

Chapitre 4

METHODES D'ETUDE DES DIFFERENCES INDIVIDUELLES

Juhel, J. (1999). Méthodes d'étude des différences individuelles. In Rossi, J.-P. *et coll.* (1999). *Méthodes d'étude en psychologie* (pp.197-234). Paris: Dunod.

Juhel, J. (2000). Metodos de estudo das diferenças individuais. In J.P. Rossi e Col., *Os métodos de investigação em psicologia* (pp.229-270). Instituto Piaget.

Dire qu'il existe des différences dans les comportements humains et que les hommes sont psychologiquement différents est une évidence. Se fixer pour objectifs d'identifier les différences psychologiques observées entre individus ou entre groupes d'individus et de comprendre les mécanismes pouvant produire ces différences est une tout autre entreprise qui nécessite des stratégies de recherche et des méthodes d'étude spécifiques. Cette ambition est celle de la psychologie différentielle, une branche de la psychologie dont les premiers développements remontent à la fin du XIX^e siècle (Huteau, 1995; Reuchlin, 1997).

1. L'ETUDE DES DIFFERENCES INDIVIDUELLES

Issues toutes les deux du courant scientifique de la psychologie, la psychologie expérimentale générale et la psychologie différentielle ont longtemps été définies par la méthode à laquelle l'une et l'autre ont recours. L'option méthodologique de la première est l'expérimentation c'est-à-dire la manipulation en laboratoire de variables indépendantes de l'individu comme par exemple la durée de présentation ou la longueur d'une liste de mots à mémoriser. L'objectif est d'expliquer les conditions d'apparition d'un phénomène psychologique donné en imputant aux seules manipulations expérimentales (les variables provoquées) la cause des effets observés. La validité des conclusions de l'expérience dépendant du soin apporté au « contrôle » des effets parasites de variables liées à la situation (au contexte, à la tâche, au matériel) ou à l'individu, ces variations indésirables sont considérées comme une forme de bruit qu'il est nécessaire de supprimer. Les variations interindividuelles notamment sont neutralisées en répartissant aléatoirement les sujets entre les différentes conditions expérimentales. C'est cette stratégie qu'adoptent les psychologues dont l'objectif est de trouver des lois générales du comportement et qui, estimant mal connaître les raisons de ces différences ou inutile de s'en préoccuper, jugent préférable de les neutraliser.

L'étude des différences individuelles s'est développée parce que certaines de ces différences paraissent survenir de manière relativement stable et cohérente chez des individus dont les conduites sont observées dans des situations diverses. Certains individus ont tendance, plus que d'autres, à réussir des tâches intellectuelles complexes, à prendre la parole en public, à s'inquiéter pour leur santé, etc. L'objet de la psychologie différentielle est donc de décrire ces différences, de prendre l'ampleur de leur dispersion. Il est aussi d'identifier les variables qui permettent d'en fournir une description résumée et/ou auxquelles on croit pouvoir attribuer la responsabilité des différences observées. Ces variables peuvent par exemple être l'âge, le sexe, les conditions de vie, etc. Mais c'est surtout à des variables individuelles latentes comme l'intelligence, les aptitudes, l'anxiété, l'extraversion, etc., que se sont intéressés les différentialistes. Pouvant prendre des valeurs différentes chez des individus différents ou chez un même individu dans des contextes apparemment semblables mais perçus différemment, ces variables individuelles sont sans doute moins assurées au plan définitoire que ne le sont les facteurs manipulés par l'expérimentaliste ou certaines variables invoquées comme l'âge ou le sexe. Leur mesure pose aussi plus de problèmes que n'en pose le codage des modalités des variables indépendantes d'une expérience de laboratoire. Pour ces raisons, les recherches en psychologie différentielle se sont d'abord attachées à définir, par la mesure et la description des relations qu'ils entretiennent entre eux, les aspects psychologiques sous lesquels les individus diffèrent de manière relativement systématique.

1.1. Prendre la mesure des différences, décrire l'organisation des conduites observables

Imaginons une situation scolaire dans laquelle des enfants d'une même classe de Cours Préparatoire (CP) doivent recopier un texte écrit au tableau. Les différences observées peuvent porter sur le temps (plus ou moins long) de réalisation de la tâche, le nombre (plus ou moins grand) de fautes de copie, la fréquence (plus ou moins élevée) des allers et retours du regard entre la feuille et le tableau, etc. Des différences inte-

individuelles dans par exemple la compréhension du texte, l'intérêt éprouvé à l'égard de la tâche ou la crainte de se tromper pourront également être mises en évidence en interrogeant les enfants d'une manière permettant de « comparer » les réponses fournies. On voit que selon le ou les critères de différenciation choisis par l'observateur, les différences ne sont pas identifiables de la même façon et peuvent renvoyer à des niveaux d'observation distincts. Ces enfants ne mettent généralement pas le même temps pour copier le texte; ils n'en ont pas une égale compréhension et n'éprouvent pas tous nécessairement un même intérêt pour la tâche qui leur est proposée. Mais si les aspects manifestes ou latents sous lesquels les enfants de cet échantillon particulier diffèrent sont multiples, certaines des différences observées relativement à un critère de différenciation déterminé (*e.g.* temps mis pour copier le texte) peuvent être assez proches de celles observées relativement à un ou plusieurs autres critères (*e.g.* le nombre d'allers et retours du regard entre la feuille et le tableau ou le nombre de fautes de copie). Une même catégorie descriptive, une même « dimension » du comportement (Reuchlin, 1963), peut alors servir à définir les différences interindividuelles relatives à plusieurs critères convergents.

Parallèlement à l'essor, auquel ils ont largement contribué, des méthodes de la mesure en psychologie (la psychométrie), les psychologues différentalistes se sont efforcés d'élaborer des instruments adaptés à la mesure des variations interindividuelles dans les conduites observables (les tests) afin d'en décrire l'ampleur chez les individus d'une catégorie déterminée. Grâce à l'emploi de l'analyse factorielle, une méthode d'analyse des associations entre variables dont nous décrirons les caractéristiques et les principes d'application plus loin (cf. section 4.1.1), les différentalistes ont aussi tenté d'élaborer des modèles « théoriques » visant à représenter formellement la structure des covariations entre les dimensions fondamentales de la conduite.

Ces recherches à la fois descriptives et inductives ont d'abord été effectuées à propos de l'organisation générale, semblable chez tous les individus, des conduites intellectuelles observables dans les tests d'intelligence. Elles ont été étendues par la suite à l'étude des intérêts professionnels puis à celle des dimensions de la personnalité. On se limitera ici à indiquer que ces recherches ont permis de donner de chacune de ces constructions psychologiques classiques une représentation générale à la fois complexe et hiérarchiquement ordonnée (pour revue Huteau, 1995; Reuchlin, 1997). La complexité de la représentation de l'intelligence par exemple provient du nombre des dimensions ou facteurs identifiés. Le fait que ceux-ci soient situés à des niveaux de généralité différents ordonne cette représentation. Lorsque l'échantillonnage des épreuves est suffisamment varié, il est ainsi possible de distinguer des facteurs d'intelligence fluide (capacité de raisonnement abstrait), d'intelligence cristallisée (compétences verbales), de visualisation spatiale (capacité à percevoir, mémoriser, transformer l'information visuo-spatiale), de connaissance, d'apprentissage et de mémoire, de production des idées (ou créativité), de vitesse perceptive et cognitive. Cette représentation hiérarchique possède aussi un caractère unitaire dû à l'interdépendance des dimensions identifiées, interdépendance alors représentée par l'hypothèse d'une dimension commune à toutes les conduites intellectuelles observées dans les situations échantillonnées. D'un très haut niveau de généralité, cette dimension commune est appelée facteur général.

Essentiellement utilisée à des fins pratiques, notamment dans les domaines de l'orientation scolaire, de la sélection des personnels industriels et militaires ou de la psychiatrie, la psychologie différentielle n'a cependant pas connu le même succès au plan théorique. S'il est vrai que les différences observées, pourvu qu'elles aient un caractère suffisant de stabilité et de cohérence, semblent témoigner de l'existence de certains mécanismes psychologiques sous-jacents et que les individus peuvent être différenciés en référence à des facteurs organisant les conduites observables, cette première génération de travaux s'est en effet révélée assez peu capable d'aller au-delà de l'identification normative des sources de variation interindividuelle. Il faut ajouter qu'en partie à cause de leurs évidentes implications sociales, les questions théoriques les plus assidument explorées par les différentalistes classiques (l'étude des facteurs de milieu et des facteurs génétiques de la différenciation interindividuelle, celle des différences entre sexes, classes sociales, « races » dans le domaine de l'intelligence, etc.) ont été l'objet de très vives controverses. La psychologie différentielle a donc

été longtemps assimilée à une forme de psychologie appliquée susceptible d'offrir quelques possibilités de prédiction au praticien mais d'un apport théorique extrêmement réduit voire même pour certains, idéologiquement contestable.

1.2. Les évolutions dans l'étude des différences individuelles

Il faudra attendre les années 60 pour commencer à voir évoluer le statut des différences individuelles en psychologie, le coup d'envoi en étant donné par l'américain Lee Cronbach dans un texte soulignant la complémentarité des méthodes expérimentale et différentielle. A l'heure actuelle, de nombreuses recherches expérimentales comme celles étudiant par exemple les processus psychologiques mis en œuvre dans l'acte lexique ou la compréhension de textes, s'appuient sur une différenciation des sujets par rapport à des variables interindividuelles mesurant le niveau intellectuel, la sensibilité phonologique (ou conscience qu'a l'enfant de la structure phonologique du langage), le niveau de compréhension de lecture, etc. C'est en appliquant conjointement ces deux méthodes à l'étude cognitive de la lecture qu'il a ainsi été possible de montrer que la conscience phonologique mesurée chez l'enfant de moyenne section de maternelle est un prédicteur valide des performances réalisées lors du début de l'apprentissage de la lecture et que le lecteur accompli rappelle et perçoit mieux les structures phonologiques que le moins bon lecteur.

C'est à peu près à cette même époque qu'en France, Maurice Reuchlin, à qui on doit d'avoir impulsé des recherches différentielles sur le développement intellectuel de l'enfant (Larivée *et al.*, 1996), souligne la portée de l'étude des différences individuelles pour l'élaboration théorique en psychologie, l'intérêt à prendre en compte des différences individuelles dans les conduites pour aménager ou remettre plus fondamentalement en question les modèles généraux élaborés à partir de conduites « moyennes ». L'aménagement du modèle Piagetien par les psychologues cherchant à intégrer en une approche développementale différentielle la perspective génétique opératoire au point de vue différentialiste en est une excellente illustration. Les résultats issus de ces travaux importants montrent en particulier que le développement cognitif est multidimensionnel et que, comme le suggère Lautrey (1990), des modes privilégiés de traitement de l'information sont associés à des différences individuelles liées à la disponibilité de processus vicariants (Reuchlin, 1978) pouvant fournir une même réponse adaptative ainsi qu'à des différences situationnelles. Un nombre croissant de psychologues voit d'ailleurs dans les variations observées chez un même individu au cours du temps (la variabilité intra-individuelle) un principe explicatif général du fonctionnement et du développement individuels (Lautrey, 1997; Ohlmann, 1995; Richelle, 1995).

