe, perpendiculairement à l'axe. La quantité des informations expliquées par l'axe est le rapport entre la longueur projetée sur l'axe et la véritable longueur du vecteur. Si un vecteur se trouve proche de l'axe, ce rapport est élevé et la plupart des informations du vecteur sont expliquées par l'axe. Le vecteur AB est proche de l'axe et le rapport entre la projection AB' et le vecteur lui-même, AB, est élevé. Le vecteur AC est éloigné de l'axe et le rapport entre sa longueur projetée AC' et le vecteur lui-même, AC, est faible.

je fis appel à l'analyse factorielle, j'avais pris quatorze mesures prélevées sur les ossements de vingt-deux espèces de pélycosaurès (reptiles fossiles identifiables au voile de peau, supporté par une rangée d'épines osseuses, qu'ils ont sur le dos ; on les confond souvent avec les dinosaures, mais ce sont en fait les ancêtres des mammifères). Ma première composante principale expliqua 97,1 % des informations sur l'ensemble des quatorze vecteurs, ne laissant que 2,9 % sur les axes suivants. Mes quatorze vecteurs formaient un amas extrêmement dense (se chevauchant pratiquement tous) ; le premier axe traversait le centre du faisceau. La longueur du corps de mes pélycosaurès allait de moins de soixante centimètres à plus de trois mètres trente. Ils se ressemblaient tous beaucoup et les gros animaux avaient des mensurations plus grandes pour la totalité des quatorze os. Tous les coefficients de corrélation des os entre eux étaient très élevés ; en fait, le plus faible était un énorme .912 [lire 0,912]. Ce qui n'est guère surprenant. Après tout, les gros animaux ont de gros os et les petits de petits os. Je peux interpréter ma première composante principale comme un facteur de taille résumé, condensant ainsi (avec une perte minimale d'informations) mes quatorze mesures originales en une seule dimension interprétée comme un accroissement de la taille du corps. Dans ce cas, l'analyse factorielle a permis à la fois la *simplification* en réduisant le nombre de dimensions (en passant de quatorze à une) et l'*explicitation* par une interprétation biologique raisonnable du premier axe comme facteur de taille.

Oui mais — et quel mais ! — avant de nous réjouir et d'exalter l'analyse factorielle et d'en faire la panacée permettant de comprendre les systèmes complexes de corrélation, il nous faut reconnaître que son emploi est soumis aux mêmes précautions et est sujet aux mêmes objections que celles qui sont apparues pour les coefficients de corrélation. Dans les sections suivantes, je vais traiter de deux problèmes majeurs.

#### L'ERREUR DE RÉIFICATION

La première composante principale est une abstraction mathématique qui peut être calculée pour chaque matrice de coefficients de corrélation ; ce n'est pas une « chose » possédant une réalité physique. Les « factorielistes » sont souvent tombés dans le piège tentateur de la *réification*, c'est-à-dire qu'ils ont accordé une *signification physique* à toutes les fortes composantes principales. Parfois cela est justifié : je pense que c'est à juste titre que j'ai interprété mon premier axe pélycosaurien comme un facteur de taille. Mais une affirmation de ce type ne peut jamais venir des mathématiques seules, il y faut nécessairement l'apport de connaissances sur la nature physique des mesures elles-mêmes. Car les systèmes absurdes de corrélations possèdent tout

pareillement des composantes principales, et celles-ci peuvent fort bien expliquer plus d'informations que des composantes significatives ne le font dans d'autres systèmes. L'analyse factorielle d'une matrice de corrélations de cinq variables sur cinq, comprenant mon âge, la population du Mexique, le prix du gruycère, le poids de ma tortue apprivoisée et la distance moyenne entre les galaxies pendant ces dix dernières années fera apparaître une forte composante principale. Celle-ci — les corrélations étant fortement positives — expliquera probablement un pourcentage d'informations aussi important que le premier axe de mon étude sur les pélycosauruses. À la différence qu'il n'aura, dans ce cas, pas la moindre signification physique.

