

**Les compétences à apprendre.
Evaluation chez l'adulte.**

Daniel Pasquier

Evaluer la compétence à apprendre de l'adulte.

Evaluer la compétence à apprendre de l'adulte.

Daniel Pasquier

Du même auteur :

Pasquier, D. (1986). *Comprendre l'échec scolaire, Essai théorique de psychologie de l'éducation*. Bourges : chez l'auteur.

Pasquier, D. (1990). *L'évaluation en pédagogie. Du bon usage des tests*. Issy-les-Moulineaux : Editions E.A.P.

Pasquier D. (1992). *Agir pour la réussite scolaire*. Paris : Hachette.

Préface

par Claude Lemoine (1)

Avec "Evaluer la compétence à apprendre de l'adulte" de Daniel Pasquier, il faut saluer d'abord la sortie d'un ouvrage situant théoriquement, analysant et validant un test de psychologie en français. Cela se fait trop rare et renvoie aux critiques dévastatrices sur les tests en général avancées dans les années 1960 par des intellectuels anti-sélection. Supprimez les tests, la différenciation sociale n'en existera pourtant pas moins. Elle prendra même des formes moins contraintes et plus féroces. Tout se passe comme si des procédés sans filet, sans contrôle et sans transparence étaient plus acceptables qu'une démarche rigoureuse, méthodique et attentive à ses limites.

Mais "évaluer la compétence à apprendre" se situe à un niveau de difficulté plus élevé. Il ne s'agit pas seulement de mesurer une compétence ordinaire au sens d'une habileté particulière à faire ou à comprendre quelque chose qui relèverait d'un savoir ou d'une forme d'intelligence. L'objectif est de découvrir si quelqu'un va pouvoir apprendre. On se situe d'emblée à un niveau "méta", qui déterminerait tous les autres. L'attention ne porte plus sur un existant mais sur la capacité à évoluer dans le temps et à découvrir ce que l'on ne sait pas encore.

Ne risque-t-on pas non seulement d'apprendre qu'on ne sait pas grand chose mais en plus qu'on ne pourra rien y changer, quoi qu'on fasse ? Ou au contraire découvrira-t-on qu'on a des possibilités insoupçonnées, non reconnues jusqu'à ce jour et donc laissées pour compte ? Dans l'usage courant, l'éducabilité ne sert-elle pas à décider si un salarié est capable ou non de suivre une formation, ce qui fait du terme un couperet, en situant une personne du bon ou du mauvais côté de la barrière ?

On pressent d'emblée les enjeux individuels et sociaux. Cependant les tests ne sont pas critiquables pour l'effort de rigueur et de méthode qu'ils exigent mais par l'utilisation mentale et sociale qui en est faite et qui a bien peu à voir avec un quelconque niveau de scientificité. Peut-on miser sur un tel, l'envoyer en formation et avoir une chance suffisante de rentabiliser l'investissement envisagé ? On perçoit bien que ce genre de question de casino, où l'on prend ses semblables pour une marchandise jetable, s'oppose radicalement à la conception d'une personne perfectible, en devenir, susceptible de changer et d'innover, chère à tous les grands pédagogues.

Face à cette question importante et sensible, Daniel Pasquier se situe résolument du côté pédagogique et dans le courant de la remédiation, et cela par l'intérieur même de la mise en place de son test. Paradoxalement, la question brûlante qu'il pose le pousse à revoir la façon même de concevoir l'usage du test. Celui-ci ne peut plus se suffire à créer des catégorisations positives ou négatives dans lesquelles on place les sujets, et qui déterminent leur statut social. Il peut devenir en lui-même une méthode pour faire découvrir à ceux qui le remplissent les moyens dont ils disposent pour évoluer. Le test devient alors non plus une mesure d'un état statique, intangible, mais l'instrument par lequel l'intéressé prend conscience de ses possibilités, et par là même se transforme. C'est ce que l'auteur appelle l'évaluation dynamique.

La démarche proposée dans l'ouvrage est très méthodique. Une première partie, théorique, explicite ce qu'est exactement un test, comment il est élaboré, comment il convient de l'utiliser et sous quelles conditions. Puis un chapitre traite des tests d'intelligence, des fondements de leur construction, de leur conception factorielle et de leur limite de validité, notamment en terme de prévision. Ce détour permet de mieux situer l'originalité de l'évaluation dynamique du potentiel d'apprentissage qui se rapproche d'une évaluation formative et non pas sommative, ce qui suppose une mise en situation d'apprentissage. Un long chapitre traite de cette dynamique et de la manière de la discerner, chez l'enfant et chez l'adulte. Puis une seconde partie présente le test dans tous les détails ainsi que

les expériences relatives à sa validation. L'intérêt réside surtout dans un double relevé, avant - après, séparé par une phase d'apprentissage proprement dit. Dans ce sens, l'éducabilité de vise pas à poser la question de savoir si quelqu'un est éduicable, mais à rechercher de quelle façon on peut l'aider à apprendre et à évoluer.

Mais une autre orientation pointe également dans cette démarche. Non seulement celle-ci ne se veut pas un couperet à un instant donné, mais elle ne se limite pas non plus à un simple diagnostic fondé donné par un expert. Elle vise aussi à dégager les pistes conduisant le sujet à de nouvelles acquisitions, c'est-à-dire à repérer par lui-même les formes de sa compétence à apprendre, au cours du test et des échanges pendant l'apprentissage même. On rejoint ici le courant des conceptions éducatives de l'orientation (Blanchard & Zerwetz, 1998) et l'un des objectifs des bilans de compétences qui consiste à permettre à l'intéressé de repérer ses capacités dans le cadre d'un dispositif méthodique proposé par le conseiller (Lemoine, 2002). Dans cette perspective le test devient un moyen de développement personnel au sens où il invite la personne à découvrir sa façon d'apprendre et à en prendre conscience pour en tirer parti pour elle-même. On pourrait alors se demander s'il faut encore appeler test cette démarche qui, à la réflexion, constitue une méthode validée de repérage des compétences à développer (RCD).

Il est remarquable à cet égard que la présentation très complète et pédagogique de l'ouvrage fait écho à celle de la démarche détaillée de la méthode qui nous est proposée, et qui peut devenir un moyen de connaissance, pour l'intéressé lui-même, de la façon dont il apprend. Nul doute que cet ouvrage ne peut qu'enrichir la réflexion et la pratique professionnelles des psychologues et des conseillers qui interviennent auprès d'un large public en quête de connaissance sur soi et sur ses compétences.

(1) Professeur de psychologie à l'Université Lille 3, président de l'Association Internationale de Psychologie du Travail de Langue Française (AIPTLF), co-rédacteur de la revue Psychologie du Travail et des Organisations.

Blanchard, S. & Zerwetz, M. (1998). L'émergence des conceptions éducatives de l'orientation en France : quelques jalons historiques. *Questions d'Orientation*, 3, 27-70.

Lemoine, C. (2002). *Se former au bilan de compétences*. Paris, Dunod.

Première section : théorie

Les principes qui fondent l'évaluation dynamique du potentiel d'apprentissage datent des années trente. Cette forme d'évaluation ne s'est développée jusqu'à ce jour que de manière marginale et par à-coups. En France, l'évaluation dynamique reste encore à ce jour une pratique rare.

Il serait certainement passionnant de s'interroger sur les raisons de cette mise à l'écart, de cet abandon, de ce désintérêt pour une technique d'évaluation des capacités d'apprentissage de toute première importance. Très succinctement, on peut avancer quelques hypothèses pour rendre compte de cette pauvreté créative des psychologues français.

Du côté des praticiens, il est vrai que les techniques d'évaluation scientifique, les tests, furent considérés comme suspects et plus ou moins laissés de côté depuis 1968. Les formateurs se sont plutôt laissés bercer par la vague des pédagogies dites non-directives et l'entretien ou le questionnaire à chaud ont tenu lieu de procédures d'évaluation. Les psychologues scolaires ont eu tendance à opter pour une approche psychanalytique des difficultés d'apprentissage, un peu trop systématiquement vues comme symptômes de névroses infantiles. Les conseillers d'information et d'orientation ont, de leur côté, mis l'accent sur leur fonction d'information sur les différents cursus scolaires et de formation et sur l'éducation des choix. Ils ont laissé aux seuls professeurs le soin de jauger les capacités intellectuelles au travers des échelles de notations scolaires bien que toutes les études docimologiques démontrent l'inquiétant manque de fiabilité de ce type de notation (Pasquier, 1990).

Quant à la demande des entreprises, elle se centre toujours davantage sur l'opérationnalité immédiate du salarié. Ou encore, elle dérive vers des formes de sélection subjectives peu ou non validées (par exemple la graphologie ou les

questionnaires de personnalité), ou elle ne se concrétise que par un simple entretien d'embauche.

Du côté des universitaires une raison domine les autres : l'emprise de la psychologie anglo-saxonne est telle que, *grosso modo*, la recherche européenne en psychologie se nourrit de la traduction des textes américains. "On peut... légitimement se demander si les Européens se contentent de reproduire les méthodes, voire de discuter les modèles nés en Amérique, sans avoir de capacité d'innovation propre." (Lévy-Leboyer, 1994)

Bien sûr, Binet, dès 1911, postule l'éducabilité de l'intelligence et dans les années trente, Rey et Ombredane opérationnalisent la mesure du potentiel d'apprentissage. En solitaire, Hurtig reprend ces travaux dans les années soixante, mais il ne sera pas suivi. Il faut attendre les années quatre-vingt dix pour qu'une amorce d'intérêt se manifeste suite aux travaux des auteurs anglo-saxons qui avaient redécouvert, en censurant toute référence au marxisme, Vygotski dans les années soixante.

L'évaluation objective semble susciter un regain d'intérêt, mais il faut bien garder à l'esprit qu'un retard de trente années ne se rattrapera pas en quelques mois. Les risques de dérapage sont grands de voir se faire et se vendre n'importe quel produit d'évaluation ou de positionnement comme on dit maintenant, outils bricolés à la hâte, et si possible présentés sur ordinateur, l'effort étant concentré sur le "packaging" plus que sur la qualité technique du contenu.

Dans le cadre de ce renouveau de l'évaluation, il est remarquable que la grande majorité des psychologues et des formateurs utilise pour évaluer les capacités d'apprentissage des épreuves statiques, c'est-à-dire des épreuves qui ne font appel à aucun apprentissage.

Il est vrai que longtemps psychologues et enseignants ont recherché chez l'individu sa valeur intellectuelle vraie, stable et invariante. On retrouve facilement derrière la pratique du test statique l'idée que l'intelligence est un don, du Ciel ou de la Nature, de l'Hérédité ou du Milieu... Ce don positionnerait une fois pour toutes, en justifiant cette position définitivement, la vraie place de l'individu dans la stratification sociale.

De ce point de vue, est exemplaire le déplacement de sens qui a accompagné l'expression "bas niveau de qualification -B.N.Q.-", de la simple catégorisation administrative à la catégorisation nosologique "expliquant" l'exclusion sociale. "Chaque intelligence a sa place sociale ; une place sociale pour chaque intelligence", pourrait-on dire.

Par rapport à ce postulat du fondement de l'ordre social- et de son maintien-, comme décalque de l'ordre "naturel" des intelligences, l'idée d'évaluation dynamique, tout comme celle de remédiation cognitive, constitue une hérésie. En effet, et il est temps de l'annoncer, une évaluation dynamique introduit une séquence d'apprentissage dans la passation du test.

Il peut résulter de cet apprentissage un effet de reclassement : tel individu mal placé au test peut rejoindre les rangs médians, voire supérieurs suite à l'apprentissage. En conséquence, la position ne peut plus être considérée comme naturellement immuable puisqu'elle devient sensible à un apprentissage par médiation sociale.

Admettre l'effet de reclassement revient à abandonner le monde fixiste et stable des certitudes dogmatiques pour entrer dans un domaine de relativisation des conduites cognitives par des mises en relations et des contextualisations systémiques et multiformes. Passer à une pratique d'évaluation dynamique demande donc pour le formateur et le psychologue une formation à une autre manière de considérer la question de l'évaluation de l'intelligence.

Le présent ouvrage propose une contribution à ce projet. Il s'adresse aux psychologues, aux formateurs et à l'ensemble des professionnels de la gestion des ressources humaines qui souhaiteraient travailler à des évaluations plus justes sur le plan métrologique de la mesure et plus justes sur le plan de l'équité sociale.

C'est un ouvrage de sensibilisation, d'information, d'explication des principes théoriques et méthodologiques qui sous-tendent l'évaluation dynamique de l'éducabilité. La première partie est un texte de rappel sur les tests et la psychométrie. La seconde partie décrit les différents types de tests statiques utilisés pour évaluer l'intelligence et se termine

sur la question du pronostic de la réussite des apprentissages et des adaptations professionnels. La troisième partie constitue le cœur de l'ouvrage puisqu'on y trouve exposés les principes de l'évaluation dynamique et les principales études menées auprès des enfants et des adultes.

Bien entendu, l'intention générale n'est pas d'opposer les évaluations statiques et dynamiques. En fonction des situations et des problématiques, les deux approches peuvent présenter leurs intérêts et la justification de leur usage : c'est un rapport de complémentarité qui est à définir.

De la sorte, et le psychologue, et le formateur pourront élargir chacun leur panoplie d'outils d'évaluation scientifique à leur disposition. L'évaluation dynamique de l'éducabilité n'est pas une panacée universelle et il faudrait bien se garder des effets de mode. C'est une approche délicate à mener si l'examineur souhaite véritablement recueillir des informations qualitatives sur le mode de fonctionnement du sujet apprenant. La qualité des informations ainsi recueillies dans une perspective diagnostique conditionne directement la possibilité de poser des objectifs et des modalités de remédiation des conduites cognitives.

Les aspects métrologiques ne sont pas incompatibles avec l'évaluation dynamique et peuvent se concrétiser dans l'élaboration d'un pronostic de réussite. D'une manière générale, l'évaluation dynamique du potentiel d'apprentissage propose une mise en situation qui va aider le sujet examiné à parvenir à une meilleure connaissance de lui-même.

L'examineur pourra stimuler les prises de conscience relatives à l'ensemble des aspects des conduites cognitives et conatives mises en oeuvre à l'occasion des différents temps de l'évaluation dynamique : dans la situation de résolution spontanée de la tâche, dans la situation d'apprentissage d'éléments méthodologiques nécessaires à la résolution de cette tâche, dans la situation de transfert de ces éléments méthodologiques.

Ces prises de conscience une fois réalisées peuvent servir d'éléments d'information, à conjuguer aux autres sources d'information, pour améliorer la qualité et la pertinence des

conseils ou des décisions dans les domaines du recrutement, de la mobilité interne ou externe, du bilan des compétences personnelles et professionnelles, de la formation continue des adultes, de la recherche en psychopédagogie...

L'évaluation dynamique de l'éducabilité ne clôt pas le débat sur la définition de l'intelligence : elle le renouvelle plutôt en réintroduisant à côté de l'étude des processus de résolution, celle des processus d'apprentissage et de transfert.

Première partie : les tests

L'évaluation traverse l'ensemble des champs de socialisation et des périodes de la vie : bilan de santé, évaluation scolaire, évaluation sportive, évaluation professionnelle, bilan financier...

Par rapport aux évaluations ordinaires plus ou moins marquées de subjectivité le test se conçoit comme un outil scientifique offrant toutes garanties de la fiabilité de la mesure.

Ce souci de justesse et de stabilité de la mesure différencie le test de l'évaluation pédagogique (Pasquier, 1990a) : n'a-t-on pas calculé qu'il faudrait faire intervenir cent vingt-sept correcteurs pour obtenir la valeur vraie d'une dissertation philosophique et treize pour un devoir de mathématiques, pourtant réputées sciences exactes !

Toutefois, il faut bien se garder de croire que l'usage des tests représente une panacée universelle. Leur utilisation correcte nécessite la conjonction de plusieurs conditions préalables et de différentes précautions à observer en cours de passation et d'interprétation. Ces conditions et précautions sont référées tant au plan technique qu'au plan déontologique.

Le premier chapitre propose une définition des tests et le suivant expose les principes majeurs qui en conduisent l'élaboration. Le troisième chapitre est consacré à la présentation des domaines d'application des tests et des conditions à respecter pour une utilisation efficace et respectueuse de l'individu. Il se termine par l'exposé, dans ses grandes lignes, de la démarche interprétative des tests.

Pour aller plus loin dans le domaine de la psychologie du travail et des organisations :

Bernaud J.L. & Lemoine C. (Eds) (2000), *Traité de psychologie du travail et des organisations*, Paris : Dunod.

Pour aller plus loin dans le domaine de l'évaluation psychologique :

Huteau M. (Ed.) (1994), *Les techniques psychologiques d'évaluation des personnes*, Issy-les-Moulineaux : E.A.P.

Pour aller plus loin dans le domaine de la psychométrie :

Dikes P., Tournois J., Flieller A., Kop J.L. (1994), *La Psychométrie*, Paris : P.U.F.

Pour aller plus loin dans le domaine des tests en pédagogie :

Pasquier D. (1990), *L'évaluation en pédagogie. Du bon usage des tests*, Issy-les-Moulineaux : E.A.P.

Chapitre 1

Définition des tests.

Le mot "test" est dérivé du latin *testum* qui désigne un vase de terre. Du temps des alchimistes, le mot "têt" dénomme un pot de terre servant à l'essai de l'or. En 1890, Mc K. Cattell dans un article de la revue *Mind* dé-crit des épreuves standardisées sous le nom de "mental tests". Dès 1895, le mot "test" sera réintroduit dans la langue française, dans le sens anglais d'examen, par Bi-net et Henri.

Une définition reconnue du mot test fut proposée par l'Association internationale de Psychotechnique : "C'est une épreuve définie, impliquant une tâche à remplir, identique pour tous les sujets examinés, avec une technique précise pour l'appréciation du succès ou de l'échec, ou pour la notation numérique de la réussite. La tâche peut comporter une mise en oeuvre, soit de connaissances acquises (test pédagogique), soit des fonctions sensori-motrices ou mentales (test psychologique)." (Piéron, 1979)

La pratique des tests s'inscrit dans le domaine plus large de la psychotechnique ou de la psychométrie, domaine qui introduit la quantification et la mesure dans l'approche des processus psychologiques.

Pour distinguer le test d'une épreuve quelconque comme un devoir scolaire ou un sujet d'examen ou de concours, un critère joue un rôle clé : la standardisation. La standardisation fonde la définition générale du test : épreuve définie, tâche identique pour tous les sujets, correction strictement critériée et codifiée.

La standardisation fixe l'ensemble des paramètres de l'épreuve : contenu, questions, consignes, procédure et temps de passation, mode de correction et d'attribution de points... d'une manière définitive.

Considérées du point de vue des conditions extérieures, les passations d'un même test sont, par construction, identiques d'une fois à l'autre, d'un lieu à un autre, d'un sujet ou d'un groupe de sujets à l'autre. L'examineur qui met en oeuvre une épreuve standardisée doit respecter une attitude de neutralité bienveillante.

Neutre, il ne peut donner aucune indication non prévue dans le texte à dire pour lancer le test (les consignes). Concrètement, il ne doit ni rien faire ni rien dire qui puisse mettre le sujet sur la voie de la réponse. Cette neutralité peut choquer le formateur qui fait passer l'épreuve à des élèves ou à des stagiaires qu'il connaît bien : ce n'est pas chose facile de laisser des élèves à l'abandon face à des difficultés qu'ils ne parviennent pas à surmonter. Mais il doit prendre conscience que s'il se laisse aller à fournir des indications opératoires en réponse aux sollicitations verbales (questions) ou posturales (attitudes de découragement...) des sujets, le résultat du test reflétera plutôt ses compétences pédagogiques à apporter des aides efficaces que les capacités effectives du sujet à résoudre un problème ou à exécuter une tâche de manière autonome.

La bienveillance, complémentaire de la neutralité, s'exerce afin que le sujet puisse donner le meilleur de lui-même : si tel est bien le cas, alors la mesure indiquée par le test se rapproche de la mesure vraie des capacités potentielles du sujet. Dans le cas contraire, si le sujet reste stressé ou inhibé, la performance enregistrée à l'aide du test risque de sous-estimer ses capacités latentes.

Si l'attitude de neutralité renvoie l'examineur au respect le plus strict de la consigne, l'attitude bienveillante induit chez l'examineur la manifestation des divers encouragements possibles susceptibles de stimuler le meilleur investissement des capacités du sujet dans la tâche que le test propose. L'attitude bienveillante implique de prendre un soin tout particulier à la préparation préalable à la séance de test afin de limiter tous les risques de perturbations liés à des défauts d'organisation matérielle.

La qualité de l'accueil et le rappel des finalités de l'évaluation participent à la mise à l'aise du candidat et à son

implication dans l'épreuve : un sujet à la fois détendu et concentré, tel est le but poursuivi par l'attitude de bienveillance.

Respect des consignes, neutralité et bienveillance s'imposent absolument pour assurer la fiabilité du résultat chiffré de l'évaluation. Cette fiabilité du résultat conditionne l'utilisation de l'étalonnage, c'est-à-dire la possibilité de positionner socialement la performance du sujet par rapport à l'ensemble des performances obtenues par une population de référence, la population parente.

La présence d'un étalonnage distingue le test métrique de l'épreuve pédagogique ordinaire et amène à différencier la note brute et la note standard. Tout comme la note pédagogique, la note brute évalue la qualité des réponses par rapport au contenu de l'épreuve : selon un barème précis de correction, l'examineur attribue des points.

Le total des points obtenus par un sujet, sa note brute, renvoie, par le biais de l'étalonnage, à sa note standard qui indique le classement de la performance par rapport à la distribution des performances attribuées à la population parente.

Les tests se distinguent des autres types d'évaluation selon un second critère propre à la mesure : leurs qualités métrologiques sont contrôlées. Comme tout autre instrument de mesure, un test doit se montrer sensible, fidèle et valide.

"Un test est dit avoir une plus ou moins grande finesse discriminative selon qu'il peut classer les individus en un nombre plus ou moins grand d'échelons." (Piéron, 1979)

Une épreuve qui concentre les résultats vers les notes hautes est une épreuve trop facile, et inversement une épreuve qui concentre les résultats vers les notes basses se montre trop difficile. Un bon test différencie les sujets qui le passent sur l'ensemble de l'échelle de notation.

Il n'est pas de mesure directe de la fidélité d'un test. Trois moyens détournés peuvent être utilisés pour montrer la stabilité de la mesure. Deux concernent la stabilité de la mesure dans la durée. La technique du test-retest consiste à donner l'épreuve une seconde fois à un même groupe de sujets. Bien entendu, le fait de repasser la même épreuve peut faciliter

l'exécution d'une tâche devenue plus familière et se traduire par une augmentation générale des scores (effet de retest).

Afin de limiter un éventuel effet de retest, une seconde technique consiste à utiliser en retest une forme parallèle à l'épreuve originale. On conserve la structure de chaque item mais en l'habillant avec d'autres contenus. Toutefois, lorsque des variations de résultats se produisent, il est délicat de distinguer entre le manque de fidélité ou un manque de similitude entre les deux formes.

Il faut bien remarquer qu'un effet d'apprentissage lié au retest n'est pas un handicap dans l'évaluation de la fidélité de la mesure quand on ne considère que la stabilité des classements des sujets.

La troisième technique, celle du pair-impair, concerne la stabilité de la mesure en fonction du contenu des items. C'est pourquoi on parle également d'homogénéité interne à ce sujet. Elle consiste à dichotomiser l'épreuve en deux sub-tests, l'un regroupant les items pairs et l'autre les items impairs, puis à comparer les scores et les classements obtenus à chacun des deux sub-tests.

"La validité d'un test exprime le degré de liaison entre le rendement du sujet dans le test et son rendement dans une autre activité, que le test est censé prévoir." (Piéron, 1979)

De fait, Piéron définit ici la validité prédictive d'une épreuve, validité qui consiste à mettre en relation de manière diachronique le test pris comme prédicteur et un critère mesuré ultérieurement afin d'élaborer une prévision sous la forme d'une probabilité de réussite.

La validité concourante (ou validité empirique, externe) met en relation de manière synchronique les résultats obtenus au test et les résultats obtenus avec un critère externe (autre test, résultats scolaires, examen de fin de stage, évaluation professionnelle...).

Enfin, la validité de contenu s'apprécie par une analyse qualitative des épreuves, analyse qui doit permettre de conclure sur le degré de cohérence interne de l'épreuve par rapport à l'objet affiché de la mesure.

Plusieurs types de classification des tests ont été proposés. Par exemple, le test individuel se donne en situation duelle alors que le test collectif s'applique à un groupe de sujets. Les épreuves papier-crayon mettent en jeu des supports concrets banaux alors que la passation d'une batterie de sécurité nécessite un matériel électronique sophistiqué. L'utilisation de l'ordinateur envahit également le domaine des tests dont la passation et le traitement se retrouvent plus ou moins automatisés.

Par rapport à la nature de l'objet de l'évaluation, on distingue les tests de connaissances centrés sur les savoirs, les tests d'intelligence centrés sur l'efficacité des processus cognitifs, les tests de personnalité centrés sur les traits psychologiques de la personne.

Pour mieux comprendre le fonctionnement d'un test, il convient de considérer les principes généraux de son élaboration.

Chapitre 2

Élaboration des tests

L'élaboration d'un test procède d'un double échantillonnage : échantillonnage des contenus et échantillonnage de population. La présentation qui suit se situe à un niveau notionnel et expose les grandes étapes de l'élaboration d'un test sans entrer dans le détail technique des procédures statistiques.

Le lecteur peut visualiser ces étapes sur la figure n°1. Imaginons qu'un auteur doive construire un test de connaissances en langue anglaise. Il ne pourra pas concrètement poser une question sur chacun des éléments de la langue dans ses différents aspects lexiques et grammaticaux.

Il devra d'abord délimiter le domaine des contenus à explorer, par exemple l'anglais technico-commercial. Dans ce domaine linguistique défini, l'auteur isolera ensuite les points de maîtrise les plus représentatifs de la maîtrise de l'ensemble.

Ce choix des connaissances à interroger ne relève pas d'un simple arbitraire : l'auteur doit être en mesure d'en démontrer la pertinence en explicitant les critères qui fondent la validité de contenu de l'épreuve. Il peut également s'en remettre au tirage au sort pour retenir les points de maîtrise qui seront objets de l'évaluation.

Une fois les contenus clés déterminés et validés comme représentatifs de l'objet d'évaluation, l'auteur rédige les questions (les items) du test. Il choisit le type de tâche qui introduira chaque contenu, la modalité de présentation de l'énoncé, la forme de la réponse.

Dans le présent exemple, l'une des tâches pourrait consister à dénommer des objets techniques. L'énoncé se présenterait alors sous la forme d'objets dessinés (modalité picturale), ou bien sous la forme d'un texte de définition

(modalité verbale)... L'élève, l'étudiant, la secrétaire..., les sujets pouvant subir l'épreuve seraient invités soit à écrire le terme technique désignant l'objet, soit à cocher le terme exact dans une liste de termes proposés, soit à relier une série de dessins et une liste de termes...

Dès ce moment de la description des étapes de l'élaboration du test, apparaît la nécessité de relativisation. En effet, on peut raisonnablement imaginer que le choix d'autres contenus, d'autres types de tâches, d'autres modalités d'énoncés, d'autres formes de réponses... amènerait un même sujet à produire une performance sensiblement différente. Par conséquent de cette nécessaire relativisation, des résultats de tests ne se présentent pas *in abstracto* ; ils s'accompagnent toujours de la référence du test qui les a produits

Une fois les supports du test mis au point, l'auteur va procéder à la définition puis à l'échantillonnage de la population de référence (la population parente). Il utilisera dans cet objectif des techniques semblables à celles qui sont mises en oeuvre pour mener les sondages d'opinions. Dans l'exemple évoqué, vouloir faire passer le test à tous les technico-commerciaux relève d'une ambition démesurée.

De la même manière qu'un institut de sondage n'interroge jamais les millions de français en âge de répondre, l'auteur du test va constituer un échantillon représentatif de la population parente limité à quelques centaines d'individus. Il pourra procéder par tirage au sort (échantillon au hasard), ou bien par quota (échantillon proportionnel) s'il veut s'assurer de la proportion d'hommes et de femmes par exemple. La procédure par tirage au sort implique la connaissance nominative de tous les sujets de la population parente, ce qui est très souvent impossible à réaliser.

Une fois l'échantillonnage de la population parente terminé, il reste à faire passer le test aux individus qui composent l'échantillon. La correction des épreuves est ensuite menée à l'aide de critères et d'un barème suffisamment précis et finement codifiés. La notation répond aux mêmes règles de standardisation que celles qui codifient la passation, de telle manière que deux correcteurs indépendants du test d'un même

sujet doivent arriver à l'attribution d'une même note. Cet aspect rigide de la correction est susceptible de gêner des enseignants ou des formateurs qui souhaiteraient se former à l'utilisation de tests dans la mesure où ils perdent toute initiative dans l'appréciation de la qualité de la réponse. Le total des points ainsi attribués selon la justesse des réponses produites par un sujet constitue la note brute obtenue par ce sujet.

L'auteur, une fois recueilli l'ensemble des notes brutes des sujets composant l'échantillon, procèdera à leur étalonnage. Il va présenter les notes brutes sous la forme d'une distribution segmentée en classes d'effectifs selon deux grandes options.

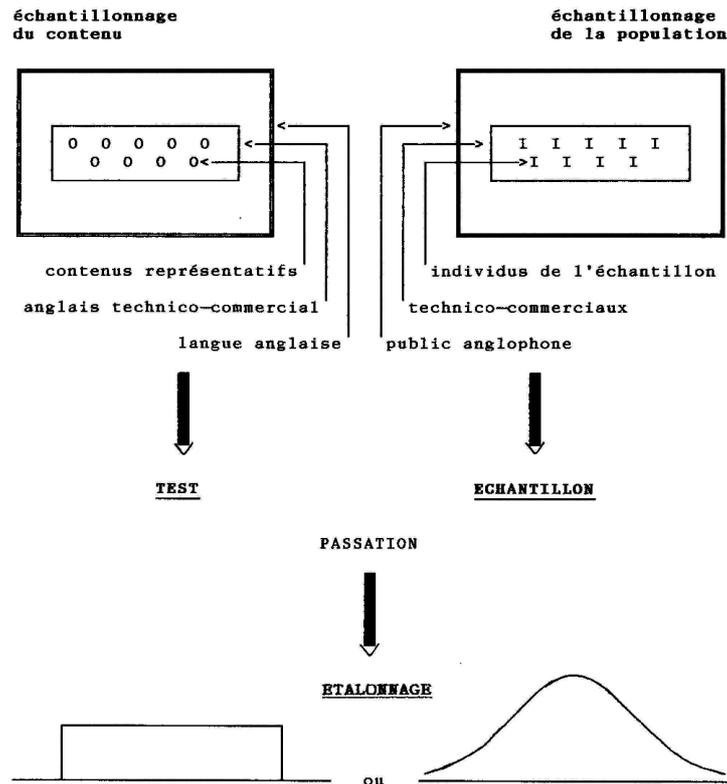
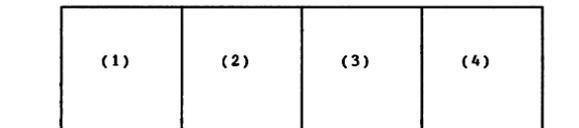


FIGURE N°1: étapes de l'élaboration d'un test

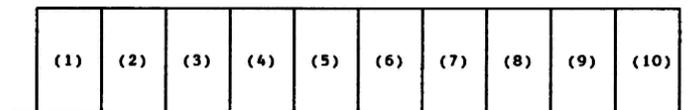
La première option consiste à choisir une distribution rectangulaire dans laquelle chaque classe statistique regroupe une même proportion de l'effectif de l'échantillon. Par exemple, dans un quartilage (voir figure n°2) on segmente la distribution des notes en quatre classes regroupant chacune 25% de l'effectif. Dans un décilage (voir figure n°3), la suite ordonnée des notes brutes est segmentée en dix classes regroupant chacune 10% de l'effectif.

La seconde option est statistiquement plus complexe et fait appel à des distributions normalisées selon le modèle de la courbe de Gauss : les limites des classes sont déterminées par une fraction d'écart type (voir figure n°4).



classe (1): 25% des notes les plus basses
classe (2): 25% des notes moyennes basses
classe (3): 25% des notes moyennes élevées
classe (4): 25% des notes les plus élevées

FIGURE N°2: quartilage



classe (1): 10% des notes les plus basses,...
classe (10): 10% des notes les plus élevées

FIGURE N°3: décilage

Dans ce cas, les effectifs par classe sont inégaux : plus importants dans les classes centrales, ils diminuent dans les classes des deux extrémités de la dispersion. Attention au piège tendu par les mots. Dire d'une distribution de notes qu'elle est normale ne signifie pas qu'elle corresponde à une loi naturelle relative à la distribution des aptitudes et qu'en conséquence toute autre forme de distribution serait anormale. Qualifier une distribution de normale appartient au champ sémantique de la statistique. Cela veut dire que la distribution des notes est

unimodale, symétrique (courbe en forme de cloche) et qu'elle est conforme à l'équation de Laplace-Gauss. Certains auteurs présentent leurs résultats sous les deux formes de distribution.

Concrètement, pour saisir ce qui fait l'originalité d'un test, il faut se représenter le passage de la note brute à la note standard. La note brute, c'est la somme des points attribués aux bonnes réponses. Comme une note pédagogique, la note brute porte sur la justesse des contenus des réponses, à la différence près que critères et barème sont rigoureusement précisés et appliqués par tous les examinateurs de la même manière.

La lecture de l'étalonnage permet de situer la note brute dans une classe statistique de la distribution standardisée. Les auteurs utilisent comme note standard le numéro d'ordre de la classe où se situe la note brute du sujet. La note standard correspond donc à une classe statistique, à un rang dans un classement, reflet d'un positionnement social de la performance. Dire d'un sujet que sa note standard est 7 sur un décilage revient à dire que sa performance est supérieure à 60% des performances des sujets de la population parente.

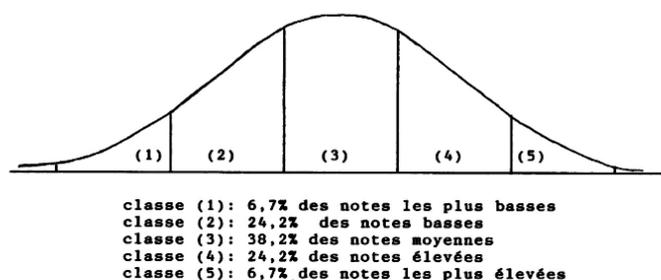
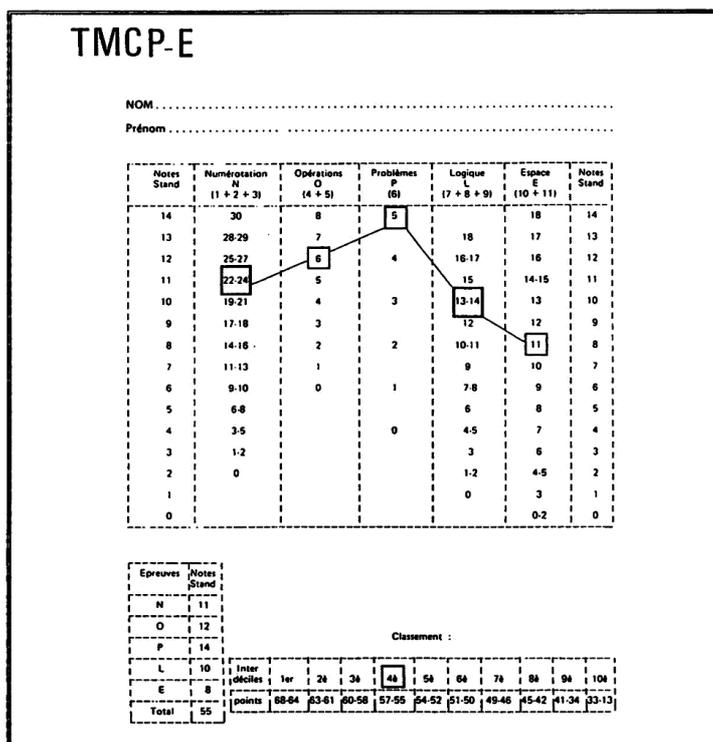


FIGURE N°4: distribution normalisée en 5 classes

Le test, et c'est là sa supériorité sur des épreuves pédagogiques ordinaires, fait passer d'une évaluation centrée sur les contenus des réponses (note brute) à un positionnement social de la performance par rapport à la population parente du sujet (note statistique standard). Cette note standard constitue une référence externe qui garantit en partie l'objectivité de

l'évaluation, dans la mesure où les consignes de passation, de correction et de notation ont été bien respectées.

Grâce aux notes standards, il devient possible d'établir un profil des performances obtenues à différentes épreuves étalonnées sur le même échantillon de sujets. Dans un profil, on compare des positions dans les classements indépendamment des caractéristiques des notes brutes qui rendent ces comparaisons impossibles quand elles n'ont pas des moyennes et des dispersions identiques.



DOCUMENT N°1: Profil obtenu au Test de mathématiques pour le cours préparatoire (Pasquier, Thébaud, 1982)

Avec l'aimable autorisation des E.A.P.

L'illustration proposée (voir document n°1) présente le profil pédagogique obtenu par un élève à une épreuve de mathématiques pour le cours préparatoire. On voit le passage des notes brutes aux notes standards : pour l'élève considéré, une note brute de 22 à 24 points en numération donne une note standard de 11, 6 points en opérations donnent 12, 5 points en problèmes donnent 14... On observe que la performance de cet élève se positionne très bien pour la résolution des problèmes, mais plus médiocrement pour la résolution d'exercices relatifs au domaine spatial.

Les notes standards par type d'exercices peuvent ensuite être additionnées et ce total donne lieu à un étalonnage de la performance sur l'ensemble de l'épreuve de mathématiques. Dans le cas présent, la note standard 4 indique que l'élève a mieux travaillé que 30% des élèves de cours préparatoire.

Ces principes généraux guident l'élaboration d'un test quels que soient le type d'épreuve et son domaine d'application. Le positionnement social de la performance sera d'autant plus fiable que le respect de ces principes généraux est assuré. De plus, la fiabilité métrologique du résultat conditionne la possibilité d'une interprétation.

Chapitre 3

Domaines, conditions d'utilisation et interprétation des tests

Les tests sont utilisés dans différents domaines de la pratique des psychologues. Dans l'approche clinique, en psychologie individuelle, les différents types de tests servent à explorer les diverses composantes de la globalité psychologique du sujet.

Dans le système éducatif, les psychologues scolaires et les conseillers d'orientation psychologues peuvent apprécier le rendement scolaire de l'élève en construisant le rapport entre le niveau scolaire objectif à l'aide d'un test de connaissances et le niveau d'effcience intellectuelle objectif à l'aide d'un test d'intelligence.

Les questionnaires d'intérêts personnels, scolaires et professionnels, les épreuves dites de personnalité... facilitent l'élaboration d'un conseil d'orientation. Un examen clinique approfondi et étayé à l'aide de tests choisis en fonction de la nature de chaque cas sera nécessaire pour isoler les sources de difficultés d'apprentissage et envisager la meilleure approche rééducative.

Dans le monde du travail, l'usage des tests s'intègre dans les procédures de recrutement, dans la gestion des ressources humaines, dans les dispositifs de mobilité interne, dans le nouveau cadre du bilan de compétences personnelles et professionnelles...

Enfin, dans le domaine de la formation professionnelle, l'utilisation des tests s'inscrit dans le cadre de la psychopédagogie des jeunes et des adultes. Ce secteur se complexifie au fil des ans par l'installation, en amont des formations qualifiantes, de dispositifs dits de préformation, d'insertion, d'orientation, de mise à niveau...

Ces dispositifs dont les dénominations changent au fil des campagnes impulsées par le Ministère du Travail, de l'Emploi et de la Formation Professionnelle ont en commun de s'adresser principalement aux jeunes sortis du système scolaire sans qualification, ou à des adultes étiquetés comme "bas-niveaux de qualification". L'objectif affiché au terme de ces dispositifs consiste à préparer les groupes de stagiaires à suivre, avec les meilleures chances de réussite, une formation professionnelle qualifiante.

Par rapport à cet objectif de qualification, il devient important de pouvoir poser un pronostic qui prévoit la probabilité de réussir telle ou telle formation. L'utilisation des tests participe à l'établissement de ce pronostic. D'autre part, de nouveaux usages des tests se précisent en termes de diagnostic psychopédagogique.

Il s'agit, pour le formateur et le psychologue, de mettre en évidence les conduites cognitives du sujet, ses modes de résolution d'un problème ou d'exécution d'une tâche, ses manières d'apprendre... Les prolongements de ce diagnostic psychopédagogique centré sur les processus cognitifs se déclinent dans deux directions complémentaires.

Soit le formateur va entreprendre d'adapter son mode de transmission des contenus de formation aux caractéristiques du fonctionnement cognitif et conatif de chacun de ses stagiaires. Soit le formateur va entreprendre de modifier la conduite cognitive du stagiaire, lorsqu'il estime nécessaire de la rendre perfectible, par le biais d'une approche définie en termes d'éducabilité.

Imaginons qu'à la suite d'un diagnostic mené avec différents tests il apparaisse qu'un stagiaire éprouve des difficultés relatives à traiter des données codées dans une modalité verbale. Soit le formateur va s'adapter à cette caractéristique du stagiaire en évitant de lui transmettre des données verbales : il les lui traduira systématiquement en données figuratives par exemple. Soit le formateur va projeter de modifier le mode de fonctionnement du stagiaire dans le sens d'un élargissement de ses outils cognitifs en lui apprenant à traiter des données verbales.

Ces pratiques de remédiation qui définissent le champ de l'éducabilité amènent les formateurs à s'intéresser à l'utilisation des tests, voire à en introduire l'usage dans leur pratique. Jusque-là réservée aux seuls psychologues, on observe une extension de l'utilisation de certains tests aux formateurs, extension qui ne va pas sans poser quelques risques.

En premier lieu, un nouveau marché peut amener certains auteurs ou éditeurs à proposer hâtivement à la vente des épreuves qui ne présentent pas toutes les garanties métrologiques nécessaires. En second lieu, des organismes de formation peuvent être tentés de créer leurs propres outils d'évaluation dans le cadre d'une approche naïve qui ne prendrait pas en compte les constats les plus élémentaires de la docimologie. En troisième lieu, on peut craindre de la part de certains formateurs non formés à leur usage une pratique sauvage des tests aux conséquences fâcheuses pour les stagiaires.

Cette interrogation amène à définir explicitement les conditions à respecter pour une utilisation heuristique des tests dans le domaine de la formation des adultes. La première de ces conditions concerne les critères de choix des épreuves. Dans le domaine de la psychopédagogie, il convient d'utiliser des tests à visée psychopédagogique : tests de connaissances scolaires, techniques, professionnelles, tests d'évaluation des capacités ou du potentiel d'apprentissage... D'autre part, le test choisi doit correspondre au groupe de stagiaires qui va le subir. Il faut donc vérifier que le test dont on envisage l'usage présente un étalonnage effectué sur un échantillon de la population parente du groupe de stagiaires. Par exemple, utiliser avec des adultes une épreuve de calcul étalonnée sur un échantillon de collégiens n'a ni sens ni légitimité.

La seconde condition concerne la raison qui déclenche la mise en oeuvre d'une évaluation objective à l'aide d'un test. Dans le sillage des événements de mai 1968, les psychologues et les formateurs ont eu tendance à rejeter toute forme d'évaluation en général et les tests en particulier. On constate un retour de balancier qui devrait permettre de rattraper les vingt années de retard prises sur les auteurs anglo-saxons. Toutefois,

il ne faudrait pas basculer dans l'extrême inverse et abuser de l'usage des tests simplement "pour voir". Une passation de tests, situation d'évaluation stressante, se justifie par un enjeu social extérieur à l'évaluation : décision de recrutement, conseil d'orientation, aménagement d'un parcours de formation, inflexion d'une trajectoire socioprofessionnelle, adaptations pédagogiques, évaluation d'une action de formation... Cet enjeu externe au test doit être partagé par le sujet qui subit l'épreuve. En effet, cet enjeu donne du sens à l'évaluation et stimule l'investissement dans la tâche.

Une troisième condition touche à l'obligation pour l'utilisateur de respecter les règles d'usage définies par l'auteur du test. Ces règles figurent dans le manuel technique qui accompagne l'épreuve et en précise les éléments de standardisation. Les consignes de passation, de correction, de notation, d'étalonnage et d'interprétation sont à respecter d'une manière stricte. Il ne s'agit pas, pour le formateur ou le psychologue, de simplement se montrer "bête et discipliné". L'effort de rigueur dans l'application des consignes augmente le degré de précision et de fiabilité de la mesure, et tout gain obtenu dans la fiabilité quantitative procure des degrés de liberté supplémentaires dans l'interprétation qualitative. En effet, quelles conjectures élaborer, quelles conclusions tirer quand la base chiffrée se montre hasardeuse et incertaine ?

Une quatrième condition implique les comportements à tenir par le formateur ou le psychologue afin de préserver le test d'une usure trop rapide. Cette préservation passe par la prise de précautions en termes de confidentialité des questions et des réponses : les bonnes réponses ne sont jamais communiquées aux stagiaires, le matériel de test, les brouillons... sont systématiquement récupérés à l'issue de chaque passation. Il ne s'agit pas là d'une réaction "paranoïaque", mais quelle valeur garderait un test si tout un chacun pouvait en apprendre les réponses à l'avance ? Dans le même ordre d'idée, la bonne tenue dans l'utilisation des tests passe par le respect des réglementations relatives au copyright.

Une cinquième condition amène le formateur ou le psychologue à respecter les règles déontologiques telles qu'elles

sont formulées, par exemple par les textes qui régissent la mise en oeuvre du bilan de compétences personnelles et professionnelles. D'une manière générale, rien ne doit être fait qui puisse porter un préjudice quelconque au sujet qui passe l'épreuve. Les résultats de l'évaluation appartiennent de manière exclusive à l'évalué et ils doivent lui être restitués de manière explicite. L'examineur est soumis au secret professionnel, et la transmission des résultats à un tiers ne peut se faire sans l'accord formel de l'intéressé.

En dernier lieu, et cette condition englobe les précédentes, il ne saurait y avoir utilisation de tests sans formation préalable de l'examineur. Cette formation ne peut se limiter au seul apprentissage des modalités de passation. Elle doit aborder les éléments les plus essentiels qui concernent l'interprétation des résultats du test. Dès 1911, Binet, concepteur du premier test d'intelligence, prenait soin de préciser les choses : "Il est clair que cette méthode de mesure ne peut être mise entre les mains du premier venu ; elle exige du tact, du doigté, une expérience des causes d'erreur à éviter, surtout une notion claire des effets de la suggestion ; de plus, elle n'a rien d'automatique ; on ne peut pas la comparer à une bascule de gare sur laquelle il suffit de monter pour que la machine délivre notre poids imprimé sur un ticket... Les résultats de notre examen n'ont pas de valeur s'ils sont séparés de tout commentaire ; ils ont besoin d'être interprétés." (Binet, 1911)

Quelques temps plus tard, Wallon enfonçait le clou : "Dans l'emploi des tests s'observent tous les degrés de la sagacité et tous les degrés de la niaiserie. Il n'y a pas de procédés qui permettent à la science de se développer automatiquement. L'intelligence est nécessaire. Lorsque l'engouement suscité par le succès ou par l'apparente facilité d'une méthode fait oublier cette vérité première, la recherche utile est vite ensevelie sous un fatras de stupidités." (Wallon, 1982)

Deux citations de cette clarté et de cette force seront tout juste suffisantes pour endiguer les nombreux risques de dérives qui guettent le psychologue ou le formateur au moment de l'interprétation des résultats chiffrés. L'examineur peut

facilement glisser du positionnement social de la performance sur l'étalonnage statistique au jugement de valeur globalisant et définitif porté sur la personne.

C'est la performance qui est classée et non le sujet. Confondre la performance effective et la valeur de son producteur conduit tout droit à des interprétations circulaires et tautologiques qui n'expliquent rien : la performance est bonne parce que le sujet est doué ; le sujet est doué parce que la performance est bonne ! Ce type d'assimilation réductrice entre le producteur et son oeuvre conduit également à certaines formes de fatalismes pédagogiques ou, selon l'expression de Feuerstein, d'approche passive acceptante du déficit ou du handicap supposé (Feuerstein, Rand, Hoffman et Miller, 1980). Selon cette approche, quand la performance au test est faible, alors les aptitudes du sujet sont faibles et il faut adapter la formation à ce bas niveau, réduire le degré d'exigence, l'intensité et le nombre des contraintes propres à tout apprentissage...

Maintenir la distinction performance-sujet constitue un point primordial pour interpréter le résultat du test en interrogeant les relations fonctionnelles qui se sont établies entre les caractéristiques formelles de la tâche proposée par les items du test et les connaissances et procédures psychologiques mises en oeuvre par le sujet. Réfléchir en termes d'interaction tâche-fonctionnement du sujet ouvre, dans le cadre de l'évaluation formative, des perspectives dans deux directions. Dans le sens de la compréhension du mode de fonctionnement de la personne, cette réflexion va alimenter l'intelligibilité des comportements et des conduites observées. Une fois la logique des conduites convenablement abstraite et définie, il devient possible de définir le contenu d'un conseil ou de fixer des objectifs de remédiation.

