

## **L'évaluation de l'intelligence : état actuel et tentatives de renouvellement**

Jacques Lautrey<sup>1</sup>

Les premiers tests construits pour évaluer l'intelligence auront bientôt cent ans : la première échelle de développement intellectuel (Binet et Simon, 1905) et la première analyse factorielle (Spearman, 1904) ont en effet été publiées au début du siècle dernier. Les tests d'intelligence les plus utilisés actuellement sont des versions révisées, améliorées, de ces premiers tests et reposent donc sur une conception de l'intelligence qui aura bientôt un siècle. Ce point de vue paraîtra peut-être excessif. On peut lui opposer que la distinction entre QI verbal et QI de performance, introduite par Wechsler à la fin des années trente, à l'occasion de la construction du Wechsler-Bellevue, et l'approche multifactorielle mise en œuvre par Thurstone à la même époque, peuvent être considérées comme des évolutions importantes en direction d'une conception multidimensionnelle de l'intelligence. Même si l'on admet cette objection, les tests les plus utilisés actuellement reposent sur une conception de l'intelligence qui a plus d'un demi-siècle.

Or depuis un demi-siècle - et a fortiori depuis un siècle - les idées sur l'intelligence ont beaucoup évolué. Pour ne citer que les deux contributions les plus marquantes, la théorie de Piaget et la théorie du traitement de l'information ont fortement renouvelé les idées sur l'intelligence et son développement. Le phénomène curieux, sur lequel je voudrais m'arrêter aujourd'hui, est que ce renouvellement des théories sur l'intelligence ne se soit pas accompagné d'un renouvellement correspondant des instruments d'évaluation de l'intelligence. Pourquoi les évolutions théoriques n'ont-elles pas engendré de tests d'intelligence capables de supplanter les tests psychométriques classiques, c'est à dire les tests inspirés des idées de Binet et de Wechsler, ou de Spearman et de Thurstone. Il y a là un paradoxe qui servira de fil directeur à mon propos.

Je ferai d'abord un rapide survol de l'état actuel de l'approche psychométrique classique, puis des tentatives de renouvellement, avant de revenir dans la conclusion à ce paradoxe et à son explication.

---

<sup>1</sup> La boratoire Cognition et Développement, UMR CNRS 8605, Université René Descartes – Paris V, institut de Psychologie, 71 Avenue Edouard Vaillant, 92 774 – Boulogne-Billancourt Cedex. E-mail : lautrey@psycho.univ-paris5.fr

## 1. Etat actuel de l'approche psychométrique classique de l'intelligence

Sous cette rubrique, c'est à l'état actuel des tests factoriels et des échelles de QI que nous nous intéresserons. Dans la suite du propos, les expressions " approche classique " ou " tests classiques " référeront toujours à ces deux types de tests. Nous commencerons par les tests factoriels. Si ceux-ci ont peu changé, les idées sur la structure factorielle de l'intelligence ont connu quelques évolutions, d'ailleurs liées à l'évolution des méthodes d'analyse factorielle. Du côté des échelles d'intelligence, auxquelles nous viendrons ensuite, il y a eu beaucoup de révisions des outils classiques, mais peu de réelles nouveautés.

### 1.1. La structure factorielle de l'intelligence.

Où en est-on du débat sur le caractère général ou multiple de l'intelligence ? Comme on sait, la controverse qui a fait rage un moment entre les partisans du facteur général mis en évidence par Spearman et les partisans des facteurs multiples mis en évidence par Thurstone, s'est résolue par l'intégration de ces deux conceptions dans un modèle hiérarchique de la structure factorielle de l'intelligence.

#### *1.1.1. Les modèles hiérarchiques*

Les travaux de Burt et Vernon (cf. Vernon, 1961) et ceux de Cattell et Horn (cf. Cattell, 1971), entre autres, ont contribué à montrer que les parts de variance expliquées par ces différents facteurs pouvaient être fractionnées en strates hiérarchiquement ordonnées. Burt et Vernon procédaient de haut en bas : après avoir extrait la variance du facteur général de sa batterie de tests, Vernon montrait que la variance restante se partageait d'abord entre deux grands facteurs de groupe, un facteur qu'il appelait Verbal-Education et un facteur qu'il appelait Kinesthésique-Moteur. Une fois retirée la variance de ces deux grands facteurs de groupe, la variance restante se partageait entre des facteurs plus spécifiques correspondant aux facteurs primaires de Thurstone. Cattell et Horn procédaient de haut en bas : après avoir extrait la variance expliquée par les facteurs multiples correspondant aux facteurs primaires de Thurstone, ils faisaient une analyse factorielle de second ordre visant à extraire la variance commune aux facteurs primaires (et responsable de leurs intercorrélations). Ils trouvaient dans une seconde strate plusieurs facteurs généraux, notamment un facteur général d'intelligence fluide, un facteur général d'intelligence cristallisée et un facteur général d'intelligence visuo-spatiale, puis plus tard d'autres encore. Mais la structure à laquelle ils aboutissaient ne comportait pas, comme chez Burt et Vernon, de facteur général unique coiffant l'ensemble de la pyramide. Il y avait donc accord sur le caractère hiérarchique de la

structure factorielle de l'intelligence. Il y avait aussi convergence sur les contenus de certains facteurs. En examinant les contenus des facteurs que comportaient ces deux modèles, on pouvait en effet voir que la première strate comportait à peu près les mêmes facteurs primaires (verbal, fluidité verbale, numérique, etc.), et que dans la deuxième strate, les contenus de certains grands facteurs de groupe étaient assez proches : le facteur d'intelligence cristallisée de Cattell et Horn saturait à peu près les mêmes facteurs primaires que le facteur Verbal-Education de Burt et Vernon. De même, le facteur Gv (intelligence visuo-spatiale) des premiers saturait à peu près les mêmes facteurs primaires que le facteur Kinesthésique-Moteur des seconds.

Il y avait pourtant, entre ces deux modèles, des contradictions qui sont restées longtemps non résolues. La structure factorielle de Burt-Vernon comportait trois strates alors que celle de Cattell et Horn n'en comportait que deux (la strate manquante étant ici la plus élevée, celle du facteur général). En revanche, la seconde strate du modèle de Burt et Vernon ne comportait pas d'équivalent du facteur d'intelligence fluide trouvé par Cattell et Horn.

### *1.1.2. Vers l'intégration des différents modèles factoriels*

Cette contradiction a été levée par les travaux de Gustaffson (1984) qui, en utilisant les possibilités d'analyse factorielle confirmatoire offertes par le logiciel LISREL, a montré que le modèle hiérarchique s'ajustant le mieux à ses données était un modèle à trois strates, dans lequel le facteur général qui figure dans la troisième strate explique la même part de variance que le facteur d'intelligence fluide figurant dans la seconde. En d'autres termes, le facteur d'intelligence fluide (Gf) de Cattell et Horn explique la même part de variance que le facteur d'intelligence général (G) de Burt et Vernon. La contradiction entre les deux modèles s'explique alors par la méthode d'analyse factorielle utilisée : en faisant leur analyse hiérarchique du haut vers le bas, Burt et Vernon isolaient d'abord le facteur général et, de ce fait, ne trouvaient pas le facteur Gf dans la seconde strate ; en faisant leur analyse hiérarchique en allant du bas vers le haut, Cattell et Horn obtenaient le facteur Gf dès la seconde strate et, de ce fait, ne trouvaient pas de variance propre à un facteur G situé dans une troisième strate.

Le modèle hiérarchique à trois strates identifié par Gustaffson dans l'analyse factorielle confirmatoire de ses propres données a été corroboré un peu plus tard par la méta-analyse réalisée par Carroll (1993). Cet auteur a réanalysé toutes les analyses factorielles de tests d'intelligence dont il a pu se procurer les données (460 études) et montré que toutes les structures factorielles trouvées sont compatibles avec un modèle hiérarchique à trois

strates dans lequel on trouve, dans la première strate, une bonne trentaine de facteurs primaires équivalents à ceux de Thurstone, dans une seconde strate, six grands facteurs de second ordre, les plus fréquemment représentés étant les facteurs d'intelligence fluide (Gf), d'intelligence cristallisée (Gc), de mémoire et apprentissage (Gm), de visualisation (Gv). Le facteur Gf sature surtout les facteurs primaires d'induction, de déduction, qui eux-mêmes saturent les tâches de raisonnement n'exigeant pas de connaissances particulières, comme les matrices progressives de Raven, les séries numériques, les tests piagétiens, etc.. Le facteur Gc sature principalement les différents facteurs primaires du domaine verbal, comme le vocabulaire ou la compréhension de lecture. Le facteur Gm sature les facteurs primaires relatifs à l'empan mnémonique, la mémoire associative, etc. Le facteur Gv sature les différents facteurs primaires relatifs à la visualisation, aux relations spatiales spatiales, à la vitesse de cloture, etc. Les autres facteurs de second ordre sont moins connus : Représentation auditive (Gu), Récupération en mémoire à long terme (Gr), Rapidité cognitive (Gs) et vitesse de traitement (Gt). Les facteurs de cette seconde strate ont des intercorrélations qui donnent lieu à un unique facteur général, de troisième ordre, isolant la variance qui leur est commune. Dans le modèle hiérarchique de Carroll comme dans celui de Gustaffson, c'est bien le facteur d'intelligence fluide qui partage le plus de variance avec le facteur général (les facteurs de la seconde strate ont été présentés ci-dessus dans l'ordre décroissant de leur saturation en facteur général). Il n'est pas possible ici d'entrer davantage dans le détail de la structure factorielle à trois strates proposée par Carroll, mais une adaptation en langue française de ce modèle pourra être trouvée ailleurs (Huteau et Lautrey, 1999).