Dans les études sur l'intelligence, l'analyse factorielle a été appliquée aux matrices de corrélations des tests mentaux. On peut, par exemple, faire subir dix tests à cent personnes. Chaque insertion significative dans notre matrice de corrélations de dix sur dix est un coefficient de corrélation entre les notes de deux tests subis par chacune des cent personnes. On sait depuis l'apparition des tests mentaux — et cela ne surprendra personne — que la plupart de ces coefficients de corrélation sont positifs, c'est-à-dire que les personnes qui obtiennent de bons résultats dans un type d'épreuves tendent, en moyenne, à réussir également aux autres. La plupart des matrices de corrélations pour les tests mentaux renferment de manière prépondérante des corrélations positives. Cette observation fondamentale a servi de point de départ à l'analyse factorielle. Charles Spearman a pratiquement inventé la technique, en 1904, comme un outil servant à rechercher les causes à partir des matrices de corrélations des tests mentaux.

La plupart des coefficients de corrélation dans la matrice étant positifs, l'analyse factorielle doit montrer une première composante principale raisonnablement forte. Dès 1904, Spearman calcula indirectement cette composante et en tira une déduction capitale, et erronée, dont a beaucoup souffert l'analyse factorielle. Il la réiffa en en faisant une « entité » et tenta d'en donner une interprétation causale sans équivoque. Il l'appela *g*, ou intelligence générale, et imagina qu'il avait ainsi isolé une qualité unitaire de base de toute activité mentale cognitive, qualité qui pourrait s'exprimer sous la forme d'un nombre unique et serait susceptible d'être utilisée pour classer les individus sur une échelle unilinéaire selon leur valeur intellectuelle.

Le facteur *g* de Spearman — la première composante principale de la matrice de corrélation des tests mentaux — n'a jamais atteint le rôle prédominant que joue la première composante dans de nombreuses études sur la croissance (comme pour mes pélycosauruses). Au mieux, *g* explique 50 à 60 % de toutes les informations de la matrice de tests. Les corrélations entre les tests sont généralement beaucoup plus faibles que les corrélations entre deux organes d'un corps en croissance. Dans la plupart des cas, la corrélation la plus élevée dans une

matrice de tests est loin d'atteindre la valeur la plus basse de ma matrice pélycosaurienne, 912.

Bien que la force de *g* ne puisse jamais se comparer à celle de la première composante principale de certaines études sur la croissance, je ne considère pas son respectable pouvoir explicatif comme l'effet du hasard. Des raisons causales sous-tendent les corrélations positives de la plupart des tests mentaux. Mais quelles sont ces raisons ? Il n'est pas possible de déduire ces raisons d'une forte première composante principale, pas plus qu'il n'est possible de déduire la cause d'un seul coefficient de corrélation à partir de son ampleur. On ne peut pas réifier *g*, en faire une « chose » à moins d'être en possession d'informations convaincantes, indépendantes du fait de la corrélation elle-même.

La situation pour les tests mentaux ressemble à ce cas hypothétique, présenté plus haut, de la corrélation entre lancers de balle. La liaison est forte et nous avons tout lieu de la considérer comme non fortuite. Mais nous ne pouvons pas déduire la cause à partir de cette seule corrélation et cette cause est certainement complexe.

Le facteur *g* de Spearman est particulièrement sujet à des interprétations ambiguës, ne serait-ce qu'à cause du fait que les deux hypothèses causales les plus contradictoires sont parfaitement compatibles avec lui : 1) le *g* traduit un niveau héréditaire d'acuité mentale (certains réussissent bien la plupart des tests parce qu'ils sont nés plus intelligents) ou 2) le *g* enregistre les avantages et les préjudices du milieu (certains individus obtiennent de bons résultats à la plupart des tests parce qu'ils ont eu une bonne scolarité, ont été correctement alimentés durant leur jeunesse, ont été élevés par des parents pleins d'attention dans une maison où les livres ne faisaient pas défaut). Si l'existence du facteur *g* peut théoriquement s'expliquer d'une façon purement héréditariste ou purement environnementaliste, c'est que sa simple présence ne peut pas conduire à sa réification. La tentation de la réification est puissante. L'idée qu'on a détecté quelque chose de « sous-jacent », derrière un vaste ensemble de coefficients de corrélation, quelque chose de peut-être plus réel que les mesures superficielles elles-mêmes, est assez grisante. C'est l'essence de Platon, la réalité éternelle, abstraite, qui se cache derrière les apparences. Mais c'est une tentation à laquelle il nous faut résister, car elle est un vieux préjugé de notre entendement, et non l'expression d'une vérité de la nature.