Par exemple, la pratique des tests met en évidence qu'un sujet élabore mentalement la réponse de manière correcte mais qu'il se trompe fréquemment dans l'effection, c'est-à-dire dans l'acte d'extériorisation, de codage et de communication de cette réponse. A partir de ce constat, il devient possible de travailler à la définition puis à l'opérationnalisation d'un objectif de

remédiation spécifique à l'intériorisation des boucles correctives. En effet, si le psychopédagogue parvient à rendre le sujet capable de se représenter la réponse concrète, de l'auto-évaluer et de la corriger mentalement avant de la produire, il met en place une régulation anticipatrice qui limitera le risque d'erreur.

Un autre risque se traduit dans l'abstraction réductrice et globalisante qui consiste à ne plus voir la personne qu'au travers de l'attribution d'une seule catégorie nosologique. Par exemple, réduire l'ensemble du fonctionnement psychologique de la personne à un seul nombre comme le Q.I. est tout à la fois dangereux et illégitime. Dangereux car le nombre pris comme absolu enferme le sujet dans une catégorie qui peut sceller son destin social, sans autre forme de questionnement, de la classe de perfectionnement au C.A.T. en passant par la S.E.S. Illégitime car un résultat de test est relatif aux composantes de la situation de mise à l'épreuve. Tout résultat chiffré doit donc s'accompagner de la définition de ces composantes : contenus de l'épreuve, type de raisonnement induit par la tâche, modalités de codage des informations, dispositions psychologiques du sujet le jour de l'évaluation...

Un risque dans l'interprétation réside dans la non prise en considération du niveau de la performance obtenue par le sujet. Un résultat élevé établi dans de bonnes conditions de passation ne peut être imputé qu'aux capacités réelles du sujet. Par contre, un mauvais résultat ne traduit pas forcément une incapacité. Il existe de multiples raisons qui peuvent empêcher une personne d'investir ses capacités dans une épreuve donnée, à un moment donné, dans des circonstances données. L'absence d'enjeu ou un enjeu trop stressant, le manque de familiarité avec les contenus qui empêche la prise de signification de la tâche, de mauvaises conditions de passation, tous types de perturbations occasionnelles... peuvent inhiber la mobilisation énergétique et/ou le déploiement des capacités nécessaires à la résolution des items de l'épreuve.

Le mauvais résultat ne permet de conclure de manière univoque, mais il ouvre la voie à un questionnement sur la nature et les sources possibles des difficultés constatées. Ce

questionnement peut déboucher sur d'autres observations et évaluations pour vérifier telle ou telle conjecture, dans le cadre d'un diagnostic progressif qui fait passer de la vérification d'une hypothèse h à la vérification d'une autre hypothèse h+1 déterminée en fonction du résultat de la vérification de h. L'interprétation procède d'une double mise en relation. Dans une perspective de cohérence interne, les données chiffrées sont comparées entre elles et rapprochées des données qualitatives issues des diverses observations recueillies en cours de passation : traces graphiques, postures, attitudes, questions, verbalisations...

La structuration finale du tableau clinique du sujet se construit à partir d'un cadre théorique choisi *a priori* par l'auteur du test et/ou par l'examineur. Comme le souligne Wildöcher (1983) à propos du modèle psychanalytique de la personnalité, et on peut élargir l'idée à tout modèle interprétatif : "[Le modèle]... n'est pas le résultat d'une approche empirique mais un cadre conceptuel destiné à décrire la manière dont s'articulent les actes mentaux." Il ne faut pas prendre le modèle pour la réalité et bien évidemment, en conséquence des divergences dans le choix des critères définissant le modèle, un même ensemble de résultats sera présenté différemment selon les examineurs. L'interprétation ne peut en rester là. La seule recherche de la cohérence interne revient à élaborer une réalité psychologique du sujet dans un univers conceptuellement monoréférencé, fini et clos.

La seconde articulation relève de la recherche de cohérence externe. L'examineur tente de corréler les conclusions de l'analyse des résultats aux données externes à la situation d'évaluation : éléments d'anamnèse, caractéristiques de la situation actuelle du sujet, nature et dynamique du projet, problématiques affective et sociale, évaluations professionnelles...

L'articulation des deux types de cohérence ouvre véritablement l'accès à une compréhension à la fois analytique et synthétique du sujet pris dans sa plus grande globalité et dans les différents aspects de son mode de fonctionnement interne et relationnel. Le principal danger de l'interprétation réside dans le

dogmatisme inhérent au mode de penser unilatéral d'un examinateur qui plaquerait un modèle théorique figé sur une réalité qui sera, si nécessaire, déformée pour entrer dans le cadre. Cette forme de mécanisme de défense empêche la coordination des points de vue et la réduction des contradictions sur un plan de globalité supérieur. La maîtrise de l'interprétation se construit progressivement, par l'expérience, par la discussion des cas entre collègues, par la compilation des signes cliniques pertinents... Cette question de l'interprétation s'applique à tous les types d'évaluation, y compris celle mise en oeuvre à l'aide des tests d'intelligence.

Deuxième partie

Les tests d'intelligence

Les épreuves conçues et utilisées pour évaluer les capacités intellectuelles se réfèrent à trois courants théoriques principaux.

Le premier chapitre présente les épreuves classiques. Au début du siècle, Binet initialise la conception et l'application des épreuves dites composites, mosaïques, hochepots, dans le cadre théorique de l'intelligence globale. Un peu plus tard, les travaux de Spearman amorcent la théorie de l'intelligence générale et inspirent la création des épreuves factorielles.

Dans le chapitre suivant seront évoquées les épreuves opératoires dérivées de la théorie constructiviste de l'orthogénèse élaborée par Piaget.

Chacun de ces types d'épreuves est utilisé pour étayer la prédiction de la réussite des apprentissages. Ces pratiques feront l'objet du dernier chapitre de cette partie. Elles seront examinées sous l'angle de leurs avantages en termes d'objectivation d'un niveau de capacités, et sous l'angle de leurs limites.

Pour aller plus loin dans le domaine des tests d'intelligence :

Huteau M., Lautrey J. (1999), *Evaluer l'intelligence. Psychométrie cognitive*, Paris : P.U.F.

Huteau M. (Ed.) (2000), *Les figures de l'intelligence*, Paris : E.A.P.

Chapitre 4

Les tests d'intelligence globale et d'intelligence générale

Binet et Simon furent les premiers auteurs à résoudre le problème de l'évaluation psychométrique de l'intelligence. Leur épreuve, le *Binet-Simon*, s'adresse aux enfants et le résultat s'exprime en termes de mois d'avance ou de retard de développement en comparant la performance de l'enfant à la performance moyenne des pairs de son âge. Un regard sur le contenu des items de la *Nouvelle échelle métrique de l'intelligence*, la N.E.M.I., version améliorée du *Binet-Simon*, illustre parfaitement la conception de l'intelligence globale de Binet : "Peu importe les tests pourvu qu'ils soient nombreux." L'idée générique consiste à juxtaposer des tâches variées, de natures différentes, de contenus différents, présentées dans différentes modalités de codage... Dans l'extrait de la liste des items proposée (document n°2), on voit que l'item 37 "Comptage de cubes" renvoie à l'analyse perceptive d'une configuration spatiale, que l'item 38 "Répéter 4 chiffres à rebours" fait appel à la mémoire à court terme, que l'item 41 "Vocabulaire" mobilise les connaissances lexicales, que l'item 50 "Phrases absurdes" fait plutôt appel à une logique du bon sens...

L'intelligence se définit concrètement dans ce cadre comme la capacité d'un sujet de répondre correctement à un ensemble de problèmes et de tâches variés, c'est-à-dire comme une capacité globale intégrant un ensemble de fonctions pouvant s'additionner, se compenser... Binet, au-delà de sa célèbre boutade : "L'intelligence, c'est ce que mesure mon test" ... a défini à plusieurs reprises sa conception globaliste de l'intelligence, à l'occasion des mises au point successives de son

test (1905, 1908, 1911) : "...son organe essentiel est le jugement, autrement dit le bon sens, le sens pratique, l'initiative, la faculté de s'adapter.", ou "Compréhension, invention, direction et censure, l'intelligence tient de ces quatre mots.", ou encore "C'est surtout par la totalité (de son intelligence) qu'un individu donne sa valeur. Nous sommes un faisceau de tendances, et c'est la résultante de toutes ces tendances qui s'exprime dans nos actes et fait que notre existence est ce qu'elle est. C'est donc cette totalité qu'il faut savoir apprécier."

- 37 - Comptage de cubes, 10 pts
- 38 - Répéter 4 chiffres (rebours)
- 39 - Phrases en désordre
- 40 - Ressemblances, 4 points
- 41 - Vocabulaire, 25 mots
- 42 - Dessins de mémoire
- 43 - Séries de nombres, 2 pts
- 44 - Ingéniosité, 1 sur 3
- 45 - 3 mots en une phrase
- 46 - Séries de mots, 3 pts
- 48 - Interpréter une gravure
- 50 - Phrases absurdes, 5 pts

DOCUMENT N°2: Extrait de la liste des items de la N.E.M.I.
(Zazzo et col., 1966)

Avec l'aimable autorisation des E.A.P.

Le test de Binet fut le point de départ d'une génération d'épreuves mosaïques, hochepots, composites mélangeant des

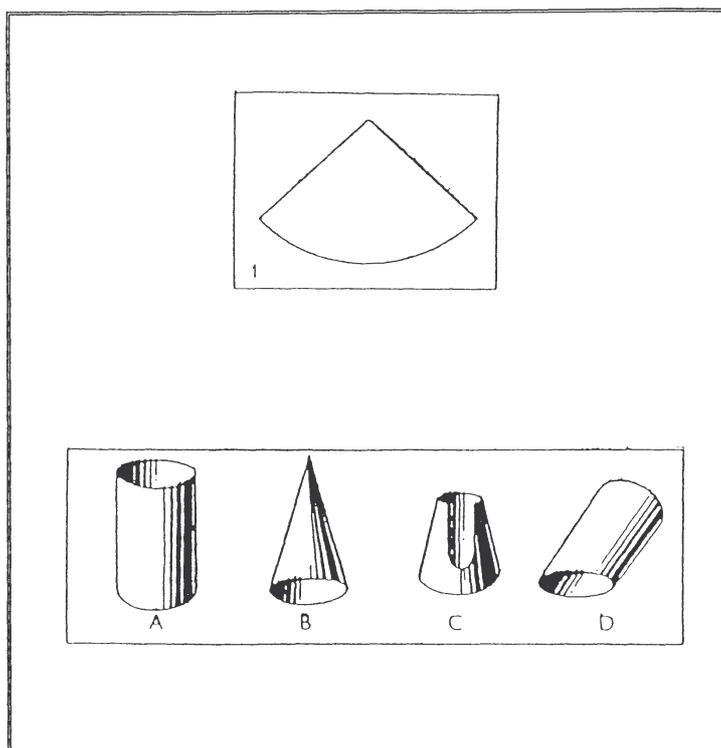
tâches hétérogènes. Le *Binet-Simon* fut amélioré en France par Zazzo et ses collaborateurs sous la forme de la N.E.M.I. (Zazzo et col., 1966). Aux Etats-Unis, le *Binet-Simon* fut révisé par Terman et Merrill (Terman, Merrill, 1959) et il inspira largement Wechsler dans l'élaboration de ses trois échelles pour les jeunes enfants, les enfants et les adultes. De passation individuelle, ces échelles sont utilisées en psychologie scolaire et en psychologie clinique. D'autres épreuves s'inspirant d'une approche globale servent dans les domaines de la psychologie du travail et de la formation professionnelle continue.

<p>Un homme possède une voiture d'une valeur de Fr. 9000.—. Elle est assurée pour les deux tiers de sa valeur. La taxe d'assurance est de Fr. 2.25 pour Fr. 100.— de valeur assurée. Quel est le montant de la taxe ?</p>	<p>Deux piliers circulaires ont respectivement 1.50 mètre et 0.625 mètre de diamètre. Quelle est la différence — en centimètres — de leur diamètre ?</p>
<p>Une courroie fait faire 9 tours à une poulie et 24 tours à une autre poulie. Combien de tours fait la seconde poulie pendant que la première fait 27 tours ?</p>	<p>Trois ouvriers ont reçu une augmentation de salaire, respectivement de Fr. 0.75, Fr. 1.— et Fr. 1.25 par heure. A eux trois, ils ont touché Fr. 630.— d'augmentation. Ils avaient travaillé 7 heures par jour et 6 jours par semaine. Pour combien de semaines ont-ils touché cette augmentation ?</p>

DOCUMENT N°3: Item numériques de la B.G.A. (Boss et al., 1960)

Avec l'aimable autorisation des E.A.P.

Les psychologues factorialistes, depuis Spearman, ont choisi une perspective inverse de celle de Binet, celle de l'intelligence générale. Le facteur, c'est d'abord un indice statistique calculé par une méthode d'analyse mathématique : "Élément commun à certaines catégories d'opérations mentales, pouvant rendre compte, pour part ou totalement, de la parenté d'épreuves fondées sur ces opérations (d'après la corrélation des résultats de ces épreuves dans un groupe d'individus). L'existence des facteurs est dégagée d'une analyse statistique ; leur nature est supposée d'après la saturation en tel ou tel d'entre eux des différentes épreuves." (Piéron, 1979)



DOCUMENT N°4: Item spatial de la B.G.A. (Boss et al., 1960)

Avec l'aimable autorisation des E.A.P.

Sans entrer dans les détails techniques de l'analyse factorielle, (voir par exemple Cibois, 1983), on peut, *grosso modo*, se construire une représentation notionnelle de ce type d'analyse. Lorsqu'il fait passer un ensemble de tests différents à un groupe de sujets, le chercheur trouve souvent que les sujets les mieux classés sur une épreuve ont tendance à bien se placer dans toutes les autres, les sujets moyens restant plutôt moyens partout, les plus faibles ayant tendance à le rester dans tous les classements.

NOTES GLOBALES											
APTITUDES	NOTES BRUTES	NOTES ÉTALONNÉES									
		--		-		0			+		++
		1	2	3	4	5	6	7	8	9	
		4%	6,6%	12,1%	17,5%	19,6%	17,5%	12,1%	6,6%	4%	
Raisonnement déductif (R1)											
Raisonnement inductif (R2)											
Raisonnement analogique (R3)											
Raisonnement pratique-technique (R4)											
Spatial											
Problèmes											
Opérations											
Attention											
Orthographe											
Compréhension verbale											
Efficacité scolaire (ES)											
Efficacité intellectuelle générale (EIG)											

Document n°5 : feuille de profil de la NV7.
(Bernaud, Priou, Simonet, 1991)
Avec l'aimable autorisation des E.A.P.

Cette relative communauté intra-individuelle des classements aux différentes épreuves a amené un auteur comme Spearman (1927) à postuler l'existence d'un facteur psychologique général de l'intelligence, le facteur G, les éventuelles variations étant imputées à des facteurs spécifiques

aux différentes épreuves. Toutefois, le passage du plan de l'analyse statistique au plan de la consistance psychologique n'est ni simple ni évident. L'identification des variables psychologiques latentes dont l'expression se manifesterait dans le facteur G reste à l'ordre du jour.

Sur un plan plus formel, le modèle d'analyse hiérarchique de Burt-Vernon, actuellement le plus largement reconnu, permet, suite à l'abstraction de l'influence du facteur G, d'isoler des facteurs régionaux, des facteurs de groupes (Vernon, 1950). "Avec le modèle factoriel hiérarchique, les épreuves psychométriques se répartissent... entre deux grands facteurs de groupe, dont l'un est généralement interprété comme Verbal-Numérique, et l'autre comme Spatial-Mécanique." (Lautrey, De Ribaupierre, Rieben, 1990)

Le facteur Verbal-Numérique renverrait aux opérations sur le discontinu nécessitant un traitement digital, analytique d'informations discrètes propres aux domaines logicomathématique et logico-propositionnel. Concrètement, ce facteur joue dans des épreuves de logique verbale ou propositionnelle, de suites numériques, de calculs arithmétiques...(voir document n°3).

Le facteur Spatial-Mécanique, quant à lui, renverrait aux opérations sur le continu nécessitant un traitement analogique, global d'informations se codant de manière imagée, propres aux domaines spatio-temporels. Ce facteur de groupe jouerait principalement dans les épreuves impliquant l'espace, le temps, les relations causales, les relations mécaniques... (voir document n°4)

A contrario des épreuves hochepots, les épreuves factorielles sont toujours homogènes du point de vue de la tâche. Les épreuves les mieux représentatives du facteur G font essentiellement appel à des opérations d'induction et de déduction, sur un matériel figuratif, *a priori* moins culturellement marqué. Par exemple, le test *Progressive Matrices* de Raven demande d'abstraire les relations qui unissent les éléments donnés pour inférer l'élément manquant.

En fonction de l'objectif de l'évaluation, le psychologue pourra choisir d'appliquer soit une épreuve de facteur général,

soit une épreuve faisant appel à un facteur spécifique. Des ensembles plus sophistiqués de batteries factorielles permettent à la fois de positionner sur les étalonnages la performance générale du sujet et de dresser son profil en mettant en relation les niveaux de performances dans chacune des épreuves spécifiques. En psychologie du travail, on peut utiliser par exemple les batteries de Bonnardel (1972) ou de Boss, Cardinet, Maire et Muller (1960)...La batterie NV7 de Priou, Bernaud et Simonnet (1992) est destinée à l'évaluation des publics de bas niveaux... (voir document n° 5). Les épreuves d'évaluation de l'intelligence globale ou de l'intelligence générale s'appuient sur des théories relativement imprécises : qu'est-ce que la globalité de l'intelligence ? Quel contenu psychologique donner au facteur G ? Ces questions restent largement ouvertes. *A contrario*, les tests opératoires sont dérivés d'une théorie hautement structurée : le constructivisme piagétien.

Remarque :

Les quelques épreuves citées le sont simplement à titre d'exemples ; pour approfondir la connaissance des épreuves existantes, il convient de s'adresser aux éditeurs spécialisés.

Chapitre 5

Les tests opératoires

Différents auteurs se sont inspirés de la théorie de Piaget pour élaborer des épreuves d'évaluation de l'intelligence connues sous le terme générique de tests opératoires. En quelques mots, l'immense travail d'observation active et de théorisation du maître genevois se résume dans l'orthogénèse. Piaget introduit l'idée majeure que l'intelligence se construit dans le jeu des interactions qui lient l'enfant à son environnement physique et humain. Il décrit dans le détail trois grandes périodes qui se dérouleraient selon un ordre immuable en tous lieux et en tous contextes culturels.

A la période de l'intelligence sensorimotrice succède la période de l'intelligence pré-opératoire puis de l'intelligence opératoire concrète qui sera elle-même englobée dans la période de l'intelligence opératoire formelle. Chaque période se subdivise en stades et sous-stades. L'objectif de ce chapitre n'étant pas de décrire ce schéma explicatif de la genèse de l'intelligence, le lecteur intéressé pourra se reporter au résumé qu'en a rédigé Piaget lui-même (Piaget 1972, chap. 2).

Selon Piaget, un stade se caractérise par une structure d'ensemble des opérations de pensée. La structure se réfère à la logistique. Par exemple, pour les opérations dites concrètes s'appliquant sur les objets ou leurs représentations mentales, la structure est définie comme un groupement avec ses caractères logiques. Ainsi, un groupement additif des classes comprendra une opération directe, une opération inverse, une opération associative, une opération identique, une opération spéciale ou tautologie.

Selon la théorie, un sujet dont le développement s'équilibre à la phase finale de l'élaboration d'un stade doit parvenir à résoudre les problèmes mettant en jeu les opérations

qui caractérisent la structure d'ensemble de ce stade. D'où l'intérêt manifesté par certains auteurs d'évaluer le stade opératoire atteint par un sujet à un moment donné de son développement.

Ainsi Longeot a-t-il mis au point l'*Echelle de développement de la pensée logique* (E.P.L.), de passation individuelle, et les *Tests d'opérations formelles* (T.O.F.), de passation collective. L'I.N.E.T.O.P. a élaboré l'*Echelle collective de développement logique* (E.C.D.L.) et Schircks un *Test de raisonnement*. Ces épreuves proposent des items se référant à différents types d'opérations mentales décrites par Piaget. Par exemple, le *Test de raisonnement* propose des items relatifs aux opérations propositionnelles, à la proportionnalité, à la dissociation de facteurs, aux probabilités, à la combinatoire (voir document n°6).

Suite à la correction, une note brute permet de passer à une note étalonnée de la même manière que pour les tests psychométriques. En plus, cette note renvoie à un stade de développement. Par exemple, avec l'E.C.D.L., le niveau de développement d'un sujet obtenant entre 3 et 8 points est situé au stade concret (voir document n°7).

Différentes séries d'études ont amené un auteur comme Longeot à relativiser la portée générale du concept de stade : "Des stades définis par une structure mathématique générale (fonction, groupement, groupe, treillis) et gouvernant les conduites cognitives, impliquent que des enfants placés dans des situations différentes y donnent des réponses de même niveau. Or, on remarque :

1. des différences considérables dans l'âge auquel un sujet réussit des épreuves exigeant la même structure opératoire
 2. des sujets de même âge se situent à des stades différents dans des épreuves différentes
 3. les corrélations entre les épreuves opératoires sont moyennes ou faibles
 4. les applications que l'on voulait tirer des stades dans les domaines de la pédagogie et de l'examen psychologique deviennent impossibles en raison des trois points précédents."
- (Longeot, 1990)

Ecrivez tous les mots imaginables avec les 4 lettres B A S F
(même s'ils n'ont pas de sens). En voici deux : B F S A , A S F B .

.....

.....

.....

.....

.....

DOCUMENT N°6: Item de combinatoire du Test de raisonnement
(Schircks, 1970)

Une étude menée par Lautrey illustre cette remise en cause de l'existence d'un stade général, explicatif de l'ensemble des conduites d'un même sujet. Les cinq épreuves de l'E.P.L. ont été données à un groupe de 210 sujets de 9 à 16 ans :

1. épreuve d'opérations combinatoires dans le cas de permutations ;
2. épreuve faisant appel aux notions de proportion et de probabilité ;
3. épreuve faisant appel à la logique des propositions ;
4. épreuve exigeant de coordonner deux systèmes de références distincts dans l'espace ;

5. épreuve de conservation poids et volume associée à une épreuve de dissociation poids et volume.

Si un même et seul stade caractérisait l'ensemble de la conduite cognitive d'un sujet alors les réponses de ce sujet devraient situer son développement au même stade pour les cinq épreuves. Ceci n'est vérifié que pour 16% des sujets, 46% présentant des écarts d'un stade entre deux épreuves au moins, 33% des écarts de deux stades, 4% des écarts de trois stades, 1% des écarts de quatre stades (Lautrey, 1980).

Stade	Total
inférieur au concret	≤ 2
concret	3 - 8
intermédiaire	9 - 13
formel A	14 - 17
formel B	≥ 18

DOCUMENT N°7: Tableau de passage de la performance au stade dans l'E.C.D.L. (INETOP, 1990)

Avec l'aimable autorisation des E.A.P.

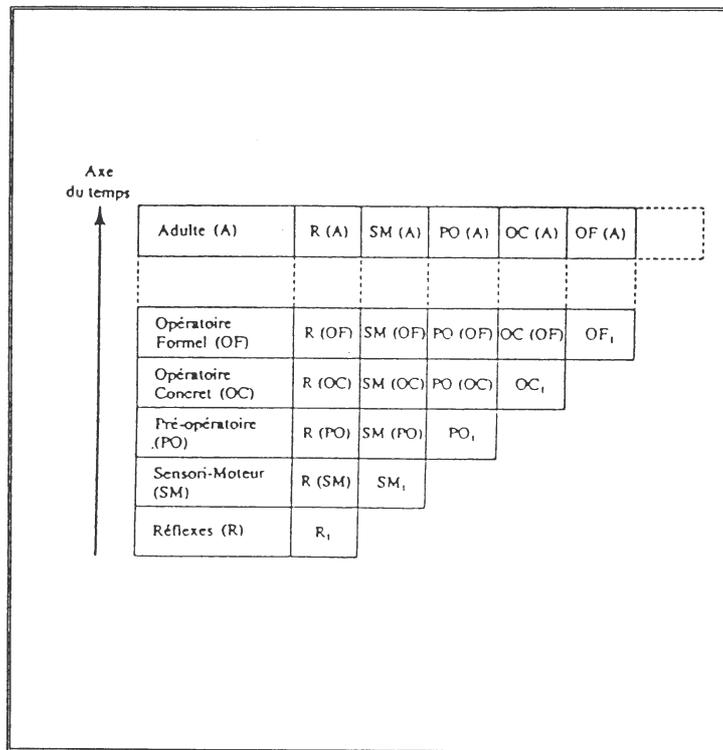
D'autre part, des analyses factorielles de batteries mixant épreuves psychométriques classiques et épreuves opératoires montrent que ces dernières se comportent comme les premières. "...dans un cas comme dans l'autre, les tâches ne sont pas homogènes et se regroupent en plusieurs ensembles distincts (ce qui donne lieu à des facteurs de groupe), ...leurs inter-corrélations sont néanmoins toutes positives (ce qui donne lieu à un facteur général." (Lautrey 1990)

Il ressort de tout cela que si les épreuves opératoires conservent un intérêt pour le psychologue, le pédagogue et le formateur, cet intérêt ne résiderait plus dans la possibilité de situer un sujet à un stade de développement général, mais plutôt dans la centration sur les aspects fonctionnels de l'intelligence en action dans une situation de résolution de problème.

On constate une utilisation des tests opératoires dans le domaine de la formation des adultes. Cela ne va pas de soi. Le premier problème à résoudre est celui des étalonnages : seul le *Test de raisonnement* de Schircks a été élaboré et étalonné auprès d'adultes en formation. Les autres épreuves citées viennent du domaine de la psychologie de l'enfant et ne s'accompagnent pas d'étalonnages pour les adultes. Leur utilisation directe auprès d'adultes ne peut se justifier. D'autre part, le concept de stade de développement n'a aucun sens pour des adultes chez lesquels la mise en place des structures cognitives est terminée. Dire d'un adulte qu'il fonctionne au stade concret relève d'une confusion entre une psychologie génétique qui rend compte du développement de l'enfant et une psychologie différentielle qui cherche à expliquer les différences entre les adultes. En tirer des conclusions quant à son orientation vers l'exercice des métiers manuels relève de la caricature grossière.

Pour pallier ces inconvénients, les concepts opératoires appliqués aux adultes ont sensiblement évolué. Des auteurs comme Schircks et Laroche (1970) ont préféré le concept de niveau opératoire, de niveau de stratégie à celui de stade. Vermersch est allé le plus loin dans l'utilisation des concepts piagétiens appliqués à la description du fonctionnement cognitif des adultes. Il décrit des registres de fonctionnement présents à

l'état latent dans l'équipement cognitif de l'adulte et mobilisables en fonction des caractéristiques des situations à maîtriser. (voir document n°8) "A chaque âge l'enfant ou l'adulte disposent d'une pluralité de possibilités pour régler leur conduite.



DOCUMENT N°8: Les registres de fonctionnement disponibles (Vermerch, 1979)

Avec l'aimable autorisation de la revue Education Permanente.

La mise en jeu d'une de ces classes d'instruments définit le registre du fonctionnement du sujet. Ainsi les instruments sensori-moteurs mis en jeu de façon isolée définissent le registre sensori-moteur. Mais si leur mise en jeu est

subordonnée à une planification (cas d'un menuisier réalisant un assemblage par exemple) le registre de fonctionnement sera défini par les instruments intellectuels réglant cette planification (fonctionnement opératoire concret ou formel). (Vermersch, 1979)

Mais peut-on tirer les tests opératoires vers ce cadre interprétatif sans risquer de "bricoler" ? Ne serait-il pas préférable de reprendre à sa base une élaboration qui tienne compte des études empiriques et des remaniements théoriques les plus récents ? Les tests opératoires ont eu, d'un point de vue fonctionnel, l'immense mérite de commencer à sensibiliser le psychologue et le formateur sur les processus de résolution en référence à un type d'opération mentale.

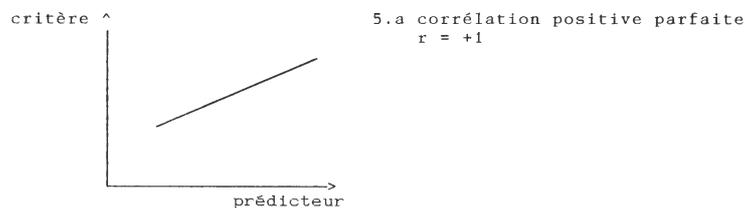
Cette amorce d'une approche individuelle et différentielle de l'évaluation mériterait de plus larges développements. C'est là un enjeu d'importance en ce qui concerne les diagnostics en vue de déterminer les objectifs et les contenus de remédiation cognitive véritablement adaptés aux besoins. Une approche opératoire actualisée pourrait également participer à la recherche d'une meilleure validité des pronostics réalisés avec les tests d'intelligence.

Chapitre 6

Les tests d'intelligence et la prédiction de la réussite des apprentissages professionnels

Les tests d'intelligence sont utilisés pour élaborer des prévisions de réussite d'un individu qui souhaite aborder un cursus de formation professionnelle ou occuper tel ou tel poste de travail. Techniquement, on met en relation le résultat d'un test pris comme prédicteur et le résultat d'une évaluation pédagogique (notes, examens...) ou d'une appréciation professionnelle (entretien, grille d'observation, indices de performance...) ; pris comme critère. La force de cette liaison, qui exprime le niveau de validité prédictive du test, se calcule à l'aide d'un coefficient de corrélation.

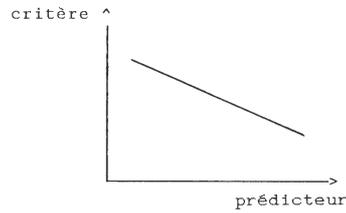
Ce coefficient, calculé sur les performances au test et à l'évaluation pédagogique ou professionnelle d'un même groupe de sujets, concrétise le degré de communauté entre les classements obtenus à ces deux épreuves, prédicteur et critère. Dans la mesure où ces deux classements sont identiques, la connaissance de l'un est suffisante pour connaître l'autre. Dans ce cas d'une corrélation positive parfaite, le coefficient prend la valeur +1 (voir figure n°5.a).



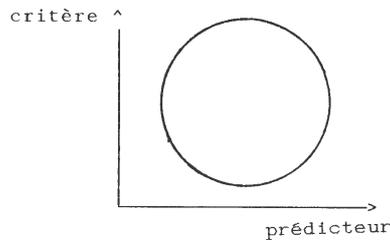
La comparaison des classements peut faire ressortir une correspondance des rangs parfaitement inverse : le premier sur

le prédicteur est le dernier sur le critère, le second sur le prédicteur est l'avant-dernier sur le critère, le troisième sur le prédicteur est l'antépénultième sur le critère... Dans ce cas d'une corrélation négative parfaite, le coefficient de corrélation prend la valeur -1 (voir figure 5.b).

5.b corrélation négative parfaite
 $r = -1$

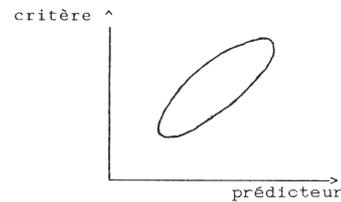


Il se peut qu'aucun lien statistique ne puisse être établi entre les deux variables, auquel cas le coefficient prend la valeur 0 (voir figure n° 5.c). Assez souvent on obtient des corrélations partielles entre les deux variables et le coefficient prend alors une valeur supérieure à 0 et inférieure à +1 (voir figure 5.d). Le nuage de points présente alors une forme caractéristique, fuselée à la manière d'un cigare.



5.c absence de corrélation
 $r = 0$

5.d corrélation positive partielle
 $0 < r < 1$



Des tables de significativité permettent de conclure si la liaison calculée est due à un hasard d'échantillonnage ou à l'effet systématique d'un facteur latent. Deux paramètres entrent en jeu dans l'utilisation de ces tables : l'effectif et le risque d'erreur (voir document n°9).

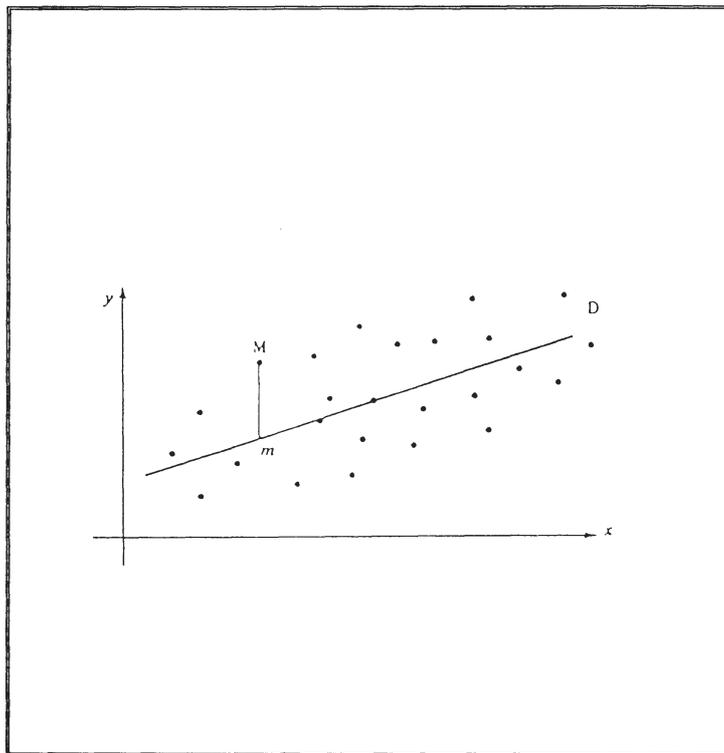
v \ P	P			
	0, 10	0, 05	0, 02	0, 01
1	0, 9877	0, 9969	0, 9995	0, 9999
2	0, 9000	0, 9500	0, 9800	0, 9900
3	0, 8054	0, 8783	0, 9343	0, 9587
4	0, 7293	0, 8114	0, 8822	0, 9172
5	0, 6694	0, 7545	0, 8329	0, 8745
6	0, 6215	0, 7067	0, 7887	0, 8343
7	0, 5822	0, 6664	0, 7498	0, 7977
8	0, 5494	0, 6319	0, 7155	0, 7646
9	0, 5214	0, 6021	0, 6851	0, 7348
10	0, 4973	0, 5760	0, 6581	0, 7079
11	0, 4762	0, 5529	0, 6339	0, 6835
12	0, 4575	0, 5324	0, 6120	0, 6614
13	0, 4409	0, 5139	0, 5923	0, 6411
14	0, 4259	0, 4973	0, 5742	0, 6226
15	0, 4124	0, 4821	0, 5577	0, 6055
16	0, 4000	0, 4683	0, 5425	0, 5897
17	0, 3887	0, 4555	0, 5285	0, 5751
18	0, 3783	0, 4438	0, 5155	0, 5614
19	0, 3687	0, 4329	0, 5034	0, 5487
20	0, 3598	0, 4227	0, 4921	0, 5368
25	0, 3233	0, 3809	0, 4451	0, 4869
30	0, 2900	0, 3494	0, 4093	0, 4487
35	0, 2746	0, 3246	0, 3810	0, 4182
40	0, 2573	0, 3044	0, 3578	0, 3932
45	0, 2428	0, 2875	0, 3384	0, 3721
50	0, 2308	0, 2732	0, 3218	0, 3541
60	0, 2108	0, 2500	0, 2948	0, 3248
70	0, 1954	0, 2310	0, 2737	0, 3017
80	0, 1829	0, 2172	0, 2565	0, 2830
90	0, 1726	0, 2050	0, 2422	0, 2673
100	0, 1638	0, 1948	0, 2301	0, 2540

DOCUMENT N°9: Table r de Bravais-Pearson (Fischer, 1947)

Avec l'aimable autorisation des P.U.F.

Par exemple, on lit sur la table que pour 15 couples d'observations, donc à $v = 13$, le coefficient sera considéré comme significatif d'une liaison statistique entre les deux variables s'il est au moins égal à 0,4409 avec un risque d'erreur de 10%, et s'il est au moins égal à 0,6411 avec un risque

d'erreur de 1%. Si on prend 62 couples d'observations, donc à $n = 60$, ces valeurs, aux mêmes risques, chutent à 0,2108 et à 0,3248.



DOCUMENT N°10: Droite des moindres carrés
(Langouët, Porlier, 1981)

Avec l'aimable autorisation de ESF éditeur.

Une corrélation statistiquement établie significative n'implique pas forcément une liaison de causalité directe entre les deux variables. Par exemple, sur un échantillon d'enfants d'âges variés, il est facile de montrer un lien entre le poids et l'âge mental. Bien entendu, le poids ne détermine pas l'âge mental, pas plus que l'âge mental ne détermine le poids. C'est le

jeu d'une troisième variable, qu'on peut qualifier de variable latente, qui explique la covariation des deux premières : l'effet de la maturation. De la même manière, un élève ne réussit pas son examen parce qu'en début d'année il a réalisé une bonne performance sur un test prédictif : ce sont ses capacités intellectuelles et ses motivations qui expliquent sa réussite aux deux types d'épreuves.

Le deuxième étalonnage concerne la probabilité de passage en classe de sixième en fonction de la note totale.

Note totale	70-58	57-49	48-41	40-30	29-22	21-11
probabilité de passage	100 %	96 %	91 %	80 %	65 %	57 %

Par exemple, un élève ayant une note totale de 25 points a, statistiquement parlant, 65 chances sur 100 de passer en classe de sixième.

DOCUMENT N°11: Etalonnage prédictif du TP6 (Pasquier, 1985)

Avec l'aimable autorisation des E.A.P.

Lorsqu'il existe un lien statistique significatif entre un prédicteur et un critère, on peut calculer la valeur probable du second connaissant le premier par le biais de la droite des

moindres carrés dont l'équation est de la forme $y = ax + b$ (voir document n° 10).

.38 à .54 pour l'essai professionnel,
.53 pour les batteries multifactorielles,
.41 à .43 pour les Assessments Centers,
.41 à .43 pour les évaluations données par les collègues et
l'encadrement,
.25 à .45 pour les tests d'aptitudes mentales générales,
.24 à .38 pour les inventaires biographiques,
.17 à .26 pour les références professionnelles,
.14 à .23 pour les interviews (entretiens),
.15 pour les tests de personnalité,
.10 pour les tests d'intérêts,
.15 pour les inventaires d'autodescription,
.00 pour la graphologie.
(d'après Lévy-Leboyer - 1987)

DOCUMENT N°12: Validité prédictive des outils du psychologue
(d'après Lévy-Leboyer, 1987)

Avec l'aimable autorisation des P.U.F.

Une autre forme de présentation consiste à utiliser la technique du "cut-off" ou tronçonnage. Par exemple, on sait qu'en 1993, en région Centre, 381 apprentis se sont présentés au CAP de cuisinier et 190 ont été admis, soit environ un sur deux (C.E.R.I.F., O.N.I.S.E.P.). Si l'on veut déterminer à l'aide d'un test prédicteur les chances de réussite au CAP cuisine, on appliquera cette proportion de 50% sur la distribution des notes du test afin de déterminer le seuil en-deçà duquel les chances de

réussite s'amointrissent. Plus finement, il est possible de segmenter la distribution des notes du prédicteur en fonction de l'évolution du taux de réussite sur le critère (voir document n°11).

Les tests ont pu avoir mauvaise presse et pourtant, quand on compare leur valeur prédictive à celle d'autres outils, techniques ou méthodes, on constate qu'ils donnent des résultats intéressants. Lévy-Leboyer (1987) a proposé un tableau récapitulatif des corrélations obtenues par différents auteurs entre les prédicteurs utilisés en psychologie et les critères de réussite dans la tenue de postes de travail (voir document n°12).

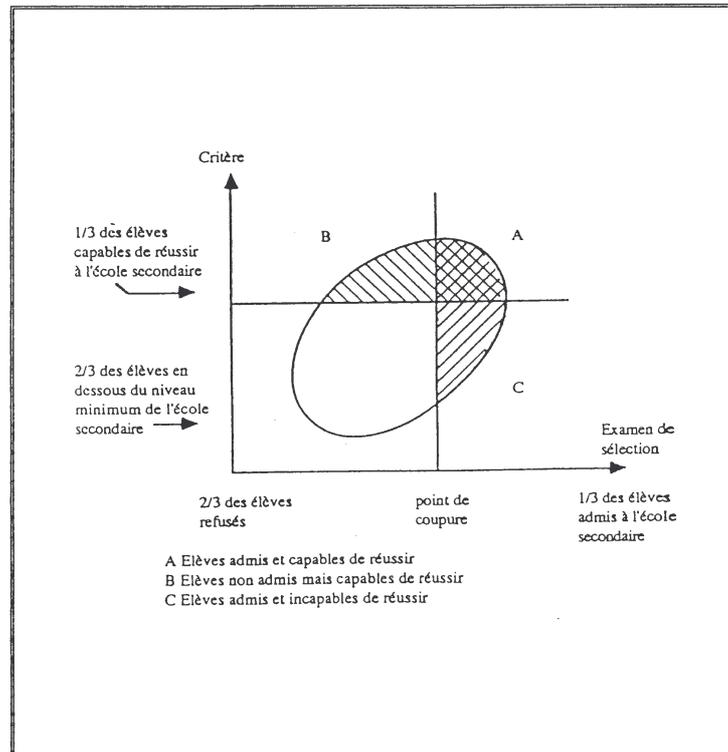
Trois constats peuvent être tirés de la lecture du document 12. En premier lieu, aucune technique à elle seule n'est suffisante pour assurer une prévision sans risque. En conséquence, il apparaît souhaitable d'avoir recours à plusieurs techniques, si possible relativement indépendantes l'une de l'autre et chacune présentant un bon degré de validité. Par exemple, un essai professionnel sera complété par la passation d'une batterie multifactorielle et par un inventaire biographique...

En second lieu, force est de constater que la fréquence des pratiques n'a pas de lien avec leur efficacité abordée de manière scientifique. Ainsi, la graphologie, largement utilisée dans les procédures de recrutement, n'est pas démontrée valide sur le plan de la prédictivité.

En troisième lieu, les validités prédictives ne sont pas identiques selon l'outil d'évaluation mis en oeuvre. D'une manière générale, il semblerait que les épreuves les moins valides touchent plutôt au domaine des variables conatives, variables de personnalité, d'intérêts... A l'autre extrémité, les épreuves les plus valides se rapprochent des mises en situation professionnelle réelles ou simulées. Les meilleurs tests sont aussi valides que ces mises en situation professionnelle, mais il est vrai avec des corrélations qui restent limitées (dans le domaine scolaire, ce type de corrélations prédicteur-critère atteint des degrés de validité nettement plus élevés).

Une application rigide de corrélations faibles peut conduire à des résultats de sélection aberrants. Par exemple,

d'après Cardinet (voir document n°13), une liaison prédicteur-critère pédagogique de 0,30 conduit à des erreurs de choix importantes. "La procédure de sélection empêche plus d'élèves de recevoir une éducation supérieure lorsqu'ils en seraient capables (B), qu'elle n'en accepte (A). Plus on est sévère à l'examen pour limiter les erreurs (C), plus on rejette de candidats capables (B). La proportion d'élèves faibles reste très élevée parmi les élèves admis (proportion de C parmi C + A), de sorte que l'avantage est minime pour l'enseignement : les classes restent presque aussi hétérogènes." (Cardinet, 1986)



DOCUMENT N°13: Résultat d'une sélection pour une validité de 0,30 (Cardinet, 1986)

Extrait de *Evaluation scolaire et mesure* de J. Cardinet, De Boeck-Wesmael, Bruxelles, 1988. Avec l'autorisation de l'éditeur.

Il est vrai que la relation aptitude-réussite, qui fonda la psychologie industrielle des années vingt, n'est plus à prendre au pied de la lettre dans la mesure où les psychologues ont réalisé l'importance des conditions du milieu sur l'investissement des capacités dans l'activité professionnelle. Par exemple, on a montré que des cadres "plutôt doués" présentaient des difficultés quand ils étaient sous-employés et que, réciproquement, des cadres "moins doués" mais sur-employés se démotivaient de la même manière. Il est donc clair que la réussite professionnelle ne dépend pas des seules aptitudes intellectuelles ; les variables motivationnelles jouent un rôle majeur dans la libération de l'investissement. Ces variables dépendent en grande partie des caractéristiques des contextes des organisations, par exemple du type de prise en considération de la personne par la hiérarchie. Pour affiner la validité prédictive des épreuves, il y a donc intérêt à définir du mieux possible les variables de contextes, puis à les mettre en relation avec la plasticité des aptitudes.

A ces formes d'explication de la relative faiblesse des corrélations par des facteurs environnementaux, il convient de rajouter un type de critique qui touche à l'aspect sommatif et statique des tests traditionnels. "Les tests classiques mettent l'accent sur ce qui ne change pas chez le sujet : le classement dans l'échantillon ou le Q.I. n'ont d'intérêt que s'ils caractérisent les enfants de manière stable ; c'est la condition "sine qua non" de la validité des épreuves utilisées. Conçues pour repérer ce qui reste invariant au cours du développement, ces épreuves sont peu propices à inspirer une action éducative axée par définition sur les transformations." (Huteau, Lautrey, 1978)

D'autres auteurs sont encore plus radicaux dans les réserves qu'ils expriment, et Rey (1934) critique vigoureusement l'utilisation des tests statiques. Certes, un sujet qui réussit une épreuve démontre qu'il a des capacités et qu'il sait les utiliser. Mais un mauvais résultat ne renseigne pas sur le devenir cognitif du sujet : peut-être a-t-il des capacités mais il ne sait pas bien les utiliser dans toutes les situations qu'il rencontre ; peut-être commence-t-il seulement à développer le type de capacité requise par le test et il conserve une marge de

progression importante qu'il va prochainement actualiser... La critique la plus percutante souligne l'aspect paradoxal de l'évaluation statique : "C'est vraiment curieux que nous utilisions les tests d'intelligence principalement pour prédire la capacité d'apprentissage alors qu'aucun de nos tests n'implique un apprentissage, à la place ils nous donnent plutôt une coupe de ce qui a été appris." (Vernon, 1950)

Le test classique indique un niveau de réussite, c'est-à-dire un état. Un apprentissage ou une adaptation relève d'une logique de transformation. On ne peut *a priori* inférer le sens et l'amplitude d'une transformation à partir de la seule connaissance d'un état. Ce n'est pas parce que le thermomètre marque 18°C que je peux prévoir si la température dans une heure sera restée stable, ou aura augmenté, ou aura diminué.

La modération des corrélations exprimant la validité prédictive des tests classiques peut provenir de ce décalage entre la situation d'évaluation propre au test-prédicteur (résolution de problèmes ou exécution de tâches sans apprentissage préalable) et la situation d'évaluation propre à l'épreuve critère (résolution de problèmes ou exécution de tâches suite à des apprentissages spécifiques).

L'évaluation dynamique, dans la mesure où elle introduit une phase d'apprentissage dans la procédure du test, opère un rapprochement entre les situations évaluatives propres au prédicteur et au critère.

Troisième partie

L'évaluation dynamique du potentiel d'apprentissage

Les épreuves classiques et les épreuves opératoires se réfèrent à une évaluation sommative qui se concrétise dans le positionnement social de la performance du sujet sur une ou plusieurs dimensions (profil de résultats). L'évaluation dynamique se rattacherait plutôt à une procédure d'évaluation formative telle qu'elle se définit dans le champ de la pédagogie. L'évaluation formative intervient au terme de chaque tâche d'apprentissage pour informer le formateur et l'apprenant du degré de maîtrise atteint. De plus, ce type d'évaluation cherche à repérer et à identifier la nature et la source des difficultés d'apprentissage. A partir de ces constats, il devient possible de proposer à l'élève des stratégies d'apprentissage pertinentes pour lever ses difficultés et lui permettre de progresser. La régulation peut également porter sur un réajustement des objectifs et des moyens pédagogiques mis en oeuvre pour les atteindre. (De Landsheere, 1979).

Alors que l'évaluation sommative ne prend en compte que la quantité et la conformité de la réponse (nombre total de bonnes réponses produites par le sujet), l'évaluation formative va se centrer sur le processus cognitif qui conduit l'élaboration de la réponse. Dans le domaine des tests, l'évolution va également dans le sens d'une prise en compte de ce qui se passe dans la "boîte noire" ; c'est à dire que le psychologue, comme le formateur, va pouvoir se centrer sur le processus mental de la résolution d'un problème ou de l'apprentissage de l'exécution d'une tâche.

Concrètement, pour atteindre ce processus fonctionnel, toutes les formes d'évaluation dynamique introduisent une mise en situation d'apprentissage du sujet. C'est l'observation et la

description des comportements et des conduites du sujet en train de résoudre ou d'apprendre à résoudre qui fournira l'ensemble des informations utiles pour conjecturer l'orientation et le mode de fonctionnement de ses capacités.

L'évaluation dynamique repose sur une idée relativement ancienne, exprimée dans les années trente par deux précurseurs : Vygotski et Rey (chapitre septième). Plus ou moins laissée de côté par les psychologues, l'évaluation dynamique suscite un renouveau d'intérêt à partir des années soixante et donne lieu à une série d'études menées auprès des enfants (chapitre huitième). L'utilisation de l'évaluation dynamique avec des publics adultes tout venant semble beaucoup plus récente et conserve un caractère largement exploratoire (chapitre neuvième).