Au plan méthodologique, le développement de nouvelles méthodes dites de modélisation structurale (Bacher, 1987, 1988; Dickès, 1994) a considérablement enrichi les possibilités de formulation et de mise à l'épreuve d'hypothèses fonctionnelles pouvant rendre compte des différences psychologiques entre individus. Reprenons, pour illustrer ce point essentiel, l'exemple des différences observées dans le temps mis par des enfants pour copier un texte écrit au tableau. Si ces différences sont en relation, à des degrés divers, avec des différences de compétence en lecture, des différences dans la crainte de mal faire, des différences dans l'intérêt éprouvé à l'égard de la tâche, etc. ou certains aspects contextuels, elles dépendent plus fondamentalement de la manière avec laquelle toutes ces variables interagissent. Des différences dans par exemple l'intérêt porté en général aux situations scolaires ont pu contribuer à produire des différences dans la compétence en lecture. Ces dernières peuvent elles-mêmes compenser ou au contraire accentuer les différences dans la crainte de mal faire. Quant à la crainte de mal faire, elle est sans doute accrue, toutes choses restant égales par ailleurs, par la présence d'un observateur étranger à la classe. Une explication simpliste qui verrait dans telle ou telle variable fondamentale, tel ou tel processus de base, l'origine des différences observées dans le temps mis par les enfants pour copier le texte ne peut donc satisfaire. Si l'on ambitionne de comprendre comment sont générées ces différences, il est alors préférable d'étudier le système de variables dont le fonctionnement contribue à produire les différences observées. Comme nous le verrons plus loin (section 4.3), les modèles structuraux offrent à cet égard des perspectives particulièrement intéressantes.

2. LA MESURE DES DIFFERENCES INDIVIDUELLES

Le psychologue qui étudie les différences entre individus ou celles observées chez un même individu à différentes occasions doit d'abord recueillir des observations en relation avec les constructions psychologiques auxquelles il s'intéresse (la conscience phonologique chez l'enfant du début de CP, l'anxiété chez le patient hospitalisé, les perceptions de contrôle chez la personne âgée, etc.). Pour « mesurer » ces constructions psychologiques, il fait appel à des procédures qui opérationnalisent une certaine façon d'en définir les aspects pertinents. Ces aspects peuvent éventuellement différer selon la situation ou l'échantillon considérés. L'anxiété pourra par exemple être mesurée différemment selon que les individus assistent à un match de football, passent un examen à l'Université, ou s'approprient à subir une intervention chirurgicale. De la même manière, on ne mesurera pas la conscience phonologique de la même façon chez des élèves de moyenne section ou de CP. Les procédures utilisées pour mettre en évidence les différences individuelles devant nécessairement être contrôlées, les tests qu'emploie à cet usage la psychologie différentielle sont aussi construits en fonction de principes techniques et méthodologiques dont certains valent également pour les tâches de laboratoire et épreuves que les recherches différentielles contemporaines leur préfèrent parfois.

2.1. Des observations à leur représentation

Comme indiqué précédemment, l'utilisation d'une épreuve faisant apparaître des différences individuelles permet au psychologue de recueillir des données d'observation qui, une fois représentées sous une forme conventionnelle, fournissent un contenu empirique à la construction psychologique étudiée. Cette opération de mesure, c'est ainsi qu'on l'appelle, mérite d'être brièvement analysée car elle implique un ensemble de choix dont il faut bien prendre conscience.

Que faut-il en effet entendre par mesure d'une variable psychologique? On peut très brièvement tenter de répondre à cette question essentielle en disant avec Dickès et ses collègues (1994) que la mesure d'une variable psychologique implique de: a) choisir théoriquement la variable à mesurer en fonction de l'ensemble des propositions théoriques qui concernent ses relations avec d'autres constructions psychologiques; b) spécifier les observations pertinentes, ce qui pose le problème du choix du dispositif et de l'échantillonnage des items; c) transformer les observations (les réponses aux différents items de l'épreuve) en informations analysées par un modèle de mesure définissant certaines propriétés des observations qui permettent de décrire correctement l'organisation des données. Le problème fondamental de la validité conceptuelle de l'épreuve (Juhel, 1994) est présent tout au long des deux premières étapes: recherche de validation nomologique ou inter-concepts lors de la première étape, recherche de validation interne ou intra-concept lors de la seconde. La dernière étape de passage de données d'observation à des données provenant d'observations (des scores) est également importante car la représentation des données d'observation peut différer selon le modèle - l'échelle - de mesure qui leur est appliqué.

La manière de représenter par des nombres des données d'observations ne sera en effet pas la même selon que l'on emploiera une échelle nominale (représentation sous forme de catégories différentes et exclusives: par exemple le genre), une échelle ordinale (représentation sous forme ordonnée ou hiérarchisée: par exemple des réponses à une échelle de préférences) ou une échelle d'intervalles (représentation sous forme ordonnée postulant l'égalité des écarts entre deux éléments ordonnés consécutifs: par exemple le nombre d'erreurs calculé sur une suite d'exercices ou le temps mis pour répondre). C'est parce que les différences sont relatives au niveau d'approximation de la mesure que les choix effectués à chacune de ces différentes étapes conditionnent fondamentalement la manière avec laquelle ces différences peuvent être représentées et, en conséquence, comprises.

2.2. Dispositifs de mesure

L'univers des instruments permettant de constater l'existence de différences individuelles est bien évidemment infini. Il comporte les multiples épreuves que les psychométriciens ont construites pour mesurer les différences interindividuelles dans un nombre extrêmement varié de domaines. C'est le cas notamment des tests d'intelligence, de mémoire, d'attention, ou des questionnaires de personnalité, d'anxiété, de dépression. Certains paradigmes expérimentaux élaborés pour étudier des processus psychologiques généraux, par exemple dans le domaine cognitif, peuvent également être employés pour mettre en évidence des différences inter- et intra-individuelles. Les résultats obtenus à l'aide d'un même instrument, un test de raisonnement par exemple, n'auront cependant pas la même signification si celui-ci est appliqué aux sujets expérimentaux d'une étude sur la compréhension de textes, à des candidats à la recherche d'un emploi, à des patients cérébro-lésés ou à des personnes âgés. Il faut donc bien insister sur le fait que la signification attachée à un instrument quel qu'il soit, dépend des individus à qui on l'administre et du cadre interprétatif dans lequel le psychologue juge théoriquement pertinent d'insérer la construction psychologique (ou le processus) ainsi opérationnalisée.

Très succinctement, deux grandes catégories de tests visant à mesurer les différences interindividuelles peuvent être distinguées selon que les instruments élaborés sont employés dans l'étude des différences dans le domaine intellectuel ou dans celle des différences dans les modes d'orientation et de contrôle des conduites (Reuchlin, 1997). A passation individuelle ou collective, ces tests, le plus souvent au format papier-crayon, peuvent être parfois informatisés.

A la première catégorie correspondent plutôt les tests d'intelligence (Huteau et Lautrey, 1997), les batteries d'aptitudes, les tests de créativité et de raisonnement, les tests utilisés dans le domaine de l'éducation (tests de lecture, de mathématiques, de connaissances scolaires, etc.), les échelles de développement intellectuel ou certains tests employés en neuropsychologie (tests de mémoire, du fonctionnement exécutif). On pourrait ranger dans la seconde catégorie les inventaires d'auto-évaluation de la personnalité, les mesures de motivation, d'intérêts, de valeurs et d'attitudes, les tests mesurant les styles personnels ou styles de conduite (Huteau, 1987) ainsi que divers instruments utilisés le plus souvent en psychologie de la santé.

Longtemps conçus en fonction de considérations pratiques par rapport à des critères externes, ces instruments reposent assez explicitement sur l'hypothèse de stabilité des différences interindividuelles. Leurs constructeurs s'attendent notamment à ce qu'une même épreuve administrée à plusieurs reprises dans le temps ou des épreuves « parallèles » appliquées aux mêmes individus produisent les mêmes variations. Guidée par des principes méthodologiques et techniques, la mise au point de ces dispositifs vise donc essentiellement à améliorer la qualité de mesure de la construction psychologique opérationnalisée par le test (*cf.* chapitre 5).

Des différences individuelles peuvent aussi être observées avec des paradigmes n'étant pas, comme les tests, spécifiquement construits pour les mettre en évidence mais conçus plutôt pour tester une hypothèse explicative d'un phénomène psychologique donné. Certains de ces paradigmes expérimentaux peuvent donc être utilisés dans l'étude des différences individuelles d'autant que les modèles auxquels ils font référence, souvent ceux de la psychologie cognitive, sont solidement définis.

Cooper (1976) s'interroge ainsi sur les processus mis en œuvre dans la comparaison perceptive de formes simples. Faisant implicitement l'hypothèse que tous les sujets procèdent de la même manière, elle construit un dispositif dans lequel on présente sur un écran d'ordinateur une forme (la « cible ») à laquelle succède, quelques secondes plus tard, une seconde forme (la « sonde ») identique ou pas à la cible. Le degré de dissimilarité des formes différentes est manipulé expérimentalement (de « peu » à « très » différentes) en modifiant la position d'un ou de plusieurs sommets de la forme. Les 8 sujets participant à l'expérience doivent juger aussi précisément et rapidement que possible de l'identité d'un grand nombre de paires cible-sonde.

Cooper, qui fait l'hypothèse que la décision des sujets repose sur un processus de comparaison analytique centré sur l'examen des positions des sommets de la sonde par rapport à celles de la cible, s'attend donc, pour les jugements de différence, à une diminution du temps de réponse (TR) avec l'augmentation du degré de dissimilarité entre les formes.

Cette hypothèse s'avère être conforme aux résultats observés pour les jugements de différence sur l'ensemble des données des sujets (figure 4.1). Une analyse plus fine permet cependant d'identifier des différences individuelles correspondant à deux patterns stables de variabilité. On voit en effet que le TR de certains sujets (par ailleurs plus rapides) n'est pas affecté par les manipulations du degré de dissimilarité. Ce résultat peut s'expliquer par la mise en œuvre d'un processus de comparaison globale ou holistique consistant à « superposer mentalement » la représentation de la cible à celle de la sonde. Même sur une tâche aussi simple que celle-ci, des différences dans les stratégies utilisées (analytique *versus* global) peuvent donc être constatées.