#### LA ROTATION ET LA NON-NÉCESSITÉ DES COMPOSANTES PRINCIPALES

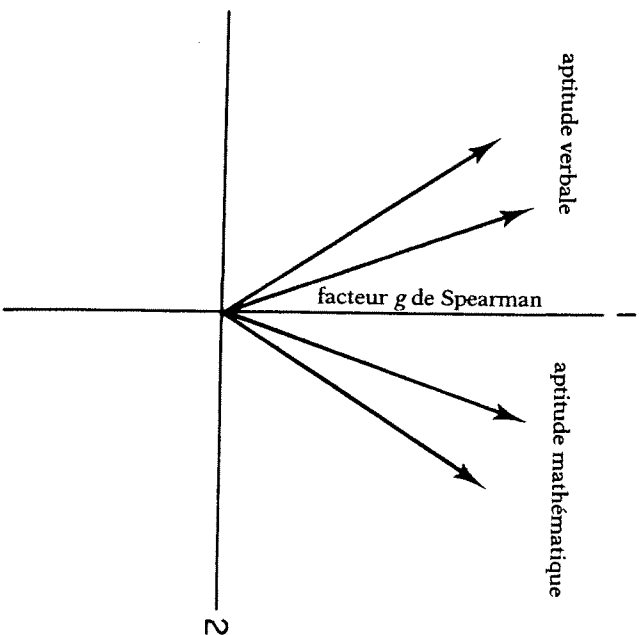
Un autre argument, plus technique, démontre clairement pourquoi les composantes principales ne peuvent pas être automatiquement réifiées comme entités causales. Si les composantes principales représentaient le seul moyen de parvenir à la simplification d'une

matrice de corrélations, on pourrait avec légitimité lui accorder un statut spécial. Mais elles ne représentent qu'une méthode parmi d'autres pour insérer des axes dans un espace pluridimensionnel. Les composantes principales ont une disposition géométrique précise, spécifiée par le critère utilisé pour les élaborer, à savoir que la première composante principale doit expliquer une quantité maximale des informations d'un ensemble de vecteurs et que les composantes suivantes doivent toutes être mutuellement perpendiculaires. Mais ce critère n'a rien de sacro-saint : les vecteurs peuvent être expliqués dans n'importe quel ensemble d'axes placés à l'intérieur de leur espace. Les composantes principales permettent dans certains cas de nous éclairer, mais d'autres critères se révèlent souvent plus utiles.

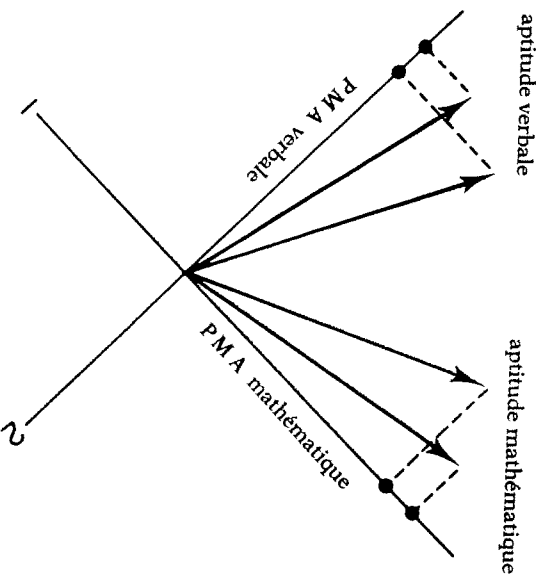
Considérons la situation suivante dans laquelle on peut être amené à préférer un autre mode de placement des axes. Dans la figure 5.6, je présente des corrélations entre quatre tests mentaux, deux sur l'aptitude verbale et deux sur l'aptitude arithmétique. Deux faisceaux (*clusters*) sont évidents, même si tous les tests sont positivement corrélés. Admettons que nous voulions isoler ces faisceaux par l'analyse factorielle. Si nous utilisons les composantes principales, il n'est pas certain que nous puissions les reconnaître. La première composante principale (le *g* de Spearman) passe pile au centre entre les deux faisceaux. Elle n'est proche d'aucun vecteur et explique une quantité approximativement égale de chacun d'eux, ce qui, par là même, masque l'existence des deux faisceaux. Cette composante est-elle une entité ? Une « intelligence générale » existe-t-elle ? Ou le facteur *g* n'est-il, dans ce cas, qu'une simple moyenne sans signification, basée sur l'amalgame erroné de deux types d'informations ?

