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

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 chomskyienne 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 chomskyienne 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-kinesthérique, 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 & Szerminska, 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 prédispose 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 ambigüe, 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 mnemonique 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 MATHEMATIQUES

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 étudiant 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 piagétiennes 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 récipients 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 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 proportionalité 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 perpective 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 corrèlent avec les test 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 6 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 corrèlent 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'hiérarchie, 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 impliqué 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 cité (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 corréler 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 untiaire, 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 réflètent 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 entrainement spécifique. La question de la possibilité d'entrainement des processus cognitifs impliqués dans la résolution de problèmes mathématiques et de l'existence de limites hériditaires 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 semantiques 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 disctincts 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 alt., 1992). Dans une de ces études (Geary & Burlingham-Dubree, 1989), les chercheurs ont administré les *Wechsler Preschool and Primary Scale of Intellignece* (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.

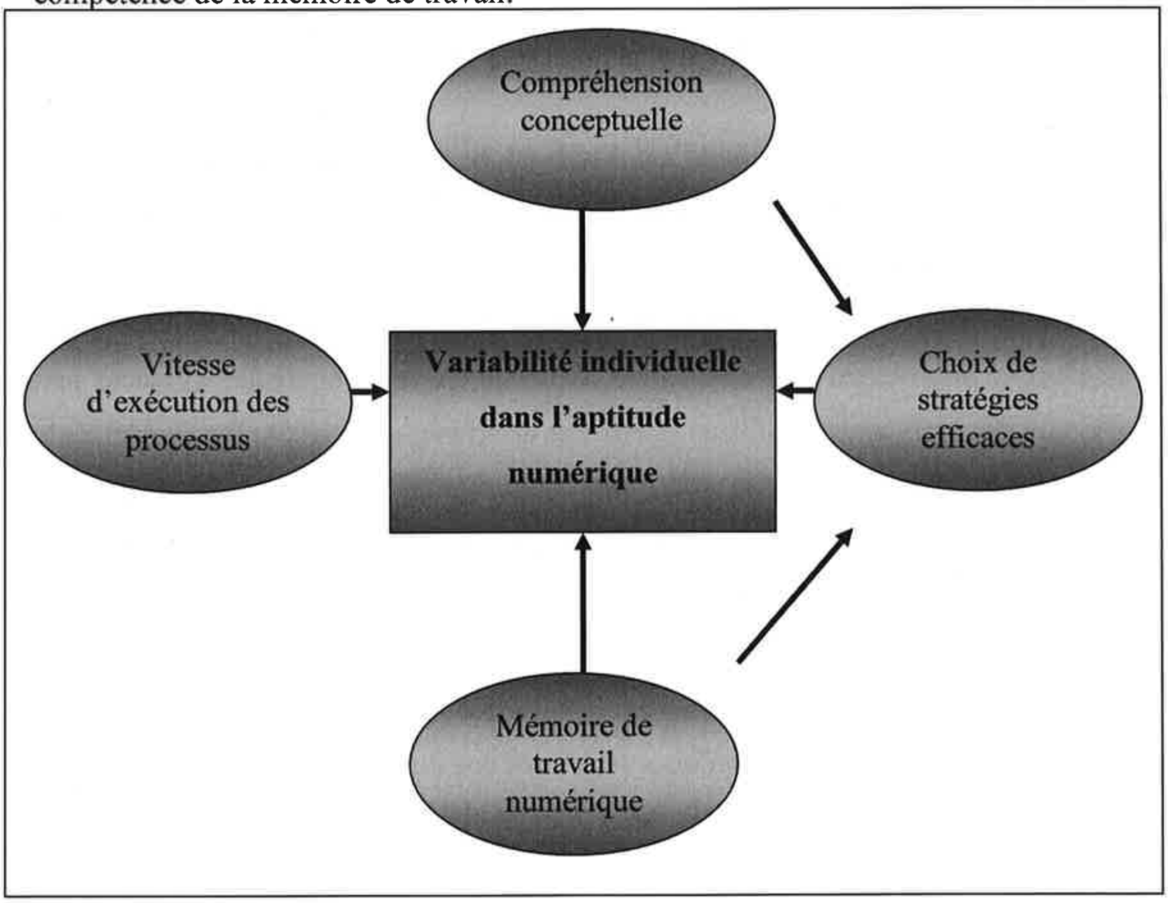


Figure 3.1 : Facteurs influançant l'aptitude numérique chez les jeunes enfants

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 opération 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 recherche 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 cours 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 conclusion 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.

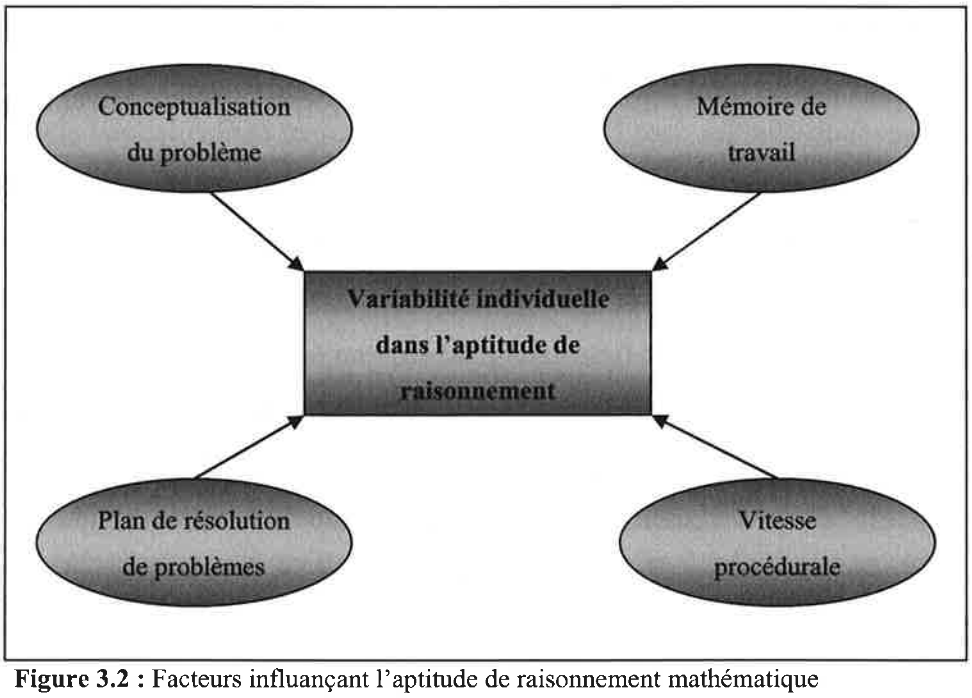


Figure 3.2 : Facteurs influançant l'aptitude de raisonnement mathématique

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, souplement 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 essayent 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 retenu 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.

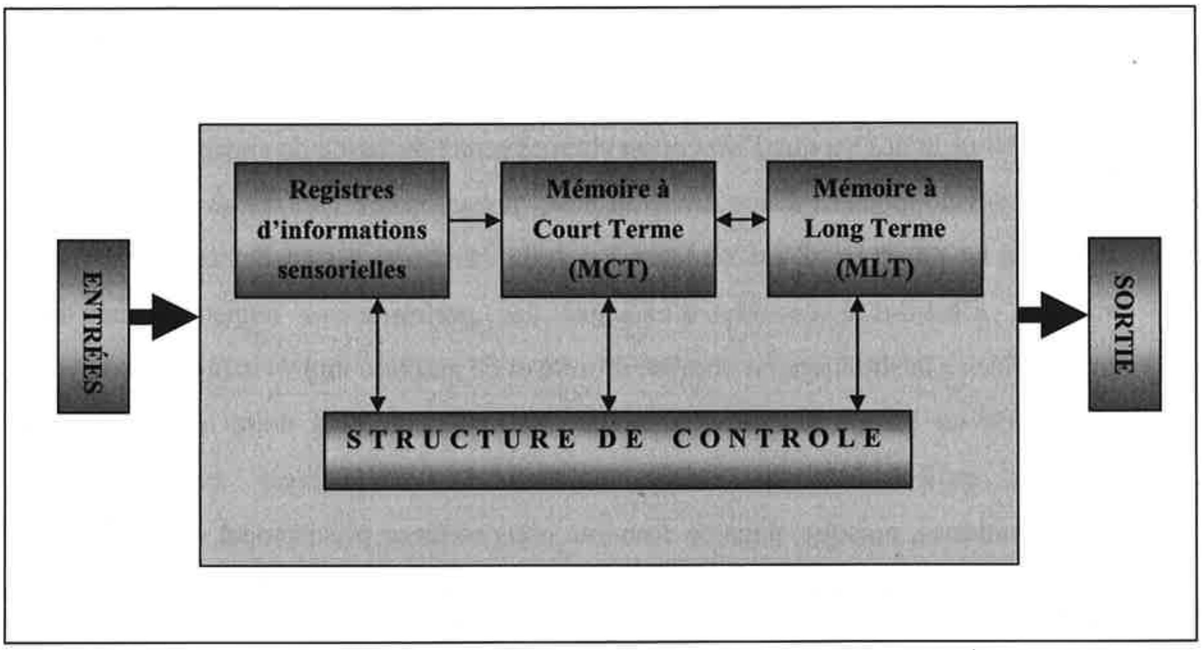


Figure 3.3 : Schéma de l'architecture cognitive proposé en 1969 par Atkinson et Shiffrin. (D'après Lemaire, 1999)

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 classifier 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 présupposée 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 1effectuées à partir de déclarations explicites
- être capable de prédire le développement ultérieur de tendences 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 classifier 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. Pellerey

M. Pellerey (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 abilités 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 iinférieur.

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 PERCEEPTION 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 (*multidimensioal scaling*) est de représenter des données de proximité sous formes de distances, dans un espace d'un nombre restreint de dimensions (Dickes et alt., 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

Tableau 1.1 : Distances kilométriques entre 21 villes françaises

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

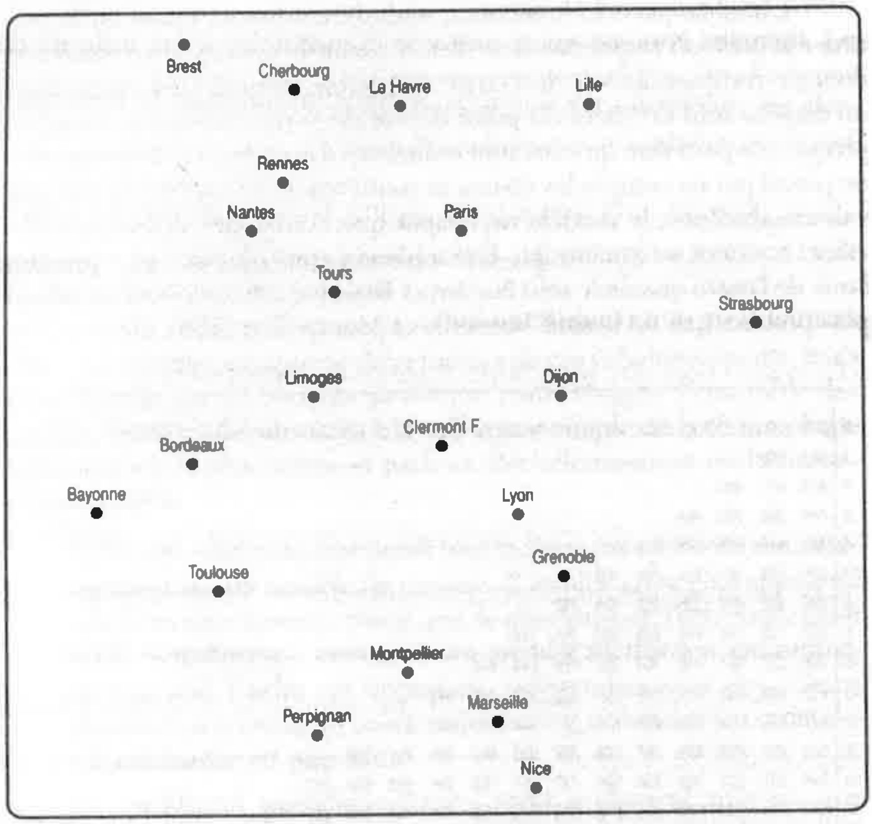


Figure 1.1 : **Configuration bidimensionnelle issue de l'EMD des distances routières entre 21 villes françaises.**

Figure 5.2 : Configuration bidimensionnelle des distances routières (Tournois 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 entrainer 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 pourrons 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 multidimesionel 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 allégement 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 15items, 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 indépendemment aux développementx 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 spetites 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 repectivement *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 réflection 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 soumis 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, qu'elle 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 caracteristiques $X_1,\ldots,X_p$, 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 alt., 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 asendante 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 asendante se présente comme un ensemble de recettes empiriques, dont seule une petite partie est adaptée à un type spécifié 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 asendante 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 étants 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 $(x_{i1}, \ldots, x_{ip})$ de i et $(x_{j1}, \ldots, x_{jp})$ 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_{il}^{s} et x_{jl}^{s} standardisés : cela revient à déclarer chacune des p variables également importantes dans la classification, plutôt que de priviligier 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-occurence d'un attribut de la co-absence de l'attribut, la première étant souvent plus significative que la seconde (la co-occurence 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-occurences 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.

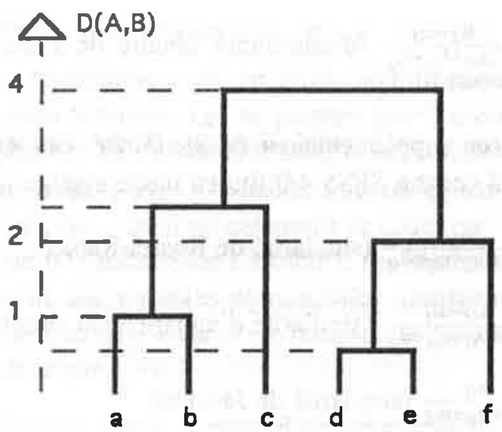


Figure 5.3 : Exemple de dendogramme (Bavaud et al., 1996)

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 *beween-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

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

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

• Enfin, la méthode centroide (*centroid clustering)* consiste à prendre simplement

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

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 dendogramme 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 corrèlent 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 rappellons 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 confimatoire 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 généralisé 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 :

$$s_{ji} = c_{j1}x_{li} + c_{j2}x_{l2} + \ldots + c_{jq}x_{qi},$$

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 (corrélation nulle 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 géneral. 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 variablilité 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 :

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

Pour interpréter les composantes Y_j, on utilise les variables X_i pour lesquels $|corr(X_i, Y_j)|$ 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 $corr^2(X_i, Y_j)$ mesure la proportion de variabilité de X_i expliquée par Y_j ; on peut ainsi décéler 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 theoriques. 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 $\sum_{j=1}^{r} corr^2(X_i, Y_j)$, mesure la proportion de la variabilité expliquée par les r premières composantes. Les variables avec $corr^2(X_i, Y_j)$ 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 $\begin{pmatrix} corr(X_i, Y_1) \\ corr(X_i, Y_2) \end{pmatrix}$ 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 interpreter, 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 $n_1, n_2, \ldots, n_m$ 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 $\mathrm{Var}(X)$ (évalué sur l'échantillon complet de $n = n_1 + n_2 + \cdots + n_m$ individus) se décompose exactement en deux parties, à savoir

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

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

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

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 :

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

Pour faire cela, on lui oppose l'hypothèse alternative complémentaire « $H_1 : H_0$ fausse », dans le cadre théorique où $X_i \sim \mathsf{N}\ (\mu_i, \sigma^2)$. 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 :

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

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

$$d \geq \mathsf{F}_{1-\alpha}[m-1, n-m].$$

Remarques :

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

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

2) Un rapport d significatif (i.e. suffisemment 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 = *Sum of Squares*) et de « Carrés moyens » (MS = *Mean Squares*). Ils sont définis par

$$SST = n\,\mathrm{Var}(X), \quad SSB = n\,\mathrm{Var}_B(X), \quad SSW = n\,\mathrm{Var}_W(X),$$

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

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 $j = 1,\ldots,a$ indice les niveaux du facteur A, $k = 1,\ldots,b$ 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 $\mathsf{N}\ (0,\sigma^2)$, 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 $l = 1,\ldots,c$ (on parle alors de design balancé) ; on a alors affaire à un total de $n = abc$ scores x_{jkl}. La moyenne du groupe jk est $\bar{x}_{jk} = \frac{1}{bc}\sum_l x_{jkl}$. La moyenne de la ligne j est $\bar{x}_{j\cdot} = \frac{1}{bc}\sum_{kl} x_{jkl}$. La moyenne de la colonne k est $\bar{x}_{\cdot k} = \frac{1}{ac}\sum_{jl} x_{jkl}$. La moyenne totale est $\bar{x} = \frac{1}{abc}\sum_{jkl} x_{jkl}$.

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

$$\sum_{jkl}(x_{jkl} - \bar{x})^2 = bc\sum_j(\bar{x}_{j\cdot} - \bar{x})^2 + ac\sum_k(\bar{x}_{\cdot k} - \bar{x})^2$$
$$+c\sum_{jk}(\bar{x}_{jk} - \bar{x}_{j\cdot} - \bar{x}_{\cdot k} + \bar{x})^2 + \sum_{jkl}(x_{jkl} - \bar{x}_{jk})^2.$$

ce que l'on abrévie (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 $SS\,\text{int}$ et $ab(c-1)$ pour SSW. On définit les carrés moyens (MS = *Mean Squares*) comme les sommes de carrés divisés par leur degrés de liberté correspondants :

$$MS1 = \frac{SS1}{a-1} \qquad MS2 = \frac{SS2}{b-1}$$
$$MS\,\text{int} = \frac{SS\,\text{int}}{(a-1)(b-1)} \qquad MSW = \frac{SSW}{ab(c-1)}.$$

Trois hypothèses sont alors à tester, à savoir :

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

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

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

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

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

Par exemple, si $d^B > \mathsf{F}_{1-\alpha}[b-1, ab(c-1)]$, 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 $a = 3, b = 2$).

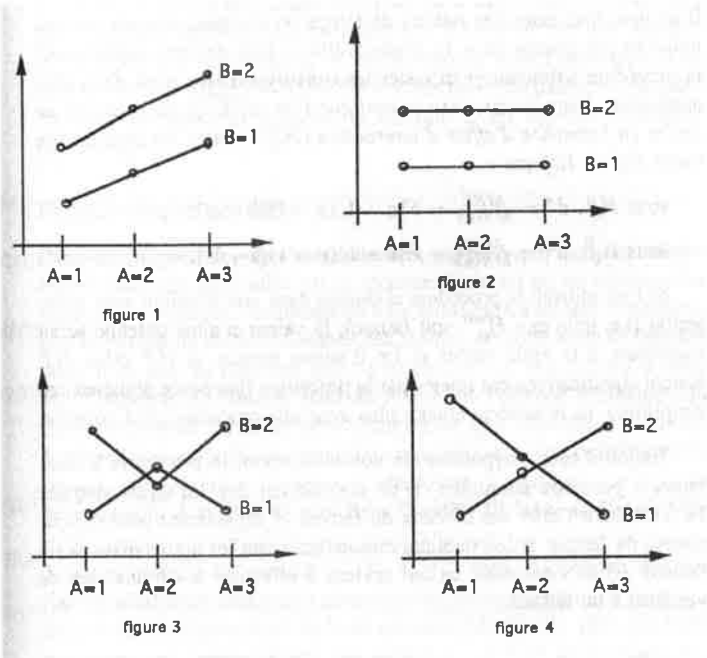


Figure 5.4 : Exemples de diagrammes d'interaction (Bavaud et al., 1996)

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

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

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

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.

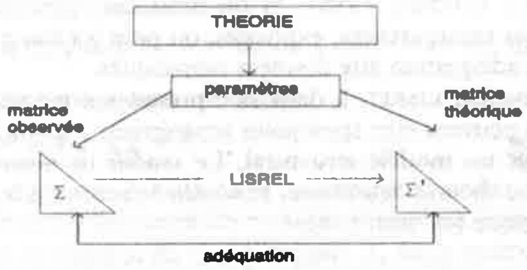


Figure 5.5 : Représentation générale du LISREL (Dickes et al, 1994)

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 $x_1, x_2, \ldots, x_q$, on désire trouver un ensemble de variables latentes (facteurs) $\xi_1, \xi_2, \ldots, \xi_n$ 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 $x_1, x_2, \ldots, x_8$ 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 :

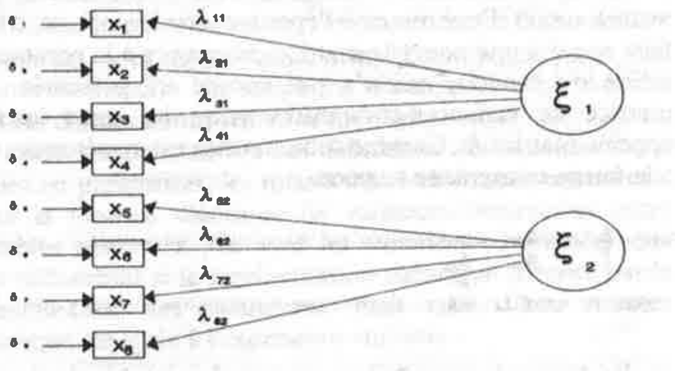


Figure 5.6 : Représentation d'un modèle factoriel avec deux facteurs orthogonaux (Dickes et al., 1994)

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 $\lambda_{11},...,\lambda_{41},\lambda_{52},...,\lambda_{82}$ 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 $\lambda_{51},...,\lambda_{81}$ comme égaux à zéro. De même, on suppose que les éléments $\lambda_{12},...,\lambda_{42}$ 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 traductions 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 :

$$\Sigma' = \Lambda_x \Phi \Lambda_x^{\tau} + \Theta_\delta,$$

Λ_x étant la matrice des λ_x et Λ_x^{τ} 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.

Chapitre VI

ANALYSE DES TAXONOMIES EXISTANTES

6.1 Démarche méthodologique

Dans un premier temps, nous allons effectuer une représentation spatiale de nos propres données, en nous servant d'un modèle d'échelonnement multidimensionnel (*multidimensional scaling*) dont l'objectif consiste à représenter des données de proximité sous forme de distances, dans un espace euclidien possédant un nombre restreint de dimensions ; cette procédure nous permettra d'avoir une première approche des possibilités de regroupement psychologiquement signifiants en 2,3 ou 4 dimensions.

Nous positionnons ensuite nos items par rapport aux taxonomies existantes, en les classant dans les catégories de ces taxonomies, d'après les définitions données par les auteurs. Nous présentons les graphiques montrant la configurations spatiale du regroupement des items d'après les différentes catégories, ce qui nous donne un premier aperçu sur la capacité des différentes taxonomies de réaliser une classification en dimensions séparables, correspondant à nos données.

Il faut cependant remarquer qu'il est difficile de juger de la composition factorielle d'un test seulement en analysant les scores obtenus par une certaine population. Le problème est que, compte tenu de son procédé de fabrication, de ses présupposés théoriques et d'autres aspects, un test mesure souvent plus d'un facteur, respectivement un facteur différent de celui qu'il est censé mesurer. Des tests qui portent sur des séries de nombres, par exemple, peuvent mettre l'accent sur des raisonnements inductifs, des raisonnements déductifs ou bien des raisonnements quantitatifs, selon le genre de problèmes qu'ils contiennent. De même, des tests portant sur les analogies verbales sont supposés mesurer des raisonnements inductifs, seulement si les analogies présentées correspondant à des problèmes relativement difficiles ; sinon ils pourraient mesurer seulement l'étendue du vocabulaire. L'analyse factorielle est seulement capable de classer et d'identifier des facteurs séparés, si la batterie de test contient un échantillon adéquat d'items mesurant chaque facteur. Dans notre cas, les épreuves standardisées qu'on analyse ont été construites surtout d'après des critères de conformité au programme scolaire. Cela implique qu'elles ne contiennent pas forcément assez d'items de chaque type pour pouvoir identifier tous les facteurs influançant la compétence mathématique des élèves se situant dans la tranche d'âge de 12 ans.