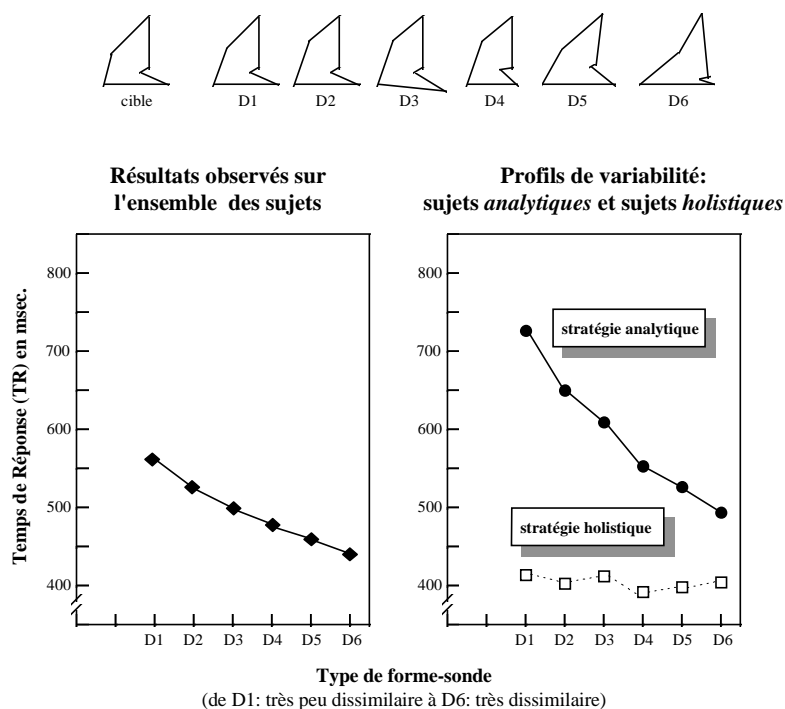


Figure 4.1 - Différences individuelles mesurées par le paradigme de comparaison visuelle (d'après Cooper, 1976).

Des expériences employant ce type de paradigme ont permis également de montrer, dans une perspective intra-individuelle cette fois, qu'au fur et à mesure que les sujets « analytiques » apprennent les cibles à comparer, certains d'entre eux peuvent tendre à adopter préférentiellement une stratégie de comparaison globale.

2.3. La référence à des normes

Dire d'un élève de CP que son regard a effectué 32 allers et retours lors de la copie d'un texte écrit au tableau ou qu'un patient devant subir une intervention chirurgicale a répondu 11 fois positivement sur les 20 items d'un questionnaire d'anxiété est en soi peu informatif. L'interprétation de ces scores isolés nécessite en effet que l'on sache comment se répartissent, s'ordonnent les comportements d'individus « semblables » (des élèves de CP, des patients hospitalisés) confrontés à ces mêmes situations (copier un texte, répondre aux items d'un questionnaire). La norme de référence est ici intra-groupe en ce sens que connaissant la dis-

tribution des résultats sur l'échantillon de référence, sous réserve cependant que celle-ci soit suffisamment étalée, on peut comparer des performances différentes en interprétant la position relative des individus dans la distribution. Cette opération n'est bien sûr possible que si les procédures d'application de l'instrument et de représentation des données d'observation (le calcul des scores individuels) sont rigoureusement les mêmes pour tous les individus d'un échantillon représentatif de la population étudiée. Il est donc essentiel de définir les normes par rapport auxquelles on standardisera l'épreuve.

La notion de standardisation est à comprendre par rapport au dispositif que l'on applique à un échantillon représentatif d'individus d'une population donnée. Les règles de présentation des items, les conditions d'administration (notamment la consigne), le choix de la métrique permettant de classer les individus et les principes de notation du test, du questionnaire ou de la tâche de laboratoire, doivent en effet être parfaitement spécifiés à l'avance. Différents aspects de l'épreuve comme la durée de passation, le nombre, la nature et la difficulté des questions posées, les pondérations affectées à chacune d'entre elles, etc., peuvent être manipulés afin d'imposer, par convention, une forme donnée, le plus souvent gaussienne, à la distribution des scores bruts. Il faut signaler à ce propos que cette « normalisation » est moins effectuée pour des raisons d'ordre psychologique - on ignore quelle est la validité de l'hypothèse de normalité des caractères psychologiques - que pour les propriétés statistiques particulièrement intéressantes des indices servant à résumer les distributions normales. Il peut être utile enfin de disposer de normes empiriques précisant la distribution des mesures effectuées sur un échantillon représentatif de la population de référence: l'épreuve est alors dite étalonnée.

2.4. Le problème de l'incertitude des mesures

La description des différences dépendant de la valeur des estimations fournies par l'instrument, il est important de connaître le degré d'incertitude attaché aux mesures effectuées.

La théorie classique des tests fait l'hypothèse que le score observé au test est influencé par deux sources: la caractéristique psychologique que l'on cherche à mesurer (le score « vrai ») et l'erreur. L'erreur peut être systématique comme certains aspects liés à l'instrument dont il est fait l'hypothèse qu'ils s'expriment de la même façon chez tous les individus de l'échantillon concerné. Elle peut aussi être non systématique ou aléatoire. Une façon immédiate de concevoir théoriquement ce que l'on appelle la fidélité de l'instrument est donc de définir celle-ci comme le pourcentage de variance qu'ont en commun le score vrai et le score observé. Plus ce pourcentage est élevé, plus la précision de la mesure est importante. Connaissant la fidélité de l'instrument, on peut alors estimer simplement la précision des mesures effectuées (l'erreur-standard de mesure) et calculer les intervalles de confiance associés à chacun des scores individuels.

En pratique cependant, on ignore bien entendu quelle est la variance vraie. Pour estimer celle-ci, on s'appuie sur le postulat que les différences entre scores individuels mesurés avec des tests qui classeraient les sujets exactement de la même façon, sont dues à l'erreur de mesure. La fidélité de l'instrument pour un échantillon donné peut alors être estimée en calculant la corrélation entre deux tests parallèles (ou pourcentage de variance vraie), entre des parties différentes d'un même test (homogénéité interne) ou entre deux applications successives du même test (stabilité test-retest). Lorsque les données sont nominales, on peut faire appel à des juges pour leur codage et calculer ensuite la fidélité inter-juges (coefficient kappa de Cohen). Bien que sa valeur soit fonction du nombre d'items du test ou du paradigme employé, un « bon » coefficient de fidélité doit être au moins de l'ordre de 0.70 ce qui signifie que la variance commune au score observé et au score vrai est de 70%.

Développés plus récemment et désormais applicables grâce aux possibilités de calcul offertes par les ordinateurs, d'autres modes d'approche de ce problème commencent progressivement à se faire connaître. Basé sur l'analyse de variance, le modèle de la généralisabilité (Cardinet, 1997) permet en particulier de diffé-

rencier les sources de la variance d'erreur (effets d'échantillonnage des sujets, des items, des conditions d'observation, etc.) et d'aboutir à l'estimation de la proportion de variance vraie dans la variance totale (coefficient intra-classe). Il peut être utilisé pour déterminer les marges d'erreur des mesures absolues effectuées au moyen de paradigmes expérimentaux (temps de réaction, nombre d'erreurs, etc.) ou des mesures recueillies à l'aide de dispositifs d'évaluation individualisée de l'apprentissage c'est-à-dire lorsque le sujet n'est comparé qu'à lui-même.

3. LA DESCRIPTION DES DIFFERENCES INDIVIDUELLES

3.1. Les résumés statistiques

La description des différences individuelles par rapport à une variable donnée rend nécessaire l'emploi de techniques statistiques univariées permettant d'en résumer les caractéristiques essentielles. Cette élaboration ne présente pas les mêmes caractéristiques et n'offre pas les mêmes possibilités d'analyse selon que les données sont qualitatives ou quantitatives.

Les données sont qualitatives ou catégorisées lorsque les différentes valeurs observées sont regroupables en catégories. Ces catégories peuvent être mutuellement exclusives et exhaustives. C'est le cas des données obtenues lorsqu'on demande à des individus d'indiquer parmi deux ou plusieurs stratégies celle qu'il met en œuvre préférentiellement pour résoudre un problème donné. Elles peuvent être aussi ordonnées comme lorsque des individus évaluent sur une échelle de Likert (de « pas du tout » à « beaucoup ») l'intérêt éprouvé à l'égard d'une tâche. Les distributions des variables qualitatives ou catégorisées sont rectangulaires (une fréquence par éventualité) et peuvent être caractérisées par la catégorie modale (catégorie où le nombre d'observations est le plus important) ou médiane (catégorie contenant l'observation médiane), par l'entropie (indice de dispersion nul si toutes les valeurs sont dans la même catégorie et maximum si celles-ci sont équiprobablement réparties) ou par les quantiles (valeurs déterminant des intervalles dans la distribution).

Les données quantitatives sont particulièrement utilisées dans l'étude des différences individuelles. Qu'elles soient continues (le temps mis pour résoudre une tâche) ou discrètes (le nombre de questions résolues), elles peuvent être résumées par des indices de tendance centrale (la moyenne ou la médiane, moins sensible aux valeurs extrêmes), de variabilité (l'écart-type, la variance) et de forme (symétrie, aplatissement). Les mesures de dispersion des scores individuels autour de la tendance centrale sont bien sûr essentielles pour le psychologue différentialiste. Lorsque la distribution des scores est approximativement normale, l'écart-type peut en particulier être très utile pour comparer la variabilité de plusieurs groupes d'individus ayant la même moyenne ou pour servir d'unité de comparaison entre individus.

Voici par exemple les distributions normales telles qu'elles pourraient théoriquement être observées chez de jeunes adultes et des personnes âgées saines de plus de 60 ans et de même niveau culturel à qui on aurait administré des épreuves de mémoire des épisodes spécifiques de la vie quotidienne et passée (ou mémoire épisodique). Les scores sont ici exprimés en écarts réduits à la moyenne pour chacune des deux populations. Comme on le voit sur la figure 4.2, 50 % des individus de chaque groupe d'âge ont un score au moins égal à la moyenne du groupe, 15.87 % ont un score à plus (resp. à moins) d'un écart-type de la moyenne, 2.28 % ont un score à plus (resp. à moins) de deux écarts-types de la moyenne et 0.14 % ont un score à plus (resp. à moins) de trois écarts-types de la moyenne. Il est intéressant de noter qu'en moyenne, les individus âgés ont un niveau de performance inférieur à celui des jeunes mais que certains âgés sont plus performants que certains jeunes.