On peut discerner le faisceau verbal et le faisceau arithmétique sur la seconde composante principale (appelée « facteur bipolaire », car certaines projections *y* sont positives et d'autres négatives lorsque les vecteurs sont situés de part et d'autre de la première composante principale). Dans ce cas, les tests verbaux se projettent sur le côté négatif de la deuxième composante et les tests arithmétiques sur le côté positif. Mais il est possible que nous ne parvenions pas du tout à déceler ces faisceaux si la première composante principale domine dans tous les vecteurs. Car les projections sur la deuxième composante seront alors faibles et la configuration peut fort bien ne pas apparaître (voir fig. 5.6).

Dans les années 1930, les analystes mirent au point des méthodes pour venir à bout de ce dilemme et pour reconnaître des faisceaux de vecteurs que les composantes principales cachent souvent. Ils y parvinrent en faisant pivoter les axes factoriels qui quittèrent leur orientation de composantes principales pour prendre de nouvelles positions. Ces rotations, établies par plusieurs critères, avaient pour but commun de placer les axes près des faisceaux. Dans la figure 5.7, par exemple, le critère utilisé place les axes près des vecteurs occupant une position extrême ou périphérique dans l'ensemble total. Si, main-



5.6 Analyse des principales composantes de quatre tests mentaux. Toutes les corrélations sont élevées et la première composante principale, le facteur *g* de Spearman, exprime la corrélation globale. Mais les facteurs de groupe pour l'aptitude verbale et l'aptitude mathématique ne sont pas bien expliqués dans ce type d'analyse.



5.7 Axes factoriels après leur rotation, pour les quatre mêmes tests mentaux présentés dans la figure 5.6. Les axes sont à présent placés près des vecteurs situés à la périphérie des faisceaux. Les facteurs de groupe pour l'aptitude verbale et l'aptitude mathématique sont à présent bien identifiés (voir les projections élevées sur les axes indiquées par les pointillés), mais *g* a disparu.  
 PMA = *Primary Mental Ability*, \* aptitude mentale primaire \*, notion proposée par Thurstone. Voir pp. 337 et suivantes.

tenant, nous « expliquons » tous les vecteurs sur ces axes pivotés, nous mettons aisément en évidence les faisceaux, car les tests arithmétiques ont une projection élevée sur l'axe 1 et faible sur l'axe 2, tandis que les tests verbaux ont une projection forte sur 2 et faible sur 1. Mais, en outre, *g* a disparu. On ne trouve plus de « facteur général » d'intelligence, plus rien qui ne puisse être réifié comme un nombre unique exprimant l'aptitude globale. Et cependant on ne perd pas d'information. Après leur rotation les deux axes expliquent autant d'informations dans les deux vecteurs qu'ils le faisaient dans leur position antérieure de composantes principales. Ils se contentent de distribuer différemment les mêmes informations sur les axes. Comment alors prétendre que *g* est une entité réelle s'il ne représente qu'un des nombreux moyens de placer les axes par rapport à un ensemble de vecteurs ?