Le modèle hiérarchique à trois strates de Carroll est celui qui réunit actuellement le plus large consensus sur la structure factorielle de l'intelligence. Il intègre à la fois le facteur général de Spearman et la structure multifactorielle de Thurstone et il concilie les modèles hiérarchiques de Burt et Vernon d'une part et de Cattell et Horn d'autre part. Il distingue des formes d'intelligence variées, qui correspondent aux différents facteurs de second ordre jusqu'ici identifiés. La présence de ces multiples facteurs traduit une relative indépendance entre les différentes formes d'intelligence distinguées et rend compte du fait qu'il est possible d'être performant dans l'une d'entre elles sans l'être nécessairement dans les autres. Il existe néanmoins une tendance statistique à ce que les sujets les plus performants dans l'une de ces formes d'intelligence tendent à l'être aussi dans les autres. C'est ce dont rend compte le facteur général. Quel est toutefois le degré de généralité de ce facteur général ? C'est une question sur laquelle nous reviendrons dans la partie consacrée aux tentatives d'élargissement du concept d'intelligence.

## 1.2. Les échelles d'intelligence

Conçues à l'origine pour donner une évaluation globale, quantifiée par un quotient intellectuel, les échelles d'intelligence ont évolué vers des évaluations plus analytiques du développement. Wechsler a introduit dans ses échelles la distinction entre QI verbal et QI performance, ainsi que la possibilité de comparer les scores des différentes sous-échelles en les standardisant. Depuis ces modifications, les échelles d'intelligence n'ont pas connu d'évolutions importantes mais seulement des révisions. Celles-ci consistent à adapter leur contenu et leur présentation aux évolutions culturelles des sociétés dans lesquelles elles sont utilisées. Ces révisions périodiques sont aussi nécessaires pour tenir compte de l'augmentation régulière du QI avec les générations. Dans les dernières décennies, cette augmentation est de l'ordre de 3 points tous les dix ans. Ce phénomène donne lieu à des interprétations contradictoires (cf. Flynn, 1987 ; Flieller, ce volume), mais quelle qu'en soit l'interprétation, il contraint à réétalonner les tests d'intelligence assez souvent.

Seules les échelles d'intelligence qui sont régulièrement révisées ont survécu. C'est le cas des échelles de Wechsler. A titre d'exemple, l'échelle pour enfants, la WISC, a été créée par Wechsler en 1949, révisée en 74 (WISC-R) et de nouveau révisée en 91 (WISCIII). A chaque révision de l'échelle américaine, une nouvelle adaptation française en a été faite (l'adaptation française de la WISC III a été publiée en 1996). L'adaptation américaine du Binet-Simon, le Stanford-Binet a aussi été régulièrement révisée et est aussi très utilisée aux Etats-Unis. Le Binet-Simon n'est par contre plus utilisé en France faute d'avoir été révisé à temps. La seule révision qui en ait été faite est celle qui a été dirigée par Zazzo en 1966 et a abouti à la Nouvelle Echelle Métrique de l'Intelligence (NEMI). La NEMI n'a toutefois pas été révisée depuis et, pour les raisons indiquées plus haut, une échelle qui n'a pas été révisée depuis 35 ans devient obsolète.

Par rapport à ces ancêtres que sont les échelles de Binet et de Wechsler, la seule nouvelle échelle d'intelligence qui soit parvenue à s'imposer est la K-ABC (Kaufman Assessment Battery for Children), une batterie de tests qui évaluent le développement de l'intelligence dans la période d'âge de 2 à 12 ans. Elle a été publiée en 1983 aux USA par A. et N. Kaufman et adaptée en France en 93 Kaufman & Kaufman, 1993). Ses auteurs ont voulu renouveler l'évaluation de l'intelligence en la faisant porter sur deux grands types de traitement de l'information, le traitement séquentiel et le traitement simultané. Cette distinction s'inspire de celle que faisait Luria entre un traitement "successif" dont le siège serait dans les aires fronto-temporales et un traitement simultané situé dans les aires pariéto-occipitales. La

K-ABC comporte donc une sous-échelle évaluant le développement du traitement séquentiel et une sous-échelle évaluant le développement du traitement simultané. Chacune de ces deux sous-échelles donne lieu à un QI et il est aussi possible, en les réunissant, de calculer un QI global, comme dans la WISC. Par ailleurs, une troisième sous-échelle évalue les connaissances (vocabulaire, information, arithmétique, compréhension de lecture). Par rapport à la WISC, la K-ABC présente donc un certain nombre de spécificités : elle évalue l'efficacité de deux grands types de processus (séquentiels et simultanés) au lieu d'évaluer les performances dans deux grands domaines distingués par les contenus symboliques sur lesquels portent les traitements (verbal et non verbal) ; elle fait une distinction plus claire entre les processus et les connaissances en introduisant une troisième échelle spécifique aux connaissances ; enfin, la passation se fait en temps libre et n'introduit donc pas dans l'évaluation, l'effet de la vitesse d'exécution.

Compte tenu de ses propriétés spécifiques, la K-ABC constitue-t-elle un renouvellement dans l'évaluation de l'intelligence ? Mesure-t-elle autre chose que les tests classiques ? Cette question sera abordée un peu plus loin, après avoir examiné les tentatives de renouvellement et notamment la réinterprétation des tests classiques dans le cadre de la psychologie cognitive.

## **2. Les tentatives de renouvellement**

Les évolutions théoriques dans la conception de l'intelligence ont suscité des tentatives de renouvellement des tests classiques. C'est ainsi que la théorie de Piaget a inspiré la construction de tests opératoires. C'est ainsi également que le développement de la psychologie du traitement de l'information a inspiré des renouvellements méthodologiques visant à identifier les processus cognitifs responsables des différences de performance dans les tests d'intelligence. Enfin, on a reproché aux tests classiques de n'évaluer qu'une des différentes formes d'intelligence et quelques tentatives ont eu lieu pour élargir cette évaluation à d'autres formes. Ces différentes tentatives seront maintenant rapidement passées en revue.

### 2.1. Les approches piagétienne et neo-piagétienne

#### *2.1.1. Les tests piagétiens*

La théorie de Piaget a profondément renouvelé les idées sur le développement de l'intelligence. Piaget et ses collaborateurs ont mis au point, pour éprouver cette théorie, une quantité de situations expérimentales qui étaient autant de situations d'évaluation du développement de l'intelligence. Plusieurs chercheurs ont eu l'idée d'adapter ces situations,

de les standardiser, d'améliorer leurs qualités psychométriques, de telle sorte qu'elles deviennent des échelles d'évaluation du stade de développement de la pensée logique de l'enfant ou de l'adolescent. On sait qu'en France, ces travaux ont été entrepris par Longeot (1969, 1974), avec l'objectif initial de vérifier une hypothèse formulée par Reuchlin (1964) sur les relations entre la notion piagétienne de stade et la notion psychométrique de facteur général. C'est ainsi que Longeot a mis au point une échelle de développement de la pensée logique (EPL) à passation individuelle et des tests à passation collective (TOF) destinés à évaluer le passage au stade des opérations formelles. Une autre échelle collective du développement logique (ECDL), elle aussi centrée sur le passage des opérations concrètes aux opérations formelles, a été mise au point par J. Hornemann. Une batterie d'épreuves centrée sur la période des opérations concrètes, plus spécialement tournée vers la construction du nombre (UDN) a été élaborée par Meljac et Lemmel (1999). Des tests piagétiens ont aussi été construits à l'étranger, notamment en Angleterre et aux Etats-Unis.

Ces nouveaux tests ont un moment paru pouvoir supplanter les échelles de développement classiques pour plusieurs raisons. Tout d'abord parce que les tests piagétiens avaient un fondement théorique. Les conduites observées pouvaient, de ce fait, être interprétées dans le cadre de la théorie qui avait guidé la construction du test. Ensuite, parce que dans ces tests, le sujet était situé par rapport à un critère plutôt que par son rang dans la population de référence : il avait ou n'avait pas encore atteint tel ou tel stade de développement. Contrairement au QI, cette forme de caractérisation est appelée à changer au cours du développement et comporte donc moins de risques de stigmatisation. Enfin, cette forme de diagnostic était supposée renseigner sur le type de raisonnement dont un enfant est capable, à tel moment précis du développement, ce qui devrait faciliter l'articulation du traitement pédagogique avec le traitement psychologique.

La suite des recherches sur la théorie de Piaget et l'expérience acquise avec ces tests ont conduit à relativiser la distinction avec les tests classiques. Des décalages beaucoup plus importants que la théorie ne le laissait prévoir sont apparus entre des items faisant en principe appel à une même structure opératoire. Un même sujet pouvait être caractérisé par des stades de développement différents selon le test opératoire considéré et, souvent, selon les items considérés dans un même test. Cette difficulté inattendue a souvent conduit à revenir au score total obtenu à l'épreuve pour déterminer le stade de développement atteint : si le sujet obtient entre  $n_1$  et  $n_2$  points à l'épreuve, il est au stade concret ; s'il obtient entre  $n_2$  et  $n_3$  points, il est au stade préformel, etc. Cette façon de procéder revient à prendre en compte la performance globale et estompe la différence entre les tests piagétiens et les tests classiques.

