

Biographie de l'auteur

Postulats de l'auteur

Hypothèses

Démarche de l'auteur

Résumé de l'ouvrage

Actualité de la question et analyse personnelle

Biographie de l'auteur

Claude LEVY-LEBOYER est professeur émérite de Psychologie du travail à l'Université René-Descartes Paris V. Elle est Past President de l'Association Internationale de Psychologie Appliquée et fondateur de l'ENOP (European Network of Organizational Psychology). Outre des activités de consultant auprès de grandes entreprises, elle poursuit des recherches dans le cadre de l'Institut de Recherches et d'Applications en Psychologie du Travail.

Les ouvrages de l'auteur

RH Les apports de la psychologie du travail

La motivation dans l'entreprise

La psychologie du travail

Le 360°

La gestion des compétences

Postulats de l'auteur

Le recrutement et plus généralement la gestion des ressources humaines sont aujourd'hui un élément essentiel à la bonne réussite de l'entreprise. Il est important de parvenir à une bonne adéquation « homme/poste », ce qui nécessite de disposer de bonnes méthodes d'évaluation des candidats potentiels. Or, ce n'est pas parce qu'une méthode est intuitive ou usuelle qu'elle est nécessairement valide et qu'elle saura prédire le succès professionnel futur. C'est pour cela qu'il est nécessaire « d'évaluer l'évaluation » à l'aide de critères précis et objectifs, et de procéder de manière scientifique en s'appuyant notamment sur des méthodes statistiques. Les qualités métriques d'une évaluation sont un préalable à son utilisation. Ces différents points font l'objet et l'utilité de nombreuses recherches sur le sujet qui peuvent aider les professionnels du recrutement dans leurs choix.

Hypothèses

Devant l'importance de l'étape de sélection des candidats, de nombreuses questions se posent aux responsables des ressources humaines. L'auteur tente d'apporter des réponses aux plus importantes d'entre elles.

Quelles sont les demandes auxquelles l'évaluation est appelée à répondre ? Quelles doivent être les qualités métriques des tests et questionnaires d'évaluation ? Quelle est la nature des informations qu'apportent les méthodes existantes et comment les utiliser ? Sur quels critères peut-on faire un choix entre les différentes méthodes ? Quelle est la validité des différentes méthodes de sélection ? Quel est le rôle des aptitudes et des traits de personnalité ? Que penser des entretiens, des observations en situation ? Comment faire le point des compétences acquises ?

Démarche de l'auteur

Dans une première partie, l'auteur fournit les bases théoriques nécessaires à la compréhension de l'ouvrage. Elle énonce les principes et les concepts qu'il faut connaître pour choisir une ou des méthodes d'évaluation et pour évaluer la qualité de l'évaluation c'est-à-dire pour en estimer la pertinence, l'utilité et la valeur des indications. Elle aborde tout d'abord l'analyse du travail dont elle souligne l'importance dans la phase amont du recrutement, puis décrit la théorie de la notation professionnelle et les différentes définitions de réussite professionnelle qui président à tout choix de méthode d'évaluation. Dans un troisième chapitre Claude LEVY-LEBOYER décrit les qualités métriques nécessaires à toute méthode d'évaluation en insistant sur l'importance de la validité de contenu. Enfin l'auteur traite des processus de décision et du développement individuel en les mettant en perspective avec les critères des méthodes d'évaluation utilisées.

Dans une seconde partie, l'auteur met en application les notions théoriques présentées dans la première partie. Claude LEVY-LEBOYER énumère en effet de la manière la plus exhaustive possible les différentes catégories de méthodes d'évaluation existantes, en explique le contenu, et donne des éléments sur leur signification et leur valeur au regard des critères de validité définis dans la première partie. L'auteur s'attache à présenter en complément une synthèse des études critiques et/ou scientifiques dont chaque méthode d'évaluation fait l'objet en décrivant lorsque c'est nécessaire les protocoles expérimentaux qui conduisent à l'estimation de la validité et du caractère prédictif des méthodes. Au fil des différents chapitres, Claude LEVY-LEBOYER passe en revue le « trio classique » CV, entretiens, références, puis aborde des méthodes plus avancées comme les tests d'aptitudes cognitives, les tests de personnalité et les essais professionnels et tests de situation. A chaque fois, Claude LEVY-LEBOYER décrit de manière objective les avantages, les inconvénients mais aussi les limites des méthodes d'évaluations étudiées en s'appuyant sur l'état de l'art de la recherche dans le domaine.

Résumé

PARTIE 1: COMMENT CHOISIR UNE MÉTHODE D'ÉVALUATION ?

Choisir une méthode d'évaluation se fait en trois étapes : tout d'abord, l'analyse du travail (1^{er} chapitre), l'évaluation du succès professionnel (2^{ème} chapitre) et les qualités métriques que doivent posséder les méthodes d'évaluation (3^{ème} chapitre - le plus abstrait mais aussi le plus important). Le chapitre 4 traite de l'utilisation des méthodes d'évaluation pour la prise de décision. Enfin, le chapitre 5 aborde la qualité des méthodes d'évaluation du point de vue de la personne évaluée.

Chapitre I - Une étape essentielle : l'analyse du travail

L'analyse du travail qui doit précéder le recrutement est aussi une étape capitale pour la conception d'une méthode d'évaluation. Il faut parvenir à relier les aptitudes humaines et les activités de travail afin de trouver une méthode d'évaluation valide. L'analyse du travail est le moyen de déterminer ce qui doit être prédit par la méthode d'évaluation.

Il existe trois objectifs de l'analyse du travail : construire des essais professionnels, aider à choisir les prédicteurs pertinents qui composent une méthode d'évaluation et créer des groupes de métiers qui auront des prédicteurs communs, facilitant ainsi les analyses et évaluations.

L'analyse du travail débouche sur une description de poste ou de fonction qui comporte six rubriques distinctes (Smith et Robertson, 1986) : les objectifs du poste, ses responsabilités (ce que le titulaire doit faire), les relations avec les autres (collègues ou clients), les conditions physiques du travail (horaires, exigences physiques, éventuellement risqués) et les conditions de rémunération et promotion.

Les Méthodes

Il existe deux approches possibles, l'une dirigée vers le travail et les tâches à accomplir, l'autre orientée vers le travailleur et ses activités.

La méthode la plus fréquemment utilisée est l'entretien ouvert avec les titulaires du poste à analyser et leur hiérarchie. Cette méthode est facile à mettre en œuvre mais prend du temps ; elle est féconde en terme d'informations recueillies mais leur synthèse systématique est difficile. On peut alors utiliser un questionnaire portant sur les activités les plus importantes et/ou les plus fréquentes. Pour faciliter ce travail, il est possible d'utiliser des questionnaires existants comme le PAQ (Position Analysis Questionnaire) de McCormick qui contient 189 descriptions d'activités regroupés en 5 catégories. Il a été construit à partir d'une recherche exhaustive et fournit une échelle de notation fiable (corrélations élevées) et exploitable (descriptions précises avec notation comparative).

Le questionnaire utilisé doit être fidèle afin de mesurer « exclusivement ce qu'on veut lui faire mesurer ». Cette fidélité peut être contrôlée en vérifiant que les résultats sont constants lorsque :

- l'analyse du même poste est faite deux fois par le même analyste
- l'analyse du même poste est faite par deux analystes différents
- l'analyse de postes équivalents est faite par le même analyste.

La formation de l'analyste apparaît comme l'élément essentiel dans la fidélité de cette méthode (Smith et Robertson 1986).

Il existe également d'autres méthodes d'analyse comme la tenue d'un journal personnel par les titulaires du poste pour consigner leurs activités. Cette méthode est facilement accessible mais lourde à réaliser. Une autre méthode consiste à faire observer directement le travail par un tiers. En apparence facile, elle omet tout ce qui n'est pas observable comme le travail mental.

On peut aussi recueillir des descriptions concrètes de comportement qui jouent un rôle clé sur la qualité du travail effectué, en analyser le contenu et dégager enfin les aspects le plus importants du travail (Flanagan 1954). Cette méthode dite des « incidents critiques » reste cependant incomplète et difficilement exploitable.

Enfin, la méthode de la « répertory grid » consiste à dégager les aspects spécifiques d'un poste une utilisant la méthode de triangulation c'est-à-dire de confrontation entre trois éléments significatifs (il peut s'agir de postes, activités, compétences d'individus). Poussant loin l'analyse, cette méthode requiert une formation du notateur.

L'analyse de poste n'est pas une science exacte. Elle comporte des erreurs et approximations dues aux faits suivants : les descriptions des observateurs sont souvent incomplètes, elles donnent beaucoup de détails qui nécessitent de prendre du recul, et enfin, elles rendent compte d'une situation existante plutôt que d'une manière idéale de faire.