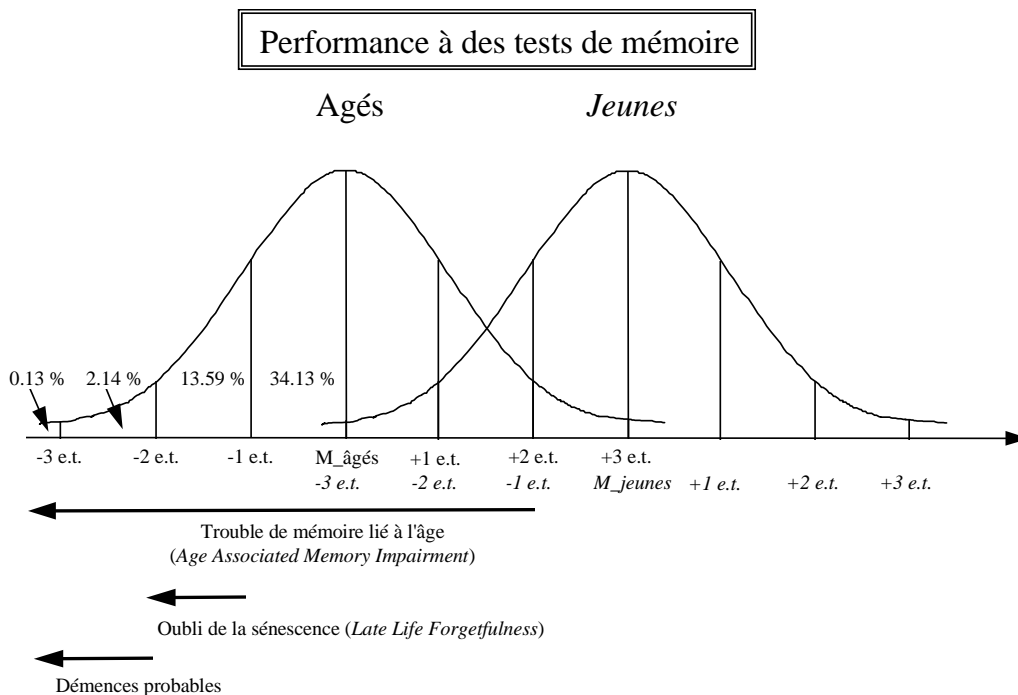


Figure 4.2 - Distributions théoriques des scores à des tests de mémoire épisodique administrés à des individus jeunes et âgés.

Cet exemple montre aussi que les troubles de mémoire de la personne âgée dans la vie de tous les jours peuvent être définis différemment selon que l'on établit les normes par rapport aux individus jeunes ou par rapport aux individus âgés normaux. Le « trouble de mémoire lié à l'âge » (AAMI), ici défini comme un déficit de plus d'un écart-type sans limite inférieure par rapport aux individus jeunes, n'est pas « l'oubli de la sénescence » (LLF) qui correspond à un déficit de mémoire entre -1 et -2 écarts-types par rapport à la moyenne des sujets âgés.

3.2. La description des liaisons entre variables

Les différences individuelles mises en évidence chez les mêmes individus à l'aide de deux ou de plusieurs instruments peuvent être plus ou moins semblables. On dispose de divers coefficients d'association permettant d'estimer le degré de ressemblance entre distributions ou degré de liaison entre variables (Reuchlin, 1976). Le plus connu est sans nul doute le coefficient de corrélation linéaire de Bravais-Pearson, utilisé quand les variables sont quantitatives. De nombreux autres coefficients comme par exemple les coefficients de contingence ou de corrélation de rang ont été développés pour les données de niveau nominal ou celles de niveau ordinal. D'autres coefficients permettent d'évaluer la force des associations entre des variables dont le niveau de mesure n'est pas le même (corrélation polychorique, corrélation polysérielle).

Le coefficient de corrélation de Bravais-Pearson, dont les valeurs s'étendent conventionnellement de -1 à +1, mesure à la fois l'amplitude et le signe de l'association linéaire entre deux variables. Une forte ressemblance entre distributions, une forte association entre variables, se traduit par un coefficient proche de 1 en valeur absolue. Un coefficient nul décrit l'absence de liaison ou l'indépendance entre variables (par exemple lorsque les deux distributions sont aléatoires). Le signe de la corrélation indique le sens de la liaison entre les variables, indépendamment de son amplitude. La corrélation peut être positive comme par exemple celle observée entre le temps de copie et le nombre d'allers et retours du regard entre la feuille et le tableau: le temps mis pour copier le texte tend à être d'autant plus long que le nombre d'allers et retours est important.

Elle peut être négative comme par exemple la corrélation entre le temps de copie et l'intérêt à l'égard de la tâche: le temps de copie tend à être d'autant plus court que l'intérêt est élevé.

L'amplitude d'un coefficient de corrélation ne doit pas être confondue avec son degré de significativité statistique. La significativité n'est en effet pas seulement fonction de la force de la liaison; elle dépend des seuils que l'on s'est fixés et par rapport auxquels on acceptera de rejeter l'hypothèse d'absence de liaison ou de ne pas accepter l'hypothèse de liaison entre les deux variables. Elle dépend aussi de la taille de l'échantillon. Une liaison de 0.40 observée entre deux variables mesurées sur un échantillon de 20 personnes n'est pas significative au seuil de .05 mais le devient sur un échantillon de 80 individus!

Le coefficient de corrélation au carré r^2 (ou coefficient de détermination) est souvent employé pour mesurer la taille de la relation entre deux variables x et y . Cet usage, qui repose sur l'hypothèse d'un modèle de prédiction de y à partir de x où toute la variance de x serait incluse dans la variance du critère y , n'est cependant justifié que si l'on purge la variance de x de la variance liée à l'erreur de mesure. La taille de l'effet est alors le carré de la corrélation corrigée en fonction de la fidélité de la mesure (corrélation désatténuée). On notera aussi que lorsque l'on soupçonne une variable z de contribuer à la corrélation observée entre deux variables x et y , il peut être intéressant de calculer leur corrélation partielle, estimée indépendamment des relations que x et y entretiennent avec z . Mais il est incorrect d'utiliser r^2 si le modèle sous-jacent est un modèle de variance commune. Lorsqu'on fait l'hypothèse qu'une même variable latente est responsable de la covariance entre les deux variables mesurées, le pourcentage de variance qu'ont en commun x et y est r et non r^2 .

4. STRATEGIES D'ANALYSE DES DIFFERENCES

4.1. L'étude structurale des associations

Le psychologue qui, au-delà du constat de l'amplitude des variations individuelles mesurées à l'aide d'un ensemble de variables sur un échantillon donné, souhaite décrire le plus économiquement possible l'organisation des relations observées, dispose d'un éventail assez large de méthodes statistiques aux logiques d'application différentes. Si l'objectif est d'identifier les sources de variations ou variables latentes hypothétiques par rapport auxquelles les individus se différencient, l'analyse factorielle appliquée à la matrice des associations entre variables observées est la méthode structurale la mieux adaptée. On peut aussi, s'il s'agit de former des variables composites regroupant les variables qui différencient les individus de l'échantillon « à peu près de la même façon », utiliser l'analyse en groupes ou l'analyse en composantes principales. L'analyse en groupes de la matrice des proximités entre individus sera employée pour former des groupes correspondant à des profils différents de variabilité intra-individuelle c'est-à-dire pour définir une structure de types au sein de laquelle se différencient les individus.

4.1.1. L'identification des facteurs et la description de leur organisation

Les méthodes d'analyse factorielle comme l'analyse en facteurs communs, l'analyse en composantes principales ou l'analyse des correspondances peuvent être utilisées dans ce but. L'objectif de ces méthodes statistiques est de fournir une représentation géométrique des associations constatées entre variables observées au moyen d'un nombre plus réduit de variables hypothétiques ou variables latentes appelées facteurs. L'une des méthodes permettant d'atteindre cet objectif, l'analyse en facteurs communs (pour présentation Dickès, 1996), est particulièrement intéressante dans l'étude des différences individuelles.

On se bornera très rapidement à indiquer que le modèle en facteurs communs repose sur un principe de décomposition linéaire de la variance de chaque variable observée x en une part de variance commune - ou communauté - « expliquée » par une combinaison linéaire des facteurs communs à l'ensemble des variables

et une part de variance unique (variance d'erreur et variance spécifique non partagée avec les facteurs communs). Les pondérations associées à chaque facteur commun sont appelées les saturations; ces coefficients, sortes de corrélations partielles entre chacune des variables et chacun des facteurs, sont les coordonnées de chaque variable dans l'espace des facteurs. Etant inconnues avant l'analyse, les communautés de chaque variable observée sont estimées itérativement à partir de valeurs fixées initialement. Les facteurs communs étant par hypothèse seuls responsables des variables observées, seule la variance commune à l'ensemble des variables observées est ensuite analysée. On dispose de plusieurs méthodes pour estimer les saturations qui permettent de reconstruire les corrélations observées entre les variables.

L'analyse en facteurs communs peut être employée dans une logique exploratoire descriptive (ou inférentielle si l'on veut généraliser à d'autres sujets - ou à d'autres variables - de la population), pour trouver la « meilleure » combinaison linéaire possible des facteurs expliquant la variance commune. L'objectif est de trouver une solution qui explique un pourcentage de variance commune aussi élevé que possible pour un nombre des facteurs aussi réduit que possible. On complète souvent l'analyse en effectuant une rotation des facteurs c'est-à-dire en cherchant une solution plus contrastée dans laquelle la variance expliquée est redistribuée par rapport aux facteurs afin d'en faciliter l'interprétation psychologique.

Matrice de corrélations

	LO	TS	SL	V	A	VF
LO	1.000					
TS	0.173	1.000				
SL	0.178	0.271	1.000			
V	0.078	0.057	0.101	1.000		
A	0.161	0.122	0.156	0.182	1.000	
VF	0.068	0.104	0.152	0.142	0.152	1.000

Matrices factorielles

Communauté	<u>Avant rotation</u>		<u>Après rotation orthogonale</u>		
	Facteur 1	Facteur 2	Facteur 1'	Facteur 2'	
LO	.122	0.345	-0.052	0.304	0.172
TS	.276	0.453	-0.266	0.521	0.071
SL	.278	0.502	-0.162	0.495	0.183
V	.177	0.295	0.300	0.047	0.418
A	.213	0.407	0.217	0.186	0.422
VF	.115	0.309	0.139	0.157	0.300
Variance commune expliquée					
		0.154	0.043	0.112	0.085
- /variance totale		15.4 %	4.3 %	11.18 %	8.52 %
- /variance commune		68.14 %	19.11 %	49.47 %	37.78 %

Tableau 4.1 - Matrice de corrélations entre les 6 variables et matrices factorielles à 2 facteurs communs avant et après rotation Varimax.

Les résultats d'une analyse en facteurs communs appliquée aux corrélations observées entre des épreuves de logique opératoire (LO), de transformation spatiale (TS), de séries de lettres (SL), de vocabulaire (V), d'arithmétique (A) et de vérification de phrases (VF) administrées à un échantillon de 8685 jeunes adultes peuvent servir à illustrer les principales étapes de la démarche qui vient d'être brièvement décrite. Les communautés des variables sont estimées dans un premier temps de l'analyse (tableau 4.1). On notera la faiblesse de ces dernières (la variance expliquée par les facteurs communs à l'ensemble des variables ne repré-

sente que 22.60 % de la variance totale) due à la faiblesse des corrélations observées sur cet échantillon. L'étape suivante est celle de l'extraction des facteurs. Les deux premiers facteurs extraits expliquant 87.55 % de la variance commune (15.4 % de la variance totale ou 68.14 % de variance commune pour le premier, 4.3 % de la variance totale ou 19.11 % de la variance commune pour le second), cette solution est donc préférée à une solution comportant 3 facteurs.