Dans le manuel de l'EPL de Longeot, une feuille de cotation proposée par Gibello, a été introduite par la suite pour comptabiliser les décalages intraindividuels de stade de développement entre sous-tests. Celle-ci permet une description plus analytique de l'allure du développement, mais laisse entier le problème de l'interprétation des hétérochronies observées. Celles-ci ne peuvent évidemment recevoir d'interprétation dans le cadre de la théorie de Piaget. Quant à l'interprétation proposée par Gibello (1984), elle est d'inspiration déféctologique : les hétérochronies de développement sont considérées comme les symptômes de divers syndromes pathologiques. Mais s'il est vrai que certaines pathologies peuvent provoquer une hétérochronie du développement cognitif, ce serait une erreur d'en conclure qu'une hétérochronie du développement est, pour autant, le signe d'une pathologie. D'autres travaux ont montré que les hétérochronies étaient la règle dans le développement cognitif normal (cf. Lautrey, 1980 ; Lautrey, de Ribaupierre et Rieben, 1986). Cette constatation ruine un des espoirs qu'avaient suscité les tests piagétiens, l'espoir de pouvoir diagnostiquer le stade – au singulier - de développement cognitif d'un enfant, indépendamment du domaine considéré, et d'en inférer le type de traitement pédagogique – au singulier – adapté à son mode de raisonnement.

Les analyses factorielles de batteries d'épreuves piagésiennes, ont par ailleurs montré une assez grande proximité entre la structure factorielle de ces épreuves et celle des tests factoriels classiques. Ces analyses ont en général mis en évidence deux grands facteurs obliques, l'un saturant les épreuves du domaine infralogique, l'autre saturant les épreuves logico-mathématiques. Par leur contenu, ces deux facteurs paraissent pouvoir être assimilés à deux des facteurs trouvés avec les batteries de tests factoriels, respectivement le facteur de visualisation spatiale et le facteur d'intelligence fluide. Ces deux facteurs étant corrélés entre eux, il est en général aussi possible d'extraire un facteur général de réussite aux tests piagésiens. Dans les études où les sujets ont passé à la fois des tests piagésiens et des tests factoriels classiques ou des échelles d'intelligence, le score en facteur général (ou le score total) dans les épreuves piagésiennes s'avère fortement corrélé au score en facteur général (ou au score total) ou au QI dans les tests d'intelligence classiques (Lautrey, à paraître).

Des différentes observations résumées plus haut, on peut conclure que les tests piagésiens évaluent à peu près la même chose que les tests d'intelligence classiques. Ils couvrent néanmoins un spectre moins large car ils n'évaluent pas l'intelligence cristallisée. Ils ont aussi probablement une moins bonne fidélité (la fidélité test-retest de l'EPL n'a à notre connaissance jamais été évaluée) du simple fait qu'ils comportent en général moins d'items que les tests classiques. En contrepartie, leur principal apport est de permettre d'évaluer



beaucoup plus finement que ne le font les tests classiques le développement de la pensée logique.

Néanmoins, force est bien de constater que ni en France ni dans aucun des autres pays où ils sont à la disposition des psychologues praticiens, les tests inspirés de la théorie de Piaget n'ont supplanté les tests classiques. L'utilisation des tests piagétiens reste dans tous les pays assez marginale. Dans la mesure où il s'en vend peu, les éditeurs hésitent à s'engager dans les frais que suppose leur révision, et si ces tests ne sont pas révisés, leur chance de concurrencer sérieusement les tests classiques – plus utilisés et donc plus souvent révisés – diminue d'autant.

### *2.1.2. Les épreuves neo-piagésiennes*

Les difficultés rencontrées avec la théorie de Piaget ont donné naissance à un courant de recherche qui, tout en conservant la perspective constructiviste de Piaget et la notion de stade de développement, en a abandonné les notions d'équilibration et de structure opératoire. Dans l'approche neo-piagésienne, les stades sont expliqués par l'évolution développementale de la capacité de traitement. Il existe des variations entre les auteurs quant à la nature de cette capacité de traitement : ressource attentionnelle ou puissance mentale pour les uns (par ex. Pascual-Leone, 1987), empan de la mémoire de travail pour les autres (par ex. Case, 1985), mais le point commun de toutes les théories neo-piagésiennes est d'admettre le caractère limité de la capacité de traitement et l'évolution de cette limite avec l'âge, sous l'effet de la maturation et de l'exercice. Une synthèse récente des théories neo-piagésiennes pourra être trouvée ailleurs (de Ribaupierre, 1997).

Diverses épreuves ont été mises au point dans le cadre de ce courant de recherche pour évaluer la capacité de traitement. A titre d'exemple, on peut citer deux épreuves mises au point par Case (1985) pour évaluer l'empan de la mémoire de travail dans le stade qu'il appelle " dimensionnel " et qui correspond en gros à la période opératoire chez Piaget. L'une, qu'il a appelée " Mr Cucumber " et que de Ribaupierre et Bayeux (1994) ont rebaptisée " Mr Cacahuète ", consiste à présenter à l'enfant une image dans laquelle des gommettes sont placées à différents endroits sur le corps d'une sorte de clown (dont la forme du corps ressemble, selon les goûts, à un concombre ou à une cacahuète). Après avoir laissé l'enfant observer cette figure pendant autant de secondes qu'elle comporte de gommettes, l'expérimentateur la masque et lui présente une figure vierge sur laquelle il doit, avec les gommettes qu'il a à sa disposition, reproduire la configuration observée. Un item est considéré comme réussi lorsque l'enfant a placé le nombre exact de gommettes dans leurs

positions exactes. La complexité des items augmente avec le nombre de gommettes placées sur le corps du clown. Le nombre de gommettes que l'enfant est capable de repositionner correctement dans cette épreuve évalue l'empan de sa mémoire de travail. Un autre épreuve, dénommée "Counting span" par Case (1985), consiste à présenter au sujet des cartes sur lesquelles sont collées des gommettes en nombre variable. La tâche de l'enfant est de compter les gommettes collées sur la première carte, de retenir ce nombre, puis de compter les gommettes sur la seconde carte, de retenir ce nombre, etc., puis de rappeler ces différents nombres de gommettes après avoir compté la dernière carte. Cette épreuve est plus conforme que la précédente à la définition généralement acceptée de la mémoire de travail, une forme de mémoire qui assure la gestion simultanée du stockage (ici, retenir les nombres de gommettes des cartes déjà comptées) et du traitement de l'information (ici compter les gommettes de la carte suivante). La complexité des items varie selon le nombre de cartes à traiter et l'empan de la mémoire de travail est évalué par le nombre maximum de cartes rappelées correctement.

Une bonne dizaine d'épreuves de ce genre ont été mises au point dans le cadre des théories neo-piagetiennes pour évaluer la capacité centrale de traitement. Elles ont le plus souvent été utilisées dans des recherches visant à montrer l'accroissement de la capacité de traitement avec l'âge, ou à montrer que telle ou telle notion piagétienne, par exemple la conservation, ne pouvait être acquise que lorsque l'empan de la mémoire de travail correspondait au nombre schèmes à activer simultanément pour réussir la tâche. Ces épreuves sont restées du domaine de la recherche et n'ont pas non plus supplanté, chez les praticiens, les échelles classiques d'évaluation du niveau de développement intellectuel.

## 2.2. L'approche du traitement de l'information

### *2.2.1. Les tentatives de décomposition des processus cognitifs en jeu dans les tests d'intelligence*

En s'appuyant sur les différences individuelles de performance dans les tests, l'approche psychométrique classique a mis en évidence de grandes dimensions de l'activité cognitive (les facteurs), mais elle n'a pas permis d'accéder aux processus mentaux qui sous-tendent ces performances. L'approche du traitement de l'information a par contre apporté des modèles et des méthodes capables d'identifier les processus en jeu dans les tâches cognitives. Le développement de ces modèles et de ces méthodes a provoqué, vers la fin des années soixante-dix, le développement d'un courant de psychologie cognitive différentielle dont l'objectif était de s'appuyer sur la psychologie du traitement de l'information pour

identifier les processus sous-jacents aux grands facteurs de l'intelligence mis en évidence par l'approche psychométrique. On pourra trouver ailleurs (Huteau et Lautrey, 1999 ; Lautrey, 1995, 1996) une présentation plus détaillée de ce courant de recherche.

Une des démarches suivies pour atteindre cet objectif est la méthode corrélacionnelle. La première étape consiste à formuler des hypothèses sur les processus en jeu dans un test représentatif du facteur que l'on souhaite étudier. Pour chacun des processus dont on fait l'hypothèse qu'il intervient, la seconde étape consiste à trouver dans l'arsenal de la psychologie cognitive, un paradigme expérimental réputé isoler ce processus particulier. La troisième étape consiste à faire passer à un même échantillon de sujets le test et le paradigme expérimental retenus. S'il existe une corrélation entre la performance dans le test et l'efficacité dans le paradigme expérimental retenu, on en infère que les différences individuelles observées dans le test peuvent être expliquées, au moins en partie, par les différences d'efficacité (temps d'exécution ou nombre d'erreurs) dans le processus visé. C'est la démarche qui a été suivie par Hunt (1985) pour chercher si les différences individuelles dans les processus de codage phonologique, de codage sémantique, ou de balayage de la mémoire à court terme, pouvaient expliquer les différences de performance dans les tests de facteur verbal. Cette même démarche a aussi été suivie pour étudier d'autres facteurs. C'est ainsi, par exemple, qu'a aussi procédé Jensen pour tenter d'expliquer le facteur général d'intelligence par les différences individuelles dans la vitesse de traitement de l'information (La vitesse de traitement de l'information étant ici évaluée par le paradigme de Hick).