6.2 L'échelonnement multidimensionnel de nos données

Nous avons utilisé le logiciel SPSS pour faire un échelonnement multidimensionnel des similarités entre les items constituant la banque d'items de notre logiciel et provenant des données de l'épreuve standardisée en mathématique de novembre 1996, faisant partie de la nouvelle procédure de passage entre le primaire et le post-primaire au Luxembourg. Nous avons effectué un échelonnement en 2, 3 et 4 dimensions. Le logiciel SPSS nous donne deux indicateurs d'adéquations du modèle, le stress et le RSQ.

- Le stress est calculé d'après la formule de Kruskal (1964a, 1964b). Si le stress est inférieur à 0,2 les données peuvent être adéquatement représentées par le modèle.
- Le RSQ représente la proportion de la variance des dissimilarités expliquée par les distances relatives des points dans l'espace euclidien. Un RSQ proche de 1 est donc favorable.

En deux dimensions, on trouve un stress de 0,13363 et un RSQ de 0,94611. La représentation des items dans l'espace euclidien à deux dimensions est la suivante :

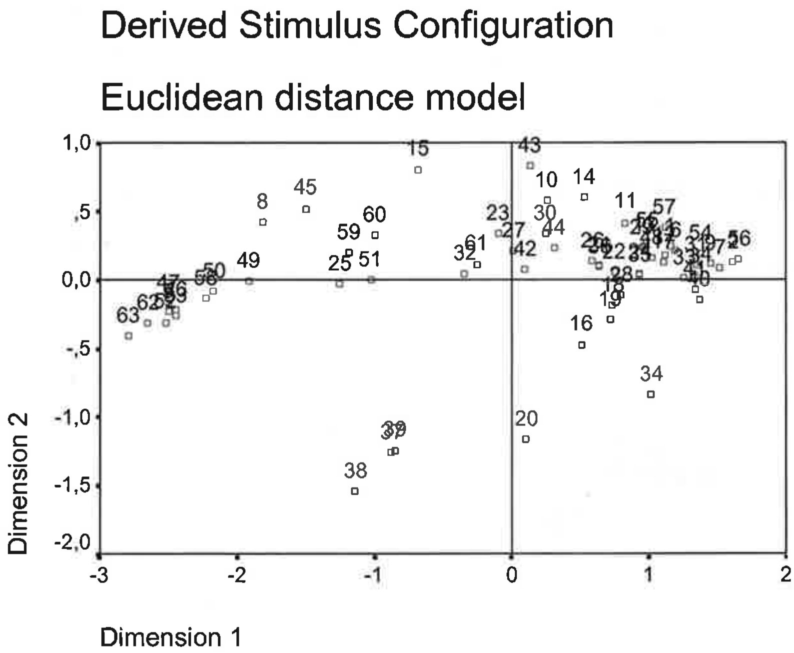


Figure 6.1 : Représentation des items en deux dimensions.

En trois dimensions, on trouve un stress de 0,10915 et un RSQ de 0,95979. La représentation des items dans l'espace euclidien à trois dimensions est la suivante :

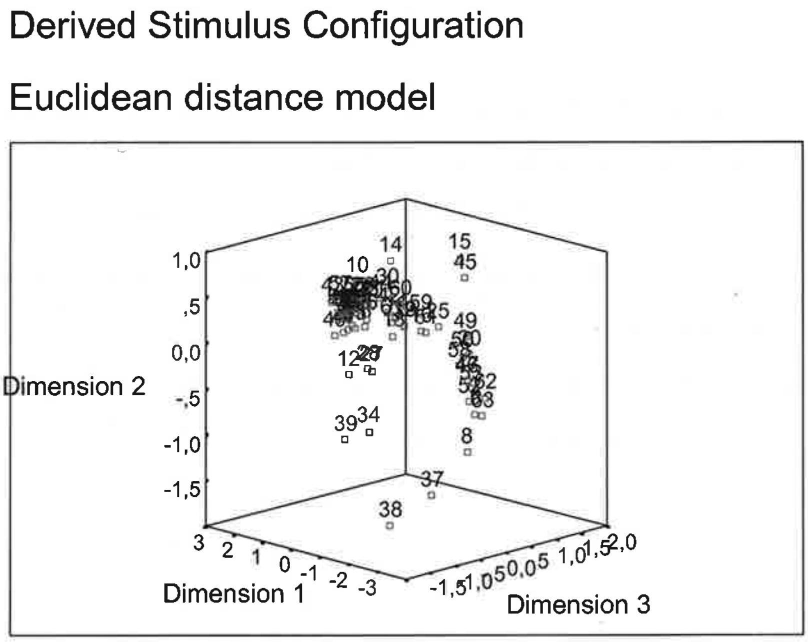


Figure 6.2 : Représentation des items en trois dimensions.

En quatre dimensions, le stress est de 0,09144 et le RSQ de 0,96814. Puisqu'il n'est pas possible de faire une représentation graphique en quatre dimensions de façon satisfaisante, nous donnons ici la représentation des items dans les trois premières dimensions de l'espace euclidien à quatre dimensions.

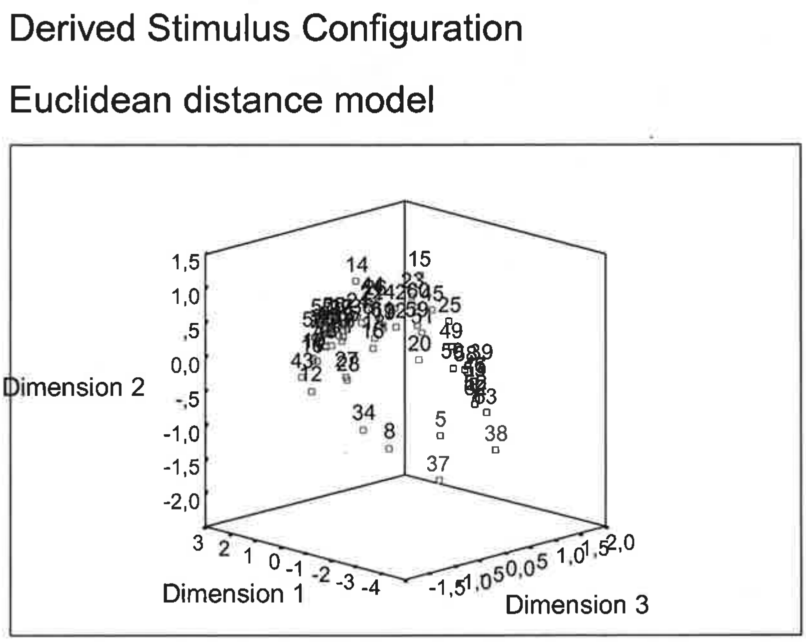


Figure 6.3 : Représentation des items dans les 3 premières de 4 dimensions.

Les trois solutions donnent donc un RSQ proche de 1, ce qui veut dire que la variance des similarités est presque complètement expliquée.

Pour les solutions à 2 et 3 dimensions, le stress indique un niveau d'adéquation satisfaisant (< 0,2) du modèle, pour la solution à 4 dimensions, l'adéquation est même bonne (< 0,1).

Les 3 modèles sont donc adaptés à la situation. Nous allons utiliser dans la suite de ce chapitre un modèle d'échelonnement à deux dimensions. En effet, les deux autres modèles impliquent des représentations plus difficiles à saisir, sans pour autant améliorer ni la précision, ni l'interprétabilité des résultats. Nous présentons ci-dessous les coordonnées des différents items dans l'espace euclidien à deux dimensions.

Item	1.coordonnée	2. coordonnée
1	1,1582	0,2396
2	1,6168	0,1305
3	1,3397	0,0205
4	1,4177	0,0323
5	-2,5006	-0,2283
6	1,1949	0,2115
7	1,5212	0,0901
8	-1,8101	0,4191
9	1,4547	0,1161
10	0,2632	0,5767
11	0,8303	0,4173
12	1,1175	0,1796
13	0,9669	0,2487
14	0,5310	0,6027
15	-0,6826	0,8044
16	0,5129	-0,4748
17	1,1073	0,1277
18	0,7362	-0,1896
19	0,7209	-0,2895
20	0,1093	-1,1628
21	0,6452	0,1044
22	0,7520	0,0526
23	-0,0952	0,3405
24	0,9352	0,0505
25	-1,2490	-0,0256
26	0,5888	0,1390
27	0,0127	0,2107
28	0,7990	-0,1115
29	0,9496	0,2363
30	0,2530	0,3374
31	1,3468	0,1106

32	-0,3398	0,0437
33	1,2602	0,0146
34	1,0204	-0,8393
35	0,9381	0,0364
36	0,6418	0,1004
37	-0,8736	-1,2594
38	-1,1368	-1,5362
39	-0,8460	-1,2403
40	1,3709	-0,1420
41	1,3404	-0,0708
42	0,0985	0,0741
43	0,1348	0,8314
44	0,3124	0,2358
45	-1,4973	0,5211
46	-2,4458	-0,2189
47	-2,5009	-0,1631
48	1,0247	0,1614
49	-1,9157	-0,0112
50	-2,1733	-0,0789
51	-1,0262	0,0064
52	-2,5223	-0,3061
53	-2,4486	-0,2625
54	1,3776	0,1909
55	0,9875	0,2809
56	1,6574	0,1464
57	1,1209	0,3764
58	-2,2280	-0,1308
59	-1,1826	0,2035
60	-0,9950	0,3269
61	-0,2493	0,1118
62	-2,6533	-0,3093
63	-2,7946	-0,4101

Tableau 6.1 : Coordonnées des items dans l'espace à deux dimensions.

6.3 Analyse des taxonomies existantes

Nous avons élaboré un tableau donnant l'appartenance des items aux différentes catégories des taxonomies présentées dans le chapitre quatre, ensuite nous représentons les items de notre banque de données dans les classes des différentes taxonomies dans l'espace à deux dimensions, d'après le modèle de base de l'échelonnement multidimensionnel. Remarquons qu'on a laissé de côté la taxonomie établi par Carroll, parce qu'elle est construite d'une manière que la plupart des items appartiennent à plusieurs classes.

6.3.1 Appartenance de nos items aux différentes taxonomies analysées

	Taxonomies						
Items		ISAM				TIMSS	
	Bloom[1]	Com.[4]	cont.[5]	NLSMA[2]	Pellerey[3]	com.[6]	cont.[7]
1	7	2	1	1	1	2	1
2	7	2	1	1	1	2	1
3	7	2	1	1	1	2	1
4	7	2	1	1	1	2	1
5	1	2	1	1	1	1	1
6	7	2	1	1	1	2	1
7	7	2	1	1	1	3	1

[1] Bloom : 1 = connaissance de détails, 4 = traduction, 5 = interprétation, 7 = application, 8 = analyse d'éléments, 9 = analyse de relations, 10 = analyse de principes d'organisation, 14 = jugement en termes d'évidence interne.
[2] NLSMA : 1 = calcul, 2 = compréhension, 3 = application, 4 = analyse.
[3] Pellerey : 1 = connaissance d'éléments isolés, 2 = connaissance avec compréhension de concepts, 3 = construction de concepts et d'organisations, 4 = résolution de problèmes, 5 = jugement.
[4] ISAM comportement : 1 = rappel et reproduction, 2 = calcul et manipulation de symboles, 4 = interprétation de données symboliques, 7 = résolution de problèmes mathématiques, 8 = résolution de problèmes non mathématiques, 9 = analyse des problèmes et détermination des opérations nécessaires.
[5] ISAM contenu : 1 = arithmétique, 2 = algèbre, 3 = géométrie, 5 = analyse, 6 = aspects généraux.
[6] TIMSS comportement : 1 = savoir, 2 = utiliser des procédures de routine, 3 = résoudre des problèmes, 4 = raisonner.
[7] 1 = nombres entiers, 2 = fractions et proportionalité, 3 = mesure, estimation et sens des nombres, 4 = représentation de données, analyse et probabilités, 5 = géométrie, 6 = structures, relations et fonctions.

8	7	2	1	1	1	3	1
9	7	2	1	1	1	3	1
10	7	2	1	1	1	3	1
11	7	2	1	1	1	2	1
12	7	2	1	1	1	2	1
13	7	2	1	1	1	2	1
14	5	7	1	2	2	2	1
15	5	7	1	2	2	3	3
16	4	7	5	2	2	3	2
17	4	7	5	2	2	3	2
18	4	7	5	2	2	3	2
19	4	7	5	2	2	3	2
20	4	7	5	2	2	3	2
21	4	7	5	2	2	3	2
22	4	4	7	2	2	3	3
23	4	4	7	2	2	3	3
24	4	4	7	2	2	3	3
25	4	4	7	2	2	3	3
26	4	4	7	2	2	3	3
27	7	2	1	1	1	2	1
28	7	2	1	1	1	2	1
29	7	2	1	1	1	1	1
30	9	7	1	2	2	4	1
31	7	2	1	1	1	4	2
32	9	7	1	3	3	4	2
33	9	4	5	2	2	1	2
34	9	4	5	2	2	3	2
35	9	8	1	2	2	4	6
36	9	8	1	2	2	4	6
37	5	7	1	3	5	1	1
38	5	7	1	3	5	1	1

39	5	7	3	2	5	1	5
40	5	7	3	2	5	1	5
41	5	7	3	2	5	1	5
42	9	7	3	3	3	4	5
43	9	2	1	3	3	3	1
44	8	2	5	2	2	1	3
45	14	8	3	2	2	3	5
46	4	4	5	2	2	3	2
47	4	4	5	2	2	3	2
48	4	4	5	2	2	3	4
49	4	4	5	2	2	3	4
50	9	9	3	4	4	2	5
51	7	2	3	3	2	2	5
52	7	7	3	3	2	2	5
53	9	9	3	4	4	3	5
54	7	2	2	1	1	3	4
55	7	2	2	1	1	3	4
56	7	1	3	1	1	1	5
57	7	2	3	1	1	2	3
58	4	1	3	1	1	1	5
59	7	1	3	2	2	4	5
60	9	7	3	3	3	2	3
61	5	1	3	2	2	1	5
62	10	9	6	4	4	4	4
63	10	9	6	4	4	4	4

Tableau 6.2 : Appartenance des items aux différentes dimensions des taxonomies

6.3.2 Echelonnement multidimensionnel de nos données dans les catégories des différentes taxonomies

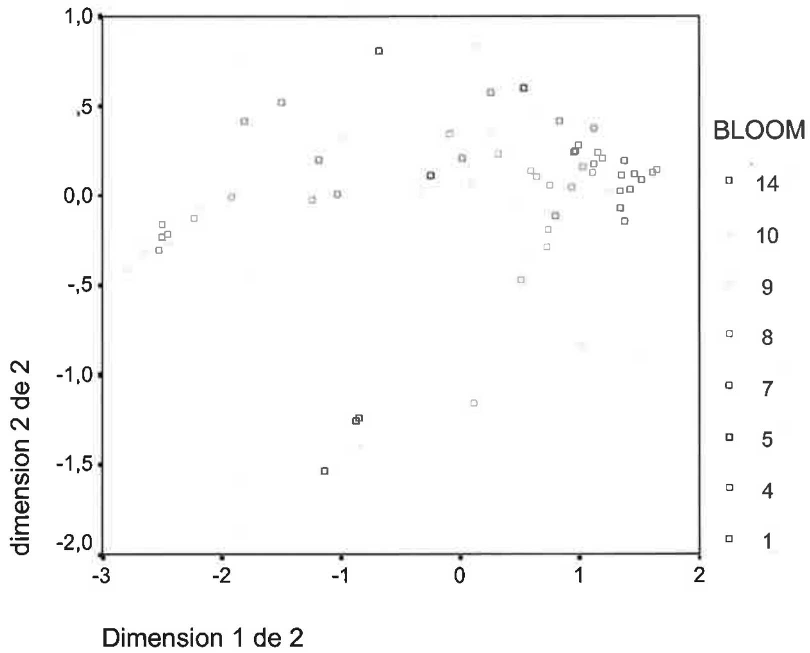


Figure 6.4 : Représentation des classes de la taxonomie de Bloom.

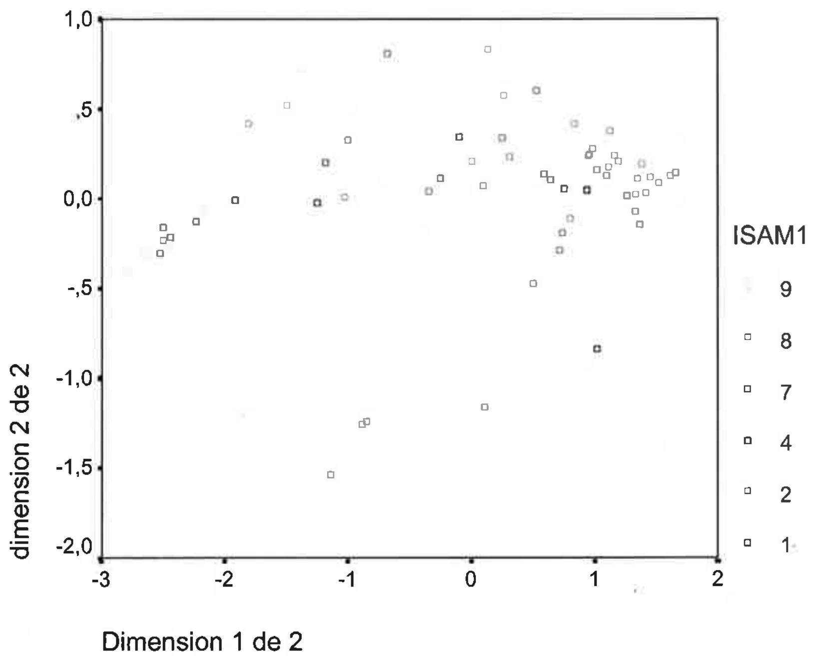


Figure 6.5 : Représentation des classes cognitives de la taxonomie de l'ISAM.

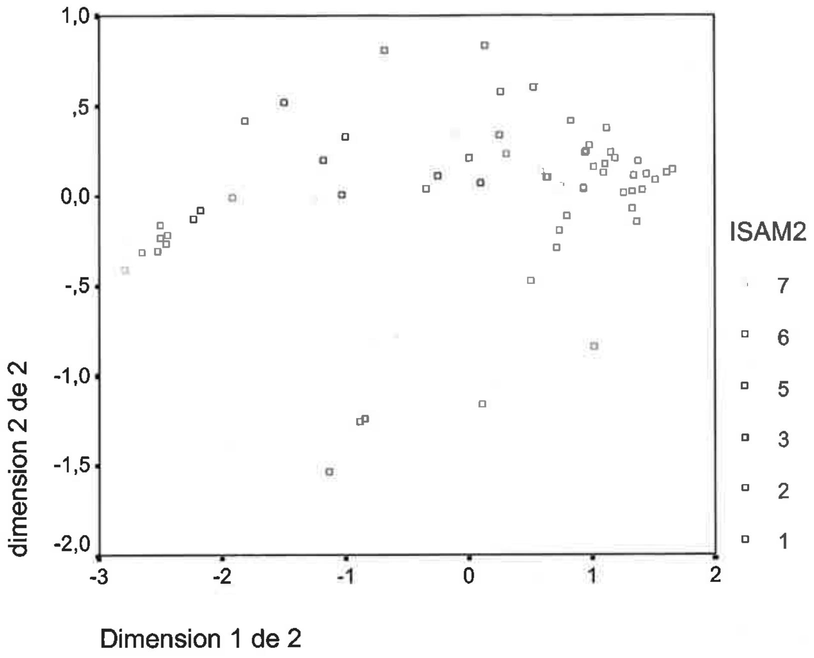


Figure 6.6 : Représentation des classes de contenu mathématique de la taxonomie de l'ISAM.

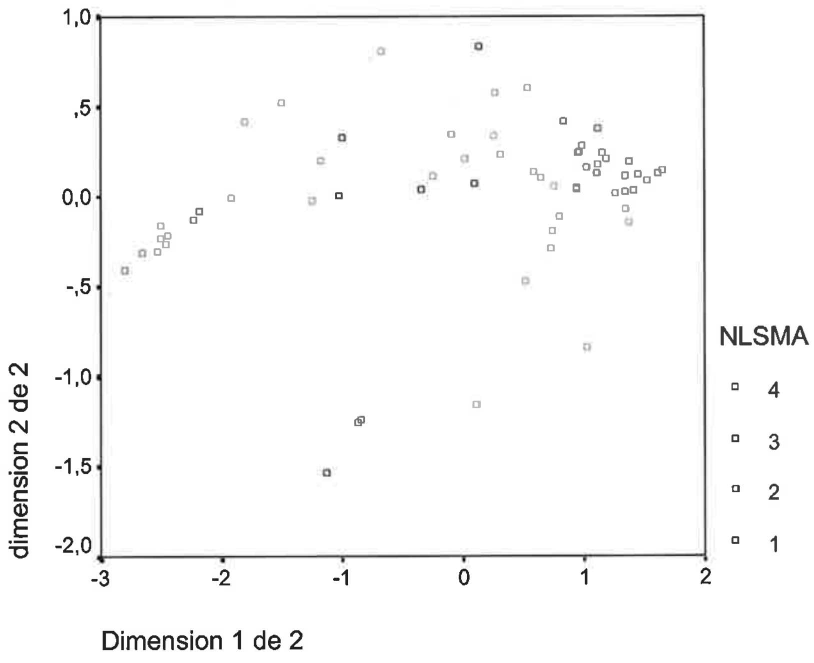


Figure 6.7 : Représentation des classes de la taxonomies de l'NLSMA.

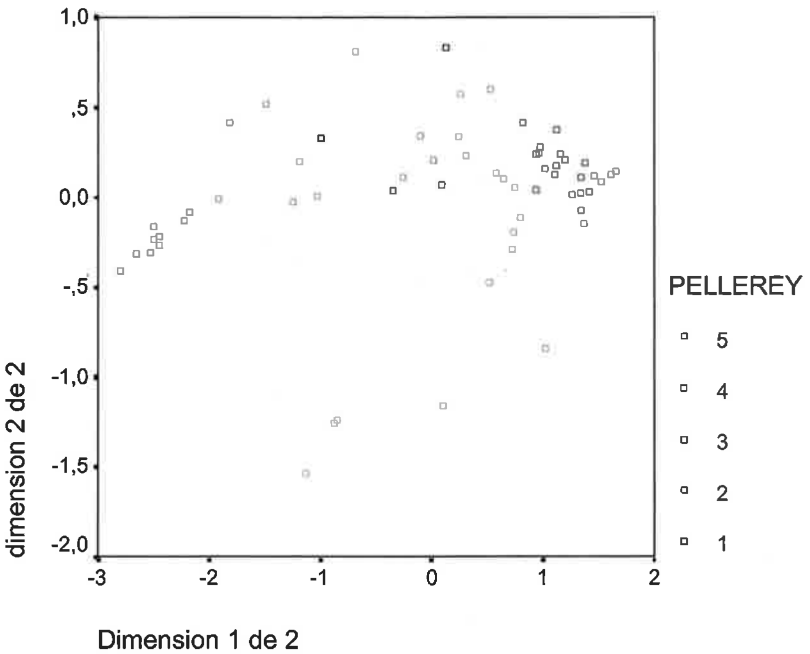


Figure 6.8 : Représentation des classes de la taxonomie de Pellerey.