On présente dans le tableau 4.1. la solution finale obtenue après rotation orthogonale des facteurs. Les saturations de chacune des variables de l'étude sur chacun des facteurs permettent de construire la représentation bidimensionnelle des variables observées, avant et après rotation orthogonale (figure 4.3).

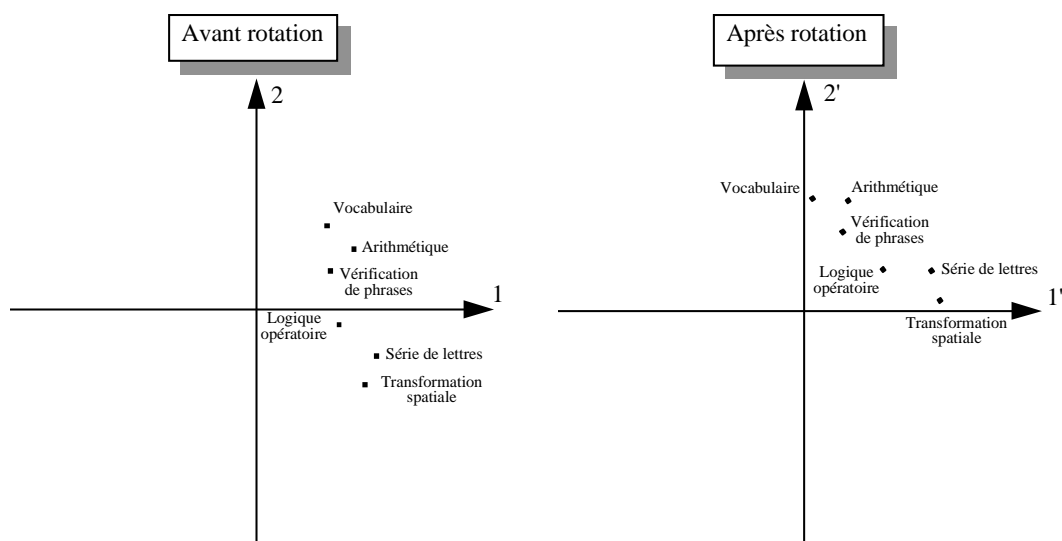


Figure 4.3 - Représentation des variables par rapport aux facteurs, avant et après rotation orthogonale.

Le facteur 1' obtenu après rotation explique maintenant 49.47 % de la variance commune, le facteur 2' en explique 37.78 % (le pourcentage de variance commune expliquée par l'ensemble des deux facteurs reste le même qu'avant rotation). En raison de ses saturations élevées avec les variables de logique opératoire, de transformation spatiale et de série de lettres, le facteur 1' peut être interprété comme un facteur de raisonnement. Le facteur 2' est interprétable pour sa part comme un facteur verbal.

L'analyse en facteurs communs peut aussi être utilisée dans une logique hypothético-déductive ou confirmatoire (Dickès, 1996) quand, après avoir spécifié des hypothèses sur le nombre de facteurs à extraire, sur les relations entre les variables latentes et sur celles entre les variables latentes et les variables observées, on cherche à évaluer la plausibilité de cet ensemble d'hypothèses théoriques (ou modèle) en déterminant le degré d'accord entre les corrélations reproduites sous hypothèses et les corrélations observées.

On adopte cette stratégie confirmatoire pour étudier à l'aide du programme LISREL (Bacher, 1987, 1988; Dickès et al. 1994) la structure factorielle des corrélations entre 9 épreuves visuo-spatiales appliquées à 92 étudiants. Il s'agit d'épreuves informatisées de mémoire visuelle et d'identification perceptive ainsi que de plusieurs tests papier-crayon. Le modèle hiérarchique mis à l'épreuve fait l'hypothèse d'un facteur général sur lequel saturent toutes les variables et de deux facteurs communs de niveau intermédiaire (figure 4.4).

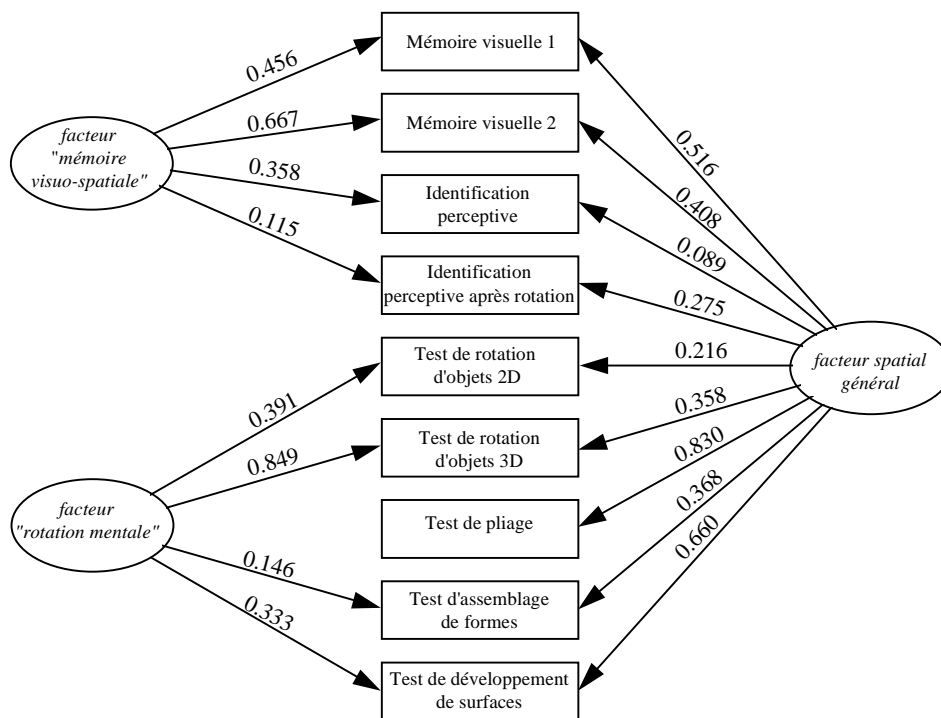


Figure 4.4 - Graphe du modèle hiérarchique d'analyse factorielle confirmatoire (d'après Juhel, 1990)

Le facteur « mémoire visuo-spatiale » est mesuré par 4 épreuves (Mémoire visuelle 1 et 2, Identification perceptive et Identification perceptive après rotation) dont les saturations sont à estimer par le programme (on dit qu'elles sont libérées). Les saturations sur ce même facteur de Rotation d'objets 2D et 3D, Pliage, Assemblage de formes et Développement de surfaces sont fixées à 0. Elles sont par contre libérées sur le facteur « rotation mentale ». On fait également l'hypothèse que les facteurs ne sont pas corrélés.

L'adéquation de la représentation théorique est vérifiée après estimation des paramètres en fonction des contraintes spécifiées. L'estimation de la saturation de Pliage sur le facteur « rotation mentale » étant très proche de 0, on spécifie un nouveau modèle dans lequel cette saturation est fixée à 0. Un test statistique permettant d'infirmer ou non l'hypothèse nulle de l'égalité entre la matrice de corrélations reproduite à partir des contraintes du modèle et la matrice de corrélations observée, est effectué par le programme. La probabilité associée de 0.40 indique que l'hypothèse nulle ne peut être rejetée et que l'hypothèse d'un facteur général et de deux facteurs de niveau intermédiaire mesurés comme indiqué sur la figure 4.4 est compatible avec les observations. On remarquera que les estimations des saturations peuvent être facilement utilisées pour calculer, pour chaque variable observée, la proportion de variance expliquée par chacun des facteurs. Pour Mémoire visuelle par exemple, $(0.516)^2 \times 100$ soit 26.63 % de variance totale sont expliqués par le facteur spatial général alors que $(0.456)^2 \times 100$ soit 20.79 % le sont par le facteur « mémoire visuo-spatiale ». La proportion de variance unique de Mémoire visuelle est donc de $100 - (26.63 + 20.79)$ soit 52.58 %.

4.1.2. La constitution de groupes

Les analyses en groupes ou en clusters sont des techniques de classification permettant de former des taxonomies ou des typologies. Elles sont généralement utilisées de manière exploratoire pour identifier des sous-groupes homogènes d'individus (ou de variables) au sein d'un échantillon donné. La première étape consiste à mesurer le degré de similitude du profil de variabilité de chaque individu (resp. de la distribution de chaque variable) avec celui de chacun (resp. de chacune) des autres. Les coefficients utilisés vont de la corrélation aux mesures d'association pour variables nominales ou ordinales. Divers algorithmes peuvent

être ensuite appliqués à la matrice des coefficients de proximité. Les algorithmes hiérarchiques par exemple regroupent progressivement les individus (resp. les variables) en groupes de taille de plus en plus importante de telle manière que la variance intra-groupe reste aussi faible que possible. Un diagramme en arbre ou dendrogramme sert à représenter la structure hiérarchique des proximités entre individus (resp. variables). Les notions de similitude et d'homogénéité étant relatives aux critères statistiques servant à les définir, les classifications auxquelles ces méthodes permettent d'aboutir peuvent différer selon le coefficient et surtout selon l'algorithme utilisés.

Imaginons à titre d'exemple qu'on veuille constituer plusieurs groupes d'individus, homogènes au regard du profil de leurs réponses aux échelles d'anxiété, d'extraversion et d'empathie d'un questionnaire de personnalité afin d'évaluer l'effet, pour chacun de ces groupes, d'une même technique d'aide à l'élaboration du projet professionnel. Pour mesurer la proximité entre chaque paire de patients, on a employé le carré de la distance euclidienne [la proximité entre les patients 1 et 2 est $(39-36)^2 + (71-34)^2 + (18-15)^2 = 1387$, celle entre les patients 1 et 13 est 94]. On a ensuite appliqué une analyse en groupes (méthode de Ward) à la matrice de proximités entre patients (figure 4.5).