Les corrélations trouvées par cette méthode ont rarement dépassé ce que l'on a appelé la "barrière des .30". Autrement dit, aucun des processus élémentaires cernés par les paradigmes expérimentaux de la psychologie cognitive n'explique une part de variance substantielle de la performance dans des tâches complexes. Ce résultat met en cause certains des postulats qui sous-tendaient cette approche, par exemple celui de séquentialité et d'additivité des processus en jeu, celui selon lequel l'efficacité d'un processus peut être évaluée isolément de celle des autres processus intervenant dans la même tâche, ou indépendamment du contenu sur lequel il porte (Lautrey, 1996). Les différences de performance dans les tâches intellectuelles complexes semblent tenir davantage à la qualité des représentations et à l'orchestration des différents processus en jeu qu'à l'efficacité de tel ou tel processus élémentaire considéré isolément. C'est un point sur lequel nous aurons l'occasion de revenir en conclusion.

### 2.2.2. *La réinterprétation cognitive des facteurs de l'intelligence : quelques hypothèses*

Si l'approche corrélacionnelle, ou l'approche composantielle - qui en est une variante un peu plus sophistiquée dont on pourra trouver ailleurs une présentation sommaire (Lautrey, 1995) - ont donné des résultats plutôt décevants, le développement de la psychologie cognitive n'en suggère pas moins quelques hypothèses générales sur les processus sous-jacents aux différents facteurs isolés par l'approche psychométrique.

Nous évoquerons ici quelques unes de ces hypothèses en nous limitant au facteur général (g), et aux trois plus connus des facteurs figurant dans la seconde strate du modèle de Carroll (1993): le facteur d'intelligence fluide (Gf), le facteur d'intelligence cristallisée (Gc), et le facteur d'intelligence visuo-spatiale (Gv). Suivant Gustaffson (1984), nous assimilerons, dans cette interprétation, g à Gf.

#### 2.2.2.1. L'intelligence fluide

Le facteur Gf était qualifié de " fluide " par Cattell et Horn car dans leurs batteries d'épreuves, il saturait les tests qui évaluaient l'efficacité des opérations cognitives de base en faisant aussi peu que possible appel à des connaissances. Ils considéraient que ce facteur était principalement sous-tendu par l'efficacité des mécanismes neurobiologiques sous-jacents aux opérations de la pensée.

Du point de vue de la psychologie cognitive actuelle, le facteur Gf pourrait être interprété comme l'expression des différences individuelles dans la capacité attentionnelle ou dans l'empan de la mémoire de travail (selon le cadre conceptuel dans lequel on se place). Quelques données expérimentales peuvent étayer ce point de vue. Kyllonen et Christal (1990) ont fait passer à de larges échantillons de sujets une batterie d'épreuves comportant des tests de raisonnement et des épreuves de mémoire de travail. L'analyse factorielle de cette batterie a permis d'extraire deux facteurs, l'un saturant les tests de raisonnement et l'autre les épreuves de mémoire de travail. La corrélation de .80 trouvée entre ces deux facteurs conduit les auteurs à se demander si les tests de raisonnement sont autre chose que des tests de mémoire de travail. Sachant que le facteur de raisonnement est fortement saturé par le facteur gf, qui est lui-même assimilable au facteur g, on est conduit à conclure que les différences individuelles dans l'empan de la mémoire de travail (ou dans la capacité attentionnelle) contribuent fortement à la variance expliquée par les facteurs Gf et g. Un autre argument empirique peut être trouvé dans une recherche de Carpenter, Just, & Shell (1990) sur les processus cognitifs en jeu dans le test des Matrices Progressives de Raven. On sait que dans les batteries factorielles, ce test est un des plus saturés en facteur Gf et donc en

facteur  $g$  aussi. L'analyse des processus par lesquels les sujets résolvent les items des Matrices Progressives, puis la mise au point de programmes simulant les procédures de résolution observées chez les sujets, ont conduit Carpenter et al. à la conclusion qu'un élément déterminant de la qualité de la performance était le nombre de buts et de sous-buts que les sujets (ou le programme de simulation) pouvaient gérer simultanément tout en analysant les dessins qui figurent dans les lignes et les colonnes des matrices. Les items des Matrices sont de bons exemples de ces tâches qui demandent peu de connaissances (savoir distinguer un carré d'un losange, des hachures verticales de hachures horizontales, etc.). La résolution des items exige par contre que les différences repérées entre les figures comparées soient gardées en mémoire immédiate (stockage) pendant que continue le processus de comparaison (traitement); à un autre niveau de généralité, dès qu'une règle de transformation en ligne ou en colonne a été trouvée, elle doit aussi être conservée en mémoire (stockage) pendant que les autres règles de transformation sont recherchées (traitement). Cette gestion simultanée de l'activité de traitement des figures, et de stockage des produits intermédiaires de ce traitement, correspond exactement à la définition qui a été donnée un peu plus haut de la mémoire de travail.

Il faut néanmoins apporter plusieurs précisions sur l'hypothèse qui vient d'être avancée. Tout d'abord, la mémoire de travail (ou la capacité attentionnelle) n'est sans doute pas unitaire. Il y a actuellement débat entre les partisans d'un unique "réservoir" central de ressources cognitives (capacité attentionnelle ou place en mémoire de travail, selon le cadre théorique dans lequel on se place) et les partisans de plusieurs "réservoirs", un par domaine. Les faits empiriques actuellement disponibles me paraissent corroborer l'hypothèse d'une pluralité de ressources (cf par exemple Shah et Miyake, 1996). Bien que le rapprochement ne soit jamais évoqué dans la littérature sur la mémoire de travail, ce débat rappelle irrésistiblement celui qui a eu lieu entre partisans d'un facteur général d'intelligence et partisans de facteurs multiples. Les données empiriques manquent pour étayer ce point de vue, mais si une batterie comportant des épreuves de mémoire de travail de domaines variés (verbal, spatial et logique par exemple) était passée par un échantillon de sujets, je parierais volontiers que le modèle factoriel le mieux ajusté serait un modèle hiérarchique avec un facteur général de mémoire de travail et trois facteurs de groupe. Ceci permet de préciser l'hypothèse développée plus haut : c'est le facteur général de mémoire de travail qui devrait contribuer fortement à l'explication du facteur général d'intelligence. Ce facteur général de mémoire de travail devrait correspondre à une ressource cognitive centrale déterminant en partie le nombre d'opérations cognitives qui peuvent être délibérément activées ou inhibées

simultanément. Les différences individuelles dans cette capacité centrale devraient jouer un rôle crucial dans la résolution de tâches complexes de domaines assez variés et contribuer ainsi à l'apparition d'un facteur général. Seconde précision, le facteur général d'intelligence ne relève sans doute pas d'un mécanisme unique et les différences individuelles de mémoire de travail (ou de ressources attentionnelles) ne suffisent probablement pas à en rendre compte (un exemple en sera donné plus loin). Dernière précision : cette interprétation ne préjuge en rien du degré de généralité du facteur général (nous reviendrons sur ce dernier point dans la partie consacrée aux tentatives d'extension de la notion d'intelligence).

#### 2.2.2.2. L'intelligence cristallisée

La forme d'intelligence sous-jacente au facteur Gc était qualifiée de " cristallisée " par Cattell et Horn car elle reposait selon eux sur l'organisation des connaissances. Le facteur Gc sature néanmoins essentiellement des tests verbaux et des tests de connaissances et c'est ce qui permet de dire qu'il est très semblable au facteur " Verbal-Educationnel " de Burt-Vernon. En dénommant cette forme d'intelligence " cristallisée " plutôt que " verbale ", Cattell et Horn mettaient délibérément l'accent sur l'organisation des connaissances plutôt que sur le format symbolique par lequel celles-ci sont exprimées. En s'appuyant sur les données actuelles de la psychologie cognitive, on peut interpréter le facteur Gc comme l'expression du degré de structuration de la mémoire déclarative, essentiellement la mémoire sémantique, mais aussi la mémoire épisodique. Depuis Schank et Abelson (1977), les modélisations des réseaux reliant les nœuds conceptuels par diverses formes de liens ont évolué et ces modélisations sont maintenant plutôt de nature connexionniste. Il reste que l'organisation conceptuelle peut varier par sa richesse (nombre de nœuds conceptuels), sa complexité (nombre de relations unissant ces nœuds conceptuels), sa structuration (organisation de ces relations entre elles), etc. En prolongeant l'idée initiale de Cattell et Horn, on peut considérer que l'intelligence cristallisée est une forme d'intelligence dont l'efficacité dépend étroitement de la richesse, de la complexité et du degré de structuration de l'organisation conceptuelle. Il se trouve que dans l'intelligence humaine, les concepts sont généralement représentés par des mots, les relations entre les concepts par des expressions verbales, etc. La structuration conceptuelle et la structuration du langage sont donc indissociablement liées. La lecture accroît le vocabulaire ainsi que la structuration des concepts et cette dernière facilite en retour la compréhension de lecture. Il n'est donc pas surprenant que le facteur d'intelligence cristallisée sature les différentes sortes de tests dits " verbaux " y compris ceux de compréhension de lecture. Néanmoins, le terme d'intelligence " cristallisée ", même s'il n'est

pas le meilleur, met à juste titre l'accent sur le rôle de l'organisation conceptuelle dans le fonctionnement de cette forme d'intelligence.