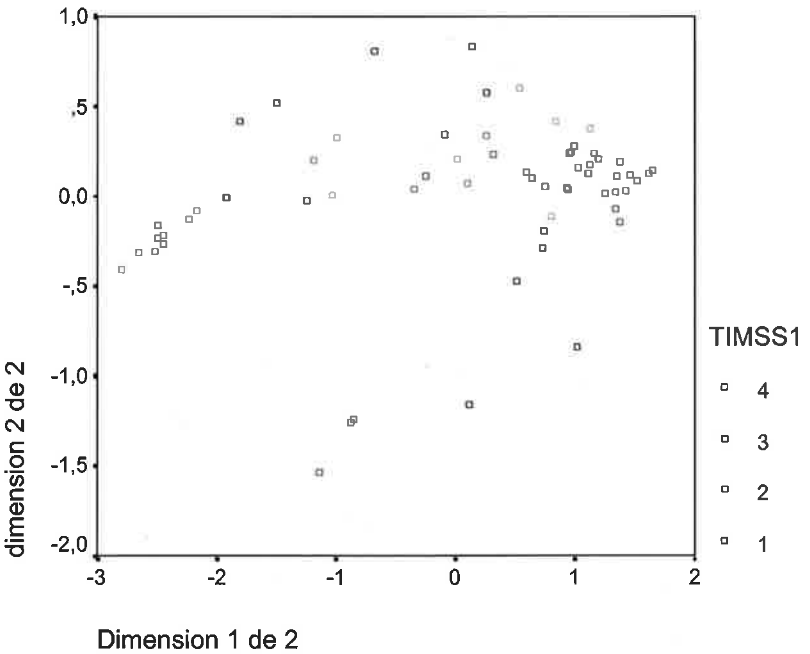


Figure 6.9 : Représentation des classes cognitives de la taxonomie du TIMSS.

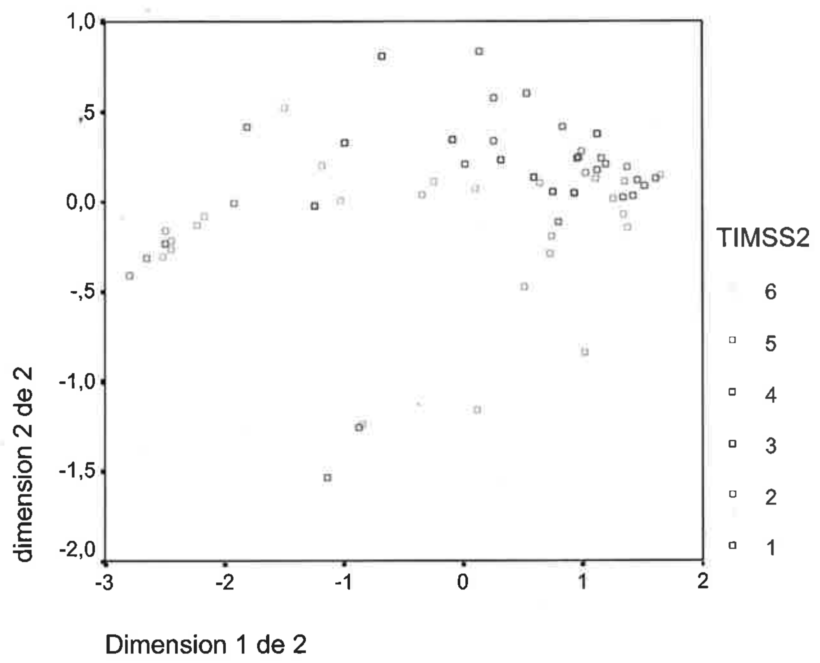


Figure 6.10 : Représentation des classes de contenu mathématique de la taxonomie du TIMSS.

Nous voyons que les taxonomies établies par Bloom, l'ISAM et le TIMSS s'appliquent assez mal aux élèves qui ont participé à l'épreuve standardisée luxembourgeoise de novembre 1996. Les taxonomies établies par Pellerey et l'NLSMA, par contre, sont plus pertinentes par rapport à nos données. La taxonomie de l'NLSMA donnerait même un très bon résultat, si les catégories de compréhension et d'application étaient réunies en une seule.

Chapitre VII

ANALYSE DES RESULTATS DE L'EPREUVE STANDARDISEE

7.1 Analyse hiérarchique ascendante

Nous avons utilisé le logiciel SPSS pour faire une analyse hiérarchique ascendante de la banque d'items d'après laquelle nous avons construit notre logiciel. La base de données pour cette analyse a comporté comme variables les 63 items qui constituent la banque d'items de notre test. Ces variables sont des variables dichotomiques ; en effet, nous avons codé par 0 une mauvaise réponse à l'item considéré et par 1 une bonne réponse et ceci pour chacun des 3590 élèves qui ont effectué la version papier-crayon du test.

Comme indice de dissimilarité, nous avons utilisé la distance euclidienne au carré pour variables binaires, procédure qui fournit l'indice de dissimilarité d'après la méthode standard pour des variables binaires. Nous avons demandé au logiciel de nous donner l'appartenance des items aux groupes constitués pour un nombre de groupes (*clusters*) variant entre 2 et 9. Nous avons effectué cette analyse au moyen de deux méthodes différentes de constitution des groupes, à savoir par la méthode du saut moyen entre les groupes, ainsi que par la méthode de Ward.

Les schémas d'agglomération obtenus par ces 2 méthodes sont les suivants :

Agglomeration Schedule

Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	46	47	142,000	0	0	15
2	2	56	369,000	0	0	3
3	2	7	483,500	2	0	4
4	2	9	529,333	3	0	6
5	54	55	631,000	0	0	16
6	2	40	655,500	4	0	7
7	2	33	701,800	6	0	9
8	62	63	728,000	0	0	15
9	2	4	731,833	7	0	10
10	2	41	779,714	9	0	11
11	2	3	806,500	10	0	13
12	59	60	811,000	0	0	34
13	2	31	828,556	11	0	16
14	35	36	862,000	0	0	31
15	46	62	871,000	1	8	17
16	2	54	879,200	13	5	19
17	46	53	885,000	15	0	20
18	21	24	890,000	0	0	24
19	2	17	900,333	16	0	22
20	46	52	921,600	17	0	28
21	18	19	956,000	0	0	32
22	1	2	975,923	0	19	23
23	1	29	988,643	22	0	25
24	21	26	1010,000	18	0	27
25	1	6	1022,667	23	0	26
26	1	57	1057,875	25	0	29
27	21	22	1066,000	24	0	35
28	46	50	1073,000	20	0	30
29	1	13	1080,177	26	0	31
30	46	58	1096,429	28	0	36
31	1	35	1096,778	29	14	33
32	18	20	1102,000	21	0	50
33	1	48	1129,250	31	0	35
34	59	61	1132,500	12	0	47
35	1	21	1141,810	33	27	38
36	5	46	1146,875	0	30	39
37	27	28	1162,000	0	0	42
38	1	12	1185,000	35	0	41
39	5	49	1200,778	36	0	55
40	23	25	1210,000	0	0	53
41	1	11	1219,654	38	0	46
42	27	30	1260,000	37	0	46
43	32	42	1265,000	0	0	45
44	14	15	1272,000	0	0	57
45	32	44	1281,500	43	0	48
46	1	27	1312,407	41	42	49
47	51	59	1328,000	0	34	48
48	32	51	1359,500	45	47	53
49	1	34	1360,767	46	0	51
50	16	18	1368,667	0	32	51
51	1	16	1372,048	49	50	54
52	37	38	1381,000	0	0	60
53	23	32	1399,286	40	48	59
54	1	10	1407,543	51	0	57
55	5	8	1442,500	39	0	56
56	5	45	1464,182	55	0	60
57	1	14	1475,472	54	44	58
58	1	43	1479,789	57	0	59
59	1	23	1513,470	58	53	62
60	5	37	1600,833	56	52	61
61	5	39	1612,643	60	0	62
62	1	5	1917,122	59	61	0

Tableau 7.1 : Schéma d'agglomération des groupes d'après la méthode du saut moyen

Agglomeration Schedule

Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	46	47	71,000	0	0	47
2	2	56	255,500	0	0	3
3	2	7	516,333	2	0	4
4	2	9	802,000	3	0	9
5	54	55	1117,500	0	0	38
6	40	41	1470,000	0	0	18
7	62	63	1834,000	0	0	15
8	33	34	2225,000	0	0	55
9	2	4	2625,200	4	0	12
10	59	60	3030,700	0	0	32
11	35	36	3461,700	0	0	43
12	2	3	3901,833	9	0	14
13	21	24	4346,833	0	0	24
14	2	31	4810,357	12	0	18
15	53	62	5281,690	0	7	17
16	18	19	5759,690	0	0	25
17	52	53	6238,357	0	15	26
18	2	40	6723,000	14	6	19
19	2	17	7235,900	18	0	28
20	22	26	7750,900	0	0	24
21	28	29	8271,400	0	0	27
22	1	6	8824,900	0	0	28
23	12	13	9384,900	0	0	33
24	21	22	9951,900	13	20	54
25	18	20	10527,233	16	0	59
26	50	52	11117,634	0	17	35
27	27	28	11710,467	0	21	39
28	1	2	12307,400	22	19	40
29	48	57	12911,900	0	0	38
30	23	25	13516,900	0	0	52
31	5	58	14122,400	0	0	35
32	59	61	14742,233	10	0	58
33	11	12	15366,900	0	23	43
34	32	42	15999,400	0	0	37
35	5	50	16632,930	31	26	47
36	14	15	17268,930	0	0	53
37	32	44	17912,430	34	0	46
38	48	54	18565,930	29	5	40
39	27	30	19232,096	27	0	49
40	1	48	19906,637	28	38	55
41	49	51	20588,137	0	0	44
42	37	38	21278,637	0	0	56
43	11	35	22007,770	33	11	49
44	45	49	22745,604	0	41	50
45	10	16	23486,604	0	0	46
46	10	32	24242,803	45	37	48
47	5	46	25005,041	35	1	60
48	10	43	25789,174	46	0	53
49	11	27	26591,762	43	39	54
50	8	45	27397,928	0	44	51
51	8	39	28232,428	50	0	52
52	8	23	29071,428	51	30	56
53	10	14	29914,344	48	36	57
54	11	21	30765,379	49	24	57
55	1	33	31617,559	40	8	61
56	8	37	32589,504	52	42	58
57	10	11	33591,855	53	54	59
58	8	59	34625,660	56	32	60
59	10	18	35755,137	57	25	61
60	5	8	37625,742	47	58	62
61	1	10	39597,949	55	59	62
62	1	5	47707,172	61	60	0

Tableau 7.2 : Schéma d'agglomération des groupes d'après la méthode de Ward.

Ces schémas indiquent la chronologie de la constitution des groupes. Expliquons brièvement ces tableaux. La première colonne est constituée tout simplement du numéro des différentes étapes. La deuxième et la troisième colonne indiquent les numéros des 2 groupes qui se sont agglomérés au cours de l'étape considérée, tandis que la quatrième colonne donne le coefficient de dissimilarité entre ces 2 groupes. La cinquième et la sixième colonne indiquent les étapes au cours desquelles les 2 groupes agglomérés dans l'étape courante ont été traités pour la première fois et la dernière colonne renvoie à la prochaine étape, au cours de laquelle le nouveau groupe formé va de nouveau entrer en jeu.

Si l'on compare les deux méthodes, on voit que les 5 premières étapes sont exactement pareilles. Dans la sixième étape, par contre, le regroupement d'après la méthode du saut moyen entre les groupes implique le regroupement des groupes numéro 2 et 40, alors que d'après la méthode de Ward, on regroupe les groupes 40 et 41. Cela ne change cependant rien au résultat final, si on se limite à un nombre de groupes compris entre 2 et 9. En effet, en utilisant la méthode du saut moyen entre les groupes, le groupe 41 est aggloméré lors de la dixième étape au groupe formé lors de la sixième étape et en utilisant la méthode de Ward, le groupe 2 est aggloméré lors de la dix-huitième étape au groupe formé précédemment.

Pour décider quel est le nombre réaliste de groupes, on examine les sauts dans les coefficients de dissimilarité. Or, dans les deux méthodes, le seul saut un peu plus grand se produit entre la 61-ème et la 62-ème étape (de 1612,643 à 1917,122 pour la méthode des sauts moyens et de 39597,949 à 47757,172 pour la méthode de Ward. Cela indique que la structure naturelle associée à nos données est une classification en deux groupes.

L'ordinateur nous sort également la répartition des items en groupes sous l'hypothèse d'existence de 2 à 9 groupes. Les tableaux 7.3 et 7.4 donnent cette répartition pour la répartition effectuée d'après la méthode des sauts moyens entre groupes, respectivement d'après la méthode de Ward.

Ensuite, nous allons montrer les dendogrammes représentant graphiquement ces résultats pour les deux méthodes de calcul utilisées.

Cluster Membership

Case	9 Clusters	8 Clusters	7 Clusters	6 Clusters	5 Clusters	4 Clusters	3 Clusters	2 Clusters
1	1	1	1	1	1	1	1	1
2	1	1	1	1	1	1	1	1
3	1	1	1	1	1	1	1	1
4	1	1	1	1	1	1	1	1
5	2	2	2	2	2	2	2	2
6	1	1	1	1	1	1	1	1
7	1	1	1	1	1	1	1	1
8	3	2	2	2	2	2	2	2
9	1	1	1	1	1	1	1	1
10	1	1	1	1	1	1	1	1
11	1	1	1	1	1	1	1	1
12	1	1	1	1	1	1	1	1
13	1	1	1	1	1	1	1	1
14	4	3	3	1	1	1	1	1
15	4	3	3	1	1	1	1	1
16	1	1	1	1	1	1	1	1
17	1	1	1	1	1	1	1	1
18	1	1	1	1	1	1	1	1
19	1	1	1	1	1	1	1	1
20	1	1	1	1	1	1	1	1
21	1	1	1	1	1	1	1	1
22	1	1	1	1	1	1	1	1
23	5	4	4	3	3	1	1	1
24	1	1	1	1	1	1	1	1
25	5	4	4	3	3	1	1	1
26	1	1	1	1	1	1	1	1
27	1	1	1	1	1	1	1	1
28	1	1	1	1	1	1	1	1
29	1	1	1	1	1	1	1	1
30	1	1	1	1	1	1	1	1
31	1	1	1	1	1	1	1	1
32	5	4	4	3	3	1	1	1
33	1	1	1	1	1	1	1	1
34	1	1	1	1	1	1	1	1
35	1	1	1	1	1	1	1	1
36	1	1	1	1	1	1	1	1
37	6	5	5	4	4	3	2	2
38	6	5	5	4	4	3	2	2
39	7	6	6	5	5	4	3	2
40	1	1	1	1	1	1	1	1
41	1	1	1	1	1	1	1	1
42	5	4	4	3	3	1	1	1
43	8	7	7	6	1	1	1	1
44	5	4	4	3	3	1	1	1
45	9	8	2	2	2	2	2	2
46	2	2	2	2	2	2	2	2
47	2	2	2	2	2	2	2	2
48	1	1	1	1	1	1	1	1
49	2	2	2	2	2	2	2	2
50	2	2	2	2	2	2	2	2
51	5	4	4	3	3	1	1	1
52	2	2	2	2	2	2	2	2
53	2	2	2	2	2	2	2	2
54	1	1	1	1	1	1	1	1
55	1	1	1	1	1	1	1	1
56	1	1	1	1	1	1	1	1
57	1	1	1	1	1	1	1	1
58	2	2	2	2	2	2	2	2
59	5	4	4	3	3	1	1	1
60	5	4	4	3	3	1	1	1
61	5	4	4	3	3	1	1	1
62	2	2	2	2	2	2	2	2
63	2	2	2	2	2	2	2	2

Tableau 7.3 : Répartition des items en groupes d'après la méthode du saut moyen entre les groupes.

Cluster Membership

Case	9 Clusters	8 Clusters	7 Clusters	6 Clusters	5 Clusters	4 Clusters	3 Clusters	2 Clusters
1	1	1	1	1	1	1	1	1
2	1	1	1	1	1	1	1	1
3	1	1	1	1	1	1	1	1
4	1	1	1	1	1	1	1	1
5	2	2	2	2	2	2	2	2
6	1	1	1	1	1	1	1	1
7	1	1	1	1	1	1	1	1
8	3	3	3	3	3	3	2	2
9	1	1	1	1	1	1	1	1
10	4	4	4	4	4	4	3	1
11	5	5	5	4	4	4	3	1
12	5	5	5	4	4	4	3	1
13	5	5	5	4	4	4	3	1
14	4	4	4	4	4	4	3	1
15	4	4	4	4	4	4	3	1
16	4	4	4	4	4	4	3	1
17	1	1	1	1	1	1	1	1
18	6	6	6	5	5	4	3	1
19	6	6	6	5	5	4	3	1
20	6	6	6	5	5	4	3	1
21	5	5	5	4	4	4	3	1
22	5	5	5	4	4	4	3	1
23	3	3	3	3	3	3	2	2
24	5	5	5	4	4	4	3	1
25	3	3	3	3	3	3	2	2
26	5	5	5	4	4	4	3	1
27	5	5	5	4	4	4	3	1
28	5	5	5	4	4	4	3	1
29	5	5	5	4	4	4	3	1
30	5	5	5	4	4	4	3	1
31	1	1	1	1	1	1	1	1
32	4	4	4	4	4	4	3	1
33	7	1	1	1	1	1	1	1
34	7	1	1	1	1	1	1	1
35	5	5	5	4	4	4	3	1
36	5	5	5	4	4	4	3	1
37	8	7	3	3	3	3	2	2
38	8	7	3	3	3	3	2	2
39	3	3	3	3	3	3	2	2
40	1	1	1	1	1	1	1	1
41	1	1	1	1	1	1	1	1
42	4	4	4	4	4	4	3	1
43	4	4	4	4	4	4	3	1
44	4	4	4	4	4	4	3	1
45	3	3	3	3	3	3	2	2
46	2	2	2	2	2	2	2	2
47	2	2	2	2	2	2	2	2
48	1	1	1	1	1	1	1	1
49	3	3	3	3	3	3	2	2
50	2	2	2	2	2	2	2	2
51	3	3	3	3	3	3	2	2
52	2	2	2	2	2	2	2	2
53	2	2	2	2	2	2	2	2
54	1	1	1	1	1	1	1	1
55	1	1	1	1	1	1	1	1
56	1	1	1	1	1	1	1	1
57	1	1	1	1	1	1	1	1
58	2	2	2	2	2	2	2	2
59	9	8	7	6	3	3	2	2
60	9	8	7	6	3	3	2	2
61	9	8	7	6	3	3	2	2
62	2	2	2	2	2	2	2	2
63	2	2	2	2	2	2	2	2

Tableau 7.4 : Répartition des items en groupes d'après la méthode de Ward.

Dendogramme d'après la méthode du saut moyen entre les groupes :

```
                         Rescaled Distance Cluster Combine

  C A S E      0         5        10        15        20        25
Label     Num  +---------+---------+---------+---------+---------+

Q_18A      46   -+-------------------+
Q_18B      47   -+                   I
Q_24A      62   -----------------+---+
Q_24B      63   -----------------+   +-----+
Q_20D      53   ---------------------+     I
Q_20C      52   ---------------------+     +-+
Q_20A      50   ---------------------------+ I
Q_22C      58   ---------------------------+ +-------+
Q_1E        5   -----------------------------+       I
Q_19B      49   -----------------------------+       +---+
Q_1H        8   -------------------------------------+   I
Q_17       45   -------------------------------------+   I
Q_12B      37   -----------------------------------+-----+-------+
Q_12C      38   -----------------------------------+     I       I
Q_12D      39   -----------------------------------------+       I
Q_6C       23   -------------------------------+---+             I
Q_6E       25   -------------------------------+   I             I
Q_9B       32   -------------------------------+-+ +---+         I
Q_13       42   -------------------------------+ +-+   I         I
Q_15       44   ---------------------------------+ I   I         I
Q_23A      59   -------------------+-------+       I   I         I
Q_23B      60   -------------------+       +-----+ I   I         I
Q_23C      61   ---------------------------+     +-+   I         I
Q_20B      51   ---------------------------------+     I         I
Q_3A       14   -------------------------------+-----+ I         I
Q_3B       15   -------------------------------+     I I         I
Q_4G       21   ---------------------+---+           I I         I
Q_6D       24   ---------------------+   +-+         I I         I
Q_6F       26   -------------------------+ +-+       I I         I
Q_6A       22   ---------------------------+ I       I I         I
Q_11A      35   ---------------------+-----+ I       I I         I
Q_11B      36   ---------------------+     I I       I I         I
Q_21A      54   -------------+-------+     I I       I I         I
Q_21B      55   -------------+       I     I I       I I         I
Q_1B        2   -------+-+           I     I I       I +---------+
Q_22A      56   -------+ +-+         I     I I       I I
Q_1G        7   ---------+ +---+     I     I I       I I
Q_1I        9   -----------+   I     I     I I       I I
Q_12E      40   ---------------+-+   I     I I       I I
Q_10A      33   ---------------+ I   I     I I       I I
Q_1D        4   -----------------+-+ +-+   I I       I I
Q_12F      41   -----------------+ I I I   I I       I I
Q_1C        3   -------------------+-+ I   I I       I I
Q_9A       31   -------------------+ I I   I +-+     I I
Q_4B       17   ---------------------+ +-+ +-+ I     I I
Q_1A        1   -----------------------+ I I I I     I I
Q_7C       29   -----------------------+ +-+ I I     I I
Q_1F        6   -------------------------+ I I +-+   I I
Q_22B      57   -------------------------+ I I I I   I I
Q_2C       13   ---------------------------+ I I I   I I
Q_19A      48   ---------------------------+ I I +-+ I I
Q_2B       12   -----------------------------+ I I I +-+
Q_2A       11   -------------------------------+ I I I
Q_7A       27   -----------------------------+-+ I I I
Q_7B       28   -----------------------------+ +-+ I I
Q_8        30   -------------------------------+   I I
Q_10B      34   -----------------------------------+ I
Q_4C       18   -----------------------+---+       I I
Q_4E       19   -----------------------+   +-------+-+
Q_4F       20   ---------------------------+       I I
Q_4A       16   -----------------------------------+ I
Q_1J       10   -----------------------------------+ I
Q_14A      43   -------------------------------------+
```

Dendogramme d'après la méthode de Ward :

```
Rescaled Distance Cluster Combine

    C A S E      0         5        10        15        20        25
  Label     Num  +---------+---------+---------+---------+---------+

  Q_18A      46   -+---+
  Q_18B      47   -+   I
  Q_24A      62   -+-+ I
  Q_24B      63   -+ I +-----+
  Q_20D      53   ---+ I     I
  Q_20C      52   ---+ I     I
  Q_20A      50   ---+-+     I
  Q_1E        5   ---+       +-------------------------------------+
  Q_22C      58   ---+       I                                     I
  Q_23A      59   ---+       I                                     I
  Q_23B      60   ---+-+     I                                     I
  Q_23C      61   ---+ I     I                                     I
  Q_12B      37   ---+-+-----+                                     I
  Q_12C      38   ---+ I                                           I
  Q_6C       23   ---+-+                                           I
  Q_6E       25   ---+ I                                           I
  Q_19B      49   ---+-+                                           I
  Q_20B      51   ---+ I                                           I
  Q_17       45   -----+                                           I
  Q_1H        8   -----+                                           I
  Q_12D      39   -----+                                           I
  Q_10A      33   -+---+                                           I
  Q_10B      34   -+   I                                           I
  Q_12E      40   -+-+ I                                           I
  Q_12F      41   -+ I I                                           I
  Q_1B        2   -+ I I                                           I
  Q_22A      56   -+ I I                                           I
  Q_1G        7   -+-+ +-----+                                     I
  Q_1I        9   -+ I I     I                                     I
  Q_1D        4   ---+ I     I                                     I
  Q_1C        3   ---+ I     I                                     I
  Q_9A       31   ---+ I     I                                     I
  Q_4B       17   ---+ I     I                                     I
  Q_1A        1   ---+-+     I                                     I
  Q_1F        6   ---+       I                                     I
  Q_21A      54   -+-+       +-------------------------------------+
  Q_21B      55   -+ I       I
  Q_19A      48   ---+       I
  Q_22B      57   ---+       I
  Q_4C       18   ---+       I
  Q_4E       19   ---+---+   I
  Q_4F       20   ---+   I   I
  Q_3A       14   ---+-+ I   I
  Q_3B       15   ---+ I I   I
  Q_9B       32   ---+ I +---+
  Q_13       42   ---+-+ I
  Q_15       44   ---+ I I
  Q_1J       10   -----+ I
  Q_4A       16   -----+ I
  Q_14A      43   -----+-+
  Q_4G       21   ---+ I
  Q_6D       24   ---+-+
  Q_6A       22   ---+ I
  Q_6F       26   ---+ I
  Q_7B       28   ---+ I
  Q_7C       29   ---+ I
  Q_7A       27   ---+-+
  Q_8        30   ---+ I
  Q_11A      35   ---+-+
  Q_11B      36   ---+ I
  Q_2B       12   ---+ I
  Q_2C       13   ---+-+
  Q_2A       11   ---+
```

La méthode de Ward aboutit plus rapidement à une structure visuelle claire en deux ou trois groupes. Le niveau d'agrégation final est cependant le même dans les deux méthodes.

