

Dégagement d'une taxonomie en mathématiques et valorisation d'un logiciel d'évaluation

Rapport final d'un projet de recherche effectué dans le cadre d'une bourse de formation-recherche du Ministère de la Culture, de l'Enseignement Supérieur et de la Recherche

Jang SCHILTZ

Centre Universitaire du Luxembourg
Département des Sciences

Août 2001

Ce travail a été fait sous la direction de Paul Dickes, professeur à l'Université Nancy II, en collaboration avec Romain Martin, psychologue-chercheur de la cellule de recherche en évaluation à l'ISERP.

Nous remercions vivement la direction du Lycée Michel Rodange pour la permission d'effectuer nos tests avec six classes de septième de cet établissement ainsi que son service de psychologie et d'orientation scolaire de pour son assistance lors de l'exécution de ces tests.

TABLE DES MATIERES

CHAPITRE I : PRESENTATION DU PROJET DE RECHERCHE	1
1.1 Objectifs	1
1.2 Théories et procédés proposés.	2
1.3 Intérêt scientifique	4
1.4 Résultats escomptés	5
 CHAPITRE II : L'ETAT DE LA RECHERCHE FONDAMENTALE.....	 7
2.1 La psychologie cognitive.	7
2.1.1 Les précurseurs psychométriques	7
2.1.2 L'intelligence artificielle et le traitement de l'information	8
2.1.3 Les sciences cognitives	8
2.1.4 L'apport de la neuropsychologie	9
2.2 La psychologie développementale.	10
2.2.1 Piaget	10
2.2.2 Vigotski	10
2.2.3 La théorie de la Gestalt	10
2.3 L'apport de la psycho-linguistique aux sciences de l'éducation.	11
2.3.1 La notion de compétence chez Chomsky	11
2.3.2 L'acception pédagogique courante	12
2.3.3 L'extrapolation des critères de Chomsky	13
 CHAPITRE III : APPLICATION AUX MATHEMATIQUES	 14
3.1 Psychologie de l'éducation et psychologie développementale	14
3.1.1 Bilan de la recherche	14
3.1.2 L'apprentissage des mathématiques	19

3.2 L'approche psychométrique	23
3.2.1 L'aptitude numérique	23
3.2.2 Le raisonnement mathématique	26
3.2.3 Autres compétences mathématiques	28
3.2.4 L'étude de J.B. Carroll	30
3.2.5 Conclusion	37
3.3 L'approche cognitive	39
3.3.1 L'aptitude numérique	41
3.3.2 Le raisonnement mathématique	46
3.3.3 L'étude du traitement de l'information	49
3.4 La notion de compétence en mathématiques	51

CHAPITRE IV : APERCU SUR LES TAXONOMIES EXISTANTES 52

4.1 La taxonomie de Bloom	53
4.2 L'Etude Internationale des Performances en Mathématiques	62
4.3 L'étude Longitudinale nationale des aptitudes mathématiques	63
4.4 La taxonomie de M. Pellerey	64
4.5 La taxonomie de la TIMSS	66
4.6 La taxonomie de John B. Carroll	67
4.6 Synthèse	70

CHAPITRE V : METHODES MULTIVARIEES D'ANALYSE STATISTIQUE ... 71

5.1 L'échelonnement multidimensionnel	71
5.2 L'analyse hiérarchique ascendante.	80
5.2.1 Introduction	80
5.2.2 Les indices de dissimilarité	82
5.2.3 La constitution des groupes	84
5.3 L'analyse factorielle	86
5.3.1 Généralités	86
5.3.2 L'analyse en composantes principales	88
5.3.3 Le nombre de facteurs	91
5.3.4. Le choix de la solution factorielle et les rotations	92
5.4. L'analyse de la variance	93
5.4.1 Introduction	93
5.4.2 L'analyse de variance à un facteur (One-Way ANOVA)	94
5.4.3 L'analyse de variance à deux facteurs (Two-Way ANOVA)	97

5.5 Le LISREL	101
5.5.1 Généralités	101
5.5.2 Le modèle mathématique	102
5.5.3 Les représentations multifactorielles	104
CHAPITRE VI : ANALYSE DES TAXONOMIES EXISTANTES	108
6.1 Démarche méthodologique	108
6.2 L'échelonnement multidimensionnel de nos données	109
6.3 Analyse des taxonomies existantes	114
6.3.1 Appartenance de nos items aux différentes taxonomies analysées	114
6.3.2 Echelonnement multidimensionnel de nos données	117
CHAPITRE VII : ANALYSE DES RESULTATS DU TEST	121
7.1 Analyse hiérarchique ascendante	121
7.2 Analyse factorielle	129
7.3 Proposition d'une taxonomie	134
CHAPITRE VIII : ANALYSE CONFIRMATOIRE DE NOTRE TAXONOMIE ...	137
8.1 Analyse de la taxonomie par le LISREL	137
8.2 Analyse de la variance	140
8.3 Interprétation psychologique de notre taxonomie	146
8.4 Conclusion	147
CHAPITRE IX : ETUDE DE VALIDATION DE NOTRE LOGICIEL	150
9.1 Introduction	150
9.1.1 Hypothèses	150
9.1.2 Construction de la forme expérimentale du test	151
9.1.3 Description du dispositif expérimental	151

9.2 Comparaison des scores et capacité de classification	152
9.3 Etude de l'effet de l'ordre de passation	156
9.4 Analyse de quelques items	159
9.5 Conclusion	160

ANNEXE : LA BANQUE D'ITEMS DU LOGICIEL.	163
--	------------

BIBLIOGRAPHIE	180
--------------------------------	------------

Chapitre I

PRESENTATION DU PROJET DE RECHERCHE

Notre projet de recherche se compose de deux parties distinctes. D'abord, le test adaptatif informatisé en mathématiques, développé au cours du projet de recherche postdoctoral BFR 98/027, est soumis à une étude de valorisation permettant de démontrer que ses qualités de puissance, de flexibilité et d'économie d'utilisation ne vont pas au dépens des qualités psychométriques classiques inhérents aux tests papier-crayon.

Une deuxième partie du projet de recherche est consacrée au dégagement d'une taxonomie en mathématiques. L'étude de la littérature existante permet de proposer une taxonomie hypothétique dont l'adéquation est vérifiée par rapport aux données obtenues par la création d'un test expérimental, au moyen de différents procédés statistiques complémentaires entre eux. L'interprétation de la structure cognitive latente tient compte de la notion chomskyenne de compétence et de la conception de l'intelligence multiple de Howard Gardner et s'inscrit dans l'évolution actuelle des sciences cognitives.

1.1 Objectifs

Au cours du projet de recherche postdoctoral intitulé "Application d'un modèle stochastique à la construction d'un test adaptatif informatisé en mathématiques" (BFR 98/027), nous avons appliqué un modèle IRT (item response theory) à la construction d'un

test adaptatif informatisé en mathématiques. Le point de départ était l'analyse et le traitement de données recueillies en novembre 1996, lors de la passation de tests standardisés pour le passage primaire post-primaire.

En adaptant les algorithmes mathématiques adéquats, nous avons construit un logiciel permettant de présenter des items, sélectionnés par l'ordinateur, adaptés au niveau de la compétence actuelle du sujet, de telle sorte que le test ne soit ni trop facile, ni trop difficile, épargnant à la fois la frustration et l'ennui à l'élève.

Le but du nouveau projet de recherche est double

- 1) Faire une étude de valorisation de ce logiciel et l'adapter aux besoins différenciés des instituteurs.
- 2) Construire une taxonomie des compétences en mathématiques

Le but de l'étude de valorisation est de montrer que le test adaptatif informatisé a les mêmes qualités psychométriques (capacité de discrimination, validité, fidélité, objectivité) qu'un test "papier-crayon" classique et que ses qualités supérieures de puissance, de flexibilité et d'économie d'utilisation ne vont pas au dépens des qualités psychométriques classiques.

Le but de la taxonomie est de trouver la structure latente des compétences entrant en action lors de la résolution des épreuves en mathématiques et de créer une forme expérimentale permettant de vérifier, au moyen d'une analyse hiérarchique ascendante, d'une analyse factorielle et d'un échelonnement multidimensionnel si les difficultés des items dépendent effectivement des facteurs latents trouvés.

1.2 Théories et procédés proposés

Du point de vue théorique, l'étude de valorisation est basée sur les données actuelles de la psychométrie (Dickes et al. 1994) et de la méthodologie expérimentale (Brink et Wood 1998, Adèr et Mellenbergh 1999).

Pour comparer les qualités du test informatisé à celles du test papier-crayon, nous présentons les deux formes à un échantillon de 6 classes d'élèves de sixième primaire, c'est-à-dire à environ 120 élèves. Afin de neutraliser l'effet de l'ordre de passation et de répétition, ils

sont attribués au hasard à 2 sous-groupes dont l'un commence par le test papier-crayon et l'autre par le test informatisé. Nous analysons les correspondances et les différences entre les deux situations, au moyen des statistiques inférentielles, pour mettre en évidence la spécificité du test informatisé et pour montrer que son économie d'application ne va pas au dépens de ses qualités psychométriques.

L'établissement de taxonomies est un problème très actuel aussi bien en psychologie cognitive que dans les sciences de l'éducation. Elle est à la base des nouvelles sciences cognitives qui constituent ce nouveau secteur du savoir qui, touchant à la fois à la philosophie, à la psychologie, à l'anthropologie et à la biologie, cherche à élucider les principes qui règlent la connaissance humaine, à la lumière des nouvelles révolutions scientifiques et techniques nées de l'informatique, de la cybernétique et des neurosciences. La structure latente des compétences est dégagée au moyen de l'analyse hiérarchique ascendante, de l'analyse factorielle ainsi que de procédés statistiques multivariés de la deuxième génération comme l'échelonnement multidimensionnel et le Lisrel (Linear Structural Relations), permettant d'en étudier la hiérarchie. En effet, plusieurs aptitudes indépendantes entre elles peuvent contribuer à la résolution d'un même problème et un même problème peut être résolu de manière différente selon les sujets.

Alors que l'analyse hiérarchique ascendante permet de constituer un classement hiérarchisé psychologiquement signifiant, l'analyse factorielle permet de dégager des facteurs de second ordre et l'échelonnement multidimensionnel de représenter les données dans l'espace en les regroupant suivant la plus petite dimensionnalité possible.

Le Lisrel sert à tester l'adéquation entre une structure théorique supposée et la structure immanente aux données recueillies. L'interprétation théorique de la structure des compétences tiendra compte des données actuelles de la psychologie cognitive.

Après un aperçu historique sur la recherche en psychologie cognitive, en psychologie développementale et en éducation des mathématiques, nous analysons les taxonomies existantes.

L'analyse de la littérature historique et actuelle nous permet de proposer une taxonomie théorique pouvant éventuellement s'appliquer à nos données. Pour établir cette taxonomie, nous tenons compte également des besoins de la pratique pédagogique, qui se dégageront d'une concertation avec le groupe d'instituteurs, professeurs, inspecteurs et psychologue-chercheurs qui élaborent les épreuves standardisées dans le cadre de la procédure de passage primaire-postprimaire.

Pour vérifier l'adéquation de notre taxonomie aux données empiriques, nous construisons un test expérimental comportant plusieurs items dans chaque catégorie de la taxonomie. Notre taxonomie théorique est valable si les procédés multivariés utilisés montrent que les difficultés des items dépendent des catégories de la structure théorique latente postulée.

1.3 Intérêt scientifique

L'étude de valorisation permet de comparer les qualités psychométriques du test adaptatif informatisé avec celles des épreuves standardisées classiques. Si ces résultats sont concluants, l'utilité du test adaptatif informatisé que nous avons construit sera démontré de manière scientifique et le logiciel pourra être exploité pour l'orientation et l'évaluation formative et cognitive des savoir et savoir faire des élèves.

Par l'étude de la taxonomie, notre recherche se situe dans le cadre actuel des sciences cognitives telle quelles ont été définies ci-dessus. Le dégagement de la structure latente des compétences en mathématique nous permet de rapprocher la notion de compétence en mathématique de la compétence définie par Chomsky pour le domaine linguistique et nous pouvons examiner si elle a éventuellement des caractéristiques analogues.

En effet, d'après Chomsky (Reboul 1995), la compétence possède les caractéristiques suivantes :

Elle s'appuie sur un code qui est essentiellement restrictif ou négatif, montrant ce qui est défendu mais laissant par ailleurs une grande marge de liberté.

Elle permet de produire des performances en nombre illimité et de façon imprévisible.

Les performances produites doivent être cohérentes entre elles et adaptées à la situation.

La compétence présuppose à la fois des savoir et savoir-faire et des capacités sans s'y réduire.

Les résultats de notre recherche peuvent donc apporter une contribution à la question débattue actuellement en psychologie cognitive, si la conception chomskyenne de la compétence, développée pour le domaine linguistique, peut s'élargir au domaine des mathématiques.

1.4 Résultats escomptés

Cette étude pourra apporter des résultats intéressants, aussi bien du point de vue pratique que théorique.

Du point de vue pratique, la valorisation scientifique de notre logiciel permettra de passer à la phase d'application pédagogique. La taxonomie pourra servir de guide pour l'élaboration des futures épreuves standardisées, élaborées dans le cadre du passage primaire-postprimaire. En effet elle permettra des tests plus conformes au fonctionnement cognitif des élèves. D'autre part, elle donnera une base scientifique à l'évaluation formative et sommative continue des élèves en mathématiques.

Du point de vue théorique, le dégagement de la taxonomie se situe dans la discussion actuelle sur la hiérarchie des compétences en mathématiques et ouvrira la voie à des généralisations théoriques et des vérifications expérimentales ultérieures. En particulier, la possibilité de concevoir les compétences en mathématiques selon un modèle chomskyien permettra d'en approcher différemment l'enseignement de base.

De plus, cette recherche constitue une contribution originale luxembourgeoise par rapport à un problème de recherche en sciences cognitives très actuel au niveau international.

En plus des applications pédagogiques immédiates, elle permettra de passer à l'exploitation de toutes les possibilités du logiciel, comme le transfert à d'autres données et à d'autres populations.

Elle permettra en outre de développer des épreuves standardisées regroupées selon les compétences sous-jacentes à examiner, donnant la possibilité d'obtenir un profil des compétences de chaque élève et de l'aider d'une manière plus spécifique par rapport à ses faiblesses en développant des stratégies compensatoires adéquates. En effet, il ne suffit pas de dire qu'un élève est faible en mathématiques, mais il faut savoir exactement quelles compétences latentes sont en jeu. D'un autre côté, il sera plus facile de rassurer un élève démotivé ayant subi des échecs scolaires, par rapport à ses compétences latentes, si l'on dispose d'un instrument de mesure se situant à mi-chemin entre les tests d'intelligence classiques, trop éloignés de la pratique pédagogique et des épreuves scolaires standardisées, trop axés sur les programmes et ne permettant pas le dépistage des compétences réelles.

Une application particulièrement intéressante est le dépistage de “l’underachievement”, à un moment, où une évolution défavorable et un blocage pernicieux de certains élèves surdoués peuvent encore être évités.

Chapitre II

L'ETAT DE LA RECHERCHE FONDAMENTALE

Aperçu historique

2.1 La psychologie cognitive

2.1.1 Les précurseurs psychométriques

Les premières études ont consisté en recherches factorielles sur la structure de l'intelligence. Alors que Spearman a développé sa théorie bifactorielle des processus mentaux, expliquant les corrélations entre les performances intellectuelles par une combinaison d'un facteur d'intelligence générale (g) et de facteurs de groupe ou spécifiques (s) dépendant de l'épreuve en question, Thurstone (1938) lui a opposé sa théorie des facteurs multiples, extrayant 7 aptitudes mentales primaires (primary mental abilities) à savoir le calcul, la fluidité verbale, la compréhension verbale, la mémoire, le raisonnement, la représentation spatiale et la vitesse perceptive. D'autres chercheurs, comme par exemple Guilford (1956), ont dégagé un nombre beaucoup plus grand d'aptitudes élémentaires ; certains chercheurs estiment même qu'on pourra multiplier à l'infini les aptitudes spécifiques, en relation avec la nature de la tâche, ce qui soulève la question de la structure hiérarchique des aptitudes mentales.

2.1.2 L'intelligence artificielle et le traitement de l'information

Les méthodes informatiques de traitement de l'information ont permis un progrès important dans la modélisation cognitive (Richard, 1998).

A la fin des années 50, les chercheurs croyaient pouvoir étudier l'ensemble des processus cognitifs à l'oeuvre dans la résolution des problèmes, en les exprimant dans un langage formalisé permettant le calcul et la simulation à l'ordinateur, de manière à pouvoir comparer les comportements simulés avec les comportements observés. Cependant, après une période d'affirmations exagérées (années 60-70, à la suite du colloque de Dartmouth), les théoriciens de l'intelligence artificielle sont revenus à une vision plus mesurée. Ils ont compris que certains champs en sont exclus, ce qui laisse subsister la possibilité d'obtenir des résultats solides dans d'autres domaines. D'après Howard Gardner (1993), l'ordinateur est un modèle adéquat pour comprendre certains processus de la pensée, mais il n'est probablement pas le meilleur modèle pour la compréhension des processus les plus importants. Et c'est en étudiant la pensée « computationnelle » que les chercheurs ont découvert à quel point les êtres humains sont différents de l'ordinateur sériel digital, à partir du moment où il s'agit d'exécuter des tâches complexes.

2.1.3 Les sciences cognitives

A partir des années 70, les interactions entre la psychologie, la philosophie, la biologie, l'anthropologie, la linguistique et l'informatique ont contribué au développement de la discipline qu'on commençait à appeler les « sciences cognitives », cherchant à comprendre les principes qui règlent la connaissance humaine à la lumière des découvertes scientifiques et techniques récentes, dans le but de développer une théorie de la représentation générale, capable de modéliser et d'expliquer globalement le fonctionnement cognitif humain. Au cours des années 80-90, la recherche s'est concentrée sur des contenus cognitifs spécifiques et, en particulier, un corpus de recherche très important s'est développé autour du domaine des mathématiques.

D'après de Corte, Greer et Verschaffel (1996), il devient de plus en plus évident que la pensée mathématique n'est pas uniquement basée sur la logique, mais qu'il faut tenir compte d'aspects non rationnels tels que l'intuition, les croyances, les attitudes, les émotions et la pensée imagée.

La théorie de « l'intelligence multiple » est une illustration de cette conception. Howard Gardner (1993b) distingue sept domaines d'aptitudes (« frames of mind ») : les intelligences linguistique, logico-mathématique, spatiale, musicale, somato-kinesthétique, introspective (connaissance de soi) et interpersonnelle (connaissance des autres). Le caractère modulaire de ces intelligences signifierait que la capacité d'un individu dans un domaine ne laisse pas prévoir sa capacité dans un autre. En collaboration avec Robert Sternberg, Howard Gardner a élaboré un programme axé sur le développement de l'intelligence pratique, apprenant aux élèves comment s'organiser, comment réaliser une tâche et comment s'entendre avec les autres (Sternberg et alt., 1995). Si la conception de Gardner a le mérite de placer l'intelligence dans son contexte pratique et culturel, elle a cependant soulevé des critiques. La principale, c'est qu'il sous-estime les aspects de l'intelligence générale qui ne sont pas spécifiques à un domaine telle que la capacité d'avoir rapidement une vision synthétique de n'importe quel domaine nouveau ou de la rapidité du fonctionnement cognitif en général.

2.1.4 L'apport de la neuropsychologie

Etudiant les relations entre le fonctionnement cognitif d'une part et le fonctionnement et la structure du système nerveux d'autre part, cette discipline permet de tester les modèles théoriques mis au point pour rendre compte des performances des sujets neurologiquement sains. En étudiant les performances de patients dont le cerveau présente des lésions, elle obtient des indications sur les fonctions des parties cérébralement atteintes, ce qui permet de compléter les études neurophysiologiques et neuroanatomiques du système nerveux central (électroencéphalographie, imagerie par résonance magnétique, tomographie cérébrale, etc...) chez un sujet qui est en train d'accomplir une tâche cognitive. La contribution de la neuropsychologie est particulièrement intéressante pour comprendre le fonctionnement cérébral lors du

traitement de l'information (encodage, recherche, comparaison, prise de décision, organisation de la réponse) et pour étudier les niveau de vigilance et le fonctionnement de la mémoire (mémoire implicite et explicite, mémoire à court terme ou mémoire de travail et mémoire à long terme, mémoire procédurale, imagée et lexicale). Delacour (1998) présente une synthèse récente de ces contributions.

Des travaux actuels montrent à quel point les émotions affectent la cognition à tous les niveaux (cf. Le Doux, 1996 ; Halberstadt & Niedenthal, 1997).

2.2 La psychologie développementale

2.2.1 Piaget

Piaget a été un des principaux promoteurs de l'idée d'après laquelle la représentation de la réalité est construite, la connaissance résultant du processus d'adaptation qui vise à donner un sens à l'environnement. L'enfant assimile l'information avec les schèmes dont il dispose à la naissance et il accomode ces schèmes en se basant sur son expérience. Une autre contribution centrale de Piaget est que la pensée est qualitativement différente à certains moments de la vie et que l'évolution se fait par stades. L'étude de la constitution des concepts et des procédures mathématiques occupe une place importante dans l'oeuvre de Piaget (Piaget, 1947 ; Piaget & Inhelder, 1941, 1948, 1959 ; Piaget & Szermiska, 1941).

2.2.2 Vigotsky

Vigotsky a souligné l'interaction entre le développement biologique et la culture environnante dans le développement cognitif.

Des fonctions naturelles élémentaires sont transformées en fonctions mentales d'un niveau plus élevé, grâce à l'interaction avec des membres plus évolués de la communauté et grâce à l'utilisation d'outils intellectuels propres à l'environnement socio-culturel. Les théories de Vigotsky ont influencé les sciences de l'éducation en Occident ; la théorie d'après laquelle l'éducation des mathématiques consiste dans l'introduction des enfants dans la communauté des mathématiciens dérive de cette

conception. D'après Vigotsky, l'intelligence, c'est la capacité de tirer profit d'une instruction.

2.2.3 La théorie de la Gestalt

Les conceptions actuelles sur les processus en jeu dans la résolution des problèmes remontent à la théorie de la Gestalt, dont l'idée centrale est que l'esprit humain interprète les données perçues en accord avec des principes d'organisation qui le prédisposent à voir des figures prégnantes ressortant sur un fond, plutôt qu'une collection de sensations atomisées ; d'où le développement d'une conception structuralisante de la pensée. George Polya (1965), par exemple, un des premiers théoriciens de la compréhension structurale des processus en jeu dans la résolution des problèmes, s'est directement inspiré de l'oeuvre de Wertheimer (1945).

2.3 L'apport de la psycho-linguistique aux sciences de l'éducation

2.3.1 La notion de compétence chez Chomsky

Noam Chomsky (1965) a utilisé l'expression « compétence linguistique » pour souligner le fait que les enfants apprennent à parler correctement, malgré le fait que dans leur entourage, ils entendent surtout des phrases lacunaires, inachevées, incorrectes. Défendant une position innéiste et rationaliste, par opposition à l'empirisme et au behaviorisme, il voulait montrer que l'apprentissage d'une langue ne peut pas se faire uniquement par imitation, mais présuppose l'existence de structures syntaxiques permettant de comprendre et de former un nombre illimité de phrases nouvellement construites.

D'après Reboul (1994), la compétence linguistique possède donc les caractéristiques suivantes :

- 1) Elle s'appuie sur un code qui est essentiellement restrictif ou négatif, montrant ce qui est défendu mais laissant par ailleurs une grande marge de liberté.

- 2) Elle permet de produire des performances en nombre illimité et de façon imprévisible.
- 3) Les performances produites doivent être cohérentes entre elles et adaptées à la situation.

Au début, Chomsky s'est surtout intéressé à la structure syntaxique de la langue, ensuite, ses recherches se sont élargies aux relations entre la syntaxe, la sémantique et la phonologie. Aujourd'hui le terme de compétence linguistique est devenu synonyme de « compétence communicative » et s'applique tout aussi bien au domaine syntaxique que morphologique, phonologique, sémantique, stylistique et expressif. Elle inclut des facteurs non verbaux permettant de comprendre le sens d'un énoncé, tels que l'intonation, la musique, le geste.

2.3.2 L'acception pédagogique courante

D'un côté, la notion de compétence a acquis une importance capitale dans la littérature pédagogique, puisque l'acquisition des compétences est présenté comme étant le but principal de l'enseignement. D'un autre côté, les pédagogues utilisent le terme de manière ambiguë, le confondant souvent avec les capacités, les savoirs, et les savoir-faire. Reboul (1995) souligne que la compétence présuppose des capacités, des savoir et des savoir-faire, mais ne s'y réduit pas. Les définitions données sont bien plus souvent négatives que positives. La compétence est une construction théorique inobservable en elle-même ; elle peut cependant être estimée à partir des performances observables et mesurables.

Selon Reboul (1995), une performance témoigne d'une compétence quand elle est conforme à un code imprévisible, libre, cohérente avec d'autres performances de même type et adaptée à la situation. La compétence peut se situer à différents niveaux de généralité. Bien qu'il existe des compétences fondamentales impliquant l'acquisition de certaines habiletés nécessaires pour participer à la culture humaine, telles que l'aisance dans la manipulation du code verbal, imagé, spatial, la constitutions des schèmes psycho-moteurs, des bases du traitement de l'information, de la classification et de la formation des concepts, les pédagogues s'intéressent

surtout à des compétences spécifiques de niveau intermédiaire, s'inscrivant dans un code précis, pour lesquelles la transférabilité n'est pas démontrée. Selon Delannoy et Passegand (1992) la seule compétence transférable, c'est l'attitude métacognitive qui permet à l'élève de juger des stratégies d'apprentissage appropriées aux différentes matières, c'est-à-dire la compétence dans l'acquisition et l'organisation mnémonique des connaissances nouvelles.

Lorsque les programmes officiels parlent de compétences transversales, le terme est généralement utilisé de manière impropre.

2.3.3 L'extrapolation des critères de Chomsky

Une question intéressante pour la recherche en sciences cognitives et en sciences de l'éducation est de savoir si les caractéristiques de la compétence chomskyenne se rencontrent dans des domaines non linguistiques exigeant des compétences spécifiques, tels que les mathématiques, les sciences, les sports, les arts plastiques, la musique, etc...

Puisque chez Chomsky, la compétence se rapproche de l'art de bien juger et qu'il insiste surtout sur l'aspect créateur et dynamique du concept, il serait profitable de refaire son travail dans d'autres domaines. Revenir à la précision, à la complexité et à la richesse de l'acception chomskyenne du terme de compétence permettrait éventuellement de définir des buts pédagogiques plus fructueux que si l'on décrit des hiérarchies de compétences calquées sur les programmes scolaires existants.

Chapitre III

APPLICATION AUX MATHÉMATIQUES

3.1 Psychologie de l'éducation et psychologie développementale

3.1.1 Bilan de la recherche

L'éducation des mathématiques figure parmi les premiers sujets étudiés en psychologie de l'éducation. Dans les années 20, les théories de Thorndike influençaient beaucoup les programmes d'arithmétique élémentaire des écoles. Thorndike faisait partie de l'école du behaviorisme dont le principe central était qu'une théorie psychologique devrait être bâtie seulement sur la base d'observations du comportement ; des processus mentaux, comme, par exemple, les pensées, les sentiments et les intentions sont exclus, car ils ne sont pas observables.

Le behaviorisme était dans une position dominante, surtout aux Etats-Unis, depuis le début du vingtième siècle, jusqu'à la révolution cognitive dans les années 60. Dans un contexte historique, le behaviorisme se situe à l'intérieur d'un mouvement de quête générale visant la précision et l'objectivité scientifique, dont les manifestations incluent également le positivisme logique et des tentatives de créer des fondements logiques inébranlables en mathématiques (De Corte, Greer & Verschaffel, 1996). Le désir de fournir des bases scientifiques à la psychologie menait à une grande rigueur méthodologique et à l'élaboration de procédures servant à quantifier et mesurer les variables psychologiques.