Pour tenir compte de l'évolution des métiers, l'analyse de poste doit accepter l'existence de plusieurs façons d'accomplir une tâche et de plusieurs profils adéquats pour remplir la mission. Elle doit comparer ces différentes manières de réaliser le travail et en retirer la meilleure. Elle doit aussi tenir compte des aspects contextuel (comme le travail en équipe de plus en plus répandu), et des aspects comportementaux et relationnels. Comme il n'existe pas de méthode universelle parfaite, il faut choisir la méthode la plus adaptée au poste à évaluer et qui la rendra donc la plus fiable. Si les analystes sont bien formés et utilisent un vocabulaire commun et précisément défini, la fiabilité sera d'autant plus grande.

L'utilisation de l'analyse du travail

Comment utiliser l'analyse du travail ? La première façon consiste à utiliser le tableau des activités de travail pour construire directement des techniques d'évaluation constituées par des échantillons des tâches effectuées dans le poste. On peut alors bâtir des méthodes de test dites « en situation » qui peuvent atteindre de fortes corrélations (0.40 à 0.55).

La deuxième utilisation consiste à faire la liste des principales caractéristiques nécessaires dans le poste et à mettre cette liste en relation avec la méthode d'évaluation à utiliser. On en déduit des prédicteurs de succès professionnels qui devront être validés par la suite.

On peut également utiliser l'analyse du travail dans le cadre de la théorie de la « généralisation de la validité » qui montre que la validité prédictive, loin d'être limitée à un poste, peut s'étendre à des groupes de professions qu'il convient de construire soigneusement ; l'analyse du travail permet de réaliser cette classification.

La classification des métiers

La classification des métiers peut être utilisée dans le cadre de la validation et de l'utilisation des prédicteurs. Les méthodes de classification sont choisies en fonction de l'usage final (orientation professionnelle, formation, etc.) mais toutes sont fondées sur des critères descriptifs des tâches et métiers composant l'ensemble à classer. Ces caractéristiques peuvent être variées mais doivent impérativement être objectives, pouvoir être évalués, et pouvoir être confrontés aux opinions de plusieurs juges. On retrouve souvent 4 catégories : nature du poste, résultats du travail accompli, conduites impliquées dans le travail et aptitudes et qualités requises. Enfin, il est nécessaire que le système de classification retenu soit exhaustif (il concerne tous les postes) et exclusif (un poste appartient à une seule famille).

Quelles conditions pour généraliser la validité

Pour pouvoir généraliser la validité d'une méthode d'évaluation (démontrée antérieurement) à l'ensemble des postes d'une même catégorie, il faut selon Pearlman créer des familles de postes telles que « les validités spécifiques des méthodes d'évaluation soient généralisable à l'intérieur de la famille mais pas entre les familles ». C'est le cas en particulier lorsque les familles de postes sont constituées à partir de l'analyse du travail (Algera 1987).

A quel niveau d'agrégation doit se faire la classification ? Le débat reste ouvert entre les défenseurs d'une agrégation « moléculaire » (Pearlman, Schmidt et Algera) et les promoteurs d'un modèle fondé sur la mesure des seules grandes aptitudes cognitives (Hunter 1987) pour prédire efficacement la réussite professionnelle. La vérité se situe probablement entre les deux et dépend fortement du type de position à évaluer.

Des thèmes de recherche pour l'avenir

L'analyse du travail est donc essentielle pour l'évaluation. On se rend compte actuellement de la nécessité de poursuivre des recherches pour améliorer le processus permettant de relier l'analyse du travail aux qualités requises et les méthodes de mesure de ces qualités.

Tout d'abord, il serait intéressant d'approfondir le débat entre la validité spécifique et la validité généralisable, notamment sur les conditions de la généralisabilité que l'on pourrait mieux comprendre avec les outils informatiques (influence des facteurs situationnels dans les prédicteurs par exemple).

Ensuite, il faudrait plus s'intéresser aux schémas mentaux impliqués dans l'analyse du travail qui sont relativement absent dans les descriptions des tâches.

Enfin, il devient nécessaire de tenir compte de la nature dynamique des postes, d'intégrer l'évolution de leur contenu et d'anticiper les changements à venir (organisation, compétences, etc.).

Chapitre II - Notation professionnelle : comment définir concrètement la réussite ?

Les objectifs

Pour mettre en place un programme d'évaluation, il faut d'abord choisir des méthodes pertinentes pour prendre des décisions de gestion du potentiel humain. Il est ensuite nécessaire de vérifier la validité des méthodes choisies fondée sur la comparaison entre le diagnostic préalable à la décision d'affectation de poste et la réalité de la réussite ultérieure dans le poste. Il est donc nécessaire de pouvoir mesurer la réussite professionnelle.

Ces indicateurs de succès doivent être soumis à l'examen critique permettant de s'assurer qu'ils remplissent des critères indispensables de fiabilité, de réalisme et de discrimination des candidats.

En outre, cette réflexion permet souvent de faire préciser ce que l'entreprise attend de ses salariés en termes de conduites professionnelles.

Gérer les ressources humaines

Les notations professionnelles ont plusieurs objectifs : validation des outils de sélections mais aussi outils de gestion des ressources humaines de plus en plus importants dans les entreprises (systèmes de rémunération, promotion, etc). On peut s'interroger sur la capacité des notations professionnelles à distinguer les performances dues aux seuls individus de celles qui sont liées aux moyens fournis et au contexte organisationnel. Ces notations restent cependant l'outil utilisé pour les discussions avec les hiérarchiques sur les possibilités de développement. Lors de l'élaboration d'une méthode d'évaluation, on ré-examinera le

système de notation utilisé, ce qui aura également comme conséquence de le clarifier au sein de l'entreprise, effort toujours bénéfique.

En premier lieu, on traduira les objectifs organisationnels en objectifs personnels concrets. Cela permet de poser les objectifs individuels comme contribution aux objectifs globaux. En second lieu, on déterminera en fonction de ce critère de succès les besoins et actions de formations.

Quels problèmes ?

Le choix d'un indicateur de succès professionnel soulève des problèmes techniques et théoriques. Le critère doit-il être unique ou multidimensionnel ? D'abord favorable à un critère unique dans les années 1950, l'état de la recherche s'accorde aujourd'hui sur la nécessité d'utiliser des descripteurs multiples sous peine de simplifier exagérément la réalité et la validation des méthodes.

Cependant, lorsque l'évaluation doit fonder une décision entre plusieurs éventualités, il faut un critère unique, éventuellement composite de plusieurs sous critères, qu'il peut être utile de pondérer. On peut également introduire des seuils sur les sous critères pour tenir compte des impératifs demandés sur certains aspects de la notation.

Autre problème théorique : le changement de la performance des individus dans le temps donne une signification différente à un système de notations. Enfin, le critère peut être « contaminé » par les conditions de travail (qualité des équipements par exemple).

En conséquence, aucun critère aussi bon soit-il ne peut être parfait et tout critère néglige forcément un aspect de la réussite professionnelle.

Les indicateurs objectifs

Bien qu'en apparence indiscutables, ces indicateurs sont les moins utilisés. Les indicateurs objectifs (comme ceux liés à une mesure de production par exemple) sont finalement difficiles à obtenir et pas assez fiables. D'autre part, un certain nombre de métiers ne se prêtent pas à ce type de critère. Enfin, il est très difficile de définir une norme différenciant réellement les « bons », les « moyens » et les « mauvais » éléments. D'autres types d'indicateurs objectifs comme l'absentéisme, l'instabilité professionnelle et le rythme de progression dans la carrière doivent tous être utilisés avec précaution souvent parce qu'ils rendent plutôt compte des facteurs externes (comme les conditions de travail, les facteurs organisationnels) que de la performance individuelle.

Il est d'autant plus nécessaire d'examiner les relations précises de ces critères avec les tâches résultant de l'analyse du travail, de s'assurer de leur non contamination et de vérifier que leur fidélité est satisfaisante.

Les indicateurs subjectifs

Ce sont les plus utilisés en pratique quel que soit l'usage final (primes, promotions, formations). Parmi les méthodes utilisées, la plus simple est le classement des individus, dont les recherches ont montré qu'elles sont fidèles et utiles lorsqu'on recherche une classification globale du personnel.

Les échelles de notations

Les échelles de notation visent à corriger les erreurs que l'on observe lors de la notation par un individu d'un autre individu. Ces erreurs sont dues à l'indulgence du notateur, l'effet de « halo » et l'erreur de « tendance centrale ».

L'effet d'indulgence peut être traité en forçant le notateur noter avec des pourcentages fixés de meilleurs et de moins bons. L'effet de halo est caractérisé par la prédominance d'une impression générale chez le notateur, impression qui va influencer (de façon positive ou négative) toutes les notes. Pour y remédier, on peut former les notateurs sur les effets de halo et leur apprendre à noter de manière spécifique les différents comportements. On peut également adapter le système de notation.