7.2 Analyse factorielle

Nous avons également utilisé le logiciel SPSS pour faire une analyse factorielle de la banque d'items de notre logiciel. En partant de la même base de données que pour l'analyse hiérarchique ascendante, nous avons fait une analyse factorielle, utilisant la méthode d'analyse en composantes principales comme méthode d'extraction, puis une méthode de rotation oblique. La méthode de rotation oblique était appropriée dans notre cas, parce que, d'après les considérations théoriques du chapitre 3, il est évident qu'on doit obtenir des facteurs qui sont corrélés entre eux ; en effet, toutes les études effectuées jusqu'ici à ce sujet, tant en psychologie de l'éducation qu'en psychologie cognitive ont montré que les différentes aptitudes mathématiques ne sont pas indépendantes les unes des autres.

On obtient le tableau suivant, représentant les valeurs propres des différentes composantes, avant la rotation dans l'ordre décroissant, ainsi que le pourcentage de la variance qui peut être expliqué par chacune des composantes.

Total Variance Explained

Component	Initial Eigenvalues Total	% of Variance	Cumulative %	Rotation Total[a]
1	9,317	14,789	14,789	7,283
2	1,874	2,975	17,763	6,102
3	1,552	2,463	20,226	4,937
4	1,499	2,379	22,605	
5	1,416	2,247	24,852	
6	1,339	2,125	26,977	
7	1,312	2,083	29,060	
8	1,229	1,950	31,011	
9	1,188	1,886	32,897	
10	1,174	1,863	34,760	
11	1,122	1,781	36,540	
12	1,107	1,757	38,298	
13	1,091	1,731	40,029	
14	1,076	1,708	41,737	
15	1,048	1,663	43,400	
16	1,024	1,625	45,025	
17	1,006	1,596	46,621	
18	,997	1,582	48,203	
19	,983	1,560	49,764	
20	,978	1,553	51,316	
21	,952	1,511	52,827	
22	,944	1,498	54,326	
23	,938	1,489	55,815	
24	,929	1,474	57,289	
25	,917	1,455	58,744	
26	,904	1,434	60,179	
27	,896	1,422	61,600	
28	,871	1,382	62,983	
29	,866	1,375	64,357	
30	,859	1,364	65,721	
31	,848	1,345	67,066	
32	,838	1,331	68,397	
33	,815	1,293	69,690	
34	,808	1,282	70,972	
35	,798	1,267	72,240	
36	,793	1,259	73,498	
37	,783	1,243	74,742	
38	,773	1,227	75,968	
39	,754	1,197	77,165	
40	,746	1,184	78,349	
41	,740	1,174	79,523	
42	,732	1,162	80,685	
43	,719	1,142	81,827	
44	,710	1,127	82,954	
45	,702	1,115	84,068	
46	,699	1,109	85,178	
47	,683	1,084	86,262	
48	,659	1,045	87,307	
49	,653	1,036	88,344	
50	,633	1,005	89,349	
51	,625	,992	90,341	
52	,615	,976	91,318	
53	,607	,963	92,280	
54	,588	,933	93,213	
55	,583	,925	94,138	
56	,567	,899	95,038	
57	,562	,893	95,930	
58	,533	,846	96,776	
59	,528	,838	97,614	
60	,504	,800	98,414	
61	,453	,720	99,134	
62	,428	,679	99,813	
63	,118	,187	100,000	

Extraction Method: Principal Component Analysis.

a. When components are correlated, sums of squared loadings cannot be added to obtain a total variance.

Tableau 7.5 : Répartition de la variance totale sur les différentes composantes.

On voit qu'il existe 17 composantes qui ont une valeur propre supérieure à 1. D'après la règle de Kaiser, cela voudrait dire qu'on est en présence de 17 facteurs, ce qui correspondrait à des facteurs très spécifiques, alors que nous nous intéressons à des facteurs plus généraux, correspondant à ceux de la strate moyenne de Carroll.

Pour arriver à un nombre de facteurs plus adapté, nous analysons le diagramme suivant, qui représente la taille des valeurs propres sous forme graphique.

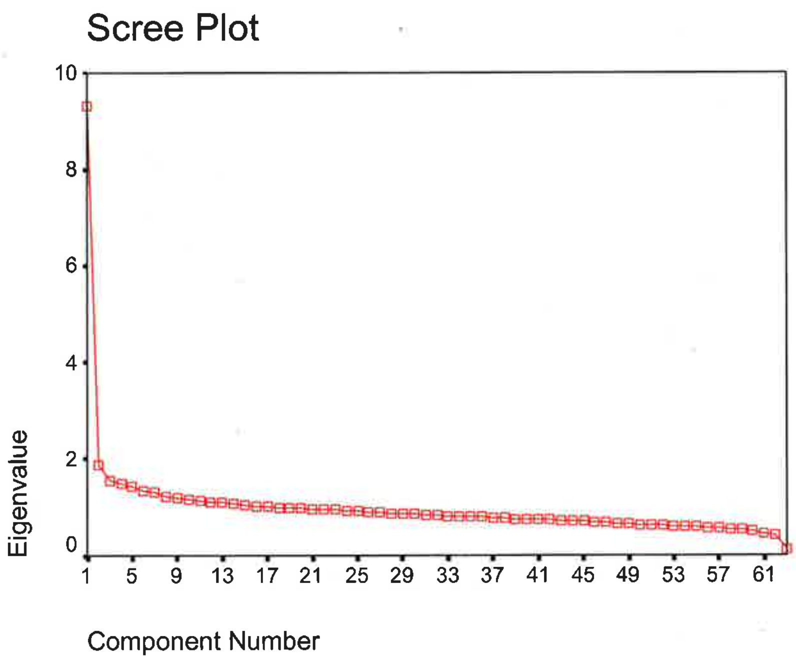


Figure 7.1 : Représentation de la taille des valeurs propres.

On voit qu'à partir de la troisième composante, la pente de la courbe devient presque nulle. Cela indique que seuls les trois premiers facteurs sont importants. On peut donc supposer qu'on est en présence de trois facteurs principaux qui représentent aedéquatement les compétences en mathématiques.

Ces facteurs sont corrélés entre eux de la manière suivante :

Component Correlation Matrix

Component	1	2	3
1	1,000	,459	-,349
2	,459	1,000	-,304
3	-,349	-,304	1,000

Extraction Method: Principal Component Analysis.
Rotation Method: Oblimin with Kaiser Normalization.

Tableau 7.6 : Matrice de corrélation entre les 3 facteurs postulés.

Il est intéressant de remarquer que le troisième facteur est négativement corrélé avec les deux autres, ce qui signifie que les aptitudes utilisées pour traiter les différents items sont partiellement exclusives les unes des autres.

La saturation des items selon les 3 facteurs postulés est quant à elle indiquée par le tableau 7. En analysant ce tableau, on arrive à l'interprétation suivante des trois facteurs :

• Facteur 1 :
La résolution des items saturés en facteur 1 présuppose une bonne aptitude de visualisation spatiale. Celle-ci s'exerce soit directement sur les figures géométriques, soit sur la série des nombres étalés mentalement dans l'espace.
Dénomination proposée : Représentation spatiale.

• Facteur 2 :
La résolution des items saturés en facteur 2 présuppose l'application de procédures de calculs apprises.
Dénomination proposée : Connaissances procédurales de calcul.

• Facteur 3 :
Les items saturés dans ce facteur, qui est corrélé négativement avec les deux précédents, présupposent la découverte de solutions inédites, donc l'utilisation du raisonnement.
Dénomination proposée : Raisonnement mathématique.

La bipolarité constatée signifie que les élèves utilisent soit des schèmes appris, soit un moyen de pensée plus créatif, correspondant à la capacité de résolution de problèmes. L'opposition bipolaire entre ces deux types d'aptitude peut être interprétée en termes d'intelligence fluide et d'intelligence cristallisée, opposition apparue depuis longtemps dans les testes d'intelligence classiques.

Si nous rapprochons notre taxonomie de celle des compétences cognitives de Carroll, notre premier facteur pourrait se situer au niveau du facteur 2V (perception visuelle), notre deuxième facteur pourrait correspondre à 2C (intelligence cristallisée)

Pattern Matrix [a]

	Component		
	1	2	3
46	,785	-,239	
47	,782	-,225	
59	,581		
60	,509		
58	,482		
53	,478		
61	,470	,134	
49	,470		
52	,464		
50	,452		
51	,446		-,129
63	,361		
45	,338		
62	,335		
32	,297	,198	-,280
5	,295		
44	,271	,195	-,139
43	,237	,189	
57	,234		
14	,206	,206	
17	,206	,152	-,155
8	,185	,176	
55		,529	
13		,473	
54		,453	,155
9		,431	
12		,419	
35		,377	-,184
3		,365	
11		,362	
27	,129	,358	-,131
7		,355	
24		,349	-,339
10	,135	,348	
1		,335	
36	,114	,334	-,229
29		,331	-,183
6		,310	
4		,300	
31		,299	
30	,183	,287	-,132
28		,284	-,169
2		,251	
48	,191	,198	
15	,146	,187	-,157
19			-,552
33		,104	-,546
18			-,527
20			-,515
34	-,199		-,443
21	,107	,249	-,351
40			-,332
25	,269	,114	-,318
16	,212		-,276
22	,148	,150	-,266
26	,128	,252	-,256
39	,130		-,254
42	,229	,185	-,252
38	,114		-,239
23	,171	,184	-,230
41			-,217
37	,154		-,200
56	,146		,147

Extraction Method: Principal Component Analysis.
Rotation Method: Oblimin with Kaiser Normalization.

a. Rotation converged in 12 iterations.

Tableau 7.7 : Saturation des items selon les 3 facteurs postulés.

et 2Y (mémoire et savoir), et notre troisième facteur pourrait correspondre exactement à 2F (intelligence fluide).

Certains items sont saturés à la fois dans les facteurs 1 et 2 ; ils sont donc résolus par beaucoup d'élèves par une combinaison de la visualisation spatiale et des connaissances procédurales, alors que l'utilisation du raisonnement semble plutôt exclure d'autres stratégies de résolution.

7.3 Proposition d'une taxonomie

Une démarche exploratoire, basée sur l'utilisation conjointe de plusieurs procédés multivariés, à savoir l'échelonnement multidimensionnel, permettant de classer nos items par rapports aux taxonomies existantes, l'analyse hiérarchique ascendante et l'analyse factorielle, servant à dégager la structure factorielle latente de notre banque de données, nous a fourni une taxonomie théorique des compétences en mathématiques, correspondant à la dénomination de nos trois facteurs de second ordre :

a) la représentation spatiale
b) les connaissances procédurales de calcul
c) le raisonnement mathématique.

Les coordonnées des items dans l'espace à deux dimensions fournies par échelonnement multidimensionnel nous permettent de faire la représentation spatiale suivante des facteurs trouvés :

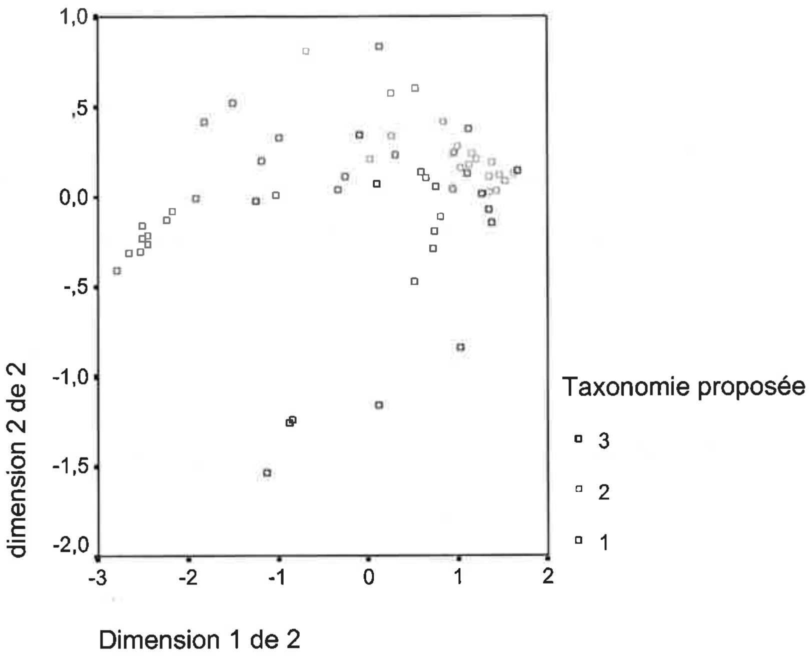


Figure 1 : Représentation de la structure factorielle dégagée.

La représentation graphique des facteurs devient plus satisfaisante, si les items suivants, saturés presque aussi fortement dans deux facteurs différents, sont reclassés de la manière suivante :

$$23 \rightarrow 2, \quad 15 \rightarrow 1, \quad 57 \rightarrow 2, \quad 17 \rightarrow 2.$$

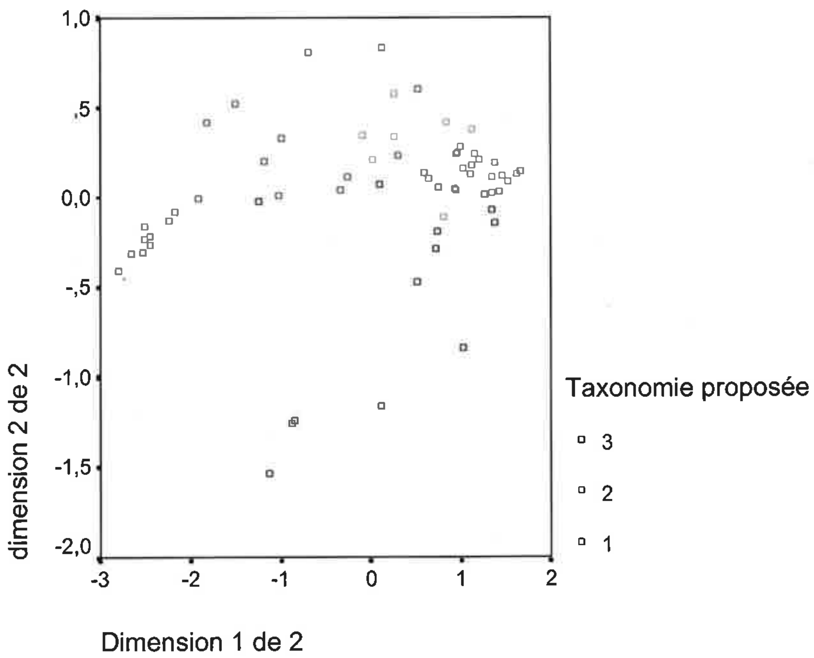


Figure 2 : Représentation de la structure factorielle modifiée.

Le résultat de notre analyse factorielle est psychologiquement plausible. Il est en accord avec la taxonomie des compétences cognitives de Carroll, qui correspond à une synthèse de la recherche sur les tests d'intelligence. D'autre part, nous avons vu que notre banque de données se positionne très bien par rapport à la taxonomie de l'NLSMA.

Ce qui est surprenant dans nos résultats, c'est le fait que le facteur de raisonnement est corrélé négativement avec les deux autres facteurs. Nous allons discuter ce point dans la suite de notre démarche confirmatoire.

Chapitre VIII

ANALYSE CONFIRMATOIRE DE NOTRE TAXONOMIE

8.1 Analyse de la taxonomie par le LISREL

Pour vérifier l'adéquation de notre taxonomie aux données empiriques, nous avons effectué une analyse d'un test expérimental autre que celui utilisé pour constituer notre banque de données et nous avons examiné, au moyen du logiciel LISREL, si l'on y retrouve les trois facteurs constituant notre taxonomie avec les mêmes relations entre eux.

Comme test expérimental, nous avons choisi l'épreuve standardisée en mathématiques de mars 1997 de la procédure de passage primaire-postprimaire. Ce choix avait l'énorme avantage que nous disposions déjà des résultats de presque 3600 élèves, alors que cela aurait pris beaucoup trop de temps et de moyens de construire nous-même un tel test et surtout de le faire passer par un groupe tellement grand d'élèves.

Nous testons ensuite sur notre modèle factoriel, obtenu par l'analyse factorielle des items de la banque de données de notre logiciel, en appliquant le LISREL sur ce nouveau test expérimental. Puisque, dans la vérification d'un modèle factoriel, le modèle structural du LISREL n'est pas de mise, nous avons seulement besoin de son modèle de mesure.

Notre modèle de mesure comporte comme variables observées exogènes les réponses aux 43 items contenus dans le test expérimental et comme variables latentes indépendantes correspondantes les trois catégories de notre taxonomie, à savoir la représentation spatiale, les connaissances procédurales de calcul et le raisonnement mathématique.

L'hypothèse que nous voulons tester est que la matrice des covariances des 43 items de notre test expérimental peut être expliquée par l'existence de 3 facteurs latents liés entre eux par la matrice de covariance obtenue lors de l'analyse factorielle des items de la banque de données de notre logiciel.

En utilisant la terminologie habituelle du LISREL, nous notons nos variables exogènes $x_1, x_2, \ldots, x_{43}$, les facteurs latents ξ_1, ξ_2, ξ_3, les résidus des variables exogènes $\delta_1, \delta_2, \ldots, \delta_{44}$ et les résidus des facteurs latents $\zeta_1, \zeta_2, \zeta_3$. En tenant compte du fait que les trois catégories de notre taxonomie ne sont pas indépendantes, nous obtenons la représentation suivante de notre modèle de mesure :

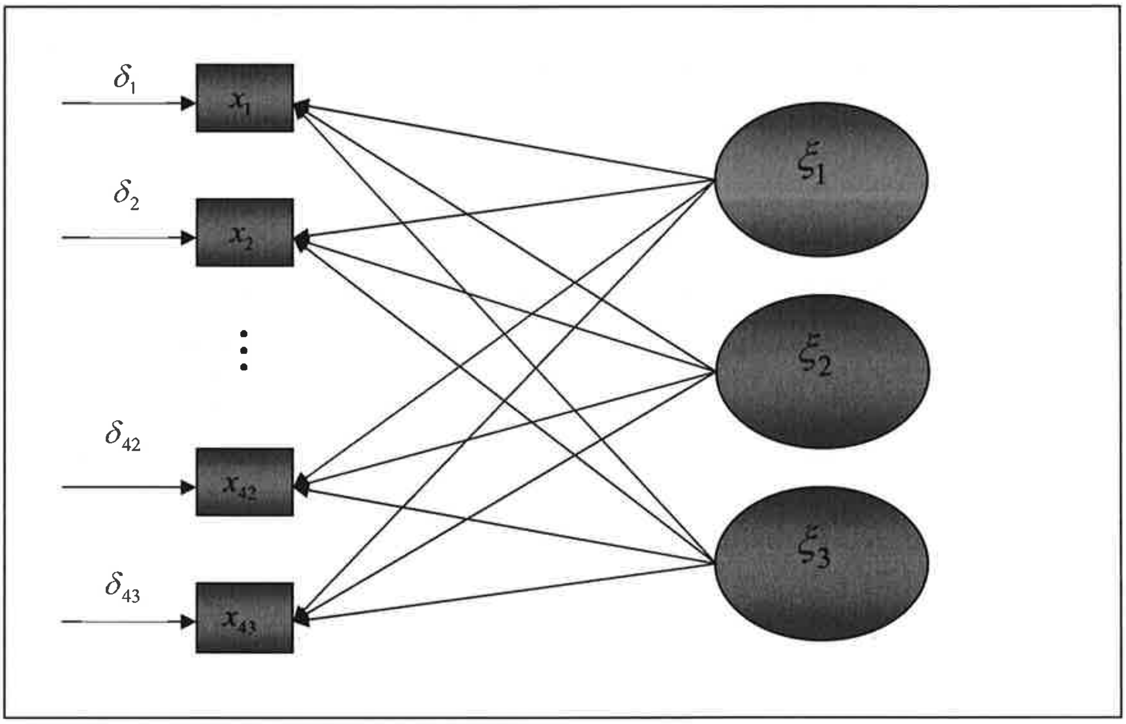


Figure 8.1 : Représentation de notre modèle de mesure

En notant Λ_x la matrice factorielle des variables latentes indépendantes de dimension 43×3, Φ leur matrice de covariance de dimension 3×3 et Θ la matrice de covariance des résidus des variables exogènes de dimension 43×43, nous obtenons l'équation de mesure

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

où x désigne le vecteur de dimension 43×1 comportant les variables x_i, $1 \le i \le 43$ et δ le vecteur de même dimension comportant les résidus δ_1, $1 \le i \le 43$.

Comme les catégories de notre taxonomie sont intercorrélées entre elles et qu'elles sont toutes définies a priori par l'ensemble des 43 items, nous ne pouvons pas fixer une partie des paramètres de la matrice Λ_x comme étant égaux à zèro ; nous déclarons donc cette matrice comme matrice libre. Pour des raisons d'unicité de la solution calculée par l'ordinateur, il faut fixer cependant au moins un élément de la matrice. Nous posons donc l'élément $\lambda_{x,43,3}$ comme étant égal à 1. Cet élément correspond en effet à une question 26 du test, qui nécessite clairement un raisonnement.

La matrice Φ par contre est entièrement fixée, puisque nous supposons qu'elle doit être égale à la matrice de covariance entre nos trois facteurs, obtenue par l'analyse factorielle des items constituant la banque de données de notre logiciel. Elle est par conséquent égale à

$$\Phi = \begin{pmatrix} 0,984 & 0,721 & 1,491 \\ 0,721 & 1,164 & 0,267 \\ 1,491 & 0,267 & 2,773 \end{pmatrix}.$$

Pour la matrice Θ finalement, nous utilisons le modèle classique, c'est-à-dire nous supposons que les erreurs de mesure sont aléatoires et non corrélées. Cette matrice est donc une matrice diagonale.

Pris comme cela, le logiciel LISREL infirme notre modèle à un niveau de 95%, mais il l'accepte à un niveau de 80 %. Cela n'est pas très étonnant. En effet, l'analyse factorielle qu'on a effectué sur les réponses aux items de notre banque de données avairt déjà montré que les trois facteurs retenus n'expliquent qu'environ un tiers de la variance totale et qu'à côté, il existe un très grand nombre de facteurs négligeables qui pris ensemble ont un poids de deux tiers. Comme l'explication de la variance est un critère important du LISREL pour le calcul de son indice d'adéquation, nous ne pouvions pas espérer un résultat excellant.