### 2.2.2.3. L'intelligence visuo-spatiale

Dans une revue de question très documentée sur le facteur Gv, le facteur de visualisation spatiale, Lohman (1988) concluait que les différences individuelles d'efficacité dans les tests spatiaux tenaient surtout aux différences dans la qualité des représentations visuo-spatiales (plus que dans la vitesse d'exécution des processus en jeu par exemple). Les travaux de psychologie cognitive sur les processus d'imagerie mentale permettent de formuler maintenant quelques hypothèses sur les propriétés qui pourraient sous-tendre ce que Lohman appelait la qualité des représentations visuo-spatiales. Il semble que le caractère analogique de la représentation imagée tiende au fait qu'elle met en œuvre certains des circuits neuronaux qui sont en jeu dans la perception visuelle. Il a notamment été démontré par des techniques de neuro-imagerie que certaines des aires primaires visuelles, qui sont normalement activées par les neurones en provenance de la rétine au cours de la perception, sont aussi activées, en l'absence de toute perception, lorsque les sujets forment une image mentale (Kosslyn et al., 1993). Or les aires visuelles primaires en question ont la particularité d'être organisées de façon rétinotopique (en d'autres termes, la disposition spatiale des neurones de ces aires correspond à la disposition spatiale des cellules de la rétine qui leur envoient leurs afférences). De ce fait, la représentation imagée d'un objet de grande taille active une surface plus grande, dans ces aires, que la représentation imagée d'un objet de petite taille, tout comme lors de la perception visuelle d'un objet. Bien que ce fait soit encore l'objet d'expérimentations et de discussions, il apporte d'ores et déjà un éclairage intéressant sur les mécanismes par lesquels la représentation imagée conserve un isomorphisme spatial avec la perception visuelle. Les représentations imagées qui sont instanciées dans ces aires peuvent par ailleurs recevoir dans les aires associatives deux sortes de codage, un codage catégoriel et un codage métrique. Le codage catégoriel est un codage assez grossier du type dessus/ dessous, à gauche / à droite, etc. le codage métrique est par contre un codage qui conserve les distances euclidiennes. L'hémisphère gauche paraît être principalement en charge du codage catégoriel et l'hémisphère droit principalement en charge du codage métrique (Kosslyn et al., 1989). Il est vraisemblable que comme le suggère Martin (1998), la précision du codage métrique intervient dans la qualité de la représentation spatiale. Il semble donc que les propriétés spatiales et métriques des images mentales leur soient conférées par les caractéristiques structurales des systèmes neuronaux impliqués dans la perception visuelle (mais sans doute

aussi dans la motricité). Des différences dans la richesse, la précision, la structuration, des connexions dans ces réseaux peuvent rendre compte des différences dans la qualité des représentations spatiales. La performance dans des tâches spatiales complexes qui impliquent la représentation de transformations, des rotations mentales par exemple, paraît tenir davantage à la qualité des représentations sur lesquelles portent les transformations qu'à la vitesse d'exécution des processus de transformation eux-mêmes (vitesse de rotation par exemple).

Ici encore, il faut apporter plusieurs précisions aux hypothèses qui viennent d'être avancées. La résolution de problème spatiaux ne requiert pas nécessairement la formation d'images mentales. On sait de longue date que les sujets peuvent employer des stratégies différentes pour résoudre les tâches visuo-spatiales. Schultz (1991) a proposé de classer ces stratégies en trois grandes catégories : déplacement imaginaire de l'objet (par exemple, imaginer la rotation d'un objet dans l'espace), déplacement imaginaire de soi-même (imaginer son propre déplacement dans l'espace pour se représenter ce que l'on verrait d'un autre point de vue), et analyse des caractéristiques (repérer différentes caractéristiques de l'objet, par ex. l'ordre de parcours de certains repères, et opérer un raisonnement logique sur les changements de position de ces caractéristiques au cours de la transformation). Un exemple récent de recherche mettant en évidence ces différences individuelles de stratégie dans une tâche spatiale peut être trouvé dans Eme et Marquer (1999). Toutes ces stratégies de résolution ne font pas appel à la représentation imagée. La stratégie analytique (fondée sur l'analyse de caractéristiques) fait appel à un raisonnement abstrait, généralement effectué sur des propositions codées verbalement. Dans ce cas, la qualité de la représentation réside probablement dans le nombre de caractéristiques repérées, dans la finesse des discriminations verbales utilisées pour les dénommer, etc. La stratégie analytique donne lieu, chez ceux qui les maîtrisent bien, à d'aussi bonnes performances que les stratégies imagées. Pour les sujets qui utilisent la stratégie analytique, les processus de résolution mis en œuvre dans les tests spatiaux ne sont pas très différents de ceux qu'ils mettent en jeu dans les tests de raisonnement logique. Dans la mesure où les performances ne tiennent pas compte des stratégies par lesquelles ces performances sont obtenues, il est probable qu'une part de la variance commune entre le facteur Gv et le facteur Gf, donc une part de la variance dont rend compte le facteur g, tient au fait que certains sujets utilisent la même stratégie (ici un raisonnement propositionnel) pour résoudre des problèmes de domaines différents. On peut supposer que les corrélations entre facteurs seraient moins marquées, et donc la part de variance du facteur g plus réduite, si les sujets pris en compte dans les analyses factorielles



étaient seulement ceux qui utilisent des processus différents dans des domaines différents, par exemple une stratégie imagée dans les problèmes spatiaux et une stratégie propositionnelle dans les problèmes de raisonnement logique. C'est une des raisons pour lesquelles il était dit plus haut que le facteur général ne peut être probablement pas être expliqué par un mécanisme unique.

Ce qui précède permet d'énoncer de façon plus précise les deux points sur lesquels porte l'interprétation cognitive du facteur Gv avancée plus haut. Le premier est que la spécificité du facteur Gv tient à ce que les tâches qu'il sature sollicitent fortement la représentation imagée (autrement dit, il n'y aurait moins de distinction entre Gv et Gf si tous les sujets employaient la stratégie analytique dans tous les items des tests spatiaux). Le second est que lorsque les sujets s'appuient sur des représentations imagées pour résoudre les tests visuo-spatiaux, la qualité de leurs représentations (qui semble déterminer en bonne partie les différences individuelles de performance) dépend de la richesse et de la structuration des réseaux neuronaux qui confèrent leurs propriétés spatiales et métriques aux images mentales.

#### 2.2.2.4. Retour sur les tests piagétiens

Les hypothèses interprétatives qui viennent d'être avancées à propos des principaux facteurs de l'intelligence peuvent être appliquées aux autres catégories de tests. Concernant les tests piagétiens, nous avons vu que, selon les théories neo-piagétiennes, le franchissement de nouveaux stades de développement cognitif est sous-tendu par le développement de la capacité centrale de traitement (empan de la mémoire de travail ou capacité attentionnelle selon les auteurs). Le facteur général des batteries de tests piagétiens peut donc être interprété, comme le facteur général (ou le facteur Gf) des batteries de tests classiques d'intelligence, c'est à dire comme équivalent à un facteur général de mémoire de travail ou de capacité attentionnelle (en apportant à cette interprétation les mêmes précisions que dans l'interprétation du facteur Gf quant à la pluralité des formes de mémoire de travail ou de ressource attentionnelle, un point qui n'a pas reçu suffisamment d'attention dans les travaux neo-piagétiens). L'interprétation des deux grands facteurs de groupe généralement trouvés avec les épreuves piagétiennes suit la même logique. Le facteur logico-mathématique est assimilable à Gf (et donc à g); les travaux de Longeot (1969) étayaient cette interprétation. Le second facteur, qui sature les épreuves du domaine infralogique, défini par Piaget comme celui du continu, peut être assimilé au facteur Gv et donc recevoir la même interprétation. Les mécanismes cognitifs sous-jacents seraient ici ceux qui assurent la caractère analogique de la représentation imagée (cf. Lautrey & Chartier, 1987).

### 2.2.2.5. Retour sur les échelles d'intelligence

Terminons par la réinterprétation cognitive des échelles d'intelligence, et notamment par la réinterprétation des sous-échelles de la K-ABC, qui est l'échelle d'intelligence la plus récente. Les épreuves qui composent l'échelle de processus séquentiels sont, de fait, des épreuves de mémoire de travail portant sur différents supports symboliques (mots, gestes, nombres) et le facteur qui les sature est donc assimilable à *gf* et interprétable comme un facteur général de mémoire de travail. Les épreuves qui composent l'échelle de processus simultanés sont, de fait, des épreuves visuo-spatiales et le facteur correspondant peut être assimilé à *Gv*. Quant à l'échelle de connaissances, qui comporte à la fois des épreuves verbales et des épreuves de connaissances, elle correspond exactement à la définition du facteur d'intelligence cristallisée. Au total, le K-ABC évalue donc à peu près la même chose que les batteries factorielles ou les échelles d'intelligence classiques. Ses spécificités principales sont, comme nous l'avons vu plus haut, d'être donnée en temps libre et de mieux distinguer que dans la WISC les épreuves correspondant à l'intelligence cristallisée et à l'intelligence fluide. Il existe aussi dans la WISC des épreuves évaluant le facteur *Gf*, c'est à dire, selon notre interprétation, la mémoire de travail ou la capacité attentionnelle, mais elles sont peu nombreuses et ne forment pas une échelle spécifique. Il s'agit de l'épreuve de mémoire des chiffres (surtout celle des chiffres à l'envers), qui est dans l'échelle verbale, et du code, qui est dans l'échelle de performance. Kaufman (1975), a d'ailleurs été le premier à montrer que l'analyse factorielle de la WISC pouvait mettre en évidence, en plus du facteur verbal et du facteur de performance, un troisième facteur saturant ces deux épreuves et qu'il interprétait comme un facteur de résistance à la distractibilité, ce qui est assez proche de la capacité attentionnelle.