Cependant, la théorie de Thorndike, ainsi que d'autres théories behavioristes plus récentes, échouent généralement à expliquer la structure logique de la discipline

des mathématiques. L'éducation basée sur les principes behavioristes était souvent fragmentée en parties indépendantes qui pouvaient être apprises par un renforcement approprié mais qui n'avaient pas de cohésion interne.

Il y avait assez rapidement des chercheurs comme Judd (1928) et Brownell (1935), qui s'opposaient aux essais initiaux d'appliquer le behaviorisme à l'éducation, argumentant qu'une instruction efficace des mathématiques doit être basée sur la compréhension des concepts de base. Par contre, même si Brownell donna des exemples montrant comment enseigner des compétences spécifiques de façon à ce que les étudiants les comprennent, il ne développa pas de théorie complète à ce sujet.

Malgré la perte de son statut comme système théorique et philosophique dominant dans les cercles académiques, le behaviorisme persiste à influencer le système éducatif et la mentalité des gens en général, ainsi que des responsables devant prendre les décisions politiques et administratives, en particulier. R.B. Davis (1990) a constaté qu'au niveau le plus important, à savoir celui des écoles et des enseignants, l'une des deux approches qui dominaient à l'époque aux Etats-Unis, était celle d'un enseignement qui exige une détermination très explicite de ce que les enseignants veulent que les élèves apprennent, puis une exposition très claire de ce contenu, un nombre considérable d'exercices portant exactement sur ce contenu et enfin, des tests de contrôle sur le même contenu, ce qui représente clairement un héritage du behaviorisme. De plus, le développement du behaviorisme a conduit à l'opinion persistante qu'il faut apprendre des contenus hiérarchisés à l'intérieur des mathématiques, ce qui a eu des effets nocifs, d'après L.B. Resnick (1987).

Au cours des années 60 émergeaient plusieurs théories cognitives et développementales qui permettaient de fournir un lien plus direct entre la théorie du corps du savoir et l'enseignement des mathématiques. Les théories de Ausubel (1968), Brunner (1960, 1966) et Gagné (1965) se concentraient explicitement sur la structure du contenu de ce qui est à apprendre. Dans une perspective différente, Piaget (Piaget 1952 ; Piaget & Inhelder, 1956 ; Piaget, Inhelder & Szeminska, 1960) proposait une théorie du développement des fondements des concepts des nombres entiers, de la mesure et de la géométrie. Piaget a apporté de nouvelles contributions pendant toute sa vie, et sa théorie a encore évolué dans les années 70. Son travail empirique et son analyse théorique liés à tous les grands sous-domaines des mathématiques ont été

fondamentaux et continuent de stimuler des découvertes, des critiques et des réactions.

D'un point de vue contemporain, on peut dire que la contribution principale de Piaget à l'éducation des mathématiques est sa démonstration de la complexité de la pensée des enfants et la mise en évidence des différences qualitatives de la pensée à des étapes diverses de leur développement (De Corte, Greer & Verschaffel, 1996). Son mérite d'avoir inspiré la conception de l'enfant comme constructeur actif de son savoir est reconnu universellement. De plus, il a introduit des innovations méthodologiques, par exemple l'utilisation des entretiens cliniques. De cette façon, il a énoncé l'élargissement des méthodes d'enquête, caractéristiques de la recherche récente en psychologie cognitive et en éducation des mathématiques.

Piaget est cependant généralement critiqué pour la raison qu'il n'a pas assez insisté sur les facteurs sociaux et culturels et plus spécifiquement, qu'il ne s'intéressait pas aux aspects pédagogiques (De Corte, Greer & Verschaffel, 1996). D'après Vergnaud (1990), il n'a jamais étudié le processus « enseigner-apprendre », ni en salle de classe, ni à domicile.

Même si Piaget ne s'est pas explicitement intéressé au problème de l'instruction et même s'il n'a pas étudié directement le développement des concepts et des procédures qui composaient la partie principale des programmes d'étude en mathématiques, les opérations logiques qu'il décrivait semblaient être sous-jacentes aux concepts de base enseignés à l'école primaire. La formation de concepts tels que la conservation, la transitivité et la classification paraissait indispensable à la compréhension de la plupart des concepts liés aux nombres entiers et à la mesure. Un certain nombre d'études essayait d'établir empiriquement comment les opérations logiques de Piaget sont liées à l'apprentissage des concepts de base en mathématiques (Carpenter, 1980 ; Hiebert & Carpenter, 1982). Bien que toutes les études aient montré que la performance aux tâches piagétienne corrèle avec la performance en arithmétique, les opérations logiques de Piaget n'ont généralement pas été jugées utiles pour expliquer les aptitudes des enfants à apprendre la plupart des concepts mathématiques de base (Hiebert & Carpenter, 1982).

Les théories d'Ausubel, Bruner et Gagné pourraient être appliquées plus directement à l'élaboration des programmes d'étude que celles de Piaget. Même si les positions d'Ausubel et Bruner sont probablement plus conformes à l'orientation

théorique actuelle que le neo-behaviorisme initial de Gagné, c'est Gagné qui indiquait le plus clairement comment on pourrait analyser les programmes d'étude en mathématiques et faire de la recherche dans ce domaine (Case & Bereiter, 1982). L'analyse des tâches (*task analysis*) de Gagné fournit une structure méthodologique permettant d'organiser systématiquement une matière d'enseignement selon une hiérarchie de principes, concepts et compétences. L'analyse des tâches a été utilisée dans la conception de plusieurs projets mathématiques élémentaires, comme, par exemple des plans d'instruction individuels (Lindvall & Bolvin, 1967) et des projets de développement des procédures mathématiques (Romberg, Harvey, Moser & Montgomery, 1976).

Un autre théoricien du développement, dont les travaux s'appliquent plus directement à l'éducation est Vigotsky. D'après Bruner (1962), on peut même dire que sa conception du développement est en même temps une théorie de l'éducation. Son travail a été essentiel pour le développement de la psychologie russe en générale et pour les théories de l'éducation mathématique en particulier (cf. Davydov, 1990 ; Kruetskii, 1976). Des idées théoriques se situant dans la tradition de son oeuvre se sont également répandues à l'ouest, ces dernières 15 années (Newman, Griffin & Cole, 1989 ; Van Oers, 1990 ; Wertsch, 1985). Vigotsky proposait une interaction entre le développement biologique et culturel, au moyen de laquelle les fonctions mentales développées naturellement sont transformées en des fonctions mentales supérieures, grâce à l'interaction avec des membres plus expérimentés d'une culture. Bruner (1962) a caractérisé la conception de l'intelligence de Vigotsky comme une « capacité de profiter de l'enseignement ».

Les innovations méthodologiques radicales de Vigotsky sont inséparables de ses analyses théoriques fondamentales et de ses critiques des autres écoles psychologiques. Il inventait des méthodes expérimentales permettant d'observer le développement produit par l'enseignement. Cette méthodologie est actuellement utilisée par des chercheurs en éducation, tels que Steffe (1991), lorsqu'il s'agit de travailler avec des étudiants individuels. Dans le contexte de la salle de classe, ces approches sont à la base d'expériences pédagogiques qui jouent actuellement un rôle proéminent dans la recherche didactique en mathématiques (cf. Cobb, Wood & Yackel, 1991 ; Cobb, Wood, Yackel & McNeal, 1992 ; Lampert, 1986 ; Mahler, Davis & Alston, 1991).

La résolution de problèmes constitue clairement un domaine de chevauchement entre la psychologie cognitive et l'éducation mathématique. L'école psychologique de la théorie de la Gestalt était la source essentielle pour la recherche et la théorisation des processus en jeu lors de la résolution de problèmes. L'idée clé de cette école, qui s'intéressait au début surtout à la perception, est que l'esprit interprète les données venant des sens selon des principes d'organisation, de sorte que l'être humain perçoit des formes (Gestalt) plutôt qu'une collection de perceptions atomisées. Cette idée a été étendue plus tard à la pensée en général et à la résolution de problèmes, avec l'accent mis sur la structure. Pour des raisons de pertinence et de commodité, il était naturel que la plus grande partie du travail des psychologues de la Gestalt concernât des sujets mathématiques. Wertheimer (1945), par exemple, a cité l'exemple historique du grand mathématicien Gauss, qui, comme jeune élève, avait simplifié le calcul impliqué dans l'addition des nombres d'une suite arithmétique, grâce à sa compréhension de la structure d'une telle suite. Mais la source d'idées la plus importante concernant la résolution de problèmes mathématiques est le travail de George Polya (1945, 1954a, 1954b, 1962, 1965). Même s'il était un mathématicien plutôt qu'un psychologue, il était influencé par la théorie de la Gestalt et son analyse profonde de la compréhension structurale et des processus de découverte en mathématique ont eu une grande influence sur la recherche en éducation des mathématiques.

Polya essayait de présenter les mathématiques selon leur ordre de développement naturel. Il croyait qu'il faut savoir comment un théorème a été découvert pour pouvoir vraiment le comprendre (cf. Albers & Alexanderson, 1985 ; Tymoczko, 1986). Dans *How to solve it* (1945), il présenta un certain nombre de directives heuristiques pour la résolution de problèmes. Beaucoup d'essais ont été faits pour utiliser ces directives, d'une manière plus ou moins directe, dans le but d'enseigner des compétences générales de résolution de problèmes. Stanic et Kilpatrick (1980) s'opposent à ce genre de distortion, où « une heuristique devient une compétence, une technique et même, paradoxalement, un algorithme ». Polya lui-même affirmait qu'il n'existe pas de raccourci pour apprendre l'art de résoudre des problèmes : comme dans d'autres arts, il faut l'apprendre en pratiquant et en réfléchissant sur la manière de le faire, sous la direction d'un enseignant expert en la matière.

3.1.2 L'apprentissage des mathématiques

Wittrock (1974) était un des premiers à signaler les implications importantes de la psychologie cognitive dans la recherche en éducation des mathématiques.

A partir des années 80, l'accent de la recherche s'est déplacé et l'on tient compte des processus cognitifs internes. L'analyse rationnelle des tâches, basée sur une analyse logique d'experts a évolué en direction d'une analyse empirique, qui se concentre sur ce que les enfants font vraiment quand ils résolvent des problèmes en situation naturelle (Romberg & Carpenter, 1986).

L'une des hypothèses principales au cours des années 80 est que les enfants construisent activement leurs connaissances grâce à l'interaction avec leur environnement et grâce à la réorganisation de leur propre concept mental (Steffe, von Glasersfeld, Richards & Cobb, 1983). Bien que l'éducation affecte clairement ce que les enfants apprennent, elle ne détermine pas complètement le processus d'apprentissage. Les enfants ne sont pas des récipiends passifs ; ils interprètent la matière exigée, l'organisent et l'assimilent à la lumière de leur propre structure mentale (Wittrock, 1974). Déjà au cours des années 70, certains chercheurs ont souligné que les enfants inventent leur propre mathématique (Resnick, 1976) et qu'ils viennent à l'école avec des systèmes mathématiques informels bien développés (Ginsburg, 1977). Quelques-uns des meilleurs exemples de la capacité d'invention des enfants viennent de la recherche sur leurs stratégies pour effectuer des additions et des soustractions.

Ces études suggèrent qu'il n'est pas nécessaire de différer la résolution des problèmes jusqu'à ce que les techniques de calcul soient maîtrisées et qu'elle pourrait être beaucoup plus intégrés dans les programmes d'étude que c'est le cas actuellement. En fait, il a été proposé d'utiliser les problèmes pour développer les concepts de l'addition et la soustraction plutôt que d'enseigner d'abord les techniques de calcul et de les utiliser ensuite pour résoudre des problèmes (Carpenter & Moser, 1983).

Jusqu'à il y a une dizaine d'années, on disait souvent, qu'à l'école primaire, les mathématiques doivent être limitées à des buts cognitifs modestes, en raison de l'immaturité développementale des enfants qui imposerait des limites cognitives strictes à leur capacité d'apprentissage. D'après Davis & Goffree (1991), cet argument

est clairement contredit par les recherches des dernières années. Du projet coréen Chisanbop à l'étude californienne SEED, chaque fois que des enfants ont eu l'occasion d'apprendre les mathématiques véritables, la plupart d'entre eux se sont bien débrouillés. Les auteurs sont d'avis que la vision réductrice des mathématiques élémentaires reflète peut-être les buts limités qu'une société fixe pour ces citoyens, ainsi que la capacité limitée des systèmes d'éducation, mais elles ne reflètent sûrement pas le degré de fonctionnement cognitif des élèves. Bien sûr, des limites cognitives existent, mais elles correspondent surtout à la restriction des expériences qu'un enfant peut avoir fait à un moment déterminé de sa vie. Si l'on tient compte de ces limitations d'une façon adéquate, la plupart des enfants sont capables de faire de grands progrès. Actuellement, les écoles primaires qui offrent la possibilité d'apprendre en quoi un raisonnement mathématique consiste vraiment sont rares dans le monde.

L'hypothèse de loin la plus répandue est, qu'à l'école primaire, les mathématiques consistent en des règles de calcul et des algorithmes qui doivent être mémorisés. La méthode courante consiste à les présenter ou démontrer aux élèves, après quoi ceux-ci répètent les règles et les algorithmes jusqu'à ce qu'ils soient capables de le faire correctement (David & McKnight, 1980). Cette interprétation de ce que « connaître » ses mathématiques veut dire, même si elle est très répandue, presque jusqu'à l'université, est en fait très controversée. Un tel savoir ou de telles capacités procédurales dénuées de sens ont déjà été rejetés par Bruner & Erlwanger (1973), Alderman, Swinton & Braswell (1979) et beaucoup d'autres, au moins depuis Dewey.

Au cours des quatre dernières décades peuvent être identifiées au moins trois écoles de pensée importantes concernant la nature de l'apprentissage des mathématiques, ainsi que de nombreuses variations et alternatives moins répandues.

Premièrement, il y a la doctrine didactique de l'apprentissage par coeur, décrite ci-dessus. Deuxièmement la doctrine de la structure mathématique abstraite (*abstract mathematical structure*) qui affirme que la vraie nature des mathématiques ne peut pas se révéler si l'on essaye de retracer les chemins chronologiques de leur développement, avec ses détours, ses impasses et ses concepts bizarres, mais qu'il faut au contraire commencer par les concepts abstraits simples et précis de l'analyse moderne, comme la notion d'ensemble, de groupe, d'espace vectoriel et ainsi de suite.

Ce point de vue était très populaire dans les années 60 et le début des années 70. De nos jours, il a perdu du terrain par rapport à la troisième école, celle du point de vue constructiviste.

L'école constructiviste, basée en partie sur les travaux de Piaget et exposée plus récemment, entre autres, par Papert (1980) suppose que les connaissances sont connectées entre elles et que de nouvelles représentations peuvent être acquises seulement à condition de pouvoir être reliées à des représentations qui existent préalablement. Les idées mathématiques, en particulier, sont basées en dernière analyse sur des expériences de l'enfance, comme par exemple l'expérience de l'horizontalité par opposition à la verticalité, le comportement des fluides sous l'influence de la gravité, le comportement d'une collection d'objets physiques quand on les déplace etc... . Cette hypothèse implique que certaines choses sont apprises facilement et d'autres non, mais le critère de simplicité provient du développement biologique et non de concepts abstraits tels que « groupe de transformation » ou « ensemble ». Pour les constructivistes, la simplicité est basée sur les connaissances antérieures, qui sont le plus souvent concrètes ou du moins empiriques. Du point de vue constructiviste, il est important que les enfants aient fait des expériences appropriées, fonctionnant comme fondement de l'apprentissage ultérieur.

Si l'on juge d'après le critère de l'habileté générale en calcul, il est clair, que même dans un grand nombre de sociétés évoluées, les compétences en mathématiques élémentaires sont loin d'être universellement répandues (Davis & Goffree, 1991). Mais, il existe un critère plus pertinent ; en effet, les meilleurs programmes expérimentaux modernes ont souvent montré que des enfants de 8 à 10 ans sont capables d'assimiler des mathématiques réelles, expérimentant avec la pensée mathématique, l'analyse mathématique, la conjecture mathématique, l'amoncellement de preuves, la conception de lignes de raisonnement, la recherche et la découverte de régularités et de motifs clés, et qu'ils en tirent un grand profit. Ce type d'éducation n'est malheureusement offert qu'à un petit nombre d'élèves dans quelques pays, comme, par exemple, la Hongrie, les écoles d'élite des États-Unis et surtout le Japon, qui, du point de vue de l'éducation mathématique est un des pays les plus efficaces du monde. A l'âge de 6 ans, on y commence par la mesure directe des dimensions de différents objets physiques (cf. Hashimoto Y. & Sawada T., 1979). Au Japon, on attribue beaucoup d'importance à la signification des objets et des symboles

mathématiques, l'éducation ne s'y fait pas par simple mémorisation. Le programme d'étude pour les enfants de 8 ans prévoit le début de l'étude de l'algèbre ; la multiplication et la division sont connues à cet âge là ; les enfants sont familiers avec les fractions décimales ; ils ont quelque connaissance de l'idée de fonction et sont capables de résoudre des problèmes en utilisant le soroban (l'abacus japonais). A neuf ans, les enfants japonais ont des connaissances considérables en géométrie, manipulant des concepts comme la mesure d'angles, la perpendicularité, l'aire, le volume, et ainsi de suite, et ils sont familiers avec de nombreuses formes géométriques en deux et en trois dimensions. Leur compétence en arithmétique inclut le maniement des nombres décimaux et des fractions. Ils savent représenter des fonctions à l'aide d'expressions algébriques, de graphes et de tables de valeurs.

A la fin de l'école primaire, un élève japonais a des connaissances considérables en calcul arithmétique, en géométrie, en mesure de temps, de distances, d'argent, d'aires, de volumes, d'angles et en notions élémentaires d'algèbre. Les enfants sont capables de concevoir leurs propres algorithmes, ils maîtrisent l'utilisation de leur abacus et comprennent la signification de ce qu'ils font. Ils ont également quelques connaissances en statistiques descriptives, ils sont familiers avec la proportionnalité directe et inverse et ont quelque compétence en géométrie descriptive. Ils ont commencé à manier des quantités discrètes et continues et connaissent les concepts d'axe de symétrie, de rapport, de proportion, de prisme, de cylindre, de cône, de pyramide, et ainsi de suite. On attend également d'eux de faire preuve d'un peu d'originalité et d'invention.

D'après des études américaines, le haut niveau dans l'éducation des mathématiques était une des raisons essentielles de la haute productivité industrielle du Japon. Au début des années 80, les Etats Unis, dont l'économie était alors en très mauvaise position, lancèrent un vaste programme éducatif, avec le but de devenir le premier pays mondial pour la qualité de l'enseignement des mathématiques, jusqu'en l'an 2000.

3.2 L'approche psychométrique

L'approche psychométrique de la recherche sur les processus cognitifs consiste à essayer de comprendre la variabilité individuelle dans les aptitudes humaines (Geary, 1994).

Historiquement, cette approche a procédé par l'administration de nombreux tests papier-crayon à des individus et des groupes d'individus, pour identifier ensuite, au moyen d'analyses factorielles et de techniques analogues, des groupes d'épreuves ayant quelques propriétés communes (Spearman, 1927 ; Thurstone, 1938). Ces études permettent de dégager des facteurs de second ordre, décrivant la corrélation entre les différentes épreuves. Un facteur d'aptitudes représente un système de savoir conceptuel ou de processus cognitifs que tous les tests qui le définissent ont en commun. Avant de pouvoir affirmer avec certitude que les tests d'aptitudes correspondent à un ensemble défini de processus cognitifs, il faut démontrer que les mêmes tests se retrouvent dans un grand nombre d'études. Ce critère est appelé stabilité factorielle ou invariance du facteur (Thurstone, 1938). Si ce critère est vérifié, le facteur pourrait représenter un domaine de base de la compétence humaine ; les aptitudes représentées par un facteur peuvent être distinctes d'autres aptitudes cognitives pour des raisons tout aussi bien éducatives et culturelles que biologiques. Ils sont donc indépendants de la controverse nature/culture.

Nous allons présenter quelques études psychométriques portant sur des tests numériques ou mathématiques, dans le but d'identifier des domaines d'aptitudes mathématiques de base et d'analyser les changements développementaux dans ces domaines.

3.2.1 L'aptitude numérique

Dans la littérature psychométrique traditionnelle, l'aptitude numérique est mesurée à l'aide de tests d'arithmétique (Thurstone, 1938). Ces études sont importantes pour deux raisons. Premièrement, si l'on peut démontrer que le facteur correspondant à l'aptitude numérique est présent dans toutes les études, alors cette découverte plaide en faveur de la conception que l'arithmétique représente un

domaine fondamental de la compétence humaine (Gelman & Gallistel, 1978 ; Wynn 1992). Deuxièmement, dès que la stabilité de ce facteur aura été établie, des plans factoriels permettront d'étudier le développement des compétences associées. En particulier, ces techniques peuvent être employées pour déterminer l'âge où l'aptitude en arithmétique émerge comme compétence distincte et pour étudier comment la relation entre l'aptitude numérique et, par exemple, l'aptitude en raisonnement mathématique, évolue avec l'âge de l'enfant.

En ce qui concerne la stabilité du facteur, le facteur numérique ou d'aptitude numérique a été identifiée dans des douzaines d'études psychométriques (voir p.ex. Canisia, 1962 ; Chein, 1939 ; Coombs, 1941 ; French, 1951 ; Goodman, 1943 ; Thurstone, 1938 ; Thurstone & Thurstone, 1941). Un facteur d'aptitude numérique distinct a été trouvé dans études avec des étudiants Américains, Chinois et Philippins (Guthrie, 1963 ; Vandenberg, 1959), aussi bien que dans des analyses avec des groupes séparés d'hommes et de femmes d'âges différents (Dye & Very, 1968). En fait, Pawlick (1966) soutenait que la « compréhension verbale (V) » et l'« aptitude numérique (N) » sont les deux facteurs cognitifs les mieux confirmés. French (1951), dans un examen complet des facteurs d'aptitudes, déclare que le facteur d'aptitude numérique est le plus clairement défini de tous. Même Spearman (1927), qui était pourtant d'avis que les différences individuelles dans les capacités humaines sont le mieux expliquées par des différences en facteur d'intelligence générale (g), écrivait que des compétences arithmétiques de base correspondent à un facteur de groupe très bien défini.

Dans les études fondamentales, le facteur d'aptitude numérique est saturé le plus en tests de calcul arithmétique (p.ex. des tests qui exigent des opérations de multiplication complexes) et en tests nécessitant une compréhension conceptuelle des relations entre les nombres et les notions arithmétiques, mais pas en tests contenant simplement des nombres comme stimulus (Thurstone, 1938 ; Thurstone & Thurstone, 1941). Au fond, le facteur d'aptitude numérique semble contenir toutes, ou presque toutes, les compétences arithmétiques de base (Geary & Widaman, 1987). Du point de vue de la perspective psychométrique, on peut dire que l'aptitude arithmétique ou numérique représente une compétence unique (Thurstone, 1938). En d'autres mots, les études psychométriques suggèrent que les capacités arithmétiques se combinent et définissent un seul domaine, qui se distingue des autres domaines, comme celui de la perception spatiale. Cela ne veut pas dire que des différences individuelles en

intelligence générale n'influencent pas les performances en arithmétique ou en d'autres domaines, mais qu'il existe un ensemble de processus et de concepts qui expliquent les différences individuelles en arithmétique, au-delà de l'influence de l'intelligence générale.

La deuxième question qu'il convient de se poser est à quel moment du développement, le facteur d'aptitude numérique commence à émerger, ou autrement, à quel âge on peut identifier les compétences en arithmétique comme compétences distinctes. Dans une étude longitudinale de la structure factorielle de l'intelligence à l'aide de l'échelle de Wechsler pour les enfants (*Wechsler Intelligence Scale for Children*), Osborne et Lindsey (1967) ont trouvé des preuves en faveur de l'existence d'un facteur d'aptitude numérique plus ou moins distinct, pour un échantillon d'enfants de la maternelle. Le facteur était défini par les sous-tests d'arithmétique et de mémoire immédiate des chiffres du WISC, ainsi que par des items quantitatifs du sous-test d'informations (p.ex. Combien de pennies faut-il avoir pour posséder un nickel ?). Ce résultat suggère que pour des enfants de la maternelle, l'aptitude numérique comporte le comptage, l'arithmétique simple, la mémoire de travail pour les nombres et des connaissances générales sur des relations quantitatives. Dans un examen des aptitudes de base des enfants de la maternelle, Meyers et Dingman (1960) ont trouvé que les compétences numériques peuvent être identifiées comme une aptitude distincte à partir de l'âge de 5 à 7 ans. Cela ne veut pas dire que le facteur d'aptitude numérique ne se manifeste pas comme facteur distinct avant cet âge, mais plutôt que les tests psychométriques standards ne peuvent être utilisés avant l'âge de la maternelle, parce que ce sont généralement des tests papier-crayon. Le facteur d'aptitude numérique a également été identifié pour des échantillons d'enfants et d'adolescents, à l'école primaire, à l'école secondaire et dans l'enseignement supérieur (Dye & Very, 1968 ; Osborne & Lindsey, 1967 ; Thurstone & Thurstone, 1941), ainsi que pour des personnes âgées (Schaie, 1983).

Même si le facteur d'aptitude numérique a été identifié pour toutes les tranches d'âges, il semble y avoir deux changements développementaux dans les tests qui le définissent. Pour des enfants en troisième année d'école primaire, par exemple, le facteur d'aptitude numérique regroupe les tests d'arithmétiques et de mémoire numérique immédiate (Osborne & Lindsey, 1967), mais pour des adolescents plus âgés et des jeunes adultes, les tests qui définissent ce facteur sont presque exclusivement de nature arithmétique (Thurstone & Thurstone, 1941). Le fait que des

tests de la mémoire de travail pour les nombres et les tests d'arithmétique sont groupés pour des enfants de la maternelle et de l'école primaire, est probablement dû à la composante mnésique des procédures arithmétiques élémentaires, telles que le comptage (Widaman & Little, 1992). Pour des enfants de l'école élémentaire, Little et Widaman (In press) ont trouvé que la mémoire de travail pour les nombres contribue directement à la performance dans des tests d'aptitude numérique, tandis que Geary et Widaman (1992) n'ont pas trouvé une telle relation directe pour des adultes. En d'autres mots, les tests de la mémoire numérique corréleront avec les tests d'arithmétique pour les jeunes enfants parce que la mémoire numérique est importante pour les premières compétences arithmétiques. Quand, plus tard, les procédures élémentaires deviennent plus automatiques, c'est-à-dire effectués sans effort, le facteur d'aptitude numérique se réduit correspond exclusivement à la compétence arithmétique.

La seconde tendance au changement s'observe dans la relation entre la performance dans des tests d'aptitude numérique et des tests de raisonnement mathématique. En gros, pour des échantillons relativement peu compétents, comme des enfants de 12 ans, les tests d'aptitude numérique et de raisonnement mathématique sont plus ou moins corrélés : les enfants qui ont une compétence plus grande en arithmétique élémentaire ont également une compétence plus grande en raisonnement mathématique (c.f. Chein, 1939 ; Dye & Very, 1968 ; Murray, 1949 ; Thurstone & Thurstone, 1941). Pour des échantillons d'adolescents plus âgés et de jeunes adultes, où presque tous les sujets maîtrisent l'arithmétique de base, il n'y a plus aucune relation entre l'aptitude numérique et le raisonnement mathématique (Dye & Very, 1968). Pichot (1965) avait déjà souligné que les facteurs de groupe trouvés avec la méthodologie de Thurstone se différencient avec l'âge. Le facteur quantitatif trouvé chez les jeunes enfants se dédouble en facteur numérique et en facteur de raisonnement. Selon cet auteur, la structure intellectuelle se différencie avec l'âge par une sorte de perte de la plasticité et une augmentation de la spécialisation.