De plus, le fait d'avoir fixé complètement la matrice de covariance entre les trois facteurs latents est une hypothèse très contraignante. Si nous laissons libre la matrice Φ, le LISREL accepte notre modèle à un niveau de 90 %.

Compte tenu de ces explications, nous pouvons conclure que notre modèle est confirmé par le LISREL. En effectuant une analyse factorielle sur les résultats de notre test expérimental confirmatoire, nous obtenons d'ailleurs une solution à trois facteurs et une matrice de corrélation entre les facteurs qui correspond assez bien à celle donné dans le tableau 7.6.

8.2 Analyse de la variance

Nous avons finalement comparé notre taxonomie à toutes celles que nous avons testées en effectuant une analyse de la variance multivariée. Nous avons testé l'effet de trois facteurs sur la matrice de dissimilarité des différentes taxonomies, à savoir les 2 paramètres « difficulté » et « discrimination » de nos items provenant du modèle de Birnbaum (variables A et B) (voir Schiltz, 1996), ainsi que la probabilité de réussite, établie empiriquement pour chaque item (variable P). Ces trois variables sont de niveau de mesure d'intervalle et comportent un nombre élevé de modalités. Ceci nous a amené à les prendre comme covariants plutôt que comme facteurs et nous avons ainsi en fait effectué une analyse de la covariance.

Comme variable dépendantes nous avons pris les catégories des différentes taxonomies. Nous avons donc examiné l'effet des trois covariants sur la capacité des différentes taxonomies de classer nos items sous des facteurs de second ordre, identifiés par des études préliminaires et psychologiquement parlants.

Comme nos variables dépendantes sont intercorrélées entre elles, nous avons utilisé un modèle multivarié, plutôt que d'analyser la situation à l'aide de plusieurs procédures univariées.

Malheureusement nous n'avons pas pu calculer les effets interactifs des trois facteurs sur les variables dépendantes, en raison du nombre trop élevé de modalités, situation, qui se présente d'ailleurs fréquemment en analyse de la covariance. Nous avons par conséquent effectué des analyses unifactorielles multivariés pour chacun des trois covariants.

Pour pallier à cet inconvénient, nous avons examiné les relations des trois facteurs entre eux. Gardons en mémoire que leur mode de construction n'est pas le même. Alors que les variables *A* et *B* sont les paramètres des items d'après le modèle de Birnbaum et dépendent donc uniquement des items, la variable *P* a été construite de manière empirique et dépend également de l'échantillon des sujets examinés.

Calculons les corrélations entre les trois covariants :

Correlations

			prob reussite	discrimination	difficulté
Spearman's rho	Correlation Coefficient	prob reussite	1,000	-,215	-,942**
		discrimination	-,215	1,000	,316*
		difficulté	-,942**	,316*	1,000
	Sig. (2-tailed)	prob reussite	,	,090	,000
		discrimination	,090	,	,012
		difficulté	,000	,012	,
	N	prob reussite	63	63	63
		discrimination	63	63	63
		difficulté	63	63	63

**. Correlation is significant at the .01 level (2-tailed).

*. Correlation is significant at the .05 level (2-tailed).

La corrélation négative hautement significative entre *P* et *A* est tout à fait plausible. En toute logique, la probabilité de réussite des items doit être inversement proportionnelle à leur difficulté. D'un autre côté, la probabilité de réussite n'est pas liée au pouvoir de discrimination des items, ce qui signifie que, pour un élève donné, le fait qu'un item soit réussi ou non, ne dépend pas du pouvoir de l'item de discriminer entre lui et d'autres élèves, mais de la manière dont il est capable de le résoudre.

Ce qui est surprenant, c'est la corrélation positive entre *A* et *B*, alors que d'après le modèle IRT, *A* et *B* devraient être des paramètres indépendants (voir

Schiltz, 1996). Ceci probablement dû au fait que notre banque d'items provenant des procédures de passage primaire-postprimaire est mal équilibrée. Pour des raisons pédagogiques, elle comprend trop d'items de calcul faciles et faiblement discriminants.

Les résultats des tests des effets intergroupes (*test of between-subjects effects*) sur les catégories des différentes taxonomies sont présentés sur les pages suivantes.

Pour le covariant *A* (« indice de difficulté »), l'analyse de la variance donne un effet significatif au seuil de 5 % pour notre taxonomie, ainsi que pour celle du contenu mathématique de l'Etude Internationale des Performances en Mathématiques (ISAM2), celle de l'Etude Longitudinale Nationale des Aptitudes Mathématiques (NLSMA), et celles du contenu mathématique et de performance cognitive de la Troisième Etude Internationale des Mathématiques et des Sciences (TIMSS1 et TIMSS2).

Pour le covariant *B* (« indice de discrimination »), l'analyse de la variance fournit un effet significatif au seuil de 1 % pour notre taxonomie, ainsi que pour celle des aptitudes comportementales de l'Etude Internationale des Performances en Mathématiques (ISAM1), celle de l'Etude Longitudinale Nationale des Aptitudes Mathématiques (NLSMA), ainsi que celle de Pellerey et un effet significatif au seuil de 5 % pour celle du contenu mathématique de l'Etude Internationale des Performances en Mathématiques (ISAM2).

Pour le covariant *P* (« probabilité de réussite »), finalement, l'analyse de la variance donne un effet significatif au seuil de 1 % pour notre taxonomie, ainsi que pour pour celle des aptitudes comportementales de l'Etude Internationale des Performances en Mathématiques (ISAM1), celle de l'Etude Longitudinale Nationale des Aptitudes Mathématiques (NLSMA), ainsi que celle de Pellerey et un effet significatif au seuil de 5 % pour celle de la performance cognitive de la Troisième Etude Internationale des Mathématiques et des Sciences (TIMSS2).

Examinés séparément, nos trois covariants ont donc une influence significative sur la structure classificatrice de notre taxonomie, c'est-à-dire qu'elle répond aux exigences que nous avons posées. Parmi les autres taxonomies examinées, seule la taxonomie celle de l'Etude Longitudinale Nationale des Aptitudes Mathématiques (NLSMA) remplit également ces conditions.

Tests of Between-Subjects Effects

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Noncent. Parameter	Observed Power[a]
Corrected Model	Taxonomie proposée	2,507[b]	1	2,507	4,340	,041	4,340	,536
	TIMSS2	12,760[c]	1	12,760	5,016	,029	5,016	,597
	TIMSS1	4,408[c]	1	4,408	5,041	,028	5,041	,599
	BLOOM	1,278[d]	1	1,278	,256	,615	,256	,079
	ISAM1	4,901[e]	1	4,901	,711	,402	,711	,132
	ISAM2	25,716[f]	1	25,716	6,814	,011	6,814	,729
	NLSMA	2,716[g]	1	2,716	4,703	,034	4,703	,569
	PELLEREY	6,86E-02[h]	1	6,86E-02	,054	,816	,054	,056
Intercept	Taxonomie proposée	39,518	1	39,518	68,406	,000	68,406	1,000
	TIMSS2	11,461	1	11,461	4,505	,038	4,505	,551
	TIMSS1	17,616	1	17,616	20,146	,000	20,146	,993
	BLOOM	221,320	1	221,320	44,355	,000	44,355	1,000
	ISAM1	78,675	1	78,675	11,409	,001	11,409	,914
	ISAM2	6,298	1	6,298	1,669	,201	1,669	,246
	NLSMA	11,906	1	11,906	20,620	,000	20,620	,994
	PELLEREY	32,901	1	32,901	26,125	,000	26,125	,999
A	Taxonomie proposée	2,507	1	2,507	4,340	,041	4,340	,536
	TIMSS2	12,760	1	12,760	5,016	,029	5,016	,597
	TIMSS1	4,408	1	4,408	5,041	,028	5,041	,599
	BLOOM	1,278	1	1,278	,256	,615	,256	,079
	ISAM1	4,901	1	4,901	,711	,402	,711	,132
	ISAM2	25,716	1	25,716	6,814	,011	6,814	,729
	NLSMA	2,716	1	2,716	4,703	,034	4,703	,569
	PELLEREY	6,86E-02	1	6,86E-02	,054	,816	,054	,056
Error	Taxonomie proposée	35,239	61	,578				
	TIMSS2	155,176	61	2,544				
	TIMSS1	53,338	61	,874				
	BLOOM	304,372	61	4,990				
	ISAM1	420,654	61	6,896				
	ISAM2	230,221	61	3,774				
	NLSMA	35,221	61	,577				
	PELLEREY	76,820	61	1,259				
Total	Taxonomie proposée	274,000	63					
	TIMSS2	643,000	63					
	TIMSS1	454,000	63					
	BLOOM	2935,000	63					
	ISAM1	1670,000	63					
	ISAM2	811,000	63					
	NLSMA	298,000	63					
	PELLEREY	388,000	63					
Corrected Total	Taxonomie proposée	37,746	62					
	TIMSS2	167,937	62					
	TIMSS1	57,746	62					
	BLOOM	305,651	62					
	ISAM1	425,556	62					
	ISAM2	255,937	62					
	NLSMA	37,937	62					
	PELLEREY	76,889	62					

a. Computed using alpha = ,05

b. R Squared = ,066 (Adjusted R Squared = ,051)

c. R Squared = ,076 (Adjusted R Squared = ,061)

d. R Squared = ,004 (Adjusted R Squared = -,012)

e. R Squared = ,012 (Adjusted R Squared = -,005)

f. R Squared = ,100 (Adjusted R Squared = ,086)

g. R Squared = ,072 (Adjusted R Squared = ,056)

h. R Squared = ,001 (Adjusted R Squared = -,015)

Figure 8.3 : Effets de la source de variation « difficulté » sur les différentes taxonomies

Tests of Between-Subjects Effects

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Noncent. Parameter	Observed Power[a]
Corrected Model	Taxonomie proposée	7,520[b]	1	7,520	15,177	,000	15,177	,970
	TIMSS2	4,311[c]	1	4,311	1,607	,210	1,607	,239
	TIMSS1	2,505[d]	1	2,505	2,766	,101	2,766	,374
	BLOOM	,639[e]	1	,639	,128	,722	,128	,064
	ISAM1	62,637[f]	1	62,637	10,528	,002	10,528	,891
	ISAM2	17,252[g]	1	17,252	4,409	,040	4,409	,543
	NLSMA	12,139[h]	1	12,139	28,703	,000	28,703	1,000
	PELLEREY	10,234[i]	1	10,234	9,365	,003	9,365	,854
Intercept	Taxonomie proposée	162,440	1	162,440	327,828	,000	327,828	1,000
	TIMSS2	424,765	1	424,765	158,354	,000	158,354	1,000
	TIMSS1	349,298	1	349,298	385,714	,000	385,714	1,000
	BLOOM	2122,384	1	2122,384	424,461	,000	424,461	1,000
	ISAM1	1245,632	1	1245,632	209,368	,000	209,368	1,000
	ISAM2	533,130	1	533,130	136,251	,000	136,251	1,000
	NLSMA	258,478	1	258,478	611,186	,000	611,186	1,000
	PELLEREY	300,132	1	300,132	274,668	,000	274,668	1,000
B	Taxonomie proposée	7,520	1	7,520	15,177	,000	15,177	,970
	TIMSS2	4,311	1	4,311	1,607	,210	1,607	,239
	TIMSS1	2,505	1	2,505	2,766	,101	2,766	,374
	BLOOM	,639	1	,639	,128	,722	,128	,064
	ISAM1	62,637	1	62,637	10,528	,002	10,528	,891
	ISAM2	17,252	1	17,252	4,409	,040	4,409	,543
	NLSMA	12,139	1	12,139	28,703	,000	28,703	1,000
	PELLEREY	10,234	1	10,234	9,365	,003	9,365	,854
Error	Taxonomie proposée	30,226	61	,496				
	TIMSS2	163,625	61	2,682				
	TIMSS1	55,241	61	,906				
	BLOOM	305,011	61	5,000				
	ISAM1	362,919	61	5,949				
	ISAM2	238,684	61	3,913				
	NLSMA	25,798	61	,423				
	PELLEREY	66,655	61	1,093				
Total	Taxonomie proposée	274,000	63					
	TIMSS2	643,000	63					
	TIMSS1	454,000	63					
	BLOOM	2935,000	63					
	ISAM1	1670,000	63					
	ISAM2	811,000	63					
	NLSMA	298,000	63					
	PELLEREY	388,000	63					
Corrected Total	Taxonomie proposée	37,746	62					
	TIMSS2	167,937	62					
	TIMSS1	57,746	62					
	BLOOM	305,651	62					
	ISAM1	425,556	62					
	ISAM2	255,937	62					
	NLSMA	37,937	62					
	PELLEREY	76,889	62					

a. Computed using alpha = ,05

b. R Squared = ,199 (Adjusted R Squared = ,186)

c. R Squared = ,026 (Adjusted R Squared = ,010)

d. R Squared = ,043 (Adjusted R Squared = ,028)

e. R Squared = ,002 (Adjusted R Squared = -,014)

f. R Squared = ,147 (Adjusted R Squared = ,133)

g. R Squared = ,067 (Adjusted R Squared = ,052)

h. R Squared = ,320 (Adjusted R Squared = ,309)

i. R Squared = ,133 (Adjusted R Squared = ,119)

Figure 8.4 : Effets de la source de variation « discrimination » sur les différentes taxonomies

Tests of Between-Subjects Effects

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Noncent. Parameter	Observed Power[a]
Corrected Model	Taxonomie proposée	11,299[b]	1	11,299	26,063	,000	26,063	,999
	TIMSS2	10,633[c]	1	10,633	4,123	,047	4,123	,515
	TIMSS1	1,045[d]	1	1,045	1,124	,293	1,124	,181
	BLOOM	4,59E-02[e]	1	4,59E-02	,009	,924	,009	,051
	ISAM1	55,321[f]	1	55,321	9,115	,004	9,115	,844
	ISAM2	15,490[g]	1	15,490	3,930	,052	3,930	,496
	NLSMA	17,723[h]	1	17,723	53,485	,000	53,485	1,000
	PELLEREY	14,576[i]	1	14,576	14,269	,000	14,269	,961
Intercept	Taxonomie proposée	3,454	1	3,454	7,966	,006	7,966	,793
	TIMSS2	104,449	1	104,449	40,504	,000	40,504	1,000
	TIMSS1	56,041	1	56,041	60,290	,000	60,290	1,000
	BLOOM	288,923	1	288,923	57,670	,000	57,670	1,000
	ISAM1	345,293	1	345,293	56,890	,000	56,890	1,000
	ISAM2	130,768	1	130,768	33,175	,000	33,175	1,000
	NLSMA	85,743	1	85,743	258,756	,000	258,756	1,000
	PELLEREY	88,069	1	88,069	86,213	,000	86,213	1,000
P	Taxonomie proposée	11,299	1	11,299	26,063	,000	26,063	,999
	TIMSS2	10,633	1	10,633	4,123	,047	4,123	,515
	TIMSS1	1,045	1	1,045	1,124	,293	1,124	,181
	BLOOM	4,59E-02	1	4,59E-02	,009	,924	,009	,051
	ISAM1	55,321	1	55,321	9,115	,004	9,115	,844
	ISAM2	15,490	1	15,490	3,930	,052	3,930	,496
	NLSMA	17,723	1	17,723	53,485	,000	53,485	1,000
	PELLEREY	14,576	1	14,576	14,269	,000	14,269	,961
Error	Taxonomie proposée	26,447	61	,434				
	TIMSS2	157,303	61	2,579				
	TIMSS1	56,701	61	,930				
	BLOOM	305,605	61	5,010				
	ISAM1	370,235	61	6,069				
	ISAM2	240,447	61	3,942				
	NLSMA	20,213	61	,331				
	PELLEREY	62,313	61	1,022				
Total	Taxonomie proposée	274,000	63					
	TIMSS2	643,000	63					
	TIMSS1	454,000	63					
	BLOOM	2935,000	63					
	ISAM1	1670,000	63					
	ISAM2	811,000	63					
	NLSMA	298,000	63					
	PELLEREY	388,000	63					
Corrected Total	Taxonomie proposée	37,746	62					
	TIMSS2	167,937	62					
	TIMSS1	57,746	62					
	BLOOM	305,651	62					
	ISAM1	425,556	62					
	ISAM2	255,937	62					
	NLSMA	37,937	62					
	PELLEREY	76,889	62					

a. Computed using alpha = ,05

b. R Squared = ,299 (Adjusted R Squared = ,288)

c. R Squared = ,063 (Adjusted R Squared = ,048)

d. R Squared = ,018 (Adjusted R Squared = ,002)

e. R Squared = ,000 (Adjusted R Squared = -,016)

f. R Squared = ,130 (Adjusted R Squared = ,116)

g. R Squared = ,061 (Adjusted R Squared = ,045)

h. R Squared = ,467 (Adjusted R Squared = ,458)

i. R Squared = ,190 (Adjusted R Squared = ,176)

Figure 8.5 : Effets de la source de variation « probabilité de reússite » sur les différentes taxonomies

8.3 Interprétation psychologique de notre taxonomie

Nous nous étions également proposé d'examiner si les trois catégories de seond ordre de notre taxonomie, à savoir la représentation spatiale, les connaissances procédurales de calcul et le raisonnement mathématique, peuvent être interprétés en termes de compétences au sens chomskyien du terme. Rappelons les caractéristiques d'une telle définition des compétences (Reboul, 1995) :

- Elles s'appuient sur un code qui est essentiellement restrictif ou négatif, montrant ce qui est défendu mais laissant par ailleurs une grande marge de liberté.
- Elle permettent de produire des performances en nombre illimité et de façon imprévisible.
- Les performances produites doivent être cohérentes entre elles et adaptées à la situation.

Nous avons établi que notre taxonomie ne dépend pas seulement du contenu des items, mais qu'il y a une interférence significative des facteurs covariants « difficulté », « discrimination » et « probabilité de réussite ». En tenant compte de nos résultats et en particulier de l'influence significative de la difficulté et de la probabilité de réussite des items, nous pouvons aisément présupposer, qu'à un degré de difficulté élevé des items, les trois catégories de notre taxonomie correspondent à des compétences authentiques, dans le sens où elles permettent de résoudre des problèmes en nombre illimité, de manière imprévisible et créatrice , mais en se basant sur un code établi et en tenant compte de la situation.

Lorsque les problèmes sont faciles, par contre, l'on peut concevoir qu'ils n'exigent pas l'entrée en action de réelles compétences, mais uniquement de savoirs et de savoir-faire ; qu'ils puissent donc être résolus de manière purement mécanique par des schémas appris. Le comportement de nos élèves dans les situations d'examen réelles peut d'ailleurs aisément corroborer cette interprétation.

Il est facile de comprendre que le facteur « raisonnement mathématique » puisse correspondre à une compétence au sens chomskyien. En est-il également ainsi pour la représentation spatiale et pour les connaissances procédurales de calcul ? Nous pensons que c'est théoriquement possible, mais qu'en pratique, au niveau de la tranche d'âge de nos sujets, les problèmes d'arithmétique et de géométrie élémentaires sont rarement résolus en faisant appel à des compétences réelles.

Rappelons à ce sujet un résultat tout à fait intéressant de notre analyse factorielle : le facteur de raisonnement mathématique est corrélé négativement avec la représentation spatiale et les connaissances procédurales de calcul. Celà peut signifier que les mêmes problèmes peuvent être résolus par des stratégies différentes et que les élèves qui réfléchissent personnellement ont moins tendance à employer aveuglément des schémas appris.

Pour trancher la question si nos trois catégories peuvent être considérées comme des compétences authentiques à cette tranche d'âge, il faudrait faire une étude avec des épreuves plus difficiles que celles utilisées dans la réalité scolaire quotidienne, ne pouvant pas être résolues de manière mécanique, mais nécessitant la découverte de stratégies originales.

8.4 Conclusion

Etablir une taxonomie en mathématiques implique qu'il faut tenir compte de la recherche fondamentale en sciences cognitives, en psychologie développementale et en psychologie de l'éducation.

Un aperçu sur la littérature consacrée aux résultats de la recherche fondamentale et appliquée dans ces différents domaines nous a permis de donner une idée de la complexité du problème. Une taxonomie est seulement valable pour une certaine tranche d'âge et une certaine culture, puisque la structure cognitive sous-jacente tend à changer et à se différencier, sous l'impact de la maturation biologique et des influences socio-culturelles et éducatives multiples.

Sur les aptitudes mathématiques détectées depuis la première moitié du vingtième siècle, c'est-à-dire l'aptitude numérique et l'aptitude de raisonnement mathématique, un assez grand consensus semble régner actuellement parmi les chercheurs. A partir de l'adolescence moyenne, l'aptitude numérique semble être définitivement distincte de l'aptitude de raisonnement mathématique. D'autres aptitudes, moins bien connues, englobant des facteurs affectifs, sont actuellement au centre de l'intérêt des chercheurs.

L'analyse des taxonomies existantes nous a montré que leur utilité comme système de classification de nos propres données est de valeur inégale. Un premier essai de traiter nos données par l'échelonnement multidimensionnel nous a montré qu'une solution de petite dimensionnalité conduit à une représentation graphique satisfaisante et explique la presque totalité de la variance de la dissimilarité des items.

Nous avons ensuite élaborée une taxonomie personnelle en combinant différents procédés statistiques multivariées à savoir l'échelonnement multidimensionnel, l'analyse hiérarchique ascendante et l'analyse factorielle.

Cette taxonomie a été vérifiée dans une dernière étape par la construction d'un test expérimental comportant plusieurs items dans chaque catégorie de la taxonomie théorique proposée, en utilisant le LISREL (*Linear Structural Relations*). Grâce à cette procédure nous avons démontré que l'adéquation entre la matrice de covariance observée et théorique était suffisante. Ensuite, nous avons effectué une analyse de la variance, pour examiner l'influence des deux paramètres du modèle de Birnbaum, ainsi que du facteur empirique de la probabilité de réussite des items sur les différentes taxonomies de la littérature, ainsi que sur notre taxonomie personnelle. Nous avons ainsi montré que sa structure factorielle dépend effectivement des trois facteurs examinés et qu'elle répond donc aux exigences posées.

Nous avons proposé une interprétation de nos trois catégories de second ordre, à savoir le raisonnement mathématique, la représentation spatiale et les connaissances procédurales de calcul, en terme de compétences chomskyennes, en tenant compte du degré des difficultés des items.

Notre recherche a cependant été limitée par la banque de données dont nous disposions. Pour pouvoir situer notre classification de compétences en mathématiques par rapport au fonctionnement cognitif général, il serait intéressant de passer une batterie de tests plus vaste, comprenant à la fois des épreuves de type scolaire et des épreuves non scolaires, correspondant à celles utilisées dans les tests d'intelligence classiques. De même, on pourrait inclure des épreuves permettant de mesurer l'interférence des facteurs émotionnels et motivationnels sur les performances en mathématiques. Comme nous l'avons montré dans notre partie théorique, l'interaction entre le fonctionnement affectif et intellectuel est au centre de la recherche actuelle en sciences cognitives et en sciences de l'éducation.

Cette piste d'études représenterait une continuation prometteuse de notre recherche.

Chapitre IX

ETUDE DE VALIDATION DE NOTRE LOGICIEL

9.1 Introduction

Dans ce chapitre, nous soumettons le test adaptatif informatisé en mathématiques, développé au cours du projet de recherche postdoctoral BFR 98/027, 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 inhérentes aux tests papier-crayon.