L'effet de « tendance centrale » provient des difficultés des notateurs à noter un individu ce qui les conduit à donner de nombreuses notes moyennes. On peut y remédier en demandant de noter sur la base de descriptions pertinentes des comportements, éventuellement pondérées en fonction de l'analyse du travail, ou avec l'aide de questionnaires à choix multiples.

Les fiches de notation sont les méthodes les plus utilisées dans les entreprises.

Des méthodes plus sophistiquées ont été mises au point avec des descriptions de comportement regroupées de façon cohérente, validées par des panels d'experts avant d'être mélangées et soumises à notation. Deux échelles, l'une comportementale (BARS - Behaviourally Anchored Rating Scales) et l'autre fondée sur l'observation (Behavioural Observation Scales), ont été mises au point pour améliorer la fiabilité des fiches de notations. Elles utilisent la méthode des incidents critiques de Flanagan qui permet de relier les comportements à ce qui se passe réellement sur le terrain.

La notation, fonction de la situation ?

La notation n'échappe pas aux intérêts personnels des notateurs. Les aspects politiques sont plus importants au niveau des cadres qu'au niveau des échelons inférieurs car la performance contextuelle des cadres est centrale. Elle est également plus sujette à interprétation.

Chapitre III - Qualités métriques des méthodes d'évaluation

Il existe de nombreuses méthodes pour évaluer les individus et prédire leur performance future. Les méthodes qui ne possèdent pas les qualités métriques essentielles doivent être radicalement éliminées. Les caractéristiques métriques des méthodes d'évaluations sont sa fidélité, sa sensibilité et son pouvoir discriminant, sa signification et sa validité par rapport aux caractéristiques requises.

Fidélité

C'est la première qualité d'une méthode d'évaluation, définie comme la capacité à fournir des résultats constants non seulement dans le temps à l'occasion de plusieurs mesures successives mais aussi lorsque l'on change d'utilisateur de la méthode.

Pour mesurer la fidélité, la méthode test-retest consiste à utiliser deux fois de suite la même technique dans les mêmes conditions expérimentales. La valeur de la corrélation entre les deux séries de mesures représente la fidélité de la méthode. L'inconvénient principal de la méthode est le manque de spontanéité dans les réponses du deuxième test, car les sujets ont déjà vu le test une première fois.

Une seconde méthode consiste à fabriquer deux versions équivalentes (mais différentes) du même questionnaire, et conduire les deux tests avec ces deux versions successivement.

Une condition de fidélité est la cohérence interne de la méthode d'évaluation. On l'obtient en divisant la méthode en deux parties, donnant deux séries de scores distincts. La cohérence interne de la méthode est mesurée par la corrélation entre les deux scores obtenus. On veillera nécessairement à la pertinence de la méthode de division de l'évaluation pour éviter des biais parasites liés par exemple à des différences de difficulté entre les deux parties. On notera également l'inconvénient du à la réduction de la taille des deux échantillons qui diminue statistiquement la valeur de la fidélité.

Une autre méthode consiste à calculer la corrélation entre chaque élément du test (item) et le score total. En supposant que les items ont la même variabilité (moyenne et variance), on peut calculer une fidélité théorique dans le cas où les réponses aux items sont dichotomiques c'est-à-dire classées comme « bonnes » ou « mauvaises » (Kuder et Richardson 1937).

Discrimination

Les méthodes d'évaluation doivent évidemment discriminer les candidats entre eux. Cette discrimination est donc relative à la population évaluée. Les questions doivent être ni trop faciles, ni trop difficiles pour les candidats pour pouvoir retirer des informations utiles.

A quoi sert un étalonnage ?

Il est nécessaire de disposer d'une norme permettant de comparer les résultats des sujets soumis aux tests à un groupe de référence qui a les mêmes caractéristiques que les candidats évalués. Ces normes de références donnent un sens à l'évaluation et prouvent que la méthode utilisée est bien discriminante pour le groupe concerné.

Ces normes peuvent par exemple se présenter sous la forme de centiles (indicateurs du rang d'un score donné au sein d'une population donnée). En l'absence de ces normes de référence, il est nécessaire de la créer en faisant passer l'épreuve par un groupe adéquat et calculer les centiles correspondant aux différents score.

Les autres manières de présenter ces normes sont les **scores standards** qui mesurent l'écart à la moyenne en termes d'écart types de la distribution des scores bruts, les **stanines** qui fournissent la même information en fixant la moyenne à 5 et l'écart-type à 2, et les **notes T** qui fixent la moyennes à 50 et l'écart type à 10. Comme ces normes fixent les deux principaux paramètres des distributions, elle permettent de comparer directement entre elles des distributions de formes initiales différentes et non comparable, et donc in fine d'établir des profils cohérents et lisibles faisant ressortir les points forts et les points faibles des différents candidats.

Stratégie et techniques de validation

Une technique d'évaluation doit également être pertinente, c'est-à-dire utile pour la décision à prendre. On définit l'utilité de deux manières : l'adéquation des informations obtenues avec le poste ou la fonction d'une part, et la qualité de la prédiction fournie par la méthode d'évaluation d'autre part. Cette qualité est obtenue en comparant le pronostic de la méthode avec la réussite professionnelle ultérieure.

Il existe trois types de validité, définis par l'usage que l'on souhaite faire de la méthode.

- La validité prédicteur/critère mesure la relation entre le prédicteur et le critère de succès professionnel
- La validité de contenu qui correspond au degré de finesse de l'évaluation (échantillonnage des comportements à pronostiquer)
- La validité de construction qui mesure la capacité de la méthode à mesurer la variable psychologique à apprécier conformément à un modèle théorique existant.

Validité de contenu

Elle dépend fortement de l'analyse du travail et implique que la performance soit clairement décrite, les objectifs de la méthode clairement formulés et l'échantillonnage du domaine professionnel fait de manière adéquate.

Pour s'assurer de la validité de contenu d'une méthode, il faut :

- Conduire une analyse du travail déterminant les comportement et aptitudes essentiels au poste
- Créer des programmes d'évaluation quantifiables et cohérents avec l'analyse du travail
- Evaluer la pertinence de la technique d'évaluation par rapport à l'activité professionnelle en tenant éventuellement compte de l'avis d'experts, notamment du point de vue de l'échantillonnage du domaine professionnel.

Validité de construction

Elle entre en jeu lorsque les qualités à mesurer ne sont pas d'ordre opérationnel mais plutôt relatives à des attributs comme l'intelligence, la fluidité verbale ou la mémoire ; ou encore lorsqu'on s'intéresse à la personnalité. Mesurer ces qualités abstraites conduit ainsi à s'interroger sur la définition que l'on donne à ces qualités.

Il existe quatre méthodes pour mesurer la validité de construction :

- vérifier la bonne corrélation entre la nouvelle technique d'évaluation et des techniques plus anciennes
- l'analyse factorielle qui vise à dégager des facteurs qui rendent compte des relations existant entre plusieurs mesures à partir de corrélations simultanées entre plusieurs tests et épreuves
- l'analyse factorielle sur les sous scores d'une seule épreuve pour mettre en évidence l'existence d'éventuels facteurs généraux importants communs aux différentes sous épreuves
- le procédé dit « multitrait-multiméthode » (Campbell et Fiske 1959) qui vise à montrer que des tests différents mesurant les mêmes facteurs corrélerent entre eux mais pas avec des tests qui mesurent d'autres facteurs. Ce procédé permet de savoir si la variance des mesures est due à la méthode utilisée ou au contenu des tests.

L'analyse des items

L'analyse des items cherche à comprendre pourquoi certaines méthodes ont une validité de construction insuffisante et à proposer des correctifs adéquats. On analyse pour cela les items (questions, problèmes,

échelles d'attitudes, grilles d'évaluation,...) et on calcule la corrélation entre chaque item et le score total. Les items à faible corrélation doivent être soit éliminés soit retravaillés et améliorés. De même, on éliminera les items non discriminants qui parasitent la méthode.

L'aspect prédictif

Il consiste à vérifier de manière expérimentale que la prédiction faite s'est trouvée confirmée. Cette approche est adaptée quand on ne sait pas de manière précise et exacte ce que l'on souhaite mesurer. C'est la validité de la relation entre l'évaluation et le comportement ultérieur qui est centrale ici.