Cette réinterprétation des échelles du K-ABC ne sont pas incompatibles avec les hypothèses neuropsychologiques qui ont inspiré ses auteurs. L'échelle de processus séquentiels est supposée cerner les processus localisés dans le lobe frontal et c'est bien dans cette zone du cerveau que l'on situe généralement le siège du contrôle attentionnel et de la capacité générale de la mémoire de travail. L'échelle de processus simultanés est supposée cerner les processus localisés dans l'hémisphère droit. Il ne serait pas exact de dire que les processus de représentation imagée sont localisés dans l'hémisphère droit, les choses sont beaucoup plus complexes, mais le codage métrique, dont le rôle paraît essentiel dans la représentation analogique, paraît bien avoir lieu principalement dans l'hémisphère droit et ceci peut être rapproché du rôle généralement attribué à l'hémisphère droit dans le traitement

global, puisque selon Kosslyn et al (1989), cette forme de codage repose sur la forme de continuité qui 'engendrent les larges recouvrements d'unités réceptrices à champ large.

### 2.2.3. Discussion

Les travaux inspirés par la psychologie cognitive n'ont pas conduit à la construction d'instruments d'évaluation qui auraient supplanté les tests d'intelligence classiques. L'utilisation des tests inspirés par la théorie de Piaget reste marginale et les batteries de tâches inspirées des paradigmes expérimentaux de la psychologie cognitive pour évaluer l'efficacité des différents processus élémentaires supposés intervenir dans les tâches complexes sont restées du domaine de la recherche (cf. Roznowski, 1993, pour une tentative de mise au point d'une telle batterie).

La psychologie cognitive a par contre permis de réinterpréter les conduites observées dans les tests d'intelligence et à en donner une autre lecture, sans pour autant avoir réussi, pour l'instant, à proposer d'instruments plus intéressants pour les psychologues praticiens. Cette réinterprétation, du moins celle avancée plus haut unifie les différentes sortes de tests d'intelligence et réduit les dimensions évaluées à quelques composantes fondamentales de l'activité intellectuelle : l'empan de la mémoire de travail, la richesse de l'organisation conceptuelle, la qualité des représentations visuo-spatiales, etc.

Il faut nuancer cela en le limitant au domaine de l'évaluation de l'intelligence générale. La psychologie cognitive a par contre permis de construire des instruments de diagnostic de processus plus spécifiques, comme ceux en jeu dans le calcul ou la lecture, mais ceci sort du propos de ce chapitre.

### 2.3. L'élargissement du concept d'intelligence

Quel est le degré de généralité du facteur général d'intelligence ? Est-il général à toutes les formes d'intelligence ? Le débat entre partisans d'un facteur général et les partisans d'intelligences multiples continue sous d'autres formes. Il existe en effet un certain nombre de tentatives pour élargir le concept d'intelligence à d'autres formes d'adaptation cognitive, d'autres formes d'intelligence, que celles évaluées par les tests classiques.

La théorie la plus diffusée à l'heure actuelle sur les intelligences multiples, du moins dans les sciences de l'éducation, est celle de Gardner (1983). Cet auteur considère que l'on a affaire à une forme d'intelligence particulière lorsqu'un certain nombre de critères sont remplis. Ces critères sont, entre autres, l'existence de créateurs géniaux ayant manifesté des talents précoces dans le domaine, l'existence de localisations cérébrales spécifiques à cette

forme d'intelligence (se traduisant notamment par le fait que la lésion de cette zone n'affecte que cette forme d'intelligence), l'existence de cas d' "idiots savants" ou d' "autistes géniaux", c'est à dire de sujets manifestant une capacité extraordinaire dans un domaine, mais des capacités intellectuelles médiocres ailleurs (par exemple les cas de calculateurs prodiges), etc. En appliquant ces critères, Gardner pense pouvoir identifier sept formes d'intelligence qu'il considère comme indépendantes entre elles : les intelligences logico-mathématique, langagière, spatiale, musicale, kinesthésique, interpersonnelle, et intrapersonnelle. On peut voir des analogies entre, certains des facteurs d'intelligence - fluide, cristallisée, visuo-spatiale, auditive - et certaines des formes d'intelligence distinguées par Gardner - respectivement logico-mathématique, langagière, spatiale, musicale. Le fait que les recherches psychométriques aient montré l'existence de corrélations entre ces facteurs pourrait, si l'on admettait cette analogie, réfuter l'hypothèse d'indépendance entre les différentes formes d'intelligence correspondantes. Gardner récuse toutefois les tests et la méthodologie psychométrique comme moyen d'évaluation des différentes formes d'intelligence qu'il distingue. L'évaluation doit selon lui être faite par les maîtres, par des méthodes qualitatives qui excluent la standardisation, dans le décours d'un programme d'enseignement conçu pour solliciter ces différentes formes d'intelligence et révéler les talents précoces. A notre connaissance, aucune donnée quantifiée fiable n'a été recueillie dans le cadre de ce courant de recherche, qui serait susceptible de mettre à l'épreuve l'hypothèse d'indépendance entre les différentes formes d'intelligence distinguées. Les conditions que met Gardner à l'évaluation de ces différentes formes d'intelligence rendent, de fait, l'hypothèse d'indépendance irréfutable, ce qui est de nature à assurer une certaine longévité à la théorie...

Si la théorie des intelligences multiples met en question l'existence d'un facteur général d'intelligence, elle n'est donc pas de nature à en apporter la réfutation empirique. Il existe toutefois d'autres tentatives d'élargissement du concept d'intelligence, qui visent aussi à montrer l'existence de formes d'intelligence peu liées à celle évaluée par les tests classiques, et qui ont des fondements empiriques plus solides. C'est ainsi qu'ont été mis au point des instruments d'évaluation de l'intelligence sociale, de l'intelligence émotionnelle ou de l'intelligence pratique.

L'intelligence sociale est la forme d'intelligence qui permet de comprendre autrui (ses pensées, ses sentiments) et d'agir efficacement sur lui (obtenir son adhésion, modifier son comportement) en situation d'interaction sociale. Divers instruments d'évaluation ont été mis au point pour cerner cette forme d'intelligence, mises en situation, questionnaires, épreuves

objectives. On pourra trouver ailleurs un état de la question sur l'évaluation de l'intelligence sociale (de Bonis & Huteau, 1994) et un échantillon des épreuves utilisées (Pelachano, 1994).

L'intelligence émotionnelle est la capacité à connaître et à réguler ses propres émotions ainsi que celles des autres, et à utiliser cette information pour guider la réflexion et l'action. Des épreuves d'évaluation ont été construites pour cerner différentes dimensions de cette forme d'intelligence, comme le degré d'attention accordé à ses propres émotions, la capacité d'identification des émotions chez soi-même et autrui, la clarté des émotions, etc (Salovey & Sluyter, 1997).

Les faibles corrélations généralement observées entre la performance dans les tests d'intelligence classiques et les évaluations de l'adaptation dans la vie quotidienne et dans la vie professionnelle ont conduit certains auteurs à considérer que cette forme d'adaptation reposait sur une forme d'intelligence spécifique, dite "pratique", qui est souvent contrastée avec l'intelligence dite "académique" mesurée par les tests classiques et requise par les apprentissages scolaires. Selon Sternberg, Wagner, Williams, Horvath, (1998), la forme de connaissance sur laquelle opère l'intelligence pratique serait essentiellement "tacite". De même que l'intelligence académique permet l'acquisition rapide et l'utilisation efficace de connaissances formelles (celles qu'enseigne l'école et que requièrent les tests), l'intelligence pratique permet l'acquisition rapide et efficace de connaissances tacites (celles que requièrent l'adaptation à la vie quotidienne et l'expertise professionnelle). Ces auteurs attribuent trois caractéristiques essentielles à la connaissance tacite. Tout d'abord, elle est procédurale, c'est à dire totalement intriquée dans l'action, acquise par et pour l'usage qui peut en être fait. Elle peut être modélisée par un système de production (Sternberg et al, 1998) ou par un réseau connexionniste (Wagner, 1990). Ensuite, la connaissance tacite est instrumentale. Elle n'a de sens que pour atteindre les buts que les individus valorisent. Enfin, elle est acquise sans l'aide directe des autres, sans enseignement explicite. C'est à cette dernière caractéristique que renvoie le qualificatif "tacite". Divers instruments ont là aussi été mis au point pour évaluer cette forme d'intelligence. Il s'agit en général d'exposés de cas qui peuvent se présenter dans la pratique d'une profession donnée (par exemple celle de cadre d'entreprise), et pour lesquels plusieurs réactions possibles sont proposées. Les sujets doivent juger sur une échelle d'estimation, pour chaque cas exposé, le degré auquel cette réaction est adaptée à la situation, et leurs réponses sont comparées à celles qui ont été données par des juges reconnus pour leur expertise dans le domaine considéré. Un échantillon de ces instruments d'évaluation peut être trouvé dans Sternberg et al (1998).