9.1.1 Hypothèses

Le but de l'étude de valorisation est de montrer que le test adaptatif informatisé donne des résultats comparables du point de vue de l'évaluation des performances des élèves, tout en réalisant un gain de temps important et en épargnant aux élèves des situations trop difficiles, donc décourageantes, respectivement trop faciles, donc ennuyeuses. Nous nous sommes intéressés aux résultats globaux, à la capacité de classification et à la capacité de dépistage des difficultés individuelles des élèves. Nous avons donc voulu tester les hypothèses suivantes :

Hypothèse 1 : Les scores et les taux de réussite obtenus dans les deux conditions sont comparables.

Hypothèse 2 : L'ordre de classification est sensiblement le même.

Hypothèse 3 : La capacité de dépistage des difficultés individuelles des élèves est comparable.

De plus, nous avons analysé l'influence de l'ordre de passation de ces deux tests, ce qui nous fournit notre quatrième hypothèse, basée sur les qualités intrinsèques du test adaptatif informatisé, qui font que la situation d'apprentissage n'est pas la même.

Hypothèse 4 : L'ordre de passation a une influence sur les résultats.

9.1.2 Construction de la forme expérimentale du test

Pour commencer, nous avons changé notre logiciel selon les proposition des instituteurs afin de faciliter son emploi et de permettre l'utilisation des procédés de statistiques inférentielles prévus. Le logiciel crée maintenant pour chaque élève un fichier Word dans lequel sont notés, outre son nom et prénom les numéros des items qu'il a résolu, les réponses qu'il a données, ainsi que sa note finale.

9.1.3 Description du dispositif expérimental

Pour comparer les qualités du test informatisé à celles du test papier-crayon, nous avons présenté les deux formes à un échantillon de 6 classes d'élèves de septième du Lycée Michel Rodange, ce qui a fait un total de 123 élèves. Afin de neutraliser l'effet de l'ordre de passation et de répétition, nous avons attribué ces élèves au hasard à 2 sous-groupes de même taille, dont l'un devait commencer par le test papier-crayon et l'autre par le test informatisé. Les élèves qui ont commencé par le test papier-crayon ont effectué le test sur ordinateur deux mois plus tard et vice versa. Malheureusement, des problèmes d'installation de notre logiciel sur les ordinateurs du Lycée ont eu pour conséquence que, pendant les deux premières séances, tous les élèves d'une classe ont dû faire la version papier-crayon. Ceci explique que, finalement, les deux sous-groupes ont des effectifs différents.

De cette manière, nous avons pu constituer un fichier avec 123 sujets, comportant comme variables la note du test sur ordinateur (variable *note*), la note du test papier-crayon (variables *notepap*), ainsi que deux fois 58 variables dichotomes indiquant si les élèves ont bien répondu aux différents items des deux tests.

Pour compléter le test sur ordinateur, les élèves ont dû répondre à un nombre d'items compris entre sept à quatorze, alors que le test papier-crayon a nécessité la réponse à la totalité des 58 items. Le test sur ordinateur réalise donc effectivement une économie de temps appréciable. Ce résultat permet donc de prévoir un gain de temps appréciable si l'on devait utiliser des batteries de tests plus larges.

Nous avons scindé l'étude statistique en trois parties. Dans la première partie, nous analysons s'il existe une différence significative entre les scores obtenus par les tests sur ordinateurs, respectivement sous la forme papier-crayon (hypothèse 1) ainsi qu'entre les classements effectués (hypothèse 2), dans la deuxiéme partie, nous avons analysé l'effet de l'ordre de passation sur les scores (hypothèse 4) et dans la troisième partie, nous avons analysé certains items, pour voir s'il existe, entre les deux tests, des différences dans le taux de réussite à ces items particuliers (hypothèse 3).

9.2 Comparaison des scores et capacité de classification

Nous commencons par quelques statistiques descriptives des variables *note* et *notepap* qui contiennent respectivement les notes des tests sur ordinateur et des tests papier-crayon.

Descriptive Statistics

	Mean	Std. Deviation	N
NOTEPAP	,6520	,6140	123
NOTE	,5507	,4671	123

Figure 9.1: Statistiques descriptives des scores

On constate que les moyennes sont assez proches (0,652 pour la version papier-crayon contre 0,5507 pour la version ordinateur) et que le test papier-crayon, qui admet une moyenne un peu plus élevée, a également une variance, donc une dispersion plus grande. Nous examinerons plus loin si ces différences sont significatives.

Ensuite, nous établissons la matrice de corrélations entre les deux situations expérimentales.

Correlations

		NOTEPAP	NOTE
Pearson Correlation	NOTEPAP	1,000	,493**
	NOTE	,493**	1,000
Sig. (2-tailed)	NOTEPAP	,	,000
	NOTE	,000	,
N	NOTEPAP	123	123
	NOTE	123	123

**. Correlation is significant at the 0.01 level (2-tailed).

Figure 9.2 : Corrélations entre les 2 situations expérimentales

La corrélation entre les résultats du test sur ordinateur et du test papier-crayon est hautement significative (seuil 0,01). Le pouvoir de classification de notre test informatisé est donc comparable à celui du test papier-crayon. L'hypothèse 2 a donc pu être vérifiée.

Puis, nous effectuons un test de comparaison des moyennes des deux variables de scores. Les histogrammes de ces deux variables ci-dessous suggèrent une distribution normale.

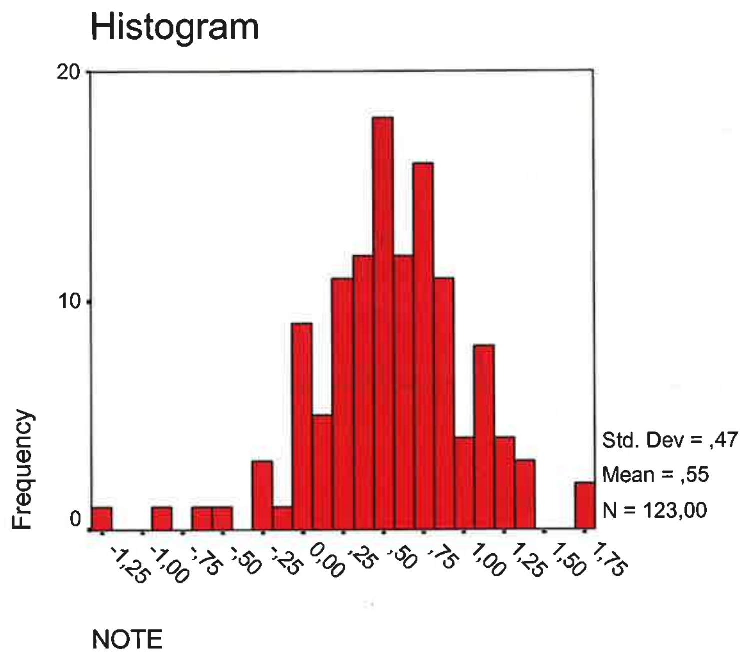


Figure 9.3 : Histogramme de la variable *note*

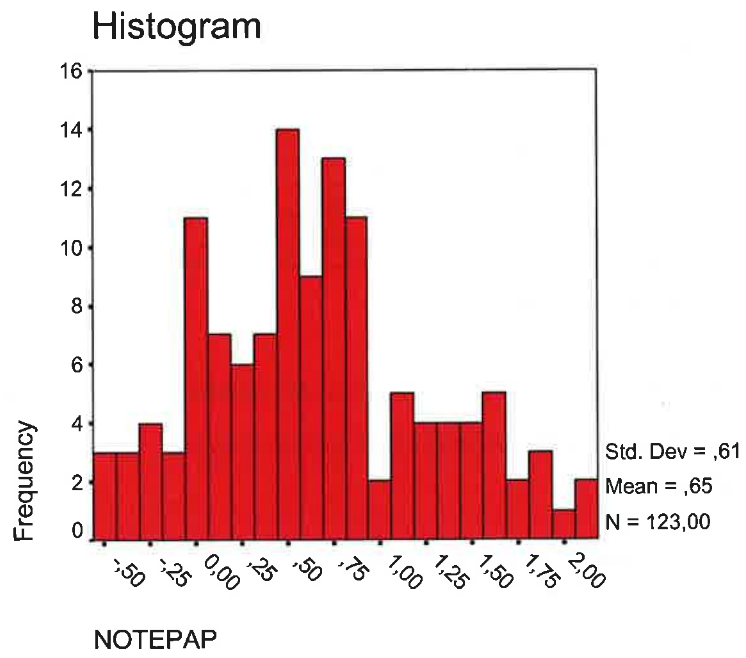


Figure 9.4 : Histogramme de la variable *notepap*

Pour vérifier cette impression, nous effectuons un test de Kolmogorov-Smirnov, qui confirme effectivement l'adéquation suffisante à la distribution normale, puisque les indices de significativé sont de 0,654 pour la variable *notepap* et de 0,348 pour la variable *note*.

One-Sample Kolmogorov-Smirnov Test

		NOTEPAP	NOTE
N		123	123
Normal Parameters[a,b]	Mean	,6210	,5345
	Std. Deviation	,6296	,4835
Most Extreme Differences	Absolute	,064	,081
	Positive	,064	,046
	Negative	-,051	-,081
Kolmogorov-Smirnov Z		,734	,934
Asymp. Sig. (2-tailed)		,654	,348

a. Test distribution is Normal.

b. Calculated from data.

Figure 9.5 : Résultat du test de Kolmogorov-Smirnov

Le test de comparaison des moyennes approprié est donc le test de Student pour échantillons appareillés.

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
					95% Confidence Interval of the Difference				
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper			
Pair 1	NOTEPAP - NOTE	8.65E-02	,6320	5.48E-02	-2.2E-02	,1949	1,578	122	,117

Figure 9.6 : Résultat du test de comparaison des moyennes

Il n'y a pas de différence significative entre les moyennes obtenues dans la situation expérimentale du test sur ordinateur et du test papier-crayon.

Finalement, nous comparons encore la fréquence de réussite dans les deux tests. On trouve un taux d'échec de 13,8 % pour la version papier-crayon contre 8,9 % pour la version ordinateur. En effctuant un test du chi deux, on trouve une distance de $\chi^2 = 2,65$, ce qui, pour une loi à un degré de liberté, n'est pas significatif au seuil de 5%. Les résultats et le pourcentage de réussite ne sont pas significativement différents dans la situation ordinateur et la situation papier-crayon.

L'hypothèse 1 est donc également vérifiée.

9.3 Etude de l'effet de l'ordre de passation

Dans cette partie, nous analysons l'effet de l'ordre de passation des deux tests sur les résultats. Pour cela, nous avons crée une variable dichotomique appelée *groupe* qui vaut 1 pour les élèves qui ont d'abord passé le test sur ordinateur et 2 pour ceux qui ont commencé par le test papier-crayon.

Group Statistics

	ordre de passage	N	Mean	Std. Deviation	Std. Error Mean
NOTE	ordinateur-papier	42	,5674	,4080	6,30E-02
	papier-ordinateur	81	,5420	,4972	5,52E-02
NOTEPAP	ordinateur-papier	42	1,0093	,5580	8,61E-02
	papier-ordinateur	81	,4668	,5601	6,22E-02

Figure 9.8 : Statistiques descriptives des scores suivant la variable *groupe*

En examinant les moyennes des scores suivant les deux sous-groupes ainsi créés, on remarque que la note moyenne du test sur ordinateur ne varie pas beaucoup suivant l'ordre de passation des tests (0,5674 pour ceux qui ont commencé avec le test sur ordinateur contre 0,5420 pour ceux qui ont commencé avec le test papier-crayon), alors que celle du test papier-crayon est plus que deux fois plus grande (1,0093 contre 0,4668) pour le sous-groupe des élèves qui ont commencé avec le test sur ordinateur.

Le test de Student confirme cette impression.

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Mean	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
NOTE	Equal variances assumed	,801	,373	,285	121	,776	2,541E-02	8,916E-02	-,1511	,2019
	Equal variances not assumed			,303	98,506	,762	2,541E-02	8,376E-02	-,1408	,1916
NOTEPAP	Equal variances assumed	,415	,520	5,100	121	,000	,5425	,1064	,3319	,7531
	Equal variances not assumed			5,106	83,362	,000	,5425	,1062	,3312	,7538

Figure 9.9 : Résultat du test de comparaison des moyennes

En effet, le test de Student montre effectivement que si la situation expérimentale « ordinateur » est présentée en premier lieu, elle ne conduit pas à une différence significative entre les moyennes, alors que, si la situation « papier-crayon » est présentée d'abord, il y a une différence hautement significative entre les moyennes. Cela permet de conclure qu'il y a eu un effet d'apprentissage positif de l'ordinateur à la situation papier-crayon, mais non vice-versa.

Nous effectuons aussi une analyse de la variance multivariée pour vérifier ce résultat, en prenant comme variables dépendantes les variables *note* et *notepap* et comme facteur la variable dichotomique *groupe*.

Le choix de la méthode a été dicté par les considérations suivantes : Notre plan expérimental correspond en fait à un plan factoriel (ou systématique) mixte (Delhomme & Meyer, 1997).

Deux groupes indépendants (facteurs intersujets) de sujets ont été soumis à deux conditions expérimentales (facteurs intrasujets). En effet, si le facteur *groupe* est constitué par l'ordre de passation, tous les sujets n'ont pas participé à ses deux modalités, ce qui nous permet de les séparer en deux groupes indépendants, selon les modalités 1 et 2. Par contre, tous les sujets ont été soumis aux deux modalités du facteur *méthode* (genre du test).

L'analyse de la variance à effectuer correspond donc en fait à la figure d'une analyse de la variance bifactorielle de type mesures mixtes (*two-factor mixed means analysis of variance*) (cf. Turner & Thayes, 2001). Ce plan comprend 4 conditions expérimentales issues du croisement de deux variables à deux modalités chacune, chaque sujet ayant seulement été soumis à la moitié des conditions possibles.

Or, le logiciel SPSS for Windows 7.5 n'est pas directement adapté à ce type de plan expérimental. Pour utiliser le modèle « mesures répétées », il faudrait que tous les sujets aient été soumis aux quatre conditions expérimentales, ce qui n'est pas le cas dans notre étude.

Nous avons donc été réduit à effectuer une analyse unifactorielle multivariée, nous permettant de mesurer l'effet du facteur intersujet *groupe* sur les variables dépendantes constituées par les scores obtenus dans les deux conditions expérimentales ordinateur et papier-crayon.

Tests of Between-Subjects Effects

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Noncent. Parameter	Observed Power[a]
Corrected Model	NOTE	1.79E-02[b]	1	1.79E-02	.081	.776	.081	.059
	NOTEPAP	8.140[c]	1	8.140	26.015	.000	26.015	.999
Intercept	NOTE	34.039	1	34.039	154.827	.000	154.827	1.000
	NOTEPAP	60.262	1	60.262	192.596	.000	192.596	1.000
GROUPE	NOTE	1.79E-02	1	1.79E-02	.081	.776	.081	.059
	NOTEPAP	8.140	1	8.140	26.015	.000	26.015	.999
Error	NOTE	26.602	121	.220				
	NOTEPAP	37.860	121	.313				
Total	NOTE	63.915	123					
	NOTEPAP	98.293	123					
Corrected Total	NOTE	26.620	122					
	NOTEPAP	46.000	122					

a. Computed using alpha = .05

b. R Squared = .001 (Adjusted R Squared = -.008)

c. R Squared = .177 (Adjusted R Squared = .170)

Figure 9.10 : Résultat de l'analyse de variance

Le résultat confirme les résultats précédents. Il existe bien une influence significative de l'ordre de passation sur les notes du test papier-crayon, mais pas sur les notes du test sur ordinateur. En d'autres termes, nous avons pu mettre en évidence un effet d'apprentissage de l'ordinateur sur la situation papier-crayon, mais non inversément. L'hypothèse 4 a donc été vérifiée ; l'ordre de passation a un effet significatif sur les résultats dans le sens où il y a un effet d'apprentissage positif si les items sont d'abord présentés par l'ordinateur.

Au-delà de la contribution théorique à l'étude de validation du test adaptatif informatisé, ce résultat permet de souligner l'utilité pédagogique d'un outil de ce type. Pour l'interpréter, il faudrait tenir compte de la composante motivationnelle. Une étude englobant les facteurs non intellectuels intervenant dans les épreuves de performance aurait dépassé le cadre de ce travail, mais en observant les élèves et en écoutant leurs commentaires, il nous a paru évident que le test sur ordinateur a suscité plus de motivation intrinsèque parmi eux. Etudier les processus affectifs et motivationnels associés à la situation ordinateur et ne se rencontrant pas de la même manière dans la situation papier-crayon constituerait un prolongement intéressant de notre étude.

9.4 Analyse de quelques items

Finalement, nous avons encore examiné s'il y a des différences entre la fréquence de réussite aux items particuliers effectués sur papier ou sur ordinateur. Comme l'ordinateur administre lui-même les questions lors du test adaptatif et que chaque élève doit répondre en principe à une série d'items adaptée à sa capacité, il est clair que tous les items n'ont pas été présentés assez souvent pour pouvoir être soumis à une analyse statistique. Nous nous sommes donc limité aux cinq items qui ont été administré à presque tout le monde, à savoir les items numéros 26, 32, 38, 47 et 52. Ce sont des items de difficulté moyenne qui discriminent bien entre les sujets. Nous avons fait un test de comparaison de la médiane de Wilcoxon pour analyser s'il y a une différence significative du score de réussite de ces items si on les effectue sur ordinateur, respectivement sur papier.

Test Statistics[c]

	P26 - Q26	P32 - Q32	P38 - Q38	P47 - Q47	P52 - Q52
Z	-1,260[a]	-1,886[b]	-,447[a]	-,849[a]	-1,000[a]
Asymp. Sig. (2-tailed)	,208	,059	,655	,396	,317

a. Based on positive ranks.

b. Based on negative ranks.

c. Wilcoxon Signed Ranks Test

Figure 9.11 : Résultat du test de Wilcoxon

Le résultat de ce test est tout à fait positif. Au seuil de 5 %, il n'y a pour aucun item une différence significative entre la réussite sur ordinateur respectivement sur papier, ce qui montre que les élèves de septième n'ont en général pas plus de problèmes de résoudre un test de mathématiques sur ordinateur que sur papier et que, si l'on fait une analyse différentielle par rapport aux items réussis et non réussis, il n'y a pas de différence entre les deux méthodes de présentation.

Remarquons cependant que pour l'item 32, le résultat est moins concluant que pour les quatre autres. En fait, l'item 32 était assez mal adapté pour un test sur ordinateur. C'était l'item suivant :

Item numéro 32 :

In der Samenhandlung hängt folgende Preisliste aus:

25 Erdbeerpflanzen:	200 F
20 g Gemüsesamen:	35 F
25 Kohlpflanzen:	87 F
Zwiebeln pro Kilogramm:	80 F

Der Lehrling hat für einen Kunden folgende Rechnung aufgestellt. Vervollständige sie!

Rechnung Nr: 3587 vom 10. April 1996 für Frau Irene Müller

Artikel		**Betrag**
3 Tüten Kunstdünger	$3 \cdot 275$	**825**
0,5 kg Zwiebeln		.....

Les instructions étaient de ne taper que ce qui devrait se trouver à l'endroit des cinq points. La bonne réponse était donc de taper simplement le nombre 40. Il y avait cependant un certain nombre d'élèves qui ont tapé $2 \times 40 = 80$, ce que l'ordinateur a considéré comme mauvaise réponse.

L'hypothèse 3 a donc également été vérifiée : la probabilité de réussite des items particuliers et donc la possibilité de dépistage des difficultés individuelles est comparable dans les 2 situations.

9.5 Conclusion

Les qualités psychométriques intrinsèques d'un test adaptatif informatisé dépendent de son mode de construction et sont liés aux modèles IRT. Nous les avons largement exposés dans notre rapport du projet de recherche postdoctoral BFR 98/027 (Schiltz, 1999).

Dans notre étude de valorisation nous n'avions plus besoin de revenir sur ce point, mais nous avons voulu comparer l'efficience de notre logiciel en matière d'évaluation des performances des élèves avec celle de la situation standard papier-crayon à laquelle ils ont été habitués depuis le début de leur scolarité. Nous sommes partis de l'idée que, si les capacités de mesure des performances, de classification et de discrimination se montrent comparables, notre logiciel a prouvé sa supériorité grâce à ses qualités de puissance, de flexibilité et d'économie inhérents à sa construction et par ses effets psychologiques favorables, liés au fait qu'il évite de décourager et d'ennuyer inutilement les élèves. Nos attentes générales ont été formulées en quatre hypothèses opérationnelles, qui ont été soumises à la vérification statistique.

Nos quatre hypothèses ont été vérifiées, en comparant notre test informatisé à la forme standard papier-crayon, ce qui constitue une confirmation indirecte de la validité interne de notre test. Les limites temporelles assignées à notre étude ne nous ont pas permis d'étudier la validité externe de notre test, en utilisant par exemple les scores scolaires obtenus par les élèves à la fin de l'année ou les résultats d'un test d'intelligence classique.

Puisque ces résultats sont concluants, l'utilité du test adaptatif informatisé que nous avons construit a été 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.

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 de réaliser 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.

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 représentation spatiale, en connaissances procédurales et en raisonnement mathématiques, 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.

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.

En général, les liens entre la recherche fondamentale, par exemple au niveau de l'établissement d'une structure hiérarchique du fonctionnement cognitif et entre la recherche appliquée en pédagogie, par exemple au niveau de l'élaboration d'épreuves d'évalluation, devraient être plus serrés.

Notre recherche permet de nombreux prolongements ultérieurs dans le domaine du fonctionnement cognitif et dans le domaine des applications pédagogiques.

Annexe

LA BANQUE D'ITEMS DU LOGICIEL

On donne ci-dessous la liste des items retenus pour faire partie de la banque d'items du logiciel.

Item numéro 1 :

Berechne im Kopf.

51+37+53+149=

Item numéro 2 :

Berechne im Kopf.

53+527=

Item numéro 3 :

Berechne im Kopf.

90:5=

Item numéro 4 :

Berechne im Kopf.

1999-430=

Item numéro 5 :

Berechne im Kopf.

10·7·0·20=

Item numéro 6 :

Berechne im Kopf.

400·12=

Item numéro 7 :

Berechne im Kopf.

21+9:3=

Item numéro 8 :

Berechne im Kopf.

(8+40):4=

Item numéro 9 :

Berechne im Kopf.

609+……=700

Item numéro 10 :

Berechne im Kopf.

……-429=2999

Item numéro 11 :

Mache als Tafelrechnung am freigelassenen Platz.

17321-2735-87

Item numéro 12 :

Mache als Tafelrechnung am freigelassenen Platz.

27·483

Item numéro 13 :

Mache als Tafelrechnung am freigelassenen Platz.

5208:14

Item numéro 14 :

Mache einen Überschlag! Kreuze jeweils die Zahl an, welche dem Resultat der angeschriebenen Rechnung am nächsten kommt.

72·31

O 210 O 2100 O 750 O 21000 O Keine Ahnung

Item numéro 15 :

Mache einen Überschlag! Kreuze jeweils die Zahl an, welche dem Resultat der angeschriebenen Rechnung am nächsten kommt.

8023:98

O 8000 O 80 O 8 O 800 O Keine Ahnung

Item numéro 16 :

Setze das richtige Zeichen (>, <, =).

1,91……1,191

Item numéro 17 :

Setze das richtige Zeichen (>, <, =).

0,01……0,010

Item numéro 18 :

Setze das richtige Zeichen (>, <, =).

$\frac{7}{12} \ldots\ldots \frac{7}{11}$

Item numéro 19 :

Setze das richtige Zeichen (>, <, =).

$\frac{7}{12} \ldots\ldots \frac{5}{8}$

Item numéro 20 :

Setze das richtige Zeichen (>, <, =).