Le principe consiste à comparer à un critère de succès professionnel le pronostic fait avec la méthode que l'on veut valider. La première méthode consiste à appliquer la technique à valider à des personnes en poste. On obtient un indice de validité concurrente, qui présente l'avantage d'être rapide à mettre en œuvre mais qui présente un inconvénient lié à la non-représentativité d'un éventuel groupe de candidats par le groupe interrogé (qui a déjà été sélectionné, qui a acquis de l'expérience et qui peut être moins motivé à répondre qu'un vrai groupe de candidats).

La validité prédictive, qui prévoit un intervalle de temps entre le relevé des informations de réussite et le relevé des informations prédictives, fait disparaître ces inconvénients. Mais elle ne peut être réellement mise en œuvre qu'avec un échantillon de population homogène et assez grand, et demande un critère de succès professionnel fiable, ce qui fait que la faisabilité de la validité prédictive est donc faible. Ces deux raisons peuvent également conduire l'évaluation de la validité prédictive/critère d'une méthode à la rejeter alors qu'elle est valide.

La validité synthétique

Elle consiste à déduire la validité des instruments d'évaluation d'un poste à partir d'une décomposition logique du poste en éléments de base, de la recherche d'instruments valides pour ces éléments de base et de la recombinaison de ces validités en un tout. L'avantage de cette technique de validation est qu'elle peut concerner de petits échantillons. Elle est économique en temps ; surtout, elle permet d'adapter les programmes d'évaluation aux changements de contenu des postes de travail, facteur très important à l'heure actuelle.

La méta analyse

Ces méthodes permettent de résumer et d'analyser une série de recherche portant sur le même thème et sert ainsi de support aux professionnels dans l'évaluation. Cependant, les efforts pour réussir à calculer la validité moyenne de méthodes prédictives ont donné des résultats décevants (validités moyennes faibles et très dispersées). Pour résoudre ce problème, Schmidt et Hunter (1981) ont développé une méthode de méta-analyse destinée à prendre en compte les grandes sources d'erreurs dans le calcul de la validité prédictive : erreurs dues à la taille de l'échantillon, à la fidélité du critère, à la variance réduite de l'échantillon et à la fidélité des méthodes de prédiction. Le processus de généralisation de la validité consiste à estimer la variance des coefficients de corrélation attribuable à ces quatre sources d'erreurs et à estimer les « validités vraies », c'est-à-dire corrigées de ces quatre sources d'erreurs. Les « validités vraies » sont alors deux fois plus élevées que les validités observées.

Il reste cependant trois sources d'erreurs que les chercheurs considèrent comme impossible à évaluer : la contamination du critère, les erreurs matérielles de calculs et les différences de contenu entre les prédicteurs. La variance résiduelle corrigée n'est donc pas toujours nulle. Schmidt et Hunter considèrent alors que ces trois sources expliquent 25% de l'erreur globale et proposent d'utiliser cette valeur comme seuil de variance résiduelle pour considérer que la validité prédictive est généralisable.

Chapitre IV - La prise de décision et l'évaluation de son utilité

On rencontre deux types de situation dans la prise de décision suite à une évaluation. Soit un grand nombre de candidats sont évalués pour être affectés à des postes stables dans le temps, soit un petit nombre de candidat sont évalués pour un poste à responsabilité importante constituant une étape dans une carrière.

Quelles que soient les méthodes de décision utilisées, plusieurs informations sont recueillies et doivent donc être agrégées. Deux méthodes existent : la synthèse **clinique** et la méthode **actuarielle**.

La synthèse clinique

La synthèse clinique implique qu'un expert tire une conclusion personnelle à partir de sa connaissance de l'ensemble d'un dossier. La synthèse clinique peut également se faire sur la base de discussions entre plusieurs personnes et intégrer des informations chiffrées (notes de tests) avec des données plus qualitatives (impressions suite à entretiens par exemple).

Cette méthode est relativement subjective mais c'est celle qui est la plus utilisée, notamment dans les cabinets de recrutement ou l'agrégation des données ne repose pas sur des bases aussi claires et solides qu'on pourrait le souhaiter, ce qui conduit à une sous-utilisation des informations disponibles.

Son succès s'explique cependant par l'image qu'elle donne d'un processus plus humanisé. De plus, comme elle repose sur l'expérience et l'intuition, elle n'a pas besoin d'être justifiée de manière rigoureuse ce qui lui donne un attrait certain.

Les méthodes actuariales

Ces méthodes reposent sur une formulation mathématique destinée à tenir compte de manière aussi efficace que possible des éléments de la décision. Elles sont plus précises et fournissent des décisions plus cohérentes entre elles que la synthèse clinique.

Les méthodes actuariales les plus simples consistent à calculer la moyenne éventuellement pondérée des différentes notes obtenues, après leur avoir donné la même échelle. Ces pondérations peuvent être calculées de manière rigoureuse et optimale si on connaît la validité des différentes épreuves donnant les notes chiffrées, à l'aide d'un calcul de régression multiple qui donne la combinaison optimale des coefficients de pondération.

Elle présente cependant l'inconvénient de supposer qu'une déficience dans un domaine peut être compensée par des qualités dans un autre, en raison du système de calcul de moyenne pondérée. Pour pallier cet inconvénient, on peut utiliser le système de seuils sur une ou plusieurs des notes pour modéliser des exigences minimum du poste dans tel ou tel domaine (méthode des seuils multiples).

Comment choisir une méthode ?

Il y a plusieurs façons de choisir une méthode d'évaluation et aucune ne s'impose systématiquement dans tous les cas. Pour choisir la méthode, il faut tenir compte du type de décision à prendre, du nombre de postes à pourvoir et du niveau des emplois ciblés.

La problématique du choix vise en fait à réduire les « faux positifs » (qui vont conduire à des erreurs de recrutement ou d'affectation) et les « faux négatifs » (rejet de candidats qui auraient été des succès). D'où la nécessaire réflexion sur l'utilité de la méthode.

Les deux autres angles d'évaluation sont les aspects financiers (coût de la méthode /bénéfice financiers) et les aspects humains (utilité des informations pour l'évalué, prévention des échecs, détection des potentiels).

Utilité des méthodes d'évaluation

Il est nécessaire de comparer l'utilité des méthodes d'évaluation pour décider de leur utilisation. Il faut pour cela posséder des données précises sur chacune des méthodes à évaluer (validité, coût, ...) et il faut enfin élaborer des données permettant de mesurer l'importance de la qualité des décisions, traduite si possible en termes financiers pour l'entreprise. L'analyse de l'utilité doit se faire en fait chaque fois que l'organisation doit prendre des décisions de sélection (recrutement, promotion etc.).

La méthode la plus simple pour évaluer l'utilité d'une méthode d'évaluation consiste à faire passer l'examen « à blanc » à une population de nouvelles recrues sans utiliser la méthode pour sélectionner. On classe ensuite les individus en deux catégories : les échecs et les succès professionnels, ce qui donne le taux de recrutements réussi sans le critère. On recalcule le taux sur la population qui aurait été sélectionnée avec application du critère et on mesure l'utilité de la méthode par l'amélioration de ce taux de recrutements réussis.

L'autre problématique consiste à fixer les bons seuils pour déterminer les meilleures conditions dans lesquelles doivent se dérouler les tests pour permettre de réduire au maximum les erreurs de recrutement. Plus la barre sera haute, plus on réduira les faux positifs par exemple. De même, plus le seuil définissant la réussite professionnelle est exigeant et plus le taux de recrutement de faux positifs sera élevé. Les conditions dans lesquelles se déroulent les tests sont donc déterminantes de l'utilité d'une méthode d'évaluation.

Les facteurs de l'utilité

L'utilité d'une méthode d'évaluation est fonction de trois paramètres :

- La validité de la méthode (coefficient de corrélation).
- Le pourcentage de bonne décision prise sans utilisation de la méthode d'évaluation
- Le nombre de personnes à retenir par rapport au nombre de candidats. Une sélectivité très élevée (taux de sélectionnés très faible) peut compenser une faible validité de la méthode (corrélation inférieure à 0.20).

L'utilité de la méthode sera d'autant plus forte que le nombre de candidats sera élevé et d'autant moins forte que ces candidats seront tous de bonne qualité.

Utilité et validité

Il existe aujourd'hui des tables qui permettent de lire directement le pourcentage de bonnes décisions en fonction de la validité. Elles montrent la variation de l'utilité lorsque varient les seuils de rejet ou le taux de réussite professionnelle ce qui permet de s'affranchir de la méthode expérimentale impliquant nécessairement le recrutement, « pour voir », de « faux négatifs ».

On peut également améliorer le procédé en construisant des tables donnant le gain de performance en fonction de la validité du prédicteur, du taux de sélection et de la variance de la performance, ce qui permet d'éviter la simplification exagérée entre « bons et mauvais ».