3.2.2 Le raisonnement mathématique

En plus du facteur d'aptitude numérique, un facteur de raisonnement mathématique a été identifié dans toutes les études effectuées grâce à l'analyse

factorielle (Canisia, 1962 ; Dye & Very, 1968 ; Ekstrom, French & Harman, 1976 ; French, 1951 ; Goodman, 1943 ; Thurstone, 1938 ; Vandenberg, 1959 ; Very, 1967). Dans les différentes études, ce facteur a obtenu beaucoup de désignations différentes, comme par exemple raisonnement arithmétique, raisonnement général et raisonnement mathématique. Pour des raisons de cohérence, nous allons toujours l'appeler facteur de raisonnement mathématique. Les tests qui définissent ce facteur exigent typiquement la capacité de trouver et d'évaluer des relations quantitatives et de tirer des conclusions sur la base de ces informations. Ainsi, l'aptitude de raisonnement mathématique décrite par les études psychométriques semble concorder avec ce qu'on appelle d'habitude la compétence de résoudre des problèmes mathématiques. Les études psychométriques peuvent nous aider à mieux comprendre la compétence de résoudre les problèmes mathématiques de plusieurs façons. Comme c'est le cas pour l'aptitude numérique, les études psychométriques peuvent être utilisées pour déterminer à quel niveau le raisonnement mathématique émerge comme facteur distinct. Puis, l'approche psychométrique peut être utilisée pour évaluer la relation entre l'aptitude de raisonnement mathématique et d'autres aptitudes numériques et de raisonnement.

Thurstone et Thurstone (1941) ont trouvé dans une étude sur 1154 enfants de 12 ans environ que les tests de raisonnement mathématique corrélaient avec les tests d'aptitude numérique classiques. Ce résultat, ainsi que d'autres résultats d'études semblables (Guthrie, 1963) suggèrent que les facteurs d'aptitude numérique et de raisonnement mathématique ne représentent pas nécessairement des aptitudes différentes pour la plupart des enfants à l'école primaire. Dans une autre étude à large échelle, Dye et Very (1968) ont administré une batterie de tests psychométriques à 358 élèves d'école secondaire et à 198 étudiants. Ces tests comportaient entre autres des mesures d'aptitude numérique, de vitesse de perception, de raisonnement verbal, de raisonnement déductif, de raisonnement mathématique et d'aptitude d'estimation. Ils ont trouvé un facteur de raisonnement mathématique, à la fois pour les élèves masculins et pour les élèves féminins à tous les âges. Alors que, pour les étudiants, ce facteur était indépendant du facteur d'aptitude numérique, il y avait une corrélation modérée pour les plus jeunes élèves d'école secondaire, probablement parce que un grand nombre d'items de raisonnement mathématiques présupposaient la capacité de résolution des problèmes d'arithmétique appliquée. Comme de bonnes compétences en arithmétique facilitent la résolution de tels problèmes, il est naturel que des tests

d'aptitude numérique et de raisonnement mathématique aient une faible corrélation à un âge où tous les sujets ne maîtrisent pas encore complètement l'arithmétique élémentaire (Geary & Widaman, 1992). Toutes les études suggèrent que les systèmes cognitifs sous-jacents à l'aptitude numérique respectivement au raisonnement mathématique commencent à diverger au cours de la scolarité.

Ceci laisse ouverte la question de savoir si l'aptitude de raisonnement mathématique est distincte d'autres formes de raisonnement (p.ex. le raisonnement déductif et inductif). Dye et Very (1968) n'ont pas trouvé de relation significative entre la performance dans des tests de raisonnement verbal et déductif et celle dans des tests de raisonnement mathématique. Ce résultat suggère, que pour des adolescents et de jeunes adultes, l'aptitude de raisonner dans un domaine quantitatif, est distincte de l'aptitude de raisonner dans des domaines non quantitatifs. Néanmoins, d'autres études ont trouvé que des tests de résolution de problèmes arithmétiques et algébriques corrèlent parfois avec des tests de raisonnement verbal et d'autres formes de raisonnement non quantitatif (Ekstrom, French & Harman, 1979 ; French, 1951 ; Vandenberg, 1959). L'ensemble de ces résultats suggère que le raisonnement mathématique se développe initialement à partir d'une compétence représentée par un facteur d'aptitude numérique et à partir de compétences de raisonnement plus générales, ce qui nous renvoie aux discussions sur l'existence et la nature d'un facteur d'intelligence générale (g), susceptible d'être remplacé avec l'âge par des facteurs de groupe de plus en plus différenciés (Pichot, 1965). Un système distinct de compétence en raisonnement mathématique émerge seulement chez les adolescents à la fin de l'école secondaire, ayant beaucoup d'expériences en mathématiques (Very, 1967).

3.2.3 Autres compétences mathématiques

Les facteurs d'aptitude numérique et de raisonnement mathématique représentent le seul système de capacités mathématiques, se retrouvant dans la plupart des études basées sur l'analyse factorielle. D'autres facteurs mathématiques ont cependant été identifiés dans certaines études. Il n'est pas clair pourquoi ils n'ont pas été trouvés aussi régulièrement que l'aptitude numérique et le raisonnement mathématique. Une raison semble consister dans le type de tests utilisé. Thurstone et

Thurstone (1941), par exemple, ont trouvé un facteur de comptage de points, qui était défini par des items évaluant la compétence de compter rapidement et justement des configurations de points. De tels items de comptage de points n'ont pas généralement été inclus dans les autres études psychométriques. Certaines études ont aussi mis en évidence des facteurs de flexibilité numérique ou bien d'estimation (Canisia, 1962 ; Very, 1967).

Le facteur de comptage de points identifié par Thurstone et Thurstone (1941) pour un échantillon d'élèves de 12 ans est intéressant, parce que les tests qui définissent ce facteur ressemblent beaucoup aux tests qui sont utilisés de nos jours dans les expériences de comptage de points (*subitizing*). Dans l'étude originelle, la performance numérique était indépendante de la performance au comptage de points. Ce résultat suggère, que pour des jeunes élèves de l'enseignement secondaire, les capacités en arithmétique ne sont pas liées aux capacités associées avec le comptage de points, qui sont le comptage et l'estimation (Mandler & Shebo, 1982). Les enfants comptent souvent pour résoudre des problèmes arithmétiques, d'où il résulte que les tests de comptage et d'arithmétique corrélaient pour des enfants de la maternelle (Osborne & Lindsey, 1967), alors que le comptage n'était certainement plus beaucoup utilisé par les élèves de l'enseignement secondaire confrontés aux problèmes arithmétiques proposés dans l'étude de Thurstone et Thurstone (Ilg & Ames, 1951). Le fait qu'il n'y pas de corrélation entre le comptage de points et l'aptitude numérique suggère que les compétences de comptage de points pourraient être distinctes des aptitudes arithmétiques, contrairement à ce qui ressort des recherches du début des années 90, affirmant que le comptage de points et l'arithmétique émergent du même système de comptage préverbal inné (Gallistel & Gelman, 1992). Ceci ne veut pas dire que le comptage de points ne soit pas utilisé dans le développement initial des compétences arithmétiques. Ce résultat suggère plutôt que les systèmes cognitifs sous-jacents au comptage de points et à l'arithmétique pourraient se différencier plus tard, par exemple au début de l'adolescence.

Le facteur de flexibilité numérique identifié par Very (1967) et Canisia (1962) implique la capacité de manipuler, d'arranger et de comparer des nombres, sans faire d'opérations arithmétiques. Des tests associés à ce facteur comportent par exemple des items de comparaison du style « plus grand que – moins grand que » d'ensembles de nombres, des items demandant d'écrire le plus grand nombre possible de nombres vérifiant certaines conditions (p.ex. être pair ou impair) ou des items de complétion de

séries de nombres (p.ex 2,4,6, ...). Canisia a trouvé que le facteur de flexibilité numérique était indépendant du facteur d'aptitude numérique et du facteur de raisonnement mathématique pour un groupe d'élèves de lycée.

Very (1967) a identifié un facteur d'estimation mesurant la capacité de faire des estimations quantitatives. Thurstone (1938) a trouvé qu'un test d'estimation quantitative ne corrélait pas avec les tests d'aptitudes numériques traditionnels, mais qu'il corrélait avec des tests de raisonnement mathématique. Very (1967) argumentait que le facteur d'estimation contient en effet la compétence de créer et d'évaluer rapidement de nouvelles hypothèses afin de tirer les bonnes conclusions, ce qui implique que l'estimation nécessite des capacités de raisonnement. La vitesse avec laquelle on effectue des comparaisons quantitatives est en elle-même distincte des aptitudes d'arithmétique et de raisonnement mathématique. Cependant, dans les performances aux tests, la capacité de faire de bonnes estimations quantitatives pourrait cependant ne pas se distinguer de l'aptitude de raisonnement mathématique.

En conclusion, certaines études psychométriques suggèrent qu'il existe des domaines mathématiques mal explorés, impliquant des compétences moins bien mis en évidence que l'aptitude numérique et le raisonnement mathématique. Ces domaines paraissent inclure des aptitudes associées avec le comptage de points, la flexibilité numérique, l'aptitude d'estimation quantitative. De ces trois domaines mathématiques potentiels, les preuves en faveur des compétences associées au comptage de points et à la flexibilité numérique sont plus fortes que les preuves en faveur d'un facteur d'estimation quantitatif. (Thurstone & Thurstone, 1941 ; Very, 1967). Le comptage de points et la numérique semblent être distinct de l'aptitude numérique et du raisonnement mathématique, au moins pour des individus compétents. Les processus associés au facteur d'estimation par contre pourraient simplement refléter la compétence de raisonnement mathématique (Petitto, 1990).

3.2.4 L'étude de J.B. Carroll

Carroll (1993) a repris les données de quelque 480 études apparues dans la littérature psychométrique et il les a analysées à l'aide d'analyse factorielles exploratives afin de dégager une taxonomie des capacités cognitives. Ses résultats ont confirmé les modèles de structure hiérarchique des capacités cognitives qui ont été proposés déjà

plus tôt par des chercheurs comme Vernon (1961), Cattell (1971) et Cattell et Horn (1978). Dans sa synthèse, Carroll suggère une théorie à trois strates dans laquelle les compétences cognitives peuvent être classifiées hiérarchiquement selon le niveau de généralité en facteurs généraux, larges et étroits. Comme le montre le diagramme de la page suivante, au sommet de l'hierarchie, le facteur d'intelligence générale, qu'on retrouve dans plus ou moins tous les tests de compétence cognitive, compose la strate la plus élevée. Une deuxième strate se compose de facteurs « larges » relativement peu nombreux qu'il a appelés intelligence fluide, intelligence cristallisée, mémoire et savoir, perception visuelle, perception auditive, capacité de remémoration, vitesse de pensée, respectivement vitesse d'exécution. Sur la troisième strate, au niveau le plus bas, on trouve un nombre relativement grand (environ 65) de facteurs « étroits » qui représentent des capacités spécifiques dans des domaines variés.

Les études factorielles ne disent en général pas grand chose sur les sources des facteurs (par exemple, à quel degré ils sont influencés par les caractéristiques génétiques ou par apprentissage), sur leur stabilité et leur degré de constance au cours de la vie. On peut cependant affirmer que la plupart des facteurs sont présents à presque tous les stades de la vie, même si les niveaux d'aptitude qu'ils déterminent peuvent augmenter ou diminuer avec le temps.

Passons en revue ceux parmi ces facteurs qui peuvent être pertinents pour le raisonnement mathématique ou pour résoudre des problèmes de nature mathématique.

a) Le facteur d'intelligence générale g

Dans les analyses factorielles exploratoires, le facteur d'intelligence générale g provient de la découverte que les facteurs de premier ordre sont substantiellement corrélés, de manière à ce qu'un seul facteur parvienne à expliquer le mieux leurs intercorrélations. Dans des analyses factorielles confirmatoires, on peut postuler un facteur général, indépendant des autres facteurs et confirmé par l'analyse (Gustafsson, 1988). Par l'un ou l'autre type d'analyse, un facteur général est trouvé très fréquemment dans les études cognitives.

Il y a eu beaucoup de spéculations et de publications sur la nature de g . Spearman a pensé que g est impliqué dans les opérations cognitives dès que l'individu doit (a) appréhender une expérience et y penser, (b) déduire ou trouver des relations

parmi des stimuli ou (c) trouver des correspondances. De nos jours, le facteur général est souvent interprété comme représentant la complexité maximale ou la difficulté générale des tâches qu'un individu possédant un niveau donné de g peut accomplir et par conséquent la quantité de manipulation mentale consciente requise par ces tâches (Jensen, 1980 ; Marshalek, Lohman & Snow, 1983). Ceci n'explique cependant pas tous les résultats. Des personnes exceptionnelles, par exemple, capable d'effectuer des tâches arithmétiques très complexes, comme de trouver la racines 23-ème d'une nombre à 201 chiffres, n'ont pas toujours des niveaux particulièrement élevés de compétence générale (e.g. Jensen, 1990).

Il est cependant difficile d'aboutir à des conclusions sur la nature de g en étudiant des cas isolés. La question peut être mieux résolu en analysant des tests hautement chargés en g . La plupart de ces tests nécessitent des raisonnements détaillés et complexes sur des similitudes, des comparaisons, le sens de mots et de phrases difficiles, des relations et des implications logiques, des problèmes quantitatifs et ainsi et suite. Certains tests nécessitent aussi de larges connaissances de principes ou de faits, ainsi que l'aptitude d'appliquer ces principes ou faits à une variété de problèmes, sans tenir compte de leur complexité. Ceci pourrait signifier que g représente l'aptitude générale d'assimiler et d'appliquer les connaissances. Des bons résultats dans un test mesurant g pourraient aussi présupposer une mémoire de travail étendue et capable (Carpenter, Just et Shell, 1990 ; Kyllonen et Christal, 1989) et des capacités de trouver la stratégie adéquate à la résolution des problèmes.

L'importance pratique du facteur g pour les mathématiques est bien illustré par l'étude de Stanley (1974) qui montre que des élèves précoces en mathématiques sont également caractérisés par un haut niveau en facteur g . D'autre part, elle montre que des élèves qui ont un QI bas (le QI est en général un bon indicateur de g) ont des difficultés considérables pour apprendre l'arithmétique élémentaire ; et même s'ils sont capables d'apprendre des opérations mathématiques élémentaires, ils ont besoin de beaucoup plus de temps pour le faire.

Tous ces résultats doivent cependant être complétés par le fait que le facteur g , comme la plupart des autres compétences cognitives est soumis au développement. Le niveau de l'aptitude correspondante tend à augmenter continuellement de l'enfance jusqu'au début de l'âge adulte, mais la vitesse de croissance dépend des individus.

b) L'intelligence fluide

Un certain nombre de facteurs de la strate inférieure, notamment l'induction, le raisonnement séquentiel et le raisonnement quantitatif sont corrélés fortement, mais, d'après le modèle décrit par Carroll, ce n'est pas seulement par l'influence du facteur g , mais également par celle du facteur de deuxième niveau appelé intelligence fluide par Cattell (1971). L'intelligence fluide représente l'aptitude générale de la pensée se manifestant dans l'induction (trouver des règles ou des généralisations qui expliquent ou gouvernent des configurations de stimuli données), du raisonnement séquentiel déductif (effectuer des raisonnements logiques qui nécessitent plusieurs pas) et des raisonnements quantitatifs (faire un raisonnement qui implique des concepts quantitatifs). L'intelligence fluide est hautement corrélée avec g et il y a même certains auteurs qui affirment que ces deux facteurs sont identiques (Gustafsson 1988 ; Undheim et Gustafsson, 1987). Carroll est néanmoins d'avis que ce sont deux facteurs différents, parce que la corrélation entre eux n'est pas toujours parfaite et qu'on peut donc les estimer indépendamment. Ainsi, l'intelligence fluide représente une compétence qui est spécifiquement en relation avec des raisonnements associés à des concepts logiques et quantitatifs.

Il est intéressant que ces raisonnements peuvent être classifiés en sous-groupes impliquant des processus inductifs, déductifs et des processus quantitatifs. De plus, on peut identifier un quatrième facteur de niveau inférieur appelé raisonnement piagétien, qui permet d'accomplir des tâches impliquant des phénomènes périodiques ou de conservation, comme celles étudiés par Piaget (Flavell 1977).

Les facteurs qui constituent l'intelligence fluide semblent caractériser les aspects majeurs de la pensée mathématique. En effet, comme on l'a vu dans les sections précédentes, beaucoup de chercheurs qui ont étudié la pensée mathématique, affirment que résoudre des problèmes mathématiques implique fréquemment les processus séparés d'induction, de déduction et de conceptualisation mathématique (Nesher & Kilpatrick, 1990). Hadamard (1954) a même proposé de classer les chercheurs en mathématiques en deux groupes, selon qu'ils sont plutôt inclinés à utiliser l'induction et l'intuition ou plutôt la déduction et le raisonnement.

Un autre facteur faisant partie de l'intelligence fluide est le facteur de vitesse de raisonnement qui représente la vitesse de résoudre des problèmes, par contraste avec le niveau d'aptitude, permettant de les résoudre en un temps indéfini. Il existe

bien sûr une corrélation entre le niveau d'aptitude et la vitesse de résolution. En effet, des individus qui ont une compétence plus grande à résoudre des problèmes mathématiques ont tendance à les résoudre plus rapidement que ceux qui ont une faible compétence dans ce domaine, mais la corrélation est loin d'être parfaite. Cela implique que des tests de raisonnement dans lesquels est imposé une limite de temps pour les résoudre évaluent souvent mal l'aptitude de résoudre des problèmes comme telle.

c) L'intelligence cristallisée

Un second facteur de deuxième strate s'appelle intelligence cristallisée (Cattell, 1971). Elle influence un certain nombre de facteurs de niveau inférieur qui concernent surtout le langage, aussi bien oral qu'écrit et aussi bien réceptif (entendre et lire) que productif (parler et écrire). En fait, ce facteur pourrait très bien s'appeler compétence linguistique, s'il ne contenait pas aussi des compétences qui sont acquises par l'expérience ou apprises à l'école. C'est pourquoi l'intelligence cristallisée est souvent caractérisée comme étant l'aptitude qui représente les aspects de l'intelligence générale acquis par expérience, apprentissage et instruction.

On peut mentionner un certain nombre de compétences linguistiques qui ont un rapport avec la pensée mathématique. Le développement du langage caractérise le niveau de compétence général d'un individu en ce qui concerne les langues, particulièrement dans les premières années de sa vie ; il mesure jusqu'à quel point, à un temps donné, la personne a assimilé la phonologie, le vocabulaire et la grammaire de sa langue maternelle. Il est bien connu, qu'en effectuant des raisonnements mathématiques, les gens utilisent généralement leur langue maternelle pour désigner les nombres (Cohen, 1994). Plus généralement, le développement linguistique individuel joue un rôle dans la résolution de problèmes mathématiques. La même remarque s'applique au facteur compréhension de l'écrit qui est d'habitude hautement corrélé avec le facteur développement du langage.

Les autres facteurs de niveau inférieur faisant partie de l'intelligence cristallisée ne sont pas spécifiques à l'apprentissage des mathématiques et nous n'allons pas les analyser en détail.

d) La mémoire

Un certain nombre de compétences relatifs à la mémoire sont suffisamment corrélés entre eux pour définir la strate de deuxième niveau appelé mémoire et savoir. Les facteurs de troisième niveau incluent l'envergure de la mémoire, la mémoire associative et la mémoire significative.

L'envergure de la mémoire est souvent testée en donnant aux sujets une série de chiffres ou lettres qu'ils doivent répéter dans l'ordre exact. Le fait que le facteur d'envergure de la mémoire est souvent substantiellement corrélé avec g peut être interprété d'après Carroll (1996) par le fait que ce facteur est une mesure de l'étendue de la mémoire de travail. Par conséquent, l'envergure de la mémoire peut être impliquée dans la résolution d'un problème mathématique, surtout si c'est un problème qu'il faut résoudre mentalement. D'autre part, Kyllonen et Christal (1989) affirment que les items utilisés pour tester l'envergure de la mémoire semblent tirés par les cheveux et ne correspondent pas à ce que les gens font quand ils apprennent vraiment quelque chose. Ils ont proposé des mesures plus réalistes de la mémoire de travail.

La mémoire associative est d'habitude testée en présentant aux sujets une série de paires de stimuli arbitraires qu'ils doivent observer pendant un temps relativement limité et en leur demandant ensuite de répondre à chaque stimulus par l'élément qui lui était associé dans la série d'apprentissage. La mémoire associative est impliquée dans l'apprentissage du vocabulaire d'une langue étrangère, mais elle influence également la vitesse avec laquelle un individu peut apprendre une série quelconque d'éléments, tels les nombres impliqués dans l'arithmétique élémentaire, les carrés des logarithmes d'une série de nombres ou même des formules et des équations mathématiques. Ce facteur influence alors la vitesse d'apprentissage et non la manière d'apprendre, ni la quantité maximale qui peut être assimilée. Les différences individuelles dans la vitesse d'apprentissage sont souvent très grandes tant chez les enfants que chez les adultes.

Un autre facteur qui pourrait être impliqué dans l'apprentissage des mathématiques est la mémoire significative, qu'on observe dans des tests où il faut mémoriser des idées qui ont un sens logique et non pas seulement des stimuli arbitraires. Le succès dans de tels tests semble dépendre non seulement de la faculté des individus de comprendre les idées, mais aussi de la vitesse avec laquelle ces idées peuvent être mémorisées.

e) La perception visuelle

D'après Carroll (1996), il est très probable que le facteur de second niveau appelé perception visuelle est impliqué dans beaucoup de types de raisonnements mathématiques. On pourrait aussi l'appeler compétence spatiale ; La compétence spatiale a souvent été citée (e.g. Werdelin, 1961) comme intervenant dans la pensée mathématique. En fait, la perception visuelle représente un élément commun dans une variété de facteurs d'ordre inférieur, comme la visualisation, les relations spatiales, la vitesse de perception, la vitesse et la flexibilité de la reconnaissance. Nous ne savons pas encore lesquels parmi ces facteurs sont vraiment importants pour les raisonnements mathématiques, même si certains travaux ont été entrepris dans ce sens (Gustafsson et Balke, 1993 ; Lubinski et Humphreys, 1990, Humphreys, Lubinski et Yao, 1993).

Les tests de visualisation présupposent la représentation d'une forme qu'il faut manipuler pour arriver à résoudre un problème dans l'espace visuel. Ils mesurent donc le niveau de difficulté auquel le sujet peut faire face et non nécessairement la vitesse avec laquelle il maîtrise ces problèmes. Les tests de relations spatiales par contre présentent des problèmes simples, comme identifier des lettres après différentes rotations, et ils mesurent la vitesse avec laquelle les sujets les résolvent. Dans les tests mesurant la vitesse de reconnaissance, on présente aux sujets des images qui sont partiellement cachées et on mesure le temps dont ils ont besoin pour reconnaître ce que ces images représentent. Dans les tests de flexibilité de reconnaissance, on présente au sujet une forme ou un dessin géométrique au sujet et on lui demande s'il est contenu dans un autre dessin qu'on lui montre. Les tests de vitesse de perception exigent la remémoration d'une configuration spatiale donnée et la capacité de la retrouver dans une série d'autres configurations spatiales.

Aucune de ces tâches n'est vraiment identique aux opérations perceptives requises en mathématiques, comme par exemple en géométrie ou en analyse fonctionnelle, mais il semble évident que des individus qui ont des problèmes avec des tests de perception visuelle ou d'aptitude spatiale éprouvent également des difficultés avec des opérations similaires dans des raisonnements mathématiques.

f) Les autres facteurs de la seconde strate

Les études d'analyses factorielles ont identifié plusieurs autres facteurs de la seconde strate, comme la perception auditive, la capacité de remémoration, la vitesse de pensée et la vitesse d'exécution. Ces facteurs ne semblent cependant ne pas être tellement importants pour les raisonnements mathématiques à part l'aptitude numérique de laquelle on a déjà parlé dans le paragraphe 3.2.1. De plus, les deux facteurs de vitesse pourraient corrélérer avec la vitesse d'exécution des problèmes mathématiques.

3.2.5 Conclusion

Il n'est en fait pas évident de donner une théorie des aptitudes cognitives qui jouent un rôle dans les raisonnements mathématiques. Malgré sept décades de travaux psychométriques, les relations entre les compétences dégagées à l'aide d'analyses factorielles et les performances réelles sont restées incertaines dans de nombreux domaines de la vie (Carroll, 1996). Il n'y a pas de doute que des relations existent, mais la nature exacte de ces relations n'est pas aussi bien connue qu'on pourrait le désirer. On sait également que l'aptitude mathématique n'est pas unitaire, c'est-à-dire qu'il existe plusieurs types de compétences sousjacentes aux performances mathématiques, mais il est difficile de déterminer quelles compétences sont les plus importantes pour des performances particulières.

Un certain nombre de chercheurs sérieux (e.g., Ceci, 1990; Winch, 1990) ont même récemment mis en question quelques unes des hypothèses de base de la tradition psychométrique, par exemple, l'hypothèse que les aptitudes cognitives identifiées par l'analyse factorielle représentent vraiment des caractéristiques importantes et durables des individus et également l'hypothèse que les scores obtenus dans des tests psychologiques peuvent être des mesures fiables et valides de telles caractéristiques.

Un argument pour de telles mises en questions est que les tests psychologiques sont typiquement décontextualisés, c'est-à-dire qu'on y présente des tâches qui ne sont pas enracinés dans des contextes ressemblant à ceux que les sujets rencontrent dans la « vraie vie ». Il y a de nombreux exemples de personnes qui sont capables

d'effectuer des raisonnements complexes et compliqués et n'obtiennent que des scores médiocres dans des tests qui sont censés mesurer justement la compétence de faire de tels raisonnements (e.g. Ceci and Liker, 1986)

Un autre argument est que les corrélations généralement positives de tests psychologiques sont surtout le résultat du fait les sujets diffèrent quant à la quantité, la diversité et la complexité de leurs expériences et apprentissages, ne serait-ce qu'à cause de différences de leurs environnements sociaux. Ceci indique que le facteur d'intelligence générale g , obtenu par des corrélations positives dans des tests, pourrait être surtout un objet mathématique fabriqué artificiellement. Il est soutenu de plus que les facteurs de niveau inférieur trouvés dans les analyses factorielles reflètent simplement des classes d'expériences qui tendent à être apprises ensemble, comme le sens des mots ou les nombres ou les formes spatiales. Il est en outre signalé que la spécialisation des compétences résulte souvent d'une longue pratique dans le domaine en question (Ericsson & Charness, 1994), et pas nécessairement d'une supériorité de quelque aptitude cognitive sousjacente.

Même si certains de ces arguments pourraient très bien être fondés, la plupart des chercheurs d'aujourd'hui pensent qu'ils vont trop loin. Il y a des preuves abondantes pour l'existence de compétences mentales mesurables par des tests psychologiques et dont on connaît des relations substantielles avec l'apprentissage et l'exécution de tâches mathématiques. Même si les tâches qui apparaissent dans les tests psychologiques sont décontextualisés, il existe des similarités fortes entre ces tâches et celles rencontrées dans la vie réelle. Et même si l'apprentissage dépend beaucoup des conditions et des stimulations de l'environnement social, on peut prouver que le degré dans lequel les gens peuvent tirer et tirent effectivement avantage de ces occasions dépend au moins en partie de leur constitution biologique.

De plus, aucun de ces arguments n'a un rapport immédiat avec la question si les résultats psychométriques qu'on a présenté dans ce paragraphe sont importants pour l'analyse des raisonnements mathématiques, parce que les mesures de facteurs effectués dans les études qu'on a mentionnées concernent uniquement la performance d'individus à un instant donné et n'impliquent rien sur les sources des différences individuelles ou les développements ultérieurs de leurs capacités. Ce qu'il faudrait mettre en question, par contre, est toute prédiction qui affirmerait que des sujets qui présentent de bas niveaux d'aptitude dans de tels processus sont incapables de profiter d'un entraînement spécifique. La question de la possibilité d'entraînement des

processus cognitifs impliqués dans la résolution de problèmes mathématiques et de l'existence de limites héréditaires aux raisonnements mathématiques n'est pas encore tranchée (Geary, 1994).