Pour aller plus loin dans le domaine de l'évaluation dynamique :

Lidz C.S. (Ed.) (1987), *Dynamic assessment. An interactionnal approach to evaluating learning potential*, New-York : The Guilford Press

Chapitre 7

Les précurseurs et les principes généraux de l'évaluation dynamique

L'œuvre de Vygotski ne se résume pas en quelques lignes. Dans le cadre de ce chapitre, seuls quelques aspects en seront abordés : ceux qui ont un rapport direct avec l'évaluation dynamique du potentiel d'apprentissage. En ce sens, Vygotski propose le concept de zone de proche développement : "Pour déterminer le niveau présent du développement on utilise des problèmes que l'enfant doit résoudre tout seul et qui ne sont indicatifs que par rapport aux fonctions déjà formées et venues à maturité. Mais essayons d'utiliser une nouvelle méthode. Admettons que nous ayons déterminé chez deux enfants un âge mental équivalant à huit ans. Si l'on va plus loin et qu'on essaie de voir comment les deux enfants résolvent les problèmes destinés aux âges suivants lorsqu'on leur vient en aide en leur montrant le début de la solution, etc., il apparaîtra qu'avec de l'aide, en collaboration avec un adulte, en suivant ses indications, l'un des deux résout jusqu'à des problèmes correspondant à l'âge de douze ans et l'autre des problèmes correspondant à l'âge de neuf ans. Cette disparité entre l'âge mental, ou niveau de développement présent, qui est déterminé à l'aide des problèmes résolus de manière autonome, et le niveau qu'atteint l'enfant lorsqu'il résout des problèmes non plus tout seul mais en collaboration détermine précisément la zone de proche développement. Dans notre exemple, pour le premier enfant cette zone est exprimée par le chiffre 4, pour l'autre par le chiffre 1. Pouvons-nous considérer que les deux enfants ont un niveau identique de développement mental, que l'état de développement de l'un coïncide avec celui de l'autre ? Évidemment non. Comme le montre la recherche, chez ces enfants les différences conditionnées par la disparité de leurs

zones respectives de proche développement s'avéreront beaucoup plus grande à l'école que la ressemblance due à un niveau identique de développement présent. Cela se manifestera avant tout dans la dynamique de leur développement intellectuel et dans la réussite relative au cours de l'apprentissage scolaire. La recherche montre que la zone de proche développement a une signification plus directe pour la dynamique du développement intellectuel et la réussite de l'apprentissage que le niveau présent de leur développement." (Vygotski, 1985)

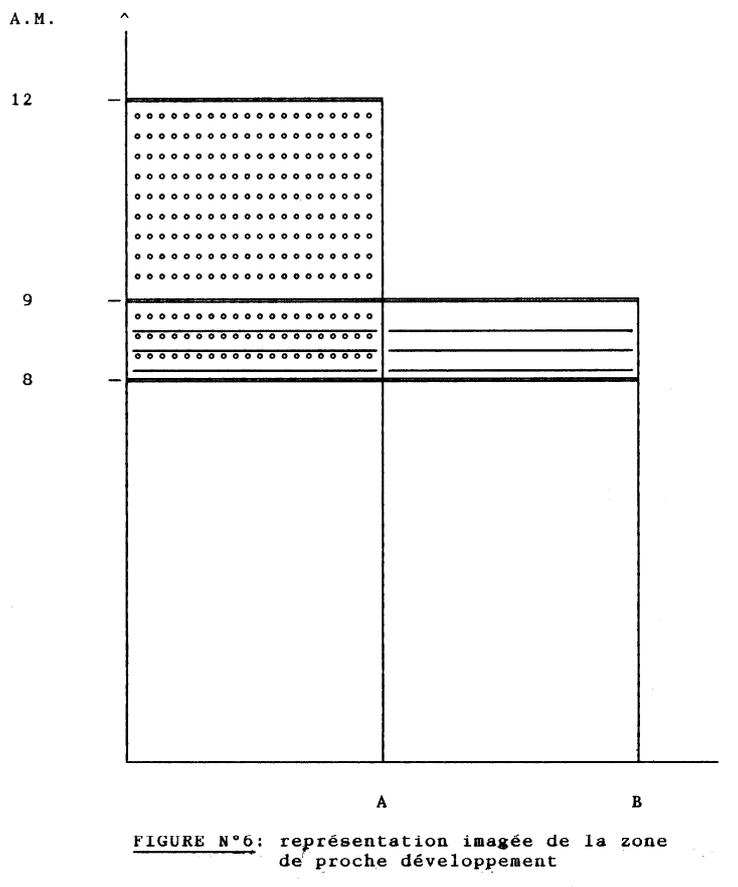
Cette citation, il est vrai un peu longue, est reproduite dans son entier car elle fonde dès 1934 le principe de base de l'évaluation dynamique. Elle inspire aujourd'hui quasiment toutes les recherches sur ce thème. La figure n°6 aide à visualiser cette définition. Face à deux enfants à évaluer, le travail du psychologue s'arrête à l'établissement de leurs âges mentaux. Dans la perspective de Vygotski, l'âge mental représente le niveau de la performance spontanée de l'enfant, atteint dans une situation de résolution autonome.

Cet âge mental n'est pas le point terminal de l'évaluation, comme dans la perspective habituelle du psychologue, mais le constat de base et le point de départ du processus d'évaluation dynamique. En effet, Vygotski introduit une seconde phase consacrée à un entraînement à résoudre les items du test, entraînement basé sur une interaction sociale, une médiation, entre le sujet novice et un expert qui lui transmet les éléments de méthode à appliquer pour réussir.

Une seconde passation de l'épreuve permet d'objectiver l'effet de la mise en situation d'apprentissage sur le niveau de performance en mesurant l'écart entre les résultats obtenus aux pré-test et les résultats obtenus au post-test. Le gain de performance que Vygotski dénomme zone de proche développement traduit la sensibilité, la réactivité des capacités cognitives du sujet placé dans une situation d'apprentissage faisant intervenir une médiation humaine.

Pour Vygotski, le développement des fonctions mentales supérieures suit une loi fondamentale : il passe obligatoirement par l'interaction sociale dans la mesure où...

"Chaque fonction psychique supérieure apparaît deux fois au cours du développement de l'enfant : d'abord comme activité collective, sociale et donc fonction interpsychique, puis la deuxième fois comme activité individuelle, comme propriété intérieure de la pensée de l'enfant, comme fonction intrapsychique." (Vygotski, in Schneuwly, Bronckart, 1985)



Concrètement, le développement d'une fonction mentale supérieure nécessite le croisement d'une fonction biologique et d'une médiation sociale d'un outil psychologique. Par exemple, le croisement du développement biologique de la mémoire et des interactions parlées entre l'enfant et son

entourage humain va permettre l'intériorisation et l'utilisation de plus en plus large et de mieux en mieux différenciée du langage.

Le langage, et d'une manière plus générale tous les systèmes de codes symboliques, propositionnels, logicomathématiques, les systèmes de valeurs... sont des élaborations culturelles apprises, transmises et réaménagées de génération en génération. Accepter cette loi fondamentale du développement comme procès d'intériorisation d'outils psychologiques historiquement et culturellement élaborés par le biais d'une médiation sociale revient pour le psychologue et pour le formateur à se donner un outil conceptuel utile au pronostic, au diagnostic et à la remédiation.

Sur le plan de l'évaluation du potentiel d'apprentissage, le modèle de Vygotski, à travers la procédure de la mise en évidence d'une zone de proche développement, offre tout à la fois des possibilités de diagnostic (aspects quantitatifs et qualitatifs de l'intériorisation), et des possibilités de pronostic (niveau et étendue de la zone de proche développement). Qui plus est, ce modèle présente une valeur heuristique de tout premier plan pour le pédagogue et le rééducateur : le niveau et l'étendue de la zone de proche développement indiquent avec précision la zone d'intervention du formateur.

Au-dessus de cette zone, sans point d'ancrage interne chez l'apprenant, le message pédagogique se dilue dans le tableau amorphe des stimuli qui participent au bruit, car non traités comme informations. En effet comme le souligne Bateson, la création d'une information répond à la perception d'une différence, et *a priori*, une différence ne se perçoit que par rapport à du connu (Bateson, 1984).

En-dessous de cette zone, le message pédagogique ne fait, au mieux, que stimuler des capacités opératoires déjà fonctionnelles. Il en résultera un gain en termes d'habileté (vitesse et précision) dans une tâche maîtrisée qui se répète à la condition qu'un enjeu suffisamment fort soutienne l'investissement (activité de bachotage par exemple). Mais aucune capacité nouvelle ne sera au rendez-vous. Vygotski argumentait ce point de vue en prenant pour exemple le cas des

enfants handicapés sur le plan intellectuel. En partant de l'hypothèse que ces enfants n'accéderont jamais à une pensée abstraite, les enseignants préfèrent les confronter à un matériel pédagogique visuel. "Il a été prouvé qu'un système d'enseignement basé exclusivement sur des moyens visuels excluant tout ce qui concerne la pensée abstraite non seulement n'aide pas l'enfant à dépasser son incapacité naturelle, mais la consolide de fait. En insistant sur la pensée visuelle, on étouffe les embryons de pensée abstraite chez ces enfants." (Vygotski, in Schneuwly, Bronckart, 1985)

En remplaçant le terme visuel par le terme concret, ce passage reste d'actualité dans le domaine de l'enseignement spécial, et d'une manière générale illustre ce que Feuerstein appelle l'approche passive acceptante (Feuerstein, Rand, Hoffman, Miller, 1980). On retrouve cette approche passive acceptante du bas niveau de fonctionnement dans tous les secteurs de l'enseignement initial, de la formation professionnelle, de l'entreprise. Cette approche postule la nécessité de posséder les capacités requises avant d'aborder une activité. L'absence de ces capacités représente l'argument majeur pour exclure l'individu de l'entrée dans cette activité. Vygotski, à la manière de Copernic, inverse radicalement les termes de l'équation : "Le seul bon enseignement est celui qui précède le développement..."

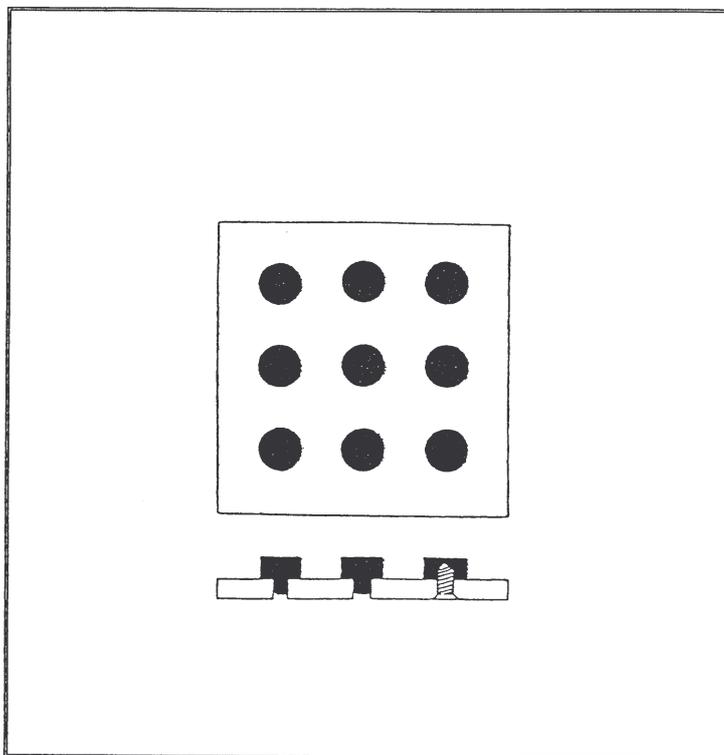
L'apprentissage donne donc naissance, réveille et anime chez l'enfant toute une série de processus de développements internes qui, à un moment donné, ne lui sont accessibles que dans le cadre de la communication avec l'adulte et de la collaboration avec les camarades, mais qui, une fois intériorisés, deviendront une conquête propre de l'enfant." (Vygotski, in Schneuwly, Bronckart, 1985) Il faut ajouter... à condition que cet enseignement se situe dans la zone de proche développement, et qu'il s'appuie sur la médiation sociale des capacités cognitives nécessaires aux nouveaux apprentissages. Dans cette perspective, la capacité ne précède plus l'apprentissage en tant que condition préalable ; au contraire, elle suit l'apprentissage en tant que produit de cet apprentissage socialement médiatisé dans une zone de proche développement.

Sur le plan du diagnostic individuel, il ne s'agit plus seulement de faire l'inventaire des acquis déjà consolidés, mais en plus de dégager et de baliser les pistes qui pourront conduire à de nouvelles acquisitions cognitives et à de nouvelles maîtrises opératoires. D'où la nécessité d'une mise en situation d'apprentissage socialement médiatisé afin de déterminer dans quelle mesure le sujet peut s'ouvrir une zone de proche développement représentative de son potentiel d'éducabilité. Ce sera ensuite à l'action éducative mise en oeuvre par l'enseignant ou le formateur de rendre progressivement le sujet autonome par intériorisation des outils cognitifs socialement médiatisés. Une fois la zone de proche développement conquise par le sujet, une nouvelle zone peut être évaluée et générer de nouveaux objectifs pédagogiques..., l'évaluation dynamique ouvrant la voie de la remédiation.

Vygotski présente son modèle à partir de la mise en oeuvre de tests d'intelligence. Rien n'empêche a priori d'étendre le paradigme de la zone de proche développement aux données proprement pédagogiques en dynamisant la passation de tests de connaissances.

Rey, psychologue suisse et inventeur de tests très proluxe, dès 1934, publie un article intitulé "D'un procédé pour évaluer l'éducabilité" (Rey, 1934). Rey explique dans cet article que pour évaluer la qualité et la rapidité de l'adaptation, la simple réponse aux items des tests n'est pas suffisante dans la mesure où elle ne fait que renseigner sur les "accommodats" déjà formés. "En vue d'un pronostic, on ne saurait... confronter, sans tomber dans l'absurde, l'existence d'un accommodat à son absence. Certes, la présence d'accommodats peut constituer une donnée significative ; en établissant que l'individu a pu et a su s'adapter dans le passé, elle laisse supposer, avec une forte probabilité, que les mêmes capacités continueront à s'adapter dans l'avenir. Rien, par contre, ne permet de soutenir la proposition inverse ; l'absence d'accommodats, à un moment donné du temps, ne signifie nullement que dans l'avenir et sous la pression des circonstances, l'adaptation ne puisse avoir lieu." (Rey, 1934)

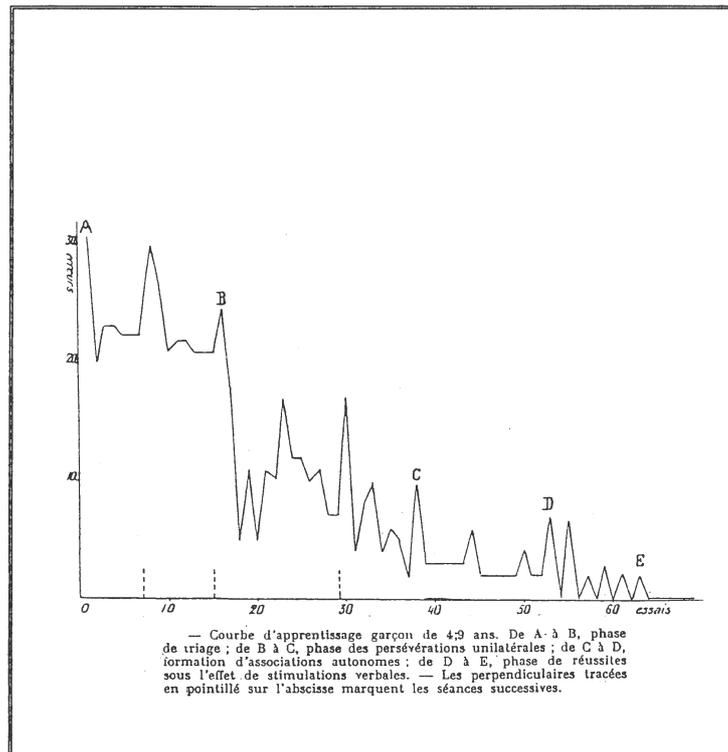
Dans cette perspective, il y a donc un intérêt majeur à compléter l'évaluation sommative des capacités préalablement acquises par l'examen des processus d'accommodation. Cela évitera de placer sur un même plan des individus ayant atteint un même niveau de performance alors que chez les uns les possibilités d'adaptation ont atteint leur maximum, que chez les autres elles conservent une marge d'évolution, et que chez les derniers elles n'ont pas encore commencé à se manifester.



DOCUMENT N°14: Labyrinthe manuel (Rey, 1934)

Pour étudier les processus d'accommodation, dans un premier temps, Rey préconise l'utilisation d'épreuves impliquant divers apprentissages. "On retient la manière et la rapidité avec

laquelle un individu s'adapte à une tâche déterminée et on compare, entre sujets, les caractères de l'éducabilité." (Rey, 1934) Dans l'article de 1934 Rey présente une épreuve originale, *Le labyrinthe manuel*, connu actuellement sous le nom de *Test des plateaux* (Rey, 1934). Le dispositif est composé de quatre planchettes de quatorze centimètres de côté. Sur chacune, neuf disques sont régulièrement disposés. Huit de ces disques sont mobiles et le neuvième est solidaire de la planchette (voir document n°14). Ce disque fixe occupe une position différente sur chacune des planchettes. Les plateaux avec leurs disques sont présentés en pile au sujet.



DOCUMENT N°15: Courbe d'apprentissage (Rey, 1934)

L'examineur demande au sujet de parvenir à saisir chaque planchette par le disque fixe, l'une après l'autre, et ce avec le nombre d'essais le plus réduit possible. L'examineur peut ensuite élaborer une courbe d'apprentissage individuelle qui visualise l'évolution du nombre des erreurs au fur et à mesure des essais (voir document n°15). Il peut également comparer la performance individuelle (nombre d'essais nécessaires pour stabiliser trois réussites successives) à la performance du groupe d'âge. Ultérieurement, Rey ira plus loin que l'élaboration d'une courbe d'apprentissage par essais et erreurs en introduisant une série d'aides actives devant conduire le sujet à la réussite.

"...le praticien donne... de l'extérieur, une discipline et un plan de travail, et c'est en observant comment le sujet s'y adapte et à quel degré de cette aide apportée il "se déclenche" que ses difficultés seront isolées. Cette découverte a certes beaucoup plus de valeur que la simple mise en évidence de l'échec au test." (Rey, 1982)

Dès cette époque Rey a souligné deux obstacles à une large pratique de l'évaluation dynamique, obstacles toujours d'actualité. Le premier obstacle concerne le temps. Le psychologue pressé n'ira pas spontanément vers des techniques lourdes, qui prennent plusieurs heures pour arriver à dresser un tableau clinique détaillé des processus qui sous-tendent la réactivité d'un sujet placé en situation d'apprentissage. Mais comme le souligne cet auteur : "Vouloir inventorier un psychisme en cinq minutes, c'est se condamner à n'en savoir que ce qu'on peut en saisir en cinq minutes." (Rey, 1934)

Le second obstacle touche à la question de l'implication du sujet dans le procès d'évaluation dynamique. Ces techniques de longue durée et qui bien souvent remettent en cause le fonctionnement spontané du sujet nécessitent un fort investissement. "L'absence d'intérêt, de mobile réel à l'action, peut donner des résultats illusoire." (Rey, 1934) En conséquence, il est inutile, voire préjudiciable, de placer une personne en situation d'évaluation dynamique sans qu'un enjeu externe majeur ne le justifie.

Dans la même période, Ombredane se situe dans ce courant de pensée : "Le test est un constat : il saisit la dérivée à un moment donné, de la fonction d'adaptation, ou éducatibilité." (Ombredane, 1936) L'évaluation doit en conséquence prendre en compte deux paramètres : le constat des acquis antérieurs mesurés par le test statique et la capacité d'apprentissage. Ombredane va donc chercher à capter d'emblée la fonction d'adaptation ou d'acquisition elle-même, c'est-à-dire l'éducatibilité du sujet.

Il conçoit alors un *Test d'éducatibilité*, épreuve d'apprentissage associatif qui restera pratiquement sans suite. Des propositions de ces grands précurseurs, il est possible de dégager les principes généraux qui fondent les pratiques d'évaluation dynamique d'aujourd'hui.

Abruptement, et par analogie avec le champ pédagogique, on peut dire qu'une procédure d'évaluation dynamique combine une évaluation sommative et une évaluation formative. L'évaluation sommative mesure l'efficacité du sujet placé en situation d'autonomie face à l'exécution d'une tâche ou à la résolution d'un problème ; l'évaluation formative évalue cliniquement et/ou psychométriquement, la réactivité du sujet placé dans une situation d'apprentissage.

Deux types de procédures peuvent être utilisées : le paradigme pré-test, apprentissage, post-test (T.A.T.) ; le paradigme apprentissage au cours du test (A.C.T.), (voir figure n°7). Plus récemment un paradigme plus économique a été proposé, apprentissage-test (A.T.) (voir *infra*, le T.E.D.E.98)

Le paradigme T.A.T. s'applique mieux aux situations d'évaluation collective. La même épreuve est généralement utilisée comme pré-test et comme post-test... Entre les deux passations s'intercale une séquence de mise en situation d'apprentissage au cours de laquelle une série d'aides méthodologiques est apportée aux sujets.

L'application du paradigme A.C.T. n'est possible qu'en situation d'évaluation individuelle. Moins économique en temps que le paradigme TAT, la procédure A.C.T. est *a priori* plus propice au recueil d'informations qualitatives sur le mode de

fonctionnement du sujet en situation d'apprentissage. Concrètement, les items sont travaillés l'un après l'autre. Une réussite spontanée permet de passer à l'item suivant alors qu'un échec déclenche la mise en oeuvre des aides par l'examineur.

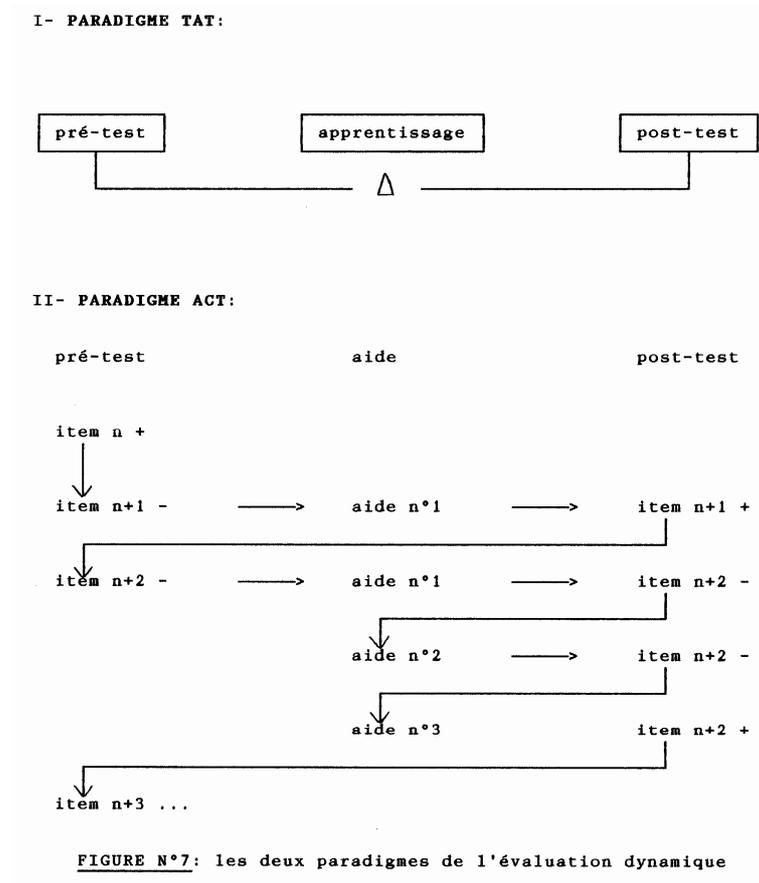
Quelque soit le paradigme, les aides peuvent prendre plusieurs formes. Dans le *Matrix Éducabilité* (Hurtig, 1966), l'examineur lit des explications verbales inspirées de celles fournies par les enfants qui réussissent la tâche.

Dans le T.E.D.E. (Pasquier, 1989c), l'aide est présentée sous la forme d'un mode opératoire écrit ; en cas de difficulté dans l'apprentissage de l'application de la méthode, un guidage opératoire pas à pas est mis en oeuvre. Dans le *Test des carrés*, (Mary, Mariel, 1978), dérivé du test des cubes de Kohs pour les écoliers de maternelle, en fonction des besoins, l'aide passe progressivement de l'explication verbale orale à la démonstration manipulatoire de la réalisation de figures modèles à l'aide de carrés de bois colorés.

Yerle, à l'occasion de son étude de la fidélité de la mesure du potentiel d'apprentissage, hiérarchise les aides en deux temps : aide à la découverte de la règle qui structure les relations internes des items (processus d'induction) ; et si nécessaire, aide à l'application de la règle (processus de déduction) pour inférer l'élément manquant d'une matrice, ou d'une série numérique (Yerle, 1992).

Lautrey a recensé différentes manières de quantifier la mesure du potentiel d'apprentissage. "Le simple gain" se calcule comme la différence des scores avant et après apprentissage. "Le gain résidualisé" se calcule comme la différence entre le score effectif obtenu au post-test et le score prévu calculé à l'aide de la droite de régression (voir chapitre 6). "Le score au post-test" mêle dans un indice unique la performance initiale et l'effet de l'apprentissage. "L'utilisation des modèles de réponse à l'item" représente la réponse la plus sophistiquée au problème de la mesure du changement, mais elle demande des nombres de sujets et d'items très importants. "L'analyse qualitative" d'aides non standardisées apportées par l'examineur s'inscrit dans une perspective clinique de l'évaluation du potentiel d'apprentissage ou de la modifiabilité structurelle cognitive.

"L'analyse quantitative" applique différents types de cotation aux aides apportées au cours de l'apprentissage pour évaluer le potentiel."(Lautrey, 1994) On peut également prendre en compte l'effet différentiel de reclassement d'un sujet au post-test, suite à l'entraînement : d'un sujet qui se situe à la même position sur les étalonnages des pré et post tests, on pourra dire qu'il présente une éducatibilité statistiquement normale, d'un autre qui améliore sa position, on dira qu'il a une éducatibilité statistiquement supérieure...



D'autres techniques sont encore utilisées comme le calcul du gain relatif, c'est-à-dire le pourcentage de progression effectivement réalisée suite à l'apprentissage par rapport à toute l'étendue de la progression possible suite au pré-test. (Pasquier, 1989) Ionescu propose le calcul d'un "quotient de gain", avec comme numérateur la "note transfert" (nombre d'items réussis après apprentissage), et comme dénominateur la "note spontanée" (nombre de réussites en situation de résolution autonome). (Ionescu et al., 1986)

Quelle que soit la procédure utilisée, et quel que soit le type d'indice de cotation retenu, les épreuves d'évaluation dynamique du potentiel d'apprentissage poursuivent des objectifs communs que listent Büchel et Paour : "...(a) décrire et comprendre les difficultés cognitives majeures. (b) mettre sur jour les potentialités d'apprentissage et de développement (notions et fonctionnements émergents) et les conditions de leur actualisation et (c) proposer un plan d'intervention éducative susceptible d'induire, à terme, l'actualisation effective des potentialités pressenties au cours de l'examen. Par ces objectifs, l'évaluation du potentiel d'apprentissage réconcilie, effectivement, mesure, compréhension, et éducation de l'intelligence." (Büchel, Paour, 1990)

Derrière ces objectifs communs, ces auteurs distinguent trois grands types de finalités. Dans un premier cas, le psychologue ou l'enseignant chercheur s'intéresse à la mise en évidence, à la description, à l'analyse d'un processus mental dans une perspective de recherche et d'élaboration de connaissances générales en psychologie de l'enfant ou de l'adulte. Il utilise alors des procédures d'évaluation dynamique du potentiel d'apprentissage, le plus souvent à caractère expérimental, comme moyen d'investigation, de découverte, d'objectivation de ce processus.

Dans un second cas, plus proches de la pratique professionnelle du psychologue, voire du formateur, intervenant en réponse à des demandes sociales de recrutement, de sélection, de diagnostic, de pronostic, d'orientation..., les auteurs d'épreuves dynamiques restent dans le cadre de la psychométrie traditionnelle. Leur objectif est bien de construire

un outil qui donne une mesure juste du potentiel d'apprentissage. Partant du postulat d'une certaine stabilité du niveau, du sens et de l'amplitude de la zone de proche développement, ces auteurs travaillent à l'amélioration des capacités métrologiques de leurs épreuves : sensibilité, fidélité et validité du potentiel d'éducabilité. Dans ces conditions, les procédures de mise en situation d'apprentissage seront rigoureusement standardisées. Il s'agit bien de se rapprocher au plus près du véritable positionnement social des capacités d'apprentissage individuelles, tout en réduisant les biais liés à l'absence de familiarité avec la tâche et aux artefacts de nature socioculturelle.

TABLEAU 1: classification des tests d'évaluation dynamique de l'éducabilité

OBJECTIF PARADIGME	MESURE DU P.A.	RECHERCHE DE LA M.S.C.
PRE-TEST APPRENTISSAGE POST-TEST		
APPRENTISSAGE AU COURS DU TEST		

Dans le troisième cas, l'approche psychométrique est écartée au profit d'une approche clinique. Le praticien ne cherche pas à établir une mesure fiable de la réactivité cognitive différentielle suite à une mise en situation d'apprentissage dont chacun des paramètres est standardisé. Il va, au contraire, jouer sur toute la gamme des aides médiatisées possibles, qu'elles soient de nature cognitive ou affective, dans le but de provoquer une modification du comportement du sujet : soit l'amener à

utiliser des capacités latentes, soit le faire entrer dans l'amorce de nouvelles maîtrises. Dans ce cadre, l'évaluation dynamique marque le premier pas de l'intervention de remédiation cognitive : il s'agit d'identifier les zones de sensibilité à certains types de médiation afin de les élargir par la suite dans l'objectif d'une optimisation de la modifiabilité structurelle cognitive. La question de la fidélité de l'épreuve s'en trouve inversée : "Ainsi, la fidélité de la tâche n'est plus considérée comme un critère pertinent de l'utilité du test ; au contraire, l'"infidélité" de la tâche, le degré avec lequel la tâche elle-même peut produire des changements chez le sujet examiné, sera le critère de sa sélection." (Feuerstein, Rand, Jensen, Kaniel, Tzuriel, 1987).

TABLEAU 2: éléments de relativisation de l'évaluation dynamique

<p>UNE EPREUVE:</p> <ul style="list-style-type: none"> - population parente - contenu - tâche - codage - questionnement - réponse - durée...
<p>UNE SITUATION D'APPRENTISSAGE:</p> <ul style="list-style-type: none"> - objet d'apprentissage - codage } de l'aide apportée - qualité } de l'aide apportée - quantité } - système de rétroaction - médiation sociale...
<p>UN SUJET:</p> <ul style="list-style-type: none"> - orientation cognitive - disposition mentale - sentiment de compétence - degré d'investissement - degré de familiarité - enjeu externe - perturbations éventuelles...

En fonction du paradigme utilisé et de l'objectif principal de l'épreuve, il devient possible de proposer un tableau pour la classification des tests d'évaluation dynamique de l'éducabilité (voir tableau n°1).

Quel que soit le paradigme utilisé et la finalité poursuivie, il est primordial d'accompagner les conclusions d'une évaluation dynamique des éléments de relativisation. C'est le seul moyen qui puisse permettre d'éviter les dérives qui conduisent, par des processus à la fois réducteurs et généralisants, à la réification d'une réalité psychologique individuelle complexe et interagissante. Il convient d'éviter que le potentiel d'apprentissage chiffré ne vienne à remplacer le Q.I. ou le facteur G dans une approche caricaturale du sujet dont les caractéristiques se retrouvent réduites à un nombre. Les éléments de relativisation concernent trois entités : l'épreuve, la situation d'apprentissage, le sujet (voir tableau n°2).

Il est nécessaire de préciser le paramétrage de l'épreuve utilisée comme support d'activité car on peut penser qu'un potentiel obtenu avec une épreuve de section de volumes sera différent pour un même individu d'un potentiel obtenu avec des suites numériques, avec une épreuve d'histoire de France, avec une épreuve de fraisage, avec une épreuve de lancer du javelot... Dans le même ordre d'idée, il est nécessaire de préciser le paramétrage de la situation d'apprentissage, un même sujet face à une même tâche pouvant réagir différemment selon que l'aide apportée sera orale ou écrite, sur support vidéo ou informatique, de type mode opératoire ou résolution d'un conflit socio-cognitif...

Enfin, et c'est probablement le facteur le plus important, il faut absolument préciser l'état des dispositions psychologiques du sujet le jour de l'évaluation. Y avait-il un enjeu externe réellement partagé ? Les conditions matérielles de la passation furent-elles satisfaisantes ? Quelle fut l'orientation cognitive du sujet par rapport à la tâche ? Quelle fut l'importance de l'investissement énergétique du sujet dans l'épreuve ? Quelle disponibilité mentale, quelle forme physique ce jour-là ?... C'est en accompagnant le résultat chiffré de la définition des différents contextes de son élaboration que

l'évaluation prendra du sens à la fois pour l'examineur et pour l'examiné.

Bien entendu, on pourrait dire qu'une contextualisation excessive de l'évaluation ne la rende inutile dans la mesure où il deviendrait impossible de procéder à une quelconque généralisation de ses conclusions. Ce raisonnement ne tient pas ; relativiser les conditions de la mesure autorise la relativisation de la généralisation des conclusions. Il y aura possibilité de généralisation, qu'il conviendra de vérifier empiriquement, à des classes de situations homomorphes à la situation d'évaluation. Si l'épreuve porte sur le raisonnement inductif, je peux prendre le risque de prévoir qu'un sujet ayant réussi possède les capacités nécessaires pour traiter d'autres situations faisant appel à ce type de raisonnement. Si l'apprentissage s'appuie sur la transmission de stratégies métacognitives, je peux prendre le risque de prévoir qu'un sujet ayant augmenté sa performance de manière significative suite à cet apprentissage pourra réussir dans des situations formatives faisant appel à ce type de stratégie. Si le sujet, le jour de la passation était perturbé par le carnet de notes que son fils lui a remis la veille, alors je ne prendrais aucun risque de généraliser quoi que ce soit d'autre que... les effets du stress.

Cette problématique du degré de généralisation à des situations à venir des conclusions tirées d'une mise en situation locale et ponctuelle rejoint la question du pronostic. Tandis que dans le cadre de l'évaluation statique on projette des acquis selon le principe que les premiers seront toujours les premiers, dans le cadre de l'évaluation dynamique on projette la différence entre deux niveaux de performance, différence induite par une mise en situation d'apprentissage.

Ce n'est pas la même logique, puisque dans le premier cas on s'appuie sur un invariant de niveau, alors que dans le second on s'appuie sur un invariant de changement de niveau. Il ne s'agit pas de dire qu'une forme est préférable à l'autre dans l'absolu, mais de montrer qu'elle est la mieux appropriée en fonction des situations et des objectifs poursuivis. Par exemple, pour prévoir une efficacité immédiate, l'évaluation statique peut sembler la meilleure solution, mais pour prévoir la probabilité

de réussite à une formation, l'évaluation dynamique devrait progressivement s'imposer. Des recherches dans ce sens sont menées principalement auprès de populations d'enfants.

Note : Le modèle de la zone de proche développement fut élaboré par Vygotski sur la base de l'utilisation de tests d'intelligence. Rien n'empêche a priori d'envisager expérimenter les paradigmes de l'évaluation dynamique en utilisant des tests pédagométriques, des épreuves de connaissances scolaires et/ou professionnelles.

Chapitre 8

L'évaluation dynamique du potentiel d'apprentissage de l'enfant

Les études ayant pour objet l'évaluation dynamique de l'éducabilité de l'enfant se sont vraiment développées depuis les années soixante. Ces différents travaux restent plutôt l'apanage des chercheurs universitaires et n'influencent encore que très faiblement les pratiques des psychologues et encore moins celles des enseignants et des formateurs. La présentation exhaustive de ces recherches est rendue impossible du fait de leur nombre. C'est pourquoi la distinction de trois types de finalités proposée par Büchel et Paour, - analyse des processus cognitifs, mesure du potentiel d'apprentissage, paramétrage d'une remédiation - (Büchel et Paour, 1990), servira de support organisateur aux brefs exposés de quelques études.

Il peut paraître outré de proposer un chapitre relatif à l'évaluation dynamique du potentiel d'apprentissage de l'enfant alors que l'ensemble de l'ouvrage est consacré à l'évaluation de celui de l'adulte. Ce débordement du cadre induit par le titre de l'ouvrage se justifie de trois points de vue. Historiquement, le modèle de l'évaluation dynamique fut élaboré dans le cadre de la psychologie de l'enfant. La somme d'études réalisée auprès des enfants est nettement plus importante que celle réalisée auprès d'adultes. En conséquence, l'expérience acquise dans le domaine de la psychologie de l'enfant peut avantageusement inspirer l'approche de l'évaluation de l'éducabilité de l'adulte. Par exemple, sur le plan méthodologique, les paradigmes méthodologiques T.A.T., A.C.T. et A.T. fournissent un modèle directement transposable d'un domaine d'application à l'autre... D'autre part, il est clair que, en s'inspirant du titre célèbre d'une collection dirigée par Zazzo "Croissance de l'enfant, genèse de l'homme", on peut retirer de la connaissance du développement

de l'enfant des principes fonctionnels qui pourront inspirer les investigations menées auprès des adultes. Certes, il ne s'agit pas de confondre la psychologie de l'enfant, être en évolution, et celle de l'adulte, être en involution. Le passage ne peut se faire directement sous peine de placages alimentant rapidement une attitude d'autant plus dogmatique que les faits y résisteront. Ainsi a-t-on pu calculer des quotients intellectuels pour des adultes, ce qui représente une ineptie majeure dans la mesure où le Q.I. mesure une vitesse de développement, procès qui n'a plus de sens chez l'adulte.

Pour illustrer la première finalité, finalité de recherche plutôt spécifique à l'enseignant-chercheur universitaire soucieux principalement d'élaborer des connaissances dites fondamentales c'est-à-dire dégagées de toutes contraintes de répondre à une demande sociale, le texte de synthèse intitulé "Une méthode d'apprentissage destinée à analyser les relations entre développement et fonctionnement cognitif" a été retenu (Orsini-Bouichou, Hurtig, Paour, Planche, 1990). Les études rapportées par les auteurs concernent la comparaison de deux populations, l'une d'enfants précoces et l'autre d'enfants retardés. Ces deux populations se situent au même niveau de développement, niveau évalué à l'aide de tests psychométriques et/ou par observation directe d'indicateurs de développement.

Ces études montrent que les comportements individuels se différencient sur les processus de régulation des phases initiales et terminales des conduites. La mise en oeuvre de réactions anticipatrices, la formulation de règles, la pratique d'auto-corrections... sont essentiellement observées chez les enfants précoces. Ainsi, la différence des performances entre les enfants serait-elle imputable non pas tant au niveau du développement atteint considéré pour lui-même qu'au type de relations qui s'établissent entre les outils cognitifs que ce niveau met à disposition de l'enfant et sa manière de les utiliser. Ces auteurs réactualisent la conception que Binet énonçait déjà en 1909 dans son affirmation de l'éducabilité de l'intelligence : "Ce qui importe pour se conduire de manière intelligente, ce n'est pas tant la force des facultés que la manière dont on s'en sert,

c'est-à-dire l'art de l'intelligence, et que cet art doit nécessairement s'affiner avec l'exercice." (Binet, 1909)

Dans une vingtaine d'expériences, l'apprentissage a été utilisé comme méthode de contrôle ou d'analyse des relations entre niveau mental et aspects fonctionnels des conduites. Par exemple, Paour montre que les enfants retardés présentent un "état de sous-fonctionnement" (Paour, 1988) qui se caractérise par une utilisation quasi stéréotypée des outils cognitifs dont ils disposent. L'apprentissage de l'utilisation de ces outils cognitifs à des classes de problèmes de plus en plus variés suivi d'entraînements préparatoires à l'accès de l'étape suivante du développement facilite l'accès effectif à cette étape par un processus généralisé à différents domaines. Ce franchissement, observé ou provoqué, d'une étape du développement s'accompagne d'une phase «...d'essais tâtonnants signalés par une grande variabilité de réponses". (id.) D'une manière générale, cette méthodologie d'intervention plaçant les sujets dans des situations d'apprentissage montre que "...le fonctionnement se présente comme un relais et un chaînon indispensable, au cours du développement". (id.)

Cette conclusion va dans le sens de l'évaluation dynamique de l'intelligence de l'enfant, et aussi de l'adulte : c'est bien en observant le fonctionnement d'un sujet adulte devant répondre à des contraintes d'apprentissage que l'on se donne une chance de pouvoir conclure quant à son potentiel d'apprentissage. Ce potentiel d'éducabilité de l'adulte est vu alors comme manière d'optimiser l'utilisation de ses capacités opératoires en s'aménageant une zone de proche apprentissage, le terme de zone de proche développement pouvant être préservé pour le domaine de la psychologie de l'enfance.

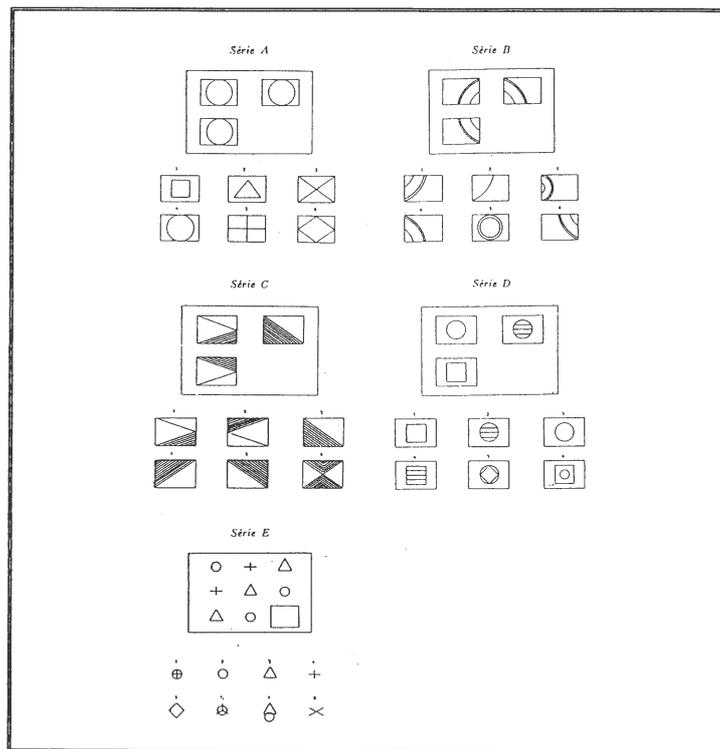
La seconde finalité recouvre le souci de la mesure du potentiel d'apprentissage. Elle se rapproche des problématiques du champ psychosocial de l'éducation, de la rééducation, du pronostic et du diagnostic individuel ou de groupes. Dans ce cadre, la mise au point de tests psychométriques pour l'évaluation dynamique de l'intelligence prend une place centrale. Les travaux de Hurtig menés autour des années soixante se situent à la charnière entre la première finalité de

recherche épistémologique et la seconde finalité qui recouvre le souci de la mesure du potentiel d'apprentissage vue comme réponse technique à une demande sociale émanant du champ de l'entreprise, de la formation, de la rééducation. Hurtig cherche à étendre les connaissances théoriques sur le développement des capacités d'apprentissage de l'enfant en gardant à l'esprit la pratique de l'examen psychologique qu'il souhaite voir élargi à l'utilisation d'épreuves standardisées d'apprentissage.

Hurtig s'inspire des anticipations de Aebli qui dérive des découvertes du constructivisme piagétien une didactique psychologique (Aebli, 1951) destinée à optimiser l'utilisation des structures déjà présentes chez l'enfant, ou d'en hâter l'apparition. Il prend en compte également les propositions de Rey et celles d'Ombredane. A ce dernier il emprunte la technique consistant à proposer comme aides à l'apprentissage des explications verbales formulées par des sujets ayant réussi la tâche, mettant en place, pour ainsi dire, un "télé-conflit-socio-cognitif". Il énonce clairement le risque pris que ces explications ne conviennent pas forcément à ceux des enfants qui ne réussissent pas la tâche. Toutefois, il justifie ses choix en termes d'"épreuves mimétiques", en ce sens que les explications verbales données se rapprochent d'une situation scolaire : "On se trouve donc dans une situation qui est, dans une certaine mesure, mimétique de la situation scolaire ; celle dans laquelle l'instituteur donne une explication au tableau, l'enfant devant par la suite, après avoir compris cette explication, l'utiliser sur un ensemble de problèmes ayant la même structure formelle." (Hurtig, 1967)

Parmi ces différents travaux, Hurtig publie une "Etude expérimentale des possibilités d'apprentissage intellectuel d'enfants débiles et d'enfants normaux" (Hurtig, 1960). Dans un premier temps, il compare les résultats de deux types d'apprentissage donnés à deux groupes de vingt élèves de différents cours élémentaires. La tâche est adaptée du test *Progressive Matrices* de Penrose et Raven, suite à une première expérience clinique de ce test ayant amené l'auteur à s'interroger à propos des effets de l'explication de la tâche sur la performance (Hurtig, 1954). Cinq séries précédées d'un

exemple sont utilisées (voir document n°16), chaque série étant homogène du point de vue de la nature des problèmes posés.



DOCUMENT N°16: Cinq types de problèmes (Hurtig, 1960)

Avec l'aimable autorisation de la revue Enfance.

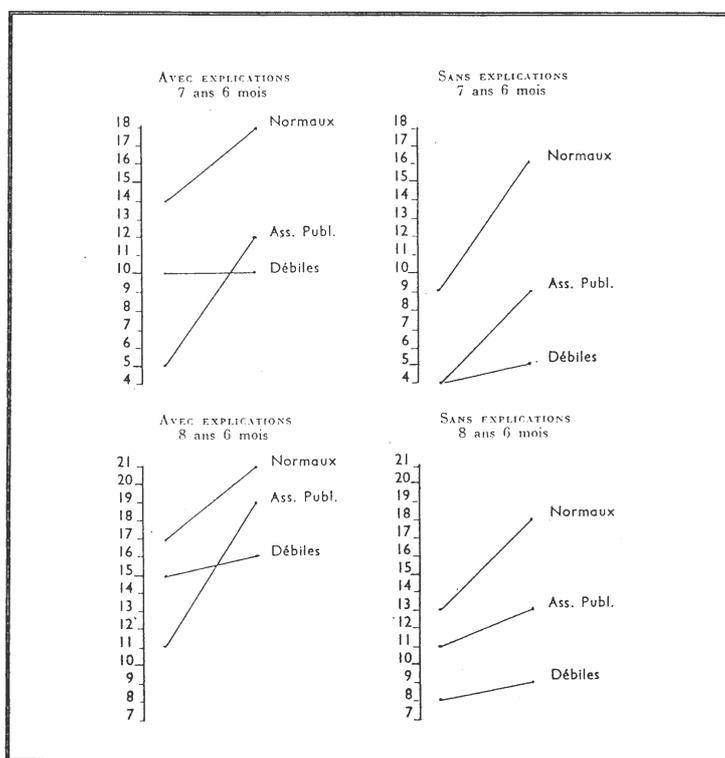
L'épreuve est donnée à quatre reprises à chacun des sujets, chaque application étant séparée d'un intervalle d'une semaine. Au premier groupe, l'épreuve est donnée à chaque fois comme une épreuve statique, simplement précédée des consignes précisant le mode de réponse attendu. A l'autre groupe, l'exemple est utilisé comme support pour expliquer le type de raisonnement impliqué dans la série correspondante (série A : dessins identiques, série B : symétrie par-rapport au

centre, série C : symétrie par rapport à un des axes, série D : raisonnement par analogie, série E : combinaison de trois éléments selon un ou plusieurs critères). Le point de repère initial est donné par une passation unique à un troisième groupe. La comparaison des scores obtenus à la première et à la quatrième passation évalue la "marge d'éducabilité" des groupes selon la modalité d'apprentissage. Les résultats vont dans le sens de l'hypothèse de départ : "Permettant à chacun d'approcher de la limite de ses "vraies possibilités", [l'apprentissage] aura un effet égalisateur : la dispersion des résultats doit diminuer, le classement des sujets doit être bouleversé par l'apprentissage." (Hurtig, 1960) Par exemple, la constance des rangs exprimée par un calcul de corrélation, est d'environ 0,70 dans la modalité sans explication et d'environ 0,30 dans la modalité avec explication. D'autre part, l'épreuve expliquée (à sa quatrième application) discrimine mieux les enfants selon leur âge et présente le meilleur pourcentage d'accord avec le classement scolaire.