Une autre façon de calculer l'utilité

En se fondant sur les moyennes de la valeur financière des salariés pour l'entreprise (obtenues à partir d'estimations fournies par leur encadrement), Schmidt *et al* ont estimé la valeur financière attachée à un meilleur recrutement. Bien que les valeurs fournies et la méthodologie utilisée aient été vivement contestées, ces travaux ont permis de relancer la réflexion sur les modèles économiques et l'utilité des évaluations. Ont été abordés les coûts secondaires engendrés par l'accroissement de productivité, l'impact fiscal sur un bénéfice plus élevé. L'aspect dynamique intégrant la durée moyenne de renouvellement du personnel a également fait l'objet de recherches complémentaires.

L'évaluation individuelle

La validité de l'instrument de mesure est une notion différente de celle de la validité de la prise de décision fondée sur un ensemble d'informations issues de méthodes d'évaluation mais aussi d'éléments plus subjectifs (impressions après entretiens, références, etc.). La prise de décision dans ce cas est souvent le fruit d'une inférence sur la façon dont va être perçu le candidat sélectionné par son entourage professionnel. Les facteurs situationnels jouent ainsi un rôle important dans l'évaluation individuelle.

En fait, les recherches sur la fidélité de la décision sont très rares et celles qui existent donnent des résultats décevants, parfois même contradictoires. Par exemple, lorsque trois décideurs donnent leur avis sur un candidat à choisir parmi trois candidats, il n'y a accord ni sur les personnalités des candidats, ni sur celui qui doit être choisi.

Les multiples causes de désaccord entre évaluateurs sont leur incompétence, le fait qu'ils défendent des idéologies qui leurs sont propres et enfin la manière avec laquelle ils organisent l'information.

Existe-t-il dans ces conditions une validité du processus d'évaluation individuelle ? L'évaluation s'appuie forcément sur le jugement des évaluateurs et chacun d'entre eux a son propre mode de fonctionnement et d'inférence. On se posera donc deux questions :

- Les informations recueillies correspondent-elles aux objectifs de l'évaluation ?
- l'intégration des informations est-elle adéquate ?

Il est difficile de répondre car le critère n'est jamais totalement fiable en raison de la fréquente contamination du critère par les conseils et les actions du recruteur, qui peut souhaiter promouvoir ceux qu'il a recrutés.

Aussi l'évaluation individuelle semble-t-elle difficile à étudier et à contrôler. Peu de recherches s'y intéressent et on doit se contenter d'énumérer les grandes étapes de ce processus pour le décrire :

- l'analyse du poste (qui reste impérative)
- le choix des outils nécessaire
- la mise en œuvre de ces outils et l'interprétation des résultats
- l'intégration des données pour arriver à la conclusion demandée
- la restitution des résultats à la personne évaluée.

Chapitre V - Le développement individuel

La notion de carrière a beaucoup évolué avec la disparition des carrières à vie et des plans de carrières. Les méthodes d'évaluation ont donc du évoluer : auparavant destinées exclusivement au moment de l'orientation des adolescents dans la vie active, elles sont aujourd'hui utilisées tout au long de la vie professionnelle et elles s'inscrivent dans un contexte dynamique.

Seul ceux qui sauront tirer partie de ce dynamisme auront une chance de « faire carrière ». De même, les organisations se doivent de gérer efficacement leurs ressources humaines.

L'acquisition de compétence représente un impératif pour l'évolution de la carrière individuelle et une obligation pour l'entreprise conduite à gérer son stock de compétence pour faire face à l'évolution rapide de son environnement.

L'évaluation acquiert donc un nouvel objectif : elle intervient dans la gestion de carrière de l'individu, et notamment dans son acquisition de nouvelles compétences.

Des méthodes spécifiques ?

Dans le cadre du développement professionnel, les exigences métriques continuent de s'appliquer mais l'évaluation et la manière de gérer ses résultats sont différentes de celles utilisées dans le cas de prise de décisions.

Ces méthodes doivent être réalisées en concertation entre l'organisation et l'individu et avec une volonté réelle de ce dernier. Le plan de développement dépend en outre des opinions que se font « les autres » de l'individu, ce qui rend nécessaire de disposer d'une vue d'ensemble sur l'environnement organisationnel dans lequel va se dérouler la carrière, notamment des données sur l'avenir de l'organisation, ses activités ou produits nouveaux, ses forces, sa culture et enfin sur le statut de l'individu dans son environnement de travail.

Deux aspects caractérisent les méthodes spécifiques à l'élaboration de plans de développement :

- elles prennent en compte la structure et les stratégies de l'organisation
- elles utilisent des informations venant de plusieurs niveaux hiérarchiques (ex. 360°).

Il existe également des centres de développement combinant tests en situations et évaluation à 360° pour mesurer les aptitudes et décrire les personnalités. Leur intérêt est de donner aux individus une visibilité sur leur atouts et handicaps, une vision réaliste sur leurs plans de développement dans l'organisation et une compréhension de la nature et de l'importance des compétences. Ces centres de développement constituent une source de motivation supplémentaire pour les individus.

Pour mesurer quoi ?

Il s'agit de dépeindre avec exactitude la réalité des perceptions réciproques et les rapports sociaux qu'elles entraînent dans un contexte organisationnel donné. On cherchera à décrire fidèlement la perception que les autres ont de l'individu même si cette perception n'est pas tout à fait exacte.

Il existera nécessairement des concordances entre les différentes opinions et des points de désaccord. A partir de là, seule la personne évaluée sera en mesure de faire un travail de réflexion sur ces données pour en tirer parti.

Dans le cas présent, les aspects importants pour une méthode d'évaluation sont la cohérence (validité de contenu), la signification des résultats obtenus, et le travail d'appropriation et d'analyse fait sur ces résultats par la personne évaluée.

Quels sont les effets de la restitution des résultats ?

Le besoin d'information sur soi est fort et les participants à des centres d'évaluations sont en général favorables à cet apport d'informations nouvelles, dans la mesure où l'objectif des exercices est transparent, et où les participants ont l'impression d'être jugés sur leurs qualités et aptitudes réelles.

La restitution des informations aux participants joue un rôle primordial notamment pour faire disparaître deux types de blocage à l'intégration de nouvelles informations par les participants. Premièrement le participant peut ne pas voir de problème particulier et ne juge pas utile de s'y intéresser. Deuxièmement, le participant se sent menacé par ce qu'il va apprendre sur lui, et perçoit des risques pour son propre développement. Il peut aussi exister un blocage de type cognitif fondé sur des expériences de formation antérieures restées sans effet, et des obstacles contingents liés à l'incapacité de l'organisation à offrir des possibilités de développement.

Les informations issues des centres de développement seront d'autant plus efficaces qu'elles seront argumentées et fondées sur des résultats interprétables et sur des comportements précis survenus pendant les tests de situation. Elles se doivent d'être constructives c'est-à-dire articulées par rapport aux objectifs individuels et présenter des possibilités de développer des compétences faibles ou de modifier des comportements inadéquats. Enfin, l'estime de soi et le soutien du supérieur hiérarchique faciliteront de manière importante l'intégration des données reçues par les participants.

DEUXIEME PARTIE : QUELLES MÉTHODES D'ÉVALUATION ?

Le rôle de la motivation

Toute évaluation a trois catégories d'objectifs : bilans des aptitudes, des compétences et de la personnalité. Mais ces compétences ne servent à rien si la motivation est absente. Or cette motivation ne se mesure pas comme une caractéristique individuelle, elle se construit. Elle se développe à travers l'expectation, c'est-à-dire conviction de l'individu qu'il est capable d'atteindre l'objectif qui lui a été assigné.

C'est l'organisation qui doit créer des conditions favorables à la motivation, condition sine que non de l'effort au travail et du développement réussi de l'individu. Mais ce n'est pas toujours le cas en pratique et les validités représentées par la corrélation entre un prédicteur et le critère se trouve modulées par la capacité de l'organisation à utiliser le potentiel des individus. Le meilleur des prédicteurs sera inefficace si le personnel se trouve placé dans des conditions de travail défavorables.

Chapitre VI - Le trio classique : Références, CV, entretiens

CV, références et entretiens sont les méthodes les plus utilisées dans les recrutements. Cette pratique est bien acceptée par les candidats. Outre la transparence apparente, les candidats apprécient de pouvoir établir un dialogue qui leur fournira des informations sur le poste à pourvoir. Les responsables en apprécient l'aspect concret.

Un premier filtre est constitué par l'examen des références écrites avant de passer aux entretiens. Il convient donc de vérifier si ces sources d'informations sont pertinentes, de connaître leur validité et d'analyser leur signification. L'ensemble des exigences décrites dans la première partie s'applique aux références, CV et entretiens.