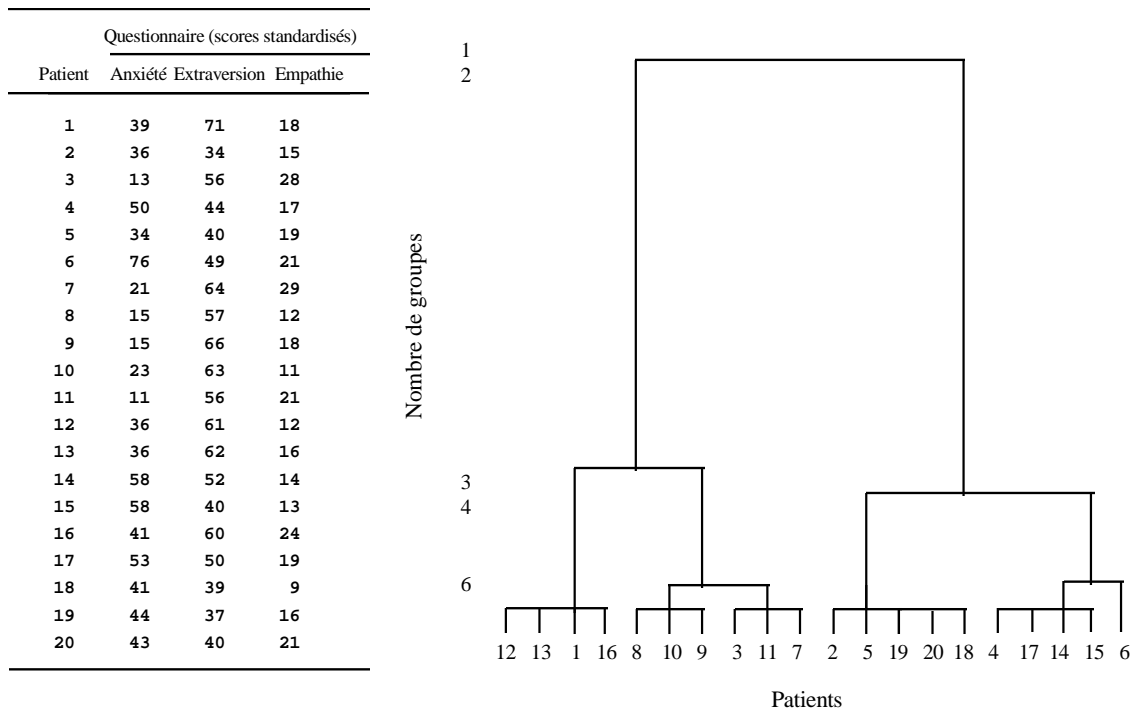


Figure 4.5 - Classification hiérarchique des 20 patients.

Les résultats de cette analyse montrent que les solutions à 3 ou à 4 groupes sont les plus appropriées. L'erreur intra-groupe étant d'autant plus importante que le nombre de groupes est réduit, la solution à 4 groupes (12, 13, 1, 16 - 8, 10, 9, 3, 11, 7 - 2, 5, 19, 20, 18 - 4, 17, 14, 15, 6), bien que moins économique, est peut-être à préférer car les groupes ainsi constitués sont plus homogènes. L'interprétation des groupes peut s'appuyer sur les fonctions composites (ici trois variables) qui permettent de les discriminer.

4.2. La recherche de prédicteurs, l'identification de facteurs de causalité

Il est habituel de prévenir des dangers d'inférer la causalité de la corrélation entre deux variables x et y . Pourtant, l'association incontestable entre le fait d'être un grand fumeur et la probabilité d'apparition d'un cancer du poumon se traduit souvent dans le langage commun par l'affirmation selon laquelle fumer provoque le cancer. Parce qu'il est sans doute plus contre-intuitif, le constat qu'une moindre assiduité au collège est associée à une amélioration de la performance au Brevet des collèges risque moins de susciter une explication qui ferait de la faible assiduité la cause de l'amélioration de la performance. Ces deux exemples montrent bien que l'observation d'une corrélation significative entre x et y , si elle est compatible avec un schéma explicatif du type $x \rightarrow y$, ne peut à elle seule en prouver la validité car de nombreuses variables intermédiaires peuvent contribuer à expliquer l'association observée. L'inférence causale n'est donc éventuellement possible qu'en référence à un système de variables englobant la liaison mise en évidence entre les deux variables. Seule l'absence de corrélation entre x et y permet de réfuter avec certitude l'hypothèse que y est expliquée par x ou d'affirmer que le raisonnement ayant conduit à l'énoncé des prédictions est erroné.

Ces remarques faites, on distinguera dans cette partie deux grandes directions de recherche dans l'analyse des différences individuelles. Dans l'investigation corrélationnelle réalisée en milieu naturel, l'objectif majeur est de repérer les liaisons à la fois fortes et stables entre des prédicteurs et un ou plusieurs critères mesurés sur chacun des individus de l'échantillon. L'affectation aléatoire des individus étant souvent très difficile, il n'est pas possible d'éprouver des hypothèses causales comme on peut le faire dans une expérience, puisque rien ne garantit qu'il n'y ait pas, sur la population, de corrélation entre les erreurs expérimentales et les effets de la manipulation. Il s'agit donc ici plus de prévision que d'explication au sens que lui donnent en tout cas les expérimentalistes. Dans une seconde direction de recherche, on préfère mettre en œuvre des procédures quasi-expérimentales de manipulation des sources de variations individuelles. L'intérêt principal des recherches de cette catégorie est qu'il devient possible de mettre à l'épreuve des hypothèses - spécifiées *a priori* - relatives aux associations entre les variables manipulées et les variables dépendantes de l'étude. Les travaux de ce type possèdent cependant des limitations contextuelles non explorées qui rendent délicate la généralisation des résultats de l'expérience.

4.2.1. L'investigation corrélationnelle

C'est à la méthodologie des enquêtes (Bacher, 1982) que peut faire appel le psychologue différentialiste qui souhaite par exemple identifier les variables les plus fortement associées à la compétence scolaire d'adolescents présentant des troubles du comportement, évaluer dans quelle mesure le fait pour un lycéen de Terminale d'avoir de bons résultats en mathématiques permet de prédire la poursuite d'études scientifiques, déterminer si l'âge, les conditions d'hébergement ou les perceptions de contrôle sont en relation avec le bien-être de la personne âgée, etc. Une fois effectué le choix des variables indépendantes et dépendantes, les données de l'enquête sont recueillies au moyen de variables mesurées chez chaque participant de l'échantillon de la population donnée. Une partie importante du travail de recherche concerne ensuite l'analyse des associations entre variables. On dispose pour cela de nombreuses techniques statistiques dont le choix repose sur le type de données recueillies (quantitatives ou catégorisées) ou sur le nombre de variables dépendantes (une ou plusieurs).

La régression multiple, dont la régression simple est un cas particulier, est une technique fréquemment employée lorsque le ou les critères que l'on cherche à prédire sont quantitatifs. Le but de cette méthode est de fournir une règle de prédiction d'un critère à partir d'un ensemble de prédicteurs intercorrélés. Son principe repose sur la constitution d'une combinaison linéaire des prédicteurs, chacun de ceux-ci étant pondéré par un coefficient mesurant la contribution propre du prédicteur à la variance du critère. Cette variable composite représente l'estimation du critère pour les prédicteurs donnés. Elle est construite par optimisation afin

que sa corrélation avec le critère soit maximum. Plus la quantité de variance du critère prédite ou « expliquée » par les prédicteurs est grande, plus faible est la variance non prédite (on dit aussi résiduelle), meilleure donc est l'estimation.

L'exemple suivant permet de saisir l'intérêt de la régression multiple dans la détermination des prédicteurs permettant de prévoir un critère. Les données analysées proviennent d'une recherche ayant pour objectif l'identification, pour un suivi longitudinal ultérieur, de certains prédicteurs de la capacité d'apprentissage de personnes âgées. Le niveau d'éducation (NE), la diversité des occupations quotidiennes (NA), le niveau de dépression (ND) et la performance à une tâche d'apprentissage (CA) ont été mesurés chez 77 personnes âgées de plus de 65 ans et vivant en maison de retraite. On a appliqué la régression multiple séquentielle à la matrice de corrélations observée (tableau 4.2). Le principe en est le suivant. Le premier prédicteur introduit dans la variable composite est celui qui présente la corrélation la plus élevée avec le critère soit NE. Le second prédicteur retenu (A) est celui dont l'introduction dans la variable composite constituée à l'étape précédente augmente le plus le pourcentage de variance expliquée. On voit que l'introduction de A conduit à une augmentation de 10.7 % de la variance expliquée (49.2-38.5). Le processus est reproduit jusqu'à ce que tous les prédicteurs apportant un minimum d'information, minimum fixé par avance, soient introduits dans la variable composite. Le programme indique lors de l'étape 4 que l'introduction de NA n'augmente pas significativement ($p > .05$) le pourcentage de variance expliquée du critère. On retient donc la solution fournie à l'étape 3 qui explique un pourcentage satisfaisant (53.3 %) de la variance de CA. Chacun des prédicteurs est pondéré par un coefficient de régression partielle pouvant, lorsqu'il est standardisé comme c'est le cas ici, être comparé aux autres. Ainsi, lorsque les autres prédicteurs sont maintenus constants, une augmentation de une unité de A prédit une diminution de 0.404 unité de la valeur estimée de CA (CA_{est}), une augmentation de une unité de ND prédit une diminution de 0.190 unité de CA_{est} , une augmentation de une unité de NE prédit une augmentation de 0.375 unité de CA_{est} .

Matrice de corrélations

	A	DO	NE	ND	CA
Age (A)	1.000				
Diversité des occupations (DO)	-0.263	1.000			
Niveau d'éducation (NE)	-0.561	0.252	1.000		
Niveau de dépression (ND)	0.017	-0.127	-0.089	1.000	
Capacité d'apprentissage (CA)	-0.618	0.344	0.621	-0.245	1.000

Régression multiple séquentielle: introduction par ordre d'« importance » de chacun des prédicteurs dans la variable composite.

	Variance expliquée (R^2)	Variable composite (standardisée)	Variance résiduelle
1. NE	0.385	$CA_{est} = 0.621 \times NE$.615
2. NE,A	0.492	$CA_{est} = 0.399 \times NE - 0.394 \times A$.508
3. NE,A,ND	0.533	$CA_{est} = 0.375 \times NE - 0.404 \times A - 0.204 \times ND$.467
4. NE,A,ND,NA	0.548	$CA_{est} = 0.358 \times NE - 0.380 \times A - 0.190 \times ND + 0.130_{ns} \times DO$.452

Tableau 4.2 - Résultats de l'analyse de régression multiple visant à identifier les meilleurs prédicteurs de la capacité d'apprentissage.

Plusieurs prédicteurs fournissant une bonne prévision de la capacité d'apprentissage ont donc pu être identifiés sur cet échantillon de personnes âgées. Toutes choses restant égales par ailleurs, l'âge qui s'accompagne d'une importante baisse de la performance à la tâche d'apprentissage, apparaît comme un « facteur de risque » primordial. Dans une moindre mesure, ce constat vaut aussi pour le niveau de dépression. Le niveau d'éducation qui est positivement relié à la performance à la tâche d'apprentissage apparaît par contre comme un « facteur de protection ». On notera enfin qu'en raison du recouvrement de la variance de DO par celle de NE, la diversité des occupations, bien que corrélée positivement à la performance, n'en est pas un bon prédicteur.