En bref, l'analyse factorielle simplifie de grands tableaux de données en réduisant le nombre de dimensions et, en échange de la perte de quelques informations, permet de reconnaître l'ordonnement d'une structure grâce à cette réduction du nombre de dimensions. En tant qu'outil simplificateur, l'analyse factorielle a rendu de grands services dans de nombreuses disciplines. Mais beaucoup d'analyses ont été au-delà de la simplification et ont été tentées de définir les facteurs comme des entités causales. L'erreur que constitue cette réification a faussé la technique dès sa naissance. Elle était « présente à la création » lorsque Spearman inventa l'analyse factorielle pour étudier la matrice de corrélation des tests mentaux et réifia sa première composante principale sous la forme du facteur *g*, ou intelligence générale innée. L'analyse factorielle peut nous aider à comprendre les causes en nous guidant vers des informations situées au-delà du caractère mathématique des corrélations. Mais les facteurs, en eux-mêmes, ne sont ni des choses ni des causes ; ce sont des abstractions mathématiques. Le même ensemble de vecteurs (voir fig. 5.6 et 5.7) pouvant être divisé en un facteur *g* et un petit axe résiduel, d'une part, ou bien en deux axes de force égale qui, sans faire du tout appel à *g*, isolent deux faisceaux distincts, l'un verbal, l'autre arithmétique, on ne peut pas prétendre que l'« intelligence générale » de Spearman soit une entité inéluctable qui sous-tende et explique nécessairement les corrélations entre les tests mentaux. Même si nous jugeons bon de considérer que *g* n'est pas le résultat du hasard, ni sa force ni sa position géométrique ne permettent de préciser ce qu'il signifie en termes de causalité — ne serait-ce que parce que ses caractéristiques s'accordent aux opinions extrémistes sur l'intelligence, qu'elles soient héréditaristes ou environnementalistes.

### Charles Spearman et l'intelligence générale

#### LA THÉORIE BIFACTORIELLE

Les coefficients de corrélation pullulent à présent comme les cafards à New York. D'une simple pression du doigt, on en produit à l'aide de la moindre calculatrice de poche. Indispensable ou non, on les considère comme faisant automatiquement partie de l'équipement de toute analyse statistique dès que plus d'un paramètre est en jeu. Dans un tel contexte, on oublie aisément qu'ils furent jadis acclamés comme une découverte sensationnelle dans le domaine de la recherche, comme un nouvel outil passionnant permettant de découvrir les structures cachées sous les mesures brutes. On peut se faire une idée de cette vague d'enthousiasme en lisant les premiers articles du grand biologiste et statisticien américain Raymond Pearl (voir Pearl, 1905 et 1906, et Pearl et Fuller, 1905). Il soutint sa thèse de docteur au chantement de siècle, puis continua, tel un enfant heureux de son nouveau jouet, à corriger tout ce qui lui tombait sous la main, de la longueur des vers de terre avec le nombre de ses anneaux (où il ne trouva pas de corrélation et supposa que l'accroissement de la longueur dépendait de l'allongement de chaque segment et non d'une augmentation de leur nombre), à la taille de la tête humaine avec l'intelligence (où il découvrit une très faible corrélation, mais qu'il attribua à l'effet indirect d'une meilleure alimentation).

Charles Spearman, éminent psychologue et statisticien brillant\*, commença à étudier les corrélations entre tests mentaux à cette époque envivante. Si l'on fait passer deux tests mentaux à un grand nombre d'individus, Spearman remarqua que les coefficients de corrélation entre eux étaient presque toujours positifs. Spearman réfléchit à ce résultat en se demandant quelle généralité plus élevée il pouvait bien impliquer. Les corrélations positives indiquaient nettement que chaque test ne mesurait pas un attribut indépendant du fonctionnement mental. Une structure plus simple se cachait derrière ces corrélations positives envahissantes ; mais quelle structure ? Spearman imagina deux propositions alternatives. En premier lieu, les corrélations positives pouvaient se réduire à un petit nombre d'attributs indépendants, les « facultés » de la phrénologie et autres écoles des

\* Spearman s'intéressait tout particulièrement aux problèmes des corrélations et inventa une mesure qui, après le  $r$  de Pearson, est la plus utilisée de celles où deux variables sont associées, le coefficient de corrélation des rangs ou coefficient de corrélation de Spearman.