Les références

Il s'agit de recueillir des informations auprès de personnes qui connaissent bien le candidat avec comme objectif de vérifier les renseignements donnés par le candidat et obtenir une opinion avertie sur ses chances de réussite dans le poste. On recherche principalement des informations sur la personnalité, les qualités sociales et la capacité à travailler avec les autres. Contrairement au cas de l'entretien ou du test qui mettent

le candidat dans une situation exceptionnelle, les références donnent des informations sur le comportement quotidien du candidat.

Les études montrent que la validité des prédictions formulées dans les références est uniformément faible. Il semble que la validité la meilleure soit obtenue lorsque les informations viennent du supérieur du candidat. Baxter *et al* ont également montré que les références sont plus corrélées avec la personne qui les donne qu'avec celle qu'elles concernent.

La faible validité s'explique par la faible fidélité des références. Pour améliorer l'utilité, il convient de travailler les questions pour les rendre plus précises, voire d'utiliser des questionnaires à choix multiples. En outre, on veillera à ne demander que ce que les personnes interrogées savent du candidat et préciser l'usage que l'on fera des informations fournies. Carroll et Nash (1972) ont construit une liste de descriptions de comportements destinés à réduire l'indulgence des références. La corrélation entre références et notations ultérieures est montée à 0.64 contre une moyenne de 0.20 pour le cas général.

En conclusion, les références devraient être utilisées comme premier filtre des candidatures inadéquates et non pour prédire la performance ultérieure.

Le CV et les biodatas

Le CV est incontournable dans l'évaluation des individus. L'inconvénient majeur de l'analyse des CV résulte des interprétations différentes propres aux notateurs qui examinent sans consigne précise les documents. Sans analyser la validité prédictive des informations biographiques, il est impossible de savoir quel mode d'interprétation est fiable et efficace. L'inventaire biographique pondéré améliore la situation mais reste insuffisant et relativement grossier car il ne repose pas sur l'analyse du travail. Les fiches biographiques avec des questions fermées à choix multiples améliorent l'efficacité mais nécessitent une mise au point attentive, en particulier en ce qui concerne le choix du critère à prédire et le choix des réponses qui seront offertes au sujet.

Les biodatas sont les informations biographiques. Pour Owen, ce ne sont pas des données isolées mais elles représentent une banque de données organisées qui permet de définir l'individu à travers les événements de sa vie passée et d'utiliser cette définition pour prédire les conduites futures. Owen souligne la « cohérence » de ces données en précisant qu'elles révèlent en fait une continuité des mêmes types de comportements et d'attitudes à travers le temps. Owen souligne également l'importance de comprendre pourquoi il existe une relation entre les informations biographiques et le comportement professionnel ultérieur.

Les biodatas, s'ils sont bien choisis, seront plus prédictifs et mieux acceptés que les tests et autres épreuves ponctuelles. Comme Hough l'a montré en 1984, l'utilisation d'informations biographiques apporte des éléments sérieux pour la prédiction de la réussite professionnelle et la prise de décision en matière de gestion du personnel.

Il existe des catégories différentes d'items de biodata selon qu'ils sont vérifiables (faits précis) ou non (opinions, aptitudes). Ces derniers sont les plus utilisés. Selon plusieurs recherches sur l'efficacité de ces méthodes, leurs avantages sont d'être bien acceptés, faciles à remplir et peu coûteuses. Mais il est permis de douter de la sincérité des sujets qui déforment leurs réponses dans le sens qui leur paraît favorable (surestimation de leur salaire, augmentation de leur expérience dans le poste, etc.). Il faut également veiller à ce que la collecte des informations ne conduise pas à une sélection « abusivement » discriminante (fondées par exemple sur l'âge ou le sexe des candidats), en ne sélectionnant - comme la loi le demande déjà aux Etats-Unis- que sur des attributs dit « modifiables ».

Enfin, il semble possible de construire des items biographiques suffisamment robustes pour être généralisables à différents métiers et différentes situations (Hunter et Hunter 1984).

Avant d'utiliser les biodatas, il est nécessaire de garder à l'esprit que la validation des items prend du temps, et que le poids des données doit être apprécié objectivement avec un dépouillement faisant l'objet de règles très précises. De ce fait, cette méthode est plus adaptée pour de grands nombres de candidats et des postes relativement stables dans le temps. Enfin, elles ne doivent jamais être utilisées seules, notamment car leur validité générale ne le permet pas.

L'entretien

C'est la plus utilisée et la moins connue de toutes les méthodes d'évaluation : nous savons mal ce qui se passe en entretien et la grande variance des situations (durées d'entretiens, entretiens de groupes individuels, nombre d'interviewer) complique encore le problème.

L'objectif de l'entretien est double : obtenir des informations sur le candidat et se faire une impression sur lui. Mais sa faible fidélité, son manque de validité, sa forte sensibilité au biais et aux distorsions en font un instrument de pronostic particulièrement discutable, comme le montrent de nombreuses recherches.

On l'utilise cependant pour trois raisons principales :

- c'est un échantillon de comportement qui donne un éclairage sur les caractéristiques sociales du candidat,
- il remplit des fonctions autres que l'évaluation (contact avec le candidat, promotion du poste à pourvoir),
- les personnes qui utilisent l'entretien pensent (à tort) qu'il est valide dans tous les cas.

Fidélité et validité de l'entretien

Plusieurs recherches tentent d'expliquer la faible fidélité de l'entretien. En particulier, lorsque l'entretien est non structuré, en l'absence de guide d'entretien par exemple, la faiblesse de la fidélité est forte et la validité est dispersée. Reynolds observe cependant une bonne corrélation lorsque l'on compare les descriptions d'un sujet données par différents interviewers. Mais ce résultat ne tient plus lorsqu'on demande aux interviewers une opinion plus qualitative que quantitative ou descriptive. De même, le désaccord existe lorsqu'il s'agit de pronostiquer le succès professionnel. Les sources de variances expliquant ce manque de fidélité, selon Guion, sont la situation dans laquelle on met le candidat, l'expérience de l'interviewer et la méthode utilisée pour mener l'entretien. Selon Guion, le bon interviewer doit bien connaître le poste à pourvoir, avoir de l'expérience, être intelligent, avoir de réelles capacités d'analyse et avoir en tête une représentation claire du candidat idéal.

L'entretien est une méthode dont l'évaluation est difficile car il est difficile d'isoler ce qui a trait à l'entretien lui-même des autres éléments utilisés pour la prise de décision. De plus les critères de succès professionnel sont souvent de faible fidélité. Une autre explication du manque de fidélité de l'entretien provient du fait que les interviewers évaluent les candidats selon un petit nombre de dimensions qui leur sont propres. C'est pourquoi il est nécessaire de compléter l'étude de la validité des entretiens par une analyse du processus cognitif qui conduit à l'évaluation. L'entretien permet en effet d'obtenir certains renseignements pertinents et utiles et à ce titre il ne doit pas être écarté.

On peut imaginer que la validité des décisions fondées sur les entretiens serait meilleure si on homogénéisait les entretiens et si les interviewers recevaient une même formation. Cela répondrait au problème des biais introduits par les interviewers (comme les préjugés liés à de similitudes entre interviewé et interviewer). Des recherches expérimentales en psychologie sociale ont permis de développer des modèles et des règles concernant la perception d'autrui, qui peuvent s'appliquer à la situation de l'entretien.

Améliorer l'entretien ?

L'entretien est et restera utilisé pour plusieurs raisons et notamment du fait qu'il apparaît valide aux yeux du candidat lui-même. Il est donc important de formuler des conseils pour accroître la validité de l'entretien. Ces conseils sont les suivants : il faut sélectionner et former les interviewers, notamment sur la nature des biais et des déformations auxquels ils risquent d'être confrontés. Il faut également leur fournir systématiquement une description précise des exigences du poste à pourvoir, et des possibilités ultérieures de carrière, ainsi qu'une liste de points à évaluer.

Entretiens structurés et situationnels

Les entretiens structurés (qui sont les plus fidèles et les plus valides) doivent :

- être précédés d'une analyse de poste précise
- porter sur de nombreuses questions et situations concernant les compétences importantes pour le poste
- noter avec l'aide de fiche de notation ou de barème documenté par des références à des comportements précis
- limiter l'accès des interviewers aux autres informations avant l'interview

- relier les entretiens aux notes obtenues aux tests

L'interview est alors plus rigoureux, moins contaminé, et le contenu de l'entretien est plus acceptable pour les candidats. Dans ces conditions, la fidélité inter-notateur est considérablement accrue par rapport à l'entretien traditionnel. Elle peut atteindre une corrélation de 0.40 et 0.56 après correction des biais de dispersion et de non-fidélité du critère (Campion 1988).

Les raisons pour lesquels les entretiens structurés sont moins utilisés sont la préférence générale pour un entretien chaleureux et convivial et de ce fait plus rassurant pour le candidat et donnant plus de pouvoir à l'interviewer.