Hurtig retrouve donc, avec trente ans de décalage, les conclusions de Vygotski selon lesquelles les indicateurs fournis par l'évaluation dynamique permettent de prévoir la qualité du développement cognitif et des apprentissages scolaires de meilleure manière que les épreuves statiques. Les travaux ultérieurs ne feront que confirmer la direction de ces conclusions. La répétition de la passation guidée amène donc un "effet de rattrapage". Hurtig étudie cet effet de rattrapage d'abord sur des groupes d'enfants récemment confiés à l'Assistance à l'Enfance à Paris. Ces enfants, sur les épreuves classiques, présentent un retard mental important. L'expérience montre que si la procédure sans explications accroît les écarts, *a contrario*, celle avec les explications les réduit. "...on a vu la marge de variation... se restreindre lorsque les conditions d'apprentissage sont favorables, augmenter lorsqu'elles sont défavorables ... L'illusion que les tests montrent un développement intellectuel à l'abri des influences sociales, relativement stable et fixe, est cependant trop persistante pour ne pas insister sur cet effet de l'apprentissage." (Hurtig, 1960)

Hurtig montre que derrière un bas niveau de "performance spontanée" peut se cacher une large marge d'éducabilité.

Enfin, le même schéma d'expérience est appliqué à des groupes de "débiles mentaux", fréquentant des classes de pré-apprentissage, âgés de 13 à 17 ans, et d'âges mentaux comparables aux âges chronologiques des enfants normaux. Quel que soit le type de passation, il y a stagnation de la performance au fil des répétitions. Le document n°17 regroupe les profils de l'ensemble des groupes.



DOCUMENT N°17: Courbes des moyennes d'apprentissage des divers groupes (Hurtig, 1960)

Avec l'aimable autorisation de la revue Enfance.

On peut rapprocher ces profils de ceux des sujets "performants", "gagnants" et "non-gagnants" proposés ultérieurement par Budoff (voir infra). Et déjà, Hurtig préconise l'utilisation d'une évaluation avec apprentissage pour distinguer entre "les cas de débilité vraie", et les cas où un retard mental masque parfois "...des déficits d'ordre social-déficits fortement atténués par un apprentissage spécifique." (id.)

Dans sa thèse, Hurtig va développer cette approche. Il stabilise son adaptation des matrices sous le sigle *Matrix Educabilité*. Il choisit le paradigme T.A.T., l'apprentissage conservant la forme d'explications verbales données à tous les sujets quels que soient les niveaux de réussite spontanée. Par cette procédure T.A.T., on retrouve un effet de reclassement des performances important suite à l'apprentissage : la corrélation passe de 0,91 pour un groupe test-retest à 0,72 pour un groupe T.A.T.

Par rapport à l'étude précédente, la variable socioculturelle est introduite sur la base de trois catégories tenant compte de la profession du père : professions manuelles, professions intermédiaires, professions libérales et cadres. L'effet de reclassement suite à l'apprentissage est très net : suite aux explications, l'influence du milieu social devient beaucoup moins sensible. "Les explications paraissent avoir une valeur de rattrapage ; le handicap social est comblé grâce à cette technique, les infériorités constatées [au pré-test], étant largement diminuées, voire supprimées." (Hurtig, 1966) Là encore, l'auteur met en relation les scores obtenus et les résultats scolaires et il confirme son résultat précédent : la corrélation entre niveau scolaire et degré de réussite spontanée est de 0,53 ; elle passe à 0,68 entre niveau scolaire et degré de réussite après l'apprentissage, la différence des corrélations étant significative à $P = 0,01$.

Enfin, une analyse fine des changements d'âge en âge, de 7 à 11 ans, permet à Hurtig de conclure dans le sens d'une zone de proche développement : "Les explications ne sont donc efficaces que pendant la période précédant immédiatement l'acquisition spontanée de ce comportement." (Hurtig, 1966)

En conclusion de ses différents travaux, Hurtig revient à la question de l'utilisation des épreuves d'apprentissage malgré les différents problèmes qu'elles posent sur le plan théorique et sur le plan de la mesure. Il insiste sur la distinction à réaliser et à maintenir entre les différents moments de l'évaluation dynamique qui renvoient à des changements de nature de la tâche : réalisation spontanée avec appui sur les expériences et acquis antérieurs, réalisation suite à des explications qui joueraient un rôle de rattrapage des lacunes liées à l'appartenance à un milieu socialement et culturellement défavorisé. D'autre part, les courbes d'apprentissage individuel ne peuvent pas être interprétées de manière simple : elles peuvent recouvrir des réalités psychologiques différentes, selon le niveau de réussite spontanée, selon le degré de familiarité avec la tâche, selon la distance au plafond que le sujet pourrait effectivement atteindre en poursuivant l'expérience suffisamment longtemps... et seule une expérience clinique permettra au praticien de progresser dans leur interprétation.

Sur le plan de la mesure, Hurtig soulève la question de l'expression de scores différentiels : simple différence des scores avant ou après apprentissage ? normalisation de la différence des scores normalisés des pré et post tests ? calcul d'un score d'amélioration relative au niveau de départ ? calcul d'un indice d'amélioration résiduelle ? Malgré ces incertitudes soulevées, Hurtig conclut à l'intérêt majeur de l'utilisation des épreuves d'évaluation dynamique : "La prise de conscience de ces problèmes devrait marquer, au contraire, le départ de recherches permettant de dépasser l'examen statique des acquisitions, et de faire avancer la psychométrie dans la voie de sa dynamisation.

Dans un examen psychologique comportant un apprentissage, il faut simultanément prendre en considération les diverses variables enregistrées : point de départ, point d'arrivée, pente de la courbe d'apprentissage, taux d'apprentissage ; ces diverses variables et leurs combinaisons devront être étalonnées. L'interprétation clinique ne pourra se faire qu'en fonction des antécédents du sujet, de son âge, ainsi que des liaisons avec les critères externes constatées entre ces

variables à divers niveaux d'apprentissage. Et l'utilisation de telles variables, par la réflexion, l'expérience et l'expérimentation qu'elle entraînerait, permettrait de dépasser les interprétations fixistes du simple constat ; habituant le psychologue aux pratiques d'une intervention contrôlée, elle lui permettrait de ne pas tomber non plus dans la mystique contraire de l'"éducabilité". (Hurtig, 1967)

Cette invitation à poursuivre dans le chemin tracé n'a guère été entendue des psychologues de France. Mary et Mariel sont probablement les seuls à avoir poursuivi dans cette voie, en travaillant sur la dynamisation de la passation des cubes de Kohs pour des enfants de quatre à six ans (Mary, Mariel, 1982). Ils retrouvent à cette occasion les grandes conclusions de Hurtig : "Dans l'ensemble, ce sont les enfants de milieu défavorisé qui progressent le plus. Tout se passe comme si ces enfants avaient des possibilités de raisonnement non mises en valeur, non exploitées par suite de la carence de leur milieu d'origine. Il leur faut une aide personnalisée pour pouvoir les actualiser." (Mary, Mariel, 1978)

Il faut se déplacer aux Etats-Unis pour prendre connaissance des travaux plus nombreux menés par Budoff et ses collaborateurs depuis 1964. Une synthèse de ces travaux a été réalisée en français par Ionescu. Budoff définit le potentiel d'apprentissage comme : "...l'habilité à apprendre et à tirer profit d'une expérience adéquate." (in Ionescu, 1985) L'intelligence est vue ici sous l'angle des capacités d'apprentissage et des capacités à traiter cognitivement l'expérience pour en tirer des acquis stables et transférables. Budoff a principalement utilisé comme épreuves les cubes de Kohs, les matrices de Raven et un test de série d'images à compléter, selon le paradigme T.A.T. L'évaluation dynamique est introduite par Budoff essentiellement pour contrebalancer les biais culturels susceptibles d'altérer la performance des enfants issus de milieux défavorisés ou présentant des handicaps. L'objectif est donc de faire la part entre les déficiences mentales véritables et les pseudo-déficiences de nature culturelle, ou sociale, ou éducative. En effet, Budoff montre que des enfants classés déficients mentaux sur la base d'un test de Q.I. sont capables

d'améliorer considérablement leur performance suite à l'entraînement, alors que l'entraînement se montrerait sans effet dans les cas où les enfants présentent des lésions cérébrales prouvées.

Budoff et ses collaborateurs apportent également un ensemble de résultats concernant la validité des indices de potentiel d'apprentissage obtenus avec les cubes de Kohs, dans des études menées avec des enfants reconnus déficients mentaux. Des liaisons significatives sont établies avec les résultats obtenus à divers tests d'intelligence classiques et, d'une manière générale, la valeur prédictive du potentiel d'apprentissage est supérieure à celle du Q.I. pour ce qui concerne tout aussi bien les connaissances scolaires, que les connaissances pratiques professionnelles, que l'adaptation sociale à moyen terme.

Selon Guthke, le gain de validité des épreuves dynamiques se confirme essentiellement avec des élèves en difficulté. Par exemple, chez des enfants à faible rendement préscolaire les corrélations comparées entre les résultats du test de Raven passé de manière statique ou dynamique et les critères de réussite scolaire donne l'avantage à la passation dynamique pour tous les critères utilisés (0,04 vs 0,55 pour la moyenne scolaire). Guthke souligne également la nécessité d'un changement de contexte socio-éducatif si l'on veut voir s'amplifier l'intérêt et la supériorité du test d'apprentissage : "Wiedl et Herrig (1978) ont constaté par exemple que la validité pronostique des tests d'apprentissage s'avérait nettement supérieure à celle des tests conventionnels si les sujets examinés reçoivent par la suite un enseignement individualisé. Cette supériorité des tests d'apprentissage est moins prononcée s'il s'agit de pronostiquer la réussite scolaire dans un enseignement conventionnel en groupe." (Guthke, 1990)

De fait, l'évaluation dynamique répond mieux aux besoins d'une évaluation psychopédagogique lisible par l'enseignant et qui puisse se prolonger concrètement au niveau de l'action éducative. En sus d'une amélioration de la qualité de la mesure, le principal intérêt de l'évaluation dynamique réside dans la qualité et la finesse d'un diagnostic individuel des

potentialités d'apprentissage, des stratégies cognitives et métacognitives, de la nature et des sources d'erreur... Ces informations qualitatives peuvent se traduire en directions d'actions d'enseignement, de formation, de remédiation...

Pour aller plus loin que le simple constat de corrélations statistiques, Budoff a travaillé à donner de la consistance psychologique à l'évaluation du potentiel d'apprentissage. Il propose une catégorisation des sujets en trois groupes selon leur profil de résultats : les sujets au score élevé réussissent d'emblée le test, les sujets gagnants augmentent leur performance suite à l'apprentissage, les sujets non-gagnants ne tirent aucun profit de l'aide apportée. Derrière ces patrons de réponses à la mise en situation d'apprentissage, Budoff a cherché à différencier les caractéristiques du fonctionnement psychologique des sujets. Des différentes études menées, il ressort que les sujets aux scores élevés et les sujets gagnants se comportent *grosso modo* de la même manière, les non-gagnants apparaissant comme désavantagés par rapport à l'ensemble des critères de comparaison utilisés.

Les non-gagnants font preuve d'une plus grande rigidité cognitive ; dans des situations expérimentales qui visent à étudier les effets d'un stress léger, les sujets aux scores élevés et les sujets gagnants parviennent à surmonter les effets du stress alors que la performance des non-gagnants subit une forte détérioration. A des épreuves explorant la dimension frustration, les gagnants produisent des réponses actives et positives et les non-gagnants des réponses passives. Dans le cadre d'approches auto-évaluatives, les non-gagnants se perçoivent comme plutôt inférieurs à leurs pairs, *a contrario* des autres sujets. Cette image de soi plutôt détériorée des non-gagnants se trouve confirmée par d'autres types d'investigations dans lesquels leurs réponses révèlent l'inadaptation, l'anxiété névrotique, et autres formes d'affections de nature socio-affectives. L'étude du niveau d'aspiration différencie également les extrêmes. Les non-gagnants ont tendance à se fixer des buts irréalistes sans parvenir à argumenter leur choix. *A contrario*, les autres sujets identifient mieux l'école comme moyen de parvenir à la situation professionnelle visée ; ils établissent plus

facilement la relation de cause à effet entre le travail fourni et le degré de réussite obtenu.

A travers l'ensemble de ces études, il se confirme que l'évaluation métrologique du potentiel d'apprentissage renvoie à des fonctionnements cognitifs, affectifs et sociaux fortement différenciés. Si Brown et Ferrara critiquent l'aspect un peu flou des descripteurs des effets de l'entraînement utilisés par Budoff, elles retiennent la pertinence de la distinction entre les gagnants et les non- gagnants (Brown et Ferrara, 1985). Suite à une visite à l'Institut de Défectologie de Moscou en 1978, Brown prit connaissance de la batterie d'épreuves développée par Lubovskii, Rozanova et Egorova dans une perspective d'analyse clinique du diagnostic des difficultés d'apprentissage, les tests standardisés ayant mauvaise presse dans l'ex-U.R.S.S. L'objectif des professeurs de Moscou était de différencier les enfants véritablement retardés mentaux de ceux qui ne présentaient que des difficultés d'apprentissage, l'hypothèse de base étant que placés en situation d'évaluation dynamique, les pseudo-déficients auraient besoin de moins d'aides que les vrais déficients pour parvenir à la réussite. D'autre part, ils seraient mieux capables de transférer les acquis de leur bref apprentissage en situation d'évaluation à des variations apportées dans la tâche, voire à des classes de problèmes.

Ces hypothèses furent empiriquement démontrées par Egorova et Lubovskii en comparant les comportements cognitifs de groupes d'enfants normaux et déficients. Sur le plan méthodologique, les professeurs de Moscou utilisent le paradigme A.C.T. : l'enfant est confronté à un item à résoudre et, en cas d'échec, l'examineur apporte des aides progressives facilitant l'accès à la solution. Ensuite, il propose une version parallèle de l'item, version à résoudre de manière autonome. Le nombre d'items réussis de manière autonome à la première présentation apporte une information sur le niveau de développement de l'enfant relativement comparable à celle donnée par un test de Q.I. L'étendue de la zone de proche développement est comprise comme l'inverse du nombre d'aides à apporter pour conduire à la réussite, ou par le niveau atteint avec le recours aux aides. Enfin, la capacité de transfert est

matérialisée par le nombre d'items parallèles réussis de manière autonome.

S'inspirant de l'approche clinique des professeurs soviétiques, Brown et Ferrera entreprirent une série d'études pour évaluer de manière expérimentale la validité de cette approche et pour étudier la faisabilité d'une standardisation des aides apportées. Ces auteurs utilisèrent des épreuves de séries de lettres à compléter construites selon des principes tirés des analyses de ce type de tâche menées par Kobovsky et Simon. Trois règles alphabétiques interviennent : on demande la lettre identique C -> C ; on demande la lettre qui suit C -> D ; on demande la lettre qui précède C -> B. Aux règles alphabétiques s'ajoute la période qui renvoie au nombre de suites imbriquées : par exemple P A O A N A M A période 2 pour la mise en relation des lettres. La procédure de type A.C.T. introduit des aides progressives pour amener à la réussite sur les items non résolus spontanément. Des séries de problèmes sont ensuite données qui éloignent de plus en plus de la situation de l'apprentissage initial selon un gradient de transfert. L'épreuve fut utilisée avec des élèves des troisième et cinquième grade. Les auteurs ont mis en relation le Q.I. des élèves avec leur rapidité d'apprentissage et l'étendue du transfert. L'analyse du graphe des résultats montre un recouvrement partiel des trois variables. Il en résulte qu'une fois sur deux la connaissance du seul Q.I. ne permet pas de prédire la vitesse de l'apprentissage ou l'étendue du transfert. De fait, un Q.I. identique peut dissimuler des profils d'apprentissage sensiblement différents, au moins cinq d'après nos auteurs qui élargissent ainsi la typologie proposée par Budoff. Les "lents" ont un Q.I. bas, ils apprennent lentement et transfèrent étroitement ; les "rapides" ont un Q.I. élevé, ils apprennent vite et transfèrent largement ; les "limités au contexte" apprennent vite et transfèrent étroitement ; les "réflexifs" apprennent lentement et transfèrent largement ; les "scores élevés", en référence à Budoff, ont un Q.I. élevé, apprennent vite et transfèrent largement.

Afin de tester la stabilité de ces profils d'apprentissage, les auteurs ont fait passer une seconde épreuve à vingt-deux sujets du groupe précédent. Cette épreuve, dérivée des matrices

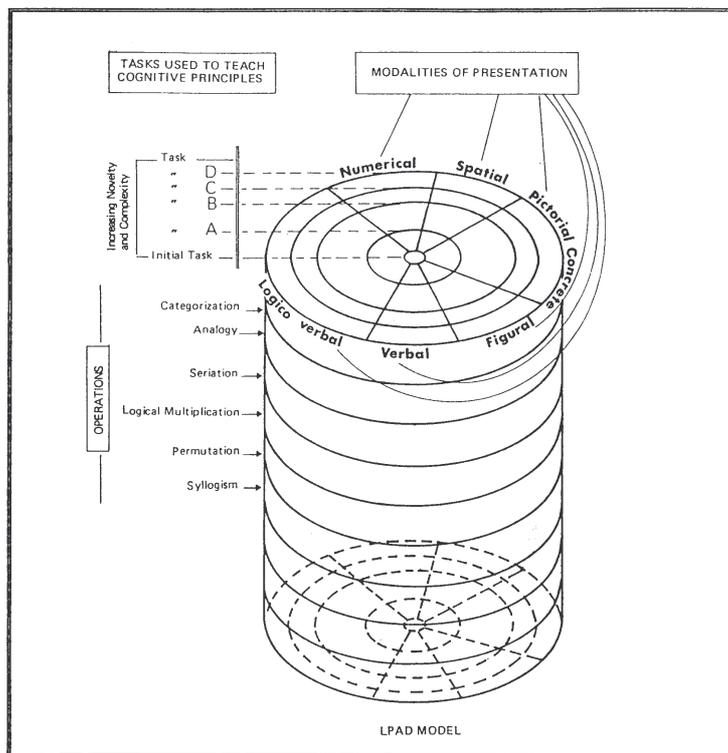
de Raven, est structurée de la même manière que la précédente, afin de dégager un gradient de transfert de l'apprentissage. Par exemple, on apprend au sujet à résoudre une première matrice à l'aide d'une opération de soustraction ; on lui apprend ensuite à résoudre une seconde matrice nécessitant la mise en oeuvre d'une opération de rotation ; enfin, on lui propose de résoudre, sans aide et par transfert d'apprentissage, une troisième matrice impliquant la mise en oeuvre combinée des deux opérations apprises (soustraction et rotation). Les auteurs affirment que soixante-seize pour cent des enfants conservent le même profil d'éducabilité dans les deux situations et dix-neuf pour cent ont légèrement fluctué du point de vue de l'efficacité de l'apprentissage (0,66 de corrélation). La capacité de transfert semble plus liée au contexte dans la mesure où le taux de concordance chute à quarante-trois pour cent et la corrélation à 0,39. Enfin, la combinaison des deux variables donne une invariance de profil pour trente-huit pour cent des sujets.

Les résultats de toutes ces études, russes, françaises ou anglo-saxonnes vont *grosso modo* dans le même sens : elles montrent, depuis Vygotski, que l'évaluation dynamique permet de distinguer sur un plan fonctionnel des individus qui se situent au même niveau sur les échelles psychométriques des épreuves statiques. Cette ouverture sur ces aspects fonctionnels est de toute première importance car elle ouvre en grand la perspective des interventions de remédiation en termes de majoration du potentiel d'éducabilité, tant dans le domaine de la rééducation orthopédagogique que dans celui de la psychopédagogie. Toutefois, des questions restent largement ouvertes. La première question concerne le type de public : principalement utilisé pour différencier la "vraie déficience mentale" de l'inhibition socio-affective des capacités cognitives, ce type d'évaluation apporte-t-il un plus pour l'évaluation des populations dites "normales" ? La seconde question est relative à la spécificité de la zone de proche développement : dans quelle mesure diffère-t-elle de manière intra-individuelle en fonction de chaque domaine de contenus ? La troisième question prolonge la précédente en termes d'études des possibilités de passage entre l'évaluation dynamique sur des

contenus relatifs aux tests et l'évaluation dynamique sur des contenus académiques ou professionnels. Élaborer ces passages représente un projet à forte plus-value heuristique en ce sens que les mêmes clés seraient opérantes pour l'établissement du diagnostic, du pronostic, du programme de la remédiation tant cognitive que scolaire ou sociale...

Dans le cadre d'une finalité de mesure, il ne faut pas oublier l'étude des qualités métrologiques des épreuves servant à l'évaluation dynamique. Dans son article de synthèse, Ionescu rapporte les études de fidélité de l'évaluation dynamique du potentiel d'apprentissage menées par Budoff et ses collaborateurs (Ionescu, Jourdan -Ionescu, 1986). Par exemple des études mettant en oeuvre des retests immédiats, à un mois et à six mois donnent des corrélations variant de 0,72 à 0,93. Ionescu lui-même a travaillé à la dynamisation de la passation du test des cubes de Kohs. Par la technique de partition pair-impair ("*split-half*"), il calcule un coefficient de fidélité égal à 0,70. Yerle a consacré une étude sur la question de la fidélité de la mesure de l'éducabilité. À partir d'un échantillon d'une centaine d'élèves de cours moyens deuxième année il utilise des épreuves de séries numériques et de matrices, conçues et/ou adaptées pour les besoins de la recherche. Le plan expérimental fait jouer deux évaluations successives du potentiel d'apprentissage, la première à l'aide des séries numériques et la seconde à l'aide des matrices. Le groupe contrôle ne passe que les pré et post tests ; le groupe pratique passe en plus des épreuves parallèles composées des items utilisés pour l'apprentissage, mais sans les aides ; le groupe expérimental bénéficie des aides sur les items des épreuves parallèles. L'auteur évalue la fidélité sur le groupe expérimental en calculant les corrélations entre les deux premiers post-tests (0,42), entre les deux seconds post-tests (0,59), entre les nombres d'aides fournies à chaque séquence d'apprentissage (0,54). Qui plus est, ces indices sont plus fortement corrélés avec le moyenne des notes scolaires que les scores obtenus au pré-test (0,70 vs 0,50).

Au travers des données proposées par Budoff, Ionescu et Yerle, on peut conclure dans le sens d'une fidélité correcte des procédures d'évaluation dynamique.



DOCUMENT N°18: Modèle du L.P.A.D. (Feuerstein, 1987)

Avec l'aimable autorisation de Reuven Feuerstein.

Dans la perspective de la troisième finalité essentiellement représentée par Feuerstein, la question de la fidélité est posée tout autrement. Feuerstein adopte et systématise une démarche équivalente à celle de Rey, avec lequel il a travaillé un certain temps, surtout lorsque Rey préconise de faire subir aux épreuves statiques, qui ne peuvent refléter que l'image d'acquisitions antérieures, des variations de

procédure, des changements dans les consignes, des répétitions... afin de se rapprocher du "processus actuel d'acquisition" (Rey, 1958).

Feuerstein présente la structuration de son dispositif d'évaluation du potentiel d'apprentissage, le L.P.A.D., sous une forme graphique (voir document n°18). Chaque tranche du cylindre représente une opération mentale et renvoie à un test spécifique. Le modèle n'est pas fini, et il est envisageable d'utiliser tous tests existants, éventuellement d'en adapter ou d'en créer de nouveaux en fonction de l'opération mentale visée. Concrètement, les épreuves les plus usitées sont les *Progressives matrices* de Raven et les *Set variations I and II* de Feuerstein, le *Labyrinthe manuel* de Rey (Test des plateaux), l'*Organisation de points* de Rey, la *Figure complexe* de Rey, le *Representational stencil design test* adapté de Grace Arthur (test de pochoirs) les *Progressions numériques* de Feuerstein, le *Test d'apprentissage de positions* de Rey, le *Test des similitudes* de Haywood, le *Test de mémoire de mots* adapté de Rey, le test *Organiser* et le *test de logique propositionnelle* de Feuerstein (Feuerstein, Rand, Haywood, Hoffman, Jensen, non daté).

Les cercles concentriques définissent les niveaux de complexité et d'abstraction pour une même opération : par exemple, passage d'une matrice à quatre éléments à une matrice à six éléments puis à une matrice à neuf éléments... Les secteurs représentent les différentes modalités de codage du contenu pour une même opération : l'inférence inductive sous-tend les matrices qui exposent le sujet à une modalité figurative, et aussi les suites de nombres présentées dans une modalité numérique

...

L'approche est individuelle et de type clinique : les scores éventuellement calculés avant et après médiation de l'examineur ne peuvent servir à comparer le sujet qu'à lui-même. Dans ces conditions, le manque de fidélité d'une épreuve devient une qualité dans la mesure où elle permet de mettre en évidence la "modifiabilité structurelle de l'individu". C'est avec juste raison que Dias pose la question de l'objectif du L.P.A.D : "Que proposent Feuerstein, Rand & Hoffman ? Est-ce la mesure

du "potentiel d'apprentissage" ou celle de la "modifiabilité cognitive" ?" (Dias, 1994)

A l'examen, il semble qu'il faille abandonner l'objectif contenu dans le titre, pour se centrer plutôt sur la recherche de l'amorce d'une modifiabilité structurelle cognitive. En effet, l'examineur médiateur clinicien se déplace dans le modèle cylindrique, il fait varier les tâches, les modalités de codage, les niveaux de difficulté... et l'ensemble des paramètres de la médiation afin de dresser l'inventaire des zones de sensibilité à la médiation, en relation avec les modalités de celle-ci, zones qui seront par la suite systématiquement élargies et généralisées par le biais d'une remédiation cognitive.

En bref, l'objectif du L.P.A.D., principalement utilisé avec des publics déficitaires, se résume dans la recherche des zones de sensibilité à la médiation sociale et à la définition des paramètres de cette médiation en fonction de chaque cas, dans le cadre de procédures d'aides non standardisées. Dans ces conditions, une passation individuelle peut prendre de cinq à vingt-cinq heures, jusqu'à ce que se produise au minimum un "micro -changement", indice d'une capacité enfouie, porte d'entrée d'un travail orthopédagogique de remédiation. Le L.P.A.D. représente donc un dispositif complexe, qui demande une connaissance approfondie de l'ensemble des tests du psychologue clinicien complétée par une habilité à se déplacer dans un système de variations, proche au fond d'une démarche de diagnostic progressif. C'est un ensemble de grand intérêt dans le domaine de l'examen clinique de l'enfant, de l'orthopédagogie, de la rééducation psycho-cognitive individuelle, afin d'affiner des diagnostics fonctionnels des possibilités d'évolution des conduites cognitives et d'amorcer un travail de remédiation avec les meilleures chances de succès. Son utilisation dans d'autres domaines, comme celui de la pédagogie, voire de la psychopédagogie, et sa mise en oeuvre par des non-professionnels de la psychologie clinique ne peut qu'apporter une série de dégradations. En effet, la formation aux tests, dans une approche clinique, nécessite une formation longue de type "bac + 5", et une pratique et une formation continues. On voit mal de quelle manière un enseignant pourrait

se former dans le cadre d'un stage court à une technique tellement sophistiquée. D'autre part, il y a le risque non négligeable d'une usure des tests du psychologue clinicien par une utilisation triviale à la fois extensive et naïve des épreuves.

Suite à cette excursion dans la littérature consacrée à l'évaluation dynamique de l'enfant, on peut en tirer quelques enseignements utiles. Essentiellement, lorsqu'on apporte une aide dans la passation d'une épreuve psychométrique, lorsqu'on dynamise le test, on obtient une amélioration des performances et un effet de reclassement, à caractère de rattrapage, qui profitent essentiellement aux enfants des milieux socialement défavorisés ; indemnes d'atteintes neurologiques ou autres troubles majeurs. Ce constat permet de distinguer entre "vrais" et "pseudo" déficients. Cette distinction recouvre des différenciations beaucoup plus larges que simplement cognitives entre ces sujets "gagnants" et "non-gagnants". D'autre part, les indices de potentiel d'apprentissage se révèlent comme meilleurs indicateurs de la réussite des apprentissages scolaires, voire de l'insertion sociale, que le simple niveau de réussite spontanée, plutôt représentatif des acquis antérieurs. Dans un souci de qualité métrologique, ces indices se montrent fidèles. Dans une perspective clinique, et dans le domaine de la déficience mentale avérée, la question de la mesure ne se poserait plus. Il est préférable, pour le psychologue clinicien, d'adopter une approche non standardisée, flexible, qui joue sur les variations de l'ensemble des paramètres de la situation évaluative, dans le but de rechercher les zones réactives à la médiation. La localisation de ces "portes d'entrées" et le dessin des "clés médiationnelles" qui en ont permis l'ouverture serviront de points d'ancrage au projet de modifiabilité cognitive de l'orthopédagogue.

Restent à voir les analogies possibles avec les travaux effectués auprès des publics d'adultes.

Chapitre 9

L'évaluation dynamique du potentiel d'apprentissage de l'adulte

Les études portant sur l'évaluation du potentiel d'apprentissage des adultes sont rares, ou bien ne font l'objet d'aucune publication. Par exemple, des psychologues et des formateurs d'adultes utilisent le L.P.A.D. de Feuerstein dans une forme de passation collective, le plus souvent en début d'application du P.E.I.F. Interrogé sur cette question, le diffuseur français de la méthode a précisé qu'à sa connaissance aucune étude ni recherche n'étaient menées à l'occasion de ces pratiques d'évaluation.

Au cours de leurs missions en ex-Congo belge, Ombredane et ses collaborateurs furent amenés à pratiquer des schémas de retest. Les premières utilisations du *Matrix-couleur* sur des populations congolaises amenèrent Ombredane à poser la question de l'appartenance culturelle : "...nous avons, dès nos premiers sondages, acquis la conviction que des tests de ce genre sont, dans une proportion qui reste à déterminer, des tests de culture autant que d'intelligence." (Ombredane, Robaye, 1953). Afin de commencer à comprendre le point de vue et la manière de penser des noirs confrontés à un test, Ombredane demanda aux sujets, à l'issue d'une première passation respectant les conditions de la consigne, d'explicitier "les chaînes de pensée" ayant conduit au résultat. "Cette explicitation avait généralement pour résultat une amélioration considérable du score." (Ombredane, Robaye, 1953). A partir de ces constats, Ombredane arrive à la conclusion qu'il serait souhaitable d'épurer le résultat du test, c'est-à-dire de pouvoir faire la part des choses entre le rendement intellectuel potentiel du sujet et les facteurs parasites qui viennent "maquiller" ce rendement. Ces facteurs parasites, essentiellement rattachés à

des différences d'appartenance culturelle, empêchent une entrée de plain-pied dans la tâche : il faudra un temps de familiarisation, d'accommodation du sujet à l'épreuve, d'où l'idée de recommencer une seconde passation.

Ombredane étudie d'abord cette question sur deux groupes de soixante élèves belges âgés de six à treize ans, dans le cadre d'une passation individuelle et en temps libre. Le premier groupe, suite à la première passation, est invité à expliciter ses choix et à corriger éventuellement ses réponses ("explicitation") ; le second groupe recommence simplement une seconde passation ("réduplication"). Dans les deux modalités, les scores progressent d'une passation à l'autre. Cette progression, globalement considérée, est indépendante de la modalité de passation, explicitation ou réduplication. Si le nombre des erreurs est équivalent, une analyse fine montre que dans la condition réduplication, la variation dans les réponses fournies est beaucoup plus grande, ainsi que le nombre de nouvelles erreurs. D'autre part, l'explicitation se montre plus efficace que la réduplication pour les items d'un degré de difficulté accessible par le sujet, c'est-à-dire pour les items situés dans une "zone d'incertitude" où réussites et échecs se superposent (zone de dispersion des résultats). Les items faciles et les items très difficiles ne sont plus sensibles au pouvoir de l'explicitation.

Cette zone d'incertitude optimale, qui rappelle la zone de proche développement de Vygotski, représente donc la zone des items à explorer sur un plan qualitatif. A partir de ces observations, Ombredane conclut sur le rôle de l'explicitation : "La découverte se fait dans le moment de la perception, l'explicitation accroche, fixe, contrôle, agit comme les épingles dans les essayages du tailleur." (Ombredane, Robaye, 1953). Il va réinjecter le résultat de cette recherche dans le cadre d'une mission de sélection de conducteurs d'engins lourds évoluant dans des exploitations de diamants de l'ex-Congo belge. Le *Matrix-couleur* est inclus dans une batterie d'une dizaine d'épreuves dont les modalités de passation furent orientées dans le sens d'une évaluation de l'éducabilité des sujets : passation multiple d'une même épreuve, démonstrations préalables,

progression du niveau de difficulté... Cette orientation du testage doit éviter l'"effet piège" qui place le sujet dans une situation plus ou moins connue et dont il doit trouver l'issue par ses propres moyens. Cette issue dépend largement des capacités acquises antérieurement et transférées dans la situation à résoudre. En situation de formation, la question prend une autre dimension : dans quelle mesure le sujet apprenant pourra-t-il intégrer les apprentissages transmis ? "Au testing de l'aptitude il y a intérêt à substituer le testing de l'éducabilité. Le problème n'est pas de savoir si un Musalampasu est venu au monde avec l'aptitude à conduire un engin lourd, mais de déterminer dans quelle mesure il sera capable de profiter des enseignements qu'on lui donnera à l'école des conducteurs d'engins lourds. Ce principe est au demeurant valable pour la psychotechnique en milieu blanc." (Ombredane, Robaye, Plumail, 1956).

La technique d'application du *Matrix-couleur* fut fixée ainsi : démonstration préalable au moyen de cinq planches en fibre permettant la manipulation et l'encastrement des pièces, passation individuelle, le sujet pointant la réponse de son choix à l'aide d'une baguette et l'examineur notant la réponse, deux réduplications consécutives suivent la première passation. Le *Matrix-couleur* fut donné à six cent quatre-vingt treize sujets choisis en fonction de l'âge (6 - 7 ans ; 8 - 9 ans ; 10 - 11 ans ; 12 - 13 ans ; 14 - 16 ans ; 17 - 30 ans) et du degré d'instruction (aucune scolarité ; 1 - 2 ans ; 3 - 4 ans ; 5 ans). Le groupe des 17 - 30 ans est subdivisé en "brousse" et "Bakwanga", ces derniers étant amenés à entrer en formation de conducteurs. Différents constats ressortent de l'analyse des résultats constitués par la moyenne des trois notes attribuées aux trois passations successives de l'épreuve. La performance est fonction de l'âge des sujets et de leur degré de scolarité. L'influence du degré de scolarité se fait sentir essentiellement à partir de la troisième année, et il y a diminution de la performance chez ceux qui ont quitté l'école il y a deux ans au moins. Quand on compare les scores obtenus à chacune des trois passations, on s'aperçoit que l'amélioration des scores dépend et de l'âge et du degré de scolarité, mais non du score à la première passation. Les trois passations se justifient par la

grande diversité des effets de reclassement. Les facteurs motivationnels semblent jouer un grand rôle dans les variations des progressions. Au fur et à mesure des passations, les patrons de réponses se rapprochent d'un schéma guttmannien principalement chez les sujets ayant le degré de scolarité le plus élevé et la motivation la plus forte : les réponses réussies deviennent de plus en plus homogènes par rapport au niveau de difficulté des items.

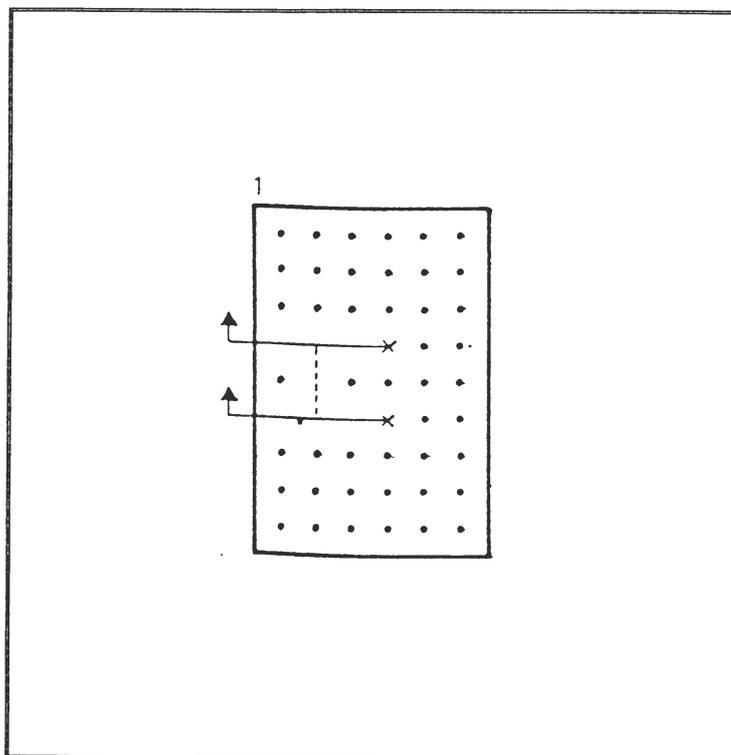
Dans une étude complémentaire, Robaye et Robaye vont plus loin dans l'analyse en croisant les facteurs de niveau de réussite à l'épreuve ("groupe inférieur" versus "groupe supérieur") et de niveau de difficulté relative de chaque item ("items trop faciles" versus "items discriminatifs" versus "items trop difficiles"). Sur les items faciles, le groupe inférieur corrige deux fois plus d'items que le groupe supérieur (0,83 vs 0,43), ce qui serait lié à "...des facteurs parasites tels que l'étrangeté de la tâche, la lenteur de la mise en marche...". (Robaye, Robaye, 1956). Pour les items discriminatifs, la tendance se renverse au profit des sujets du groupe supérieur (3,32 vs 2,17) qui corrigeraient des erreurs liées à l'inattention, ou à des variations de l'intérêt ou de la motivation. Enfin, pour le troisième groupe d'items, le groupe inférieur ne corrige plus et le groupe supérieur très peu (0 vs 0,37). En résumé, "...la simple reduplication du matrix supprime les erreurs commises pour des motifs étrangers au niveau réel atteint par les sujets ; la reduplication élimine des fautes dues à des facteurs parasites, le sujet ne "découvre" pas de nouveaux types de solution... De plus, comme les résultats sont moins affectés par ces facteurs parasites, la fidélité s'en trouve nettement améliorée." (Robaye, Robaye, 1956).

Ombredane a mené une étude de validité concurrente sur un effectif de soixante-dix huit conducteurs classés selon des critères de réussite professionnelle par un moniteur noir et un instructeur blanc (méthode de comparaison par paires). La corrélation a été calculée entre le classement professionnel et chacun des classements au *Matrix-couleur* : de la première à la troisième passation, le coefficient prend les valeurs 0,38 ; 0,48 ; 0,51. Cette étude confirme qu'un test de facteur G ne se

comporte pas de manière indépendante des aspects culturels : le résultat obtenu dépend du degré de scolarité de l'individu, c'est-à-dire de la nature de ses acquis antérieurs. Une passation multiple, même sans explications, provoque des effets de reclassement qui font penser que les résultats d'une troisième passation refléteraient mieux l'éducabilité des sujets que ceux de la première par l'élimination des effets des facteurs parasites. Toutefois, la nature de la tâche est un élément à prendre en considération. Dans une étude ultérieure, Ombredane utilise un test d'encastrement de la batterie d'habileté mécanique de Minnesota pour évaluer les capacités de discrimination spatiale d'un groupe de trois cent quatre-vingt neuf sujets noirs de l'ex-Congo belge (Ombredane, Robaye, Robaye, 1957). Ce test fait appel à la discrimination des formes, des dimensions et des orientations. Ombredane reprend un mode de passation avec reduplication. Contrairement à l'étude menée avec le *Matrix-couleur*, au fil des trois passations, la réduction des erreurs reste faible et, à âge égal, le degré de scolarité n'induit pas de différences majeures dans les résultats.

Les constats élaborés par Ombredane concernent des populations africaines *a priori* relativement éloignées des systèmes de représentations et de valeurs des adultes tout-venant de culture occidentale. Sur ces derniers, retrouve-t-on des conclusions qui vont dans les mêmes directions d'une "épuration" du résultat par élimination des facteurs parasites et par augmentation de la fidélité et de la validité ? Dans la même période, une étude américaine menée par Fattu, Mech et Kapos met en évidence l'existence, chez les adultes, de ce qui ressemble aux "zones de proche développement" chez les enfants. Ces auteurs s'intéressent à la question de la résolution de problèmes à partir du constat du vide de connaissances relatif aux processus de résolution : "Il n'est pas suffisant de décrire le "bon" sujet simplement comme celui qui résout le plus de problèmes dans une situation donnée." (Fattu, Mech, Kapos, 1954). Les auteurs souhaitent mettre en évidence, de manière contrastée, les comportements mis en oeuvre par les individus les plus efficaces et les moins efficaces. Ils souhaitent également savoir si, sous l'effet d'un entraînement, les moins

bons peuvent se rapprocher des meilleurs. L'étude porte sur 49 adultes en formation. Chacun a passé environ deux heures dans la situation expérimentale face à un appareillage constitué de trains d'engrenages.



DOCUMENT N°19: Leviers de Meili (Meili, 1972)

Avec l'aimable autorisation des E.A.P.

La tâche consiste à repérer le dysfonctionnement qui a été introduit dans ce système mécanique. La procédure comporte trois phases, chacune avec six essais : test, inter-test avec informations, aides et post-test. Le test permet de déterminer le niveau initial de la capacité de résoudre les problèmes d'engrenage. A l'inter-test, une information

audiovisuelle (bande son illustrée de diapositives) expose la nomenclature du système et est suivie de six problèmes à résoudre. Le post-test a lieu un jour plus tard et propose la résolution des six derniers problèmes suite à une seconde session d'entraînement.

Les auteurs distinguent trois groupes de sujets en fonction du profil de l'évolution de leurs résultats. Un premier groupe rassemble les sujets efficaces dans les trois séances (scores moyens : 3,5 ; 3 ; 3,8). A l'opposé, un troisième groupe rassemble les sujets peu efficaces aux trois passations (scores moyens : 1,5 ; 1,7 ; 1,9). Enfin, entre ces deux extrêmes, un second groupe concerne les sujets qui progressent sensiblement d'un essai à l'autre (scores moyens : 1,2 ; 1,8 ; 4,3). Cette forme de résultats est à rapprocher de celle que mettra en évidence Budoff quelques années plus tard avec les enfants : sujets aux scores élevés, sujets gagnants, sujets non-gagnants. On constate donc que la notion de potentiel d'apprentissage s'adresse tout aussi bien aux adultes qu'aux enfants : placés en situation d'apprentissage, la réactivité apparaît comme extrêmement différenciée.

Plus près de nous, quelques études francophones abordent la question sous différents angles. Meili, dans son test des leviers, préconise d'aider les sujets en difficulté en leur montrant les solutions. "Le fait que le sujet a toujours connaissance des solutions de tous les problèmes est caractéristique de ce test, et pour la plupart des sujets d'un genre assez nouveau. Il faut donc leur permettre de s'y adapter et d'apprendre la méthode des solutions. Cet examen s'apparente ainsi à la situation dans laquelle se trouve l'apprenti." (Meili, 1972). Concrètement, cette épreuve implique la manipulation d'un matériel : une plaque percée de trous, des barrettes perforées, des fiches qui servent à fixer les réglettes sur la plaque, et des chevilles pour articuler les réglettes entre elles. L'examineur fixe deux barrettes sur la plaque et il demande au sujet d'élaborer avec le matériel mis à sa disposition un système de telle manière qu'un mouvement donné à l'une des barrettes provoque le déplacement de l'autre dans une direction imposée (voir document n° 19).

Les douze items sont présentés successivement et en temps limité. En cas de difficulté, l'examineur fournit le matériel nécessaire à la construction du système de leviers. Si la difficulté persiste, il réalise la solution en démonstration avant de passer à l'item suivant, en s'abstenant de donner toute explication verbale. L'étalonnage porte sur des groupes de jeunes gens suisses ayant quitté l'école (14 ans, 15 ans, 16 ans, 17 ans). Il manque une étude de fidélité et l'appréciation des différentes formes de validité. De manière complémentaire, Reuchlin (1950) propose une observation du comportement pendant le test en trois types d'activité : activité organisée, activité inorganisée, activité rapide.

Debray et Dufay ont opté pour une approche exclusivement clinique de l'évaluation dynamique du potentiel d'apprentissage d'un groupe de soixante opérateurs industriels. L'évaluation dynamique intervient ici comme élément d'orientation dans un dispositif de formation devant conduire au titre de "Professionnel de fabrication" ; elle est complétée par une mesure de niveau en mathématiques, français, dessin industriel et technologie. Selon les résultats de l'évaluation, les opérateurs pourront soit accéder directement à un module de quatre cent heures de formation technique, soit commencer par une préformation distribuée sur une année. L'évaluation dynamique est retenue afin d'atteindre les capacités d'apprentissage du sujet, au-delà de son niveau de connaissance actuel. Elle prendra deux demi-journées. La démarche générale est dérivée de celle que Feuerstein préconise dans le cadre du L.P.A.D., devenu pour la circonstance E.D.P.A. (Evaluation Dynamique du Potentiel d'Apprentissage) et elle répond au souci d'explorer en général " les capacités d'apprentissage du sujet face à différents registres de raisonnement" sur quatre axes en particulier : "1 - Quelles méthodes d'apprentissage et/ou de résolution de problèmes étaient accessibles à chaque sujet ? 2 - Quelles étaient leurs capacités d'abstraction et/ou leur capacité de gérer la complexité ? 3 - Quelles aptitudes à la décentration (forme de flexibilité) leur étaient accessibles ? 4 - De quelles possibilités de vérification et de ténacité face à une tâche faisaient-ils preuve ?" (Debray & Dufay, 1994). Les exercices

sont originaux. Des contenus décontextualisés du scolaire et du professionnel sont organisés de manière à induire des modes de raisonnement et des conduites utiles aux apprentissages futurs. La première séance n'implique pas la performance : elle est consacrée à la sensibilisation des sujets aux différents types d'exercices, par des travaux alternativement en individuel, en petits groupes, en collectif... La seconde séance se centre sur l'évaluation proprement dite : le travail se fait en individuel et l'observation porte sur les capacités de transférer les apprentissages réalisés au cours de la première séance sur les nouveaux exercices. Ces derniers offrant des variations dans le niveau de difficulté, il devient possible de prendre en compte les dimensions de décentration et de ténacité. A l'issue de l'évaluation, trois groupes de sujets furent constitués : trente et un pour une entrée directe en formation, quinze pour la préformation, quatorze "cas frontières" aux résultats hétérogènes, à discuter à la lumière des résultats d'enseignement général. Un suivi du parcours des stagiaires devrait permettre de valider le système de catégorisation utilisé. Toutefois, il manque une étude de fidélité pour tester la solidité de l'attribution des catégories.

Chartier et Loarer s'intéressent à l'évaluation de publics non-francophones ou de bas niveau de qualification afin de minorer les effets parasites qui jouent dans l'évaluation statique, effets liés au degré de familiarité avec la tâche et/ou aux composantes perceptives... Le matériel est repris de l'épreuve des cubes de la W.A.I.S. et la procédure empruntée à Ionescu, l'aide verbale étant supprimée. Concrètement, tous les sujets répondent à tous les items. En cas d'échec intervient le système d'aide hiérarchisé en trois niveaux : modèle à l'échelle 1, modèle surligné, démonstration par l'examineur. Suite à une aide, le modèle initial à l'échelle 1/2 est présenté de nouveau. Une première étude de validité a été conduite à son terme : "...la prise en compte de la note de transfert dans un score global NS + NT [NS note spontanée, NT note transfert des réussites après aide.] améliore quelque peu la validité prédictive de l'épreuve par rapport à un critère externe qui est la notation des moniteurs ayant eu à superviser le travail des sujets." (Chartier & Loarer,

1994). Pour le reste, les auteurs projettent d'utiliser des traitements statistiques sophistiqués de réponse à l'item, afin de dépasser les défauts des scores de gain, en reprenant les propositions d'Embretson (1987).

Dans le cadre de la formation professionnelle des adultes, la recherche d'une dynamisation d'épreuves psychométriques a porté sur le test C.E.R.P. 2, ou C2a. Cette épreuve, créée en 1945 par Royer, est dérivée des matrices de Raven. Elle fut utilisée comme épreuve d'intelligence générale dans une batterie de sélection des candidats à la formation professionnelle. Le test est décrit par Patin et Vinatier (1963). Il se compose de quarante neuf items, dont trois exemples, présentés par ordre de difficulté croissante. Ils ont forme de matrices à compléter par réponse à choix forcé, dans un temps limité. La méthode pair-impair donne des coefficients d'homogénéité élevés : 0,79 ; 0,84 ; 0,89. La méthode par reduplication traduit une fidélité correcte : 0,67 ; 0,79 ; 0,81 ; 0,86 ; 0,93. Entre les deux passations, l'effet de retest produit une augmentation moyenne des scores d'environ 2,5 points. Les scores ont tendance à baisser avec l'âge à partir de vingt ans, processus d'involution propre aux tests de facteur G. A âge égal, ils varient dans le même sens que la durée de scolarité. Les corrélations avec d'autres épreuves sont généralement élevées : de 0,50 à 0,68 avec des épreuves d'intelligence ; de 0,45 à 0,66 avec des épreuves spatiales ; 0,57 avec une épreuve de raisonnement logique. Les corrélations avec des épreuves de connaissances sont plus faibles et plus disparates : de 0,06 à 0,45 ; et celle avec une épreuve psychomotrice se situe à 0,31. Les corrélations avec la réussite aux examens de fin de stage sont elles-mêmes fluctuantes en fonction des métiers ; elles varient de - 0,28 à 0,52.