$\frac{70}{120} \ldots\ldots \frac{10}{16}$

Item numéro 21 :

Setze das richtige Zeichen (>, <, =).

$\frac{3}{4}$ h……45min

Item numéro 22 :

Ergänze.

2 m = …… dm

Item numéro 23 :

Ergänze.

3,2 kg = …… g

Item numéro 24 :

Ergänze.

1 min 12 s = …… s

Item numéro 25 :

Ergänze.

2500 kg = 2,5 ……

Item numéro 26 :

Ergänze.

3 cm 2 mm = …… mm

Item numéro 27 :

Hier sind fünf Zahlen:

22738 **76372** **21670** **13655** **35742**

Welche von ihnen ist teilbar durch 4? ……

Item numéro 28 :

Hier sind fünf Zahlen:

22738 **76372** **21670** **13655** **35742**

Welche von ihnen ist teilbar durch 3 ?……

Item numéro 29 :

Hier sind fünf Zahlen:

22738 **76372** **21670** **13655** **35742**

Welche von ihnen ist gleichzeitig teilbar durch 2 **und** durch 5 ? ……

Item numéro 30 :

Die Zahl 6352 ist nicht teilbar durch 9. Wieviel muβ man mindestens zu der Zahl 6352 hinzuzählen damit sie durch 9 teilbar wird?

Deine Antwort: ……

Item numéro 31 :

In der Samenhandlung hängt folgende Preisliste aus:

Kunstdünger pro Tüte:	275 F
25 Erdbeerpflanzen:	200 F
20 g Gemüsesamen:	35 F
25 Kohlpflanzen:	87 F
Zwiebeln pro Kilogramm:	80 F

Der Lehrling hat für einen Kunden folgende Rechnung aufgestellt. Vervollständige sie!

Rechnung Nr: 3587 vom 10. April 1996 für Frau Irene Müller

Artikel		**Betrag**
3 Tüten Kunstdünger	3·275=	

Item numéro 32 :

In der Samenhandlung hängt folgende Preisliste aus:

Kunstdünger pro Tüte:	275 F
25 Erdbeerpflanzen:	200 F
20 g Gemüsesamen:	35 F
25 Kohlpflanzen:	87 F
Zwiebeln pro Kilogramm:	80 F

Der Lehrling hat für einen Kunden folgende Rechnung aufgestellt. Vervollständige sie!

Rechnung Nr: 3587 vom 10. April 1996 für Frau Irene Müller

Artikel	**Betrag**
0,5 kg Zwiebeln	

Item numéro 33 :

Hier sind zwei Brüche : $\frac{4}{5}$ und $\frac{5}{4}$

Welcher der beiden Brüche ist größer als 1 ? ......

Item numéro 34 :

Hier sind zwei Brüche : $\frac{4}{5}$ und $\frac{5}{4}$

Welcher der beiden Brüche liegt auf dem Zahlenstrahl am nächsten bei 1 ? ......

Item numéro 35 :

Hier ist eine Zahlenfolge. Suche die folgende Zahl.

569 548 527 506 ......

Item numéro 36 :

Hier ist eine Zahlenfolge. Suche die folgende Zahl.

14 8 16 10 20 14 28 ……

Item numéro 37 :

Ist der folgende Satz wahr oder falsch? Kreuze an!

57 ist eine Primzahl	O wahr	O falsch

Item numéro 38 :

Ist der folgende Satz wahr oder falsch? Kreuze an!

Das Produkt von zwei Primzahlen ist nie eine Primzahl	O wahr	O falsch

---

Item numéro 39 :

Ist der folgende Satz wahr oder falsch ? Kreuze an !

Die Diagonalen eines Rechteckes stehen senkrecht zueinander	O wahr	O falsch

---

Item numéro 40 :

Ist der folgende Satz wahr oder falsch? Kreuze an!

Die gegenüberliegenden Seiten eines Rechtecks sind parallel	O wahr	O falsch

---

Item numéro 41 :

Ist der folgende Satz wahr oder falsch? Kreuze an!

Ein Würfel hat 6 quadratische Seitenflächen	O wahr	O falsch

Item numéro 42 :

Hier ist die Breite eines Rechteckes:

Vervollständige das Rechteck, wenn du weißt, daß seine Länge doppelt so groß wie die Breite ist.

Item numéro 43 :

Wenn du weißt, daß $\boxed{\mathbf{72 \cdot 28 = 2016}}$ ergibt, ergänze folgende Rechenaufgabe, ohne zusätzliche schriftliche Berechnungen.

$$\mathbf{36 \cdot 28 =}$$

Item numéro 44 :

Du darfst die Ziffern $\boxed{7}$ $\boxed{2}$ $\boxed{0}$ $\boxed{9}$ und das Komma jeweils einmal benutzen.

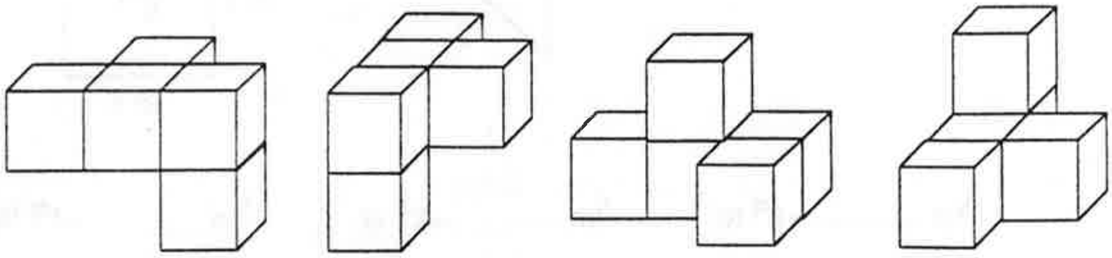
Welches ist die kleinste <u>Dezimalzahl</u>, die du mit diesen Zeichen bilden kannst?

Deine Antwort: ……

Item numéro 45 :

Hier sind 5 Würfel aus verschiedenen Blickwinkeln gesehen. Welche Figur unterscheidet sich von den anderen?

Umkreise die richtige Antwort.

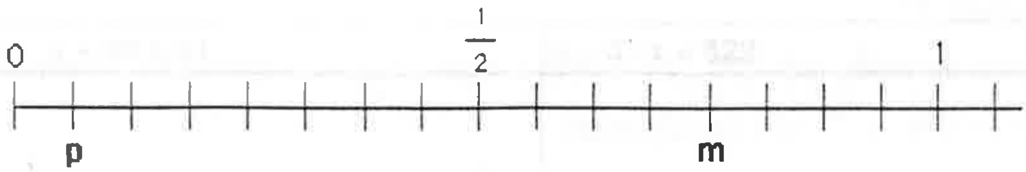


Figur 1 Figur 2 Figur 3 Figur 4

Keine Figur **1** Figur **2** Figur **3** Figur **4**

Item numéro 46 :

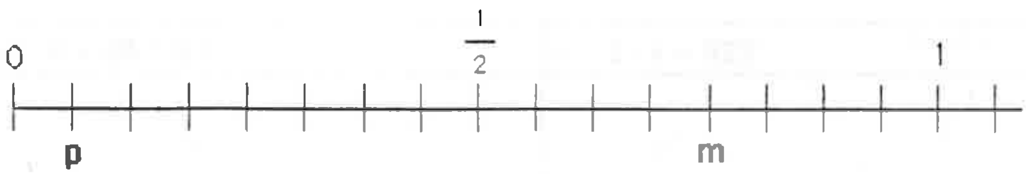
Bestimme den Bruch, welcher durch den Buchstaben m auf dem Zahlenstrahl dargestellt ist.



m = ……

Item numéro 47 :

Bestimme den Bruch, welcher durch den Buchstaben p auf dem Zahlenstrahl dargestellt ist.

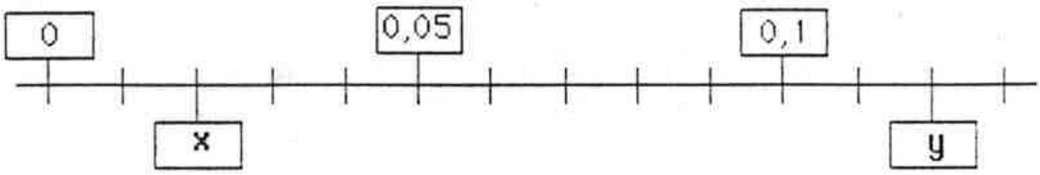


p = ……

Item numéro 48 :

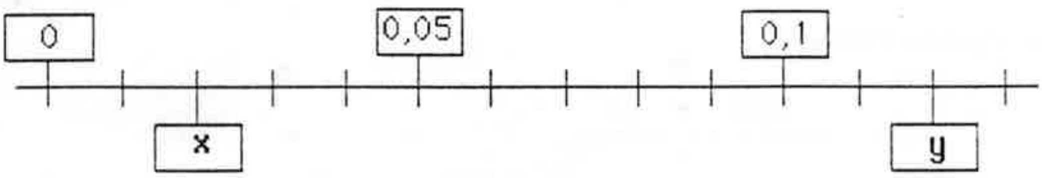
Bestimme die Dezimalzahl, welche durch den Buchstaben y auf dem Zahlenstrahl dargestellt ist?

y = ……



Item numéro 49 :

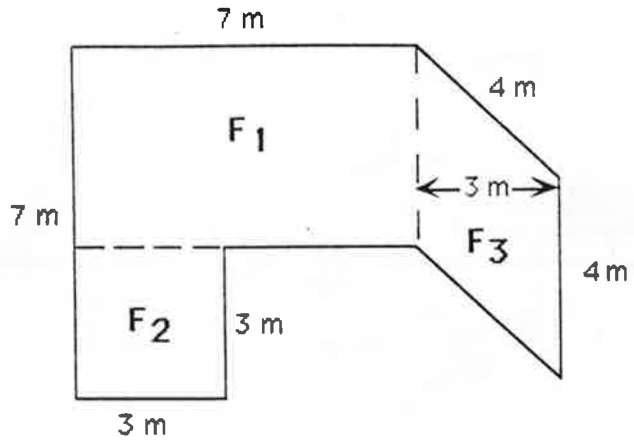
Bestimme die Dezimalzahl, welche durch den Buchstaben x auf dem Zahlenstrahl dargestellt ist?



x = ……

Item numéro 50 :

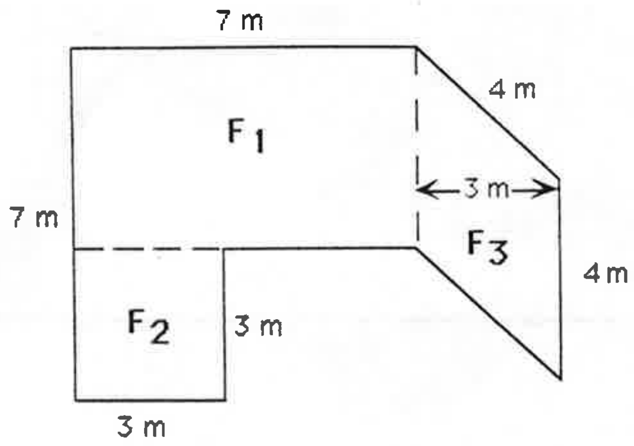
Berechne die Fläche F_1.



$F_1 =$ ...... m^2

Item numéro 51 :

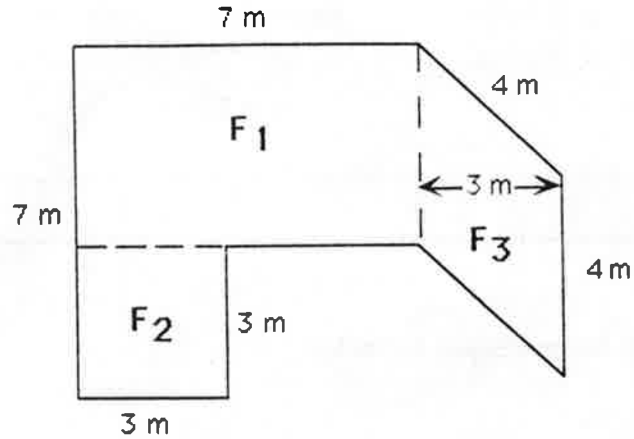
Berechne die Fläche F_2.



$F_2 =$ ...... m^2

Item numéro 52 :

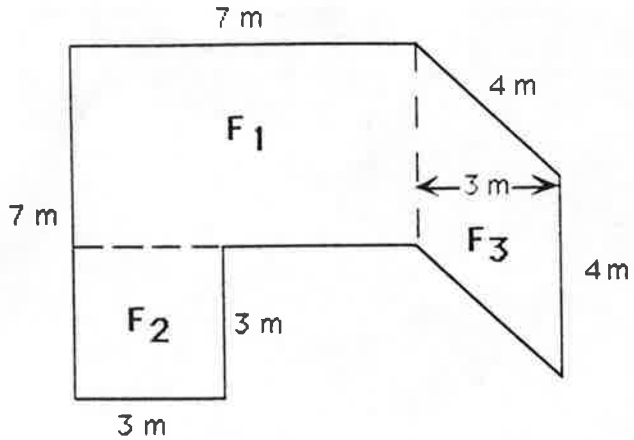
Berechne die Fläche F_3.



$F_3 = \ldots\ldots\ m^2$

Item numéro 53 :

Berechne den Umfang der Figur.



Der Umfang der Figur beträgt …… m.

Item numéro 54 :

Löse folgende Gleichung in dem vorgesehenen Kasten.

x + 26 = 81

Item numéro 55 :

Löse folgende Gleichung in dem vorgesehenen Kasten.

Item numéro 56 :

Zeichne die Strecke [AB].

Item numéro 57 :

Miß und stelle fest: Die Strecke [AB] ist …… cm lang.

Item numéro 58 :

Zeichne die Halbgerade [CA)

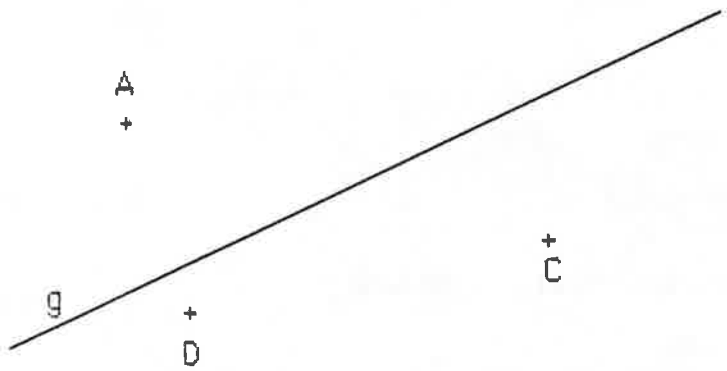
---

42) Item numéro 59 :

Konstruire zu der Geraden **g** die Senkrechte durch den Punkt **C**.

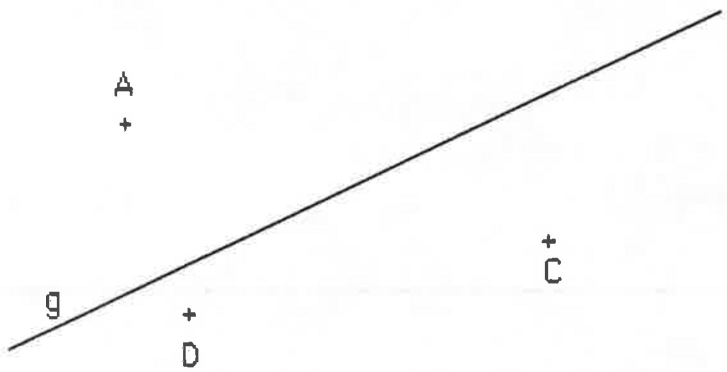
Item numéro 60 :

Zeichne, miß und stelle fest:



Der Abstand des Punktes **A** von der Geraden **g** beträgt …… cm.

Item numéro 61 :



Zeichne die Parallele zu **g** durch den Punkt **D**.

Item numéro 62 :

Diese Eintrittsscheine wurden während drei Tagen an der Kinokasse verkauft:

	Montag	Dienstag	Mittwoch
Der erste verkaufte Schein trug die Nummer	120	387	501
Der letzte verkaufte Schein trug die Nummer	386	500	713

Wieviele Eintrittsscheine wurden am Montag verkauft? ……

Item numéro 63 :

Diese Eintrittsscheine wurden während drei Tagen an der Kinokasse verkauft:

	Montag	Dienstag	Mittwoch
Der erste verkaufte Schein trug die Nummer	120	387	501
Der letzte verkaufte Schein trug die Nummer	386	500	713

Wieviele der Eintrittsscheine, die am Dienstag verkauft wurden, trugen **wenigstens eine** Ziffer 4? ……

BIBLIOGRAPHIE

Adèr H.J. & Mellenbergh G.J., (1999*), Research Methodology in the Social, Behavioural & Life Sciences*, London : SAGE Publications.

Albers D.J. & Alexanderson G.L., (1985), *Mathematical people : Profiles and interviews*, Boston : Birkhauser.

Alderman D.L., Swinton S.S. & Braswell J.S., (1979), Assessing basic arithmetic skills and understanding across curricula : Computer-assisted instruction and compensatory education, *J. Children's Math. Behav.*, **2**, p.3-28.

Ashcraft M.H., (1992), Cognitive arithmetic : A review of data and theory, *Cognition*, **44**, p.75-106.

Ausubel D.P., (1968), *Educational psychology : A cognitive view*, New York : Holt, Rinehart & Winston.

Bacher F. & Dickes P., (1994), Les méthodes statistiques en psychologie différentielle : perspectives de développement, *Psychologie Française*, **29**, p.9-15.

Bavaud F., Capel R., Crettaz de Roten F. & Müller J.P., (1996), *Guide de l'analyse statistique de données avec SPSS 6*, Genève : Slatkine.

Benzécri J.P., (1973), *L'analyse des données*, Paris : Dunod.

Brink P.J & Wood M.J., (1998), *Advanced design in nursing research*, London : SAGE Publications.

Brownell W.A., (1935), Psychological considerations in the learning and teaching of arithmetic, Dans : W.D. Reeve (ed), *The teaching of arithmetic : The tenth yearbook of the National Council of Teachers of Mathematics*, p.1-31, New York : Teachers College Press.

Bruner J.S., (1960), *The process of education*, Cambridge, MA : Harvard University Press.

Bruner J.S., (1962), Introduction, Dans : L.S. Vigotsky, *Thought and language*, Cambridge, MA : MIT Press.

Bruner J.S., (1966), *Towards a theory of instruction*, Cambridge MA : Harvard University Press.

Bruner J.S. & Erlwanger S.H., (1973), Benny's conception of rules and answers in IPI mathematics, *J. Children's Math. Behav.*, **1**, p.7-26.

Canisia M., (1962), Mathematical ability as related to reasoning and use of symbols, *Educational and Psychological Measurement*, **22**, p.105-127.

Carpenter T.P., (1980), Cognitive development and mathematics learning, Dans : R. Shumway (ed), *Research in mathematics education*, p.146-206, Reston, VA : National Council of Teachers of Mathematics.

Carpenter T.P. & Moser J.M., (1983), The acquisition of addition and substraction concepts, Dans : R. Lesh & M. Landau (eds), *The acquisition of mathematical concepts and processes*, p.7-44, New York : Academic Press.

Carpenter T.P. & Moser J.M., (1984), The acquisition of addition and substraction concepts in grades one trough three, *Journal for Research in Mathematics Education*, **15**, p.179-202.

Carpenter T.P., Just M.A. & Shell P., (1990), What one intelligence test measures : A theoretical account of the processing in the Raven Matrices Test, *Psychological Review*, **97**, p.404-431.

Carroll J.B., (1993), *Human cognitive abilities : A survey of factor-analytic studies*, New York : Cambridge University Press.

Carroll J.B., (1996), Mathematical Abilities: Some Results From Factor Analysis, Dans : R.J. Sternberg & T. Ben-Zeev (eds), *The Nature of Mathematical Thinking*, Mahwah, NJ : Lawrence Erlbaum Associates.

Case R. & Bereiter C., (1982), *From behaviorism to cognitive development : Steps in the evolution of instructional design*, Proceedings of the Conference for Educational Technology in the 80s, Caracas, Venezuela.

Cattell R.B., (1966), The scree test for the number of factors, *Multivariate Behavioral Research*, **1**, p.245-276.

Cattell R.B., (1971), *Abilities : Their structure, growth, and action*, Boston : Houghton Mifflin.

Cattell R.B. & Horn J.L., (1978), A check on the theory of fluid and crystallized intelligence with description of new subtest designs, *Journal of Educational Measurement*, **15**, p.139-164.

Ceci S.J., (1990), *On intelligence ... more or less : A bioecological treatise on intellectual development*, Englewood Cliffs, NJ : Prentice-Hall.

Ceci S.J. & Liker J.K., (1986), A day at the races : A study of IQ, expertise, and cognitive complexity, *Journal of Experimental Psychology : General*, **115**, p.255-266.

Chein I., (1939), Factors in mental organization, *Psychological Record*, **3**, p.71-94.

Chi M.T., Glaser R. & Farr M.J. (eds), *The nature of expertise*, Hillsdale, NJ : Lawrence Erlbaum Associates.

Chomsky N., (1965), *Aspects of the theory of syntax*, Cambridge, MA : MIT Press.

Cobb P., Wood T. & Yackel E., (1991), A constructivist approach to second grade mathematics, Dans : E. von Glasersfeld (ed), *Constructivism in mathematics Education*, p.157-176, Dordrecht : Kluwer.

Cobb P., Wood T., Yackel E. & McNeal G., (1992), Characteristics of classroom mathematics traditions : An interactional analysis, *American Educational Research Journal*, **29**, p.573-602.

Cohen A.D., (1994), The hierarchy of consistency : A review and model of longitudinal findings on adult individual differences in intelligence, personality and self-opinion, *Personality and Individual Differences*, **5**, p.11-25.

Coombs C.H., (1941), A factorial study of number ability, *Psychometrika*, **6**, p.161-189.

Coxon A.P.M., (1982), *The user's guide to multidimensional scaling*, London : Heinemann.

Davydov V.V., (1990), *Types of generalization in instruction : Logical and psychological problems in the structuring of school curricula* (Soviet Studies iin Mathematics Education, vol. 2), Reston, VA : National Council of Teachers of Mathematics.

Davis R.B., (1984), *Learning mathematics : The cognitive science approach to mathematics education*, Norwood, NJ : Ablex.

Davis R.B., (1990), How computer help us understand people, *International Journal of Educational Research*, **14**(1), p.93-100.

Davis R.B. & Goffree F., (1991), Mathematics : Elementary-school Programs, Dans : A. Lewy, *The International Enceclopedia of Curriculum*, p.821-825, Oxford : Pergamon Press.

Davis R.B. & McKnight C., (1980), The influence of semantic content on algorithmic behavior, *J. Math. Behav*, **3**, p.39-87.

De Corte E., Greer B. & Verschaffel L., (1996), Mathematics Teaching and Learning, Dans : Berliner, *Handbook of Educational Psychology*, p.491-549.

De Corte E. & Verschaffel L., (1988), Computer simulation as a tool in research in problem solving in subject-matter domains, *International Journal of Educational Research*, **12**, p.49-69.

Dehaene S., Tzourio N., Frak V., Raynaud L. et al., (1996), Cerebral activities during number multiplication and comparison : A PET study, *Neuropsychologica*, **34**, p.1097-1106.

Delacour J., (1998), *Une introduction aux neurosciences cognitives*, Bruxelles : De Boeck Université.

Delannoy C. & Passegand J.C., (1992), *L'intelligence peut-elle s'éduquer ?*, Paris : Hachette Education.