Chapitre VII - Aptitude cognitive et aptitudes spécifiques

Un pionnier français : Binet

Binet a été le premier à tenter de prédire les résultats scolaires par la mesure de l'aptitude mentale. L'innovation principale de Binet était la définition du concept d'aptitude et la démonstration de l'utilité de sa mesure. Il a en outre compris que pour pouvoir s'exercer, les aptitudes doivent être développées par la formation et mobilisées par l'expérience.

Binet a émis l'hypothèse de l'existence d'une aptitude mentale fondamentale, caractéristique de la réussite à l'ensemble des épreuves qu'il avait mis au point pour prédire la réussite scolaire. En multipliant les tâches tests et en généralisant à l'adulte, on observe une réalité plus complexe : certains sujets sont en effet capable d'exécuter certaines tâches spécifiques et échouent sur d'autres tâches. Cette observation a révélé l'existence de deux types d'aptitudes : les aptitudes cognitives générales et les aptitudes spécifiques.

Il faut alors déterminer quel type d'aptitudes (générale ou spécifique) est de caractère dominant, et quelles sont les relations entre les aptitudes. Il est aussi nécessaire de s'interroger sur la validité des tests d'aptitude. Enfin, il faut tenter de comprendre pourquoi la validité des tests d'aptitude se confirme.

La structure des aptitudes

La réussite dépend d'une aptitude générale (le fonctionnement cognitif) et d'une ou plusieurs aptitudes spécifiques. Leur identification et leur classement ont été fait par analyse factorielle. Le résultat montre que les aptitudes sont de portées variées selon qu'elles interviennent dans une ou un grand nombre de tâches. Les aptitudes de large portée sont peu nombreuses : on peut citer l'aptitude verbale, l'aptitude spatiale et l'aptitude numérique.

Les recherches s'accordent sur l'existence d'une structure hiérarchique des aptitudes. L'aptitude générale joue un rôle dans toutes les tâches et représente 50% de la variance des résultats de tests mentaux. Les autres aptitudes sont indépendantes entre-elles et sont plus ou moins pointues. Il existe donc quatre types de tests d'aptitude qui mesurent

- le fonctionnement cognitif uniquement
- une des aptitudes fondamentales
- un ensemble d'aptitudes relativement fondamentales
- une aptitude spécifique isolée.

Peut-on mesurer l'aptitude cognitive générale ?

L'imbrication des aptitudes rend difficile l'élaboration de tests d'aptitude cognitive pure. Selon Spearman, les tests qui se rapprochent le plus de cet objectif sont les tests d'exécution de tâches abstraites faisant appel à des déductions logiques.

On peut également utiliser des tests qui mélangent des problèmes de différentes natures pour compenser le manque d'une aptitude spécifique par l'existence d'une autre. Il existe d'autres tests appelés batteries factorielles, batteries multi aptitudes qui comportent un ensemble de tests mesurant chacun une aptitude indépendante des autres. Citons par exemple les batterie PMA, DAT et GATB. L'analyse du travail permettra

de relier la série des aptitudes fondamentales et les activités du travail quotidien. Il peut enfin être utile de construire un test spécifique pour mesurer une aptitude spécifique mais décisive à l'exécution de la tâche.

Validité des tests d'aptitude

Les résultats des recherches sur les tests d'aptitudes montrent que la moyenne des corrélations obtenues dans un ensemble d'études sur plus de 30000 personnes vaut 0.50. A l'intérieur d'une même famille de métiers, la validité des tests est constante. La validité de ces tests est ainsi démontrée et on admet que ces tests sont plus pertinents lors de l'entrée dans le monde du travail (sans autre information professionnelle).

Il existe cependant de réelles variations de la validité inter métiers. On observe que plus la complexité des tâches d'un métier augmente, plus les tests d'aptitude ont une grande validité pour ce métier. Symétriquement, plus les métiers ont une complexité faible, et plus la validité des tests psychomoteurs augmente.

Une autre source de variabilité au sein d'une même famille de métiers, et due au type de critère utilisé pour décrire le succès professionnel. Une méta analyse conduite dans le domaine de la production industrielle sur cinq critères professionnels et sept types de tests d'aptitude a montré que quatre de ces cinq critères sont bien prédits par les tests, avec quelques variations en fonction des tests choisis.

Que mesurent les tests d'aptitude cognitive ?

Une question centrale concerne le rôle des aptitudes générales dans la vie professionnelle. Interviennent-elles seulement dans l'apprentissage ou également dans l'exercice des activités professionnelles ? Dans le premier cas, on s'attend à ce que les validités des tests diminuent avec l'expérience professionnelle. La méta analyse a montré que l'aptitude mentale générale corrèle effectivement avec les connaissances professionnelles acquises, lesquelles corrèlent avec les notations professionnelles. Les recherches de Schmidt ont montré que la validité des tests d'aptitude cognitive reste en fait constante dans le temps. On peut en conclure que les tests de fonctionnement cognitif prédisent le succès professionnel parce qu'ils mesurent 1) une aptitude nécessaire à l'acquisition des connaissances requise pour faire correctement son travail et 2) une aptitude directement utile dans l'exécution du travail quotidien.

Quels progrès à venir ?

Le développement de l'informatique comme support de tests permet de remplacer les questionnaires classiques. Des logiciels de tests sont aujourd'hui commercialisés et leur validité corrèle avec les tests papier avec une validité intacte. Ils sont de plus fortement appréciés par les candidats. Les progrès plus importants concernent les tests sur mesure ou « adaptive testing » qui présentent un plus fort pouvoir discriminant, en proposant aux candidats des questions dont la difficulté est fonction du niveau du candidat, mesuré par les réponses qu'il a déjà fournies au cours du test. Enfin, l'informatique permet d'étendre l'évaluation à la mesure de la façon dont un candidat a trouvé une réponse, sans se limiter à la valeur de cette réponse, ce qui donne accès à de nouvelles variables.

Chapitre VIII - Les tests de personnalité et leurs indications

Il existe une forte demande concernant l'évaluation et la description de la personnalité, mais cette demande reste floue dans ses formulations. Les cadres et responsables du personnel peuvent cependant préciser les traits de personnalité requis par tel ou tel emploi. Dernièrement, des progrès théoriques et pratiques importants dans le domaine de la description de personnalité ont amené les praticiens à réviser le choix et l'utilisation des méthodes d'une manière qui a augmenté la validité des questionnaires.

Problèmes éthiques

Les tests de personnalités ont été considérés comme une intrusion dans la sphère intime. Peu d'entre nous souhaitent exprimer leur personnalité profonde. La question est de savoir si les tests de personnalité ne forcent pas l'individu à révéler des traits de caractères qu'il souhaite garder pour lui. Aussi devient-il inacceptable de faire passer des tests de personnalité sans en avoir vérifié au préalable l'utilité des questions et leur validité prédictive, et d'être capable de prouver l'utilité des méthodes employées.

L'autre question concerne le droit de tester la personnalité sans que l'individu évalué ne comprenne bien la portée de ses réponses. Tout examen doit donc être expliqué au sujet de manière accessible en justifiant sa validité et doit être passé dans un climat de confiance réciproque. Lors du choix d'une technique de description de personnalité, il est capital de bien préciser ce qu'on souhaite mesurer et en quoi la technique est efficace.

Problèmes théoriques et pratiques

La personnalité globale : approche « psychodynamique »

Tout comportement résulterait selon cette approche du désir de satisfaire les besoins profonds, les tendances, impulsions et instincts qui agissent au niveau de l'inconscient. C'est donc à ce niveau qu'il faut chercher les déterminants du comportement. Cette théorie génère des méthodes difficiles à valider mais est aujourd'hui communément admise.

Certaines techniques dites « projectives » ont été construites avec l'intention de décrire la personnalité globale et son fonctionnement dynamique inconscient à partir des réponses des candidats à des stimuli ambigus qu'ils doivent interpréter.

Les résultats d'efficacité prédictive de ces méthodes ont cependant été souvent décevants. Les raisons sont principalement dues aux conditions d'utilisation de ces méthodes et la fidélité de la méthode se confond avec la fidélité de l'expérimentateur lui-même.

Une des seules variantes de ces méthodes à disposer d'un cadre théorique clair est la méthode MSCS (Miner Sentence Completion Scale - Echelle de Phrases à Compléter de Miner) ; c'est la seule méthode qui a obtenu des résultats probants en ce qui concerne la prédiction du succès des cadres dans l'entreprise.

On retiendra que les tests projectifs ont une faible validité et une faible fidélité. Cela ne remet pas en cause cependant les avantages de la notion de projection si on parvient à l'utiliser dans un cadre théorique approprié.