3.3 L'approche cognitive

Au-delà des différences mesurables, cette approche cherche à identifier les processus cognitifs sous-jacents à la variabilité individuelle. La recherche, aussi bien que la conceptualisation dans l'éducation des mathématiques a bénéficié des vagues successives de la révolution cognitive en psychologie, qu'elle a influencé à son tour (Gardner, 1993a).

Au cours des années 60 et 70, la richesse et la complexité des tâches, en partie mathématiques, utilisées dans la recherche sur la résolution des problèmes augmentait prodigieusement (Newell & Simon, 1972) et des systèmes théoriques rendant compte des structures complexes du savoir étaient conçus. Une conséquence de grande portée était l'interaction entre la psychologie de l'éducation et la pratique de l'enseignement dans les écoles qui s'instaurait à cette époque (Greeno, 1980).

Les ordinateurs ont joué un rôle clé dans ces changements. La simulation sur ordinateur devenait la méthodologie standard pour une grande partie des chercheurs en théorie de traitement de l'information (Rabinowitz, 1988). Le mérite primordial de cette approche est le fait de devoir détailler une théorie suffisamment pour pouvoir l'incorporer dans un logiciel qui tourne sur ordinateur, ce qui renforce la rigueur dans la construction de cette théorie et dans l'élaboration des hypothèses nécessaires pour la vérifier empiriquement. L'influence des ordinateurs s'étendait cependant bien au-delà de leur utilisation spécifique dans la simulation, grâce à la conceptualisation de la pensée comme système de traitement de l'information (Davis, 1984). L'interaction conséquente entre les études empiriques et le développement de modèles sur ordinateur était particulièrement fructueuse, entraînant des avantages dans les deux sens ; c'est pourquoi, ce domaine de recherche montre de manière exemplaire les forces et les faiblesses de l'approche de la simulation sur ordinateur (cf. De Corte & Verschaffel, 1988).

La simulation décrite par Riley, Greeno et Heller (1983) fonctionne de la manière suivante : le chercheur qui veut analyser les composantes d'un problème construit d'abord un schéma, sous forme de réseau de quantités et de relations sémantiques internes, en termes de changement, de combinaison ou de comparaison. Basés sur l'information recueillie à partir de ce modèle, les processus de résolution opèrent alors sur des ensembles de blocks sémantiques, correspondant à l'ensemble des objets impliqués dans cette situation. Un éventail de modèles de calcul, correspondant aux différents niveaux de compétence dans la résolution de problèmes de changement, de combinaison et de comparaison, engendre des performances comparables à celles observées empiriquement.

Du point de vue de l'éducation en mathématiques, ces efforts constituent une des démonstrations les plus impressionnantes du potentiel de la simulation sur ordinateur. Comme l'indique Greeno (1987), les simulations ont le mérite de fournir des descriptions explicites de connaissances et de processus qui sont d'habitude implicites et souvent vagues. De plus, elles sont potentiellement utiles pour élaborer du matériel et préparer des interventions pédagogiques.

Il faut cependant souligner que les résultats empiriques n'ont pas tous correspondu aux modèles (Carpenter & Moser, 1984 ; De Corte & Verschaffel, 1988). De plus, l'application de ces études est limitée et elles laissent de côté des aspects psychologiques importants, telle que l'influence de variables textuelles sur la construction de la représentation initiale du problème. On a trouvé des divergences significatives quand on a comparé la performance des programmes d'ordinateur avec les résultats empiriques, à un niveau supérieur aux simples données de performance (difficulté du problème, temps nécessaire à la résolution, erreurs typiques), pour analyser les représentations sous-jacentes, les manipulations cognitives et les stratégies de résolution (De Corte & Verschaffel, 1998).

A partir de 1990, il y a eu un regain d'intérêt pour les facteurs non-intellectuels inhérents aux processus cognitifs ; en particulier, en ce qui concerne la pensée mathématique, il faut faire place à l'intuition, les croyances, les attitudes, les émotions et la pensée imagée. De Corte, Greer et Verschaffel (1996) caractérisent ce courant de seconde vague cognitive, par opposition à la première vague qui s'était exclusivement intéressée au côté logique et rationnel de la cognition.

Nous avons vu que, depuis les années 70, les psychologues cognitifs s'intéressent à l'identification des processus cognitifs sous-jacents à la performance observable dans les tests psychométriques traditionnels (voir Hunt, Lunneborg & Lewis, 1975). On montrait, par exemple, qu'une bonne performance dans des tests d'aptitude verbale est associée, entre autre, à la vitesse avec laquelle on peut récupérer les informations verbales de la mémoire à long terme. La tentative de combiner les approches psychométriques et cognitives pour mieux comprendre la variabilité individuelle dans les aptitudes intellectuelles a donné des résultats peu probants (Keating, List & Merriman, 1985). Néanmoins, l'essai de relier les études cognitives des procédures d'arithmétique (Ashcraft, 1992) aux concepts d'aptitude numérique et de raisonnement mathématique, définis de manière psychométrique, a été relativement couronnée de succès (Geary & Widaman, 1987 ; Siegler, 1988). Ces recherches ont fourni un aperçu sur les processus cognitifs et les stratégies intellectuelles sous-jacents aux différences individuelles dans les performances aux tests d'aptitude numérique et de raisonnement mathématique. Les résultats et les implications de ces études seront résumés ci-dessous.

3.3.1 L'aptitude numérique

Il y a eu plusieurs études qui ont directement examiné la relation entre les processus cognitifs associés à la résolution de problèmes arithmétiques, comme, par exemple, récupérer des concepts et procédures arithmétiques dans la mémoire à long terme, et la performance dans des tests d'aptitude numérique (Geary & Widaman, 1987, 1992 ; Widaman & alt, 1992). Dans la première de ces études, Geary et Widaman ont administré une batterie de tests papier-crayon à 1000 étudiants ; ils ont demandé aux étudiants de résoudre en outre des problèmes arithmétiques administrés par ordinateur. La batterie de tests papier-crayon incluait des tests d'aptitude numérique, de vitesse de perception (i.e la vitesse de codage ou de lecture des symboles) et d'aptitude spatiale. Le traitement statistique des données des problèmes administrés par ordinateur a permis de décomposer la résolution des problèmes en composantes procédurales élémentaires, telles que le codage des nombres (i.e. lire les nombres qui apparaissent sur l'écran), la récupération des données arithmétiques de la mémoire à long terme, le passage au sous-ensemble de symboles suivant. De plus, ces

techniques ont fourni de l'information sur la vitesse d'exécution de tous ces processus. Par exemple, en résolvant des problèmes d'addition complexes, tels que $47+79$, un sujet avait besoin de 0,9s pour passer d'une colonne à la suivante, tandis qu'un deuxième sujet avait besoin de 0,5s seulement, pour effectuer le même processus.

Ces résultats ont montré une très forte corrélation entre la vitesse avec laquelle des processus arithmétiques de base peuvent être accomplis et la performance dans les tests traditionnels d'aptitude numérique. Plus précisément, la performance dans les tests d'aptitude numérique papier-crayon était d'autant meilleure que la vitesse de récupération de la mémoire à long terme et le passage à l'opération suivante étaient plus grands. De plus, la vitesse d'exécution de ces processus arithmétiques n'était pas directement liée à la performance dans les tests de vitesse de perception et dans les tests spatiaux. Ces résultats sont importants pour plusieurs raisons. Premièrement, en ce qui concerne le domaine mathématique, ils montrent que des chercheurs cognitifs contemporains (p.ex. Ashcraft, 1992) identifient les mêmes processus sous-jacents aux différences individuelles dans les performances arithmétiques de base que les premiers chercheurs psychométriques (p.ex. Thurstone, 1938). Deuxièmement, le concept d'aptitude numérique apparaît comme étant clairement de nature arithmétique et les processus qui contribuent aux aptitudes en arithmétique semblent être distincts des processus associés aux autres aptitudes mentales, en particulier à l'aptitude spatiale.

Plusieurs études ont examiné la relation entre les processus arithmétiques et la performance aux tests arithmétiques traditionnels pour des groupes d'enfants (Geary & Brown, 1991 ; Geary & Burlingham-Dubree, 1989 ; Siegler, 1988 ; Widaman et al., 1992). Dans une de ces études (Geary & Burlingham-Dubree, 1989), les chercheurs ont administré les *Wechsler Preschool and Primary Scale of Intelligence* (WPPSI ; Wechsler, 1967) et le sous-test d'arithmétique du *Wide Range Achievement Test* (WRAT ; Jastak & Jastak, 1978) à des enfants de la maternelle et même à des sujets plus jeunes, et ils ont enregistré sur vidéo les enfants en train de résoudre une série de problèmes d'addition simples. Les bandes vidéo ont été analysées pour déterminer le type de stratégie que les enfants ont utilisé pour résoudre les problèmes d'addition, comme par exemple compter sur les doigts, et la vitesse avec laquelle ils ont exécuté les différentes opérations. Ensuite, deux dimensions cognitives furent dégagées ; la première concernait la vitesse avec laquelle les enfants avaient accès à

leur mémoire à long pour les problèmes connus. La deuxième dimension concernait le choix des stratégies appropriées. Par exemple, si un enfant se rappelait la réponse correcte d'un problème, alors le choix judicieux consistait à donner la réponse dont il se souvenait. D'autre part, s'il ne pouvait pas se souvenir de la réponse correcte, alors une meilleure stratégie était de résoudre le problème lui-même, par exemple, en comptant sur ses doigts. Pour des enfants de cet âge, compter sur les doigts est souvent préférable à l'extraction de la réponse de la mémoire, parce que cette stratégie donne des résultats plus précis. Un mauvais choix aurait été de donner une réponse un peu au hasard. Cette étude, ainsi que d'autres recherches effectuées par Siegler (1988, 1993) ont montré que les enfants variaient beaucoup dans l'adéquation du choix. Certains enfants choisissaient presque toujours la meilleure stratégie, tandis que d'autres faisaient souvent de mauvais choix. Le choix de la stratégie appropriée implique la capacité d'utiliser la procédure la plus rapide et la plus exacte pour trouver la réponse : il faut choisir une stratégie qui fournit un bon équilibre entre le temps nécessaire à la résolution du problème et la probabilité que cette stratégie fournisse la réponse correcte.

Le but principal de l'étude de Geary et Burlingham-Dubree (1989) était de déterminer la relation entre ces deux dimensions cognitives et la performance dans les tests. Ils ont trouvé une corrélation significative entre la variable « choix d'une stratégie » et les sous-tests d'arithmétique du WPPSI et WRAT. L'aptitude de faire un bon choix de stratégie est associée à une meilleure performance dans les deux tests arithmétiques. Les choix de stratégies étaient aussi plus ou moins liés à la performance dans plusieurs des sous-tests spatiaux du WPPSI, mais ils étaient indépendants des sous-tests verbaux. Chez les jeunes enfants, la relation entre le choix d'une stratégie permettant de résoudre des problèmes d'addition et l'aptitude spatiale était probablement due à l'utilisation d'information spatiales, comme par exemple l'habitude de représenter des additions sur les doigts (cf. Carpenter & Moser, 1984). La vitesse de récupération mnésique des données arithmétiques était aussi liée à la performance dans les tests arithmétiques, mais pas à la performance dans les sous-tests spatiaux ou verbaux du WPPSI. Ce que Geary et Widaman (1987) avaient démontré pour des étudiants et Widaman et alt. (1992) pour des élèves de l'école primaire est valable également pour les enfants plus jeunes : plus les données arithmétiques de base peuvent être retirées rapidement de la mémoire, plus la performance est bonne dans les tests d'arithmétique.

Du point de vue de la psychologie cognitive, cela ne signifie pas que les différences individuelles dans l'aptitude arithmétique ne soient pas liées au savoir conceptuel. En fait, Geary, Bow-Thomas Et Yao (1992) ont trouvé dans une étude d'enfants de la première classe de l'école primaire que le fait de savoir compter implique une bonne partie des différences individuelles dans les choix de stratégies pour résoudre des problèmes arithmétiques. En particulier, l'utilisation d'une procédure de comptage efficace paraissait dépendre d'une compréhension mature des concepts de comptage. Aussi bien le savoir conceptuel que la compétence de faire de bons choix stratégiques étaient liés à la performance dans un test mathématique traditionnel. La compréhension des concepts de comptage était corrélée positivement à l'utilisation de stratégies de comptage sophistiquées pour résoudre des problèmes d'addition et contribuait largement à la variance des scores aux tests mathématiques. Dans une étude plus récente, Geary, Bow-Thomas et alt. (1993) ont trouvé que l'étendu de la mémoire de travail influence également l'utilisation des stratégies de comptage chez les enfants de la maternelle. En particulier, l'utilisation de stratégies de comptage verbal, par opposition au comptage sur les doigts, était liée à une bonne compétence de la mémoire de travail.

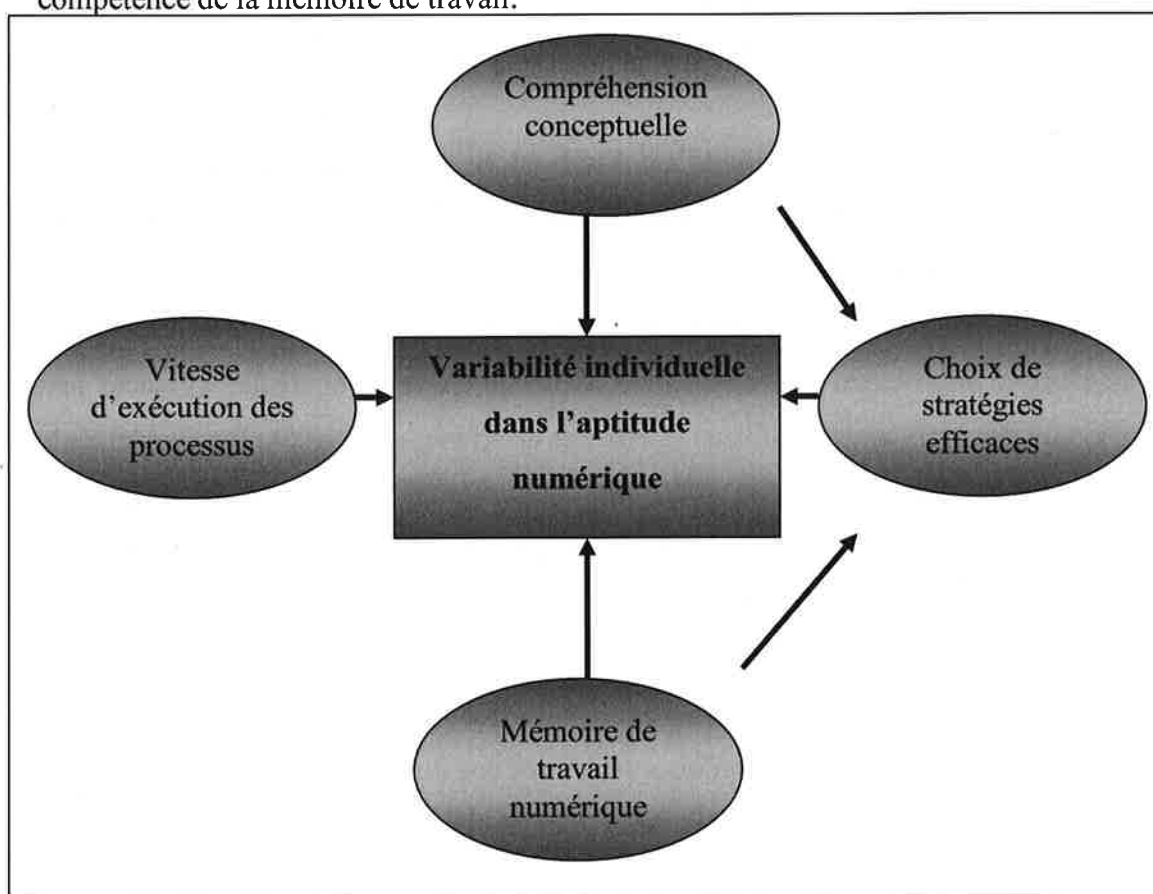


Figure 3.1 : Facteurs influençant l'aptitude numérique chez les jeunes enfants

Ces études font apparaître une continuité claire entre les études psychométriques décrites dans la section ci-dessus et les études cognitives sur les différences individuelles dans l'aptitude arithmétique de base. Les chercheurs psychométriques et expérimentaux ont identifié les mêmes sources de différences individuelles. Les deux approches suggèrent également que la source de la variabilité individuelle dans les aptitudes de base change un peu avec l'âge. Pour des enfants de la maternelle et de l'école primaire, les différences individuelles dans les aptitudes à faire de bons choix de stratégie et les facteurs qui affectent ces choix, comme le savoir conceptuel et la mémoire de travail, semblent être la source principale des différences individuelles dans l'aptitude arithmétique. La vitesse avec laquelle les opérations de base, telle que la récupération de données de la mémoire, peuvent être exécutées, influence également la variabilité individuelle, mais elle semble seulement être un facteur mineur, comparée aux différences dans la capacité stratégique et peut-être le savoir conceptuel. Au fur et à mesure que les données de base sont stockées dans la mémoire et que l'utilisation des procédures devient plus ou moins automatique, les différences de stratégie entre les individus disparaissent en grande partie. A partir de ce moment, la source primaire des différences dans les aptitudes arithmétiques est la vitesse avec laquelle on exécute les processus arithmétiques de base.

En conclusion, les enfants qui excellent en arithmétique semblent avoir une bonne compréhension des concepts numériques et arithmétiques de base et une bonne mémoire de travail numérique (Geary, Bow-Thomas & alt., 1993 ; Geary, Bow-Thomas & Yao, 1992). Les capacités conceptuelles et la mémoire de travail bien développée contribuent à la compétence dans les choix de stratégies efficaces pour résoudre des problèmes. L'utilisation efficace de stratégies de résolution de problèmes alternatives, d'autre part, semble être une source essentielle de différences individuelles dans les aptitudes arithmétiques des jeunes enfants, bien que la vitesse d'exécution des processus arithmétiques de base soit également importante (Geary & Burlingham-Dubree, 1989 ; Siegler, 1988).

Des recherches préliminaires sur l'utilisation de procédés imageants lors de la résolution de problèmes arithmétiques suggèrent que des circuits neuronaux différents sont utilisés lors de la comparaison de la grandeur de différents nombres et lors de la multiplication des nombres (Dehaene et al., 1996). L'importance des recherches en neuro-psychologie est donc cruciale dans ce domaine.

3.3.2 Le raisonnement mathématique

Même s'il y a eu plusieurs études cognitives sur les différences individuelles dans les aptitudes de raisonnement en général (Sternberg, 1977 ; Sternberg & Gardner, 1983), les études concernant le raisonnement mathématique sont relativement rares.

Geary et Widaman (1992) ont examiné les relations entre la vitesse d'exécution d'opérations arithmétiques, le maintien de ces opérations dans la mémoire de travail et la performance dans une batterie de tests psychométriques, chez 112 recrues de l'aviation militaire américaine. Les tests psychométriques évaluaient l'aptitude numérique, la vitesse de perception, le raisonnement mathématique et un facteur de mémoire. Un facteur de raisonnement mathématique était défini par deux ensembles de problèmes d'arithmétique appliquée. L'un des deux ensembles exigeait des recrues de résoudre de vrais problèmes, tandis que l'autre exigeait simplement la détermination de la suite d'opérations arithmétiques nécessaires. Le facteur de mémoire était défini par des tests de rappel de suites de nombres, présentés visuellement et verbalement pendant un court laps de temps.

Un des buts de cette étude consistait à déterminer dans quelle mesure les différences individuelles dans l'aptitude de raisonnement mathématique étaient liées à la vitesse d'exécution des procédures arithmétiques de base et à l'aptitude de faire des opérations arithmétiques mentalement dans la mémoire de travail. Les résultats répliquaient les conclusions de leur étude sur les aptitudes numériques : la vitesse d'exécution des processus de base expliquait environ 80% de la variance des aptitudes d'arithmétiques de base des jeunes adultes. Même ce facteur contribuait également à la performance aux tests de raisonnement mathématique, elle n'expliquait qu'une petite partie de la variance. La capacité d'effectuer mentalement des opérations arithmétiques, par contre, était une source importante de différences individuelles et expliquait environ 25% de la variance de l'aptitude de raisonnement mathématique chez les adultes.

Cette étude, combinée aux résultats de recherche sur l'aptitude de raisonnement général (Sternberg & Gardner, 1983), permet d'affirmer que les différences individuelles dans l'aptitude de raisonnement mathématique proviennent de quatre types de compétences cognitives. D'une part, l'aptitude de constituer mentalement des représentations de problèmes d'arithmétique appliquée ; en effet

Lewis et Mayer (1989) ont montré que cette compétence joue un rôle important chez les étudiants de premier cycle universitaire. D'autre part, la facilité avec laquelle des plans de résolution de problèmes sont développés est aussi probablement une source de différences individuelles dans les aptitudes de raisonnement mathématiques (Mayer, 1981).

La troisième source de différences individuelles dans l'aptitude de raisonnement mathématique semble être la mémoire de travail, ou l'aptitude de garder de l'information importante à l'esprit pendant qu'on effectue des opérations mathématiques. Finalement, la vitesse avec laquelle on effectue des procédures mathématiques de base est également une source de différences individuelles.

En résumé, des individus qui sont compétents en raisonnement mathématique semblent exécuter des calculs faciles rapidement et automatiquement, ils semblent être capables de garder des informations importantes à l'esprit pendant qu'ils effectuent d'autres opérations, ils développent facilement des schémas pour aider la représentation du problème et ils développent des plans de résolution adéquats.

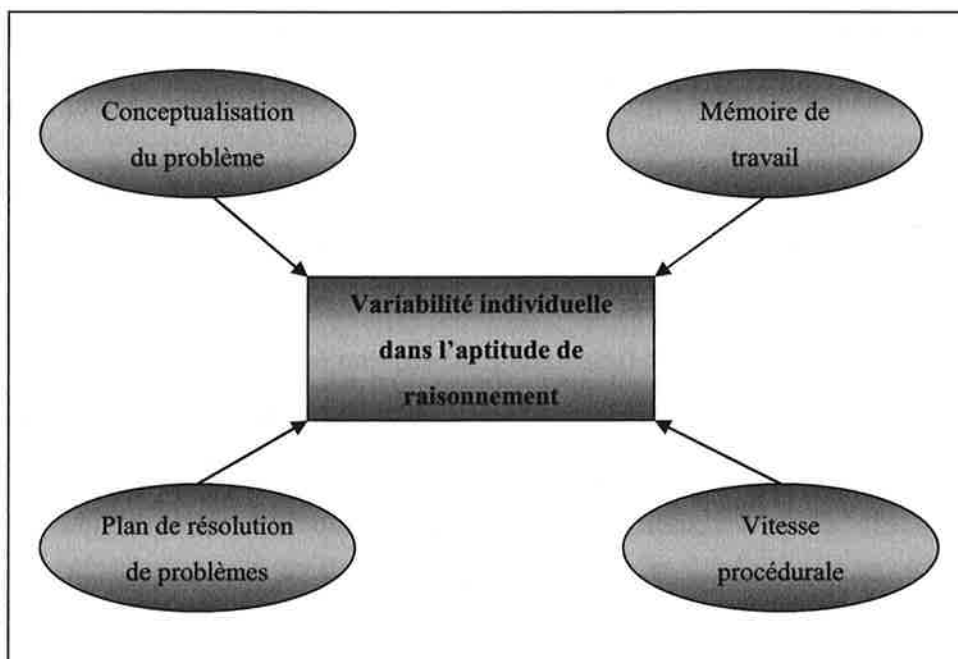


Figure 3.2 : Facteurs influençant l'aptitude de raisonnement mathématique

Chez les élèves plus âgés, ayant acquis une plus grande habileté dans la résolution des problèmes, les aptitudes en jeu sont plus complexes. D'après De Corte, Greer et Verschaffel (1996), les chercheurs ont actuellement reconnu l'importance de

4 aptitudes de base : les connaissances spécifiques, les méthodes heuristiques, les habiletés métacognitives et les dispositions affectives favorables.

Chi, Glaser et Farr (1988) ont montré que ceux qui savent résoudre les problèmes mathématiques d'une manière experte disposent d'un savoir étendu, soûplement organisé et facilement accessible.

Les méthodes heuristiques permettent de transformer le problème original de manière à ce qu'il puisse être abordé par des procédures familières, ce qui revient à poser le problème autrement (Silver, 1994).

Les processus métacognitifs entrant en jeu comprennent d'un côté la prise de conscience de son propre fonctionnement mental, d'un autre côté le contrôle de ce fonctionnement. La théorie d'autorégulation de Kuhl (1985) peut faire comprendre pourquoi les experts en raisonnement mathématique présentent un degré élevé d'auto-contrôle ; ne perdant jamais de vue leurs objectifs, ils peuvent faire les corrections nécessaires, dès qu'ils s'écartent de la bonne direction. L'activité métacognitive est loin d'être complètement éclaircie. Deux questions ne sont pas résolues : le degré de conscience ou d'inconscience de ces processus et la question de la transférabilité des activités métacognitives d'un domaine à un autre.

Selon McLeod (1990), les croyances, attitudes et émotions des élèves interfèrent avec l'apprentissage des mathématiques. Alors que les croyances et attitudes sont plutôt stables, les émotions se dissipent rapidement. Parmi les croyances, McLeod a distingué celles qui se rapportent aux mathématiques et celles qui se rapportent à soi-même et à sa propre compétence. Schoenfeld (1988) a montré que, même lorsque l'enseignement des mathématiques se fait d'après une méthode scientifiquement valable, les élèves peuvent présenter des croyances et attitudes négatives reflétant les influences du groupe environnant.

Les activités métacognitives sont spécialement susceptibles d'être altérées par les émotions, alors que la mémoire et les procédures sont moins affectées. D'autre part, la nature des réactions affectives peut changer selon le stade de résolution du problème.

Il y a une interaction complexe entre ces différents types d'aptitudes. Goldin (1992) donne une idée de cette complexité dans son modèle des différents types de représentations internes. Des chercheurs tels que Perkins, Jay et Tishman (1993) parlent d'une disposition aux mathématiques, disposition qu'on doit acquérir par une longue familiarité avec la matière et qui ne dépend pas entièrement de l'enseignement.

3.3.3 L'étude du traitement de l'information

Dans les années 70, une nouvelle approche de l'étude des processus mentaux s'est largement répandue : l'étude du traitement de l'information (*information processing*).

Ces théories décrivent les processus de la pensée à l'aide de la manipulation de symboles (Siegler, 1982) ; elles portent principalement sur le traitement et la représentation mentale de l'information et elles essaient d'atteindre un haut degré de précision dans la description de la connaissance.

La plupart des approches du traitement de l'information partagent certaines hypothèses sur l'architecture du système de traitement de l'information. Il est généralement admis qu'il existe trois degrés structuraux du système : un registre sensoriel ou d'accueil, une mémoire de travail et une mémoire à long terme.

Toute information entre dans le système par le registre sensoriel, mais elle peut y être retenue seulement pendant un court laps de temps. Pour rester dans le système, l'information doit pénétrer dans la mémoire de travail, où elle peut être combinée avec d'autres informations de la mémoire à long terme.

Toutes les opérations de traitement d'information se font dans la mémoire de travail. En d'autres termes, c'est à cet endroit que toute réflexion consciente se fait. La mémoire de travail est cependant très limitée en capacité et peut seulement traiter un petit nombre de morceaux d'information simultanément. La mémoire à long terme, par contre, est potentiellement illimitée en capacité et contient toute l'information qu'on a enregistrée au cours de sa vie. Elle est limitée par les difficultés d'accès à cette information.