4.2.2. La planification quasi-expérimentale

La complémentarité soulignée plus haut entre les méthodes expérimentale et différentielle s'illustre particulièrement bien dans les recherches quasi-expérimentales dans lesquelles des sources de variations interindividuelles sont considérées comme des variables indépendantes susceptibles d'intervenir dans les effets observés entre les manipulations expérimentales et les variables dépendantes de l'étude mesurées sur des groupes contrastés ou de sujets différenciés sous l'angle d'une ou de plusieurs variables individuelles. Les résultats de ces travaux étant le plus souvent traités par analyse de variance, le psychologue choisit souvent de comparer - ou contraster - deux niveaux au moins de chacune des sources de variations interindividuelles (jeunes vs vieux, anxieux vs non anxieux, bons lecteurs vs mauvais lecteurs, etc.). Le regroupement des sujets, opération d'autant délicate qu'elle doit s'effectuer par rapport à des variables invoquées échappant plus ou moins au contrôle du chercheur (le niveau d'efficacité, le niveau d'anxiété, etc.), ne s'impose cependant pas lorsque les données sont analysées avec une méthode comme l'analyse de régression dont l'analyse de variance peut être vue comme un cas particulier.

Les protocoles quasi-expérimentaux incluant des variables individuelles peuvent être extrêmement variés. Ils sont utilisés, de diverses manières, dans de nombreux domaines de recherche en psychologie (Reuchlin, 1992). En voici une illustration simple portant sur l'étude des relations entre cognition et anxiété.

	Niveau d'anxiété			<u>Analyses de variance</u>		
	Faible (n=21)	Moyen (n=26)	Elevé (n=25)	Effet anxiété	Effet bloc	Effet anxiété × bloc
Score moyen à l'échelle d'anxiété	4.8	8.7	13.2			
Attention soutenue						
- précision bloc 1	4.7	4.2	3.8	ns	p<0.001	ns
bloc 6	2.4	1.9	2.3			
- TR bloc1	389	371	350	ns	ns	p<.05
(msec.) bloc 6	384	370	430			
Vitesse psychom. sur les 6 blocs (coups/sec.)	7.1	6.4	6.4	ns	p<0.001	ns

Tableau 4.3 - Résultats observés aux épreuves d'attention soutenue et de vitesse pour les différents groupes (TR: temps de réponse).

Afin d'examiner le rôle possible des ressources de traitement dans les déficits de performance associés à l'anxiété sub-clinique, on mesure chez 72 sujets adultes le niveau d'anxiété (échelle d'auto-évaluation) et la performance à des tâches d'attention soutenue et de vitesse psychomotrice dont les items sont répartis en 6 blocs. Les résultats moyens observés sur les 3 groupes constitués en fonction du score des sujets à l'échelle d'anxiété sont présentés dans le tableau 4.3. On peut remarquer que si l'efficacité du traitement de l'information mesurée par la précision à la tâche d'attention soutenue décline avec le temps, aucun effet de l'anxiété sur la précision n'est observé. L'analyse de variance effectuée sur la durée du traitement de l'information mesurée par les temps de réponse met cependant en évidence un intéressant effet d'interaction. On voit en effet que le temps de traitement des sujets les plus anxieux augmente significativement au fur et à mesure que ceux-ci progressent dans la tâche. Ce résultat, qui ne peut être attribué à des différences de vitesse psychomotrice puisque celle-ci ne varie pas significativement d'un groupe à l'autre, est compatible avec l'hypothèse selon laquelle les individus dont le niveau d'anxiété est élevé doivent allouer plus de ressources attentionnelles à la tâche pour maintenir un niveau de performance comparable à celui d'individus dont le niveau d'anxiété est plus faible.

D'autres exemples d'application de ce type de méthodologie peuvent être empruntés aux recherches de psychologie de l'éducation qui tentent de déterminer comment l'interaction entre des différences psychologiques entre élèves (aptitudes, styles d'apprentissage, etc.) et des variations dans l'environnement scolaire (caractéristiques de la situation pédagogique, style d'enseignement, etc.) peut jouer sur l'efficacité des apprentissages. On forme des enseignants de mathématiques à accroître la durée des activités d'apprentissage de leurs élèves ou à promouvoir chez ceux-ci l'utilisation de stratégies spécifiques d'apprentissage puis on constitue quatre groupes de sujets en croisant le niveau de connaissances en mathématiques (fort *versus* faible) par le type d'enseignement (apprentissage de longue durée *versus* apprentissage de stratégies spécifiques). Les comparaisons entre groupes permettent de faire la preuve qu'en moyenne, les élèves ayant le plus de connaissances en mathématiques bénéficient plus d'une augmentation de la durée des activités alors que l'apprentissage de stratégies spécifiques est plus profitable aux élèves disposant de moins de connaissances. Certains autres de ces travaux utilisent aussi des plans du type « pré-test, traitement, post-test » pour évaluer, sur des groupes de sujets ou des sujets différenciés par rapport au niveau de connaissances, à la motivation, à l'estime de soi, etc. les effets d'un même traitement sur la performance. C'est ainsi que Loarer *et al.* (1995), dans une étude ayant pour objectif d'évaluer les effets d'un programme d'entraînement cognitif montrent que l'impact - globalement faible - du programme sur la performance cognitive de jeunes adultes, la performance est plus marqué pour les individus dont le niveau initial était le plus élevé.

Les recherches différentielles faisant appel à des procédures quasi-expérimentales ont cependant à faire face à un certain nombre de difficultés. La définition des variables individuelles ayant servi à contraster les groupes ou à différencier les sujets doit être conceptuellement et opératoirement aussi bien comprise que possible. Se pose en outre le problème de l'équivalence des groupes quand la stratégie de recherche nécessite l'utilisation d'un groupe contrôle. Une autre difficulté importante concerne le rôle de ces variables individuelles dans l'explication par rapport à celui de variables situationnelles manipulées expérimentalement. La manipulation des variables individuelles, notamment en milieu naturel, ne peut avoir qu'assez rarement un caractère aléatoire. On ne peut donc exclure l'hypothèse que l'effet mis en évidence entre une variable individuelle ayant la fonction de variable indépendante et une variable dépendante puisse être dû à des variables non invoquées dans l'étude. Une façon d'assurer la validité de l'interprétation est alors de répliquer l'étude sur un échantillon comparable de sujets.

4.3. La modélisation causale des relations entre variables

Nous avons vu que pour structurer et interpréter les différences dans les conduites observables des individus, les psychologues différentialistes peuvent évoquer l'influence de variables latentes d'un niveau plus ou moins élevé de généralité. Lorsque, malgré les difficultés liées à l'affectation aléatoire des sujets, la manipulation d'une ou de plusieurs de variables individuelles s'avère possible, une approche quasi-expérimentale à visée explicative peut aussi être utilisée. Une troisième voie dont l'objectif est de comprendre le fonctionnement d'un ensemble de variables dont l'interdépendance est plus la règle que l'exception peut aussi être empruntée dans l'étude des différences individuelles.

Souhaitant modéliser la structure des interactions entre des variables pertinentes au regard de la problématique concernée, de la conduite étudiée et des caractéristiques de l'échantillon, le psychologue qui étudie les différences individuelles peut faire appel à des méthodes statistiques, les modèles structuraux (Bacher, 1987, 1988; Dickès *et al.*, 1994), lui permettant de tester ses hypothèses. La démarche de modélisation structurale est une démarche hypothético-déductive qui peut, indépendamment des phases d'opérationnalisation des variables et de recueil des données, être décomposée en deux étapes principales.

L'étape de spécification consiste à énoncer des hypothèses théoriques puis à élaborer un ou plusieurs modèles susceptibles de prédire, d'expliquer l'organisation des données recueillies sur l'échantillon considéré. L'ensemble des hypothèses opérationnalisées ou modèle est ici explicité sous la forme d'une structure

organisée de variables entretenant entre elles des relations causales orientées dans un sens unique ou dans les deux sens. Les variables, qui peuvent être observées ou latentes, sont dites exogènes lorsqu'on fait l'hypothèse qu'elles se situent « à l'origine » de la chaîne de causalité. Elles sont dites endogènes lorsqu'on considère qu'elles peuvent être sous l'influence d'une ou de plusieurs autres variables.

Le ou les modèles spécifiés sont ensuite soumis à l'épreuve des faits. Chacun de ces modèles permettant de « reconstruire » une matrice de variances-covariances entre les variables mesurées sur l'échantillon, il s'agit fondamentalement de juger du degré d'ajustement du modèle, de sa qualité prévisionnelle, en comparant la matrice de variances-covariances qu'il permet de prédire à la matrice de variances-covariances observée. Statistiquement, cela revient à tester l'hypothèse nulle d'identité entre les variances et covariances prédites par le modèle et celles observées sur l'échantillon. La stratégie généralement adoptée consiste à mettre en compétition plusieurs modèles afin d'identifier parmi eux le modèle le plus vraisemblable c'est-à-dire celui dont les prédictions sont les plus compatibles avec l'organisation des données. Le degré de généralité du modèle retenu peut être éventuellement évalué en examinant la qualité de son ajustement avec des données recueillies sur un échantillon autre que celui sur lequel il a été initialement testé.

L'exemple suivant dans lequel la structure considérée ne comporte que des variables observées illustre la démarche de modélisation structurale. Dans une étude réalisée chez plus de 500 élèves de cours moyen 1^{ère} et 2^{ème} année, Rouxel (1997) cherche à comprendre comment les différences interindividuelles dans le niveau d'anxiété et dans les croyances d'efficacité interagissent avec les différences de connaissances pour produire les différences de performance à un contrôle de français. L'hypothèse principale de la recherche est que l'auto-efficacité perçue en situation d'évaluation médiatise les effets de variables cognitivo-affectives moins spécifiques à la situation sur la performance.

Les variables cognitivo-affectives sont mesurées par des questionnaires à la fois à un niveau général (propension à être anxieux en situation d'évaluation, à croire en général à son efficacité en français) et juste avant et après le contrôle (être anxieux lors du contrôle, croire à son efficacité au contrôle). Les connaissances des élèves en français sont évaluées au moyen d'un test scolaire. Après avoir testé au moyen du modèle LISREL plusieurs structures concurrentes, l'auteur retient un modèle général compatible avec l'hypothèse d'une influence directe (et indirecte par l'intermédiaire de l'anxiété) de l'auto-efficacité sur la performance. Mais deux modèles structuraux aux caractéristiques différentes sont nécessaires pour rendre compte des corrélations entre les variables de la structure selon qu'elles sont observées sur l'échantillon des élèves ayant un niveau élevé de connaissances ($n=129$) ou sur celui des élèves ayant un niveau plus faible ($n=123$).