Pour les trois formes d'intelligence qui viennent d'être évoquées – intelligence sociale, intelligence émotionnelle, et intelligence pratique – les évaluations réalisées corrélaient assez peu, voire parfois pas du tout avec la performance dans des tests de facteur g. Il est cependant difficile, dans l'état actuel de la recherche, de savoir si ces faibles corrélations sont attribuables à des insuffisances dans les qualités psychométriques de ces nouveaux instruments ou à une indépendance plus fondamentale.

Le facteur général que l'on trouve d'habitude avec une batterie de tests d'intelligence classiques est-il plus général encore et peut-il être étendu à ces autres formes d'intelligence ou bien s'agit-il de formes d'adaptation cognitive complètement indépendantes de celles mesurées jusqu'ici avec les tests classiques ? Il est possible, et même vraisemblable, que le facteur qui a été jusqu'ici considéré comme général soit en fait spécifique à l'intelligence académique. Il se peut aussi que ce facteur général soit sous-tendu par des mécanismes de contrôle attentionnel qui sont à l'œuvre dans toute tâche cognitive tant soit peu complexe, y compris dans celles qui évaluent les autres formes d'intelligence. Cette question est, dans l'état actuel des choses, encore ouverte.

### **3. Conclusion.**

Revenons maintenant au paradoxe souligné dans l'introduction. Une enquête internationale menée dans 44 pays (Oakland & HU, 1992) montre que les 5 tests d'intelligence actuellement les plus utilisés dans le monde sont la WAIS, la WISC, les Matrices Progressives de Raven, le Stanford-Binet et le DAT (Differential Aptitude Test, une batterie factorielle). Ces tests ont certes été régulièrement révisés et réétalonnés, mais tous ont été conçus il y a plus d'un demi-siècle. Comme nous l'avons vu, les idées sur l'intelligence ont évolué depuis la période dans laquelle ces tests ont été conçus, notamment sous l'influence de la théorie de Piaget et de la psychologie cognitive inspirée des modèles de traitement de l'information. Pourquoi ces évolutions théoriques n'ont-elles donné lieu à aucun instrument d'évaluation capable de supplanter les tests classiques dans la pratique de l'examen psychologique ? Et pourquoi les tests classiques ont-ils si bien résisté au temps et aux renouvellement théoriques ?

Je ne crois pas que cette bonne résistance tienne à la robustesse des théories de l'intelligence qui ont inspiré les pionniers de la psychométrie. Binet et Wechsler partageaient certes quelques idées assez robustes, qui ont joué un rôle fondateur dans la conception de leurs échelles et qui sont toujours d'actualité : 1) les différences individuelles d'intelligence doivent être évaluées dans des tâches faisant appel à ce que Binet appelait les processus

supérieurs et non dans des tâches faisant appel aux processus élémentaires ; 2) Elles doivent être évaluées à travers différents aspects de l'activité intellectuelle, par exemple le jugement, le raisonnement, la mémoire, chez Binet ; 3) Ce qui compte n'est pas tant la performance dans telle ou telle aptitude élémentaire, que ce que Wechsler appelait "l'orchestration" de ces différentes aptitudes élémentaires dans les tâches complexes. Ces quelques idées ne constituaient pas pour autant une théorie explicite et articulée de l'activité intellectuelle.

La bonne résistance des tests classiques au temps tient plutôt au caractère très pragmatique de l'approche des pionniers de la psychométrie. Binet a essayé, pendant une bonne quinzaine d'années, beaucoup de choses pour tenter de mesurer l'intelligence. Il a retenu les situations qui différençaient bien les sujets retardés des sujets sans retard avéré, les sujets jeunes des sujets plus âgés, les bons élèves de ceux que les instituteurs de l'école de la Grange aux belles trouvaient moins bons. En procédant ainsi, ces pionniers ont trouvé un moyen indirect, empirique, d'évaluer une capacité d'adaptation cognitive relativement générale. Évaluée à travers les produits d'adaptations cognitives passées, cette capacité renseigne – sous certaines conditions – sur les possibilités d'adaptation cognitive future (une condition essentielle étant que les sujets aient baigné dans la culture au sein de laquelle a été élaboré le test, et que le pronostic porte sur l'adaptation future au sein de cette même culture).

Les deux raisons qui m'incitent à penser que la capacité d'adaptation cognitive cernée par cette approche très pragmatique est assez générale sont les suivantes. La première s'appuie sur des constatations empiriques. On observe en effet que ces tests corrélaient fortement avec ceux qui ont été construits ultérieurement à partir de cadres théoriques assez différents. Par exemple, les Matrices Progressives de Raven ont été construites à partir d'un cadre théorique très différent de celui qui a guidé la construction des échelles de Wechsler. Raven s'est inspiré des idées de Spearman sur les processus en jeu dans les tâches fortement saturées en facteur g. Ses matrices sont des situations qui sollicitent à la fois, de façon très ingénieuse, les deux processus auxquels Spearman accordait un rôle central : l' "éducation" de relations (inférer les règles de transformation des figures en ligne et en colonne) et l' "éducation" de corrélats (en déduire l'élément manquant dans la case vide de la matrice). Or la corrélation entre le score aux matrices Progressives de Raven et le QI à la WAIS est de l'ordre de .80. La théorie de Piaget a conduit à imaginer des situations d'évaluation très différentes de celles qui avaient été utilisées dans les tests d'intelligence précédents. Les épreuves de conservation ou d'inclusion de classes ont été imaginées à partir d'hypothèses qui étaient parfaitement étrangères aux conceptions de Wechsler ou de Spearman. Or la corrélation entre le QI à la WISC et le score total dans des batteries d'épreuves piagétienne

est aussi de l'ordre de .70 à .80 selon les études (cf. Lautrey, à paraître). La K-ABC a été élaborée à partir d'un cadre théorique inspiré de la neuropsychologie et mettant l'accent sur l'efficacité des processus. Certaines épreuves de l'échelle de processus séquentiels sont par exemple inspirées d'épreuves que Luria avait mises au point pour examiner ses patients. La corrélation entre le QI à la K-ABC et le QI à la WISC n'en est pas moins de l'ordre de .70. La seconde raison qui m'incite à penser que les tests classiques cernent une capacité d'adaptation cognitive relativement générale est que d'un point de vue plus théorique, le cadre conceptuel de la psychologie cognitive permet d'interpréter ce que mesurent ces tests de la même manière que ce que mesurent les autres tests. Comme nous l'avons vu plus haut, tous mesurent quelques grandes dimensions de l'activité cognitive comme la richesse et la qualité de l'organisation conceptuelle (qui va de pair avec la richesse et la qualité de la structuration du langage), la qualité de la représentation visuo-spatiale, la capacité générale de la mémoire de travail (ou capacité attentionnelle). Ceci explique d'ailleurs la taille des corrélations entre les différents tests.

En résumé donc, Binet ou Wechsler ont élaboré leurs échelles d'intelligence de façon très pragmatique, mais le fait que d'autres tests, élaborés à partir de cadres théoriques plus riches et assez diverses, mesurent à peu près la même chose, leur confère – a posteriori – une validité théorique qui leur faisait défaut au départ. Par leur démarche pragmatique, ils ont permis de réunir un ensemble de faits dont toute théorie de l'intelligence devra pouvoir rendre compte. Si la capacité d'adaptation cognitive que mesurent ces tests est assez générale, elle n'est sans doute pas pour autant toute l'intelligence. Ils n'évaluent probablement que ce que nous avons appelé plus haut l'intelligence académique, c'est à dire la forme d'adaptation cognitive en jeu dans des apprentissages explicites, comme ceux qui sont faits dans le système scolaire. Mais même si ce n'est qu'une des formes de l'intelligence, il reste que dans des sociétés comme les nôtres, elle joue un rôle suffisamment crucial, pour les individus et pour les sociétés, pour que son évaluation soit un enjeu scientifique et social important.

Si les tests classiques mesurent au fond à peu près la même chose que ceux qui sont fondés sur des théories plus récentes, pourquoi les abandonner au profit de ces derniers ? Il faudrait pour cela que les tests fondés sur d'autres cadres théoriques apportent quelque chose en plus. C'était bien l'objectif des recherches appuyées sur la psychologie cognitive, qui visaient à fournir une analyse des processus qui échappe aux tests classiques. Ces recherches ont apporté des éléments permettant d'enrichir l'interprétation des conduites observées, mais elle ont échoué à fournir de nouveaux instruments d'évaluation capables de renseigner à la fois sur la performance et sur les processus par lesquels celle-ci a été obtenue. Un détour par



l'épistémologie de la psychologie peut aider à comprendre les raisons de cet échec. Reuchlin (1995) montre que depuis les débuts de son histoire, la psychologie a eu des difficultés à sortir d'une oscillation entre une approche qu'il appelle "holiste" et une approche qu'il appelle "élémentiste". Le holisme considère les entités psychiques comme des totalités non décomposables, tandis que l'élémentisme les considère comme analysables en éléments indépendants les uns des autres et, donc, isolables. Des notions comme celles de QI, ou celle de facteur, s'inscrivent dans la tradition holiste. Elles rendent compte d'une performance globale, dont on ignore sur quels éléments elle repose. Les approches corrélacionnelle et composantielle, qui se sont appuyées sur les méthodes de la psychologie cognitive pour analyser les processus en jeu dans les tests peuvent être interprétées comme un retour à l'élémentisme (qui prévalait déjà, avant Binet, dans les tentatives de mesure de l'intelligence, par exemple chez Galton ou chez Cattell). On a pensé pouvoir recomposer et comprendre les différences observées au niveau de la totalité -le QI ou le facteur- en mettant bout à bout les différences observées dans les différents processus élémentaires qui ont pu être isolés.