On dispose donc avec le C2a d'une épreuve bien rôdée et aux qualités métrologiques parfaitement définies. Suite à la modernisation des batteries de sélection, cette épreuve devenait disponible dans le champ de l'expérimentation psychopédagogique, sans risque majeur d'interférer avec les procédures de sélection. Un premier essai de dynamisation fut tenté pour évaluer les effets d'actions expérimentales menées

pour des publics de jeunes adultes de bas niveaux. "...nous avons transformé le C2a... pour en tirer deux indices : l'évaluation de la performance directe ; l'évolution des capacités auto -correctives alimentées par la présentation écrite de principes pour réussir". (Pasquier, Chiocchetti, Dupuy, Neveu, Brachet, Vernet, 1987). L'épreuve fut passée en temps libre et en réponse ouverte (dessiner la partie manquante de la matrice) par une quinzaine de jeunes stagiaires en début d'un module axé sur le réentraînement intellectuel et au terme de ce module, six mois plus tard. La moyenne des scores exprimant la performance directe, avant la présentation des principes de résolution, passe de 29,7 à 33,5 points. Par contre le taux des erreurs auto corrigées chute de 17,4 à 4,5%. Il est vrai que les items restant à corriger au cours du post test sont d'un niveau de difficulté plus élevé et on retrouverait là le constat de Robaye relatif à l'influence de la difficulté des items sur les possibilités d'autocorrection. Les stagiaires ont également avancé l'argument que le module leur avait appris à mieux contrôler leurs processus d'élaboration des réponses en cours de réalisation de la tâche. Ce premier essai ne fut guère convaincant du fait de l'hétérogénéité d'une activité de résolution suivie d'une activité d'auto-correction qui complique le transfert des apprentissages.

Pour la seconde tentative de dynamisation du C2a, le schéma T.A.T. fut retenu. Les trois séquences furent réparties sur trois journées consécutives : la première et la troisième consacrées à la passation du test en temps libre, la seconde à l'apprentissage des principes pour réussir. Dans ce nouveau cadre... "Evaluer l'éducabilité consiste à définir le degré de sensibilité d'un sujet à une expérience d'apprentissage. Ce degré de sensibilité s'exprime métriquement sous la forme d'un potentiel d'apprentissage." (Pasquier, 1989c). L'objectif était d'avancer dans l'élaboration d'un premier outil d'évaluation du potentiel d'apprentissage des adultes dans une perspective psychopédagogique de préconisation de remédiations cognitives pour des stagiaires montrant une faible sensibilité à l'expérience d'apprentissage et plus généralement d'individualisation des formations. "L'évaluation du P.A. peut se traduire en termes de

risque pédagogique, ou encore de dépendance pédagogique : plus le P.A. est réduit, et plus les apprentissages réalisés seront fonction de l'aide apportée par le formateur et inversement, plus le P.A. est élevé et plus le stagiaire pourra apprendre par ses propres moyens." (Pasquier, Bougeard, Corvaisier, Toulisse, 1989a)

Sur un plan fonctionnel, le schéma T.A.T. fait appel à la capacité du sujet de transférer en situation de résolution autonome (post-test) des apprentissages réalisés avec une aide pédagogique externe (entraînement) dans le but d'améliorer sa première performance (pré -test). Une attention plus forte a été donnée à la question du transfert des apprentissages, en référence aux travaux d'Oléron : "...l'extension de la généralité est le plus souvent limitée : des gradients de généralisation ont été fréquemment constatés, et le sujet tend à cesser de répondre quand les stimuli nouveaux sont par trop différents qualitativement ou quantitativement des stimuli qui ont provoqué les premières réponses." (Oléron, 1963). Par exemple, un élève qui a appris à tracer la hauteur du triangle à l'intérieur sera incapable de résoudre le cas du traçage à l'extérieur, restant limité au cas particulier sur lequel l'enseignement s'est appuyé. Le gradient de transfert a été segmenté en quatre niveaux. Le niveau A regroupe sept items communs au test et à l'apprentissage. Ces items ont servi de support à l'examineur pour réaliser les démonstrations : les sujets ont donc vu les réponses correctes. Il leur suffit de se les remémorer au cours du post-test, par transfert direct de la réponse. Le niveau B comprend onze items communs à l'entraînement et au test servant de supports pour mener à bien l'apprentissage des principes pour réussir repris de la première tentative (voir *supra*). Une réussite au post-test traduit donc un transfert direct du principe de résolution sur un matériel identique à celui ayant servi à l'apprentissage. Le niveau C comprend quinze items différents à l'entraînement et au test mais il est possible de les résoudre par application directe de l'un des principes de résolution enseignés. Une évolution de la performance sur ces items peut donc s'interpréter comme transfert direct d'une méthode sur des items homomorphes. Le niveau D, enfin,

regroupe seize items différents à l'entraînement et au test : on ne peut pas les résoudre par application directe de l'un des principes de résolution enseignés. Une progression sur ces items suite à l'entraînement suppose une généralisation de la démarche inductive-déductive.

TABLEAU 3: C2a-PA, gain relatif, taux d'interférence et gradient de transfert

gradient de transfert	taux d'interférences	gain relatif médian
identique 1	20%	69,5%
identique 2	10%	64,5%
semblable	30%	21,5%
différent	50%	1,5%
ensemble	10%	32,5%

Au cours de l'entraînement, cinq principes pour réussir sont enseignés. Premier principe : "Je prolonge les traits de la même manière qu'ils sont commencés" (principe de continuité). Deuxième principe : "La bonne réponse s'obtient en s'imaginant que l'on plie le dessin selon une ligne droite horizontale, qui passe par son milieu" (principe de symétrie). Troisième principe : "Pour passer d'un dessin à l'autre, j'ajoute toujours une même quantité" (principe de progression). Quatrième principe : "Dans chaque ligne, l'une des figures est égale à la somme des deux autres" (principe d'addition). Cinquième principe : "Je fais tourner la figure ou certains de ses éléments d'une manière constante" (principe de rotation). Cinq items par principe sont utilisés à l'entraînement pour apprendre à appliquer ce principe.

Les deux premiers servent de support de démonstration à l'animateur et les trois autres à l'entraînement individuel. En cas de difficultés éprouvées par un sujet, l'animateur intervient pour aider, d'abord en demandant au sujet de verbaliser sa manière de faire (recours à la métacognition) et, si nécessaire, il propose une démonstration (recours à l'imitation). Les items d'entraînement sont à compléter par le dessin de la réponse ce qui donne la possibilité de pouvoir mener une analyse qualitative des traces graphiques. Ces traces peuvent s'interpréter du point de vue des ajustements psychomoteurs, par exemple pour dessiner deux demi-cercles tangents, ou du point de vue de la constance perceptive, ou encore du point de vue de la compréhension du problème posé...

Différents scores ont été étalonnés : niveau initial obtenu au pré-test, potentiel d'apprentissage global, potentiels d'apprentissage par niveau de transfert, taux d'intégration positive (item corrigés au post-test), taux d'intégration négative (nouvelles erreurs au post-test). Les indices de potentiel sont calculés comme des gains relatifs ($\text{gain effectif} / \text{gain possible} \times 100$), mode de calcul emprunté à Higé et Martin (1979). Ce mode de calcul a le mérite de la simplicité et il est compris par les sujets, par les examinateurs, par les formateurs... Toutefois, plus le niveau initial est élevé, et moins il reste d'items à conquérir. Un même indice recouvre donc des réalités différentes : 25% de potentiel c'est gagner 2 items sur 8, ou 4 sur 16, ou 8 sur 32... Il en résulte un certain paradoxe : à P.A. constant, plus le niveau initial est bas et plus il faut gagner d'items. Paradoxe apparent quand on considère que, les items étant classés par ordre de difficulté croissante, le nombre d'items à gagner contrebalance le niveau de difficulté de chacun d'entre eux.

L'étalonnage a été construit auprès de cent dix-neuf stagiaires d'une dizaine de stages de pré-formation. Les âges s'étendent de 17 à 48 ans ; 46% de l'effectif est masculin et 54% féminin. Les niveaux de formation vont de l'I.M.PRO. au D.E.U.G. ; les expériences professionnelles, en général, ne dépassent pas une année. Cet aspect disparate de l'échantillon reflète bien l'hétérogénéité du public en quête d'emploi ou de

formation. Les niveaux initiaux de réussite au pré-test sont très dispersés, de 5 à 46 points, avec une médiane à 34,5. Le potentiel global représente la même dispersion : de - 50% à + 75%, la médiane se situant à 32,5 (voir tableau n°3). Cela veut dire qu'au post-test, si les meilleurs progressent des 3/4 dans la zone qui leur reste à conquérir, les moins bons perdent la moitié des points gagnés au pré-test.

La mise en situation d'apprentissage de l'adulte amène des résultats largement différentiels. Le P.A. évalué chez l'adulte apparaît donc comme une variable extrêmement discriminante des sujets gagnants, non-gagnants, voire "perdants".

Les auteurs lus ne signalent pas la question de la diminution de la performance suite à l'entraînement qui touche ici 10% des sujets. Cette tendance à la régression croît avec les niveaux de transfert puisqu'elle touche 50% des sujets pour les items du quatrième niveau de transfert. D'une manière générale, la valeur du P.A. médian chute au fur et à mesure qu'on se déplace dans les niveaux de transfert de 69,5% à 1,5% (voir tableau n°3) : plus on s'éloigne de la situation d'apprentissage, plus le transfert devient aléatoire et plus le risque d'interférence augmente.

Cette expérimentation dans le champ de la psychopédagogie confirme la pertinence de la variable potentiel d'apprentissage pour les adultes, du moins sur le plan de la discriminativité, de la sensibilité à la mise en situation d'apprentissage (Pasquier, 1989b, 1989c). Elle met également en évidence l'importance de la variable gradient de transfert, loi qui apparaît commune aux enfants et aux adultes, et donc *a priori* relativement insensible à l'effet du vécu expérimental. Toutefois, cette expérimentation incomplète sur certains points amène de nouvelles questions de nature métrologique : quelle fidélité de la mesure du P.A. ? quelles validités de contenu, concurrente et prédictive de cette mesure ? Sur le plan théorique, la question de l'interprétation du P.A. de l'adulte reste ouverte : quel degré de similitude ou de différence avec la zone de proche développement de Vygotski ? quel rapport entre effet de reclassement et épuration de la mesure au sens donné par

Ombredane ? quelles influences des variables différentielles de sexe, d'âge, de milieu, d'expérience professionnelle ? quel degré de généralisabilité du P.A. en fonction des domaines de contenus et de compétences ? quelle interprétation des interférences ?

Le travail de mise au point du T.E.D.E. a permis d'avancer sur quelques-uns de ces points de questionnement.

Deuxième section : le T.E.D.E.

Historiquement parlant, Le T.E.D.E., *Test d'évaluation dynamique de l'éducabilité* (Pasquier, 1989d), constitue le premier volet de la "démarche éducabilité professionnelle" proposée par Avenir & Entreprise ®, (voir figure n°8). Cette démarche fut élaborée pour répondre à des besoins d'entreprises travaillant sur des plans de qualification professionnelle de leur personnel (Briet, 1991 ; Jacquin, 1991). Le schéma se développe parallèlement pour les opérateurs et pour l'encadrement de proximité. Pour les opérateurs, suite à la passation du T.E.D.E., quatre parcours conduisent aux modules qualifiants. Le parcours I, prévu pour les personnes dont le potentiel d'apprentissage se situe autour de 0, fait passer par les quatre modules de remédiation des *Exercices pour développer l'intelligence*, les E.D.I. (Pasquier 1990b, 1990c, 1990d, 1991a). Le parcours II, pour les personnes qui apprennent bien mais qui éprouvent des difficultés à transférer de manière flexible, prévoit le passage par les deux derniers modules des E.D.I. Le parcours III est une opportunité de réentraînement au travail intellectuel pour des opérateurs dont le potentiel d'apprentissage est élevé mais qui ne sont pas allés en formation depuis longtemps. Enfin, le parcours IV conduit directement aux modules qualifiants les personnes aptes à les suivre d'emblée avec les meilleures chances de succès. Parallèlement, les agents de l'encadrement de proximité cheminent sur les modules de Formation à l'animation formative, les P.A.F. (Pasquier, 1991b, 1991c). A l'issue des parcours, les partenaires se retrouvent au poste de travail où les agents de l'encadrement de proximité stimuleront et guideront le transfert des acquis du développement personnel et de la formation qualifiante en direction de l'activité professionnelle...

Cette brève description de la démarche permet de situer la place originale du T.E.D.E. qui, bien entendu, peut s'utiliser en dehors de cette démarche globale. Dans les faits, le T.E.D.E.

s'intègre à tous dispositifs d'évaluation dont l'un des objectifs se définit dans la connaissance du potentiel d'apprentissage de la personne.

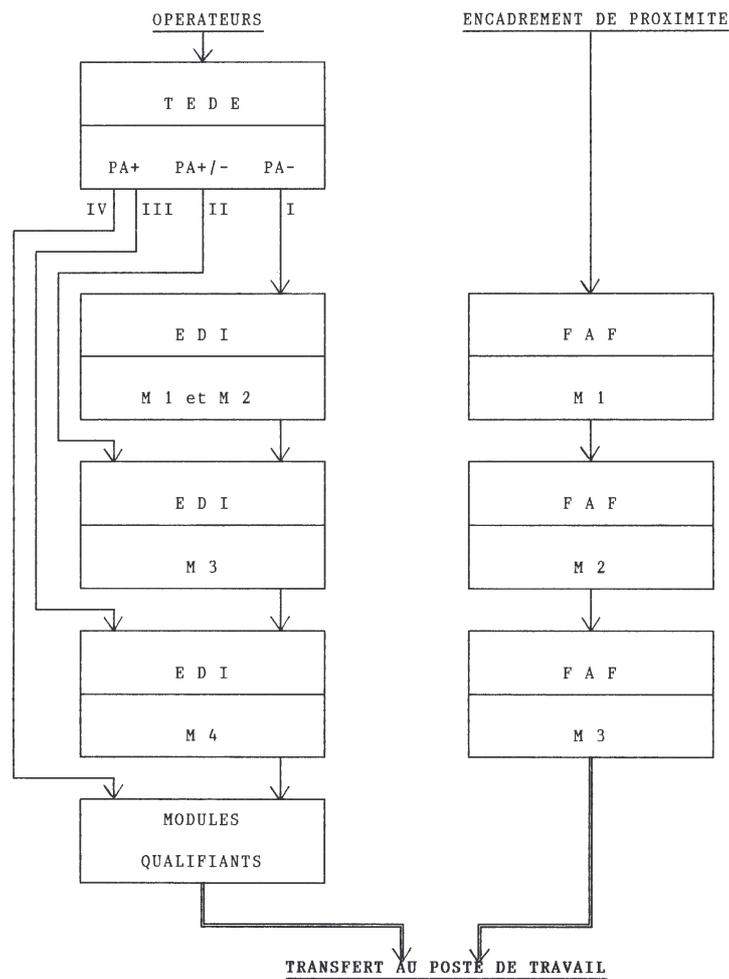


FIGURE N°8: démarche éducativité professionnelle A & E (R)

Au niveau de l'évaluation des personnes, on trouvera le T.E.D.E. intégré à des batteries de tests de recrutement ou de sélection ; à des procédures d'investigation mises en oeuvre

dans le cadre du bilan des compétences personnelles et professionnelles ; à des plans d'évaluation psychopédagogique destinés à l'élaboration de préconisations de remédiation cognitive ou de parcours de formation individualisés ; à des phases d'évaluation des potentiels dans le cadre de la gestion prévisionnelle des emplois et des compétences... Dans ce cadre de l'évaluation individuelle, le T.E.D.E. est mis en oeuvre afin de construire un pronostic de réussite dans le suivi d'une formation professionnelle, dans un changement de fonction ou de poste de travail, et d'une manière générale afin d'apprécier les chances de réussite des projets d'inflexion des trajectoires socioprofessionnelles. Éventuellement, le recueil et le traitement des observations qualitatives faciliteront la compréhension du mode de fonctionnement du sujet dans la tâche et, si nécessaire, la formulation des conjectures préparatoires à la mise en oeuvre d'un diagnostic approfondi. Enfin, le T.E.D.E. peut servir dans le cadre de plans expérimentaux dans le but d'évaluer les effets, en termes d'élargissement du potentiel d'apprentissage, d'actions de remédiation, de pré-formation, de formation professionnelle.

Le chapitre dixième sera consacré à la description du T.E.D.E. dans sa version d'origine. Le chapitre onzième présentera les différentes versions de l'épreuve, et le chapitre douzième évoquera les différentes études des qualités métrologiques du test.

Pour aller plus loin dans la connaissance du T.E.D.E. :

Pasquier D. (1991), Mesurer le potentiel d'apprentissage, *Les cahiers d'information du directeur du personnel*, N°17

Pasquier D. (1991), Regards neufs sur la capacité d'apprendre, *Les cahiers d'information du directeur du personnel*, N°18

Pasquier D. (1994), Évaluation dynamique, potentiel d'apprentissage et facteur G, in *A la recherche du*

potentiel d'apprentissage, Garbo R. et Lebeer J. (éds.),
University of Antwerp : EAMC

Pasquier D. (1994), Le test d'évaluation dynamique de l'éducabilité (T.E.D.E.), in *Les techniques d'évaluation des personnes*, Huteau M. (éd.), Issy-les-Moulineaux : EAP

Pasquier D. (1996), Un paradigme d'hier pour demain : l'évaluation dynamique du potentiel d'apprentissage, *Actes du congrès international d'actualité de la recherche en éducation et formation*, Nanterre : Université de Paris X

Pasquier D. (1997), L'évaluation dynamique en psychologie, *Le journal des psychologues*, N°145

Pasquier D., Estebe I., Jaigu J., Prévoir la réussite chez de jeunes apprentis : approche exploratoire, *Pratiques Psychologiques*, 2001-1, 99-110

Pasquier D., Estèbe I., Jaigu J., *Prévoir la réussite chez de jeunes apprentis : rapport d'expérimentation 1999-2001*, [en ligne]. Disponible : avenireentreprise.tk/bibliothèque/rapportcfa.pdf

Pasquier D., Maire P. (2000), *Individualiser les A.S.L.O.S. (Ateliers de structuration logique et spatiale) à partir de la passation du T.E.D.E.-A. (Test d'évaluation dynamique de l'éducabilité - adapté)*. [en ligne]. Disponible : inrp.fr/Acces/Biennale/5biennale/Contrib/Long/L310.htm

Chapitre 10

Le T.E.D.E. historique

Le T.E.D.E. est bâti selon un principe de recherche d'homomorphie entre le test et la situation de travail dans le but de lui garantir la meilleure validité de construction. La première homomorphie concerne la nature de la tâche du test. Plusieurs options s'offrent au créateur et les choix opérés en final ne sont pas exempts d'un certain arbitraire. La première option serait de s'appuyer directement sur des contenus techniques et professionnels dans une épreuve censée mesurer l'éducabilité professionnelle. Un tel choix, à première vue très rationnel, pose rapidement des problèmes insurmontables quand on commence à penser son opérationnalisation. Il conviendrait de construire une épreuve par métier : vu leur nombre, cela paraît une entreprise infaisable sans moyens considérables. D'autre part, les métiers évoluent plus ou moins rapidement, plus ou moins radicalement et il faudrait actualiser les contenus régulièrement pour tenir compte de ces évolutions. Enfin, il est fort probable que certains sujets aient acquis une certaine expérience du métier, ce qui leur donnerait un avantage substantiel par rapport aux sujets entièrement naïfs. Il faut donc revenir vers la tradition des épreuves le plus fréquemment dynamisées, à savoir les tests de facteur G dont le prototype reste les matrices. L'analyse cognitive d'une tâche de résolution de matrices montre un mouvement d'induction -déduction : analyse de la configuration et des relations entre ses éléments afin d'en induire la règle organisatrice ; application de la règle à l'élément manquant afin de déduire cet élément. Il est vrai que ce processus cognitif sature les activités à caractère scolaire : élaboration de règles de grammaire, élaboration de théorèmes de géométrie, formulation de lois physiques... et leurs applications dans des exercices de résolution convergente.

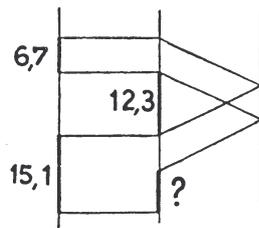


FIGURE C

2e EXEMPLE - Figure C.

On vous a donné 3 longueurs en trait fort : 6,7; 15,1; 12,3 et on vous demande de calculer une quatrième longueur en trait fort devant laquelle on a mis un point d'interrogation. Vous voyez qu'elle est la différence entre $15,1 - 6,7 = 8,4$. La réponse est **8,4**.

Il était donc inutile de se servir de la longueur donnée 12,3.

Avez-vous bien compris?

DOCUMENT N°20: Item du test Calcul des longueurs (Faverge, 1965)

Avec l'aimable autorisation des E.A.P.

Quand on essaie, de la manière la plus générale, de dégager la structure des conduites à l'œuvre dans la tenue d'un poste de travail, il apparaît assez vite que, contrairement au monde scolaire de la résolution de problèmes, l'inférence inductive-déductive ne joue pas forcément un rôle déterminant. A la limite, dans le moment même de l'exécution d'une tâche dont la maîtrise est parfaitement intériorisée, cette forme d'inférence peut ne jouer aucun rôle : le peintre qui passe un panneau mural au rouleau, le maçon qui monte du parpaing, la secrétaire qui saisit des factures, le mécanicien qui remonte une culasse..., n'ont pas recours à des conduites inductives du type

de celles qu'on trouve dans les tests. Ils mettent en oeuvre des représentations fonctionnelles de l'action "...qui assurent la planification et le guidage de cette activité." (Leplat, 1985).

Les représentations fonctionnelles sont au principe de la formation des images opératives, "...structures informationnelles spécialisées qui se forment au cours de telle ou telle action dirigée sur des objets." (id.) "L'image opérative représente... toujours une certaine information immédiatement disponible sur l'objet (information inhérente à l'image) reflétée dans la conscience du sujet et interagissant activement avec l'information-signal, c'est-à-dire avec l'information qui vient au sujet de l'extérieur au cours même de l'action." (Ochanine, 1981).

De fait, toute situation de travail renvoie au domaine des praxies, aux contraintes de l'action, à la logique des transformations, aux paramètres de l'effectuation et du contrôle. Bien évidemment, il ne s'agit pas pour autant de rejeter les processus inductifs qui jouent toujours un rôle, principalement dans les situations de diagnostic de pannes ou de dysfonctionnements... Dans une situation de travail, il y a donc toujours quelque chose à faire (dimension praxique) et quelque chose à comprendre (dimension cognitive), et la qualité du produit du travail, de fait le degré de maîtrise de la tâche, reflètera la plus ou moins bonne coordination de ces deux dimensions. On peut illustrer ce principe par l'exemple des postes de travail modernisés : l'opérateur peut n'avoir plus qu'à pianoter sur les touches du clavier de son ordinateur, mais s'il pianote au hasard, il aura très rapidement des ennuis avec son patron ! Il doit mettre en relation la frappe des touches avec un ensemble de connaissances et de procédures propres à la machine conduite, à l'ordinateur, à lui-même... Dans ces conditions, il fallait trouver une tâche qui articulât ces deux dimensions, la dimension cognitive d'une part et la dimension praxique d'autre part, dans une troisième qui les englobe, troisième dimension relative à l'opérativité. Le test de calcul des longueurs propose une tâche intéressante de ce point de vue (Faverge, 1955, 1977) : elle met en oeuvre un calcul numérique alimenté par des opérations projectives. Dans l'exemple

proposé, (voir document n°20), il faut projeter mentalement les segments 6,7 et 15,1 sur l'axe 2 pour calculer par différence la longueur du segment ?, l'information 12,3 n'étant pas pertinente.

La réussite de la tâche met bien en jeu l'articulation d'une dimension praxique mentale (opérations projectives), et d'une dimension mathématique (opération numérique). Cette épreuve date de 1946, conçue à l'origine pour être mise en relation avec la réussite en mathématiques. Elle comprend dix-huit items à résoudre en temps limité. L'homogénéité (méthode pairs-impairs) est bonne (0,71 à 0,85), et les scores de fidélité (test -retest) apparaissent moyens (0,55 à 0,66), mais probablement peu représentatifs : "En effet les sujets ont été pendant trois mois soumis à une formation scolaire intensive et ne sont pas restés les mêmes ; d'autre part, il y a peut-être aussi une sous-estimation due au fait que le test a servi à la sélection des stagiaires." (Patin, Vinatier, 1962). Le nombre limité d'items peut également intervenir dans la valeur de la fidélité estimée. Des corrélations avec d'autres épreuves ont été calculées : de 0,38 à 0,53 avec le C2a, de 0,18 à 0,49 avec le *Test Mécanique* de Rennes, de 0,31 à 0,47 avec une épreuve de développés de volumes, de 0,14 à 0,49 avec une épreuve de mathématiques. L'ensemble de ces corrélations largement positives fait penser à une bonne saturation du test en facteur G, bien qu'aucune étude factorielle n'ait été menée.

Les études de validité concourante avec des épreuves de connaissances scolaires varient de 0,39 à 0,47. Les études de validité prédictive prenant en compte l'examen de fin de stage comme critère sont dans l'ensemble satisfaisantes, avec des corrélations pouvant atteindre 0,47 (commis du bâtiment), 0,58 (formation d'opérateurs mécanographes), 0,59 (formation de conducteurs de travaux)...

Toutefois, il n'était pas possible de reprendre en l'état le test de Faverge pour en dynamiser la passation : le calcul sur des nombres décimaux est une variable essentiellement scolaire qui empêche l'utilisation du test avec des sujets a-scolaires, ou peu scolarisés, ou sortis du système scolaire depuis longtemps... D'autre part, la présence d'informations parasites complexifie la possibilité de rentrer de plain-pied dans la tâche. C'est l'une des

raisons qui, à l'occasion de l'élaboration du T.E.D.E., ont amené des modifications importantes dans la présentation des items (voir *infra*).

La question de la durée fut réglée également dans un souci d'homomorphisme. Les tests habituels sont souvent de courte durée, de l'ordre de quinze à trente minutes, alors qu'une journée de travail, en formation ou en entreprise dure huit heures. Dans le test statique, il manque donc la dimension temporelle, dimension importante qui renvoie à la tenue dans le temps du sujet confronté à une tâche relativement répétitive et à la gestion de l'investissement énergétique distribué dans le temps. La dimension temporelle a été réintroduite dans la passation du T.E.D.E. afin de se rapprocher des conditions de la réalité du monde du travail et de la formation.

Le choix du paradigme s'est porté sur T.A.T. sans trop d'hésitations dans la mesure où il permet une passation collective, situation proche des stages de formation. Au fond, on peut rapprocher, sans risque de contresens, la procédure du T.E.D.E. à une micro-formation : évaluation initiale, enseignement et apprentissage, évaluation finale. Évidemment, au lieu de se distribuer sur quelques mois, le cursus est condensé en quelques heures. Les formateurs qui utilisent le T.E.D.E. soulignent l'importance de cet aspect économique de l'usage du T.E.D.E.

Habituellement, il leur faut trois à quatre semaines pour se construire une représentation stabilisée des capacités d'apprentissage d'un stagiaire. La mise en oeuvre du T.E.D.E. leur donne le même type d'information en quelques heures, et ce dès le début du stage. Le formateur peut en conséquence commencer à individualiser son action pédagogique beaucoup plus rapidement en fonction du profil éducatif de chacun des apprenants.

Il restait à définir le paramétrage de l'aide et de l'apprentissage fournis aux sujets, toujours dans un souci d'homomorphisme test/situations de formation ou de travail. Après analyse des formations des différents secteurs professionnels, il est apparu que l'objet d'apprentissage le mieux en rapport avec la situation de l'adulte apprenant serait

l'application d'un mode opératoire. L'aide pédagogique reprendra le schéma traditionnel de la démonstration, suivie d'une interaction individuelle au cours des applications, *grosso modo* à la manière des formateurs d'adultes. Au cours de cette interaction, l'examineur met en oeuvre un guidage pas à pas de l'application du mode opératoire, guidage qui vise à ré-autonomiser le sujet dans sa dynamique d'apprentissage. C'est au cours de l'interaction que les observations qualitatives sont les plus riches car elles permettent de se rapprocher des processus fonctionnels activés chez le sujet. Par ces observations, l'examineur-médiateur distingue entre des difficultés de différentes natures : décodage des termes écrits de l'opération, compréhension de la proposition verbale (image cognitive de l'opération), passage à l'effectuation (image opérative).

En final, la procédure de passation se déroulait en trois séquences consécutives, chacune d'une durée de quatre-vingt dix minutes : le pré-test suivi d'une pause, l'apprentissage des modes opératoires suivi de la pause déjeuner, le post-test. La passation et la correction, pour un groupe de huit sujets au maximum, mobilisait un examinateur sur une journée. La procédure peut sembler longue, voire lourde, et elle ne conviendrait pas au psychologue pressé évoqué par Rey dès 1934. L'examen T.E.D.E. se prolonge obligatoirement par un entretien de restitution des résultats qui s'intègre à part entière dans le procès d'évaluation dans la mesure où la résolution de certains items pourra être reprise, et surtout parce que la dimension subjective du rapport du sujet au savoir, du rapport du sujet à lui-même dans une situation d'apprentissage pourra y être abordée.

Dans sa forme première, le T.E.D.E. comprend trente-six items répartis selon deux critères à trois modalités : le niveau de complexité et le gradient de transfert. Il est donné dans la même forme au pré et au post tests. C'est la structure arithmétique sous-jacente qui différencie le niveau de complexité : $a = b + c$; $a + b = c + d$; $a - b = c - d$. Il y a douze items par niveau de complexité (voir document n°21). Le gradient de transfert est obtenu par rapprochement entre les

items du test et les items utilisés comme supports à l'entraînement. Douze items sont identiques aux tests et à l'entraînement pour évaluer un transfert très proche faisant jouer essentiellement la mémoire. Douze items sont semblables aux tests et à l'entraînement pour évaluer un transfert proche, par reduplication de l'application du mode opératoire enseigné (primat donné à la rigueur méthodique). Enfin (voir document n°22), douze items sont différents et n'ont pas d'homologues à l'entraînement pour évaluer un transfert éloigné qui demande des adaptations dans l'application de la méthode (primat donné à la flexibilité cognitive).

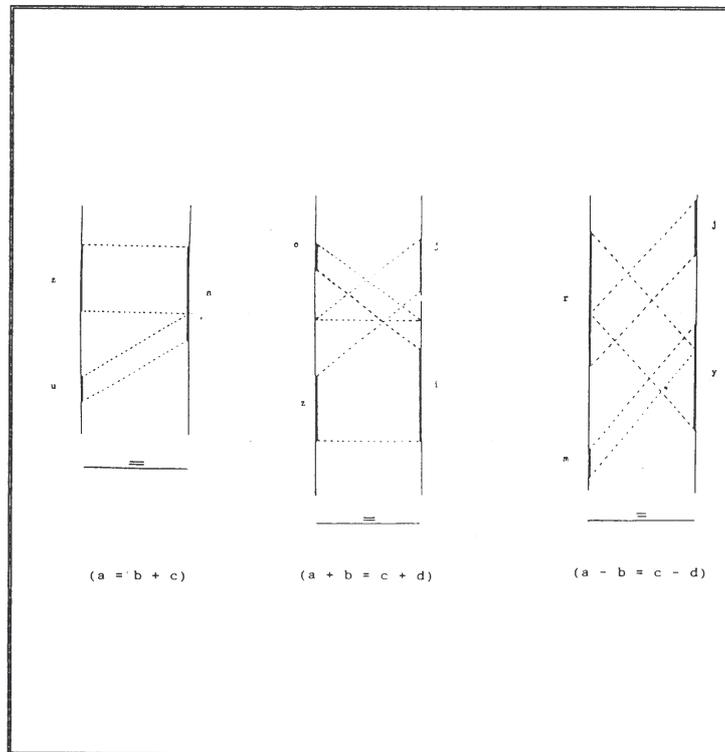
En croisant les deux critères, on obtient donc neuf groupes de quatre items. Dans chaque paquet d'items, ceux-ci ne sont pas rigoureusement classés par ordre croissant de difficulté dans la mesure où l'occurrence des événements qui créent l'histoire individuelle au quotidien respecte un certain désordre. Toutefois, de manière globale, le taux de réussite spontanée au pré-test varie en sens inverse du degré de complexité de la tâche (voir tableau n°4, étude menée sur trente-quatre protocoles mixant les deux modes de passation).

De la même manière, le taux de transfert (nouvelles réussites au post-test par rapport au nombre d'items à gagner) varie en fonction du degré de complexité et du gradient de transfert (voir tableau n°5, même *corpus* que tab. n°4).

La consigne du T.E.D.E. est double ; elle demande au sujet d'écrire l'égalité et de tracer les projections réalisées mentalement qui conduisent à la démonstration de la justesse de l'égalité par superposition des segments. On y retrouve bien l'articulation opérative entre la compréhension logico-mathématique (composer une égalité) d'une part et l'action effectrice d'autre part (tracer les déplacements), articulation qui peut conduire à la réussite de la démonstration par superposition (voir figure n°9).

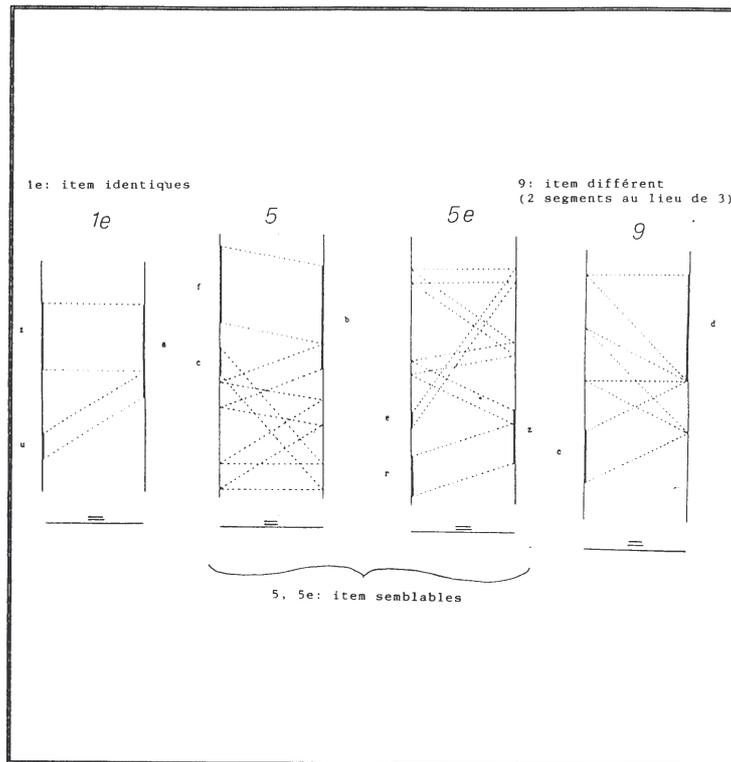
Le T.E.D.E. est utilisable avec des personnes situées aux niveaux VI à III. Toutefois, il peut arriver que, par un effet de prérequis, certaines personnes ne parviennent à rentrer dans la tâche et que, par un effet de plafonnement, d'autres parviennent à répondre correctement pratiquement à toutes les

questions dès le pré-test. L'étude d'une forme de passation adaptée aux sujets identifiés comme fonctionnant à bas niveau cognitif a démarré fin 1994. Deux étalonnages sont disponibles, l'un pour une autopathologie (projets dans le secteur tertiaire tous niveaux, niveaux IV et III), l'autre pour une passation guidée (projets de niveaux VI et V sauf tertiaire).



DOCUMENT N°21: Structuration du T.E.D.E. par niveau de complexité (Pasquier, 1989)

A quelles occasions utiliser le T.E.D.E. ? Certainement pas en toutes circonstances : dans la mesure où il permet d'évaluer un potentiel d'éducabilité, l'application du T.E.D.E. ne se justifie que dans les cas référenciables à une problématique d'inflexion d'une trajectoire socioprofessionnelle.



DOCUMENT N°22: Structuration du T.E.D.E. par gradient de transfert (Pasquier, 1989)

On ne mesure pas un potentiel pour voir, par simple curiosité ; on ne mesure pas un potentiel pour aider à choisir une orientation ; on mesure un potentiel pour évaluer la probabilité de concrétisation positive d'un projet apporté et porté par le sujet. Un enjeu externe à l'évaluation et partagé avec le sujet est donc absolument nécessaire pour libérer l'investissement dans la tâche, pour donner du sens à l'évaluation, pour orienter et comprendre la restitution des

résultats et leur utilisation comme élément de conseil dans un processus de prise de décision.

TABLEAU 4: T.E.D.E., taux de réussite spontanée et niveau de complexité

série	$a = b + c$	$a+b = c+d$	$a-b = c-d$
moyenne	61,83%	26,58%	15,16%
écart-type	24,03%	16,13%	10,46%

La problématique du projet socioprofessionnel, de son élaboration et de son évolution, s'inscrit dans plusieurs moments institutionnels. La loi de 1991 instituant le bilan de compétences personnelles et professionnelles offre un cadre légal à cette évaluation. Lorsque le projet passe par le détour d'une formation, le T.E.D.E. peut avantageusement compléter les épreuves statiques pour asseoir un pronostic de réussite. Lorsque plusieurs types de formation existent, de la formation sur le tas, au poste de travail, à la formation en centre, la structure du profil éducativité ainsi que les observations qualitatives peuvent aider au choix de la formule la mieux adaptée au cas.

Les procédures de recrutement représentent un autre contexte d'utilisation du T.E.D.E., que le recrutement soit externe ou qu'il concerne un mouvement de mobilité interne. Dans les deux cas, le psychologue va procéder au rapprochement du profil du poste et du profil éducativité du candidat. Bien entendu, une décision de recrutement ne se prend jamais sur un seul critère : le profil éducativité est une information parmi d'autres, à coordonner aux autres informations.

Bilan, recrutement... reste à décrire le troisième contexte d'utilisation du T.E.D.E. : le centre de formation. Là,

deux situations sont à distinguer : la pré-formation et la formation professionnelle. En pré-formation, le T.E.D.E, ne peut être utilisé de manière systématique, mais seulement auprès des stagiaires dont le projet professionnel devient suffisamment défini et stabilisé. Dans le cadre de la formation professionnelle, la question du projet ne se pose plus ni en termes d'orientation, ni en termes de sélection : les stagiaires sont là pour apprendre un métier précis. Le formateur peut alors utiliser le T.E.D.E. dès le début du stage. Il retirera de la passation des indications précieuses qui lui assureront la possibilité d'adapter sa prestation pédagogique aux profils éducatifs et du groupe et de chaque individu qui le compose.

TABLEAU 5: T.E.D.E., variation du taux de transfert en fonction du degré de complexité et du gradient de transfert

niveau de complexité gradient de transfert	a = b + c	a + b = c + d	a - b = c + d	ensemble
identique	86,36%	57,89%	43,26%	53,46%
semblable	72,34%	42,71%	32,43%	43,67%
différent	40,47%	18,64%	17,05%	23,56%
ensemble	56,86%	37,03%	29,94%	37,78%

Les résultats du T.E.D.E. se distribuent sur un profil éducatif étayé par l'étalonnage en forme de décalages des variables retenues : niveau de réussite initiale au pré-test, niveau de réussite terminale au post-test, potentiel d'apprentissage calculé comme gain relatif, progression relative par gradient de transfert, progression sur les seules égalités. Les informations qualitatives sont recueillies à plusieurs sources : l'observation générale du comportement au cours des trois

séquences, le recueil de données pendant les interactions duelles relatives à la manière dont le sujet apprend à appliquer un mode opératoire, la structure des résultats chiffrés par pavés d'items d'une passation à l'autre (échec-échec ; échec-réussite ; réussite-échec ; réussite-réussite), la qualité des traces graphiques.

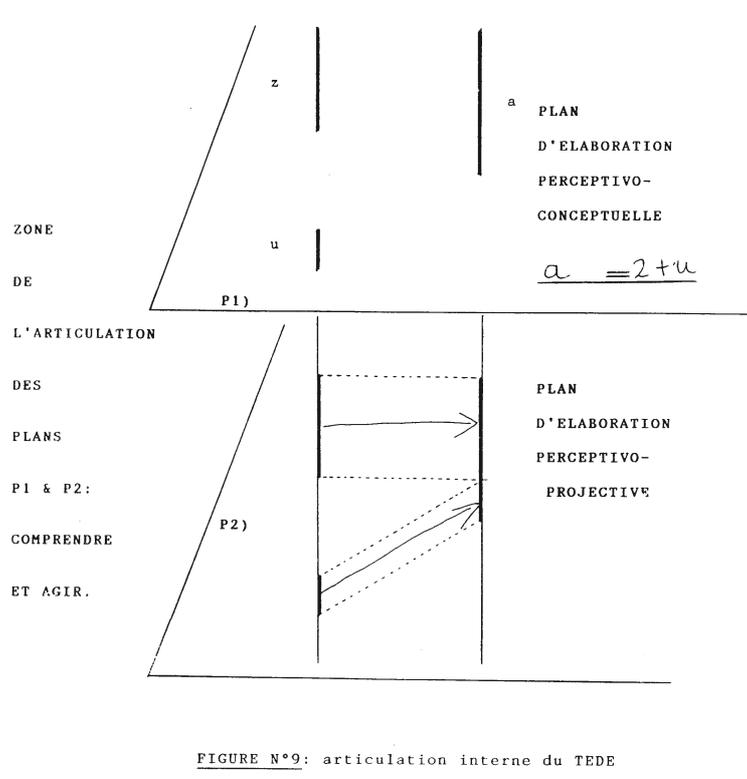


FIGURE N°9: articulation interne du TEDE

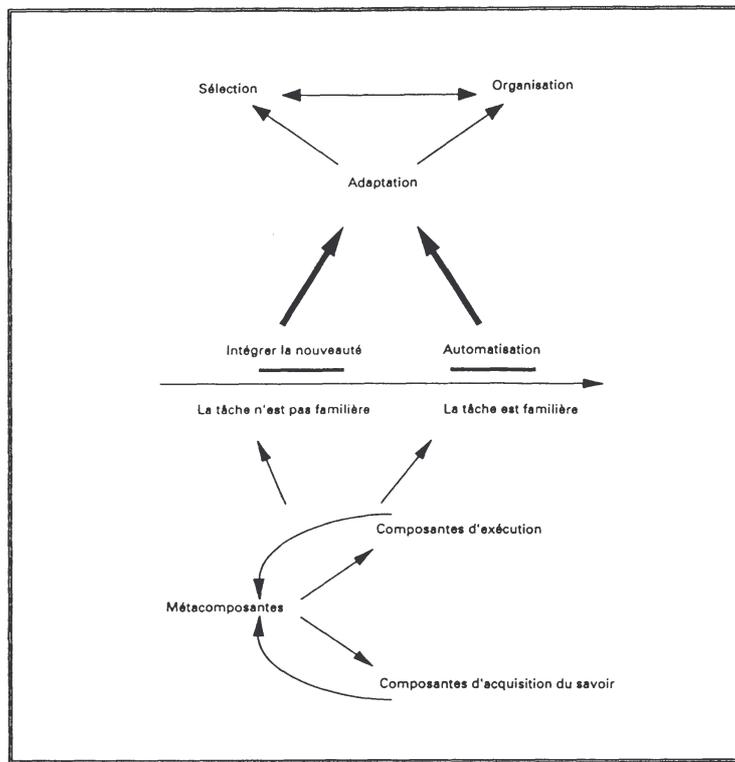
L'interprétation part du profil général des résultats et va rechercher des explications à la physionomie de ce profil dans les données du test lui-même (recherche de cohérence interne). Pour ce faire, les résultats chiffrés et leurs inter-variations seront rapprochés des informations qualitatives recueillies en cours de passation, d'abord de manière analytique descendante, si nécessaire jusqu'au niveau des traces graphiques, à la recherche des indices significatifs, puis de manière synthétique

ascendante par recoupement et regroupement des indices, et par formulation de conjectures. Ces conjectures seront travaillées au cours de l'entretien de restitution dans le souci d'une recherche de cohérence externe. Il s'agit de déterminer si la conduite observée au cours du test est un simple artefact de passation ou bien si cette conduite renvoie à une dimension stable du mode de fonctionnement du sujet. Par exemple, suite à la passation du T.E.D.E., l'examineur s'assure du degré de généralisabilité qu'il peut attribuer au fait qu'un sujet n'a pas progressé car il est revenu à son mode de fonctionnement spontané au post-test. Est-ce là un accident de parcours lié éventuellement à un manque de familiarité avec le mode opératoire proposé ? Ou bien, est-ce un trait d'égoïsme bien enraciné et soigneusement entretenu qui fait que cette personne refuse systématiquement de s'approprier une méthode externe ? Ou bien encore, est-ce une tendance à la régression cognitive et opératoire en réaction à la résistance de l'objet à traiter ? Sans entrer dans le détail des nombreuses combinatoires interprétatives qui assurent l'unicité de chaque cas, la lecture interprétative globale des résultats peut se baser sur au moins quatre corpus théoriques relativement complémentaires.

Au plus simple, on peut considérer le point de vue d'Ombredane d'une épuration du résultat du test par le simple fait de sa réduplication qui permet de neutraliser tout ou partie des facteurs parasites liés à l'appartenance culturelle, au degré d'habitude de confrontation avec les situations évaluatives et compétitives, au degré de familiarité avec la tâche, à l'intensité de la motivation, à la tonalité émotionnelle face au stress, face au nouveau...

Selon les enseignements de Piaget on peut voir les choses en termes de conflit cognitif et de compensation des perturbations (voir figure n°10). Au cours du pré-test, la personne construit son propre mode de fonctionnement spontané pour résoudre la tâche. Arrive la contrainte d'apprentissage d'une méthode externe qui va s'opposer à ce mode de fonctionnement spontané. Trois grands types d'issues à ce conflit sont envisageables selon les phases décrites par Piaget

le degré de l'intériorisation comme condition de l'amélioration de la performance renvoie à considérer les choses du point de vue d'Ochanine, celui de l'opérativité. En effet, il ne suffit pas de mémoriser une suite de propositions orales ou écrites, encore faut-il les traduire en conduites effectives et efficaces.



DOCUMENT N°23: Modèle triarchique de l'intelligence
(Sternberg, 1994)

Avec l'aimable autorisation des E.A.P.

Enfin, au niveau d'adaptabilité le plus général, on peut lire les résultats avec les concepts propres à Canguilhem. On peut dire que, dans la situation T.E.D.E., un bon potentiel d'apprentissage traduit une bonne santé mentale, c'est-à-dire une forme d'équilibration entre la normalité (adhésion coalescente à

la méthode apportée par l'expert, à la norme socialement reconnue) ; et la normativité (perception articulée des choses, capacité de se créer ses propres normes). Réussir à progresser dans les trois niveaux verticaux de complexité et dans les trois niveaux horizontaux de transfert nécessite d'intégrer correctement la méthode externe, mais aussi de l'adapter, c'est-à-dire de créer de nouvelles normes, pour dépasser les variations verticales et horizontales. "Le normal, en matière biologique, ce n'est pas tant la forme ancienne que la forme nouvelle, si elle trouve les conditions d'existence dans lesquelles elle paraîtra normative, c'est-à-dire déclassant toutes les formes passées, dépassées et peut-être bientôt trépassées". (Canguilhem, 1979).

D'autres lectures, plus partielles, sont encore possibles en termes de dépendance-indépendance à l'égard du champ au vu de difficultés spécifiques à résoudre les items où des chemins superposés apparaissent au sujet comme définitivement embrouillés. L'évolution du contrôle du sujet sur sa propre activité se met en évidence dans l'évolution de l'étendue de la dispersion aux pré et post tests, ainsi que dans la variation des réussites et des échecs d'une séance à l'autre. La comparaison entre la qualité des apprentissages réalisés pendant la séance consacrée à l'entraînement et le degré de réussite au post-test donne un indicateur du maintien de l'apprentissage dans le temps. Les autres aspects métacognitifs inhérents au processus d'apprentissage, à sa conduite et à sa régulation, peuvent être provoqués, observés et enregistrés à l'occasion des interactions duelles examinateur-sujet, par exemple quand l'examineur demande au sujet d'explicitier sa manière de procéder pour apprendre à appliquer le mode opératoire.

Il reste possible d'intégrer ces différentes facettes de l'analyse du potentiel d'apprentissage en voie de concrétisation dans la situation de test dans un modèle général comme celui proposé par Sternberg : le modèle triarchique de l'intelligence (voir document n°20). "La faculté de faire face à de nouvelles tâches et à de nouvelles situations est donc une des facettes très importantes de l'intelligence." (Sternberg, 1994). Ce que Sternberg dénomme composantes d'exécution se retrouve

essentiellement au cours des séquences pré-test et post-test alors que les composantes d'acquisition de savoirs auront la priorité à l'occasion de la séquence d'apprentissage. Les métacomposantes seront particulièrement stimulées au cours des interactions duelles et de l'entretien de restitution. Les autres dimensions du modèle de Sternberg telles que le degré de familiarité avec la tâche, l'intégration de la nouveauté et l'autonomisation, l'autogestion mentale dans ses trois fonctions d'adaptation au milieu, de sélection et d'organisation du milieu environnant peuvent se montrer pertinentes pour décrire et interpréter les conduites d'apprentissage et de transfert mises en oeuvre par le sujet à la recherche d'une issue dans la situation T.E.D.E.