Les variables endogènes des modèles de la figure 4.6 sont l'auto-efficacité en français et l'anxiété liée à l'évaluation. Les variables endogènes peuvent être dépendantes et indépendantes (auto-efficacité pré-contrôle, anxiété pré-contrôle) ou seulement dépendantes (anxiété post-contrôle). Les coefficients standardisés apparaissant au-dessus des pistes causales peuvent être interprétés comme des coefficients de régression partielle. On voit que quel que soit le niveau de connaissances, croire en son efficacité en français tend à augmenter le niveau de performance (coefficient de 0.23 ou 0.21) et à réduire indirectement le niveau d'anxiété avant le contrôle. L'hypothèse que plus le niveau d'anxiété liée à l'évaluation est élevé, plus le niveau d'anxiété transitoire avant et après le contrôle l'est aussi s'avère également compatible avec l'organisation des données. La comparaison des deux modèles met également en évidence que l'anxiété liée à l'évaluation n'a aucun effet sur la performance chez les élèves dont le niveau de connaissances est élevé alors qu'elle a à la fois un effet direct et indirect sur la performance chez l'autre groupe. L'effet direct (coefficient de parcours de 0.17) pourrait traduire le rôle motivationnel d'aide à l'investissement de l'anxiété dans la préparation du contrôle. L'effet indirect par l'intermédiaire de l'anxiété situationnelle réfléchirait pour sa part l'effet négatif classique de l'anxiété sur la performance. Le rôle des variables cognitivo-affectives dans la performance semble donc plus déterminant pour les élèves de faible niveau de connaissances qu'il ne paraît l'être pour ceux dont le niveau de connaissances est élevé (15 % *versus* 5 % de va-

riance expliquée). On a ici un exemple de structure de variables cognitivo-affectives dont le fonctionnement diffère selon le niveau de connaissances des élèves.

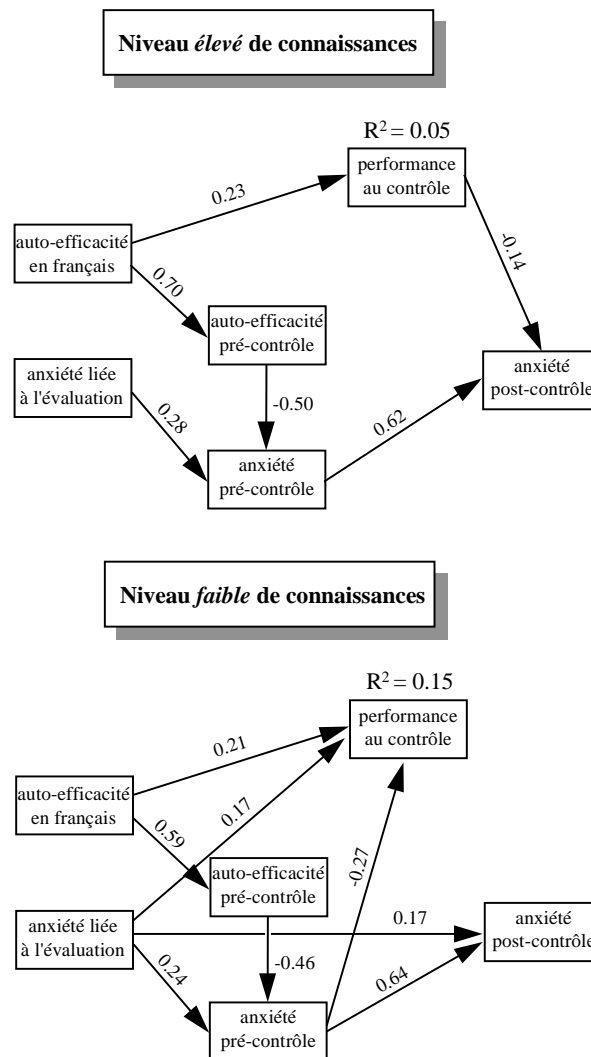


Figure 4.6 - Modèles structurels des relations entre variables cognitivo-affectives et performance au contrôle de français (d'après Rouxel, 1997).

L'adhésion à cette approche structurale et fonctionnelle des différences psychologiques entre individus a plusieurs conséquences importantes. Il faut d'abord pouvoir accepter qu'aucun, parmi les modèles spécifiés, ne puisse décrire correctement les données ou pouvoir reconnaître que des modèles différents sont également capables de rendre compte des variations observées. Le choix d'une telle option favorise cependant une approche plus dynamique des différences dans laquelle les liaisons entre variables peuvent évoluer fonctionnellement au cours du temps (Juhel, 1997). Il favorise enfin une vision moins strictement dispositionnelle des déterminants psychologiques de la conduite, certaines variables contextuelles pouvant faire partie de la structure considérée.

Un exemple récent emprunté aux recherches différentielles sur les stratégies de « faire face » aux situations aversives (le *coping*) peut servir à illustrer certains des points précédents. On considère en effet que la stratégie mise en œuvre par un individu pour résister par exemple à la douleur (s'efforcer de penser à autre chose, prendre médicament sur médicament, etc.) émerge de la transaction entre certaines caractéristiques

psychologiques de l'individu (sa propension générale à l'anxiété, ses craintes spécifiques, son expérience de situations douloureuses semblables, etc.) et l'événement douloureux vécu. Il n'y aurait donc pas une seule stratégie adaptée quel que soit l'événement douloureux ni une seule stratégie appropriée quels que soient les individus. On pourrait alors chercher à comprendre les différences observées en tentant de modéliser différents systèmes de relations entre les antécédents psychologiques de l'individu, les caractéristiques perçues de la situation, les stratégies d'ajustement préférentiellement appliquées, etc. afin de décrire l'évolution au cours du temps des liaisons fonctionnelles dont ces modèles font l'hypothèse.

5. CONCLUSION

Bien qu'il n'ait été possible ici de ne donner qu'une image très schématique des méthodes pouvant être employées dans l'étude des différences individuelles, leur diversité rend possible d'aborder de manière heuristique les questions auxquelles est amené à s'intéresser le psychologue différentialiste. La logique et les principes d'application de ces méthodes ne peuvent sans doute pas être intégrés sans effort. Mais la difficulté est sans doute moins d'en acquérir une maîtrise technique minimum que de comprendre comment le choix d'une méthode donnée permet au psychologue d'appréhender sous un angle ou sous un autre, les différences psychologiques entre individus. Celui-ci en effet risque de ne pas problématiser de la même manière une question donnée selon qu'il envisage de chercher une cause ou de comprendre un fonctionnement. Les données d'observation qu'il lui faut construire ne véhiculeront pas la même information selon l'échelle de mesure employée. Les conclusions même de sa recherche ne seront pas indépendantes des méthodes d'analyse qui lient indissociablement ces premières au substrat empirique servant à les justifier.

Il faut aussi insister sur l'évolution méthodologique fondamentale apparue ces dernières années dans l'étude des différences individuelles grâce aux possibilités offertes par les modèles structuraux. Comme on l'a indiqué, ces modèles dont l'emploi rend nécessaire une explicitation théorique *a priori* des relations entre variables mesurant des différences, permettent de tester des hypothèses causales à propos de certaines liaisons fonctionnelles entre variables. Ils peuvent donc servir à la recherche de lois générales, au sens de vérifiées pour tous les individus d'une même catégorie définie par rapport à une ou plusieurs variables individuelles, et contribuer ainsi à diversifier les modèles généraux des conduites.

Bibliographie

- Bacher, F. (1982). *Les enquêtes en psychologie* (2 vol.). Lille: PUL.
- Bacher, F. (1987, 1988). Les modèles structuraux en psychologie. présentation d'un modèle. LISREL; Première partie. *Le Travail humain*, 50, 347-370; Deuxième partie. *Le Travail humain*, 51, 273-288.
- Cardinet, J. (1997). De la théorie classique des tests à la théorie de la généralisabilité: l'apport de l'analyse de la variance. *Revue Européenne de Psychologie Appliquée*, 47, 197-204.
- Cooper, L.A. (1976). Individual differences in visual comparison processes. *Perception and Psychophysics*, 19, 433-444.
- Dickès, P. (1996). L'analyse factorielle linéaire et ses deux logiques d'application. *Psychologie Française*, 41, 9-22.
- Dickès, P., Tournois, J., Flieller, A., & Kop, J.-L. (1994). *La psychométrie*. Paris: PUF.
- Huteau, M. (1987). *Style cognitif et personnalité. La dépendance-indépendance à l'égard du champ*. Lille: PUL.
- Huteau, M. (1995). *Manuel de psychologie différentielle*. Paris: Dunod.
- Huteau, M., & Lautrey, J. (1997). *Les tests d'intelligence*. Paris: Editions La Découverte.

- Juhel, J. (1990). Analyse confirmatoire des relations entre visualisation spatiale et mémoire visuo-spatiale. *Psychologie et Psychométrie*, 11(3), 28-42.
- Juhel, J. (1994). Opérationnalisation et recherche de validation: une alternative? In M. Huteau (Ed.). *Les techniques psychologiques d'évaluation des personnes* (pp. 164-171). Issy-Les-Moulineaux: EAP.
- Juhel, J. (1997). Fonctionnement des systèmes psychologiques et modélisation de la variabilité. In J. Juhel, T. Marivain et G. Rouxel (Eds.). *Psychologie et différences individuelles: questions actuelles* (pp. 35-40). Rennes: PUR.
- Larivée, S., Normandeau, S., et Parent, S. (1996). La filière francophone de la psychologie développementale différentielle. *L'Année Psychologique*, 96, 291-342.
- Lautrey, J. (1990). Esquisse d'un modèle pluraliste du développement cognitif. In M. Reuchlin, J. Lautrey, C. Marendaz et T. Ohlmann (Dir.). *Cognition: l'individuel et l'universel* (pp. 185-216). Paris: PUF.
- Lautrey, J. (Ed.) (1995). *Universel et différentiel en psychologie*. Paris: PUF.
- Lautrey, J. (1997). Théorie de l'évolution et psychologie différentielle: vers une nouvelle phase? In M. Launay (Ed.), *Actes des XI^{èmes} Journées de Psychologie différentielle* (pp. 217-235). Université Paul-Valéry, Montpellier.
- Loarer, E., Chartier, D., Huteau, M. & Lautrey, J. (1995). *Peut-on éduquer l'intelligence?* Bern: Peter Lang.
- Ohlmann, T. (1995). Processus vicariants et théorie neutraliste de l'évolution: une nécessaire convergence. In J. Lautrey (Dir.). *Universel et différentiel en psychologie* (pp. 77-105). Paris: PUF.
- Reuchlin, M. (1963). La notion de dimension en psychologie différentielle. *Psychologie Française*, 8, 80-95.
- Reuchlin, M. (1976). *Précis de statistique*. Paris: PUF.
- Reuchlin, M. (1978). Processus vicariants et différences individuelles. *Journal de Psychologie*, 2, 133-145.
- Reuchlin, M. (1992). *Introduction à la recherche en psychologie*. Paris: Nathan.
- Reuchlin, M. (1997). *La psychologie différentielle* (nouvelle édition). Paris: PUF.
- Richelle, M. (1995). Eloge des variations. In J. Lautrey (Dir.). *Universel et différentiel en psychologie* (pp. 35-50). Paris: PUF.
- Rouxel, G. (1997). *Déterminants cognitivo-affectifs de la performance scolaire: approche interactionniste*. Thèse de Doctorat de psychologie. Université Rennes 2.