La véritable erreur de Cyril Burr

297

débuts de la psychologie. L'esprit comportait peut-être des « compartiments » séparés pour les aptitudes arithmétique, verbale et spatiale, par exemple. Spearman appela ces théories de l'intelligence « oligarchiques ». En second lieu, les corrélations positives pouvaient se réduire à un seul facteur sous-jacent, notion que Spearman qualifia de « monararchie ». Dans un cas comme dans l'autre, il reconnut que les facteurs sous-jacents, qu'ils soient en petit nombre (oligarchiques) ou unique (monararchie), n'englobaient pas toutes les informations d'une matrice de coefficients de corrélation positifs pour une grande quantité de tests. Il restait une « variance résiduelle », constituée par les informations particulières à chaque test et sans liaison les unes avec les autres. En d'autres termes, chaque test avait sa composante « anarchique ». Spearman appela la variance résiduelle de chaque test, le  $s$ , ou informations spécifiques. Ainsi, selon le raisonnement de Spearman, l'étude de la structure sous-jacente pouvait conduire à deux solutions. Ou bien à une « théorie bifactorielle » dans laquelle chaque test renferme des informations spécifiques (son  $s$ ) et traduisait les effets d'un facteur sous-jacent unique que Spearman appela  $g$  ou intelligence générale. Ou alors chaque test pouvait comprendre ses propres informations spécifiques ainsi qu'une ou plusieurs facultés prises parmi un ensemble de facultés sous-jacentes indépendantes, ce qui conduisait à une théorie plurifactorielle. Si la plus simple des deux théories, la bifactorielle, était validée, tous les attributs ordinaires de l'intelligence se résumeraient à une seule entité, véritable « intelligence générale » dont la mesure pourrait fournir un critère permettant de classer sans équivoque les individus selon leur valeur mentale.

Charles Spearman mit au point l'analyse factorielle — qui reste de nos jours la technique la plus importante dans le domaine des statistiques à variables multiples — comme un procédé destiné à trancher entre les deux théories, la bifactorielle d'une part et la plurifactorielle de l'autre, en déterminant si la variance commune dans une matrice de coefficients de corrélation pouvait être ramenée à un facteur « général » unique ou à seulement plusieurs facteurs « de groupes » indépendants. Ne trouvant qu'une « intelligence » unique, il choisit la théorie bifactorielle et, en 1904, publia un article qui, plus tard, reçut, comme titre : « Aucun événement dans l'histoire des tests mentaux ne s'est révélé d'une importance aussi capitale que la proposition faite par Spearman avec sa célèbre théorie bifactorielle » (Guilford, 1936, p. 155). Dans son exultation, et avec l'immodestie qui le caractérisait, Spearman donna à son article ce titre grandiloquent : « L'intelligence générale objectivement mesurée et déterminée ». Dix ans plus tard (1914, p. 237), il ne se tenait plus de joie : « L'avenir de la recherche dans le domaine de l'hérédité des aptitudes doit se concentrer sur la théorie des "deux facteurs" ». Elle seule semble en mesure de réduire le chaos déconcertant des faits en un ordonnancement évident. Par son

entremise, les problèmes sont éclaircis ; à maints égards, on entrevoit déjà leurs réponses ; et partout, on envisage la perspective d'une solution décisive. »

#### LA MÉTHODE DES DIFFÉRENCES TÉTRADES

Au cours de ses premiers travaux, Spearman n'a pas employé la méthode des composantes principales décrite plus haut. Il a d'abord mis au point un procédé plus simple, bien que plus fastidieux, mais mieux adapté à une époque où l'ordinateur était inconnu et où tous les calculs devaient être faits à la main\*. Il calcula toute la matrice des coefficients de corrélation entre deux tests, prit tous les groupements possibles de quatre mesures et calcula, pour chaque nombre, ce qu'il appelait la « différence tétrade ». L'exemple suivant illustre ce qu'est la différence tétrade et explique comment Spearman l'utilisa pour décider si la variance commune d'une matrice pouvait se réduire à un facteur général ou seulement à plusieurs facteurs de groupe.