Delhomme P. & Meyer T., (1997), *Les projets de recherche en psychologie sociale*, Paris : Armand Colin/Masson.

Dickes P., (1986), Du modèle exploratoire au modèle confirmatoire en sciences sociales, *Actes du Colloque « Doctrines, sciences ou pratiques sociales »*, p. 75-85 , Nancy : Presses Universitaires de Nancy.

Dickes P., (1996), L'analyse factorielle linéaire et ses deux logiques d'application, *psychologie française*, **41-1**, p.9-22.

Dickes P., Tournois J., Flieller A. & Kop J.L., (1994), *La psychométrie*, Paris : PUF.

Dye N.W. & Very P.S., (1968), Growth changes in factorial structure by age and sex, *Genetic Psychology Monographs*, **31,** p.674-685.

Ekstrom R.B., French J.W. & Harman H.H., (1979), Cognitive factors : Their identification and replication, *Multivariate Behavioral Research Monographs*, **79**(2).

Ericsson K.A. & Charness N., (1994), Expert performance : Its structure and acquisition, *American Psychologist*, **49**, p.725-747.

Fillenbaum S. & Rapoport A., (1971), *Structures in the subjective lexicon*, New York : Academic Press.

Flavell J.H., (1977), *Cognitive development*, Englewood Cliffs, NJ : Prentice-Hall.

French J.W., (1951), The description of aptitude and achievement tests in terms of rotated factors, *Psychometric Monographs*, **5**.

Gagné R.M., (1965), *The conditions of learning*. New York : Holt, Rinehart & Winston.

Gallistel C.R. & Gelman R., (1992), Preverbal and verbal counting and computation, *Cognition*, **44**, p.43-74.

Gardner H., (1993a), *Histoire de la révolution cognitive*, Paris : Payot.

Gardner H., (1993b), *Multiple Intelligence*, New York : Basic Books.

Gardner H., (1997), Developmental Views of Multiple Intelligence, Dans : G.O. Mazier (ed), *Twenty-year commemoration to the life of A.R. Luna (1902-1977)*, p.61-79, New York : Semenenko Foundation.

Gardner H., (1999), *Intelligence Reframed*, New York : Basic Books.

Geary D.C., (1994), *Children's Mathematical Development*, Washington, DC : American Psychological Association.

Geary D.C., Bow-Thomas C.C. & Yao Y., (1992), Counting knowledge and skill in cognitive addition : A comparison of normal and mathematically disabled children, *Journal of Experimental Child Psychology*, **54**, p.372-391.

Geary D.C., Bow-Thomas C.C., Fan L. & Siegler R.S., (1993), Even before formal instruction, Chinese children outperform American children in mental addition, *Cognitive Development*, **8**, p.517-529.

Geary D.C. & Brown S.C., (1991), Cognitive addition : Strategy choice and speed-of-processing differences in gifted, normal and mathematically disabled children, *Developmental Psychology*, **27**, p-398-406.

Geary D.C. & Burlinghampton-Dubree M., (1988), External validation of the strategy choice model for addition. *Journal of Experimental Child Psychology*, **47**, p.175-192.

Geary D.C. & Widaman K.F., (1987), Individual differences in cognitive arithmetic, *Journal of Experimental Psychology : General*, **116**, p.154-171.

Geary D.C. & Widaman K.F., (1992), Numerical cognition : On the convergence of componential and psychometric models, *Intelligence*, **16**, p.47-80.

Gelman R. & Gallistel C.R., (1978), *The child's understanding of number*, Cambridge, MA : Harvard University Press.

Ginsburg H., (1977), *Children's arithmetic : The learning process*, New York : Van Nostrand.

Goldin G.A., (1992), On developing a unified model for the psychology of mathematical learning and problem solving, Dans : W. Geeslin & K. Graham (eds), *Proceedings of the Sixteenth Annual meeting of the International Group for the psychology of Mathematics Education*(Vol 3), p.235-261, Durham : University of New Hampshire.

Guilford J.P., (1956), The Structure of Intellect, *Psychol. Bull.*, **53**, p.267-293.

Guilford J.P., (1967), *The Nature of Human Intelligence*, New York : Mc Graw Hill.

Gustafsson J.-E., (1988), Hierarchical models of individual differences in cognitive abilities, Dans : R.J. Sternberg (ed.), *Advances in the psychology of human intelligence*, vol.4, p.35-71, Hillsdale, NJ : Lawrence Erlbaum Associates.

Gustafsson J.-E. & Balke G., (1993), General and specific abilities as predictors of school achievement, *Multivariate Behavioral Research*, **28**, p.407-434.

Goodman C.H., (1943), A factorial analysis of Thurstone's sixteen Primary Mental Ability tests, *Psychometrika*, **8**, p.141-151.

Greeno J.G., (1980), Psychology of learning, 1960-1980 : One participant's observation, *American Psychologist*, **35**, p.713-728.

Greeno J.G., (1987), Instructional representations based on research about understanding, Dans : A.H. Schoenfeld (ed), *Cognitive science and mathematics education*, p.61-88, Hillsdale, NJ : Lawrence Erlbaum Associates.

Guthrie G.M., (1963), Structure of abilities in a non-Western culture, *Journal of Educational Psychology*, **54**, p.94-103.

Guttman L., (1953), Image theory for the structures of quantitative variates, *Psychometrika*, **18**, p.277-296.

Hadamard J., (1954), *An essay on the psyxhology of invention in the mathematical field*, New York : Dover.

Hammond S., (2000), Introduction to Multivariate Data Analysis, Dans : G.M. Breakwell, S. Hammond & C. Fife-Schaw, *Research methods in psychology*, p.372-396, London : Sage Publications.

Halberstadt J.B. & Niedenthal P.M., (1997), Emotional state and the use of stimulus dimensions in judgement, *Journal of Personality and Social Psychology*, **72**(5), p.1017-1033.

Hashimoto Y., Sawada T., (1979), *Mathematics program in Japan*, Tokyo : National Institute for Educational Research.

Hiebert J. & Carpenter T.P., (1982), Piagetian tasks as readiness measures in mathematics instruction : A critical review, *Educational Studies in Mathematics*, **13**, p.329-345.

Humphreys L.G., Lubinski D. & Yao G., (1993), Utility of predicting group membership : Exemplified by the roles of spatial visualization in becoming an engineer, physical scientist, or artist, *Journal of Applied Psychology*, **78**, p.250-261.

Hunt E., Lunneborg C. & Lewis J., (1975), What does it mean to be high verbal ?, *Cognitive Psychology*, **7**, p.194-227.

Husén T. (ed.), (1967) *International Study of Achievement in Mathematics : A Comparison of Twelve Countries*, Vols **1** et **2**, New York : Wiley.

Ilg F. & Ames L.B., (1962), Development trends in arithmetic, *Journal of Genetic Psychology*, **79**, p.3-29.

Jastak J.F. & Jastak F.R., (1978), *Wide Range Achievement Test (WRAT)*, Wilmington, DE : Jastak Associates.

Jensen A.R., (1980), *Bias in mental testing*, New York : The Free Press.

Jensen A.R., (1990), Speed of information processing in a calculating prodigy, *Intelligence*, **14**, p.259-274.

Jöreskog K.G., (1970a), A general method for analysis of covariance structures, *Biometrika*, **57**, p.239-251.

Jöreskog K.G., (1970b), Estimating and testing of simplex models, *British Journal of Mathematical and Statistical Psychology*, **23**, p.121-145.

Jöreskog K.G., (1973), A general method for estimating a linear structural equation system, Dans A.S. Goldberger & O.D. Duncan (eds), *Structural equation models in the social sciences*, p.85-112, New York : Seminar Press.

Judd C.H., (1928), The fallacy of treating school subjects as "tool subjects", Dans J.R. Clark & W.D. Reeve (eds), *Selected topics in the teaching of mathematics* (3rd Yearbook of the National Council of Teachers of Mathematics, p.1-10, New York : Columbia University, Teachers College.

Kaiser H.F., (1960), The application of electronic computers to factor analysis, *Educational and Psychological Measurement*, **20**, p.141-151.

Kaiser H.F. & Caffrey J., (1965), Alpha factor analysis, *Psychometrika*, **30**, p.1-14.

Keating D.P., List J.A. & Merriman W.E., (1985), Cognitive processing and cognitive ability : A multivariate validity investigation, *Intelligence*, **9**, p.149-170.

Krathwohl D.R., Bloom B.S. & Masia B.B, (1956), *Taxonomy of Educational Objectives. The Classification of Educational Goals : Handbook I : Cognitive Domain*, New York : David McKay Company INC.

Krusetskii V.A., (1976), *The psychology of mathematical abilities in school children*, Chicago : University of Chicago Press.

Kruskal J.B., (1964a), Multidimensional scaling by optimizing goodness of fit to a non-metric hypothesis, *Psychometrika*, **29**, p.1-27.

Kruskal J.B., (1964b), Non metric multidimensional scaling : a numerical method, *Psychometrika*, **29**, p.115-129.

Kruskal J.B. & Wish M., (1978), *Multidimensional scaling*, Beverly Hills, CA : Sage Publications.

Kuhnl J., (1985), Volitional mediators of cognition-behavior consistency : Self-regulatory processes and action versus state orientation, Dans : J. Kuhl & J. Beckman (eds), *Action control : From cognition to behavior*, p. 101-128, Berlin : Springer.

Kyllonen P.C. & Christal R.E., (1989), Cognitive modeling of learning abilities : A status report of LAMP, Dans : R.F. Dillon & J.W. Pellegrino (eds.), *Testing : Theoretical and applied perspectives*, p.146-173, New York : Praeger.

Lampert M., (1986), Knowing, doing, and teaching multiplication,*Cognition and instruction*, **3**, p.305-342.

LeDoux J.E., (1996), *The emotional brain : The mysterious underpinnings of emotional life*, New York : Simon and Schuster, Inc.

Lewis A.B. & Mayer R.E., (1987), Student's miscomprehension of relational statements in arithmetic word problems, *Journal of Educational Psychology*, **79**, p.363-371.

Lindvall C.M & Bolvin J.O., (1967), Programmed instruction in the schools : An application of programming principles in individually Prescribed Instruction. Dans : P.C. Lange (ed), *Programmed instruction : Sixty-sixth yearbook of the National Society for the Study of Education*, p.217-254, Chicago : University of Chicago Press.

Lingoes J.C., (1971), Some boundary conditions for a monotone analysis of symmetric matrices, *Psychometrika*, **36**, p.195-203.

Lingoes J.C., (1972), A general survey of the Guttman-Lingoes nonmetric program series, Dans : R.N. Shepard, A. Romney & S. Nerlove (eds), *Multidimensional scaling : Theory and applications in the behavioral sciences*, Vol 1 : *Theory*, New York : Seminar Press.

Little T.D. & Widaman K.F., (1994), A production task evaluation of individual differences in mental addition skill development : Internal and external validation of chronometric models, *Journal of Experimental Child Psychology*.

Lubinski D. & Humphreys L.G., (1990), A broadly based analysis of mathematical giftedness, *Intelligence*, **14**, p.327-355.

Maher C.A., Davis R.B. & Alston A., (1991), Implementing a "thinking curriculum" in mathematics, *Journal of Mathematical Behaviour*, **10**, p.219-224.

Mandler G. & Shebo B.J., (1982), Subitizing : An analysis of its component processes, *Journal of Experimental Psychology : General*, **111**, p.1-22.

Marshalek B., Lohman D.F. & Snow R.E., (1983), The complexity continuum in the radex and hierarchical models of intelligence, *Intelligence*, **7**, p.107-127.

Mayer R.E., (1981), Frequency norms and structural anlysis of algebra story problems into families, caegories and templates, *Instructional Science*, **10**, p.135-175.

McLeod D.B., (1990), Information-processing theories and mathematics learning : The role of affect, *International Journal of Educational Research*, **14**, p.13-29.

Meyers C.E. & Dingman H.F., (1960), The structure of abilities at the preschool ages : Hypothesized domains, *Psychological Bulletin*, **57**, p.514-532.

Murray J.E., (1949), An analysis of geometric ability, *Journal of Educational Psychology*, **40**, p.118-124.

Nesher P. & Kilpatrick J. (eds.), (1990), *Mathematics and cognition : A research synthesis by the International Group for the Psychology of Mathematics Education*, Cambridge, England : Cambridge University Press.

Newell A. & Simon H.A., (1972), *Human problem solving*, Englewood cliffs, NJ : Prentice Hall.

Newman D., Griffin P. & Cole M., (1989), *The construction zone : Working for cognitive change in school*, Cambridge : Cambridge University Press.

Osborne R.T. & Lindsey J.M., (1967), A longitudinal investigation of change in the factorial composition of intelligence with age in young school children, *Journal of Genetic Psychology*, **110**, p.49-58.

Papert S., (1980), *Children, Computers and Powerful Ideas*, New York : Basic Books.

Pawlik K., (1966), Concepts in human cognition and aptitudes, Dans : R.B. Catell (ed), *Handbook of multivariate experimental psychology*, **76**, p.678-689.

Pellerey M., (1974), Obiettivi dell'insegnamento della matematica nella scuola secondaria superiore. *Oreintamenti Pedagogici*, **29**, p.218-38.

Pellerey M., (1991), Mathematics Instruction, Dans : A. Lewy (ed), *The International Enceclopedia of Curriculum*, p.870-880, Oxford : Pergamon Press.

Perkins D.N., Jay E. & Tishman S., (1993), Beyond abilities : A dispositional theory of thinking, *Merrill Palmer Quaterly*, **39**, p.1-21.

Petitto A.L., (1990), Development of numberline and measurement concepts, *Cognition and Instruction*, **7**, p.55-78.

Piaget J., (1947), *La Psychologie de l'intelligence*, Paris : A. Colin.

Piaget J., (1952), *The child's conception of number*, New York : Humanities Press.

Piaget J. & Inhelder B., (1941), *Le développement des quantités chez l'enfant*, Neuchâtel et Paris : Delachaux et Niestlé.

Piaget J. & Inhelder B., (1948), *La représentation de l'espace chez l'enfant*, Paris : PUF.

Piaget J. & Inhelder B., (1956), *The child's conception of space*, London : Routledge and Kegan Paul.

Piaget J. & Inhelder B., (1959), *La genèse des structures logiques élémentaires*, Neuchâtel et Paris : Delachaux et Niestlé.

Piaget J., Inhelder B. & Szeminska A. (1960), *The child's conception of geometry*. New York : Basic Books.

Piaget J. & Szeminska A., (1941), *La genèse du nombre chez l'enfant*, Neuchâtel et Paris : Delachaux et Niestlé.

Pichot P., (1965), *Les tests mentaux*, Paris : PUF.

Polya G., (1945), *How to solve it*, Princeton, NJ : Princeton University Press.

Polya G., (1954a), *Induction and analogy in mathematics*, Princeton, NJ : Princeton University Press.

Polya G., (1954b), *Patterns of plausible inference*, Princeton, NJ : Princeton University Press.

Polya G., (1962), *Mathematical discovery* (Vol.1), New York : Wiley.

Polya G., (1965), *Mathematical discovery* (Vol.2), New York : Wiley.

Rabinowitz M. (ed), (1988), Computer simulations as research tools, *International Journal of Educational Research*, **12**, p.1-102.

Reboul Olivier, (1995), *Qu'est-ce qu'apprendre ?*, Paris : PUF.

Resnick L.B., (1976), Task analysis in instructional design : some cases from mathematics. Dans : D. Klahr (ed), *Cognition and instruction*, p.51-80, Hillsdale, NJ : Erlbaum.

Resnick L.B., (1987), *Education and learning to think*, Washington, DC : National Academy Press.

Reuchlin M., (1964), *Méthodes d'analyse factorielle à l'usage des psychologues*, Paris : PUF.

Reuchlin M., (1992), *Introduction à la recherche en psychologie*, Paris : Nathan.

Richard J.F., (1990), *Les activités mentales : comprendre, raisonner, trouver des solutions*, Paris : Armand Colin.

Riley M.S., Greeno J.G. & Heller J.I., (1983), Development of children's problem-solving ability in arithmetic, Dans : J.P. Ginsburg (ed), *The development of mathematical thinking*, p.153-196, New York : Academic Press.

Romberg T.A. & Carpenter T.P, (1986), Dans : M.C. Wittrock (ed), *Handbook of Research on Teaching*, New York : Macmillan.

Romberg T., Harvey J., Moser J. & Montgomery M., (1976), *Developing mathematical processes*, Chicago : Rand McNally.

Roskam E.E., (1981), Data theory and scaling : a methodological essay on the role and use of data theory and scaling in psychology, Dans : I. Borg (ed), *Multidimensional data representation : When and Why*, p.193-229, Ann Arbor, Michigan : Mathesis Press.

Schaie K.W., (1983), The Seattle longitudinal study : A 21-year exploration of psychometric intelligence in adulthood, Dans : K.W. Schaie (ed), *Longitudinal studies of adult psychological development*, p.64-135, New York : Guilford Press.

Schoenfeld A.H., (1988), When good teaching leads to bad results : The disasters of "well-taught" mathematics courses, *educational Psychologist*, **23**, p.145-166.

Schiffman S.A., Reynolds M.L. & Young F.W., (1981), *Multidimensional Scaling : Theory, Methods and Applications*, New York : Academic Press.

Schiltz J., (1999), *Application d'un modèle stochastique à la construction d'un test adaptatif informatisé en mathématiques*, Rapport établi pour le Ministère de l'Education Nationale et de la Formation Professionnelle, Luxembourg.

Shepard R.N., (1962), The analysis of proximities : Multidimensional scaling with an unknown distance function, part I & II, *Psychometrika*, **27**, p.125-140, p.219-246.

Shepard R.N., (1974), Representation of structure in similarity data : Problems and prospects, *Psychometrika*, **39**, p.373-421.

Siegler R.S., (1983), Information processing approaches to development. Dans : P. Mussen & W. Kassen (ed), *Manual of child psychology*, p.129-211, New York : Wiley.

Siegler R.S., (1988), Individual differences in strategy choices : Good students, not so good students and perfectionists. *Child Development*, **59**, p.833-851.

Siegler R.S., (1993), Adaptive and nonadaptive characteristics of low-income children's mathematical strategy use, Dans : L.A. Penner, G.M. Batsche, H.M. Knoff & D.L. Nelson (ed), *The challenge in mathematics and science education : Psychology's response*, p.341-366, Washington, DC : American Psychological Association.

Silver E.A., (1994), On mathematical problem posing, *For the Learning of Mathematics*, **14**(1), p.19-28.

Spearman C., (1927), *The abilities of man*. London : Macmillan.

Spence I. & Domoney D.W., (1974), Single subject incomplete designs for nonmetric multidimensional scaling, *Psychometrika*, **39**, p.469-490.

Stanic G.M & Kilpatrick J., (1980), Historical perspectives on problem solving in the mathematics curriculum, Dans : R. Charles & E.A. Silver (eds), *The teaching and assessing of mathematical problem solving*, Hillsdale, NJ : Lawrence Erlbaum Associates.

Stanley J.C., (1974), Intellectual precocity, Dans : J.C. Stanley, D.P. Keating & L.H. Fox (eds.), *Mathematical talent : Discovery, description and development*, p.1-22, Baltimore : Johns Hopkins University Press.

Steffe L.P., (1991), The constructivist teaching experiment : Illustrations and implications, Dans E. von Glasersfeld (ed), *Radical constructivism in mathematics education*, p.177-194, Dordrecht : Kluwer.

Steffe L.P., von Glasersfeld E., Richards J. & Cobb P., (1983), *Children's counting types : Philosophy, theory and application*. New York : Praeger.

Sternberg R.J., (1977), *Intelligence, information processing, and analogical reasoning : The compential analysis of human abilities*, Hillsdale, NJ : Erlbaum.

Sternberg R.J. & Gardner M.K., (1983), Unities in inductive reasoning, *journal of Experimental Psychology : General*, **112**, p.80-116.

Sternberg R.J., (1994), Reforming school reform : Comments on Multiple Intelligences ; The Theory in Practice, *Teachers College Record*, **94**(4).

Takane Y., Young F.W. & De Leeuw J., (1977), Nonmetric individual differences multidimensional scaling : An alternating least squares method with optimal scaling features , *Psychometrika*, **42**, p.7-67.

Thurstone L.L., (1938), Primary mental abilities, *Psychometric Monographs* (No. **1**).

Thurstone L.L., (1947), *Multiple factor analysis*, Chicago : University of Chicago Press.

Thurstone L.L. & Thurstone T.G., (1941), Factorial studies of Intelligence, *Psychometric Monographs*, **2**.

Tournois J., (1989), Multi-interprétation en analyse par échelonnement multidimensionnel (EMD), *Bulletin de Psychologie*, **388**, p.200-209.

Tournois J., Dickes P., (1993), *Pratique de l'échelonnement multidimensionnel*, Bruxelles : De Boeck Université.

Turner J.R. & Thayer J.F., (2001), *Introduction to Analysis of Variance*, London : Sage Publications.

Tymoczko T., (1986), Interlude, Dans : T. Tymoczko (ed), *New directions in the philosophy of mathematics*, p.95-98, Boston : Birkhauser.

Undheim J.O. & Gustafsson J.-E., (1987), The hierarchical organization of cognitive abilities : Restoring general intelligence through the use of linear structural relations (LISREL), *Multivariate Behavioral Research*, **22**, p.149-171.

Vandenberg S.G., (1959), The primary mental abilities of Chinese students : a comparative study of the ability of a factor structure, *Annals of the New York Academy of Sciences*, **79**, p.257-304.

Van Oers B., (1990), The development of mathematical thinking in school : A comparison of the action-psychological and information-processing approaches, *International Journal of Educational Research*, **14**, p.51-66.

Vergnaud G., (1990), Epistemology and psychology of mathematics education, Dans : P. Nesher & J. Kilpatrick (eds), *Mathematics and cognition: A research synthesis by the International Group for the Psychology of Mathematics Education*, p.14-30, Cambridge : Cambridge University Press.

Vernon P.E., (1961), *The structure of human abilities (2nd ed.)*, London : Methuen.

Very P.S., (1967), differential factor structures in mathematical ability, *Genetic Psychology Monographs*, **75**, p.169-207.

Walkey F.H., (1983), Simple versus complex factor analysis of reponses to multiple scale questiooaires, *Multivariate Behavioral Research*, **18**, p.401-421.

Wechsler D., (1967), *Manual for the Wechsler Intelligence Scale for Children*, New York : Psychological Corporation.

Werdelin I., (1961), *Geometrical ability and the space factor in boys and girls*, Lund, Suède : Gleerups.

Wertheimer M., (1945), *Productive thinking*, London : Tavistock.

Wertsch J.V., (1985), *Vigotsky and the social foundation of mind*, Cambridge, MA : MIT Press.

Widaman K.F. & Little T.D., (1992), The development of skill in mental arithmetic : an individual differences perspective, Dans : J.I.D Campbell (ed), *The nature and origins of mathematical skills*, p.189-253, Amsterdam : North-Holland.

Widaman K.F., Little T.D., Geary D.C. & Cormier P., (1992), Individual differences in the development of skill in mental addition : Internal and external validation of chronometric models, *Learning and Individual Differences*, **4**, p.167-213.

Wilson J.W., (1971), Secondary school mathematics, Dans : B.S. Bloom, J.T. Hastings & G.F. Madaus (eds), *Handbook on Formative and Summative Evaluation of Student Learning*, New-York : McGraw-Hill.

Winch C., (1990), *Language, ability, and educational achievement*, New York : Routledge, Chapman & Hall.

Wittrock M.C., (1974), A generative model of mathematics learning, *Journal for Research in Mathematics Education*, **5**, p.181-196.

Wynn K., (1992), Addition and substraction by human infants, *Nature*, **358**, p.749-750.