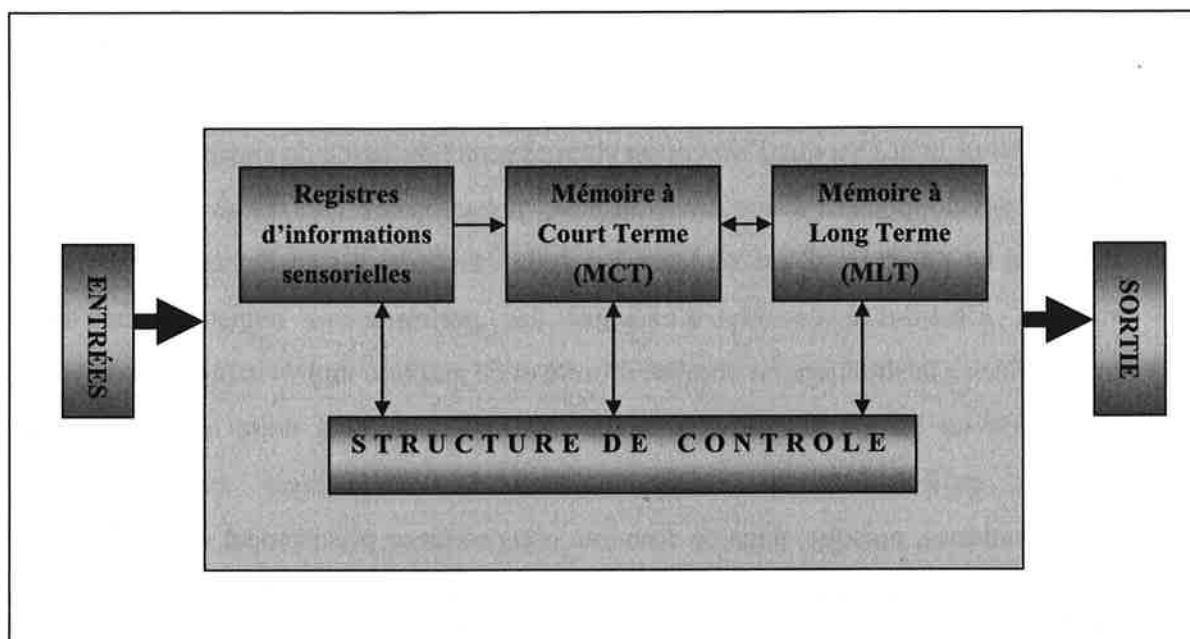


Figure 3.3 : Schéma de l'architecture cognitive proposé en 1969 par Atkinson et Shiffrin. (D'après Lemaire, 1999)

En plus de ces caractéristiques structurales de base, la plupart des théories de traitement de l'information supposent l'existence de processus exécutifs ou d'évaluation, qui contrôlent les opérations du système. Cela inclut des actes de routine comme l'utilisation de stratégies de répétition pour fixer l'information dans la mémoire à long terme ou le développement d'une heuristique générale pour la résolution de problèmes.

Les théories de l'étude du traitement de l'information peuvent être classées dans deux grandes catégories : celles qui s'occupent principalement de la nature du système de traitement de l'information et celles qui examinent comment le système de traitement de l'information fonctionne dans des situations particulières (Siegler, 1983). D'après Romberg et Carpenter (1986), la seconde permet le plus facilement de comprendre comment on apprend les mathématiques. Plutôt que de fonder leurs théories sur des simples tâches de laboratoire, comme l'ont fait les behavioristes, les chercheurs du traitement de l'information se sont souciés d'avantage des opérations cognitives complexes. Voilà pourquoi leur travail constitue une approche idéale pour l'étude de l'apprentissage des mathématiques.

3.4 La notion de compétence en mathématique

Nous avons vu que l'acception chomskyenne du terme de compétence consiste dans le développement d'une intuition juste, présupposant une vision synthétique du domaine en question. C'est seulement à cette condition qu'on devient expert en la matière, c'est-à-dire capable d'exécuter des performances respectant un code, appropriées à la situation, en nombre illimité et de manière imprévisible et libre.

En se basant sur les critères de Chomsky, on peut donc raisonnablement postuler qu'il existe un certain nombre de compétences authentiques en mathématiques, puisque, dans ce domaine nous sommes précisément confrontés à un code, à des règles restrictives laissant une grande marge de liberté permettant de produire des performances selon les modalités définies par Chomsky. Ces compétences peuvent seulement être estimées à travers les performances scolaires si l'enseignant ne pose pas simplement des problèmes de routine dans lesquels les élèves reproduisent des automatismes appris.

D'autre part, nous pouvons admettre que l'enseignement des mathématiques à l'école primaire fait appel à de nombreuses compétences fondamentales, acquises au cours des premières années de la vie, telles que la compétence conceptuelle, la compétence dans l'utilisation d'un code imagé, la compétence dans la manipulation des structures spatiales, la compétence dans la manipulation des structures temporelles, la compétence logique, la compétence combinatoire, la compétence dans la résolution des problèmes, la compétence dans le domaine de la pensée divergente etc... (Richard, 1990). Au contact de la matière enseignée, des compétences plus spécifiquement scolaires peuvent émerger. Cette question sera approfondie à la suite de l'établissement de notre taxonomie.

Chapitre IV

APERCU SUR LES TAXONOMIES EXISTANTES

Les études psychométriques ont mis en évidence qu'il existe deux domaines généraux d'aptitudes mathématiques :

- a) les nombres entiers et l'arithmétique
- b) la résolution de problèmes.

Nous avons vu que, vers l'âge de 15 ans, ces deux classes d'aptitudes sont indépendantes l'une de l'autre (Dye & Very, 1968). Cela suggère que les processus cognitifs qui contribuent à de bonnes compétences en arithmétiques sont différents de ceux qui permettent de résoudre des problèmes, du moins à partir de l'adolescence.

A côté de cette classification sommaire, il y a eu plusieurs recherches cognitives effectuées dans le but de classer les compétences en mathématiques. Depuis Tyler, on distingue généralement deux dimensions, l'une concernant le contenu mathématique (*content*) et l'autre l'activité (*activities*) c'est-à-dire les processus cognitifs en jeu.

4.1 La taxonomie de Bloom

Benjamin S. Bloom a établi une taxonomie pour le domaine cognitif en général (cf. Krathwohl, Bloom & Masia, 1956), au début des années 50. Il y distingue deux dimensions, le savoir et les compétences intellectuelles.

Pour établir la taxonomie, des chercheurs qui étaient en contact avec des enseignants, regroupaient les buts éducatifs possibles dans un vaste système, construit selon un ordre de complexité croissante. Ils le soumettaient à l'évaluation, au moyen de tests, construits selon les définitions de la taxonomie. Les buts éducatifs proviennent donc d'une discussion a priori et non pas d'une recherche psychométrique ni d'une expérimentation cognitive. Ils ont été complétés par une classification des buts éducatifs dans le domaine affectif et dans le domaine psychomoteur (cf. Krathwohl, Bloom & Masia, 1956).

LE SAVOIR

1.00 LE SAVOIR

Le savoir, (ou la connaissance), selon la définition de Bloom, est constitué de rappels de détails et de généralités, de rappels de méthodes et de processus ou de rappels de schémas, de structures ou de cadres. Pour pouvoir mesurer cette dimension, il suffit de présenter le matériel approprié à l'esprit (situation de rappel). Même si quelques modifications du matériel peuvent être exigées, cela ne constitue qu'une partie relativement peu importante de la tâche. L'étude du savoir insiste surtout sur les processus psychologiques de la mémorisation. Pour utiliser une analogie, si l'on se présente l'esprit comme un fichier, une situation de test du savoir consiste à demander au sujet de trouver dans le problème ou la tâche les signaux appropriés pour produire le plus efficacement possible l'information classée ou stockée.

1.10 Connaissance de détails

Il s'agit du rappel de détails et de morceaux d'information isolés. L'accent est mis sur des symboles ayant des référents concrets. Ce matériel, d'un niveau

d'abstraction très bas, peut fournir les éléments à partir desquels des formes plus complexes et abstraites de savoir sont construites.

1.11 Connaissance de terminologie

Il s'agit de reconnaître des référents pour des symboles spécifiques (verbaux et non verbaux), comme par exemple la connaissance du référent généralement accepté d'un symbole, la connaissance de variétés de symboles pouvant être utilisés pour un seul référent ou la connaissance du référent le plus approprié à une utilisation donnée.

Exemples :

- définir des termes techniques en donnant leurs attributs, propriétés ou relations
- être au courant d'un grand nombre de mots avec leur sens courant.

1.12 Connaissance de faits spécifiques

Il s'agit de la connaissance de dates, d'événements, de personnes, d'endroits, etc... L'information peut être très précise et spécifique ou bien approximative ou relative.

Exemples :

- se rappeler les faits les plus importants de cultures particulières
- posséder d'un savoir minimal sur les organismes étudiés en laboratoire.

1.20 Connaissance de moyens de traiter des détails

Il s'agit de la connaissance des moyens d'organiser, d'étudier, de juger et de critiquer, par exemple les méthodes d'enquête, l'ordre chronologique et les standards de jugement dans un domaine, ainsi que le modèle d'organisation par lequel les domaines eux-mêmes sont déterminés et organisés intérieurement. Ce savoir se trouve à un niveau d'abstraction intermédiaire entre le savoir spécifique et les connaissances générales. Il demande une conscience passive de la nature des matériaux plutôt qu'une expérience active d'utilisation de ces mêmes matériaux.

1.21 Connaissance de conventions

Il s'agit de la connaissance de façons caractéristiques de traiter et de présenter des idées et des phénomènes. Pour des raisons de communication et de cohérence, des travailleurs dans un domaine utilisent les usages, les styles, les pratiques et les formes les mieux adaptées à leur objectif ou aux phénomènes auxquels ils ont affaire. Il faut

remarquer que, même si ces formes et conventions sont vraisemblablement établies sur des bases arbitraires, accidentelles ou dogmatiques, elles sont retenues à cause du consentement général des individus concernés par le sujet, phénomène ou problème.

Exemples :

- être familier avec les formes et conventions de type artistique ou scientifique d'oeuvres majeurs ; comme, par exemple, la poésie, les pièces de théâtre, les articles scientifiques, etc...
- rendre des élèves conscients de la forme et de l'utilisation correctes de la parole et de l'écriture.

1.22 Connaissance de tendances et de successions

Il s'agit de la connaissance des processus, directions et développements chronologiques des phénomènes.

Exemples

- comprendre la continuité de le développement de la culture américaine telle qu'elle est illustrée par la vie de tous les jours
- connaître les tendances principales sous-jacentes au développement des programmes d'assistance publique.

1.23 Connaissance de classifications et de catégories

Il s'agit de la connaissance de classes, d'ensembles, de divisions et d'arrangements fondamentaux pour un domaine, objectif, argument ou problème donné.

Exemples :

- reconnaître les domaines de divers types de problèmes ou matériaux
- se familiariser avec un certain type de littérature.

1.24 Connaissance de critères

Il s'agit de la connaissance de critères selon lesquels des faits, principes, opinions et conduites sont testés ou jugés.

Exemples :

- être au courant des critères de jugement appropriés à un type de travail
- connaître les critères d'évaluation d'activités de récréation.

1.25 Connaissance de méthodologie

Il s'agit de la connaissance des méthodes d'enquêtes, des techniques et des procédures utilisées dans un domaine particulier, ainsi que de celles utilisées pour l'analyse des problèmes et phénomènes particuliers. L'accent est mis ici sur la connaissance de la méthode plutôt que sur la capacité de l'individu d'utiliser la méthode.

Exemples :

- connaître les méthodes scientifiques servant à évaluer des concepts de santé
- connaître les méthodes d'investigation de problèmes importants en sciences sociales.

1.30 Connaissance des généralités et des faits abstraits dans un domaine.

Il s'agit de la connaissance des schémas et modèles majeurs, d'après lesquels les phénomènes et idées sont organisés. Ce sont les grandes structures, les théories et les généralisations qui dominent un domaine ou qui sont utilisés généralement pour étudier des phénomènes ou résoudre des problèmes du domaine concerné. Cette compétence correspond au plus haut degré d'abstraction et de complexité.

1.31 Connaissance de principes et de généralisations

Il s'agit de la connaissance d'abstractions conceptuelles particulières résumant les observations de phénomènes. Ces généralisations sont utiles pour expliquer, décrire, faire des prévisions ou déterminer l'action ou la direction la plus appropriée ou la plus importante.

Exemples :

- connaître les principes majeurs qui gouvernent notre expérience avec les phénomènes biologiques
- se rappeler les généralisations les plus importantes sur des cultures particulières

1.32 Connaissance des théories et des structures

Il s'agit de la connaissance du corps de principes et de généralisations représentant un phénomène, problème ou domaine complexe d'une manière claire, arrondie et systématique. Ce sont les formulations les plus abstraites ; elles rendent compte de l'organisation d'un grand nombre de détails.

Exemples :

- se rappeler les théories majeures sur des cultures particulières
- connaître une formulation relativement complète de la théorie de l'évolution.

LES COMPÉTENCES INTELLECTUELLES

Ces compétences concernent les opérations et techniques servant à traiter des problèmes et des matériaux. Les matériaux et problèmes peuvent être de nature à n'exiger que peu ou pas du tout d'information spécialisé et technique. L'information dont on a besoin est présumée faire partie des connaissances générales de l'individu. D'autres problèmes peuvent exiger de l'information spécialisée et technique à un niveau assez élevé. La dimension des compétences intellectuelles met l'accent sur les processus mentaux d'organisation et de réorganisation nécessaires pour aboutir à un objectif particulier. Les matériaux peuvent soit être présentés soit provenir de la mémoire.

2.00 LA COMPRÉHENSION

Le niveau le plus bas de la compréhension est la compréhension implicite ; l'individu sait ce qui lui est communiqué et peut utiliser ce matériel ou cette idée sans la lier nécessairement à un autre matériel et sans en voir toutes les implications.

2.10 La traduction

La qualité de la traduction dépend du soin et de la précision avec lesquels la communication est paraphrasée ou transférée d'une langue ou d'une forme de communication dans une autre. La traduction est évaluée sur la base de la fidélité et de la précision, c'est-à-dire sur le degré de préservation du matériel original.

Exemples :

- être capable de comprendre des déclarations non littérales (métaphores, symbolismes, ironie, exagération)
- être capable de traduire du matériel mathématique verbal dans une écriture symbolique et vice versa.

2.20 L'interprétation

Il s'agit d'expliquer ou de résumer le contenu d'une communication. Alors que la traduction implique une reproduction littérale de la communication, l'interprétation implique un réordonnement, un réarrangement ou une révision du matériel.

Exemples :

- être apte à saisir le sens d'un travail à tout degré de généralité
- Savoir interpréter différents types de communication sociale.

2.30 L'extrapolation

Il s'agit d'aller au-delà des informations reçues, de déterminer les implications, conséquences, corollaires, effets, etc..., pouvant être inférées à partir de la communication originale.

Exemples :

- être apte à traiter les conclusions d'un travail en termes d'inférence immédiates effectuées à partir de déclarations explicites
- être capable de prédire le développement ultérieur de tendances actuelles.

3.00 L'APPLICATION

Il s'agit d'utiliser des abstractions dans des situations particulières et concrètes. Les abstractions peuvent avoir la forme d'idées générales, de règles, de procédures ou de méthodes généralisées. Elles peuvent également consister en principes techniques, idées ou théories, devant être rappelées et appliquées.

Exemples :

- appliquer les termes ou concepts scientifiques d'un article aux phénomènes décrits dans d'autres articles
- prédire les effets probables d'une perturbation au sein d'un état biologique auparavant en équilibre.

4.00 L'ANALYSE

Il s'agit de la décomposition d'une communication dans ses composantes élémentaires, de façon à mettre en évidence la hiérarchie des idées ou la relation entre elles. De telles analyses doivent rendre la communication plus compréhensible, indiquer ses principes d'organisation et découvrir la manière dont elle réussit à transmettre un contenu complexe.

4.10 L'analyse d'éléments

Il s'agit de l'identification des éléments contenus dans une communication.

Exemples :

- être capable de reconnaître des hypothèses non citées
- savoir distinguer les faits des hypothèses.

4.20 L'analyse de relations

Il s'agit d'extraire les connections et les interactions entre les éléments et les parties d'une communication.

Exemples :

- être apte à vérifier la cohérence d'hypothèses
- être capable de comprendre les relations entre les idées d'un passage.

4.30 L'analyse de principes d'organisation

Il s'agit de l'organisation, de l'arrangement systématique et de la structure de la communication. Cette forme d'analyse inclut la structure explicite autant que la structure implicite, ainsi que les bases, les arrangements nécessaires et la dynamique qui font de la communication une unité.

Exemples :

- savoir reconnaître la forme ou le modèle utilisé dans des oeuvres littéraires ou musicales comme moyen de comprendre leur sens
- être capable de reconnaître les techniques générales utilisées dans du matériel persuasif, comme par exemple les publicités, la propagande, etc...

5.00 LA SYNTHÈSE

Il s'agit de rassembler des éléments et des parties afin de former un tout, c'est-à-dire d'utiliser des pièces, parties, éléments, etc... et de les arranger et combiner d'une manière à constituer un modèle ou une structure nouvelle.

5.10 La production d'une communication originale

Il s'agit du développement d'une forme de communication personnelle par laquelle l'écrivain ou l'orateur essaye de transmettre ses idées, sentiments ou expériences à autrui.

Exemples :

- être capable d'écrire en utilisant une excellente organisation formelle et stylistique
- être capable de raconter une expérience personnelle d'une façon expressive.

5.20 La production d'un plan ou d'un ensemble d'opérations planifiées

Il s'agit du développement d'un plan de travail ou de l'élaboration de la suite des opérations à effectuer. Le plan devrait satisfaire les exigences de la tâche, proposée à l'auteur par autrui ou choisie par lui-même.

Exemples :

- proposer un plan expérimental pour tester des hypothèses
- planifier une situation particulière d'enseignement.

5.30 La déduction d'un ensemble de relations abstraites

Il s'agit du développement d'un ensemble de relations abstraites, dans le but de classer ou d'expliciter des données ou phénomènes particuliers, grâce à la déduction de propositions ou de relations d'un ensemble de propositions ou de représentations symboliques de base.

Exemples :

- l'aptitude de formuler des hypothèses appropriées, basées sur une analyse de facteurs impliqués et de modifier de telles hypothèses à la lumière de facteurs et considérations nouvelles
- faire des découvertes et des généralisation mathématiques.

	DIMENSIONS COGNITIVES					
	SAVOIR	COMPETENCES INTELLECTUELLES				
Niveau de complexité	Savoir	Compréhension	Application	Analyse	Synthèse	Evaluation
Elémentaire	Connaissance de détails	Traduction		Analyse d'éléments	Production d'une communication originale	Jugement en termes d'évidence interne
Moyen	Connaissance de moyens de traiter des détails	Interprétation		Analyse de relations	Production d'un plan ou d'un ensemble d'opérations planifiées	Jugement en termes de critères externes
Elevé	Connaissance des généralités et des faits abstraits dans un domaine	Extrapolation		Analyse des principes d'organisation	Déduction d'un ensemble de relations abstraites	

Tableau 4.1 : Les dimensions principales de la taxonomie de Bloom

6.00 L'ÉVALUATION

Il s'agit de faire des jugements sur la valeur qu'ont les matériaux et les méthodes, en vue d'objectifs donnés. Cette compétence implique la capacité de faire des jugements quantitatifs et qualitatifs sur l'adéquation du matériel et des méthodes par rapport à certains critères, ainsi que l'utilisation des normes d'évaluation. Les critères peuvent être fournis au sujet ou déterminés par lui-même.

6.10 Le jugement en termes d'évidence interne

Il s'agit de l'évaluation de la précision d'une communication, selon des critères internes comme l'exactitude logique et la cohérence.

Exemples :

- évaluer la précision en rapportant des faits d'une documentation
- trouver des fautes logiques dans une argumentation.

6.20 Le jugement en termes de critères externes

Il s'agit de l'évaluation de matériel, en se référant à des critères sélectionnés.

Exemples :

- comparer des théories majeures, des généralisations et des faits socio-cultures
- comparer un travail, selon des critères externes, avec des travaux d'un niveau supérieur dans le même domaine.

Malgré sa complétude, la taxonomie de Bloom a posé beaucoup de problèmes quand on essayait de l'adapter au domaine des mathématiques. Pour cette raison, les chercheurs ont élaboré différentes classifications qui conviennent mieux à ce domaine particulier de la cognition humaine.

4.2 L'Étude Internationale des Performances en Mathématiques

L'Étude Internationale des Performances en Mathématiques (*International Study of Achievement in Mathematics*) (Husen, 1967) a classifié les objectifs de l'enseignement en mathématiques, grâce à une analyse détaillée et des objectifs possibles et des objectifs réels poursuivis dans les programmes d'études des 12 pays participants. Les chercheurs sont arrivés à établir une matrice tridimensionnelle.

La première dimension spécifie le comportement que l'élève acquiert. En fait, le comportement comporte trois facettes, une cognitive, une affective et une motrice.

Les 10 aptitudes comportementales dégagées sont les suivantes :

- a) L'aptitude de se rappeler et de reproduire des définitions, notations, opérations et concepts.
- b) L'aptitude de calculer rapidement et efficacement et de manipuler des symboles.
- c) L'aptitude de traduire des données en symboles.
- d) L'aptitude d'interpréter des données qui apparaissent sous forme symbolique.
- e) L'aptitude de suivre une ligne de raisonnement ou de preuve.
- f) L'aptitude de construire une démonstration.
- g) L'aptitude d'appliquer des concepts à la résolution de problèmes mathématiques.

- h) L'aptitude d'appliquer des concepts à la résolution de problèmes non mathématiques.
- i) L'aptitude d'analyser des problèmes et de déterminer les opérations qui pourraient y être appliquées.
- j) L'aptitude de trouver des généralisations mathématiques.

La deuxième dimension est celle des différents contenus mathématiques. Elle comporte entre autres les thèmes suivants :

- a) Arithmétique
- b) Algèbre
- c) Géométrie plane, géométrie de l'espace, géométrie analytique
- d) Trigonométrie
- e) Probabilité, analyse
- f) Aspects généraux

La troisième dimension concerne l'utilisation des connaissances ou des compétences acquises. D'après M. Pellerey (1991), la raison pour la distinction qu'on y fait entre les mathématiques considérées comme un corps de savoir (mathématiques pures) et les mathématiques considérées comme un instrument (mathématiques appliquées) est probablement due à la division introduite dans l'éducation mathématique dans différents pays de la terre.

4.3 L'étude Longitudinale Nationale des Aptitudes Mathématiques

Le travail effectué par un groupe de chercheurs américains, intitulé Étude Longitudinale Nationale des Aptitudes Mathématiques (*National Longitudinal Study of Mathematical Abilities*) (Wilson, 1971) a abouti à un modèle à deux dimensions. Les facettes de la dimension comportementale sont les suivantes :

- a) Le calcul (*computation*) : cette aptitude consiste à se rappeler les faits mathématiques fondamentaux, la terminologie et les procédures nécessaires pour appliquer des algorithmes. L'accent est placé sur la simple évocation du savoir et l'accomplissement des opérations d'une manière appropriée.

- b) La compréhension (*comprehension*) : cette catégorie comporte l'aptitude de se rappeler des concepts et de faire une généralisation mathématique : l'élève doit montrer qu'il a compris les concepts et les relations qui les relient entre eux et transférer ces données.
- c) Application (*application*) : il s'agit de résoudre des problèmes familiers, similaires à ceux que l'élève avait à maîtriser pendant son processus d'apprentissage ; les élèves doivent choisir la méthode qui permet de trouver la solution et l'appliquer correctement.
- d) Analyse (*analysis*) : cette aptitude consiste à aller plus loin que ce qu'on a appris, à faire des expériences mathématiques nouvelles, à résoudre des problèmes inhabituels.

4.4 La taxonomie de M. Pellerrey

M. Pellerrey (1974) a fait une synthèse de ces premières taxonomies et a proposé de scinder la dimension comportementale en deux, en distinguant des processus de reproduction et des processus de production. Il discerne les aptitudes suivantes :

Processus de reproduction :

- a) Connaissance d'éléments isolés, non organisés en un tout, sous une forme répétitive courante. Connaissance d'organisations, mais sous forme de données purement mnémoniques. Ceci peut se passer à un degré verbal ou non verbal, comme dans le cas de graphiques, de l'application d'algorithmes, de constructions avec règle et boussole. Tous ces processus impliquent la mémoire ou la reconnaissance.

Font partie de cette catégorie :

La terminologie, les symboles.

Les principes et les règles.

Les faits, les énoncés mathématiques.

Le développement d'algorithmes.

La résolution de problèmes de routine.

Les capacités motrices, comme par exemple l'utilisation de la règle et de la boussole.

- b) Connaissance avec compréhension de concepts, d'organisations perçues comme un tout, construit précédemment. Cette aptitude se distingue de la précédente du fait qu'elle implique une plus grande activité, surtout d'analyse, de la part de celui qui apprend, sans pour autant demander un réel engagement constructif. Cette aptitude peut également se réaliser à un degré verbal ou non verbal.

Exemples :

Voir des relations, par exemple, entre une hypothèse et une conséquence ou voir des régularités.

Comprendre des organisations spatiales, temporelles ou logiques et leurs principes de structuration.

Encoder et traduire (au moyen de mots, de graphiques et de symboles).

Suivre un raisonnement.

Comprendre un problème.

Approfondir les moyens de résoudre des problèmes.

Processus de production :

- a) Construction de concepts et d'organisations, sur la base d'expériences ou de connaissances ou d'informations préalables : les associer ou les séparer au fur et à mesure qu'ils apparaissent.

Processus faisant partie de cette dimension :

Construction de concepts, de systèmes d'idées unifiées, de symboles.

Construction de schèmes, d'organigrammes, de graphiques, d'algorithmes.

Formulation d'une définition.

Construction d'une organisation logique (axiomatique).

Représentation de données ou d'informations.

Construction de jeux, résolution de problèmes simples.

- b) Résolution de problèmes. Cette aptitude concerne les problèmes véritables, au sens où leur solution n'est pas donnée par des procédés simples, comme le rappel ou la reconnaissance.

Exemples :

Deviner, supposer, formuler des hypothèses.

Trouver des objets, réels ou mathématiques, qui vérifient des conditions données.

Trouver tous les objets qui vérifient des conditions données.

Généraliser, étendre des concepts par analogie.

Trouver un modèle mathématique adéquat (parmi des modèles connus).

Construire un modèle mathématique apte à résoudre un problème.

Trouver un algorithme de résolution qui standardise un problème.

- c) Juger. Cette aptitude consiste à défendre un point de vue par argumentation ou à rejeter un jugement, une solution, etc... sur la base de principes ou critères venant à la fois de l'intérieur de l'objet en question et de l'extérieur.

Exemples :

Juger si un énoncé a un sens.

Juger si un énoncé est vrai.

Juger si un problème est bien défini, si on dispose d'assez, respectivement de trop de données, respectivement de données contradictoires.

Juger si un symbole, une définition, une solution proposée conviennent.

Juger si un raisonnement est correct.

Juger si une solution vérifie les hypothèses.

Juger si une solution est raisonnable et/ou pratique.

Juger si une solution est élégante et/ou stimulante.

4.5 La taxonomie de la TIMSS

La Troisième Étude Internationale des Mathématiques et des Sciences (*Third International Mathematics and Science Study*) (1991-1996) est l'étude internationale la plus étendue et plus ambitieuse qui ait jamais été faite dans le domaine de la comparaison de la performance d'élèves. On a comparé la performance en mathématiques et en sciences d'élèves de plus de 50 pays.

La structure du programme de la TIMSS fournissait un système unifié de catégories pour l'analyse des programmes d'études dans les différents pays. Le développement des tests d'évaluation se faisait suivant une structure deux-dimensionnelle, une dimension de contenu et une dimension de performance cognitive. La première décrivait la branche des mathématiques ou des sciences testée

et la seconde caractérisait le type de performance cognitive que l'élève devait accomplir pour résoudre l'item en question.

Les 6 catégories de contenu mathématiques pour les enfants de 12 ans étaient

- a) les nombres entiers
- b) les fractions et la proportionalité
- c) la mesure, l'estimation et le sens des nombres
- d) la représentation de données, l'analyse et les probabilités
- e) la géométrie
- f) les structures, les relations et les fonctions.

Les 4 catégories de performance cognitive pour les enfants du même âge étaient

- a) savoir
- b) utiliser des procédures de routine
- c) rechercher et résoudre des problèmes
- d) faire un raisonnement mathématique.