Au point où nous en sommes, nous savons évaluer les différences individuelles au niveau de la totalité et au niveau des processus élémentaires. Le chaînon manquant est la façon dont ces derniers sont "orchestrés" pour donner lieu à la totalité. Les modes de recombinaison qui ont été envisagés (en général des modèles séquentiels et additifs) sont manifestement inadaptés. Comme le suggère Reuchlin (1995), la solution est probablement dans une approche qui ne soit ni holiste ni élémentiste, mais structuraliste, c'est à dire capable de spécifier les relations entre éléments qui rendent les propriétés de la totalité irréductibles à celles des éléments. Toutefois, les approches structuralistes jusqu'ici adoptées dans le domaine de l'intelligence, la théorie de Piaget et l'approche factorielle, ont décrit des structures qui sont statiques et qui, pour cette raison, ne sont pas appropriées pour trouver le chaînon manquant. Il faut en effet parvenir à comprendre comment les multiples processus mentaux qui entrent en jeu dans la performance interagissent pour former un système et comment ce système évolue au cours même de l'activité intellectuelle. C'est donc une sorte de structure dynamique qu'il faut tenter d'appréhender, du type de celles que forment les systèmes dynamiques non linéaires. Ceci suppose une réorientation du cadre théorique de la recherche sur l'intelligence, mais aussi sans doute une réorientation des méthodes d'observation. La modélisation de la dynamique de la pensée exige en effet des observables appropriés, c'est à dire recueillis au niveau de l'individu, dans le déroulement même de l'activité intellectuelle, avec une très forte densité temporelle. Il est vraisemblable que l'analyse en temps réel de ces observables, et leur comparaison à des modélisations permettant de les interpréter, dépasse les

capacités de l'observateur humain et doivent être pris en partie en charge par un dispositif informatique. Quoi qu'il en soit, le paradoxe qui a servi de fil directeur à cet exposé laisse penser que seuls des instruments donnant les moyens d'évaluer à la fois la performance globale et la façon dont les processus cognitifs ont été orchestrés pour la produire, auront quelque chance de supplanter les tests d'intelligence classiques.

### **Bibliographie**

- Binet A., Simon T. (1905). Méthodes nouvelles pour le diagnostic du niveau intellectuel des anormaux. L'Année Psychologique, 11, 191-244.
- de Bonis, M., & Huteau, M. (1994). Introduction : facettes de l'intelligence sociale. Revue Européenne de Psychologie Appliquée, 44 (4), 263-266.
- Carpenter P.A., Just M.A., Shell P. (1990). What one intelligence test measures : a theoretical account of the processing in the Raven progressive matrices test. Psychological Review, 97, 404-431.
- Carroll, J.B. (1993). Human cognitive abilities. Cambridge : Cambridge University Press.
- Case, R. (1985). Intellectual Development. Birth to Adulthood. New-York : Academic Press.
- Cattell, R.B. (1971). Abilities : Their structure, growth, and action. Boston : Houghton Mifflin.
- Eme, P.E. & Marquer , J. (1999) Individual strategies in a spatial task and how they relate to aptitudes. European journal of psychology of education, 14, 89-108.
- Gardner, H. (1983). Frames of mind : the theory of multiple intelligences. New York : Basic Books. (traduction française : Les Formes de l'intelligence, Paris : Odile Jacob, 1997)
- Gardner H. (1993). Multiple intelligences. The theory in practice. New-York : Basicbooks. (traduction française : Les intelligences multiples. Paris : Retz., 1996).
- Gibello, B. (1984). L'enfant à l'intelligence troublée. Paris: Le Centurion
- Gustaffson, J.E. (1984). A unifying model for the structure of intellectual abilities. Intelligence, 8, 179-203.
- Hunt, E. (1985). Verbal ability. In R. J. Sternberg (Ed.), Human abilities: An information processing approach (pp. 31-58). New-York: Freeman and Co.
- Huteau, M., & Lautrey, J. (1997). Les tests d'intelligence. Paris: Editions La Découverte.

- Huteau, M., & Lautrey, J. (1999). Evaluer l'intelligence – Psychométrie cognitive. Paris : PUF.
- Jensen, A.R. (1987). Individual differences in the Hick paradigm. In P.A. Vernon (Ed.), Speed of information processing and intelligence (p. 101-175). Norwood, NJ : Ablex.
- Kaufman, A. (1975). Factor analysis of the WISC-R at 11 age levels between 6;6 and 16;6 years. Journal of Consulting and Clinical Psychology, 43, 135-147.
- Kaufman, A. S., & Kaufman, N. L. (1993). K-ABC. Batterie pour l'examen psychologique de l'enfant.- Manuel d'interprétation. Paris: Editions du centre de Psychologie Appliquée.
- Kosslyn, S. et al (1989). Evidence for two types of spatial representations : Hemisphere specialization for categorical and coordinate relations. Journal of Experimental Psychology : Human Perception and Performance, 15, 723-735.
- Kosslyn, S. et al (1993). Visual mental imagery activates topographically organized visual cortex : PET investigations. Journal of Cognitive Neuroscience, 5, 263-287.
- Kyllonen, P., & Christal, R. (1990). Reasoning ability is (little more than) working memory capacity ? Intelligence, 14 (4), 389-433.
- Lautrey, J. (1980). La variabilité intra-individuelle du niveau de développement opératoire et ses implications théoriques. Bulletin de Psychologie, 33, 685-697.
- Lautrey, J. (1995). Les apports de la psychologie cognitive à la compréhension des différences en matière d'intelligence et de réussite scolaire. In G. Mialaret (Ed.), Intelligences, Scolarités et Réussites. Grenoble: Editions La Pensée sauvage.
- Lautrey J. (1996). La recherche des "particules élémentaires" de l'intelligence : une impasse ? Psychologie Française, 41, 23-34.
- Lautrey, J. (à paraître). Is there a general factor of cognitive development ? In R.J. Sternberg & E. Grigorenko (Eds.), How General is the General Factor of Intelligence ? Hillsdale, NJ : Lawrence Erlbaum.
- Lautrey, J., & Chartier, D. (1987). Images mentales de transformations et opérations cognitives: une revue critique des études développementales. L'Année Psychologique, 87, 581-602.
- Lautrey, J., de Ribaupierre, A., Rieben, L. (1986). Les différences dans la forme du développement cognitif évalué avec des épreuves piagétienne. Cahiers de Psychologie Cognitive, 6, 575-613.
- Lohman, D.F. (1988). Spatial abilities as traits, processes, and knowledge. In R.J. Sternberg (Ed.), Advances in the psychology of human intelligence. (Vol. 4). Hillsdale, NJ : Erlbaum.
- Longeot F. (1969). Psychologie différentielle et théorie opératoire de l'intelligence. Paris: Dunod.

- Longeot F. (1974 / 1979). L'échelle de développement de la pensée logique.. Paris : Éditions des Etablissements d' Applications Psychotechniques
- Martin, R. (1998). Encodage spatial et intelligence. Thèse de doctorat. Document non publié. Université de Nancy.
- Meljac, C. & Lemmel, G. (1999). Manuel de l'UDN II. Paris : ECPA
- Oakland, T., & Hu, S. (1992). The top 10 tests used with children and youth worldwide. Bulletin of the International Test Commission, 19, 99-120.
- Pascual-Leone, J. (1987). Organismic processes for neo-piagetian theories : A dialectical causal account of cognitive development. International Journal of Psychology, 22, 531-570.
- Pelachano, V. (1994). Assessment of interpersonal skills across life-span. Revue Européenne de Psychologie Appliquée, 44, (4), 281-287.
- Reuchlin, M. (1964). L'intelligence : conception génétique opératoire et conception factorielle. Revue Suisse de Psychologie, 23, 113-134.
- Reuchlin, M. (1995). Totalités, Eléments, Structures en Psychologie. Paris: PUF
- de Ribaupierre, A. (1997). Les modèles néo-piagétiens : quoi de nouveau ? Psychologie Française, 42 (1), 9-21.
- de Ribaupierre, A., & Bayeux, C. (1994). Developmental change in a spatial task of attentional capacity : An essay toward an integration of two working memory models. International Journal of Behavioral Development, 17 (1), 5-35.
- Roznowski, M. (1993). Measures of cognitive processes: Their stability and other psychometric and measurement properties. Intelligence, 17, 361-388.
- Salovey, P.& Sluyter, D. (Eds). Emotional development and emotional intelligence : Educational implications. New-York : Basicbooks.
- Schank, R.C., & Abelson, R. (1977). Scripts, plans, goals, and understanding. Hillsdale, NJ : Erlbaum.
- Schultz, K. (1991). The contribution of solution strategy to spatial performance. Canadian Journal of psychology, 45, 474-491.
- Shah, P., & Miyake, A. (1996). The separability of working memory resources for spatial thinking and language processing : an individual differences approach. Journal of Experimental Psychology : General, 125, 4-27.
- Spearman, C.E. (1904). General intelligence objectively measured and determined. American Journal of Psychology, 15, 201-209.

Sternberg, R.J., Wagner, R.K., Williams, W.M., Horvath, J.A. (1998). Testing common sense. American Psychologist. Soumis

Vernon, P.E.(1961). The structure of human abilities (2<sup>ème</sup> édition). Londres : Methuen

Wagner, R.K. (1990). Products and processes of practical reasoning. International Journal of Educational Research, 14, 437-454.