Dans ce cadre interprétatif, la question de l'interprétation des interférences, relativement nombreuses puisqu'elles concernent environ 15% des sujets dans la modalité de passation guidée, mérite un instant de réflexion : comment comprendre que la mise en situation d'apprentissage provoque une chute de l'efficacité ? Le taux d'interférences croît avec le degré de complexité et le gradient de transfert (voir tableau n°6, même corpus que tab. n°4). Plusieurs cas de figures sont à envisager. Le sujet d'emblée performant et à l'aise dans la tâche, s'il ne voit pas l'intérêt de parfaire sa technique afin d'améliorer sa performance, n'investira pas l'apprentissage et se démobilisera au post-test. A l'autre extrême, pour un sujet éprouvant des difficultés importantes pour entrer dans la tâche, l'apprentissage représente une surcharge cognitive plus qu'une aide. La performance d'un sujet hyper-réflexif peut chuter au post-test, et ce même sujet peut réussir les items au cours de l'entretien de restitution : au post-test, notre sujet se trouvait entre deux modes de fonctionnement, l'ancien en voie de déstructuration, et le nouveau en voie de structuration. Enfin, d'autres effets peuvent jouer qui relèvent de la fatigabilité, de la démotivation, de la labilité de l'attention, des attitudes de bâclage ou d'abandon...

Les données qualitatives et les interprétations interne et externe sont autant de sources d'informations qui pourront alimenter la rédaction d'une synthèse générale du cas, synthèse susceptible d'apporter réponse à la problématique de départ. Ces

informations serviront éventuellement à rédiger des préconisations en termes de remédiation cognitive (Cf. démarche éducativité professionnelle, *supra*).

TABLEAU 6: T.E.D.E., variation du taux d'interférence en fonction du degré de complexité et du gradient de transfert

niveau de complexité gradient de transfert	$a = b + c$	$a + b = c + d$	$a - b = c + d$	ensemble
identique	6,14%	11,66%	28,12%	11,16%
semblable	13,63%	16,12%	24%	15,97%
différent	15,38%	27,77%	57,14%	22,07%
ensemble	10,62%	15,59%	29,68%	14,75%

Quel que soit le type de synthèse rédigée, la relativisation des résultats quantitatifs par les éléments qualitatifs en constituera le corps. Relativiser une mesure ne signifie pas la remettre en cause : si les conditions de passation sont respectées, la mesure est un constat objectif fiable et solide qui va servir de base à l'interprétation. Relativiser le résultat d'une mesure revient à lui donner de la consistance en le contextualisant. Relativiser aide à éviter la dérive observée dans l'usage du Q.I., dérive vers la réification de la personne par le moyen du nombre. Enfin, la contextualisation va donner du sens à la performance ; le T.E.D.E. n'échappe pas à cette règle fondamentale de l'interprétation des mesures. Dans leurs grandes lignes, ces règles interprétatives ont été conservées pour les versions ultérieures du T.E.D.E.

Chapitre 11

Les différentes versions du T.E.D.E.

La version originale du T.E.D.E. de 1989 a été révisée en 1993, principalement par l'introduction d'un logiciel de traitement des résultats fonctionnant sous DOS et intégrant deux étalonnages selon les niveaux de qualification ainsi que la possibilité de construire des profils de groupe, ou de situer un sujet particulier par rapport à son groupe.

Le *T.E.D.E.98*© constitue la troisième édition de l'épreuve, édition cette fois-ci entièrement refondue. Cette refonte a donc pris en compte 9 années de pratiques et d'études menées sur les résultats obtenus. Au niveau de la pratique, les utilisateurs ont pu reprocher au test une certaine lourdeur liée à un temps de passation jugé excessif (3 séances de 1h30). Une deuxième critique a porté sur le fait d'imposer un mode opératoire unique comme moyen d'apprentissage de la tâche. Si cette modalité semblait adaptée aux milieux industriels, une extension du champ d'application à des cadres moins contraignants comme le bilan de compétences nécessitait d'abandonner une telle contrainte

La refonte a donc visé prioritairement un allègement de la procédure de passation. Encore fallait-il qu'un tel allègement puisse être effectué sans porter atteinte aux qualités métrologiques de l'épreuve. C'est à ce niveau que les études menées sur les résultats se sont montrées les plus utiles. Deux constats majeurs ont pu être établis : c'est le score après apprentissage qui présente la meilleure validité prédictive ; c'est ce même score qui est le plus saturé en facteur général de l'intelligence. On peut d'ailleurs raisonnablement penser que le second constat rend compte du premier dans une certaine mesure.

A partir de là est venue l'idée d'un nouveau paradigme de passation : apprentissage-test. On gagne ainsi le temps d'une séquence de passation par rapport au paradigme original : prétest-apprentissage-post test. La justification pragmatique est simple : à quoi bon maintenir une première séquence, le prétest, qui apporte des informations moins fiables que la troisième, le post test ?

Une première série de passations conduites selon ce nouveau paradigme a montré que la nouvelle procédure était jouable, et ce de deux points de vue. Côté gestion du temps par l'examineur, les deux séquences pouvaient tenir dans une demi-journée et du temps était libéré pour travailler plus en profondeur les aspects cliniques au cours de l'entretien de restitution. Côté confort de la passation pour les bénéficiaires, les sujets se plaignaient moins de la longueur de l'épreuve.

Un certain nombre de blocages face à l'épreuve disparaissaient, principalement ceux qui relevaient d'un conflit entre le mode de résolution spontanée qui s'installait au cours de la première séance de résolution en situation d'autonomie et le mode opératoire imposé au cours de la séance d'entraînement. De plus, le fait d'entrer dans le processus évaluatif par une séquence d'apprentissage diminue sensiblement la charge émotionnelle et le stress inhérents aux situations de test.

Par contre, au niveau des scores, on perdait les notes différentielles qui traduisaient l'évolution entre les deux tests suite à l'apprentissage. Ce n'est pas dramatique dans la mesure où le score après apprentissage est le plus solide, ce qui s'explique par le fait qu'il synthétise au mieux à la fois les capacités du sujet avant la passation et l'effet spécifique de l'apprentissage sur ces capacités.

Restait à explorer un autre gain de temps, par rapport au nombre d'items cette fois-ci. En reprenant les résultats engrangés par une démarche d'élimination pas à pas, on est parvenu à la conclusion qu'il était possible de n'utiliser que la moitié des items sans perte notable de fidélité par rapport au score établi sur l'ensemble.

A partir de là, les 36 items originaux furent répartis en deux groupes de 18 chacun, de niveau de difficulté équivalent, ce qui a permis la fabrication de deux formes parallèles, la forme A et la forme B. Les résultats obtenus à l'une ou à l'autre de ces deux formes sont pratiquement équivalents. La forme B peut s'utiliser en cas de nécessité de retest, par exemple dans le cadre d'un dispositif d'évaluation des effets d'une action de remédiation cognitive.

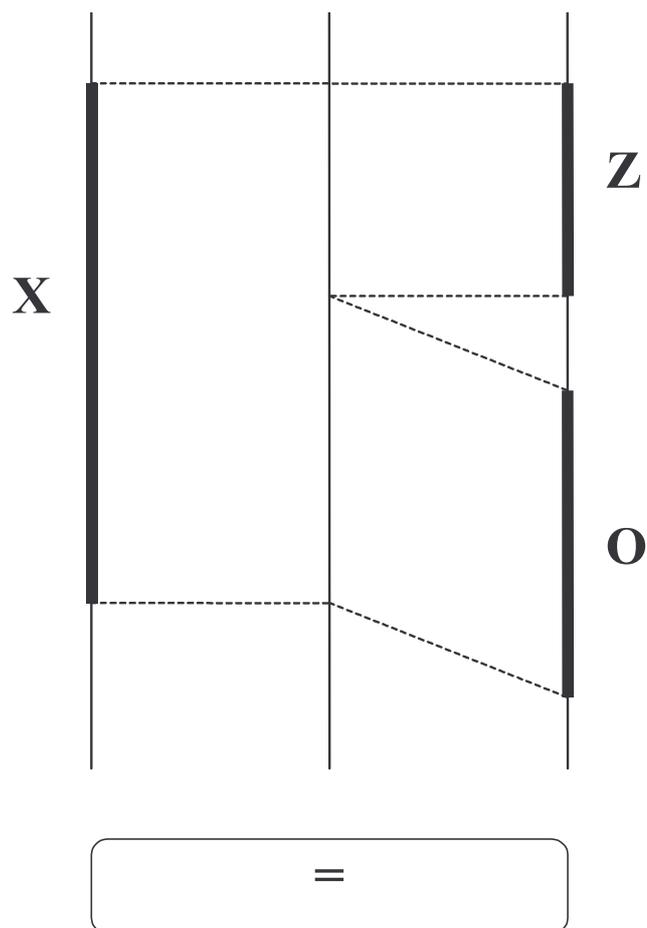
En résumé, on peut dire que la refonte complète de l'épreuve a permis de réduire la durée de passation à deux heures de temps sans perte majeure des qualités métrologiques de l'épreuve tout en augmentant le confort des usagers, examinateur et sujet.

Sur le plan du matériel, le graphisme des items a été revu en format image de *Word* (voir document n° 24) ce qui a permis d'améliorer sensiblement leur lisibilité. Après correction manuelle, les résultats bruts sont saisis dans une feuille d'Excel et leur traitement se fait automatiquement. La sortie papier comporte trois pages. La première supporte les éléments de synthèse, la seconde un visuel facilitant le positionnement de la performance pour l'ensemble de l'échantillon de 424 sujets et pour chacun des niveaux de qualification et la troisième un profil fonctionnel des différents indices partiels : vitesse, réussite pour les seuls égalités et les seuls déplacements (équilibre de la tâche), pour chacune des trois séries (gradient de complexité) et pour les items identiques, semblables et différents (gradient de transfert).

Les modalités de la séquence d'apprentissage ont évolué vers une autopassation assistée, un *Dossier d'instructions* comportant les explications nécessaires à la bonne exécution de la tâche ainsi que les corrigés des exercices du *Livret pour l'apprentissage*. En cas d'erreur, une page parallèle permet de reprendre l'apprentissage de ces exercices.

Sur le plan du matériel, le graphisme des items a été revu en format image de *Word* (voir document n° 24) ce qui a permis d'améliorer sensiblement leur lisibilité. Après correction manuelle, les résultats bruts sont saisis dans une feuille d'Excel

et leur traitement se fait automatiquement. La sortie papier comporte trois pages.



Document 24 : Nouvelle présentation des items du T.E.D.E.

La première supporte les éléments de synthèse, la seconde un visuel facilitant le positionnement de la performance pour l'ensemble de l'échantillon de 424 sujets et

pour chacun des niveaux de qualification et la troisième un profil fonctionnel des différents indices partiels : vitesse, réussite pour les seuls égalités et les seuls déplacements (équilibre de la tâche), pour chacune des trois séries (gradient de complexité) et pour les items identiques, semblables et différents (gradient de transfert).

T.E.D.E. 2000 ©

Résultats

Nom: TASM Prénom: Luce
 Âge: 24 ans Sexe: féminin
 Niveau: IV, bac... Date: 19/11/2001

Objectif de la passation:

Bilan de compétences

Positionnement:

PA
118

m= 100
 σ = 15
 N= 589

Limites de confiance	
supérieure	124
inférieure	112

(pour 90% des cas)

niveau	classe
III	4
IV	4
V	4
VI	-

(5 classes)

Synthèse:

Le potentiel d'apprentissage obtenu se situe dans la zone élevée.

Par rapport au niveau de qualification actuel, ce potentiel renvoie la zone élevée.

Relativement au sexe, à l'âge, et au niveau de qualification actuel combinés, cette performance se classe en zone élevée

Document 25 : page 1 des résultats du T.E.D.E 2000

Les modalités de la séquence d'apprentissage ont évolué vers une autopathologie assistée, un *Dossier d'instructions*

comportant les explications nécessaires à la bonne exécution de la tâche ainsi que les corrigés des exercices du *Livret pour l'apprentissage*. En cas d'erreur, une page parallèle permet de reprendre l'apprentissage de ces exercices.

Cette version du T.E.D.E.98 a été améliorée en 2000 par un renforcement de l'échantillon passant à 589 sujets. Cette dernière mouture, le T.E.D.E.2000 introduit de nombreuses informations complémentaires au niveau de la sortie des résultats.

La page 1 indique les limites de confiance inférieure et supérieure qui encadre la valeur du PA à $p=0,90$, ce qui signifie qu'il y a 90 chances sur 100 que la valeur vraie du PA se situe entre les limites indiquées. Le cadre "Synthèse" donne automatiquement les positionnements par rapport à l'ensemble de l'échantillon, par rapport au niveau de qualification actuel du sujet et par rapport à ses caractéristiques différentielles combinées (voir document n°25).

La page 2 (voir document n°26) est un visuel qui présente les positionnements du PA obtenu globalement (colonne de gauche) et par rapport aux différents niveaux de qualification. Cette présentation aide à la réflexion sur les stratégies de passage d'un niveau à un autre : passage horizontal ou bien passage en escalier via une phase préparatoire de ré-entraînement cognitif.

La page 3 permet de comparer à l'aide du "Profil fonctionnel" (voir document n°27) les sous-scores d'un sujet au niveau global afin de déterminer les éventuels points forts ou points faibles. Cette lecture du profil peut donner des éléments de compréhension de la construction du PA ou de ses limitations et déboucher sur la formulation de conseils en termes de développement personnel.

La page 4 offre un ensemble d'analyses complémentaires qui prennent en compte les données de la séquence d'apprentissage. Le "Profil apprentissage" (voir document n°28) décrit le comportement de l'apprenant dans le choix de ses réponses et le "Profil transfert" de quelle manière les acquis de l'apprentissage sont rédupliqués au cours de la séquence de test (voir document n°29)..

Légende:

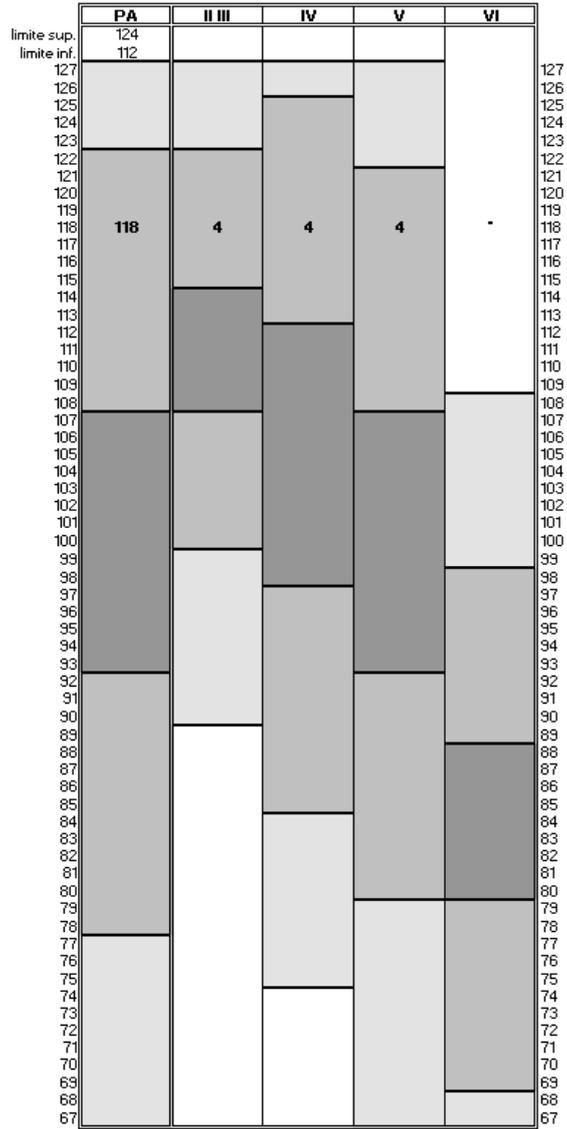
— limite des classes
 - valeur non observée
 ■ classe médiane
1, 2, 3, 4 ou 5
 numéro de la classe;
 sa position indique la
 place de la performance
 dans la classe

Effectifs théoriques:

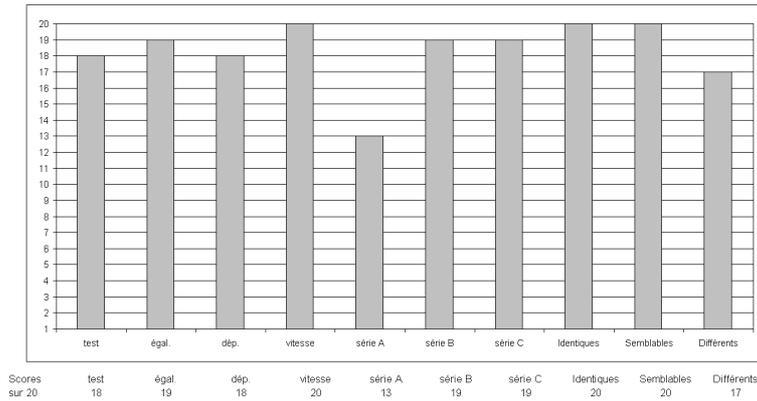
classe	n
5	6,70%
4	24,20%
3	38,20%
2	24,20%
1	6,70%

classe	n cumulés décroissants
5	100,00%
4	93,30%
3	69,10%
2	30,90%
1	6,70%

classe	n cumulés croissants
5	6,70%
4	30,90%
3	69,10%
2	93,30%
1	100,00%



Profil fonctionnel test:



Document 27 : page 3 des résultats du T.E.D.E 2000.

Sur cette page, on trouve également une présentation du PA sous forme de percentilage selon les facteurs différentiels de niveau, d'âge et de sexe où des trois combinés. En réponse à des demandes émanant d'établissements spécialisés, le T.E.D.E.-A ou T.E.D.E.-Adapté est une variante de l'épreuve originale utilisée avec les publics tout-venant, variante destinée aux publics dits spécifiques : illettrés, handicapés, désavantagés, bas niveaux... fréquentant ou non les établissements spécialisés.

Profil apprentissage:

apprentissage	réussite spontanée	erreur bien corrigée	erreur illusoire	interférence	erreur mal corrigée	erreur non traitée	omission
a1	1						
a2	1						
a3	1						
a4	1						
b1					1		
b2	1						
b3	1						
b4	1						
c1	1						
c2		1					
c3	1						
c4						1	
Total	9	1	0	0	1	1	0
%	75,00%	8,33%	0,00%	0,00%	8,33%	8,33%	0,00%

Document 28 : page 4 des résultats du T.E.D.E 2000.

L'adaptation réalisée permet une entrée et une évolution dans la tâche d'un maximum de sujets dans la mesure où l'essentiel du temps reste consacré à l'activité d'apprentissage.

Profil transfert:

transfert	réussite réitérée	gain	perte	échec réitéré	omission réitérée
a1	1				
a2	1				
a3	1				
a4	1				
b1		1			
b2	1				
b3	1				
b4	1				
c1	1				
c2	1				
c3	1				
c4		1			
Total	10	2	0	0	0
%	83,33%	16,67%	0,00%	0,00%	0,00%

Document 29 : page 4 des résultats du T.E.D.E 2000.

L'objectif de l'épreuve a été recentré sur l'identification des sujets capables, dans une situation d'apprentissage didactique, de se construire des acquis et de les transférer sur un gradient de transfert relatif aux contenus et sur un autre gradient relatif au temps. Les aspects fonctionnels et diagnostiques conservent leur valeur heuristique afin de générer toutes conjectures utiles et nécessaires à la détermination d'objectifs de remédiation.

Concrètement, seule la première série d'exercices est travaillée. Un seul item est présenté par page, ce qui permet une parfaite lisibilité et freine le "papillonnage" d'un item à l'autre. La passation est menée par groupe de 4 sujets. Une première séquence est consacrée à l'entraînement sur le mode du conflit socio-cognitif. Elle est suivie d'un test immédiat puis d'un retest à une semaine. Ce paradigme permet d'apprécier la tenue dans le temps de l'apprentissage et de différencier entre les sujets dont l'apprentissage tient et ceux dont l'apprentissage s'estompe en l'absence d'aide extérieure. Devant la diversité des publics spécifiques, il était illusoire de construire un étalonnage d'ensemble. C'est pourquoi un logiciel de traitement des résultats permet à l'utilisateur de construire son ou ses

étalonnages à partir de ses propres données recueillies auprès des publics avec lesquels il travaille.

PROFIL DIFFERENTIEL

-SUJET- Nom: DUPOND Date: 01/01/01	Prénom: Jacques Groupe: B																				
	-3σ			-2σ			-1σ			m		+1σ			+2σ			+3σ			
	(-)0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20(+)
NA										X											
NI													X								
ND													X								
<hr/>																					
δ												X									
<hr/>																					
III													X								
IID														X							
<hr/>																					
ISI													X								
ISD													X								
<hr/>																					
IDI									X												
IDD												X									
<hr/>																					
EA												X									
EI												X									
ED														X							
<hr/>																					
DA										X											
DI												X									
DD												X									

Document 30 : page des résultats du T.E.D.E Adapté.

Un profil présente l'ensemble des scores et des sous-scores, y compris la tenue dans le temps de l'apprentissage (voir document n°29). Une grille d'analyse des données quantitatives et qualitatives recueillies permet la formulation d'objectifs de remédiation et le choix qui semble le mieux adapté parmi la progression des exercices des A.S.L.O.S. et du P.E.I.F.

L'avant-dernière version du T.E.D.E. concerne les C.F.A. et les apprentis. Le T.E.D.E.-C.F.A. conserve les différents livrets, paradigme, consignes de passation et items du T.E.D.E.2000. Seuls les étalonnages et la présentation des résultats diffèrent.

L'étalonnage a été réalisé sur un échantillon de plus de 600 apprentis à leur entrée au C.F.A. Le PA est donné pour l'ensemble des apprentis de l'échantillon ainsi qu'un classement sur 100 pour chacun des cursus : C.A.P., B.E.P., B.P., Bac Pro.

S'y ajoutent les chances de réussir l'examen. Un visuel permet de résumer ces informations et de les percevoir globalement (voir document n°30).

Résultats

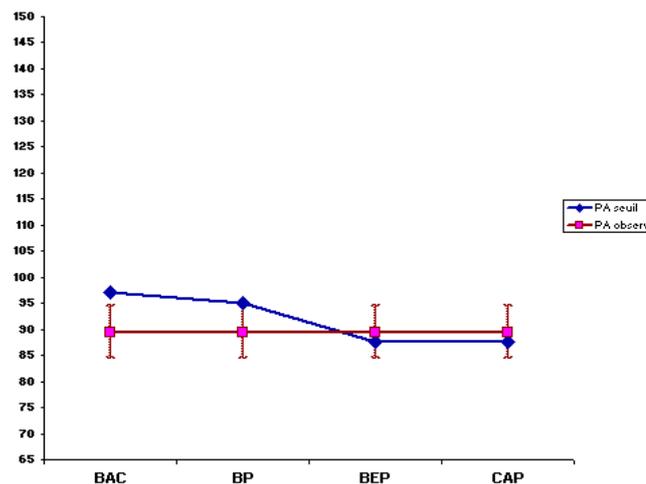
Nom: BIN Prénom: Ret
 Âge: 17 ans Sexe: 2
 Date: 19/11/2001
 Niveau préparé: CAP Branche: vente

Objectif de la passation:

Positionnement: PA = limite sup. limite inf.

	Ensemble	BAC	BP	BEP	CAP
rang/100	75	94	89	72	70
niveau	bas	très bas	bas	bas	bas
pronostic	*	56%	56%	72%	72%
n	667	76	71	183	336
m	100,00	107,50	107,41	100,44	96,49
σ	15,00	15,18	16,68	16,03	12,69

Pronostic ($p=0,20$):



Les pages suivantes présentent un profil fonctionnel afin de déterminer les points forts ou les points faibles éventuels au niveau des sous-scores du test et on retrouve également les tableaux relatifs aux profils d'apprentissage et de transfert.

La dernière version éditée par les E.A.P. rassemble dans un même ensemble T.E.D.E.6 la version pour les adultes et celle pour les apprentis. Les étalonnages ont été renforcés et des études de validation approfondies y sont développées.

Chapitre 12

Les qualités métrologiques du T.E.D.E.

Différentes études des qualités métrologiques du T.E.D.E. dans ses versions de 1989 et 1993 ont été entreprises dont certaines ont donné lieu à publications (Pasquier, 1991d, 1994a, 1994b). Ces études concernent souvent de petits effectifs dans la mesure où elles accompagnent des pratiques sociales d'évaluation du potentiel d'apprentissage. Ces études ont donc une double finalité : asseoir progressivement, par la répétition des résultats, les qualités métrologiques du T.E.D.E. ; évaluer, pour les améliorer, les pratiques de terrain.

La sensibilité ou finesse discriminative de l'épreuve est assurée par une lecture des résultats sur trois échelles discriminatives. Les notes brutes, qui indiquent le niveau de performance au pré-test et au post-test, varient de 0 à 36 points, et les indices de gain relatif de -100% à +100%. Les décilages des notes et indices bruts offrent une échelle de notes standards en dix échelons représentant chacun 10% de la population. Enfin, pour regarder les résultats avec plus de distance, il suffit de considérer les trois zones du profil : classe A pour les 30% supérieurs, classe B pour les 40% médians, zone C pour les 30% inférieurs (voir tableau n°7).

La fidélité du T.E.D.E. a été évaluée par la méthode des items pairs-impairs. Deux valeurs de fidélité ont été jusque maintenant obtenues en ce qui concerne l'indice de gain relatif sur l'ensemble de l'épreuve. La première, pour un groupe de neuf ouvrières ayant suivi la passation guidée, prend la valeur 0,90. La seconde, menée sur les protocoles d'un groupe de trente quatre stagiaires en formation professionnelle, prend la valeur 0,88 (coefficient ρ de Spearman corrigé par la formule Spearman-Brown). D'autre part, une étude selon la méthode test-retest à 7 mois d'intervalle sur un groupe de 13 sujets a

fourni les corrélations suivantes : 0,80** pour le score au pré-test ; 0,92** pour le score au retest ; 0,56* pour le gain relatif. (A titre de comparaison, pour ce même groupe de sujets, la fidélité du test *Mécanique* fut de 0,81** et celle d'un Q.C.M. de connaissances professionnelles de 0,86**).

TABLEAU 7: T.E.D.E., sensibilité

NOTES BRUTES:										
-> pour les niveaux de performance: 37 échelons										
0 point					18 points					36 points
-> pour les gains relatifs: 201 échelons										
- 100%					0%					+ 100%

DECILAGE: 10 échelons										
1	2	3	4	5	6	7	8	9	10	

3 ZONES:										
- (30%)						= (40%)				+ (30%)

La validité de contenu a été largement abordée dans le chapitre précédent, au niveau de la justification des choix opérés pour construire l'épreuve, selon un principe d'homomorphie qui relie le paramétrage des critères de l'épreuve et les critères de la tenue d'un poste de travail, en entreprise ou en formation.

Quatre études de validité concourante ont été menées en milieu industriel. Dans les deux premières, on met en relation le potentiel d'apprentissage et une évaluation des opérateurs réalisée par leur encadrement. Dans la troisième, une batterie de quatre tests, dont le T.E.D.E., a été donnée à un même groupe d'agents, ce qui a permis la réalisation de l'analyse factorielle des résultats.

La première étude, dans la phase de mise au point du T.E.D.E., porte sur neuf ouvrières, âgées de vingt-cinq à cinquante ans, travaillant sur des tours semi-automatiques. L'intervention a été présentée comme une expérience par laquelle "On teste le test", composé alors seulement de deux niveaux de complexité ($a = b + c$ et $a + b = c + d$), soit vingt-quatre items. Un entretien de restitution a été organisé suite à la passation dans un double but : communiquer les résultats à chacune des intéressées et enregistrer leur vécu de l'épreuve.

TABLEAU 8: T.E.D.E., validité concourante, opératrices sur tours semi-automatiques

T.E.D.E. critère	réussite spontanée (ST1)	gain relatif (PAT)
classement entreprise (CE)	0,49 N.S.	0,67 S.

D'une manière générale, le T.E.D.E. a été jugé intéressant par les opératrices, certaines allant jusqu'à expliquer que ce genre de travail avait stimulé chez elles le besoin de reprendre une activité de nature intellectuelle. Ce besoin est formulé par contraste avec le travail manuel quotidien. La nouveauté de la tâche a provoqué quelques effets de surprise et a suscité de l'intérêt. La démarche d'apprentissage est bien perçue et parfois directement associée aux souvenirs scolaires. L'aspect visuel du test est souligné, et les items identifiés comme les plus ardues à résoudre comportent "trop de pointillés", ce qui renvoie à la question de la dépendance-indépendance à l'égard du champ. La nécessité de maintenir une attention soutenue comme condition de l'efficacité est également pointée.

TABEAU 9: T.E.D.E., validité concourante, item pour établir le classement entreprise

CHAPITRE V: efficacité quantitative et organisationnelle

- * l'agent produit les quantités demandées dans le temps imparti
- * l'agent ne dépasse pas la quantité de déchets obligée
- * l'agent intervient immédiatement face à un constat de défaut
- * l'agent planifie sa tâche
- * l'agent aménage son environnement
- * l'agent respecte les règles d'hygiène et de sécurité
- * l'agent régule son activité pendant la tâche
- * l'agent s'estime à sa juste valeur

CHAPITRE VI: initiative en réponse aux dysfonctionnements

- * l'agent diagnostique la nature du dysfonctionnement
- * l'agent évalue les conséquences du dysfonctionnement
- * l'agent apporte une réponse adaptée au dysfonctionnement

CHAPITRE VII: compréhension des consignes

- * l'agent comprend les consignes verbales
- * l'agent comprend les consignes écrites
- * l'agent comprend les plans, les symboles et les abréviations
- * l'agent comprend les consignes complexes

CHAPITRE VIII: apprentissage de nouveaux savoirs et savoir-faire

- * l'agent assimile rapidement
- * l'agent réalise un apprentissage durable
- * l'agent utilise les nouveaux acquis d'une façon autonome dans des délais raisonnables
- * l'agent utilise ses acquis antérieurs pour faciliter ses apprentissages et leurs adaptations (transfert)

Parallèlement à la passation du T.E.D.E., le chef d'atelier et les agents de maîtrise concernés ont évalué les

opératrices sur le critère "rapidité et efficacité dans le lancement de la fabrication de nouvelles séries".

Chacun des évaluateurs élaborera son classement de manière indépendante, sans aucune concertation avec ses autres collègues. Le classement entreprise (C.E.) fut construit par cumul des estimations des évaluateurs. Le tableau de corrélations (voir tableau n°8) met en relation le critère adaptabilité professionnelle évalué par l'entreprise (C.E.) et les deux classements obtenus à partir du T.E.D.E., à savoir le degré de réussite spontanée au pré-test (ST1), et le gain de performance relatif suite à l'apprentissage (PAT). Les deux corrélations sont élevées, le PAT seul franchissant le seuil de significativité à $p = 0,05$.

L'étude suivante porte sur un groupe de cinquante-deux ouvriers spécialisés dans les métiers de l'extrusion et de la flexographie, O.S. considérés par leur hiérarchie comme les moins aptes à acquérir une qualification professionnelle. Ces ouvriers ont passé le T.E.D.E. en passation guidée. La direction de l'entreprise développant un plan de professionnalisation des salariés non-qualifiés, il fut prévu que ces O.S. suivraient une action de remédiation cognitive avec les supports E.D.I. (Pasquier, 1990b, 1990c, 1990d, 1991a) avant d'aborder les modules de formation professionnelle. Afin d'évaluer les effets de cette phase de réentraînement intellectuel, une évaluation portant sur l'évolution des capacités et des comportements des O.S. concernés fut organisée avant et après la pré-formation.

Pour les besoins de cette étude de validité concurrente, seuls les résultats élaborés avant l'entrée en pré-formation furent pris en compte. L'évaluation a été menée par le biais d'un questionnaire des agents de maîtrise et des chefs d'ateliers, questionnaire conduit de manière aveugle par un cadre de l'établissement : ce cadre ne connaissait pas l'ouvrier évalué afin d'éviter toute influence sur le contenu des réponses. L'agent de maîtrise était invité à répondre en premier et le chef d'atelier n'intervenait que lorsqu'il pensait que le jugement formulé nécessitait un ajustement. Dans cette situation de désaccord, le cadre invitait les deux évaluateurs à rapprocher leurs points de vue par échanges d'arguments.

La présente étude ne prend en compte que les dix-neuf items ayant un rapport direct avec la tenue du poste de travail (voir tableau n°9). A chaque item, l'agent de maîtrise répondait sur une échelle de fréquence d'occurrence à quatre degrés : rarement, parfois, souvent, à chaque fois. On a ainsi construit une note C.E. sur le critère "tenue du poste de travail", variant de 4 à 76 points. Cette note est mise en corrélation avec le potentiel d'apprentissage. Malgré le fait que seul l'effectif de plus faible niveau de l'entreprise soit pris en compte, la corrélation de 0,46 est très significative ($p = 0,01$) : la qualité dans la tenue du poste de travail est ici statistiquement liée au potentiel d'apprentissage exprimé par le gain relatif (PAT).

Dans la troisième étude de validité concurrente sont mis en relation les résultats donnés par le T.E.D.E. et trois autres tests auprès d'un même groupe de sujets. Cette étude a pour objectif de répondre à deux questions : le score ST1 de réussite spontanée au pré-test est-il un bon indicateur d'intelligence générale ? le potentiel d'apprentissage évalué par le PAT est-il une variable relativement spécifique ? Elle est dérivée d'une action menée dans une entreprise de taille nationale.

La direction des ressources humaines de cette entreprise organisait un mouvement de mobilité interne passant par une reconversion professionnelle. Des personnels occupant des postes destinés à disparaître furent invités à se porter candidats pour se former à un nouveau métier en phase de développement. Le dispositif de reconversion prévoyait un appel à candidatures internes suivi d'une évaluation-orientation dans le dispositif de formation. Ce dispositif permettait plusieurs parcours. Selon les résultats de l'évaluation, le candidat pouvait accéder soit à une formation d'ouvrier professionnel, soit à une remise à niveau IV, soit directement en formation de technicien.

L'évaluation portait sur quatre dimensions : les connaissances, les aptitudes, les motivations, les conduites sociales. Il y eut vingt-deux candidats à la reconversion. Quatre épreuves relatives à la mesure des capacités intellectuelles furent données aux candidats. Le *Test mécanique* (Rennes,

1952), en abrégé le M, évalue le bon sens pratique à travers la résolution de problèmes mettant en jeu des représentations d'objets physiques. Le D48 (Pichot, 1965), évalue le facteur G à travers des items de type séries à compléter présentées sous forme de suites de dominos. Une épreuve de logique propositionnelle (Voge), en abrégé LP, apprécie la capacité à raisonner sur les relations syntaxiques internes à des propositions verbales, indépendamment des contenus et réalités évoquées. Le T.E.D.E. (Parquier, 1989c) dont deux indices sont ici pris en compte : le niveau de réussite spontanée au pré-test (ST1) et le degré de réactivité à une mise en situation d'apprentissage mesurée par le gain relatif de la performance (FAT). Les caractéristiques différentielles du groupe de candidats sont les suivantes : n = 22 dont 1 femme et 21 hommes, âgés de 23 à 50 ans ; et dont 10 niveaux V, 9 niveaux IV, 3 niveaux III.

TABEAU 10: T.E.D.E., validité concurrente, matrice de corrélations

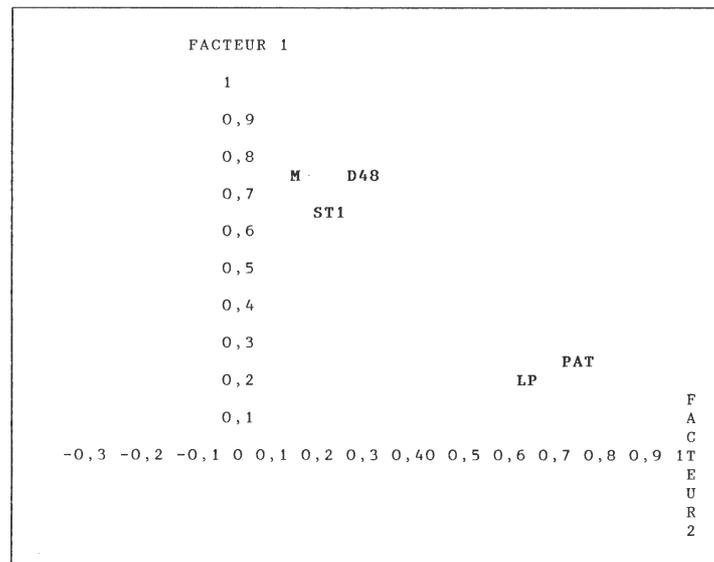
M	D48	LP	ST1	PAT	AGE
M	0,60**	0,22	0,56**	0,32	0,21
D48		0,26	0,58**	0,56**	-0,29
LP			0,40	0,56**	-0,35*
ST1				0,30	-0,32
PAT					-0,35*
AGE					

* significatif à p 0,05
 ** très significatif à p 0,01

En calculant deux à deux les corrélations entre les cinq notes psychométriques et l'âge considéré comme variable supplémentaire on obtient une matrice de corrélations (voir tableau n°10). Les corrélations entre les indices

psychométriques sont toutes positives, mais pas toutes significatives : il est donc probable que les épreuves s'agrègent par le facteur G, tout en conservant chacune une certaine spécificité. La relative indépendance du ST1 et du PAT montre que ces deux variables ne sont pas trop redondantes, et qu'en conséquence elles apportent des informations complémentaires. On peut dire que le T.E.D.E. est validé par des épreuves classiques ayant largement fait leurs preuves dans le domaine de la psychologie du travail. Quand on regarde l'influence de l'âge, on constate que seul le *Test mécanique* résiste au processus d'involution.

TABLEAU 11: T.E.D.E., validité concourante, analyse factorielle



L'analyse factorielle en facteurs principaux confirme et complète l'analyse de la matrice. Avant rotation, on trouve un facteur général qui explique 80% de la communauté des résultats. Un second facteur regroupe ST1, D48 et M d'un côté et LP et PAT de l'autre (voir tableau n°11). On peut faire l'hypothèse que le rapprochement des indices LP et PAT vient

du fait que l'aide apportée au cours de la phase d'entraînement du T.E.D.E. est présentée sous la forme d'un mode opératoire qui décline une suite de propositions verbales. On ne peut pas dire *a priori* que ce rapprochement est lié au seul mode de codage de l'information ou bien s'il inclut également le raisonnement sur les articulations syntaxiques.

En conclusion de cette étude, il se confirme que le T.E.D.E. est bien corrélé avec d'autres épreuves qui mesurent des capacités intellectuelles. Le potentiel d'éducabilité mesuré par ce test présente une certaine spécificité par rapport à la réussite spontanée au pré-test. Cette relative spécificité serait à rapprocher des aspects logico-verbaux présents dans la phase d'apprentissage de l'application du mode opératoire. C'est là une conclusion provisoire dans la mesure où les résultats des analyses factorielles, pouvant fluctuer selon la technique utilisée, demandent toujours contrevalidation.

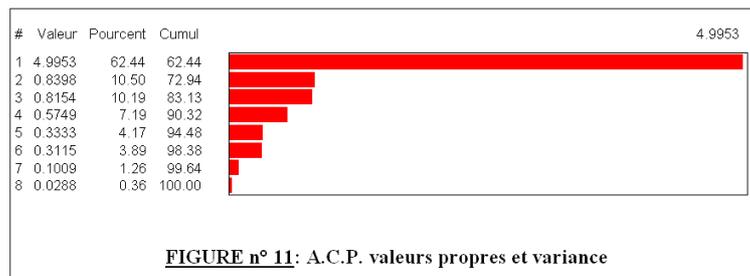
La quatrième étude de validité concourante a été menée dans une entreprise métallurgique d'emboutissage et de soudage qui produit des sous-ensembles de véhicules automobiles (Pasquier, 1996). Cette entreprise a robotisé une chaîne de fabrication. Un appel à candidature interne a été lancé auprès de l'ensemble du personnel. 88 candidats ont subi les épreuves de sélection précédant les modules de préformation et de formation. Les passations ont permis d'obtenir les indices suivants : notes au test *Mécanique* (M), à un test de reconnaissance de coupes de volumes (S) et au *T.E.D.E.* (ST1 , ST2, PAT, PA1 , PA2 et PA3).

Des variables individuelles ont été relevées dont on peut penser qu'elles aient une influence sur l'éducabilité cognitive : l'âge des sujets (AGE), leur niveau de formation professionnelle (CAP : aucune formation, CAP non réussi, CAP réussi), leur ancienneté dans l'entreprise (ANC). L'âge moyen des sujets est de 38 ans, le mode se situant à 41 /42 ans. 44,3 % ont réussi au C.A.P. ; 27,3 % ont échoué à l'examen ; 23,9 % n'ont pas de formation. Ces sujets ont une ancienneté moyenne dans l'entreprise de 13 ans environ, le mode étant à 17/18 ans.

Nous avons effectué une analyse en composantes principales avec le logiciel *STATlab* sur les indices

psychométriques, les variables AGE, CAP et ANC étant introduites *a posteriori* comme variables supplémentaires. La lecture des résultats et leur interprétation s'inspirent de G. Langouet et J.C. Porlier (1989, 3-II). L'interprétation portera essentiellement sur les facteurs afin de dégager l'éventuelle spécificité de l'évaluation dynamique.

L'examen des valeurs propres (voir figure n°11) amène, si l'on prend comme critère de choix l'explication d'au moins 80% de variance, à retenir un modèle à 3 facteurs : un facteur principal qui rend compte des 2/3 de la variance et 2 facteurs spécifiques rendant compte chacun de 10% de variance.



Les corrélations entre les composantes et les variables (voir tableau n° 15) indiquent que l'indice ST2 est le plus fortement saturé dans ce facteur général de performance cognitive.

La seconde composante est bipolaire ; elle oppose les indices statiques de niveau (S, M, ST1) et les indices dynamiques de progression (PA1, PA2, PA3, PAT), le ST2 occupant une position charnière. On trouve donc ici la confirmation d'une spécificité modérée des indices dynamiques par rapport aux indices statiques. Enfin, la troisième composante est également bipolaire. Ce facteur semble opposer le transfert direct (PA1) au transfert flexible (PA3), ST2 étant au centre. Le positionnement des variables supplémentaires confirme le lien négatif entre efficacité cognitive et âge ou ancienneté *a contrario* du niveau de formation (voir figure n°12).

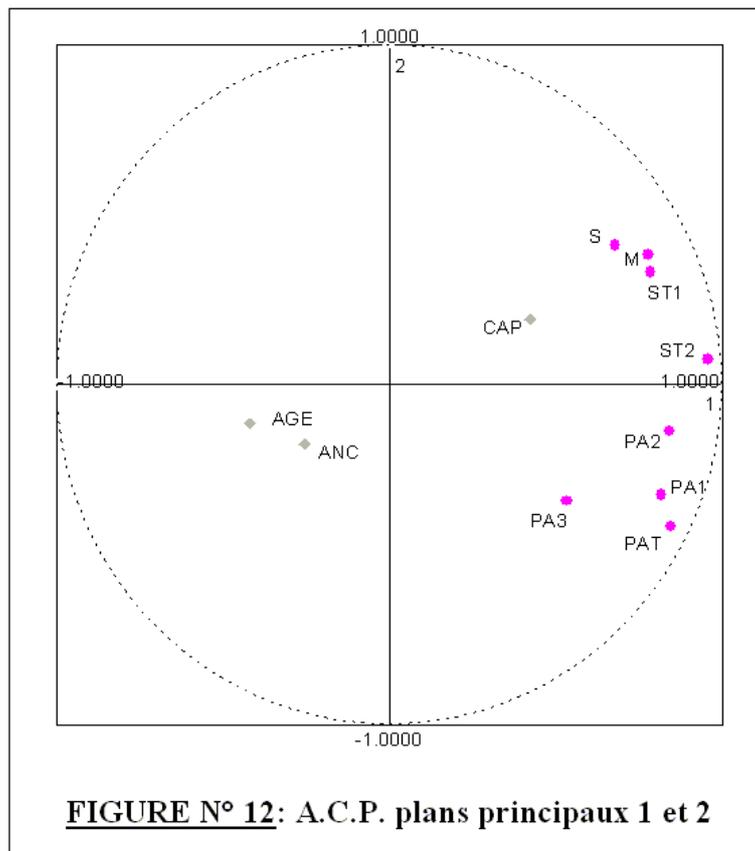
Ces résultats montrent qu'à côté d'un facteur général qui fixe bien l'épreuve dans le champ de l'évaluation cognitive, des informations complémentaires spécifiques sont apportées sur une dimension d'éducabilité et une autre de flexibilité. On accède ainsi à une meilleure connaissance du contenu additionnel des indices formels. L'interprétation des indices du T.E.D.E., en termes de niveau et en termes de réactivité à une mise en situation d'apprentissage, s'en trouve mieux précisée et gagne en consistance psychologique.

	Axe : 1		Axe : 2		Axe : 3	
	Coord	Cor	Coord	Cor	Coord	Cor
M	0.781	0.610	0.382	0.146	-0.132	0.017
S	0.680	0.462	0.410	0.168	0.221	0.049
ST1	0.785	0.616	0.334	0.111	-0.197	0.039
ST2	0.958	0.918	0.074	0.006	-0.087	0.008
PAT	0.849	0.721	-0.414	0.171	0.200	0.040
PA1	0.818	0.670	-0.321	0.103	0.353	0.125
PA2	0.844	0.712	-0.133	0.018	0.139	0.019
PA3	0.535	0.287	-0.342	0.117	-0.720	0.518

TABLEAU N°15 : A.C.P., coordonnées et corrélations

Deux études de validité prédictive sont actuellement disponibles. La première s'est déroulée durant la phase de mise au point du T.E.D.E. Menée à l'occasion d'un stage d'alphabétisation, elle ne porte que sur les douze items du premier niveau de complexité ($a = b + c$). Le stage regroupait sur une durée de quatre mois un effectif de quatorze femmes, âgées de trente à quarante-cinq ans, d'origine asiatique, essentiellement dans un objectif de perfectionnement en langue française. Afin d'éviter tout effet de réalisation automatique des attentes, l'épreuve ne sera corrigée qu'à l'issue du stage, après

que la formatrice ait procédé à l'évaluation pédagogique des effets de la formation.



Cette évaluation pédagogique a consisté à situer la qualité du déchiffrement de phrases et textes simples sur une échelle en cinq degrés : nul, un peu, moyen, bon, très bon. L'évaluation de la qualité du déchiffrement menée en début et en fin de formation a permis de distinguer trois groupes de stagiaires : celles qui n'ont pas progressé (G1), celles qui ont partiellement progressé (G2), celles qui ont pleinement atteint l'objectif de déchiffrement (G3).

Pour chaque groupe, on a calculé la moyenne des potentiels d'apprentissage, et on observe que cette moyenne varie dans le même sens que les évolutions pédagogiques (voir tableau n°12). Le rapport de corrélation prend une valeur élevée (0,83) et confirme la covariation du prédicteur potentiel d'apprentissage et du critère perfectionnement du déchiffrage.

TABLEAU 12: T.E.D.E., validité prédictive et activité lexicale de déchiffrage

groupe	G1	G2	G3	ensemble
effectif	4	5	5	14
PAT moyen	8,5	25,8	42,8	26,9
écart-type	12,2	30,3	44,9	33,7

La seconde étude de validité prédictive a pour cadre un centre de rééducation professionnelle des adultes reconnus travailleurs handicapés. Les stagiaires concernés, orientés par la CO.T.O.R.E.P. après sélection, suivaient une formation de technicien en électronique et en informatique industrielle (T.E.I.I.). Ils subirent une évaluation initiale composée du T.E.D.E. et d'une épreuve de mathématiques élaborée par les formateurs. Quatre mois plus tard, un examen partiel fut passé dans les matières suivantes : anglais, interrogations écrite et orale en électricité, épreuve pratique qui consistait à construire un circuit imprimé sur une plaque à partir d'un plan donné.

Les résultats de l'évaluation initiale furent restitués aux stagiaires et aux formateurs afin d'aménager au mieux une pratique d'aides pédagogiques et relationnelles individualisées. Cette restitution n'exclut pas un risque d'"effet Pygmalion" (Rosenthal, Jacobson, 1971), les formateurs pouvant avoir tendance, plus ou moins consciemment, à harmoniser leurs conduites et pratiques professionnelles dans le but que la prévision transmise par l'expert se réalise. Ce risque semble

toutefois limité, dans les situations où les enseignants ont eu le temps de se forger leur propre attente, avant le passage de l'évaluation psychométrique. Dans ce cas, "...la transmission des résultats d'un test prédictif aux enseignants constitue une situation qui ne peut se réduire à celle de l'effet Pygmalion puisque le taux de vérification des prédictions est insensible à la connaissance ou à l'ignorance des prédictions par le maître." (Pasquier, 1992).