4.6 La taxonomie de John B. Carroll

Comme on l'a déjà mentionné dans le chapitre précédent, John B. Carroll (1993) a réanalysé plus de 480 ensembles de données provenant de la littérature psychométrique des 60 à 70 dernières années, à l'aide d'analyses factorielles exploratives, dans le but de trouver une taxonomie des aptitudes cognitives. Il a établi une taxonomie comportant trois strates.

La strate de niveau le plus élevé comporte un seul facteur, à savoir le facteur *g* d'intelligence générale.

La strate de niveau moyen comporte les huit facteurs énumérés ci-dessous, qui possèdent tous une sous-strate avec plusieurs facteurs de niveau inférieur.

1. L'INTELLIGENCE FLUIDE

1.1 Le raisonnement séquentiel

1.2 L'induction

1.3 Le raisonnement quantitatif

1.4 La raisonnement piagétien

1.5 La vitesse de raisonnement

2. L'INTELLIGENCE CRISTALLISEE

2.1 Le développement du langage

2.2 La compréhension du langage parlé

2.3 Le savoir lexical

2.4 La compréhension de l'écrit

2.5 Le décodage de l'écrit

2.6 Les connaissances orthographiques

2.7 Le codage phonétique

2.8 La sensibilité grammaticale

2.9 L'aptitude aux langue étrangères

2.10 L'aptitude à la communication

2.11 La capacité d'écouter

2.12 La vitesse de lecture

2.13 L'aisance d'expression orale

2.14 La capacité d'écriture

3. LA MEMOIRE ET LE SAVOIR

3.1 L'envergure de la mémoire

3.2 La mémoire associative

3.3 La mémoire de rappel libre

3.4 La mémoire significative

3.5 La mémoire visuelle

3.6 L'aptitude d'apprendre

4. LA PERCEPTION VISUELLE

4.1 La visualisation

4.2 Les relations spatiales

4.3 La vitesse de reconnaissance

4.4 La flexibilité de reconnaissance

4.5 L'intégration sérielle perceptuelle

4.6 Le balayage spatial

4.7 La vitesse de perception

4.8 La formation d'images

4.9 L'estimation de longueurs

4.10 La perception d'illusions

4.11 L'alternance de perceptions

5. LA PERCEPTION AUDITIVE

5.1 3 facteurs d'audition et de paroles

5.2 Le discernement des bruits de la parole

5.3 Le discernement de bruits généraux

5.4 Le discernement de la fréquence de bruits

5.5 Le discernement de la durée de bruits

5.6 Le discernement et le jugement musical

5.7 La résistance à la déformation de stimuli auditifs

5.8 Le pistage temporel

5.9 Le maintien et la perception de rythmes

5.10 La mémoire de motifs sonores

5.11 La tonalité absolue

5.12 La localisation de sons

6. LA CAPACITE D'ABSTRACTION

6.1 L'originalité créative

6.2 L'aisance d'idées

6.3 La capacité de nommer

6.4 L'aisance d'association

6.5 L'aisance d'expression

6.6 L'aisance linguistique

6.7 La sensibilité aux problèmes

6.8 L'aisance formelle

6.9 La flexibilité formelle

7. LA VITESSE DE PENSEE

7.1 La vitesse d'exécution des tests

7.2 La capacité numérique

7.3 La vitesse de perception

8. LA VITESSE D'EXECUTION

8.1 Le temps de réaction simple

8.2 Le temps de réaction à un choix

8.3 La vitesse de traitement sémantique

8.4 La vitesse de comparaison mentale

Carroll conclut qu'il n'existe pas de capacités uniquement mathématiques, à part peut être deux, mais que la plupart des abilités cognitives entrent en action lors de la résolution de problèmes mathématiques, ce qui rend très ardu l'établissement d'une taxonomie mathématique.

4.7 Synthèse

Alors que les taxonomies de Bloom et de Carroll concernent les compétences cognitives en général, les 4 taxonomies mathématiques qui ont été présentées ci-dessus ont été établies à partir de l'analyse des programmes de mathématiques existants. Elles se sont cependant inspirées des recherches psychométriques et cognitives, au sens où elles distinguent les procédures des contenus. Il n'est donc pas étonnant qu'elles soient assez semblables en ce qui concerne les contenus scolaires et qu'elles se distinguent seulement par la mise en évidence des procédures qu'elles jugent pertinentes et qui dépendent évidemment de l'âge des élèves examinés. La distinction entre processus de production et de reproduction introduite par Pellerey correspond à la reconnaissance de l'importance du rôle de la mémoire dans les performances scolaires.

Nous allons tester ces taxonomies sur la base de nos propres données.

Chapitre V

METHODES MULTIVARIEES D'ANALYSE STATISTIQUE

Dans ce chapitre, nous introduisons quelques méthodes d'analyse statistique multivariée que nous utiliserons dans la suite.

5.1 L'échelonnement multidimensionnel

L'objectif des méthodes d'échelonnement multidimensionnel (*multidimensional scaling*) est de représenter des données de proximité sous formes de distances, dans un espace d'un nombre restreint de dimensions (Dickes et al., 1994). A partir d'un ensemble de données représentant différentes formes de proximité entre items (similarités, dissimilarités, distances psychologiques, cooccurences, correlations, ...), on cherche si la représentation des items par rapport à quelques axes de coordonnées est réalisable.

L'exemple suivant (Tournois et Dickes, 1993) donnera une idée approximative de l'objectif de ces échelonnements. Supposons que nous disposions d'une liste des distances routières entre différentes villes françaises. Est-il possible de reconstituer la carte de la France, à partir de cette seule information?

Le tableau des distances kilométriques, qui rassemble les données soumises dans l'analyse, est présentée dans le tableau ci-dessous. Ces données ne sont, pour la réalisation d'une carte, qu'approximatives, puisqu'il ne s'agit pas de distances à vol d'oiseau mais de distances routières: la distance entre les villes des zones montagneuses est, comparativement plus grande que celle entre les villes situées en plaine.

1																					
2	178																				
3	811	636																			
4	816	641	401																		
5	544	369	763	654																	
6	793	618	803	648	278																
7	809	653	991	894	284	302															
8	802	612	477	225	556	528	782														
9	953	778	734	474	606	463	765	286													
10	396	221	601	540	184	397	468	500	593												
11	775	548	886	789	179	197	105	677	660	363											
12	678	649	1194	1091	437	511	277	991	974	593	314										
13	520	491	1038	975	367	488	287	923	973	435	291	158									
14	506	331	305	310	458	589	733	381	608	296	628	889	731								
15	847	818	1363	1227	596	635	333	1102	1089	762	438	190	327	1058							
16	734	559	575	341	387	307	561	221	219	374	456	764	754	389	894						
17	442	459	1095	1054	467	649	448	1014	1073	514	452	319	161	790	488	854					
18	612	437	244	204	507	559	747	279	520	375	642	944	810	106	1080	345	889				
19	1099	924	1020	786	584	306	522	666	503	703	434	748	725	823	855	446	886	790			
20	279	250	886	845	381	659	630	805	898	305	496	399	241	581	568	679	209	680	930		
21	505	330	453	356	298	394	538	297	448	203	433	736	638	195	871	229	717	209	628	508	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21

Tableau 1.1 : Distances kilométriques entre 21 villes françaises

1. Bayonne	8. Le Havre	15. Nice
2. Bordeaux	9. Lille	16. Paris
3. Brest	10. Limoges	17. Perpignan
4. Cherbourg	11. Lyon	18. Rennes
5. Clermont-Ferrand	12. Marseille	19. Strasbourg
6. Dijon	13. Montpellier	20. Toulouse
7. Grenoble	14. Nantes	21. Tours

Figure 5.1 : Distances kilométriques entre 21 villes françaises (Tournois et Dickes, 1993)

L'analyse par échelonnement multidimensionnel a pour objectif de restituer la position respective des villes, à partir de ces distances routières, donc de restituer la carte de la France. Toutefois, comme les données dont on dispose sont erronées du point de vue de la représentation cartographique (on peut dire qu'elles sont entachées d'erreurs), l'échelonnement ne prend pas en compte les distances routières de ce tableau en tant que valeurs absolues: le modèle ne retient que l'ordre des distances entre villes ; il retient seulement que Grenoble et Lyon (105) sont plus proches l'une de l'autre que ne le sont Nantes et Rennes (106), elles-mêmes étant plus proches que ne le sont Marseille et Montpellier (158), etc

Le modèle transforme l'information d'ordre du tableau des données en des distances euclidiennes qui respectent cet ordre et cherche la position théorique des villes qui respecte ces distances. Il la traduit dans une représentation géométrique, présentée ci-dessous.

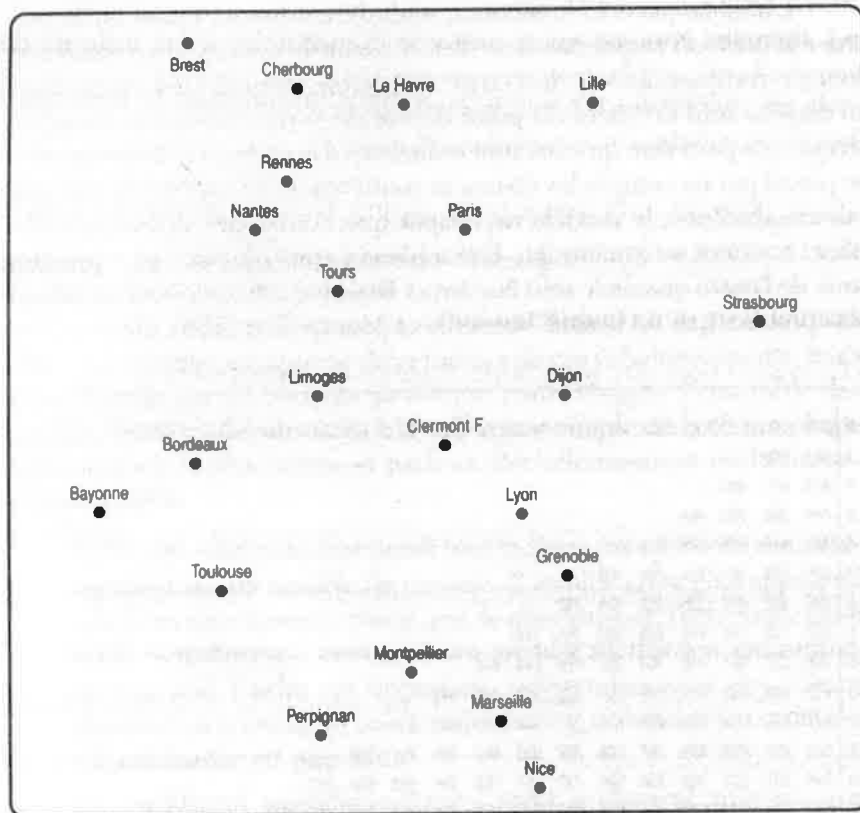


Figure 1.1 : Configuration bidimensionnelle issue de l'EMD des distances routières entre 21 villes françaises.

Figure 5.2 : Configuration bidimensionnelle des distances routières (Tounois et Dickes, 1993)

L'analyse nous informe que, pour représenter ces données, deux dimensions suffisent : c'est la dimensionnalité appropriée pour cette représentation. Le terme dimensionnalité renvoie simplement au nombre de dimensions de l'espace dans lequel sont représentées les données, c'est-à-dire au nombre d'axes de coordonnées de la solution.

Il va sans dire que dans les applications réelles, la dimensionnalité vraie est inconnue au départ ; elle doit donc être recherchée par l'analyste. L'échelonnement multidimensionnel permettra de déterminer la dimensionnalité la plus petite qui soit suffisante pour représenter correctement les données. Dans le cas présent, l'échelonnement multidimensionnel signale que les distances entre villes sont très mal

représentées sur une seule dimension : l'information de départ est falsifiée et ne peut entraîner que le rejet d'un tel échelonnement pour cause d'inefficacité. D'autre part, ces distances ne sont pas mieux représentées dans un espace tridimensionnel que dans l'espace à deux dimensions de la figure 5.1, qui constitue donc la solution retenue.

Les modèles de l'échelonnement multidimensionnel sont donc des modèles de représentation spatiale. Pour en avoir une vue d'ensemble, il est utile de distinguer trois aspects de ces modèles : les données, la transformation et la métrique (Coxon, 1982).

- Les *données* peuvent être constituées d'une (ou de plusieurs) matrice(s) de proximités, carrée(s) symétrique(s) ou asymétrique(s) ou rectangulaire(s) conditionnelle(s) ou non conditionnelle(s).
- La *transformation* peut être monotone, linéaire ou logarithmique.
- La *métrique* peut être une quelconque des distances de Minkowski.

Chaque croisement de ces caractéristiques fournit un modèle différent. Comme nous ne pourrions naturellement pas présenter ici l'ensemble de ces modèles, nous nous bornerons à expliciter le modèle de base élaboré par Shepard (1962) et Kruskal (1964a, 1964b), qui est encore à l'heure actuelle le modèle le plus largement répandu. Dans ce modèle, l'échelonnement réalise une transformation monotone des données d'entrée en distances euclidiennes, c'est-à-dire que seul l'ordre des données d'entrée est pris en compte dans la transformation. Pourtant, à la sortie, les données sont représentées dans un espace euclidien. Le résultat est un système de coordonnées des points dans un espace à n dimensions. Les analyses par échelonnement multidimensionnel sont donc des ré-échelonnements. Il y a généralement passage de données d'un faible niveau de mesure (ordinal le plus souvent) à des résultats d'un niveau de mesure plus élevé, le niveau métrique.

Si l'on se réfère au résultat de l'analyse, les échelonnements multidimensionnels sont des techniques de représentation géométrique. En effet, les données de départ sont, à l'issue de l'analyse, représentées géométriquement dans un espace euclidien.

Une remarque s'impose : ce ne sont pas nécessairement des techniques qui visent à rechercher des « facteurs » sous-jacents au phénomène étudié. Ces techniques résument l'information en la représentant une forme immédiatement perceptible.

L'appréhension visuelle d'une carte est, certes, plus facile que la lecture d'une grande matrice de données. C'est aussi la totalité de l'information qui est représentée sur la carte géographique. Ces techniques ne réduisent pas l'information, elles la condensent en une forme accessible.

D'après Tournois et Dickes (1993), les échelonnements multidimensionnels reposent sur un modèle de mesure qui présente plusieurs avantages :

- il est très peu exigeant sur les données qu'il permet de traiter : l'information ordinale suffit ;
- il est peu contraignant dans son fonctionnement : une transformation monotone suffit ;
- il est avantageux dans sa réalisation : les distances de la configuration finale sont du niveau d'intervalle et dépassent donc l'information ordinale des données de départ.

La plus grande source de diversification des modèles d'échelonnement multidimensionnel reste la diversité des données qu'ils peuvent traiter. Il s'est imposé une distinction entre données de *similarité directe* et données de *similarité indirecte*. Les similarités directes sont des données dont les observations comportent déjà la signification de similarité (ou de dissimilarité). La collecte porte donc nécessairement sur les relations entre éléments. Les similarités indirectes sont des données de similarité obtenues par le calcul d'un coefficient ou d'un indicateur quelconque de proximité (ou distance), à partir d'observations ne contenant en elle-mêmes aucun indice de similarité.

Historiquement, les échelonnements multidimensionnels étaient destinés au traitement de données de similarité directe. La pratique de la recherche en a étendu l'usage aux proximités calculées sur des profils, ce qui ouvre incontestablement les champs potentiels d'application : toutes les observations de type « stimulus unique » se prêtent à l'application d'un échelonnement multidimensionnel, après calcul d'indicateurs de profil, telle la corrélation, qui les transforment en données de proximité ou de dissimilarité.

Les similarités directes offrent cependant l'avantage de ne pas introduire de présupposés quant aux dimensions qui sous-tendent le jugement, le comportement, la perception, etc On s'est donc efforcé de forger des techniques ou, plutôt,

d'aménager des techniques de collecte déjà existantes, pour recueillir des similarités directes.

Les similarités directes étant des informations relationnelles, les techniques de collecte excluent la présentation de type stimulus unique. Les techniques les plus utilisées sont les suivantes :

- Le *tri subjectif* est très utilisé pour sa rapidité d'exécution lorsque les stimuli sont nombreux. Il permet de récolter des observations nominales : sont similaires les éléments classés dans une même catégorie.
- L'*estimation de similarité de paires* est une technique plus lourde puisqu'elle suppose la présentation de paires d'items. Le relevé de confusion (assimilation de deux stimuli présentés simultanément ou successivement) en est une variante. Plus généralement, la présentation de paires d'items, assortie d'une évaluation de leur similarité, soit dichotomique, soit ordinale, permet l'obtention de similarités directes.
- La *méthode des tétrades* (présentation de quatre stimuli) est une transposition de la technique de collecte des comparaison par paires, dans laquelle on présente des paires de paires. Cette technique a deux inconvénients majeurs : elle est très lourde et la tâche est difficile pour le sujet.
- La *méthode des triades* (présentation de trois stimuli) se veut un allègement de la précédente. On parle quelquefois de « tétrades conjointes » pour en rendre compte, considérant ainsi que trois éléments constituent une tétrade ayant deux éléments en commun.
- La technique des *rangements conditionnels* se présente comme une extension de la collecte par ordination. Prenant tour à tour chacun des items comme terme de comparaison, l'ensemble des items restants est ordonné par rapport à ce terme, en fonction de sa dissimilarité. Ainsi, s'il y a 15 items, cette technique réclame 15 ordinations de 14 stimuli. Ce type de collecte est plus simple dans sa présentation que la technique des triades et offre le même nombre de comparaisons (Roskam, 1979).

Plus généralement, l'observation de similarité est un fait courant. Par exemple, pour un ensemble de périodiques scientifiques, le relevé bibliographique du nombre de citations des différents périodiques par chacun d'entre eux permet de dresser un tableau de proximités asymétriques entre les périodiques (Coxon, 1982). La fréquence des rencontres entre les personnes, telle qu'elle se présente dans un sociogramme, est

peut-être l'exemple le plus courant de proximités observées en situations naturelles. Ainsi, en dehors du recours à des techniques spécifiques de collecte de similarités directes, l'observation permet d'obtenir des similarités de ce type, qui se prêtent à un traitement par échelonnement multidimensionnel.

Historiquement, on peut reconnaître deux grands courants de travaux qui participent indépendamment aux développements du modèle de base, le courant des *Bell Laboratories*, avec le programme de Shepard-Kruskal et le courant de Guttman-Lingoes, avec les programmes de la série SSA (*Smallest Space Analysis* ou *Similarities Structure Analysis*) (cf. Lingoes, 1972).

Le modèle de base traite une *matrice symétrique de proximités*. Il est indifférent que les données soient des dissimilarités ou des similarités, la fonction qui les relie aux distances étant seulement définie comme croissante ou décroissante selon le cas. Il est toutefois plus simple, sur le plan logique, de considérer les dissimilarités. Dans ce cas, les grandes dissimilarités doivent être représentées par de grandes distances, les petites, par de petites distances.

Le modèle réalise une *transformation monotone* des dissimilarités. Cette transformation est faiblement monotone dans le modèle de Shepard-Kruskal, strictement monotone dans le modèle de Guttman-Lingoes.

La métrique est celle de la *distance euclidienne*.

Les points (correspondant aux items) sont, à l'issue de l'application du modèle, positionnés dans un espace euclidien d'une certaine dimension.

Le coeur du modèle réside dans la transformation monotone des dissimilarités en distances : les distances euclidiennes croissent quand les dissimilarités croissent.

Alors que les distances euclidiennes résultantes réalisent nécessairement les axiomes de la métrique euclidienne, les dissimilarités de départ ne les respectent pas obligatoirement. En d'autres termes, les données ne sont pas nécessairement représentables dans un espace de petite dimension. Il y a des dérogations au modèle. Celles-ci sont appréciées par un indicateur d'inadéquation appelé *stress* dans le modèle de Shepard-Kruskal respectivement *coefficient d'aliénation* dans le modèle de Guttman-Lingoes.

L'échelonnement multidimensionnel est réalisé par une procédure itérative dont on peut trouver l'exposé mathématique chez Takane, Young et De Leeuw (1977) et une présentation détaillée, moins technique, chez Tournois et Dickes (1993).

L'échelonnement multidimensionnel travaillant dans une dimensionnalité donnée, la pratique du chercheur consiste, en fait, à obtenir plusieurs solutions (en 5, en 4, en 3, en 2 dimensions) et à retenir la solution la plus intéressante : le plus souvent, la plus petite dimensionnalité produisant un résultat qui a du sens.

Le choix de la dimensionnalité reste le plus souvent intuitif. Il repose sur la consultation de la chute des valeurs de stress (celles-ci décroissent toujours quand la dimensionnalité augmente) et sur la possibilité d'interpréter la configuration. Cette pratique paraît ressembler au choix du nombre de facteurs dans une analyse factorielle, mais elle est pourtant différente. Elle est en fait beaucoup plus simple, ne serait-ce que parce que les échelonnements multidimensionnels n'effectuent pas de solutions au-delà de six dimensions, limite posée par les logiciels. Cette limite s'explique par le fait que soit les dissimilarités sont représentables au plan métrique et, dans un tel cas, elles le sont dans une dimension petite, soit elles ne sont pas représentables et dans ce cas, il est inutile d'augmenter la dimensionnalité. En pratique, on trouve généralement une solution acceptable en deux ou trois dimensions.

Deux mises en garde doivent toutefois être faites pour le choix de la dimensionnalité :

- 1) On obtient toujours de bonnes valeurs de stress avec très peu d'éléments à échelonner. Tous les auteurs mettent en garde contre le risque de demander une solution de dimensionnalité trop élevée par rapport au nombre d'items. Par exemple, pour une solution bidimensionnelle, il faut avoir de neuf à douze items, pour une solution tridimensionnelle, de treize à dix-huit items (Spence et Domoney, 1974 ; Kruskal et Wish, 1978 ; Schiffman, Reynolds et Young, 1981 ; Coxon, 1982). En effet, la représentation géométrique des données sera toujours d'autant meilleure que la solution demandée aura plus de dimensions. Et il est toujours possible de représenter quatre points en trois dimensions avec un stress nul, et même en deux dimension, si la distance est euclidienne (Lingoes, 1971).
- 2) Une valeur de stress exceptionnellement bonne (proche de zéro) est le plus souvent le signal d'une « dégénérescence » de la solution. On dit que la solution est dégénérée quand, en présence de sous-groupe très nets dans les données, les points tendent à s'agglomérer selon ces sous-groupes. En effet, de par sa nature ordinale, l'algorithme peut toujours rapprocher les similarités intragroupe et éloigner les similarités intergroupes. Dans les cas extrêmes, le stress tend vers

zéro (Shepard, 1974). Ainsi, une très bonne valeur de stress n'est souvent que le reflet d'une dégénérescence partielle des solutions. Ce phénomène n'est pas rare dans les applications qui traitent par exemple, des données mettant en relation soi et autrui ou l'ingroup et l'outgroup. Les applications de ce type produiront toujours des stress d'apparence exceptionnelle, qui ne garantissent pas la qualité du résultat en termes de mesure. En face d'une solution dégénérée, le seul résultat acquis est la présence de sous-groupes. La pratique consiste soit à recommencer des échelonnements multidimensionnels séparés pour chacun des sous-groupes, soit à augmenter la dimensionnalité (si le nombre d'éléments à échelonner le permet), soit à réaliser un échelonnement multidimensionnel en recourant à une transformation linéaire (si la nature des données le permet).

Dans la configuration résultante, toutes les distances entre points sont significatives. Mais la configuration exprimée en distances *euclidiennes* permet la translation d'origine, la rotation orthogonale des axes, la réflexion et le changement d'échelle.

Il n'est pas rare que ces configurations prennent une structure évidente, par exemple, la forme d'un cercle, la forme d'un fer à cheval, représentant une dimension qui s'est incurvée en raison de la nature ordinale de l'algorithme, ou de deux ou trois sous-ensembles de points, comme dans les phénomènes de dégénérescence. La conduite interprétative est dès lors facilitée par l'organisation globale de la configuration. Le cas général est toutefois l'apparence de points qui ne présentent pas de structure manifeste. Dans ce cas, la conduite interprétative peut être engagée selon plusieurs voies.

- Une *approche dimensionnelle* consiste à rechercher des grands axes porteurs de signification, qui traversent l'ensemble de la configuration. Il faut toutefois garder à l'esprit que les dimensions significatives ne sont pas nécessairement portées par les axes de coordonnées et peuvent être recherchées dans n'importe quelle orientation. Cette recherche de grandes dimensions significatives peut être vérifiée, lorsqu'on dispose de l'information nécessaire, par la projection d'une dimension interprétative dans l'espace résultant de l'échelonnement multidimensionnel (Kruskal et Wish, 1978).
- Des *groupes de points* peuvent être recherchés dans cette configuration. Cette recherche peut être validée par la mise en oeuvre d'une analyse en grappes

(analyse en « clusters »), pratiquée sur la matrice de proximités qui a été soumise à l'échelonnement. On valide ainsi, par convergence, la plus grande proximité des distances intragroupe. Les résultats de cette analyse peuvent être représentés manuellement dans l'espace de l'échelonnement multidimensionnel (cf. Fillenbaum et Rapoport, 1971).

- L'interprétation peut aussi consister à rechercher des *régions* dans la configuration, c'est-à-dire des zones contiguës de l'espace, quelle qu'en soit par ailleurs la forme, dans lesquelles tous les points ont une signification commune. Cette démarche est essentiellement utilisée dans une approche confirmatoire, le plus souvent en liaison avec une phrase en facettes qui a permis de spécifier à l'avance les significations à rechercher. L'*approche en facettes* constitue le prototype d'une telle pratique.

Ces diverses voies interprétatives ne s'excluent pas. J. Tournois (1989) montre qu'elles gagnent à être pratiquées conjointement, surtout lorsque la recherche est exploratoire. La multiplicité des voies interprétatives qui peuvent être appliquées à ces configurations est une richesse qui fait de l'échelonnement multidimensionnel un outil exploratoire puissant. Dans une approche confirmatoire, la voie interprétative est naturellement guidée par les hypothèses posées.

5.2 L'analyse hiérarchique ascendante

5.2.1 Introduction

L'analyse hiérarchique ascendante (appelée également analyse en grappes ou clusters) répond au problème suivant : comment classer n individus, sur lesquels on a mesuré p caractéristiques X_1, \dots, X_p , en un certain nombre K de sous-groupes homogènes? L'analyse se déroule en deux phases :

- a) il s'agit d'abord de définir un indice de dissimilarité entre toutes les paires d'individus, d'autant plus élevé que leurs profils (sur les p caractéristiques) sont éloignés.
- b) il faut ensuite convenir d'une règle de regroupement permettant de décider si deux individus doivent appartenir au même groupe, pour un seuil donné (représentant

un niveau de précision ou de pouvoir discriminant) : à un certain niveau, deux individus peuvent être confondus dans le même groupe, alors qu'à un niveau de précision plus élevé ils seront distingués et appartiendront donc à deux sous-groupes différents. En faisant varier le seuil, les n individus au départ distincts se trouveront ultimement rassemblés dans un groupe unique. Le dendogramme constitue une représentation graphique de ce processus d'agrégation.

D'après François Bavaud (Bavaud et al., 1996), l'analyse hiérarchique ascendante est un outil un peu à part dans l'univers des méthodes statistiques :

- Elle ne comporte pas de tests d'hypothèses : l'aspect inférentiel est donc inexistant. De fait, la « justification » d'une analyse hiérarchique ascendante est son aptitude à produire des classifications qui « font sens ».
- Elle possède un grand nombre de variantes fort diverses et ceci pour le traitement des deux phases mentionnées ci-dessus ; comme la classification obtenue dépend en général fortement des options choisies, l'analyse hiérarchique ascendante se présente comme un ensemble de recettes empiriques, dont seule une petite partie est adaptée à un type spécifique de données.
- Enfin, l'accent est mis sur l'analyse des n individus plutôt que sur les p variables, comme c'est le cas habituellement. Cette approche centrée sur les individus résulte souvent non pas de l'intérêt particulier porté aux individus eux-mêmes mais bien du constat que les alternatives sont limitées lorsque l'échantillon est petit ($n < p$).