Supposons que l'on veuille calculer la différence tétrade portant sur quatre mesures relevées sur une série de souris allant des nouveau-nés aux adultes : longueur des pattes, grosseur des pattes, longueur de la queue et grosseur de la queue. On calcule tous les coefficients de corrélation par couples de variables et l'on trouve, ce qui n'a rien de surprenant, que tous sont positifs — au fur et à mesure que les souris grandissent, leurs organes deviennent plus gros. Mais on aimerait savoir si la variance dans les corrélations positives est le fait d'un seul facteur général — la croissance elle-même — ou si deux composantes séparées peuvent être isolées, en l'occurrence un facteur pattes et un facteur queue ou bien un facteur longueur et un facteur grosseur. Spearman donne de la différence tétrade la formule suivante :

$$r_{13} \times r_{24} - r_{23} \times r_{14}$$

où  $r$  est le coefficient de corrélation et où les deux indices représentent les deux mesures corrélées (ici 1 est la longueur — L — des pattes, 2 la grosseur — G — des pattes, 3 la longueur de la queue et 4 la grosseur de la queue ;  $r_{13}$  est donc le coefficient de corrélation entre la première et la troisième mesure, c'est-à-dire entre la longueur des pattes et la longueur de la queue). Dans notre exemple, la différence tétrade est :

(L pattes et L queue)  $\times$  (G pattes et G queue) — (G pattes et L queue)  $\times$  (L pattes et G queue).

Selon Spearman, des différences tétrades de zéro impliquaient

\* Le  $g$  calculé par la formule tétrade est conceptuellement équivalent et mathématiquement presque équivalent à la première composante principale décrite en pp. 289 et suivantes que l'on utilise actuellement dans l'analyse factorielle.

l'existence d'un facteur général unique alors que des valeurs soit positives soit négatives indiquaient la présence de facteurs de groupe. Supposons, par exemple, que ce soit des facteurs de groupe pour la longueur générale du corps et pour la grosseur générale du corps qui régissent la croissance des souris. On obtiendrait alors une valeur positive élevée de la différence tétrade, car les coefficients de corrélation entre une longueur et une autre longueur, ou entre une grosseur et une autre grosseur, tendraient à être plus élevés que les coefficients de corrélation entre une grosseur et une longueur. (Remarquez que le côté gauche de l'équation tétrade ne comprend que des longueurs avec des longueurs ou des grosseurs avec des grosseurs, alors que le côté droit ne comprend que des longueurs avec des grosseurs.) Mais si un seul facteur général de croissance détermine la taille des souris, les longueurs et les grosseurs doivent présenter entre elles une corrélation aussi élevée que les longueurs avec les longueurs ou que les grosseurs avec les grosseurs, et la différence tétrade doit être de zéro. La figure 5.8 montre une hypothétique matrice de corrélations pour les quatre mesures où la différence tétrade est de zéro. (Les valeurs sont tirées d'un exemple donné par Spearman dans un autre contexte, 1927, p. 74.) La figure 5.8 montre aussi une matrice hypothétique différente : avec celle-ci la différence tétrade est positive et, si d'autres tétrades présentent la même tendance, on sera amené à reconnaître la présence de facteurs de groupe pour la longueur et la grosseur.

La matrice du haut de la figure 5.8 illustre un autre point important qui s'est répercuté dans toute l'histoire de l'analyse factorielle en psychologie. Remarquez que, malgré la différence tétrade de zéro, les coefficients de corrélation n'ont pas besoin d'être (et presque invariablement ne sont pas) égaux. Ici, la grosseur des pattes avec la longueur des pattes donne une corrélation de .80 alors qu'elle n'est que de .18 pour la grosseur de la queue par la longueur de la queue. Ces différences traduisent des « saturations » variées en  $g$ , le facteur général unique lorsque les différences tétrades sont de zéro. Les mesures des pattes ont des saturations plus élevées que les mesures de la queue, c'est-à-dire qu'elles sont plus proches de  $g$  ou qu'elles le traduisent mieux (en termes modernes, elles sont situées plus près de la première composante principale dans les représentations géométriques telles que celle de la figure 5.6). Les mesures de la queue sont moins fortement saturées en  $g^*$ . Elles renferment peu de variance commune et s'expliquent surtout par leur teneur en  $s$  (informations spécifiques à chaque mesure). Revenons maintenant aux tests mentaux : si  $g$  représente l'intelligence générale, les tests mentaux les plus chargés en  $g$

\* Le terme « saturation » fait référence à la corrélation entre un test et un axe factoriel. Si un test est fortement saturé en un facteur, c'est que la plupart de ses informations sont prises en compte (expliquées) par ce facteur.