Quoiqu'il en soit, pour quinze stagiaires étaient disponibles les résultats de l'évaluation initiale et ceux de l'examen partiel. Les résultats aberrants d'un sujet furent écartés : excellent au T.E.D.E., toutes ses notes de partiel sont médiocres, même en anglais alors qu'il est bilingue dans sa famille. Dans un premier temps, la note globale du partiel est mise en relation avec les trois notes de l'évaluation initiale : note en mathématiques (Math.), niveau de réussite au pré-test (STI) et réactivité à la mise en situation d'apprentissage didactique (PAT) (voir tableau n°13).

TABLEAU 13: T.E.D.E., validité prédictive et examen partiel de stage de formation de techniciens

prédicteurs	mathématique	STI	PAT
partiel	0,17 N.S.	0,27 N.S.	0,57 S.

Les trois corrélations sont positives. Toutefois, seule la liaison PAT-partiel dépasse le seuil de significativité statistique. Dans ce cas, le PAT est le meilleur prédicteur de la réussite des apprentissages techniques, et le niveau de mathématiques le moins bon.

Considérons maintenant les liaisons entre le PAT et les différentes matières du partiel (voir tableau n°14). Toutes les liaisons sont positives, mais d'inégales valeurs. La liaison PAT-anglais reste très faible. Il est vrai que l'anglais est ici une

"matière secondaire" qui ne prend que deux heures par semaine, ce qui renvoie à la question de la motivation pour se lancer dans un apprentissage.

TABLEAU 14: T.E.D.E., validité prédictive et matières d'enseignement

matières gradient de transfert	électricité			anglais
	théorie	pratique	oral	
PAT	0,47 S.	0,75 T.S.	0 38	0,11

La liaison avec l'oral d'électricité a une valeur intéressante, mais non significative. Il est vrai que l'on sait depuis longtemps que la fidélité des interrogations orales laisse souvent à désirer. On voit ici que, pour la même matière, l'interrogation écrite franchit la barrière de la significativité statistique.

Epreuve	Effectif	Corrélation
Test de vocabulaire	15	0,28
Test de vocabulaire	26	0,50**
Test de syntaxe	15	0,49*
Bilan français	322	0,28**
Bilan mathématiques	322	0,37***
<i>Xydias</i> (mathématiques)	31	0,42**
<i>D 48</i>	22	0,58***
Logique propositionnelle	22	0,40*
<i>Test Mécanique</i>	198	0,75***
Test spatial	24	0,70***
Test spatial	88	0,58***
<i>G.M.T.</i> (gestion du temps)	32	0,22
<i>T.G.R.A.</i> (attention)	103	0,45***
<i>N.E.L.C.-I.E.</i> (internalité)	31	0,27

* pour p=0,10 limite ; ** pour p=0,05 significatif ; *** pour p=0,01 très significatif

Tableau 16: validités concourantes.

Enfin, la liaison PAT-épreuve pratique est excellente. Il est vrai que l'évaluation de l'épreuve pratique présente les meilleures conditions à la fois de fiabilité de la correction (le système fonctionne ou ne fonctionne pas), et d'investissement dans une tâche facilement assimilée à la pratique du métier appris.

De nouvelles études ont été menées avec la version 2000 du T.E.D.E. Côté sensibilité l'échelle des P.A. standards concrètement observés s'étend de 67 à 127 points, soit une étendue de quatre écarts types ou bien, d'un point de vue ordinal, de 60 positions. Le T.E.D.E.2000 montre une très forte cohérence interne, l'alpha de Cronbach prenant la valeur 0,92 ; d'autre part, chacun des items participe de façon élevée à la construction du score global. La fidélité a été évaluée par la méthode des items pairs et des items impairs ; la corrélation corrigée entre les deux sous tests ainsi obtenus s'élève à 0,94. Un retest à une semaine d'intervalle pour un groupe de 15 sujets donne également une corrélation de 0,94.

En termes de validité concourante, des corrélations ont pu être calculées en fonction de différentes applications du test sur le terrain (voir tableau n° 16). On observe des liens positifs, le plus souvent significatifs, avec les épreuves renvoyant aux aptitudes verbale, numérique, spatiale, temporelle, mécanique.

Le tableau 17 rappelle les validités prédictives obtenues avec le T.E.D.E. dans ses différentes formes pour publics tout venant ainsi qu'avec les apprentis. La validité prédictive des tests classiques se situe entre 0,25 et 0,45 (d'après Lévy-Leboyer, 1987). Comparativement, celle du T.E.D.E. oscille entre 0,40 et 0,83 avec une validité moyenne d'environ 0,52 et donc supérieure au maximum de celle des tests d'aptitudes.

La dynamisation de la passation renforcerait donc la validité prédictive de manière consistante. Il resterait à le vérifier de manière spécifique pour le T.E.D.E. en comparant des passations sans apprentissage et avec apprentissage, ce qui est pratiquement impossible dans des conditions de terrain.

La version 2000 du T.E.D.E. a permis de mener une analyse différentielle détaillée de l'échantillon composé de 589 sujets adultes tout venant ayant passé le test à l'occasion d'un recrutement, d'un bilan de compétences, d'une entrée dans une formation, de l'élaboration d'un plan de formation...

Contexte	Public	Effectif	Niveau	Critère	Corrélation	Signif.
formation	A.I.S.A. ¹	18	V	E.F.S. ²	0,40	*
hôpital	aide-soignants	20	VI	Q.C.M. ³	0,44	*
formation	T.E.I.I. ⁴	26	IV	E.F.S.	0,40	**
entreprise	opératrices sur tours semi- automatiques	9	VI ; V	évaluation encadrement	0,67	**
entreprise	opérateurs extrusion et flexographie	52	VI ; V	évaluation encadrement	0,46	***
formation	opérateurs monteurs de capteurs	20	V	note de technologie	0,73	***
alphabétisation	femmes étrangères	14	VI	déchiffrage	0,83	***
entrée en apprentissage	apprentis débutants	244	VI ; V	notes sur 2 ans	0,51	***

TABLEAU N°17 : validité prédictive du T.E.D.E.]

¹ A.I.S.A. : agent d'installation de systèmes automatisés

² E.F.S. : examen de fin de stage

³ Q.C.M. : questionnaire à choix multiple

⁴ T.E.I.I. : technicien en électronique et informatique industrielle

On remarque une sur-représentation des hommes (377 vs 192) qui par ailleurs obtiennent des résultats supérieurs à ceux des femmes (102 points vs 94 ; $z=7,11^{***}$). On expliquerait de manière classique cet écart par la dimension spatiale contenue dans le test ; il faut toutefois noter que les meilleures performances obtenues par les femmes atteignent le niveau le plus élevé, ce qui ne les distingue pas par rapport aux meilleures de celles des hommes. L'écart de performance lié au sexe, dans ces conditions, serait à rapporter à une source d'écart culturelle plutôt qu'à une source d'écart constitutionnelle.

Le lien entre potentiel d'apprentissage et âge se concrétise dans une corrélation négative ($R = - 0,33^{**}$) : le potentiel aurait donc tendance à diminuer avec l'âge, ou bien il serait sensible à un effet de cohorte.

Les P.A. moyens varient dans le même sens que les niveaux de qualification, avec une liaison élevée ($\eta = 0,53$), ce qui valide l'utilisation du test en termes de prévision de réussite dans ces différents niveaux (voir tableau n°18).

niveaux de qualification					
PA	II	III	IV	V	VI
n	21	43	120	302	103
m	112,79	110,93	105,60	100,60	84,52
σ	6,76	7,86	13,84	14,09	9,43

TABLEAU N°18 : potentiels d'apprentissage et niveau de qualification

Différents travaux de validation en cours ont montré les qualités métrologiques du T.E.D.E.-A. La fidélité des scores aux tests se montre excellente (0,90). La corrélation entre les deux tests est de l'ordre de 0,80 ; l'effet de reclassement lié au temps est donc relatif et les écarts méritent bien une analyse attentive.

Une étude menée auprès d'un groupe d'une trentaine d'élèves de C.I.P.P.A. donne des corrélations positives et significatives entre le T.E.D.E.-A et la performance obtenue au test Mécanique (0,52) ainsi qu'avec le niveau scolaire (0,45) d'une part et le niveau d'insertion probable apprécié par l'équipe éducative (0,37) d'autre part. Ces deux dernières liaisons augmentent pour les élèves dont le niveau au premier test est le plus faible (respectivement 0,52 et 0,75).

Une recherche menée en pays flamand auprès d'un groupe de 28 adolescents de l'enseignement spécial a montré l'existence de corrélations positives entre le T.E.D.E.-A et les P.M. de Raven (0,56**) et entre le T.E.D.E.-A et le T.G.R.A. qui évalue les capacités d'attention (0,57**). D'autre part, le T.E.D.E.-A se montre un meilleur prédicteur du niveau des apprentissages pré-professionnels réalisés sur l'année scolaire que les P.M. ou que les capacités attentionnelles (0,50** vs 0,40* vs 0,34).

Les études de validité sont délicates à conclure en ce sens que pour un même prédicteur, elles donnent des résultats hétérogènes. Par exemple, pour le D48, les corrélations test-critère "note d'atelier" varient de -0,13 à 0,43 (Pichot, 1965). Les résultats obtenus jusqu'à maintenant pour le T.E.D.E. sont encourageants dans la mesure où certaines liaisons atteignent des valeurs supérieures à celles obtenues avec des épreuves statiques.

Cet ensemble d'études va dans le bon sens : les indices techniques recueillis indiquent que les qualités métrologiques du T.E.D.E. sont satisfaisantes. La fidélité de la mesure semble assurée ; l'épreuve, saturée en facteur G, corrèle de solide manière avec des épreuves classiques éprouvées. D'autre part, l'avantage du T.E.D.E. en matière de validité prédictive commence à se dessiner.

Bien entendu, la fiabilité de la mesure joue comme un effet de levier qui démultiplie l'intérêt des observations qualitatives recueillies sur comment le sujet résout, comment le sujet apprend, comment le sujet transfère, comment le sujet se perçoit dans son rapport à l'apprentissage... La démarche évaluative dynamique réconcilie par complémentarité les aspects quantitatifs indispensables à l'élaboration pronostique et les aspects qualitatifs, objectifs et subjectifs, nécessaires à la compréhension du sujet.

Enfin, le travail mené avec les apprentis dans le cadre de la préparation du T.E.D.E.6 a permis de mieux connaître les relations entre les indices des différents profils et d'en tirer une typologie. L'analyse porte sur les indices des profils d'apprentissage et de transfert établis auprès de 313 apprentis : réussite spontanée (RS), erreur bien corrigée (EBC), erreur illusoire (EI), interférence (INT), erreur mal corrigée (EMC), erreur non traitée (ENT) et omission (OM) pour le profil d'apprentissage ; réussite réitérée (RR), gain (G), perte (P), échec réitéré (ER) et omission réitérée (OMR).

Un échelonnement multidimensionnel donne un indice de stress trop élevé (0,28) pour une dimension et avec deux dimensions, on parvient à un ajustement acceptable (0,10). La figure 13 positionne les variables sur ces deux dimensions sans

que pour autant une structure manifeste ne s'impose. Dans ce cas, une analyse en grappes peut aider à l'interprétation (Tournois & Dickes, 1993). Tant par la méthode des distances moyennes entre classes que par la méthode des distances moyennes intra-classes, on obtient le même dendrogramme obtenu à partir de la même matrice. A partir ce dendrogramme on peut regrouper les variables en sept ensembles. L'ensemble RS/RR rassemble les indices de réussite à l'apprentissage et au test : réussite spontanée et réussite réitérée. L'ensemble OM/OMR rapproche les indices relatifs aux omissions de l'apprentissage et aux omissions réitérées au test. L'ensemble EI/EBC se rapporte au traitement positif de l'erreur en cours d'apprentissage : erreurs bien corrigées et erreurs illusives. Juste au-dessus sur la carte, l'ensemble EMC/INT concerne les indices de traitement négatif de l'erreur : erreurs mal corrigées et interférences. Les indices du groupe ER/ENT évoquent le manque de contrôle : erreurs réitérées ou erreurs non traitées. Enfin, les variables P, indice de perte, et G, indice de gain font cavaliers seuls.

Dans cette représentation spatiale des variables des profils d'apprentissage et de transfert est-il possible de faire ressortir une dimension latente de potentiel d'apprentissage ? L'examen visuel des corrélations entre ces variables et les scores de potentiel et de niveaux scolaires à l'entrée ou à l'issue de la première année justifie cette question dans la mesure où des corrélations significatives et parfois élevées s'observent (tableau 19).

En suivant la procédure préconisée par Tournois & Dickes (1993), trois équations de régression linéaire multiple avec les corrélations -scores des profils d'apprentissage et de transfert/indices d'apprentissage- prises successivement comme variable dépendante et les coordonnées sur les axes prises comme variables indépendantes ont été calculées. Les coefficients R multiple expriment une qualité de relation satisfaisante entre les indices d'apprentissage et les coordonnées des axes pour le potentiel d'apprentissage (77% de variance) et les résultats à l'issue de la première année (85% de variance), cette qualité devenant limite pour le niveau à l'entrée.

Les coefficients de régression des équations relatives à ces deux indices ont été convertis en cosinus¹, cosinus eux-mêmes convertis en angles (tableau 35).

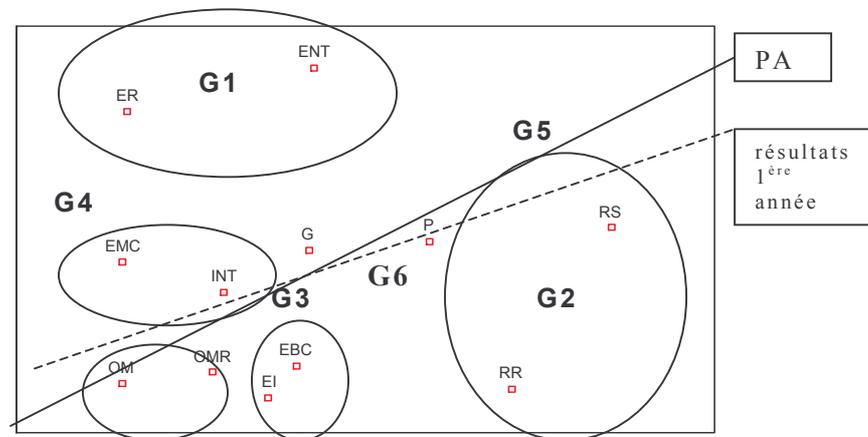


Figure 13. Echantillon des apprentis, seconde cohorte, espace bidimensionnel des indices d'apprentissage et de transfert. Les régions issues de l'analyse hiérarchique sont indiquées ainsi que les dimensions latentes interprétatives. Les groupes de sujets (G1-impulsivité ; G2-efficacité ; G3-réflexivité ; G4-déficit ; G5-rapidité ; G6-limitation) sont positionnés de manière analogique.

A partir de ces angles, on a pu tracer les deux axes sur l'espace bi-dimensionnel. Ces deux axes partagent de la même façon les points de l'espace et on peut considérer qu'ils matérialisent deux dimensions latentes très proches. «Et pour point X_i de l'EMD, on peut obtenir sa projection perpendiculairement sur la droite de régression. Elle fournit une valeur d'échelle estimée par la régression comme étant la

¹ $c_r = b_r / (b_1^2 + b_2^2)^{1/2}$ où c est le cosinus et b les coefficients de régression.

meilleure possible (Tournois & Dickes, 1993). On peut conclure que la prise en compte de toutes les relations entre toutes les variables des profils d'apprentissage et de transfert s'organisent dans un espace bi-dimensionnel dans la direction d'une dimension latente de potentiel d'apprentissage s'exprimant à la fois dans le résultat du T.E.D.E et dans les acquis scolaires de la première année.

variable profil	d1	d2	COR var/PA	COR var/ niveau entrées	COR var/ résultats 1ère année
Réussite spontanée	1,13	0,21	0,54**	0,16**	0,25**
Erreur bien corrigée	0,00	-0,34	0,04	-0,04	0,03
Erreur illusoire	-0,10	-0,47	0,01	-0,05	-0,04
Interférence	-0,26	-0,05	-0,16**	0,01	-0,15**
Erreur mal corrigée	-0,62	0,07	-0,27**	-0,07	-0,13*
Erreur non traitée	0,06	0,84	-0,21**	-0,04	-0,08
Omission	-0,62	-0,41	-0,28**	-0,10	-0,13*
Réussite réitérée	0,77	-0,43	0,86**	0,31**	0,37**
Gain	0,05	0,12	0,43**	0,11*	0,07
Perte	0,48	0,15	-0,11*	-0,17**	0,00
Echec réitéré	-0,60	0,67	-0,74**	-0,18**	-0,29**
Omission réitérée	-0,30	-0,36	-0,22**	-0,15**	-0,16**
	R multiple		0,88	0,70	0,92
	R ²		77%	49%	85%
	F		15,28	4,38	26,11
	p		0,0013**	0,047*	0,0002**
	angle 1		29°99	21°36	22°08

Tableau 19. Coordonnées des indicateurs de profil sur les axes de l'espace bidimensionnel. Corrélations (rhô de Spearman) entre ces indicateurs et les indices d'apprentissage ($p < 0,05$ * ; $p < 0,01$ **). Régressions multiples et direction (angles) d'une variable latente de potentiel d'apprentissage.

Une seconde analyse hiérarchique a été menée sur les observations en demandant un regroupement des sujets en six

classes². Afin de dresser le profil de chacun de ces groupes, les moyennes sur les indices d'apprentissage et de transfert ont été calculées pour chaque groupe, ainsi que les moyennes du potentiel d'apprentissage, du bilan pédagogique d'entrée et des résultats de la première année.

Une série d'ANOVA donne les différences entre indices très significatives dans chacun des groupes. Le tableau 20 présente l'ensemble de l'information.

Le comportement de l'apprenti du groupe 1 revient à travailler plutôt rapidement (peu d'omissions), de manière spontanée sans assurer un contrôle de sa réponse (erreurs non traitées et échecs réitérés en nombre). De cette manière plutôt impulsive, il assure environ une réponse sur deux au niveau de l'apprentissage alors qu'il transfère moins bien que la moyenne. Son PA n'atteint pas la moyenne. Ses connaissances scolaires à l'entrée au centre et ses connaissances acquises après une année le situent autour de la moyenne, en troisième rang.

Le groupe 2 est celui de l'excellence avec un PA moyen se situant à deux écarts types. Pratiquement tous les items sont traités, les réussites spontanées et réitérées atteignent des maxima pour cette cohorte et les erreurs font l'objet d'un bon contrôle. Sur le plan des connaissances scolaires à l'entrée et acquises en cours d'année, les sujets de ce groupe obtiennent les meilleurs résultats.

Le groupe 3 obtient un PA en-dessous de la moyenne. Plutôt lents (omissions et omissions réitérées), hyper-réflexifs et hyper-contrôlés, les apprentis de ce groupe réussissent rarement les exercices du premier coup et ils les refont (erreurs bien corrigées, erreurs illusoires); capables de reproduire *grosso modo* leurs réussites de l'apprentissage au test, ils peuvent également en perdre. En-dessous de la moyenne en

² La question du nombre de groupes de sujets a été tranchée pour satisfaire à deux contraintes : les groupes se distribuent sur un continuum de P.A. ; les groupes restent interprétables au regard des différents indicateurs. Un essai en 5 groupes ne satisfait pas la première contrainte et un essai en 7 groupes ne satisfait pas la seconde.

termes de connaissances scolaires à l'entrée, ils vont légèrement progresser en cours d'année.

Groupe	1	2	3	4	5	6	Ensemble
Effectif	62	43	52	51	41	36	313
Réussite spontanée	4,13	9,26	1,09	<u>0,39</u>	6,48	6,95	4,34
Erreur bien corrigée	0,71	0,64	1,72	0,37	<u>0,16</u>	1,38	0,83
Erreur illusoire	0,43	0,53	2,48	0,4	<u>0,14</u>	0,72	0,81
Interférence	0,16	<u>0,06</u>	0,53	0,96	0,23	0,15	0,37
Erreur mal corrigée	0,59	<u>0,26</u>	1,62	4	0,43	0,64	1,34
Erreur non traitée	5,40	<u>1,11</u>	1,47	4,16	4,48	1,62	3,20
Omission	0,56	0,13	3,14	1,61	<u>0,09</u>	0,64	1,11
Réussite réitérée	2,15	8,60	2,57	<u>0,33</u>	4,25	3,38	3,31
Gain	0,81	0,96	0,98	0,93	1,82	<u>0,56</u>	1
Perte	3,12	1,85	2,74	<u>0,82</u>	2,52	5,62	2,67
Echec réitéré	5,88	<u>0,55</u>	4,53	9,70	3,41	1,92	4,69
Omission réitérée	0,04	0,04	0,95	0,21	<u>0</u>	0,51	0,29
PA	94,46	130,52	96,38	<u>85,88</u>	111,33	99,7	101,69
Bilan pédagogique d'entrée	-0,12	0,62	-0,12	-0,05	0,46	<u>-0,27</u>	0,06
Résultats de la première année	0,12	0,60	-0,04	<u>-0,52</u>	0,36	<u>-0,14</u>	0,05

Tableau 20. Profils de réponses (en gras, la moyenne la plus élevée ; soulignée, la moyenne la plus basse). Les indicateurs d'apprentissages scolaires sont exprimés en notes centrées réduites.

Le groupe 4, à caractère déficitaire, rassemble les apprentis passablement désorientés par le micro-monde du T.E.D.E. : ils vont plutôt lentement (omissions), produisent peu de réussites spontanées, ne corrigent pas ou corrigent mal les erreurs et reproduisent leurs échecs au test. Dans ces conditions, leur PA se retrouve plutôt bas, à environ moins un écart type. Leur niveau de connaissances scolaires est plutôt faible à l'entrée, et l'écart va se creuser par rapport aux autres en cours d'année de manière très prononcée.

Le groupe 5 obtient un PA au-dessus de la moyenne. Les apprentis de ce groupe travaillent vite (pratiquement pas d'omissions) et ils transfèrent assez bien pouvant même gagner quelques points, mais ils ne se montrent pas aussi efficaces que ceux du groupe 2. Les niveaux de réussite académique sont au-dessus de la moyenne, en second rang.

Le groupe 6 présente un PA moyen. Bien que capables de produire des réussites spontanées en cours d'apprentissage, les apprentis de ce groupe éprouvent de grandes difficultés à transférer leurs réussites à l'apprentissage aux items du test, ce que laisse supposer la fréquence élevée des pertes. Leur apprentissage manifeste une forte tendance à rester limité au contexte de sa réalisation, tendance qui explique probablement leur faible niveau de connaissances à l'entrée et en cours d'année, leur progression étant liée en partie à la dérive pédagogique du groupe 4.

Au-delà du fait que cette typologie en six groupes positionnés sur l'espace bi-dimensionnel puisse participer à une interprétation plus fine des résultats et des pronostics, on observe que la valeur du PA semble essentiellement liée à une variable latente d'efficacité à la source de la production de réussites spontanées au cours de l'apprentissage et de réussites réitérées au cours du test. Toutefois, on ne connaît pas la contribution de chacune de ces deux conduites cognitives à la construction de la performance. La prochaine section tente de répondre à cette question.

La question relative à la part de la réussite spontanée et à la part de la réussite réitérée dans la construction de la performance d'apprentissage est une question centrale dans la

définition de la spécificité de l'évaluation dynamique par rapport à l'évaluation statique. En effet, si la réussite spontanée à l'apprentissage explique également la performance au test, on pourrait assimiler évaluation statique et évaluation dynamique et référer la performance à un classique facteur général de l'intelligence. Par contre, si la réussite réitérée explique la performance au test plutôt que la réussite spontanée, alors la spécificité du test d'apprentissage apparaît clairement : l'évaluation dynamique stimule la capacité de réitérer en situation d'autonomie ce qui a été appris avec de l'aide. Qu'en est-il exactement à partir des données recueillies auprès de 286 apprentis. On a calculé les liaisons entre réussite réitérée, réussite spontanée et performances d'apprentissage. On observe que la réussite réitérée explique mieux le PA que la réussite spontanée (74% de variance expliquée vs 29%) ; les corrélations partielles confirment également l'avantage à la réussite réitérée sur la réussite spontanée dans la construction de la performance au test (0,81** à réussite spontanée égale et 0,01 à réussite réitérée égale). Avec des corrélations moins élevées, les différences vont dans le même sens par rapport aux résultats du bilan pédagogique d'entrée (épreuves de français et de mathématiques données à l'entrée au Centre de formation, dans la même période que le T.E.D.E.) : 10% de variance expliquée contre 3% ; 0,26** vs -0,02 pour les régressions partielles. Enfin, pour ce qui est des résultats obtenus au Centre à l'issue de la première année, les résultats vont dans le même sens, à l'avantage de la réussite réitérée : 14% de variance expliquée contre 7% ; 0,26** vs -0,04 pour les régressions partielles .

La figure 14 reprend l'ensemble des corrélations entre les variables. On observe finalement que dans tous les cas la valeur explicative des performances d'apprentissage de la réussite réitérée est supérieure à celle de la réussite spontanée. Qui plus est, les valeurs des corrélations entre réussite réitérée et critères de performances pédagogiques sont pratiquement égales à celles du PA. En d'autres termes, ce qui fonde la validité prédictive du T.E.D.E. serait bien la compétence à refaire par soi-même ce qu'on a appris à faire avec des aides didactiques.

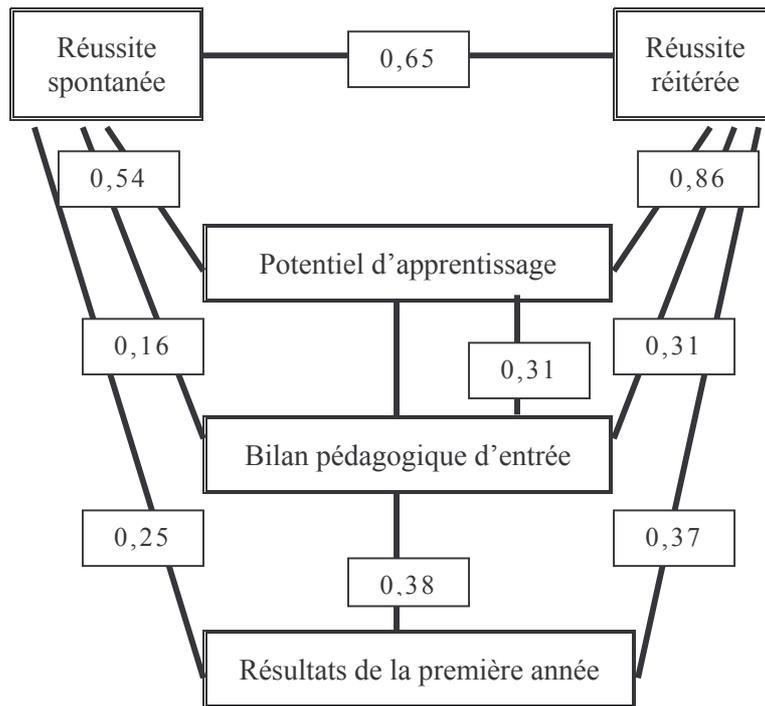


Figure 14. Liaisons réussite réitérée vs réussite spontanée et performances d'apprentissage

Le "micro-monde" du T.E.D.E. offre donc bien une mise en situation homomorphe à la situation de formation : on apprend avec l'aide du moniteur du centre de formation ou du tuteur de l'entreprise ou du maître d'apprentissage, puis on est évalué sur sa capacité à refaire seul les exercices de la progression par une procédure de contrôle ou d'examen. **Plus que la capacité de savoir faire immédiate, c'est la capacité de savoir faire différée, médiatisée par le transfert des compétences acquises qui présente une valeur explicative -et prédictive- du niveau d'apprentissage atteint ou virtuel.**

Voilà qui justifie pleinement le paradigme apprentissage-test et en cela le T.E.D.E. se démarque bien des tests classiques d'une part, et des scores différentiels du paradigme test-entraînement-retest d'autre part.

Conclusion

L'évaluation dynamique du potentiel d'apprentissage de l'adulte est donc une idée ancienne, qui s'est développée dans le temps et dans l'espace de manière sporadique. Cette idée se traduit essentiellement dans l'adaptation ou la création d'outils et de procédures d'évaluation en référence plus ou moins étroite et explicite avec les grandes théories de l'intelligence. Ce référentiel théorique hétéroclite mériterait un travail d'approfondissement, de réflexion et d'analyse, de recherche de cohérence, de modélisation dans la mesure où "Mieux fonder théoriquement l'évaluation n'implique pas seulement que chaque technique ait une validité théorique satisfaisante, mais implique aussi que ces techniques soit situées dans un cadre théorique large fournissant une vue d'ensemble des mécanismes du développement des individus et de leur adaptation aux tâches qu'ils choisissent ou qui leur sont prescrites." (Huteau, 1994)

Il est vrai que cette théorie générale du développement et de l'adaptation fait défaut depuis que les grandes théories de Freud, Piaget, Wallon ont été plus ou moins démontées sans pour autant être remplacées : les psychologues théoriciens sont toujours dans l'incapacité majeure d'expliquer comment un sujet enfant ou adulte apprend. Il n'y a pas de théorie de l'évaluation dynamique parce qu'il n'y a plus de théorie de l'apprentissage universellement reconnue.. En conséquence, le point de vue pragmatique, qui reconnaît la valeur de l'outil à l'aune de la qualité des réponses concrètes qu'il apporte (valeur d'usage), l'emporte sur le point de vue théorique d'où découlerait une application pratique. Les chercheurs universitaires ont donc là, à portée de main, un champ d'investigation de première importance qui leur offre la possibilité de faire le lien entre des pratiques de terrain et des formulations conceptuelles de connaissances plus générales.

Un deuxième plan de recherche concerne la méthodologie de l'évaluation dynamique. Le chemin parcouru

dans ce paysage a montré une diversité de conceptions et d'objectifs qu'il serait bon de reprendre et d'affiner. Ces conceptions se distribuent sur un *continuum* dont la première borne se préoccupe de l'épuration du résultat. En éliminant le maximum de variables parasites liées à l'appartenance socioculturelle, au degré de familiarité avec la tâche, à l'orientation cognitive du sujet, aux aspects motivationnels et émotionnels, le psychologue ou le formateur peuvent ainsi se rapprocher de la valeur vraie des capacités intellectuelles du sujet. La seconde borne postule que par la médiation sociale, l'évaluateur-expert va transmettre des outils psychologiques au sujet-novice, outils qui permettront à ce novice, une fois qu'il les aura intériorisés, de devenir performant dans la tâche. Les bornes intermédiaires traduisent des conceptions de rattrapage ou de déblocage, l'aide apportée au sujet ayant pour but de faciliter l'expression de toutes ses capacités latentes par le moyen d'une modification de l'image de soi.

Bien évidemment, ces objectifs renvoient à des processus cognitifs différents : dépassement des variables parasites, levée des freins à l'entrée dans la tâche et à l'investissement vers une performance, prise de conscience de ses capacités latentes et apprentissage du meilleur usage des outils cognitifs à disposition, apprentissage de nouvelles capacités, restructuration d'ensemble de ses modes de fonctionnement cognitif et conatif.

Souci métrologique de la validité de la mesure, objectif d'actualisation et de concrétisation des capacités virtuelles, ambition d'accélérer un développement, de modifier la structure des capacités en apprenant des capacités nouvelles et en modifiant la perception de soi, autant de conceptions différentes dans un même champ. L'intérêt serait ici de mieux cerner l'articulation entre l'objectif, le processus et l'outil afin que le praticien soit à même de choisir l'outil le mieux adapté à son projet parmi l'ensemble des outils disponibles.

Un autre axe de recherche méthodologique concerne la question des contenus. Pour parler dans le langage de Vygotski, il faut se demander s'il existe une seule zone de proche développement, valable pour tous les contenus et toutes les

formes d'apprentissage ou bien s'il existe une zone de proche développement par activité ou par type d'activité. Des plans de recherche expérimentaux s'imposent pour explorer cet espace d'interrogation, ainsi que pour mieux différencier l'évaluation dynamique du potentiel de l'enfant de celui de l'adulte. On a vu que des techniques élaborées dans le domaine de la psychologie éducative de l'enfant avaient pu être importées dans le domaine de la psychopédagogie de l'adulte sans autre forme d'examen. C'est là un abus d'usage caractérisé qui ne peut conduire qu'à des conclusions saugrenues, voire néfastes pour le sujet. Par exemple, il n'est pas rare de voir qu'on donne à passer l'E.C.D.L. ou les T.O.F à des adultes, épreuves étalonnées au niveau du collège. Refuser l'accès à une formation qualifiante à une personne et justifier ce refus par une conclusion du type "Votre fonctionnement se situe au stade opératoire concret" relève de la plus rigoureuse et de la plus vigoureuse des incompétences. En effet, il n'existe pas de références étalonnées avec des adultes de ces épreuves. D'autre part, le concept de stade renvoie à celui de développement, propre à l'enfant et à l'adolescent.

L'évaluation dynamique de l'enfant se situe bien dans cette perspective du développement par la maturation biologique, par l'expérience avec le monde physique, par les médiations sociales et corporelles des affects, des valeurs, des outils langagiers et intellectuels... Pour l'adulte, dont la mise en place des structures biologiques et cognitives est achevée, la question de l'évaluation dynamique du potentiel d'apprentissage se posera dans la perspective du freinage des processus d'involution de ses capacités. Dans ce sens il est facile d'imaginer des études sur l'évolution du potentiel d'apprentissage par rapport à l'involution des capacités en s'inspirant des travaux de Pichot (1965) qui utilise conjointement un test sensible au vieillissement (le D48 en rapport avec le facteur fluide) et un test qui lui résiste (épreuve de vocabulaire en rapport avec le facteur cristallisé).

En troisième lieu, il ne faut pas négliger les aspects métrologiques, et principalement la question de la validité prédictive. En effet, un avantage majeur des tests d'évaluation

dynamique du potentiel d'apprentissage serait de mieux se comporter en tant que prédicteurs de la réussite des apprentissages. Encore faudrait-il en administrer la preuve de manière la plus large possible en termes de métiers, en termes de tranches d'âge, en termes de sexe, en termes de niveau scolaire...

Les recherches de validité amèneront également à travailler le choix des différents scores utilisables dans une épreuve dynamique, selon l'objet de la mesure. On peut prendre en compte des niveaux de performances : niveau de réussite spontanée, niveau de réussite après apprentissage ; ou bien l'effet du transfert de l'apprentissage sur la performance : gains absolu ou relatif, gradient de transfert ; ou bien des variables liées directement à l'apprentissage : nombre d'essais ou nombre d'aides, durée de l'apprentissage, niveau de l'aide... Ces différents types de scores, selon le modèle statistique impliqué dans leur élaboration (de la simple somme ou différence arithmétique aux modèles de réponse à l'item), n'ont pas forcément une valeur identique en matière de validité prédictive et seule une étude comparative empirique apportera réponse à cette question.

Au plan de l'approche clinique de l'individu, là encore de nombreuses pistes de recherches s'ouvrent aux praticiens et chercheurs de bonne volonté afin d'affiner l'approche des fonctionnements individuels dans des contextes définis. La constitution et l'étude de batteries d'épreuves en fonction des problématiques individuelles, -tests psychométriques et pédagogométriques, questionnaires, mises en situations réelles ou simulées-, incluant des procédures d'évaluation dynamique permettraient d'avancer d'une part dans la qualité de la réponse apportée au problème posé par chaque sujet, d'autre part de mieux connaître les articulations en termes de complémentarité des différents outils d'évaluation disponibles.

Toujours sur le plan de l'approche clinique, ce n'est pas parce que chaque cas est unique qu'il faille s'empêcher de procéder à des échanges entre praticiens, à des études de cas comparatives, dans le but, par méta-analyse des collections de cas, de dégager des lois générales relatives aux processus

fonctionnels, représentationnels, et sur les dynamiques socio-affectives du sujet adulte placé en situation d'apprentissage. Ces lois seraient à rapprocher de patrons de réponses aux épreuves, la seule présence de ces patrons ayant valeur d'indicateur, de signe, voire de symptôme de l'existence de tel ou tel processus, sans avoir besoin de re-procéder à l'analyse complète. Il est clair également que seule une entrée par l'approche clinique permettrait d'aborder la question de la pathologie du potentiel d'apprentissage, en explorant cas par cas les déterminants possibles des différents types de patrons de réponses aberrants (interférences, attitudes égocentrées, labilité de l'attention, hyper ou hypo contrôle, perception lacunaire...). Cette approche clinique se situe finalement au cœur du champ d'études et de recherches dans la mesure où ses constats pourront alimenter la formulation d'hypothèses suivies de leurs vérifications à l'aide de plans expérimentaux.

Ces mêmes constats pourraient également être réinvestis dans le champ de la psychopédagogie des adultes. Par exemple, on peut penser utiliser la mesure du potentiel d'apprentissage comme critère d'évaluation des effets d'une action de formation. Et puis cela renvoie à la question de la remédiation cognitive : comment actualiser, comment renforcer, comment élargir le potentiel d'apprentissage de l'adulte ? Comment lui apprendre à mieux utiliser les outils cognitifs dont il dispose, comment lui apprendre à transférer ses compétences acquises dans un domaine professionnel à un nouveau domaine de reconversion ou de loisir...

Il est vrai que beaucoup d'approches pédagogiques de la remédiation ressemblent à la médecine symptomatique qui s'obstine à confondre le symptôme et les sources de la maladie. On pourrait tout aussi bien prendre des exemples dans d'autres domaines, comme le traitement de la statistique du chômage plutôt que celui des racines du chômage. Ce n'est pas parce qu'on met le thermomètre dans la glace que la fièvre du patient va s'atténuer et sa maladie disparaître. Ce n'est pas parce qu'on isole des jeunes dans des stages qu'ils ont de *facto* un emploi...

L'intervention qui vise la stimulation des capacités d'apprentissage ne va pas se centrer sur l'entraînement à passer

le test, mais s'adresser aux sources déterminant l'engagement et l'efficacité dans de nouveaux apprentissages. Parmi ces sources, on peut rapidement citer les facteurs de stimulation-inhibition des fonctions intellectuelles, l'image de soi et de sa compétence dans un type de tâche, l'empreinte socio-affective des processus de socialisation actuels et passés, les cicatrices mentales des conséquences des erreurs, la valence conative du vécu des expériences d'apprentissage antérieures, le degré de maîtrise des fonctions exécutives périphériques au niveau de la perception et de l'effectuation... Dans la mesure où l'approche clinique du sujet adulte en situation d'apprentissage au cours de l'évaluation dynamique permettra de décoder et de comprendre le jeu de ces déterminants, cette approche deviendra une source d'inspiration pour les pédagogues.

Chercher en psychopédagogie ne signifie pas vouloir toujours inventer l'eau chaude ou le fil à couper le beurre ; en effet il convient de "...discerner ce qui représente de réelles nouveautés, éventuellement de réels progrès, et pas seulement des changements de vocabulaire ou des innovations dont l'originalité n'apparaît qu'à ceux qui ont choisi d'ignorer leurs prédécesseurs." (Reuchlin, 1994)

Chercher ne signifie pas succomber aux effets de modes aussi superficielles que passagères, non plus que d'adhérer à des entreprises militantes qui convoqueraient la science pour garantir des points de vue partisans : chercher consiste à partir des problématiques et des projets concrets portés et apportés par l'adulte qui s'adresse au psychologue ou au formateur afin d'améliorer l'efficacité et le confort de la prestation sociale publique, associative ou privée.

Il reste à conclure sans rien fermer des perspectives ouvertes par le renouveau d'intérêt qui se manifeste actuellement pour l'évaluation dynamique, et tout en essayant de ressaisir l'essentiel. Cet intérêt est justifié de plusieurs points de vue pour le praticien psychologue, formateur ou gestionnaire des ressources humaines.

Des études passées en revue on retirera l'idée que l'évaluation dynamique apporte une information quantitative fidèle et plus valide en termes de prédictibilité de la réussite des

apprentissages que l'évaluation statique. L'existence d'un effet de reclassement des performances suite à l'apprentissage montre bien que la performance intellectuelle est sensible à l'apprentissage par transmission sociale d'une méthodologie de résolution. L'évaluation dynamique permet donc, dans une certaine mesure, de diminuer les biais culturels qui parasitent les situations d'évaluation.

Un autre intérêt vient de la quantité et de la qualité de l'information recueillie par une dynamisation des épreuves statiques. L'évaluation dynamique commence là où l'évaluation statique s'arrête. C'est-à-dire qu'en plus du simple niveau de réussite spontanée à l'épreuve subie, le psychologue ou le formateur accède à des informations relatives au comportement, aux conduites, aux représentations et aux affects du sujet dans une véritable situation d'apprentissage. Seule l'évaluation dynamique permet une centration sur les processus fonctionnels mis en oeuvre par le sujet apprenant. C'est un avantage majeur en ce sens que la connaissance des caractéristiques du sujet seront plus finement appréhendées et en ce sens que ces informations nuancées pourront être exploitées dans le champ psychopédagogique.

L'évaluation dynamique donne la chance d'apporter une information compréhensible par le psychologue, par le formateur et par le sujet : c'est la condition qui ouvre la voie à une psychopédagogie partenariale, l'expertise devenant accessible à tous, dans la mesure où chacun en porte et apporte sa part. Toute tendance à l'enthousiasme devant être contrebalancée par l'examen des limites et réserves d'usage, il faut savoir raison garder pour ne point basculer dans le mythe de l'éducabilité et de l'illusion de la toute puissance pédagogique : le médiateur propose, l'apprenant dispose. L'évaluation dynamique soulève autant de questions qu'elle apporte de réponses, ce qui renforce, à l'évidence, son intérêt. Questionnement théorique et, pour le praticien, questionnement méthodologique où pratiquement beaucoup reste à faire en ce qui concerne l'évaluation des adultes.

Alors continuons à faire, pour mieux dire ensuite.

Bibliographie

- Aeblin H.(1951), *Didactique Psychologique*, Neuchâtel : Delachaux et Niestlé
- Bateson G. (1984), *La nature et la pensée*, trad. française, Paris : Seuil
- Binet A. (1911), *Les idées modernes sur les enfants*, Paris, Flammarion
- Binet A., Simon T. (1983), *La mesure du développement de l'intelligence chez les jeunes enfants*, Issy-les-Moulineaux : Éditions scientifiques et psychologiques
- Bonnardel R. (1983), *Manuel des tests d'intelligence concrète*, Issy-les-Moulineaux : Éditions scientifiques et psychologiques
- Boss, Cardinet, Maire, Muller, *Manuel de la Batterie générale d'aptitudes*, Neuchâtel, Delachaux et Niestlé, 1960
- Briet S. (1991), "Remue-méninges à l'usine", *Libération*, 30-04-91
- Brown A.L., Ferrara R.A. (1985), "Diagnosing zones of proximal developpement", in Wertsch J.V. (ed), *Culture, communication and cognition, Vygotskian perspectives*, New-York : Cambridge University Press
- Büchel F.P., Paour J.L., "Introduction. Contributions à l'étude des potentiels d'apprentissage et de développement", in *European Journal of Psychology of Education*, 1990, Vol. V, n°2
- Canguilhem C. (1979), *Le normal et le pathologique*, Paris : PUF
- Cardinet J. (1986), *Évaluation scolaire et mesure*, Bruxelles : De Boeck-Wesmael
- C.E.R.I.F., O.N.I.S.E.P. Centre (1994), *L'apprentissage en région Centre*, Orléans
- Chartier D., Loarer E. (1994), "Évaluation dynamique de l'intelligence 'non-verbale par la procédure aide au cours du test : application à une population non-francophone et à des adultes de bas niveau de qualification", in Huteau M. (ed), *Les techniques psychologiques d'évaluation des personnes*, Issy-les-Moulineaux : EAP

- Cibois P. (1983), *L'analyse factorielle*, Paris : P.U.F.
- Debray R., Dufay C. (1994), "Une expérience d'évaluation dynamique du potentiel d'apprentissage en milieu industriel", in Huteau M. (ed), *Les techniques psychologiques d'évaluation des personnes*, Issy-les-Moulineaux : EAP
- De Landsheere G. (1979), *Dictionnaire de l'évaluation et de la recherche en éducation*, Paris : P.U.F.
- Dias B. (1994), "De l'approche psychométrique à l'évaluation dynamique du potentiel d'apprentissage", in Garbo R., Lebeer J. (eds), *A la recherche du potentiel d'apprentissage*, Anvers : AEMAM
- Embretson S.E. (1987), "Toward development of a psychometric approach" in Lidz C.S. (ed), *Dynamic assessment. An interactional approach to evaluating learning potential*, New-York : The Guilford Press
- Fattu N.A., Mech E., Kapos E. (1954), "Some statistical relationships between selected response dimensions and problem-solving proficiency", *Psychological Monographs : General and Applied*, , 68, 6, Whole N°. 377
- Faverge J.M. (1955), *Calcul des longueurs, Test*, Braine-le-Château : Application des techniques modernes,
- Faverge J.M. (1977), *Calcul des longueurs, Les tests de J.M. Faverge, Manuel*, Issy-les-Moulineaux : Editions Scientifiques et Psychologiques,
- Feuerstein R., Rand Y., Hoffman M.B., Miller R., (1980) *Instrumental Enrichment An Intervention Program for Cognitive Modifiability*, Baltimore : University Park Press
- Feuerstein R., Rand Y., Jensen M.R., Kaniel S., Tzurriel D. (1987), "Prerequisites for assessment of learning potentiel : the LPAD model", in Schneider-Lidz C. (ed), *Dynamic Assessment*, New-York : Guilford Press
- Feuerstein R., Rand Y., Haywood H.C., Hoffman M.B., Jensen M.R. (non-daté), *L.P.A.D. Learning Potential Assessment Device, Manual experimental version*, Jerusalem : Hadassah-Wizo-Canada-Research Institute,
- Fischer R.A.(1947), *Les Méthodes statistiques adaptées à la recherche scientifique*, trad. française, Paris : P.U.F.

- Guthke, J. (1990). Les tests d'apprentissage en tant qu'alternative ou complément des tests d'intelligence: Un bilan de leur évolution. *European Journal of Psychology of Education*, V, 2 (Special Issue), 117-133.
- Higelé P., Martin B. (1979), "Une expérience d'apprentissage d'opérations intellectuelles", *Revue Française de Pédagogie*, n°46
- Hurtig M. (1954), "Recherche sur la perfectibilité. Les effets de l'explication "Progressive Matrices 47" de Raven chez l'enfant normal et l'enfant débile", *Enfance*
- Hurtig M. (1960), "Étude expérimentale des possibilités d'apprentissage intellectuel d'enfants débiles et d'enfants normaux", *Enfance*, 4-5
- Hurtig M. (1966), *Les effets de l'explication sur les performances intellectuelles des enfants d'âge scolaire*, Doctorat 3è cycle
- Hurtig M. (1967), "Constat d'acquisitions ou pronostic d'apprentissage. Peut-on dynamiser la psychométrie ?" *Revue suisse de psychologie*, 26
- Huteau M., Lautrey J. (1978), "L'utilisation des tests d'intelligence et de la psychologie cognitive dans l'éducation et l'orientation", *L'orientation scolaire et professionnelle*, n°2
- Huteau M. (1994), "L'évaluation psychologique des personnes : problèmes et enjeux actuels", in Huteau M. (ed), *Les techniques psychologiques d'évaluation des personnes*, Issy-les-Moulineaux : EAP
- INETOP (1990), *Manuel de l'échelle collective de développement logique*, Issy-les-Moulineaux : E.A.P.
- Ionescu S., Jourdan-Ionescu. C. (1985), "L'évaluation du potentiel d'apprentissage, 1. Utilisation du test des cubes", *Bulletin de Psychologie*, tome- XXXVIII, n° 372
- Ionescu S., Jourdan-Ionescu C., Alain M. (1986), "L'évaluation du potentiel d'apprentissage, II Une nouvelle méthode de quantification ", *Bulletin de Psychologie*, , tome XL, n°380
- Jacquin S. (1991), "Quand le détonateur s'appelle LIGE", *Supplément à Partenaires*, n.°14
- Langouet G., Porlier J.C. (1981), *Mesure et statistique en milieu éducatif*, Paris : E.S.F.