Pour toutes ces raisons, l'analyse hiérarchique ascendante est un outil exploratoire dont le seul mérite est de produire un ordre qui « parle » au chercheur et lui suggère de nouvelles hypothèses et pistes d'analyse. Il est conseillé au chercheur de varier les options d'analyse sur un même jeu de données et d'écarter sans trop d'états d'âme les classifications trop peu éclairantes.

5.2.2 Les indices de dissimilarité

5.2.2.1 Généralités

La première étape d'une classification hiérarchique ascendante consiste à définir un indice de dissimilarité d_{ij} entre toutes les paires (i, j) d'individus. Un indice de dissimilarité doit satisfaire les conditions de

(1) non-négativité : $d_{ij} \geq 0$

(2) symétrie : $d_{ij} = d_{ji}$

(3) normalisation : $d_{ii} = 0$.

Si, en plus, on a

(4) l'inégalité triangulaire : $d_{ij} \leq d_{ik} + d_{kj}, \forall k$,

on parle de distance métrique ou simplement de distance. Enfin, lorsque

(5) l'inégalité ultramétrique : $d_{ij} \leq \max(d_{ik}, d_{jk}), \forall k$,

est satisfaite, on parle de distance ultramétrique. Dans un monde ultramétrique, trois objets i, j et k forment toujours un triangle isocèle, les deux côtés égaux étant plus grands que le troisième. Une distance ultramétrique est forcément métrique.

Il faut parfois considérer des indices de similarité s_{ij} , satisfaisant aux conditions $s_{ij} \geq 0, s_{ij} = s_{ji}$ et $s_{ii} = s_{\max} = \text{constante}, \forall i$. Une similarité s_{ij} peut être convertie en distance d_{ij} par $d_{ij} = s_{\max} - s_{ij}$. Pour les puristes, la transformation suivante

$$\tilde{d}_{ij} = d_{ij} + \max_k \max(d_{ij}, d_{ik}, d_{jk})$$

transforme une dissimilarité d_{ij} en distance métrique \tilde{d}_{ij} .

5.2.2.2 Scores quantitatifs

Lorsque les p variables sont quantitatives, on peut définir une distance d_{ij} entre les profils (x_{i1}, \dots, x_{ip}) de i et (x_{j1}, \dots, x_{jp}) de j , de quantité de façons, répertoriées ou non, en particulier la distance euclidienne, la distance euclidienne au carré, la

distance city block, la distance de Chebychev, les distances de Minkovski et les distances en puissances généralisée.

De plus, il est toujours possible de considérer des scores x_{ii}^s et x_{ji}^s standardisés : cela revient à déclarer chacune des p variables également importantes dans la classification, plutôt que de privilégier celles dont l'écart-type est important. On peut aussi considérer une « standardisation multivariée » (éliminant également les corrélations entre variables) avec la distance de Mahalanobis.

En outre, le carré du coefficient de corrélation entre les individus i et j peut être utilisé comme indice de similarité s_{ij} , avec $s_{\max} = 1$. Pour que cet indice de similarité fasse sens, il est nécessaire que les p variables X_k soient du même type. La distance métrique $d_{ij} = \sqrt{1 - r_{ij}^2}$ obtenue ainsi est appelée distance de Pearson.

5.2.2.3 Scores catégoriels

Pour p variables qualitatives, on se limitera principalement au cas fréquent de variables binaires, exprimant de fait la présence ou l'absence d'un attribut qu'il est commode de coder par « 1 » (présence de l'attribut) et par « 0 » (absence de l'attribut). Au besoin, on effectuera une dichotomisation des variables plurimodales, transformant une variable à m modalités en $m - 1$ variables binaires de présence/absence.

Deux individus i et j étant donnés, il est nécessaire de distinguer la co-occurrence d'un attribut de la co-absence de l'attribut, la première étant souvent plus significative que la seconde (la co-occurrence de « possède une résidence secondaire » pouvant par exemple refléter une similarité plus forte que ne le ferait sa co-absence). Il existe plusieurs possibilités de définir des distances à l'aide des nombre de co-occurrences respectivement de co-absences.

5.2.2.4 Scores de comptage

Lorsque les profils des individus sont constitués de p effectifs, il est naturel de définir la dissimilarité associée à une paire d'individus comme le χ^2 du tableau $(2 \times p)$ de contingence associé, ou, de façon presque équivalente, par $\phi^2 = \chi^2 / n$, n étant l'effectif total du tableau. L'indice de dissimilarité ainsi défini est symétrique, non-négatif, et nul si et seulement si les profils des deux individus sont proportionnels.

5.2.3 La constitution des groupes

Etant donné les distances d_{ij} entre individus, il s'agit de construire une partition dont les q classes ou groupes sont aussi homogènes que possible. La classification hiérarchique ascendante consiste à regrouper les 2 individus les plus proches, et à réitérer le processus jusqu'au regroupement complet. L'algorithme aboutit donc à un arbre de classification ou dendrogramme. L'échelle verticale est le niveau d'agrégation $D(A, B)$, i.e. la distance à laquelle les groupes A et B ont été réunis.

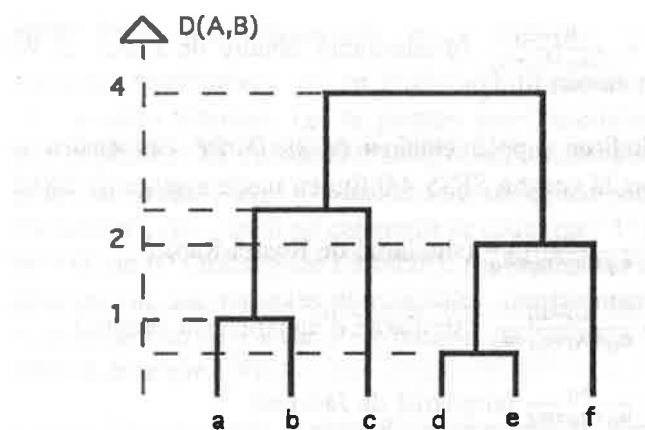


Figure 5.3 : Exemple de dendrogramme (Bavaud et al., 1996)

L'algorithme sera donc entièrement spécifié une fois que le sera la distance $D(A, B)$ entre deux groupes.

- La méthode du saut minimal (*single linkage* ou *nearest neighbourhood*) consiste à prendre

$$D_m(A, B) = \min_{i \in A} \max_{j \in B} d(i, j).$$

- La méthode du saut maximal (*complete linkage* ou *furthest neighbor*) consiste à prendre

$$D_M(A, B) = \max_{i \in A} \max_{j \in B} d(i, j).$$

A partir d'un nuage étiré de n points, l'algorithme du saut minimal tendra à ajouter les individus un à un au cluster déjà formé (chaînage), tandis que l'algorithme du saut maximal tendra à former des sous-sous-clusters de taille similaire qui seront regroupés dans des sous-clusters de taille similaire, etc ... (dendogramme homogène).

- Un compromis entre ces deux critères est fourni par la méthode du saut moyen entre les groupes (*baverage* ou *between-groups linkage*) avec

$$D_b(A, B) = \frac{1}{n_A n_B} \sum_{i \in A, j \in B} d(i, j)$$

où n_A et n_B sont les effectifs des deux groupes. Ce critère ne fait pas intervenir les dissimilarités intra-groupes : il garantit que les deux clusters A et B soient composés d'individus proches en moyenne, mais ne garantit pas l'homogénéité de chacun des deux clusters.

- La variante du saut moyen intra-groupes (*waverage* ou *within-groups linkage*)

$$D_w(A, B) = \frac{1}{(n_A + n_B)(n_A + n_B - 1)} \sum_{(i, j) \in (A \cup B)} d(i, j)$$

garantit également l'homogénéité intragroupe.

- La matrice $(p \times p)$ de variance-covariance totale T se décompose en variance interclasse et intraclasse $T = B^{(q)} + W^{(q)}$. Au début du processus d'agglomération, $q = n, B^{(n)} = T$ et $W^{(n)} = 0$. A mesure que les clusters sont formés, q et $B^{(q)}$ diminuent tandis que $W^{(q)}$ augmente. A la fin, $q = 1, B^{(1)} = 0$ et $W^{(1)} = T$. Par construction, la perte d'inertie intraclasse $W^{(q)} - W^{(q+1)}$ est positive. La méthode de

Ward (*Ward's method*) consiste à agglomérer les deux groupes A et B de façon à ce que la trace de la perte d'inertie intraclasse soit minimale. On obtient ainsi

$$D_W(A, B) = \frac{n_A n_B}{n_A + n_B} (\bar{X}_A^2 + \bar{X}_B^2),$$

où \bar{X}_A et \bar{X}_B désignent les centres de gravité des ensembles A respectivement B .

- Enfin, la méthode centroïde (*centroid clustering*) consiste à prendre simplement

$$D_C(A, B) = \bar{X}_A^2 + \bar{X}_B^2.$$

De façon générale, l'indice d'aggrégation $D(A, B)$ induit une distance D_{ij} entre individus, qui est définie comme le niveau d'agrégation minimal où i et j ont été réunis dans un même groupe. La distance D_{ij} est ultramétrique ; de fait, à tout dendrogramme est associé une structure de distance ultramétrique et vice-versa.

Or, les distances d_{ij} de départ entre individus ne sont pas ultramétriques en général ; toute classification, induisant une distance ultramétrique entre individus, crée donc une distorsion dans les données initiales.

5.3 L'analyse factorielle

5.3.1 Généralités

L'analyse factorielle est un procédé qui ramène un grand nombre de variables à un petit nombre de grandeurs causales latentes appelées facteurs. Cela se fait en regroupant les variables qui corréleront fortement entre eux sous un facteur. Ainsi des variables appartenant à des facteurs différents sont corrélées peu entre eux. Le but de l'analyse factorielle est par conséquent de déterminer des facteurs qui permettent d'expliquer le plus complètement possible les rapports observés entre les différentes variables.

Les méthodes d'analyse factorielle figurent sans doute parmi les modèles psychométriques les plus connus. L'ancienneté de ces méthodes (elles remontent au début du siècle) et leur présence dans tous les programmes intégrés d'analyses statistiques ont beaucoup contribué à leur diffusion. D'après Dickes et al. (1994), les

principes de base qui sous-tendent ces modèles ne sont pourtant pas toujours bien connus des chercheurs, ce qui donne lieu à des utilisations mal appropriées (comme le note Reuchlin (1992). Nous rappelons ces principes ici.

Il est nécessaire de parler d'analyses factorielles au pluriel pour rendre compte de la diversité des méthodes : outre l'analyse en composantes principales et l'analyse en facteurs communs et uniques évoquées plus loin, on peut citer, par exemple, l'analyse factorielle des images de Guttman (1953), l'analyse factorielle alpha de Kaiser et Caffrey (1965) ou l'analyse factorielle des correspondances de Benzécri (1973).

L'objectif de l'analyse factorielle est de condenser un ensemble d'observations qui se présentent sous la forme d'une matrice à deux modes, sujets-variables. La plupart du temps, ce sont les similarités entre les variables qui intéressent le chercheur, mais on peut également effectuer une analyse factorielle sur les sujets (on parle alors d'analyse factorielle en plan Q).

Pour pouvoir condenser les observations, il faut postuler l'existence de variables latentes appelées facteurs, qui doivent rendre compte des relations entre les variables observées (ou variables manifestes). Généralement, on a besoin de plusieurs facteurs pour rendre compte de manière satisfaisante de ces relations. Lorsque le chercheur n'a pas d'hypothèse a priori sur le nombre de variables latentes nécessaires, on dit que sa démarche est exploratoire. Dans le cas opposé, lorsque le chercheur a des hypothèses précises sur le nombre de facteurs nécessaires ou, plus généralement, sur différents paramètres de la solution factorielle, on parle de démarche confirmatoire. La frontière entre démarche exploratoire et confirmatoire n'est pas toujours très nette ; Il est rare en effet de n'avoir aucune attente par rapport aux résultats. Mais, en tout état de cause, même dans une démarche confirmatoire, l'analyse factorielle ne fournit pas de critère statistique d'adéquation permettant d'évaluer la plausibilité de telles hypothèses. Ce n'est qu'avec les modèles d'équations structurales que de telles possibilités existent.

La démarche commune à toutes les analyses factorielles peut être résumée de la manière suivante. C'est la matrice de corrélations ou de covariances entre les variables qui constitue les données d'entrée du modèle. S'agissant de corrélations, les variables devraient, en toute rigueur, être du niveau de mesure d'intervalles. Comme un tel niveau est rarement atteint avec des variables psychologiques, l'utilisation de l'analyse factorielle s'est généralisée à des variables ordinales. Mais dans ce cas,

plusieurs précautions doivent être prises ; les variables doivent comporter un nombre suffisant de modalités et leur distribution doit être proche d'une distribution normale. A partir de la matrice de corrélations, l'analyse factorielle génère deux nouvelles matrices. La première est la matrice factorielle ; elle exprime les corrélations (appelées saturations) entre les variables manifestes et les variables latentes. Les variables les plus saturées permettent de donner une interprétation psychologique aux facteurs. A partir de cette matrice, on peut recalculer une matrice de corrélations entre les variables (la corrélation entre deux variables est égale à la somme des produits des saturations de ces variables dans les facteurs de l'analyse) ; lorsque les différences entre ces corrélations et les corrélations observées sont petites, la solution factorielle est jugée satisfaisante. La seconde matrice exprime les scores des sujets dans chacun des facteurs (ces scores sont appelés scores factoriels). Les facteurs expriment donc les observations de départ sous forme condensée. Les analyses factorielles visent alors à rendre compte de variables par d'autres variables.

Selon Reuchlin (1964), le modèle général de l'analyse factorielle peut s'exprimer de la manière suivante :

$$s_{ji} = c_{j1}x_{1i} + c_{j2}x_{2i} + \dots + c_{jq}x_{qi},$$

où s_{ji} est la mesure du comportement de l'individu i dans la situation j (e.g. réponse de l'individu i à la question j d'un questionnaire), x_{qi} est le score de l'individu i dans le facteur q et c_{jq} est la saturation de la variable j dans le facteur q .

Comme l'atteste cette équation, le modèle de toute analyse factorielle est linéaire. Deux autres postulats font également partie du modèle : les scores des sujets pour les variables manifestes et pour les variables latentes ont une distribution normale (on les exprime en général sous forme de scores centrés-réduits) et les facteurs sont indépendants (corrélations nulles entre les facteurs).

5.3.2 L'analyse en composantes principales

La corrélation simple entre deux variables est une mesure de la dépendance linéaire, mais l'examen de la matrice de corrélations offre difficilement une vue d'ensemble des relations. De plus, une corrélation entre deux variables peut être

influencée par d'autres variables ; il est alors préférable de calculer la corrélation partielle qui élimine l'effet linéaire des variables restantes.

L'analyse en composantes principales transforme les variables originales en de nouvelles variables non corrélées entre elles. Les nouvelles variables sont appelées composantes principales (facteurs), elles ont la forme d'une combinaison linéaire des variables observées originales. On mesure la quantité d'information contenue dans chaque composante principale par sa variance. Les composantes principales sont hiérarchisées, c'est-à-dire que la première composante exprime la plus grande proportion de la variance totale et la dernière, la proportion la plus faible (la variance totale est la variance du nuage de points formé par les variables dans l'espace des sujets). Autrement dit, la première composante est celle qui permet de reproduire au mieux la matrice de corrélations observée. C'est pourquoi la première composante est souvent assimilée à un facteur général. Le nombre de composantes nécessaires pour reproduire exactement la matrice de corrélations est égal au nombre de variables manifestes. Si l'on désire réduire la dimension du problème sans perdre trop d'information, on ne retient que les premières composantes principales.

L'analyse en composantes principales ne s'applique que sur les données ayant une bonne structure de corrélations. La pertinence d'une analyse en composantes principales peut se vérifier à l'aide du test de sphéricité de Bartlett qui teste si toutes les corrélations sont nulles. On rejette cette hypothèse lorsque la statistique du test est trop grande. On n'applique la méthode d'analyse en composantes principales que si l'on rejette cette hypothèse.

L'analyse en composantes principales recherche des transformations linéaires non corrélées des variables initiales qui expliquent le mieux la diversité des données. Elle effectue donc une rotation orthogonale des axes originaux sur les axes majeurs.

De manière algébrique, l'analyse en composantes principales transforme les p variables originales en p nouvelles variables ; la solution est donnée par l'ensemble des p valeurs propres λ_i et des p vecteurs propres \vec{y}_i de la matrice des corrélations : les vecteurs propres indiquent la direction des axes majeurs et les valeurs propres indiquent la variance dans cette direction.

Ainsi l'analyse en composantes principales décompose la variabilité totale des données initiales en une somme de variabilité expliquée par chacune des composantes, d'après le théorème de la décomposition spectrale.

A partir des vecteurs propres et des valeurs propres, on calcule la matrice des saturations qui est composée des corrélations entre les variables initiales et les composantes principales.

La j -ème composante principale peut s'écrire de la façon suivante à l'aide des saturations :

$$Y_j = \frac{\text{corr}(X_1, Y_j)}{\sqrt{\lambda_j}} X_1 + \frac{\text{corr}(X_2, Y_j)}{\sqrt{\lambda_j}} X_2 + \dots + \frac{\text{corr}(X_p, Y_j)}{\sqrt{\lambda_j}} X_p.$$

Pour interpréter les composantes Y_j , on utilise les variables X_i pour lesquels $|\text{corr}(X_i, Y_j)|$ est grand. Si les corrélations des variables retenues sont de même signe, on dit que la composante mesure un effet de taille ; si les corrélations des variables retenues sont de signes différents, on dit que la composante mesure un effet de forme (opposition entre deux groupes de variables).

Le carré des saturations $\text{corr}^2(X_i, Y_j)$ mesure la proportion de variabilité de X_i expliquée par Y_j ; on peut ainsi déceler les variables mal expliquées par chaque composante.

L'analyse factorielle des correspondances, développée en France par Benzécri et ses collaborateurs (1973), réalise une analyse en composantes principales sur un tableau de contingence.

Remarque :

On confond parfois l'analyse en composantes principales et l'analyse en facteurs communs et uniques. Même si l'analyse en facteurs communs et uniques répond effectivement au même objectif général de condensation d'un ensemble d'observations, elle repose sur une logique différente qu'il est bon de rappeler.

Dans l'analyse en facteurs communs et uniques, les facteurs sont des variables hypothétiques censées expliquer les corrélations observées entre des variables manifestes. Comme ces facteurs ne peuvent rendre compte que de ce qui est commun à plusieurs variables, il subsiste une partie spécifique à chaque variable. Cette partie

spécifique correspond à un facteur unique (il y a autant de facteurs uniques que de variables, ce qui conduit à extraire avec cette méthode plus de facteurs que de variables manifestes). Les analyses en facteurs communs et uniques visent donc à rechercher les sources de variation communes à plusieurs variables.

5.3.3 Le nombre de facteurs

Le choix du nombre de facteurs est un problème crucial en analyse factorielle. Pourtant, ce problème n'admet pas de solution évidente : si le chercheur décide de retenir un nombre important de facteurs, il reproduira correctement la matrice de corrélations entre les variables mais l'économie réalisée sera faible. A l'inverse, en retenant peu de facteurs, il perdra de l'information.

De multiples méthodes formelles ont été proposées pour aider le chercheur dans son choix. Selon Walkey (1983), Vernon aurait déjà recensé vingt-cinq méthodes différentes en 1949, et bien d'autres sont apparues depuis. Pourtant, aucune de ces méthodes ne peut être considérée comme entièrement satisfaisante, car elles reposent toutes en partie sur des critères subjectifs. Parmi les méthodes les plus utilisées, on citera celle de Kaiser (1960) et celle de Cattell (1966b). Le critère de Kaiser repose sur l'hypothèse d'une distribution aléatoire de la variance expliquée par les facteurs, celui de Cattell, sur les différences entre les pourcentages de variance expliquée par les facteurs. Plus précisément,

- 1) La règle de Kaiser ne garde que les composantes avec une valeur propre supérieure à 1.
- 2) La règle de Catell garde les composantes situées avant un saut dans le diagramme des valeurs propres.
- 3) La règle de Joliffe demande que 80% de la variabilité soit expliquée (d'autres auteurs proposent un pourcentage variant de 70% à 90%).

Mais ces critères ne sont qu'indicatifs ; c'est au chercheur qu'il appartient de prendre une décision qui sera d'autant plus satisfaisante qu'elle prendra appui sur des bases théoriques. Pratiquement, l'analyse en composantes principales est inutile si on doit garder plus de 4 ou 5 composantes.

Ayant retenu les r premières composantes, la communalité de la variable X_i , définie par $\sum_{j=1}^r \text{corr}^2(X_i, Y_j)$, mesure la proportion de la variabilité expliquée par les r premières composantes. Les variables avec $\text{corr}^2(X_i, Y_j)$ grand, pour des composantes non retenues, seront mal expliquées après réduction de la dimension.

Il existe deux représentation graphiques liées à l'analyse en composantes principales :

- 1) Représentation des individus par leurs premiers scores factoriels ; c'est la meilleure représentation au niveau de la variance, elle peut être utile pour un test de normalité ou pour la détection de données aberrantes.
- 2) Cercle des corrélations qui représente les variables par leurs coordonnées $\begin{pmatrix} \text{corr}(X_i, Y_1) \\ \text{corr}(X_i, Y_2) \end{pmatrix}$ à l'intérieur d'un cercle de rayon 1. Plus une variable est proche du cercle, plus elle est bien expliquée par les deux premières composantes. Si on retient trois composantes, on représentera les corrélations dans une sphère de rayon 1, etc

5.3.4 Le choix de la solution factorielle et les rotations

A nombre égal de facteurs, il existe une infinité de solutions permettant de reproduire une même matrice de corrélations. Se pose alors le problème du choix de la solution factorielle. Le critère adopté est un critère de simplification. Sous le nom de «structure simple », Thurstone (1947) a proposé de choisir une solution factorielle dans laquelle un nombre maximum de variables a des saturations nulles. Cette méthode permet d'obtenir des solutions plus simples à interpréter, puisqu'on diminue le nombre de variables saturées dans chaque facteur.

Plus généralement, le problème du choix de la structure factorielle est assimilé au problème des rotations. Cette notion renvoie à une interprétation géométrique de l'analyse factorielle dans laquelle les saturations des variables sont des coordonnées

spatiales, les facteurs constituant les axes de coordonnées de ce système. Intuitivement, une rotation consiste à utiliser un autre système de coordonnées que les facteurs initiaux. Après rotation, les saturations factorielles des variables (i.e. les coordonnées) sont différentes mais la qualité de la solution est inchangée : la matrice de corrélations reproduite à partir des nouveaux axes de coordonnées (ie. les nouveaux facteurs) est exactement la même qu'avec la solution initiale.

Plusieurs méthodes de rotation ont été proposées. Une distinction importante oppose les rotations orthogonales aux rotations obliques. Dans une rotation orthogonale, on conserve l'indépendance entre les facteurs ; dans une rotation oblique, on obtient des facteurs corrélés. Dans ce dernier cas, on opère souvent une nouvelle analyse factorielle sur les facteurs corrélés pour identifier les sources de variation communes aux différents facteurs : on parle alors d'analyse factorielle de second ordre. Ici aussi, ce sont les connaissances théoriques sur le domaine étudié qui doivent contribuer à choisir la rotation la plus appropriée.

Les principales rotations orthogonales utilisés sont les rotations

- varimax qui minimise le nombre de coefficients élevés pour chaque facteur,
- quartimax qui minimise le nombre de coefficients élevés pour chaque variable,
- équimax qui est une combinaison des deux rotations précédentes.

D'après Sean Hammond (2000), un point souvent sousestimé est le choix d'un échantillon assez large. Comme l'analyse factorielle est une méthode de décomposition de la variance, on a besoin d'un échantillon dont la taille minimise les erreurs d'attribution. Pour avoir une solution factorielle fiable il est recommandé de prendre un échantillon d'au moins 200 individus si possible. D'une manière générale, il est également recommandé d'avoir au moins quatre fois plus de sujets que de variables.

5.4 L'analyse de la variance

5.4.1 Introduction

Le but de l'analyse de variance est d'identifier et de comparer les sources de variations d'une variable. Cette technique, aux ramifications et développements

multiples, est en principe utilisable chaque fois qu'il s'agit d'évaluer l'effet de variables indépendantes (catégorielles) sur une variable dépendante (quantitative). Historiquement l'analyse de la variance s'est développée de façon relativement autonome dans des contextes essentiellement expérimentaux tels que l'agronomie, la biologie, la recherche pharmaceutique, ce qui se traduit par une terminologie un peu particulière. Il s'agissait à l'origine de mesurer les effets des gènes et des divers facteurs environnementaux sur le rendement de cultures ou d'animaux domestiques.

Ainsi, chaque variable (fonctionnellement) indépendante est appelée facteur et les modalités correspondantes niveaux. Une combinaison des niveaux des facteurs est appelée traitement.

Aujourd'hui, l'analyse de variance est devenue classique dans la plupart des sciences expérimentales et est utilisée pour tester l'efficacité de nouveaux traitements entendus au sens large. Les applications de l'analyse de variance sont donc nombreuses et variées.

A un niveau général, les modèles d'analyse de variance, d'analyse de covariance et de régression linéaire sont des cas particuliers du Modèle Linéaire Généralisé, la structure de dépendance entre variables étant explicitée au moyen d'une matrice de design. Nous allons nous limiter dans cet exposé à présenter brièvement les deux situations les plus simples, les plans d'expérience à un facteur ainsi que les plans d'expérience à deux facteurs.

5.4.2 L'analyse de variance à un facteur (*One-Way ANOVA*)

Le plan d'expérience le plus simple est celui à un facteur. On l'utilise pour étudier l'effet d'un seul facteur sur une variable. Un facteur est un agent susceptible d'influencer la distribution de la variable. Pour détecter et quantifier une telle influence, on fait varier le facteur : on obtient ainsi des traitements différents.

Il s'agit en fait d'une généralisation du test de la moyenne pour plus de deux populations. On dispose de m groupes ($m \geq 3$) de tailles respectives n_1, n_2, \dots, n_m sur lesquels les scores d'une variable quantitative X ont été mesurés. Il s'agit de déterminer si les groupes sont différents, au sens que leurs moyennes, prises globalement (et non comparées deux à deux) diffèrent significativement. L'idée est de

comparer la variance intergroupe (mesurant la dispersion des moyennes de chaque groupe par rapport à la moyenne empirique générale) à la variance intragroupe (mesurant la dispersion moyenne de l'intérieur de chaque groupe) : les groupes apparaîtront comme distincts pour autant que ce rapport soit suffisamment élevé .

On peut justifier rigoureusement ce résultat intuitif : la variance totale $\text{Var}(X)$ (évalué sur l'échantillon complet de $n = n_1 + n_2 + \dots + n_m$ individus) se décompose exactement en deux parties, à savoir

$$\text{Var}(X) = \text{Var}_B(X) + \text{Var}_W(X),$$

où $\text{Var}_B(X)$ est la variance intergroupe (*between groups*) et $\text{Var}_W(X)$ est la variance intragroupe (*within groups*), définies par

$$\text{Var}_B(X) = \frac{1}{n} \sum_{i=1}^n n_i (\bar{X}_i - \bar{X})^2 \quad \text{et} \quad \text{Var}_W(X) = \frac{1}{n} \sum_{i=1}^n n_i s_i^2.$$

Ici \bar{X} désigne la moyenne totale, \bar{X}_i celle du groupe i et s_i^2 dénote la variance du groupe i .

L'hypothèse statistique que l'on tente de rejeter est celle de l'égalité entre les moyennes théoriques des m groupes :

$$H_0 : \mu_1 = \mu_2 = \dots = \mu_m \quad (\text{pas de différence entre les groupes})$$

Pour faire cela, on lui oppose l'hypothèse alternative complémentaire « $H_1 : H_0$ fausse », dans le cadre théorique où $X_i \sim N(\mu_i, \sigma^2)$. Cette hypothèse de travail, qui reste à justifier, postule que les scores X_i du i -ème groupe sont distribués normalement autour d'une moyenne μ_i inconnue et avec une variance σ^2 inconnue, identique pour chaque groupe. Concrètement, une prudence certaine est de mise en face de groupes aux variances empiriques trop disparates.