- Lautrey J. (1980), "La variabilité intra-individuelle du niveau de développement opératoire et ses implications théoriques", *Bulletin de Psychologie*, 33
- Lautrey J., De Ribaupierre A., Rieben L. (1990), "L'intégration des aspects génétiques et différentiels du développement cognitif", in Reuchlin M., Longeot F., Marendaz C., Ohlmann T. (eds), *Connaître Différemment*, Nancy : Presses Universitaires de Nancy
- Lautrey J. (1994), "L'évaluation du potentiel d'apprentissage, état de la question", in Huteau M. (ed), *Les techniques psychologiques d'évaluation des personnes*, Issy-les-Moulineaux : E.A.P.
- Leplat J. (1985), "Les représentations fonctionnelles dans le travail", Ehrlich S. (ed), *Les Représentations, Psychologie Française*, 30, ¾
- Lévy-Leboyer C., Spérandio J.C. (1987), *Traité de psychologie du travail*, Paris : PUF
- Lévy-Leboyer C. (1994), "La sélection du personnel en Europe", in Huteau M. (ed), *Les techniques psychologiques d'évaluation des personnes*, Issy-les-Moulineaux : EAP
- Langouet G., Porlier J.C. (1989), *Pratiques statistiques en sciences humaines et sociales*, ESF : Paris
- Longeot F. (1979), *Manuel de l'échelle de développement de la pensée logique*, Issy-les-Moulineaux : E.A.P.
- Longeot F., *Test des opérations formelles*, Paris, CNAM-INOP
- Longeot F., "L'évolution de la notion de stade opératoire" (1990), in Reuchlin M., Longeot F., Marendaz C., Ohlmann T. (eds), *Connaître Différemment*, Nancy : Presses Universitaires de Nancy
- Mary C., Mariel R. (1982), "Le développement intellectuel chez les enfants de 4 à 6 ans, le test des carrés", *Psychologie Scolaire*, N°40
- Mary C., Mariel R. (1978), *Le test des carrés*, Issy-les-Moulineaux : E.A.P.
- Meili R. (1972), *Le test d'intelligence mécanique, test des leviers*, Issy-les-Moulineaux : Éditions Scientifiques et Psychologiques
- Ochanine D. (1981), "Rôle de l'image opérative dans la saisie du contenu informationnel des signaux", *L'image opérative, Actes d'un séminaire et recueil d'articles de D. Ochanine*, Paris : Université de

Paris I, Centre d'Éducation Permanente, Département d'Ergonomie
et d'Écologie Humaine

- Oléron P., "Les activités intellectuelles" (1963), in Fraisse P., Piaget J. (eds), *Traité de psychologie expérimentale, VII : L'intelligence*, Paris : PUF
- Ombredane A. (1936), *Le problème des aptitudes à l'âge scolaire*, Paris : Herman
- Ombredane A., Robaye F. (1953), "Le problème de, l'épuration des tests d'intelligence étudié sur le matrix-couleur. Comparaison des techniques de réduplication et d'explicitation", *Bulletin du Centre d'Études et Recherches Psychotechniques*, 4
- Ombredane A., Robaye F., Plumail H. (1956), "Résultats d'une application répétée du matrix-couleur à une population de Noirs Congolais", *Bulletin du Centre d'Études et Recherches Psychotechniques*, 2
- Ombredane A., Robaye F., Robaye E. (1957), "Résultats d'une application selon une technique nouvelle du test des relations spatiales de Minnesota à une population de Noirs Asalampasu", *Bulletin du Centre d'Études et Recherches Psychotechniques*, VI
- Orsini-Bouichou F., Hurtig M., Paour J.L., Planche P. (1990), "Une méthode d'apprentissage destinée à analyser les relations entre développement et fonctionnement cognitifs", in Netchine-Grynberg G. (ed), *Développement et fonctionnement cognitifs chez l'enfant, Des modèles généraux aux modèles locaux*, Paris : Presses Universitaires de France
- Paour J.L., "Retard mental et aides cognitives" (1988), in Caverni J.P., Bastien C., Mendelsohn P., Tiberghien G. (eds), *Psychologie cognitive : modèles et méthodes*, Grenoble : Presses Universitaires de Grenoble
- Pasquier D., Thébaud J. (1982), *Test de mathématiques pour le cours préparatoire*, Issy-les-Moulineaux : Éditions Scientifiques et Psychologiques
- Pasquier D., *Test pour le passage en sixième*, Issy-les-Moulineaux, Éditions Scientifiques et Psychologiques, 1985
- Pasquier D., Chiocchetti J., Dupuy D., Neveu M., Brachet J., Jaigu D., Vernet J.M. (1987), "Module de réentraînement et de développement intellectuels -Compte-rendu d'expérimentation-", *Les Cahiers de l'Éducabilité*, 1, AFPA-OP2000, Bourges : AFPA

- Pasquier D., Bougeard C., Corvaisier F., Sicard P., Toulisse C. (1989), *C2a-PA, Évaluation du potentiel d'apprentissage dans une visée psychopédagogique*, Orléans : AFPA-CPR
- Pasquier D. (1989a), "Évaluation dynamique du potentiel d'apprentissage, premier outil, premiers résultats", *Méthodologie et outils de bilan, évaluation-orientation*, Paris : CLP
- Pasquier D. (1989b), "L'éducabilité à l'AFPA : perspectives de recherches et premiers résultats", *Éducation Permanente-AFPA*, n°100/101
- Pasquier D. (1989c), *Test d'évaluation dynamique de l'éducabilité*, Bourges : chez l'auteur
- Pasquier D. (1990a), *L'évaluation en pédagogie, Du bon usage des tests*, Issy-les-Moulineaux, Éditions E.A.P.
- Pasquier D. (1990b), E.D.I. *Intériorisation*, Bourges : chez l'auteur
- Pasquier D. (1990c), E.D.I. *Conceptualisation*, Bourges : chez l'auteur
- Pasquier D. (1990d), E.D.I. *Comparaison*, Bourges : chez l'auteur
- Pasquier D. (1991a), E.D.I. *Axiomatique*, Bourges : chez l'auteur
- Pasquier D. (1991b), *F.A.F. Interaction formative médiatisée.*, Bourges : chez l'auteur
- Pasquier D. (1991c), *F.A.F. Apprentissage, transfert, métacognition*, Bourges : chez l'auteur
- Pasquier D. (1991d), "Regards neufs sur la capacité d'apprendre", *Les cahiers d'information du directeur du personnel*, 18
- Pasquier D. (1992), *Agir pour la réussite scolaire*, Paris : Hachette
- Pasquier D. (1994a), "Le test d'évaluation dynamique de l'éducabilité - T.E.D.E.-", in Huteau M. (ed), *Les techniques psychologiques d'évaluation des personnes*, Issy-les-Moulineaux : EAP
- Pasquier D. (1994b), "Évaluation dynamique, potentiel d'apprentissage et facteur G", in Garbo R., Lebeer .I. (eds), *A la recherche du potentiel d'apprentissage*, Anvers : AEMAM
- Pasquier D. (1996), Un paradigme d'hier pour demain : l'évaluation dynamique du potentiel d'apprentissage, *Actes du congrès*

- Patin J., Vinatier H. (1962), "Le test de calcul de longueurs de J.M. Faverge", *Bulletin du Centre d'Études et Recherches Psychotechniques*, XI, n°2
- Patin J., Vinatier H. (1963), "Un test d'intelligence : le C.E.R.P. 2", *Bulletin du Centre d'Études et Recherches Psychotechniques*, XII, -n°2 –
- Piaget J. (1972), *Problèmes de Psychologie Génétique*, Paris : Denoël Gonthier
- Piaget J. (1975), *L'équilibration des structures cognitives, problème central du développement*, Paris : P.U.F.
- Pichot P. (1965), *Test D48*, Éditions du Centre de Psychologie Appliquée : Paris
- Piéron H. (1979), *Vocabulaire de la psychologie*, Paris : PUF
- Priou P., Bernaud J.L., Simonnet R. (1991), *Manuel de la batterie multifactorielle d'aptitudes*, Issy-les-Moulineaux : Éditions E.A.P.
- Raven J.C. (1970), *Progressive Matrices 1947, révision 1956*, Issy-les-Moulineaux : Editions Scientifiques et Psychotechniques
- Rennes P. (1952), *Test mécanique*, Paris : Éditions du Centre de Psychologie Appliquée
- Reuchlin M. (1950), *L'observation de comportement à partir des tests de Kohs et de Meili*, BINOP
- Reuchlin M., "Épreuves analytiques ou globales, cognitives ou conatives, Évolution des fondements théoriques et perspectives" (1994), in Huteau M. (ed), *Les techniques psychologiques d'évaluation des personnes*, Issy-les-Moulineaux : EAP
- Rey A. (1934), "D'un procédé pour évaluer l'éducabilité : quelques applications en psychopathologie", *Archives de Psychologie*, 24, (96)
- Rey A. (1958), *L'examen clinique en psychologie*, Paris : PUF
- Rey A. (1991), *L'évaluation de l'éducabilité cognitive. (le test des plateaux)*, Issy-les-Moulineaux : Éditions Scientifiques et Psychotechniques

- Rey A. (1982), *Examen clinique en psychologie et techniques psychométriques*, Issy-les-Moulineaux : Éditions E.A.P.
- Robaye F., Robaye E. (1956), "Détermination de la zone discriminative dans un test à item de difficulté croissante", *Bulletin du Centre d'Études et Recherches Psychotechniques*, 2
- Rosenthal A. (1971), Jacobson L., *Pygmalion à l'école*, Paris : Casterman
- Schircks A. (1970), *Test de raisonnement, Évaluation des opérations intellectuelles d'adultes en formation*, Nancy : I.N.F.A.
- Schircks A., Laroche J.L. (1970), "Étude des opérations intellectuelles chez des adultes de la promotion supérieure du travail", *Le Travail Humain*, 33, 1-2
- Schneuwly B., Bronckart J.L. (eds) (1985), *Vygotski aujourd'hui*, Neuchâtel : Delachaux et Niestlé
- Spearman C. (1927), *The abilities of man*, New-York : Mc Millan
- Sternberg R.J. (1994), "La conception triarchique de l'intelligence", in Huteau M. (ed), *Les techniques psychologiques d'évaluation des personnes*, Issy-les-Moulineaux : EAP
- Terman Merrill M. (1959), *Échelle de développement mental, adaptation française de Cesselin*, Paris : Colin Bourrelier
- Tournois J. & Dickes P. (1993). *Pratique de l'échelonnement dimensionnel. De l'observation à l'interprétation*. Bruxelles : De Boeck-Wesmael
- Vermersch P. (1979), "Une application de la théorie opératoire de l'intelligence de J. Piaget aux problèmes de formation", *Éducation Permanente*, 51, déc.
- Vernon P.E. (1950), *The structure of human abilities*, New-York : Wiley
- Vernon P.E. (1969), *Intelligence and cultural environment*, Londres : Methuen
- Voge (non-daté), *Test de logique*, Montreuil-sous-Bois : AFPA
- Vygotski L.S. (1985), *Pensée et langage*, trad. française, Paris : Messidor/Éditions sociales
- Wallon H. (1982), *La vie mentale*, Paris : Éditions sociales

- Wechsler D., *Échelle d'intelligence de Wechsler pour adultes*, édition française, Paris, E.C.P.A.
- Wildöcher D. (1983), "Le modèle psychanalytique de la personnalité", *L'orientation scolaire et professionnelle*, 12, n°3
- Yerle A. (1992), *Étude de la fidélité du potentiel d'apprentissage, mémoire de D.E.A.*, Paris : Université René Descartes, doc. multig.
- Zazzo R., Gilly M., Verba M. (1966), *La nouvelle échelle métrique de l'intelligence*, Paris : Colin Bourrelier

Table des figures

FIGURE N°1 : Étapes de l'élaboration d'un test.....	27
FIGURE N°2 : Quartilage.....	28
FIGURE N°3 : Décilage.....	28
FIGURE N°4 : Distribution normalisée en 5 classes.....	29
FIGURE N°5a : Corrélacion positive parfaite.....	61
FIGURE N°5b : Corrélacion négative parfaite.....	62
FIGURE N°5c : Absence de corrélation.....	62
FIGURE N°5d : Corrélacion positive partielle.....	62
FIGURE N°6 : Représentation imagée de la zone de proche développement.....	75
FIGURE N°7 : Les deux paradigmes de l'évaluation dynamiqu.....	84
FIGURE N°8 : Démarche éducatibilité professionnelle A&E®.....	128
FIGURE N°9 : Articulation interne du T.E.D.E.....	142
FIGURE N°10 : Le conflit cognitif dans la situation T.E.D.E.....	144
FIGURE N°11 : A.C.P. valeurs propres et variance.....	170
FIGURE N°12 : A.C.P. plans principaux 1 et 2.....	172
FIGURE N°13 : Echelonnement multidimensionnel.....	181
FIGURE N°14 : Corrélacions réussites réitérée et spontanée.....	187

Table des documents

DOCUMENT N°1 : Exemple de profil obtenu au <i>Test de mathématiques pour le cours préparatoire</i> . Pasquier, Thébault, 1982	30
DOCUMENT N°2 : Extrait de la liste des items de la N.E.M.I., Zazzo et col., 1966.....	46
DOCUMENT N°3 : Items numériques de la B.G.A., Boss, Cardinet, Maire, Muller, 1960	47
DOCUMENT N°4 : Item spatial de la B.G.A., Boss, Cardinet, Maire, Muller, 1960.....	48
DOCUMENT N°5 : Feuille de profil de la batterie NV7, Priou, Bernaud, Simonnet, 1991	49
DOCUMENT N°6 : Item de combinatoire du <i>Test de raisonnement</i> , Schircks, 1970.....	55
DOCUMENT N°7 : Tableau de passage de la performance au stade dans l' E.C.D.L., I.N.E.T.O.P., 1990	56
DOCUMENT N°8 : Les registres de fonctionnement disponibles, Vermersch, 1979.....	58
DOCUMENT N°9 : Table du r de Bravais-Pearson, Fischer, 1947	63
DOCUMENT N°10 : Droite des moindres carrés, Langouët, Porlier, 1981	64
DOCUMENT N°11 : Étalonnage prédictif du T.P.6, Pasquier, 1985	65
DOCUMENT N°12 : Validité prédictive des outils du psychologue, d'après Lévy-Leboyer, 1987	66
DOCUMENT N°13 : Résultat d'une sélection pour une validité de 0,30, Cardinet, 1986.....	68
DOCUMENT N°14 : <i>Labyrinthe manuel</i> , Rey, 1934	79
DOCUMENT N°15 : Courbe d'apprentissage, Rey, 1934	80
DOCUMENT N°16 : Cinq types de problèmes de matrices, Hurtig, 1960.....	95
DOCUMENT N°17 : Courbes des moyennes d'apprentissage des divers groupes, Hurtig, 1960	97

DOCUMENT N°18 : Modèle global du L.P.A.D., Feuerstein, 1987	107
DOCUMENT N°19 : <i>Leviens de Meili</i> , Meili, 1972	116
DOCUMENT N°20 : Item du <i>Test de Calcul de longueurs</i> , Faverge, 1955	132
DOCUMENT N°21 : Structuration du T.E.D.E. par niveau de complexité, Pasquier 1989.....	138
DOCUMENT N°22 : Structuration du T.E.D.E. par gradient de transfert, Pasquier 1989	139
DOCUMENT N°23 : Modèle triarchique de l'intelligence, Sternberg, 1994.....	145
DOCUMENT N°24 : Nouvelle présentation des items du T.E.D.E	152
DOCUMENT N°25 : Présentation des résultats du T.E.D.E.2000, synthèse	153
DOCUMENT N°26 : Présentation des résultats du T.E.D.E.2000, visuel.....	155
DOCUMENT N°27 : Présentation des résultats du T.E.D.E.2000, profil fonctionnel.....	156
DOCUMENT N°28 : Présentation des résultats du T.E.D.E.2000, profil apprentissage	156
DOCUMENT N°29 : Présentation des résultats du T.E.D.E.2000, profil transfert.....	157
DOCUMENT N°30 : Présentation des résultats du T.E.D.E.-A, profil différentiel.....	158
DOCUMENT N°31 : Présentation des résultats du T.E.D.E.- C.F.A., positionnement et pronostic	159

Table des tableaux

TABLEAU 1 : classification des tests d'évaluation dynamique de l'éducabilité	86
TABLEAU 2 : éléments de relativisation de l'évaluation dynamique	87
TABLEAU 3 : C2a-PA, gain relatif, taux d'interférence et gradient de transfert	123
TABLEAU 4 : T.E.D.E., taux de réussite spontanée et niveau de complexité.....	140
TABLEAU 5 : T.E.D.E. , variation du taux de transfert en fonction du degré de complexité et du gradient de transfert.....	141
TABLEAU 6 : T.E.D.E. , variation du taux d'interférence en fonction du degré de complexité et du gradient de transfert.....	148
TABLEAU 7 : T.E.D.E., sensibilité	162
TABLEAU 8 : T.E.D.E., validité concourante, opératrices sur tours semi-automatiques	163
TABLEAU 9 : T.E.D.E., validité concourante, items pour établir le classement entreprise	164
TABLEAU 10 : T.E.D.E., validité concourante, matrice de corrélations	167
TABLEAU 11 : T.E.D.E. , validité concourante, analyse factorielle	168
TABLEAU 12 : T.E.D.E., validité prédictive et activité lexicale de déchiffrement	173
TABLEAU 13 : T.E.D.E. , validité prédictive et examen partiel de stage de formation de techniciens	174
TABLEAU 14 : T.E.D.E. , validité prédictive et matières d'enseignement.....	175
TABLEAU 15 : T.E.D.E. , A.C.P., coordonnées et corrélations	171
TABLEAU 16 : T.E.D.E. , validité concourante du T.E.D.E.2000	175

TABLEAU 17 : T.E.D.E.2000, validité prédictive	177
TABLEAU 18 : T.E.D.E. , potentiels d'apprentissage et niveaux de qualification	178
TABLEAU 19 : Coordonnées de l'E.M.D	182
TABLEAU 20 : Profils de réponses	184

Index des auteurs cités et des thèmes abordés

A

- abandon.... 11, 20, 147
- accommodats 78
- acquis ..78, 82, 89, 98, 99, 100, 108, 110, 111, 115, 127, 131, 147, 154, 157, 182
- action effectrice... 137
- adaptabilité
 - professionnelle.. 165
- adaptation sociale. 101
- Aebelin H. 94
- affects..... 191, 195
- âge mental. 64, 73, 74, 97
- agent de maîtrise. 164, 165, 166
- aides20, 81, 82, 83, 86, 94, 103, 104, 106, 109, 116, 173, 186, 192
- alpha de Cronbach 176
- alphabétisation..... 171
- amélioration des scores 113
- analyse factorielle. 49, 162, 168
- analyse hiérarchique50, 181, 182
- analyse qualitative 22, 83, 124
- analyse statistique. 48, 50
- anamnèse 40
- ancienneté 169, 170
- anxiété névrotique 102
- appartenance
 - culturelle .. 111, 143
- application des tests17
- apprenant (l') ..14, 34, 71, 76, 113, 135, 154, 190, 195
- apprentissage... 11, 12, 13, 14, 15, 33, 35, 37, 38, 70, 71, 73, 74, 76, 77, 78, 81, 82, 83, 85, 86, 88, 89, 91, 92, 93, 94, 96, 97, 98, 99, 100, 101, 102, 103, 104, 106, 108, 109, 110, 111, 117, 118, 121, 122, 124, 125, 127, 128, 129, 135, 136, 141, 143, 145, 146, 147, 149, 150, 151, 153, 154, 157, 158, 160, 161, 162, 163, 165, 166, 167, 169, 171, 173, 174, 175, 176, 179, 180, 181, 182, 183, 185, 186, 187, 189, 190, 191, 192, 193, 195

- apprentissages
 - techniques 174
 - approche clinique .. 33, 86, 104, 109, 192, 194
 - approche passive
 - acceptante 38, 77
 - approche
 - psychanalytique .. 11
 - approche rééducative 33
 - approches auto-évaluatives 102
 - aptitudes ... 28, 38, 69, 118, 166, 176
 - arbitraire 25, 131
 - artefacts 86
 - ascendante 143
 - aspects fonctionnels 57, 93, 105, 157
 - attention soutenue 163
 - autogestion mentale 147
 - autonomisation..... 147
 - autopassation assistée 151, 153
- B**
- bâclage..... 147
 - barème 21, 26, 29
 - bas niveau de qualification 13, 119
 - Bateson G. 76
 - batteries factorielles 51
 - Bernaud J.L..... 49, 51
 - biais 21, 34, 65, 76, 86, 100, 109, 165, 195
 - bilan des compétences personnelles et professionnelles . 15, 129
 - Binet A. 12, 37, 43, 45, 46, 48, 92
 - bipolaire 170
 - Bonnardel R. 51
 - Boss 51
 - boucles correctives. 39
 - Bougeard C. 122
 - Brachet J. 121
 - Briet S. 127
 - Bronckart J.L..... 75, 77
 - Brown A.L. 103, 104, 161
 - Büchel F.P. 85, 91
- C**
- cadre (hiérarchique) .. 12, 33, 35, 38, 40, 41, 43, 45, 59, 73, 77, 85, 87, 89, 91, 93, 102, 106, 109, 112, 118, 120, 121, 129, 140, 141, 147, 151, 154, 165, 173, 179, 189
 - Canguilhem C. 145
 - capacité 12, 45, 69, 76, 77, 82, 103, 105, 109, 116, 118, 122, 146, 167, 186, 187
 - capacités intellectuelles 11,

43, 65, 166, 169,
 190
 caractère déficitaire
 185
 Cardinet J. 51, 68
 catégorie nosologique
 39
 catégorisation 13, 102,
 119
 centre de rééducation
 professionnelle.. 173
 chaîne de fabrication
 169
 Chartier D. 119
 chef d'atelier. 164, 165
 Chiocchetti J. 121
 Cibois P. 49
 classe de
 perfectionnement. 39
 classes de problèmes
 93, 103
 classification ... 23, 88
 codage 38, 169
 coefficient de
 corrélation.... 61, 62
 cohérence externe.. 40,
 143
 cohérence interne .. 22,
 40, 142, 176
 combinatoire..... 54
 communication. 38, 77
 compensation 143
 comportements 36, 38,
 72, 92, 103, 115,
 165
 composantes ... 33, 39,
 119, 146, 169, 170
 composites..... 43, 46
 conatif..... 34, 190
 conduites cognitives
 . 13, 14, 34, 54, 109,
 185
 conduites sociales 166
 confidentialité 36
 conflit socio-cognitif
 88, 157
 confort.. 150, 151, 194
 conjecture 40
 conjecturer 72
 conseil d'orientation
 33, 36
 conseiller
 d'information et
 d'orientation 11
 consigne.. 20, 111, 137
 consistance
 psychologique.... 50,
 102, 171
 contenu de l'épreuve 21
 contextes... 53, 69, 88,
 192
 contextualisation... 89,
 148
 contraintes 38, 92, 93,
 133, 183
 contribution 13, 185, 1
 contrôle . 93, 112, 133,
 146, 180, 183, 187,
 193, 1
 coordination ... 41, 133
 correction.. 19, 21, 26,
 30, 36, 54, 121, 136,
 151, 176
 corrélations partielles
 62, 186
 Corvaisier F. 122

coupes de volumes 169
 courbe d'apprentissage
 81, 99
 courbe de Gauss 28
 covariation 65, 173
 critère .. 19, 21, 22, 25,
 26, 29, 35, 40, 61,
 62, 65, 67, 68, 70,
 87, 96, 99, 101, 102,
 114, 119, 134, 136,
 137, 140, 162, 165,
 166, 170, 173, 179,
 186, 193
 cursus de formation 61
D
 De Landsheere G. 71
 De Ribaupierre A. 50
 Debray R. 118
 déchiffrage ... 172, 173
 décilage 28, 29
 décodage 136
 déduction . 50, 83, 131
 degré de complexité
 137, 147
 degré de familiarité
 avec la tâche 143,
 147
 degré de maîtrise... 71,
 133, 194
 demande sociale 92, 94
 démarche éducativité
 professionnelle. 127,
 148
 démarche
 interprétative 17
 démonstration 83, 113,
 118, 119, 124, 136,
 137
 démotivation 147
 dendrogramme 180
 déontologie 17
 dépendance-
 indépendance... 146,
 163
 déplacements 151, 153
 dérives..... 37, 88
 descendante 142
 description 26, 57, 72,
 85, 127, 129
 dessin industriel... 118
 déstructuration..... 147
 détérioration..... 102
 déterminants . 193, 194
 développement. 45, 53,
 54, 56, 57, 69, 73,
 74, 75, 76, 77, 78,
 85, 91, 92, 93, 94,
 96, 103, 115, 127,
 154, 166, 189, 190,
 191
 diagnostic . 14, 34, 40,
 76, 78, 85, 93, 101,
 103, 106, 109, 129,
 133, 157
 Dias B. 108
 Dickes P. 180, 182
 didactique
 psychologique..... 94
 différence.. 29, 76, 83,
 89, 92, 98, 99, 125,
 134, 192
 dimension cognitive
 133

dimension latente 180, 182	échantillonnage de population 25
dimension mathématique... 134	échantillonnage des contenus..... 25
dimension praxique 133, 134	échelonnement multidimensionnel 179
dimension subjective 136	éducabilité 12, 13, 14, 15, 34, 35, 78, 80, 82, 84, 88, 91, 92, 100, 105, 106, 112, 115, 121, 127, 131, 135, 138, 140, 141, 169, 171, 195, 1
discriminativité... 125	éducation des choix 11
dispersion 28, 96, 125, 146	effectation 38, 133, 136, 194
dispositions psychologiques... 39, 88	effet d'apprentissage 22
dissociation de facteurs 54	effet de prérequis. 137
distribution normalisée..... 28	effet de rattrapage.. 96
distribution standardisée..... 29	effet de reclassement 13, 98, 110, 125, 178, 195
dogmatique 92	effet de retest . 22, 120
dogmatisme 41	effet piège..... 113
domaine spatial 31	effet Pygmalion ... 173
don 12	efficacité... 23, 67, 82, 89, 105, 147, 163, 165, 170, 181, 185, 194
données figuratives. 34	efficience intellectuelle 33
droite des moindres carrés 66	égalités . 141, 151, 153
Dufay C. 118	Embretson S.E. 120
Dupuy D. 121	encadrement de proximité 127
dysfonctionnements 133	
E	
écart type 28, 176, 183, 185	
échantillon proportionnel 26	

enjeu....36, 39, 59, 76,
 81, 88, 139
 entraînement... 74, 84,
 101, 103, 115, 122,
 123, 125, 137, 146,
 150, 154, 157, 169,
 188, 193
 entretien.... 11, 12, 61,
 136, 143, 147, 150,
 163
 épistémologique 94
 épreuve 19, 20, 21, 22,
 25, 26, 31, 35, 36,
 37, 39, 45, 49, 50,
 55, 56, 69, 70, 74,
 80, 82, 87, 88, 89,
 95, 96, 104, 108,
 110, 112, 113, 114,
 117, 119, 120, 129,
 131, 134, 149, 150,
 151, 156, 157, 161,
 162, 163, 167, 171,
 173, 176, 179, 191,
 192, 195
 épreuve pratique.. 173,
 176
 épreuves factorielles
 43, 50
 épreuves hochepots. 50
 épreuves mosaïques 46
 épreuves opératoires
43, 54, 57, 71
 épreuves parallèles 106
 épuration 115, 125,
 143, 190
 équilibrage de la tâche
 151, 153
 équilibration 145
 équité sociale 13
 espace bi-
 dimensionnel ... 181,
 185
 essai professionnel. 67
 essais et erreurs 81
 établissements
 spécialisés 156
 étalonnage. 21, 27, 29,
 31, 35, 36, 38, 118,
 124, 141, 157, 158
 études docimologiques
 11
 évaluation de
 l'intelligence 13, 51,
 53
 évaluation des
 personnes 128
 évaluation dynamique
 .. 11, 13, 14, 15, 70,
 71, 72, 73, 74, 78,
 81, 82, 85, 87, 88,
 89, 90, 91, 93, 96,
 99, 100, 101, 103,
 105, 106, 107, 110,
 118, 127, 170, 186,
 189, 191, 192, 194,
 195, 1
 évaluation formative
 38, 71, 82
 évaluation
 métrologique 103
 évaluation objective
 12, 35
 évaluation
 psychopédagogique
 101, 129

évaluation scientifique 11, 14
 évaluation sommative 71, 79, 82
 évolution... 67, 71, 79, 81, 92, 109, 117, 121, 122, 140, 146, 150, 157, 165, 191
 examen clinique 33, 109
 examen partiel 173, 174
 exclusion sociale 13
 exécution d'une tâche . 22, 34, 71, 82, 132
 expert 74, 146, 173, 190
 explication 13, 69, 83, 94, 96, 98, 99, 115, 118, 142, 151, 154, 170
 explicitation . 111, 112
 extériorisation 38

F

facettes 146
 facteur cristallisé... 191
 facteur fluide 191
 facteur G ... 49, 50, 51, 88, 114, 120, 131, 134, 167, 168, 179
 facteur général 50, 57, 149, 168, 170, 171, 186
 facteur latent 63
 facteur spatial-mécanique 50
 facteur spécifique... 51

facteur verbal-numérique 50
 facteurs de groupes 50
 facteurs différentiels 156
 facteurs environnementaux 69
 facteurs motivationnels .. 114
 facteurs parasites 111, 114, 115, 143
 facteurs spécifiques 49, 170
 familiarité . 39, 86, 99, 119, 143, 190
 fatigabilité 147
 Fattu N.A. 115
 Faverge J.M. ... 133, 134
 Ferrara R.A. 103
 Feuerstein R. 38, 77, 87, 107, 108, 111, 118
 fiabilité 11, 17, 21, 31, 36, 176, 179
 fiable 31, 86, 148
 fidèle 21, 194
 fidélité. 21, 22, 83, 86, 87, 106, 107, 108, 114, 115, 118, 119, 120, 125, 134, 150, 161, 175, 176, 178, 179
 finesse discriminative 21, 161
 flexibilité.... 118, 137, 171
 fonctions exécutives 194

- fonctions mentales
supérieures 74
- formateurs . 11, 12, 13,
27, 35, 91, 111, 124,
135, 136, 173
- formation continue 15,
109
- formation
professionnelle... 33,
34, 47, 77, 120, 129,
141, 161, 165, 169
- forme parallèle..... 22
- français ... 11, 26, 100,
111, 118, 186
- frustration 102
- G**
- gagnants .98, 102, 103,
110, 117, 125
- gain 36, 74, 76, 83, 85,
101, 120, 124, 141,
150, 161, 165, 166,
167, 179, 180
- gain de performance
..... 74, 165
- généralisabilité 126
- généralisation .89, 122
- gestion des ressources
humaines ... 13, 33, 1
- globalité
psychologique 33
- gradient de transfert
. 104, 105, 122, 125,
136, 137, 141, 147,
151, 153, 157, 192
- graphe des résultats
..... 104
- graphologie 11, 67
- groupe contrôle.... 106
- groupement additif . 53
- guidage pas à pas . 136
- Güthke J. 101
- H**
- habileté..... 76
- Haywood H.C. 108
- hérédité..... 12
- heuristique 35, 106
- Higelé P. 124
- histoire individuelle
..... 137
- hochebot 43, 46
- Hoffman M.B..... 38, 77,
108
- homogénéité interne 22
- homomorphe ... 89, 122
- homomorphie 131, 162
- Hurtig M. 12, 83, 92, 93,
94, 96, 98, 99, 100
- Huteau M. 69, 189
- hyper-contrôlé 183
- hyper-réflexif 147, 183
- I**
- image cognitive ... 136
- image de soi 102, 190,
194
- images opératives. 133
- implication 21, 81
- inadaptation 102
- incapacité.. 39, 77, 189
- indice de stress 179
- indices ... 61, 101, 106,
110, 121, 124, 142,
151, 153, 161, 167,
168, 169, 170, 171,

- 179, 180, 181, 182,
183
- induction.. 50, 83, 131
- inférence inductive
..... 108, 132
- inflexion .36, 129, 138
- inhibition socio-
affective..... 105
- insertion..33, 110, 178
- instrument de mesure
.....21
- intelligence générale
43, 48, 51, 120, 166,
1
- intelligence globale43,
45
- intelligence opératoire
concrète 53
- intelligence opératoire
formelle 53
- intelligence pré-
opératoire 53
- intelligence
sensorimotrice..... 53
- intelligibilité 38
- interaction 38, 74, 136
- interaction sociale.. 74
- intériorisation. 39, 76,
78, 144
- interprétation.. 17, 31,
33, 36, 37, 39, 40,
99, 125, 142, 147,
148, 170, 171, 180,
185, 1
- invariant..... 69, 89
- inventaire
biographique..... 67
- involution.....92, 191
- Ionescu S. 85, 100, 106,
107, 119
- item22, 25, 38, 39, 45,
54, 74, 78, 83, 85,
103, 104, 106, 112,
114, 118, 119, 120,
121, 122, 123, 124,
125, 134, 135, 136,
137, 142, 146, 147,
150, 151, 157, 158,
161, 163, 166, 167,
171, 176, 183, 185,
192
- J**
- Jacobson L..... 173
- Jacquin S..... 127
- Jensen M.R. 87, 108
- jeunes 33, 34, 47, 118,
121, 193
- justesse17, 27, 29, 137
- K**
- Kaniel S..... 87
- Kapos..... 115
- L**
- labilité de l'attention
..... 147, 193
- langage 76, 190
- Langouet G. 170
- Laroche J.L..... 57
- Lautrey J..... 50, 55, 56,
57, 69, 83
- Leplat J. 133
- Lévy-Leboyer C... 12, 67,
176
- liaison statistique... 63

Loarer E. 119
 logicomathématique 50
 logico-propositionnel
 50
 logique de
 transformation.... 70,
 133
 logique des conduites
 38
 logique verbale 50
 logistique 53
 loi de 1991 140
 Longeot F. 54

M

Maire 51
 maîtrise... 25, 41, 132,
 165
 majoration du
 potentiel..... 105
 manières d'apprendre
 34
 marge d'éducabilité 96,
 97
 Mariel R. 83, 100
 Martin B. 124
 Mary C. 83, 100
 mathématiques 17, 31,
 118, 134, 173, 174,
 186
 matrice... 83, 105, 108,
 121, 167, 168, 180
 maturation 65, 191
 mécanisme de défense
 41
 Mech 115
 médiateur 109, 136,
 195, 1

médiation .. 13, 74, 75,
 76, 77, 87, 108, 109,
 110, 144, 190
 Meili R. 117
 métacognition 124
 métacomposantes . 147
 méthode test-retest 161
 métiers... 57, 120, 131,
 165, 192
 micro-monde 185, 187
 milieu .. 12, 69, 98, 99,
 100, 113, 123, 126,
 147, 162
 Miller R. 38, 77
 mise à niveau..... 33
 mises en situation . 67,
 192
 mobilisation
 énergétique..... 39
 mobilité interne 15, 33,
 140, 166
 modalité de passation
 112, 147
 modalité de
 présentation..... 25
 modalité figurative 108
 modalité numérique
 108
 modalité verbale 26, 34
 modalités de codage
 39, 45, 108, 109
 mode opératoire 83,
 88, 136, 137, 142,
 143, 144, 146, 149,
 150, 169
 modèle interprétatif 40
 modèle théorique ... 41
 modèle triarchique 146

modes de résolution 34
 modifiabilité
 structurelle
 cognitive 83, 87, 109
 modules. 127, 165, 169
 monde physique ... 191
 motivations.....65, 166
 moyenne.45, 101, 106,
 113, 120, 121, 169,
 173, 176, 183, 184,
 185
 Muller 51
N
 nature de la tâche ..99
 neutralité
 bienveillante 20
 Neveu M..... 121
 névrose infantile ... 11
 niveau d'aspiration 102
 niveau scolaire 33, 98,
 178, 192
 non-gagnants .98, 102,
 110, 117, 125
 normalité..... 145
 normativité 146
 notation 11, 19, 21, 26,
 30, 36, 119
 note brute .. 21, 27, 29,
 31, 54
 note pédagogique ..21,
 29
 note standard .. 21, 29,
 31
 novice 74, 190
O
 objectivation.... 43, 85
 objectivité 29
 observation 53, 61, 71,
 92, 118, 119, 141, 1
 occurrence.... 137, 166
 Ochanine D. 133, 145
 Oléron P. 122
 Ombredane A.12, 82, 94,
 111, 112, 114, 115,
 126, 143
 opérateurs ... 118, 127,
 134, 162
 opération... 31, 38, 48,
 50, 53, 54, 55, 59,
 105, 108, 131, 133,
 134, 136
 opérations projectives
 133, 134
 opérativité.... 133, 145
 ordinateur .12, 23, 133
 organisations 69
 Orsini-Bouichou F..... 92
 orthogénèse43, 53
 ouvriers spécialisés
 165
P
 pair-impair 22, 106,
 120
 Paour J.L. 85, 91, 92, 93
 papier-crayon 23
 paradigme . 78, 82, 83,
 88, 90, 91, 98, 100,
 103, 135, 150, 157,
 158, 188
 paramétrage.....88, 91,
 135, 162
 paramètres. 19, 63, 82,
 86, 109, 110, 133

Pasquier D. . 11, 17, 83,
 85, 121, 125, 127,
 161, 165, 169, 174
 passation collective 54,
 111, 135
 passation individuelle
 47, 54, 109, 112,
 113
 Patin J. 120, 134
 patrons de réponses
 102, 114, 193
 percentile 156
 perception 76, 112,
 146, 190, 193, 194
 perdants 125
 période 53, 82, 98,
 104, 115, 186
 perturbations .. 20, 39,
 143
 Piaget J. 43, 53, 54,
 143, 189
 Pichot P. . 167, 179, 191
 Piéron H. . 19, 21, 22, 48
 plan expérimental . 106
 Planche P. 92
 plans de recherche 191
 plasticité 69
 Plumail H. 113
 points faibles 154, 160
 points forts ... 154, 160
 population de
 référence 21, 26
 population parente . 21,
 26, 29, 35
 Porlier J.C. 170
 positionnement 12, 29,
 31, 38, 71, 86, 151,
 152, 170
 poste de travail 61,
 127, 129, 132, 140,
 162, 166, 1
 post-test 74, 82, 83,
 106, 116, 122, 124,
 125, 136, 137, 141,
 143, 146, 147, 161
 postures 40
 potentiel
 d'apprentissage 177,
 180, 182, 183
 potentiel d'éducabilité
 78, 86, 93, 138, 169
 pragmatique.. 150, 189
 pratique des tests .. 19,
 38
 praxies 133
 prédicteur.. 22, 61, 62,
 65, 66, 67, 68, 70,
 173, 174, 178, 179
 préformation.. 33, 118,
 169
 pré-test 74, 82, 85, 98,
 106, 124, 125, 136,
 137, 138, 141, 143,
 147, 161, 162, 165,
 166, 167, 169, 174
 prévision.. 22, 67, 173,
 178
 Priou P. 49, 51
 prise de conscience 14,
 99, 190
 prise en considération
 39, 69
 probabilité de réussite
 22, 90
 probabilités 54

problématique 89, 138, 140, 147
 procédures de recrutement..33, 140
 procédures statistiques 25
 procès .76, 81, 92, 136
 processus cognitifs 23, 34, 91, 190
 processus d'accommodation .79
 processus de prise de décision..... 140
 processus de régulation 92
 processus d'involution 120, 168, 191
 processus fonctionnel ... 71, 136, 193, 195
 processus inductifs 133
 processus mental ... 71, 85
 processus psychologiques.... 19
 profil apprentissage 154
 profil fonctionnel . 154
 profil pédagogique 31
 profil transfert 154
 progression..... 70, 85, 112, 113, 123, 141, 158, 170, 185, 187
 pronostic ... 14, 34, 76, 78, 85, 89, 93, 101, 106, 129, 140, 179
 proportionnalité 54
 psychologie anglo-saxonne..... 12
 psychologie de l'enfant ... 57, 85, 91
 psychologie différentielle 57
 psychologie génétique 57
 psychologie individuelle 33
 psychologie industrielle..... 69
 psychologue scolaire 11, 33
 psychométrie ... 13, 19, 85, 99
 psychopédagogie... 15, 33, 35, 105, 109, 125, 191, 193, 194, 195
 psychotechnique ... 19, 113
 publics spécifiques 157
 publics tout-venant 156
Q
 Q.I. 39, 69, 88, 92, 100, 101, 103, 104, 148
 qualification professionnelle 127, 165
 qualités métrologiques ... 21, 106, 120, 129, 149, 151, 161, 178, 179, 2
 quantification 19
 quartilage..... 28

questionnaire de
 personnalité 12
questionnaires
 d'intérêts 33
questionnement 39,
 126, 165, 195
quota 26
quotient intellectuel 92

R

raisonnement inductif
 89
Rand Y. 38, 77, 87, 108
rapport de corrélation
 173
rapport du sujet au
 savoir..... 136
Raven J.C. . 50, 94, 100,
 101, 105, 108, 120,
 178
réactivité... 74, 81, 82,
 86, 117, 167, 171,
 174, 1
réalisation
 automatique des
 attentes 171
recherche... 12, 15, 37,
 40, 59, 73, 85, 92,
 94, 106, 109, 111,
 112, 120, 131, 142,
 147, 178, 189, 190
reconversion . 166, 193
recrutement 15, 36, 67,
 85, 128, 140, 177
réduplication 112, 114,
 115, 120, 137, 143
rééducation.... 93, 105,
 109

référence... 12, 26, 29,
 59, 104, 122, 189
réflexifs 104
registres de
 fonctionnement ... 57
règles alphabétiques
 104
règles déontologiques
 36
règles interprétatives
 148
régression 83, 125,
 143, 180, 181
régulation.. 39, 71, 146
réification 88, 148
relations causales... 50
relations mécaniques
 50
relativisation ... 13, 26,
 88, 89, 148
remédiation 13, 14, 35,
 38, 39, 59, 76, 78,
 87, 91, 102, 105,
 109, 127, 129, 148,
 151, 157, 158, 165,
 193, 1
rendement scolaire . 33
Rennes P. 134, 166
réponse technique .. 94
représentatif ... 25, 134
représentations
 mentales..... 53
résistance..... 143
résolution de problème
 31, 57, 70, 115, 118,
 132, 167
restitution ... 136, 143,
 147, 150, 163, 173

retard mental ... 96, 98
Reuchlin M. 118, 194
réussite 14, 19, 34, 43,
61, 65, 66, 67, 69,
70, 74, 81, 83, 98,
99, 101, 103, 104,
110, 114, 120, 122,
125, 129, 134, 137,
140, 141, 146, 151,
153, 165, 166, 167,
169, 174, 178, 179,
180, 185, 186, 187,
192, 194, 195, 1
Rey A. 12, 69, 72, 78,
79, 81, 94, 107, 108,
136
Rieben L. 50
risque d'erreur.. 39, 63
Robaye E. 114
Robaye F. 111, 112, 114,
115, 121
Rosenthal A. 173
rotation (A.C.P.) . 105,
123, 168

S

salariés non-qualifiés
..... 165
santé mentale 145
saturé 149, 170
Schircks A. 54, 57
Schneuwly B. 75, 77
score d'amélioration 99
secret professionnel 37
sélection.... 11, 67, 85,
87, 112, 120, 128,
134, 141, 147, 169,
173

sens 13, 19, 34, 35, 36,
38, 45, 46, 57, 70,
71, 73, 86, 89, 90,
92, 93, 94, 96, 98,
105, 107, 112, 120,
125, 137, 139, 144,
148, 167, 173, 178,
179, 186, 191, 195
sensibilité . 74, 86, 87,
109, 121, 125, 161,
176
sensible..... 13, 21, 98,
177, 191, 195
série de lettres 104
série d'images 100
série numérique 83
sexe..... 126, 156, 177,
192
signe 193
significativité 63, 165,
174, 175
significativité
statistique . 174, 175
Simon T. 45, 47, 104
Simonnet R. 51
situation de travail
..... 131, 133
situations évaluatives
..... 70, 143
sondages d'opinions 26
sous-stades 53
Spearman C. . 43, 48, 49,
161, 182
stabilité 17, 21, 22, 86,
104
stade 53, 54, 55, 56,
57, 191

standardisation 19, 26,
 36, 104
 Sternberg R.J. 146
 stratégies
 métacognitives 89
 stratification sociale
 12
 stress 89, 102, 143,
 150
 structuration .. 40, 108,
 147
 structure 22, 53, 54,
 83, 94, 132, 136,
 140, 142, 180, 190
 structure arithmétique
 136
 structure manifeste 180
 sub-tests 22
 suites numériques .. 50,
 88
 superposition 137
 surcharge cognitive
 147
 symétrique 29
 symptôme 11, 193
 système éducatif 33
 systèmes de codes .. 76
T
 tautologie 53
 taux de transfert ... 137
 technologie 118
 tenue dans le temps
 157, 158
 test collectif 23
 test individuel 23
 test statique 12, 13, 69,
 82, 135
 test-retest 21, 98
 tests de connaissances
 23, 35, 78
 tests de personnalité 23
 tests d'intelligence 23,
 41, 43, 59, 61, 70,
 78, 90, 101, 1
 tests psychométriques
 54, 92, 93, 192
 théorie constructiviste
 43
 théories de
 l'intelligence 189
 tirage au sort 25, 26
 tonalité émotionnelle ... 143
 Toulisse C. 122
 Tournois J. 180, 182
 traces graphiques .. 40,
 124, 142
 traitement analogique
 50
 traitement digital ... 50
 traits psychologiques
 23
 trajectoires
 socioprofessionnelle
 s 129
 transférer les acquis
 103
 transfert 14, 15, 85,
 103, 104, 105, 119,
 121, 122, 124, 125,
 127, 136, 146, 147,
 160, 170, 179, 180,
 181, 182, 183, 187,
 192
 transfert direct 122,
 170

transfert éloigné... 137
 transfert flexible .. 170
 transmission des
 résultats37, 174
 tronçonnage 66
 type de tâche .25, 104,
 194
 types d'apprentissage
 94
 Tzurriel D..... 87
U
 unicité 143
 unimodale..... 29
V
 valeur explicative 186,
 187
 valeur heuristique.. 76,
 157
 valeur prédictive ... 67,
 101
 valeur vraie ... 17, 154,
 190
 valeurs propres 170
 valide 21, 67, 178, 194
 validité concourante
 .. 22, 114, 134, 162,
 165, 166, 169, 176
 validité de contenu 22,
 25, 162
 validité empirique ..22
 validité prédictive .22,
 61, 69, 70, 119, 134,
 149, 171, 173, 176,
 179, 186, 191, 192
 variable dépendante
 180
 variable latente 65,
 182, 185
 variable
 socioculturelle 98
 variable
 supplémentaire.. 167
 variables conatives. 67
 variables de
 personnalité..... 67
 variables
 indépendantes ... 180
 variables
 psychologiques
 latentes 50
 variables
 supplémentaires 170
 variance 170, 180, 186
 variations des
 progressions 114
 verbalisations 40
 Vermersch P..... 57
 Vernet J.M..... 121
 Vernon P.E. 50, 70
 Vinatier H. 120, 134
 vitesse76, 92, 104,
 151, 153
 Voge..... 167
 Vygotski L.S. 12, 72, 73,
 74, 76, 77, 78, 90,
 96, 105, 112, 125,
 144, 190
W
 Wallon H..... 37, 189
 Wechsler D. 47
 Wildöcher D..... 40

Y
Yerle A.....83, 106, 107

Z
Zazzo R..... 47, 91
zone de dispersion 112

zone de proche
développement... 73,
74, 76, 77, 78, 86,
90, 93, 98, 103, 105,
112, 125, 190
zone d'incertitude. 112

Table des matières

Préface par Claude Lemoine.....	7
Première section : théorie	11
Première partie : les tests	17
Chapitre 1 : définition des tests	19
Chapitre 2 : élaboration des tests	25
Chapitre 3 : domaines, conditions d'utilisation et interprétation des tests.....	33
Deuxième partie : les tests d'intelligence.....	43
Chapitre 4 : les tests d'intelligence globale et d'intelligence générale.....	45
Chapitre 5 : Les tests opératoires.....	53
Chapitre 6 : Les tests d'intelligence et la prédiction de la réussite des apprentissages professionnels	61
Troisième partie : L'évaluation dynamique du potentiel d'apprentissage.....	71
Chapitre 7 : Les précurseurs et les principes généraux de l'évaluation dynamique	73
Chapitre 8 : L'évaluation dynamique du potentiel d'apprentissage de l'enfant	91
Chapitre 9 : L'évaluation dynamique du potentiel d'apprentissage de l'adulte	111
Deuxième section : le T.E.D.E	127
Chapitre 10 : Le T.E.D.E. historique	131
Chapitre 11 : Les différentes versions du T.E.D.E.....	149
Chapitre 12 : Les qualités métrologiques du T.E.D.E	161
Conclusion.....	189
Bibliographie	197
Table des figures.....	207
Table des documents	208
Table des tableaux	210
Index des auteurs et des thèmes.....	213
Table des matières	231

Daniel Pasquier

Les compétences à apprendre. Évaluation chez l'adulte.

La question de l'évaluation des capacités d'apprentissage est revenue à l'ordre du jour, par exemple à travers l'implantation des centres de bilans de compétences. Mais comment évaluer ces capacités d'apprentissage? Par cet ouvrage, l'auteur apporte sa contribution en réponse à cette question. Suite à une présentation des principes d'élaboration et d'usage traditionnel des principaux types d'épreuves l'auteur ouvre la réflexion sur la perspective de l'évaluation dynamique. Dynamiser l'évaluation améliore les qualités de la mesure du potentiel d'apprentissage, dépoussière l'approche méthodologique routinière, stimule l'évaluateur dans un nouveau rôle de médiateur, élargit le champ de l'observation aux relations qui permettent d'expliquer la performance par le processus psychologique mis en oeuvre. Il en résulte des pronostics plus fiables et des diagnostics plus finement différenciés, ce qui facilite de manière décisive le passage entre l'évaluation et le projet, entre l'évaluation et la remédiation, entre l'évaluation et la formation, entre l'évaluation et l'adaptation au poste de travail...

Ce livre s'adresse au psychologue, au formateur, au gestionnaire des ressources humaines... soucieux d'entretenir et de renouveler sa réflexion et sa pratique professionnelle sur le thème de l'évaluation des compétences à apprendre.

Daniel Pasquier dirige le cabinet Avenir & Entreprise, conseil en gestion des ressources humaines. Ses thèmes de recherches actuels concernent l'éducabilité professionnelle des adultes, la localisation du contrôle des renforcements, la réactivité normative et la gestion des ressources attentionnelles.