Si H_0 est vraie, la distribution de la variable de décision d , défini ci-dessous, suit une loi de Fisher-Snedecor :

$$d := \frac{\frac{\text{Var}_B(X)}{m-1}}{\frac{\text{Var}_W(X)}{n-m}} \sim F [m-1, n-m].$$

On rejette donc H_0 (et accepte H_1) au niveau α si

$$d \geq F_{1-\alpha}[m-1, n-m].$$

Remarques :

- 1) $\text{Var}(X)$ est la variance totale, tandis que $\text{Var}_b(X)$ peut être interprété comme la variance expliquée (par l'appartenance à tel ou tel groupe), alors que $\text{Var}_w(X)$, qui mesure la dispersion dans les groupes, est la variance résiduelle. Comme en régression linéaire, on définit le rapport de corrélation par

$$\eta^2 = \frac{\text{variance expliquée}}{\text{variance totale}} = \frac{\text{Var}_b(X)}{\text{Var}(X)}.$$

- 2) Un rapport d significatif (i.e. suffisamment grand pour α et n donnés) indique que les moyennes des populations correspondantes aux groupes ne sont probablement pas toutes égales, mais n'indique pas quels sont les groupes dont les moyennes diffèrent significativement. Pour ce faire, on recourt aux procédures de comparaison multiples (*post hoc multiple comparaison*) ; essentiellement ces procédures ajustent le niveau de signification au contexte des comparaisons multiples, contexte qui rend fallacieux l'usage simple du test de la moyenne entre toutes les paires possibles : avec 5 groupes par exemple, il y a 10 paires de comparaisons et la probabilité (sous H_0) pour que au moins un des niveaux de signification d'un test de la moyenne soit inférieur à $\alpha = 0,05$ est de $\alpha_{\text{réel}} = 0,29$.
- 3) Dans le contexte de l'analyse de la variance, on parle souvent de « Somme de carrés » ($SS = \text{Sum of Squares}$) et de « Carrés moyens » ($MS = \text{Mean Squares}$). Ils sont définis par

$$SST = n \text{Var}(X), \quad SSB = n \text{Var}_b(X), \quad SSW = n \text{Var}_w(X),$$

$$MSB = \frac{SSB}{m-1} \quad \text{et} \quad MSW = \frac{SSW}{n-m}.$$

Alors,

$$SST = SSB + SSW \quad \text{et} \quad d = \frac{MSB}{MSW}.$$

5.4.2 L'analyse de variance à deux facteurs (Two-Way ANOVA)

Il s'agit maintenant d'examiner dans quelle mesure une variable quantitative X (la « réponse ») peut dépendre d'un premier facteur A (variable catégorielle à a modalités), d'un second facteur B (variable catégorielle à b modalités), ainsi que de déterminer, le cas échéant, si ces influences s'additionnent simplement ou non, se combinant de façon spécifique pour produire de nouveaux effets (interaction). Tout ceci est réflété par le modèle

$$X_{jk} = \mu + \alpha_j + \beta_k + \gamma_{jk} + \varepsilon_{jk},$$

où $j = 1, \dots, a$ indice les niveaux du facteur A , $k = 1, \dots, b$ indice les niveaux du facteur B , X_{jk} dénote la variable X restreinte au groupe (ou traitement) jk . Dans ce modèle, μ représente un effet global, commun pour tous les groupes, α_j représente l'effet « de ligne » dû au facteur A , β_k représente l'effet « de colonne » dû au facteur B et γ_{jk} représente l'effet d'interaction entre les facteurs : tous ces termes sont des constantes. ε_{jk} est une variable d'erreur, supposée suivre une distribution normale $N(0, \sigma^2)$, où σ^2 ne dépend pas du groupe : comme dans l'analyse de la variance à un facteur, seules les moyennes des X_{jk} et non leurs variances sont supposées pouvoir varier d'un groupe à l'autre.

On suppose, pour simplifier, que chaque groupe jk est composé du même nombre c d'individus, indicés par $l = 1, \dots, c$ (on parle alors de design équilibré) ; on a alors affaire à un total de $n = abc$ scores x_{jkl} . La moyenne du groupe jk est $\bar{x}_{jk} = \frac{1}{c} \sum_l x_{jkl}$. La moyenne de la ligne j est $\bar{x}_j = \frac{1}{bc} \sum_{kl} x_{jkl}$. La moyenne de la colonne k est $\bar{x}_k = \frac{1}{ac} \sum_{jl} x_{jkl}$. La moyenne totale est $\bar{x} = \frac{1}{abc} \sum_{jkl} x_{jkl}$.

La variance (à un facteur de normalisation près) se décompose comme

$$\begin{aligned} \sum_{jkl} (x_{jkl} - \bar{x})^2 &= bc \sum_j (\bar{x}_j - \bar{x})^2 + ac \sum_k (\bar{x}_k - \bar{x})^2 \\ &\quad + c \sum_{jk} (\bar{x}_{jk} - \bar{x}_j - \bar{x}_k + \bar{x})^2 + \sum_{jkl} (x_{jkl} - \bar{x}_{jk})^2. \end{aligned}$$

ce que l'on abrégie (dans l'ordre) par

$$SST = SS1 + SS2 + SS_{\text{int}} + SSW.$$

Les degrés de liberté associés aux différents termes sont $a - 1$ pour $SS1$, $b - 1$ pour $SS2$, $(a - 1)(b - 1)$ pour SS_{int} et $ab(c - 1)$ pour SSW . On définit les carrés moyens ($MS = \text{Mean Squares}$) comme les sommes de carrés divisés par leur degrés de liberté correspondants :

$$\begin{aligned} MS1 &= \frac{SS1}{a - 1} & MS2 &= \frac{SS2}{b - 1} \\ MS_{\text{int}} &= \frac{SS_{\text{int}}}{(a - 1)(b - 1)} & MSW &= \frac{SSW}{ab(c - 1)}. \end{aligned}$$

Trois hypothèses sont alors à tester, à savoir :

H_0^A : il n'y a pas d'effet dû au facteur A : $\alpha_j = 0, j = 1, \dots, a$.

H_0^B : il n'y a pas d'effet dû au facteur B : $\beta_k = 0, k = 1, \dots, b$.

H_0^{int} : il n'y a pas d'effet d'interaction entre les facteurs A et B : $\gamma_{jk} = 0$, pour tous j, k . Cela signifie que tel niveau du facteur A affecte uniformément les individus, quel que soit le niveau du facteur B (et réciproquement).

Les hypothèses alternatives sont simplement les négations de H_0^A , H_0^B et H_0^{int} . Les variables de décision sont les rapports des carrés moyens :

Sous $H_0^A, d^A = \frac{MS1}{MSW} \sim F [a - 1, ab(c - 1)]$.

Sous $H_0^B, d^B = \frac{MS2}{MSW} \sim F [b - 1, ab(c - 1)]$.

Sous $H_0^{\text{int}}, d^{\text{int}} = \frac{MS_{\text{int}}}{MSW} \sim F [(a - 1)(b - 1), ab(c - 1)]$.

Par exemple, si $d^B > F_{1-\alpha}[b - 1, ab(c - 1)]$, on rejette H_0^B au niveau α .

Contrairement à ce qui peut se passer dans certains modèles loglinéaires hiérarchiques, tous les cas de rejet ou de non-rejet peuvent en principe se produire : il se peut par exemple que H_0^A et H_0^B ne soient pas rejetées alors que H_0^{int} le soit, ou que H_0^A soit rejetée alors que H_0^{int} et H_0^B ne le soient pas, etc

Le diagramme d'interaction consiste à représenter en ordonnée \bar{x}_{jk} , les moyennes par groupe (traitement), les a modalités du premier facteur étant représentées en abscisse et les points correspondant à la même modalité du second facteur étant reliés par un trait (le rôle des facteurs pouvant être échangés). Les figures ci-dessus illustrent 4 situations typiques (avec $a = 3, b = 2$).

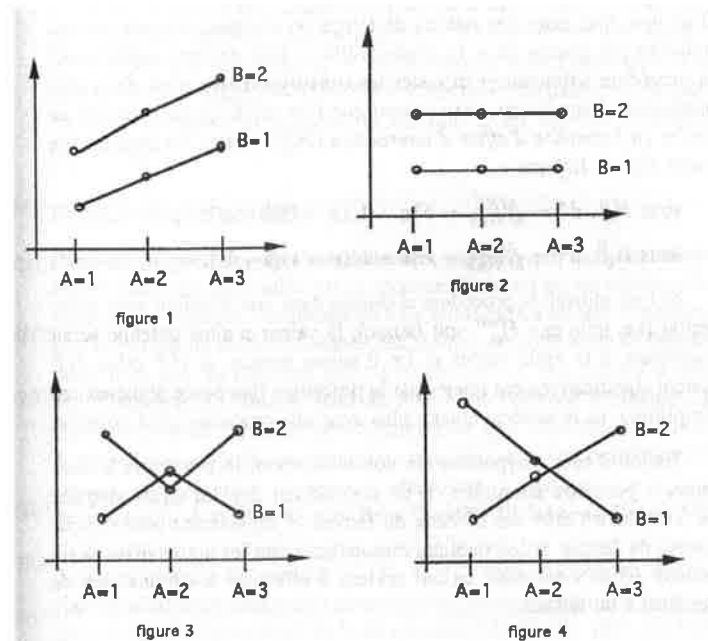


Figure 5.4 : Exemples de diagrammes d'interaction (Bavaud et al., 1996)

Ces diagrammes permettent de visualiser les effets éventuels de chacun des facteurs et de leur interaction :

Figure 1 : effet de A , effet de B , pas d'effet d'interaction.

Figure 2 : pas d'effet de A , effet de B , pas d'effet d'interaction.

Figure 3 : pas d'effet de A (en moyenne), effet de B (en moyenne), effet d'interaction

Figure 4 : pas d'effet de A (en moyenne), pas d'effet de B (en moyenne), effet d'interaction.

On voit qu'un effet d'interaction entraîne le non-parallélisme des segments. La significativité des effets (de A , de B d'interaction) dépend par contre également des effectifs et des variances intragroupe : elle ne peut être uniquement déterminée à partir du diagramme d'interaction.

Il se peut que pour des raisons de temps ou d'argent il y ait un seul individu par groupe ($c = 1$), d'où $MSW = 0/0$ devient indéterminé : la procédure habituelle pour tester les trois hypothèses n'est alors plus applicable. Dans ce cas, pour autant que l'on ait de bonnes raisons de croire en l'absence d'effets d'interaction (H_0^{int} vraie), on peut encore tester H_0^A et H_0^B par

$$\text{Sous } H_0^A, d^A = \frac{MS1}{MS_{\text{int}}} \sim F [a-1, (a-1)(b-1)].$$

$$\text{Sous } H_0^B, d^B = \frac{MS2}{MS_{\text{int}}} \sim F [b-1, (a-1)(b-1)].$$

Si l'on utilisait la procédure ci-dessus dans une situation avec interaction (i.e. telle que H_0^{int} soit fausse), la valeur p ainsi obtenue serait supérieure à la vraie valeur p . En d'autres termes, si H_0^A et/ou H_0^B étaient significativement rejetés par la procédure (supposée abusivement simplifiée), ils le seraient encore plus avec une procédure plus correcte.

Toujours sous l'hypothèse de non-interaction, la procédure à deux facteurs peut être simplifiée en ne considérant que les effets simples, i.e. en testant l'effet des niveaux du facteur A séparément pour chaque niveau du facteur B (en oubliant momentanément les autres niveaux du facteur B) et vice-versa, ce qui revient à effectuer $a + b$ analyses de variance à un facteur.

5.5 Le LISREL

5.5.1 Généralités

La pratique de routine des analyses factorielles est actuellement largement facilitée par les moyens automatiques de calcul qu'offrent les logiciels statistiques. Mais, d'après Dickes (1996), cette facilité a aussi ses inconvénients. Lorsqu'il utilise ces programmes, le chercheur est en quelque sorte contraint d'adopter les modèles tels qu'ils sont proposés et d'accepter les contraintes qui y sont associées. L'utilisation de ces programmes conduit ainsi trop naturellement à une démarche a-théorique que l'on peut qualifier d'exploratoire : le chercheur dispose de données auxquelles il applique un modèle d'analyse fourni par un logiciel et il interprète ensuite les résultats. Les méthodes d'analyses structurales de relations linéaires introduisent une rupture par

rapport à une telle démarche en proposant, par contraste, une approche confirmatoire (Dickes, 1986). Le LISREL (*Linear structural Relationships*), introduit par Jöreskog (1970, 1973) repose sur des systèmes d'équations structurales linéaires et il englobe et généralise les méthodes linéaires classiques : méthodes de régressions, analyses en pistes causales, analyses canoniques, analyses factorielles. D'après Dickes et al. (1994), le LISREL a changé la démarche des chercheurs en sciences sociales, car il exige une explication a priori des liens entre la théorie et les données : on pose d'abord les hypothèses, puis on regarde si elles sont infirmées ou confirmées par le modèle. De plus, il présente la possibilité d'exprimer sous forme de modèles les représentations théoriques les plus subtiles et notamment les représentations factorielles. En fait cette méthode ne se contente même pas d'intégrer les modèles existants : elle permet la mise à l'épreuve de bien d'autres modèles dont le nombre est seulement limité par l'imagination du chercheur (Dickes, 1986).

La logique fondamentale du LISREL est de mettre à l'épreuve, à partir du lien observé entre des variables empiriques (exprimé le plus souvent en termes de variances-covariances), des hypothèses inférées à partir d'une théorie concernant les relations de dépendance et/ou d'interdépendance entre des variables observées et/ou des variables latentes, par l'intermédiaire de la manipulation de paramètres.

Soit Σ la matrice de variances-covariances entre les variables observées et Σ' la matrice des variances-covariances théorique. Le point de départ consiste à élaborer une représentation théorique concernant les données. Une telle représentation théorique est susceptible d'être concrétisée par le choix des paramètres. D'après Dickes et al (1994), toute la richesse du modèle d'équations structurales est contenue dans la variété des paramètres et la flexibilité de leur gestion. En fonction de la représentation théorique et en fonction de la matrice de variances-covariances observée on peut calculer une matrice de variances-covariances théorique. Cette matrice est obtenue par des procédures itératives (en particulier par un algorithme de maximum de vraisemblance) qui permettent de rendre minimum les différences entre Σ et Σ' . Si l'adéquation entre les deux matrices est satisfaisante, on estime que la représentation théorique sied aux données. En revanche, si l'adéquation est insatisfaisante, on considère que la représentation théorique n'est pas adéquate aux données, ce qui conduit soit à la rejeter, soit à la modifier.

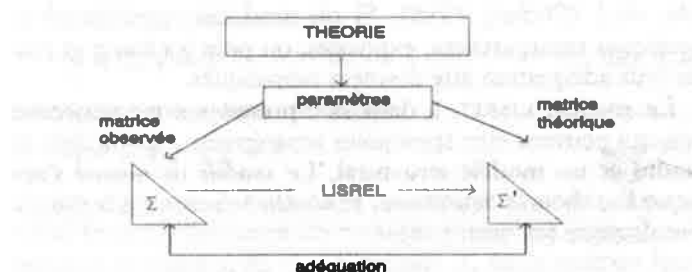


Figure 5.5 : Représentation générale du LISREL (Dickes et al, 1994)

Les données sont constituées par la matrice de variances-covariances calculée entre les variables observées. Sa diagonale contient les variances des variables. Une matrice de variances-covariances a la même signification qu'une matrice de corrélations : elle exprime les relations linéaires entre les variables. Dans le modèle LISREL, on préfère cependant la matrice de variances-covariances à la matrice de corrélations, car on a ainsi la possibilité d'évaluer des hypothèses sur les variances.

Le fait d'utiliser des covariances met bien en évidence qu'il s'agit de modèles linéaires. Ceci suppose que les variables d'origine possèdent au moins les propriétés des échelles d'intervalles et que la relation entre deux variables quelconques est linéaire. Le LISREL impose donc des contraintes fortes aux variables d'origine. Ces contraintes peuvent cependant être relâchées dans la pratique par divers procédés.

La propriété la plus importante du LISREL est la possibilité qu'il donne au chercheur de tester des théories. La démarche est confirmatoire, en opposition à la démarche exploratoire généralement appliquée dans les analyses linéaires multivariées (Dickes, 1986). Si l'on rend ces représentations théoriques transparentes, explicites, on peut les tester et évaluer leur adéquation aux données empiriques. Ceci est réalisé en libérant la métrique, en manipulant les paramètres et en multipliant les contrôles d'adéquation (Bacher et Dickes, 1994).

5.5.2 Le modèle mathématique

Le modèle LISREL comporte deux composantes interconnectées mais qui peuvent être appliquées séparément : un modèle de mesure et un modèle structural. Ces

modèles établissent le lien entre les quatre types de variables représentés dans le tableau ci-dessous.

	Variable indépendante	Variable dépendante
Variable observée	x	Y
Variable latente	ξ	η

Les variables latentes sont inférées à partir des variables observées. Elles correspondent aux concepts théoriques. Les variables observées sont celles qui sont produites à partir du dispositif d'observation ou d'expérimentation ; elles sont supposées centrées, c'est-à-dire de moyenne nulle. Les variables indépendantes ne sont influencées par aucune autre. On les appelle aussi variables exogènes. Les variables dépendantes sont celles qui sont influencées par des variables indépendantes ou par d'autres variables dépendantes. On parle également de variables endogènes. Le modèle complet du LISREL est exprimé, sous forme matricielle, par les trois équations linéaires suivantes :

Modèle de mesure pour x :

$$x = \Lambda_x \xi + \delta.$$

Modèle de mesure pour y :

$$y = \Lambda_y \eta + \varepsilon.$$

Modèle d'équation structurale :

$$\eta = B \eta + \Gamma \xi + \zeta.$$

Les deux premières équations constituent le modèle de mesure et établissent le lien entre les variables observées (x ou y) et des variables latentes (ξ ou η). La matrice Λ_x est la matrice factorielle des variables latentes indépendantes (elle exprime les relations entre les variables observées et les variables latentes) et la matrice Λ_y est la matrice factorielle des variables latentes dépendantes. Les matrices ε et δ sont les résidus, qui sont composés de deux parties : les erreurs de mesure des variables x et y ainsi que les parties spécifiques de ces variables non partagées avec les facteurs communs.

La troisième équation constitue le modèle structural. Elle met en relation les variables latentes endogènes et exogènes ; elle sert à tester des hypothèses concernant des liens de dépendance entre les variables latentes. Le modèle est suffisamment souple pour tester des liens de dépendance entre plusieurs variables dépendantes à la fois. Les éléments de la matrice B représentent les effets directs des variables latentes η sur d'autres variables latentes η et les éléments de la matrice Γ représentent les effets directs des variables latentes ξ sur les variables latentes η . La matrice ζ est une matrice de résidus.

Ces trois équations constituent le modèle structural complet. Mais, en général, la première équation suffit pour tester un modèle psychométrique. En fait, le chercheur a la possibilité d'exprimer ses hypothèses de recherche en intervenant sur huit matrices différentes dont tous les éléments constituent les paramètres du modèle. Dans les applications, certains des éléments des matrices sont fixés (des valeurs préétablies sont assignées aux paramètres), ou contraints (les paramètres contraints sont inconnus mais égaux à un ou à plusieurs autres paramètres), ou libres (les paramètres libres ne sont ni connus, ni égaux à d'autres). Les paramètres contraints sont fixés par le modèle et les paramètres libres sont estimés par une fonction d'estimation.

A partir des différents paramètres, fixés par le chercheur et estimés par le modèle, on peut calculer une matrice de variances-covariances théorique. La différence entre les matrices de variances-covariances théorique et observée est appelée matrice des résidus. On dispose d'un certain nombre d'indicateurs statistiques qui permettent d'évaluer l'importance des résidus. Si ceux-ci sont négligeables, la représentation théorique est considérée comme s'appliquant aux observations.

5.5.3 Les représentations multifactorielles

Le LISREL sert en particulier à tester le modèle factoriel, plus précisément des hypothèses théoriques concernant le nombre de facteurs et leur structure. Il s'agit d'une exploitation confirmatoire qui exige du chercheur des connaissances a priori sur ses données. Cette approche suppose que les variables observées puissent être classées en termes de contenu, de processus théoriques mis en oeuvre, ou encore de conditions expérimentales.

Dans ce modèle, les facteurs sont considérés comme les variables latentes qui mesurent le(s) concept(s) psychologique(s) sans erreur. D'après Dickes et al. (1994), l'idée de base de l'analyse factorielle peut être formulée dans les termes du modèle LISREL de la manière suivante : pour un ensemble de variables observées x_1, x_2, \dots, x_q , on désire trouver un ensemble de variables latentes (facteurs) $\xi_1, \xi_2, \dots, \xi_n$ dont le nombre soit inférieur au nombre de variables observées. Les facteurs sont censés rendre compte des intercorrélations (interdépendances) existant entre les variables observées. Si on retire l'influence des facteurs, on suppose qu'il n'existe plus aucune corrélation entre les variables.

Supposons donc que les variables observées x dépendent de plusieurs variables latentes ξ , que les erreurs de mesure δ ne soient pas corrélées avec les variables latentes ξ , que les coefficients qui expriment les effets des variables latentes sur les variables observées soient représentés par la matrice Λ_x et que la moyenne des erreurs δ soit égale à zéro.

Si l'on a par exemple huit variables observées x_1, x_2, \dots, x_8 qui forment deux facteurs communs ξ_1 et ξ_2 , le modèle de mesure peut être écrit sous forme matricielle de la façon suivante :

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \\ x_8 \end{pmatrix} = \begin{pmatrix} \lambda_{11} & 0 \\ \lambda_{21} & 0 \\ \lambda_{31} & 0 \\ \lambda_{41} & 0 \\ 0 & \lambda_{52} \\ 0 & \lambda_{62} \\ 0 & \lambda_{72} \\ 0 & \lambda_{82} \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} + \begin{pmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \\ \delta_5 \\ \delta_6 \\ \delta_7 \\ \delta_8 \end{pmatrix}.$$

Cette équation matricielle correspond à la représentation ci-dessous :

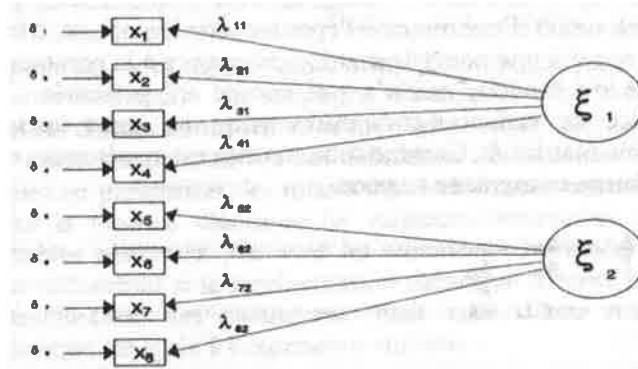


Figure 5.6 : Représentation d'un modèle factoriel avec deux facteurs orthogonaux (Dickes et al., 1994)

Dans le modèle, la représentation théorique à tester peut être paramétrisée ainsi :

- Le premier choix opéré concerne le nombre de facteurs. On suppose que deux variables latentes permettent de rendre compte des interdépendances observées.
- Le second choix consiste à assigner des paramètres fixes et libres aux éléments de la matrice factorielle Λ_x . On suppose que les quatre premières variables observées x_1 à x_4 définissent le premier trait latent et les quatre variables observées suivantes x_5 à x_8 le second. On veut obtenir des estimations $\lambda_{11}, \dots, \lambda_{41}, \lambda_{52}, \dots, \lambda_{82}$ de ces éléments : on les déclare libres. On suppose que les variables x_5 à x_8 ne servent pas à définir le premier trait latent. Ceci conduit à fixer les éléments $\lambda_{51}, \dots, \lambda_{81}$ comme égaux à zéro. De même, on suppose que les éléments $\lambda_{12}, \dots, \lambda_{42}$ sont égaux à zéro.
- Le troisième choix concerne les interrelations entre les variables latentes. On peut supposer que les variables sont indépendantes les unes des autres : il s'agit là d'une représentation de deux facteurs indépendants (ou orthogonaux). Pour mettre à l'épreuve cette hypothèse, il faut faire appel à une nouvelle matrice, nécessaire à la paramétrisation du modèle, qui n'a pas encore été présentée : la matrice des variances-covariances entre les traits latents, appelée matrice Φ . Comme cette matrice est symétrique, elle a la forme triangulaire suivante :

$$\begin{pmatrix} \phi_{11} & \phi_{21} \\ \phi_{21} & \phi_{22} \end{pmatrix}.$$

La diagonale contient les variances des variables latentes et l'élément ϕ_{21} exprime la covariance entre les deux facteurs. Comme on suppose qu'il n'y a pas de corrélation entre les facteurs, cet élément est fixé à zéro. En revanche, si on

supposait que les variables latentes étaient interdépendantes, l'élément ϕ_{21} serait laissé libre.

- Enfin, une dernière matrice nécessite d'être définie : la matrice de variances-covariances entre les erreurs de mesure δ . Cette matrice, désignée par Θ est également symétrique ; elle a donc la forme suivante :

$$\begin{pmatrix} \theta_{11} & \theta_{21} & \theta_{31} & \theta_{41} & \theta_{51} & \theta_{61} & \theta_{71} & \theta_{81} \\ \theta_{21} & \theta_{22} & \theta_{32} & \theta_{42} & \theta_{52} & \theta_{62} & \theta_{72} & \theta_{82} \\ \theta_{31} & \theta_{32} & \theta_{33} & \theta_{43} & \theta_{53} & \theta_{63} & \theta_{73} & \theta_{83} \\ \theta_{41} & \theta_{42} & \theta_{43} & \theta_{44} & \theta_{54} & \theta_{64} & \theta_{74} & \theta_{84} \\ \theta_{51} & \theta_{52} & \theta_{53} & \theta_{54} & \theta_{55} & \theta_{65} & \theta_{75} & \theta_{85} \\ \theta_{61} & \theta_{62} & \theta_{63} & \theta_{64} & \theta_{65} & \theta_{66} & \theta_{76} & \theta_{86} \\ \theta_{71} & \theta_{72} & \theta_{73} & \theta_{74} & \theta_{75} & \theta_{76} & \theta_{77} & \theta_{87} \\ \theta_{81} & \theta_{82} & \theta_{83} & \theta_{84} & \theta_{85} & \theta_{86} & \theta_{87} & \theta_{88} \end{pmatrix}.$$

Sur la diagonale, on trouve la variance des erreurs. Les éléments non diagonaux représentent les covariances entre les erreurs. Dans le modèle classique, on considère que ces éléments sont égaux à zéro, c'est-à-dire que les erreurs de mesure sont aléatoires et non corrélées. Les possibilités de paramétrisation de cette matrice constituent un avantage du LISREL sur les méthodes factorielles traditionnelles, car on peut mettre à l'épreuve des hypothèses spécifiques concernant l'interdépendance des erreurs de mesure. Pour simplifier, on suppose cependant ici que ces éléments sont fixés à zéro.

Après cette traduction des définitions des concepts théoriques en paramètres, les quatre matrices permettent de calculer la matrice théorique de variances-covariances entre variables observées. Ce sont les variances-covariances que l'on obtiendrait si la représentation théorique était vraie. Cette matrice théorique est égale à l'expression suivante :

$$\Sigma' = \Lambda_x \Phi \Lambda_x^r + \Theta_\delta,$$

Λ_x étant la matrice des λ_x et Λ_x^r sa transposée.

Les paramètres libres sont estimés par une fonction de maximum de vraisemblance. Les matrices de variances-covariances théorique et observée sont alors comparées. Si les résidus sont négligeables, la représentation théorique est adéquate aux données. Sinon elle est rejetée ou modifiée.