Traits ou situations : les questionnaires de personnalité

Il s'agit de questionnaires qui ont fait l'objet d'analyses factorielles et permettent d'obtenir un profil de personnalité. Ils reposent sur la stabilité supposée des traits de personnalité et la possibilité de les inférer à partir de questions sur les comportements passés ou habituels des sujets. C'est l'analyse factorielle qui permet de préciser la manière dont les réponses se regroupent et d'en inférer l'existence de traits que des grilles de dépouillement permettent de noter.

La validité apparaît néanmoins comme très faible au vu d'une méta-analyse faite par Hunter. La première raison est la transparence des questions et l'existence de candidats avertis. Les parades sont une rédaction moins naïve des questions posées et l'introduction de questions destinées à constituer une échelle de mensonge ou particulièrement sensibles au désir de faire bonne impression. L'autre raison est la part trop importante laissée à l'interprétation dans la formulation des questions.

Enfin, l'environnement et la manière dont il est manipulé peuvent exercer une forte influence sur le comportement ce qui contribue à la faible fidélité des tests de personnalité.

Validité des questionnaires de personnalité

Utilisés dans de bonnes conditions (analyse de poste), les questionnaires de personnalité donnent une validité appréciable lorsque les scores obtenus sont vus de manière globale. On peut synthétiser les résultats des tests autour de cinq traits de personnalité : extraversion, stabilité émotionnelle, convivialité, conscience professionnelle et ouverture intellectuelle. Différents aspects de la personnalité peuvent déterminer différents aspects de l'efficacité professionnelle. De plus il faut tenir compte dans les tests du contexte situationnel qui peut intervenir au niveau des performances et des comportements.

L'auto-évaluation

Dans l'état des connaissances sur l'autoévaluation, on sait que l'image de soi joue un grand rôle dans la motivation et qu'elle se construit à partir du jugement des autres sur nos actes et notre performance. En raison de la différence d'interprétation de chacun des libellés des tests, et de la déformation due au désir de « désirabilité sociale » observé chez les candidats, l'auto-évaluation est un mauvais instrument

de la description de l'image de soi. Elle doit plutôt être réservée comme outil de gestion des ressources humaines et des motivations individuelle que comme méthode de sélection.

Chapitre IX - Et les compétences ? Des essais professionnels aux tests de situation et aux « centres d'évaluation »

Un essai professionnel est un échantillon de travail, une tâche choisie représentative de ce que le candidat devra faire dans son activité professionnelle. Les essais professionnels sont parfois appelés « tests situationnels » lorsque le sujet est observé par des personnes formées et munies de grilles de notation. Les essais professionnels construits pour être strictement similaires aux tâches professionnelles ont une validité fondée sur l'hypothèse que le comportement en tests sera répété en situation réelle et qu'un bon échantillon de comportement permettra de prédire l'ensemble des comportements de même catégorie.

Les avantages des essais professionnels sont leur forte validité apparente. Ils sont bien acceptés par les candidats car ils se sentent jugés équitablement et se rendent compte concrètement du travail qui les attend. Ils sont donc motivés à réussir ces tests, bien plus que dans le cas des tests classiques. On peut observer une baisse du taux de turnover dans ces cas.

De plus les études de validité ont montré qu'ils sont de bons prédicteurs du succès professionnel. Les inconvénients sont le coût et la lourdeur de ces essais, puisqu'ils sont spécifiques du poste à pourvoir.

Il existe quatre catégories d'essais professionnels que nous allons détailler dans les prochains paragraphes.

Les tests psychomoteurs

Il s'agit de tests d'aptitudes motrices telles que les manipulations d'objets (tests de dextérité). Bien que long et fastidieux à mettre au point, ils ont une validité très significative. Les méta-analyses récentes ont montré que la validité des tests psychomoteurs est généralisable à travers diverses situations.

Les épreuves de connaissance et essais de formation

En raison des diplômes qui attestent des qualifications des candidats, l'utilisation des épreuves de connaissance est rare, bien qu'elle donne des bons résultats lorsque les recruteurs les utilisent.

Les tests de formation consistent à évaluer la capacité des candidats à maîtriser une nouvelle activité et sont utilisés pour les postes à faible qualification. Ces tests présentent une bonne validité, supérieure à la validité des tests d'aptitude classiques.

Les tests de situation

Ces tests consistent à observer les candidats lorsqu'ils sont confrontés à une situation très proche de celle qu'ils rencontreront dans la vie professionnelle réelle. On a montré que ces tests avaient une très forte corrélation.

Les tests de groupe

Dans ces tests plusieurs candidats au même poste doivent effectuer un exercice collectif et aboutir à une décision ou réalisation commune. Cette tâche est en réalité un prétexte pour observer le comportement social des candidats. La validité est élevée pour les postes de cadres et la fidélité inter-notateurs est grande et augmente si les notateurs ont reçu une formation préalable.

Les centres d'évaluation

Proches des essais professionnels, ils sont plus riches et plus complet, plus coûteux également, mais plus valides et efficaces.

Les centres d'évaluation réalisent l'évaluation d'un groupe de cinq à huit sujets ne se connaissant pas, de même âge et de niveau de qualification équivalent. Ces évaluations impliquent la présence d'un psychologue et de un à trois évaluateurs supplémentaires, psychologues ou non. Les non-psychologues ne doivent pas être les supérieurs hiérarchiques des candidats évalués et ont suivi une formation adéquate. Ils peuvent cependant appartenir au service qui recrute.

Les centres d'évaluation comportent toujours un ou plusieurs tests de situation, des tests classiques, plusieurs entretiens, chacun avec un objectif précis. Ils sont toujours suivis d'une réunion de synthèse où les décisions sont prises en commun par un groupe d'évaluateurs, et d'un entretien de feedback fait par le psychologue avec chaque candidat.

Les centres d'évaluation peuvent avoir un objectif spécifique (recrutement pour un poste donné, sélection à l'entrée d'une formation...) ou un objectif plus large (recherche de potentiels d'encadrement) mais ces objectifs doivent toujours être clairement définis.

Quelles compétences ?

L'expression des compétences à évaluer est utile pour l'entreprise car elle la force à définir une stratégie claire. Relative à l'entreprise, elle s'exprime dans un vocabulaire qui peut donc lui être spécifique. On admet que pour un centre d'évaluation, cinq à sept compétences à évaluer suffisent ; il ne doit jamais y en avoir plus de dix.

La liste des compétences sert à créer la grille du programme, élaborée pour permettre une évaluation selon plusieurs angles pour tenir compte de l'impact situationnel. C'est le rapprochement des essais et scores qui débouche sur l'évaluation du candidat.

On veillera à ce que les listes des points à évaluer ne soit pas trop détaillée mais plus globale, avec des catégories de comportements auxquels le candidat doit avoir recours. Deux notateurs doivent évaluer le candidat. Ensuite, une discussion destinée à éliminer les aspects subjectifs a lieu entre les notateurs.

La validité des centres d'évaluation

Cette méthode est très utilisée, et notamment aux Etats-Unis. Bray et Grant ont fourni la meilleure analyse pour évaluer ces centres d'évaluation. Sur une période de huit ans, ils ont montré que 75 à 82% des candidats promus avaient été correctement évalués par le centre d'évaluation. La prédiction des non-promotion est encore meilleure puisqu'elle s'élève à 95%. Les validités de cette méthode sont comparables quel que soit le critère utilisé et s'élève à 0.44.

Enfin la forte validité apparente des centres d'évaluation aux yeux des candidats est un autre avantage de cette méthode.

... et les critiques

Les critiques adressées aux centres d'évaluation sont le coût pour l'entreprise et le temps demandé notamment aux candidats et aux notateurs. Il faut cependant savoir les mettre en regards des gains financiers apportés.

Il existe de plus un effet d'influence des résultats des tests des centres d'évaluation sur le comportement, dont les résultats peuvent augmenter la motivation ou la combativité des candidats bien notés.

Le critère utilisé (promotion des candidats) a été contesté par de nombreux chercheurs. Plus grave encore, Schmidt et Hunter dans une méta-analyse ont montré que la validité incrémentielle de l'arsenal des centres d'évaluation par rapport aux seuls tests de fonctionnement cognitif est très faible.

L'homogénéité des résultats pour mesurer un même trait psychologique dans différents exercices a également fortement été remise en cause. Il semblerait en fait que l'on note plutôt une impression générale qu'un trait spécifique ce que Russel a confirmé en 1985 dans une analyse factorielle.

Enfin, pour améliorer la validité des centres d'évaluation, il est nécessaire de ne pas se laisser séduire par leur côté ludique ni de les considérer comme une variante de la notation professionnelle. Pour ce faire, une standardisation des procédures améliorerait certainement les résultats (définition précise des dimensions notées, guide de notations, etc.).

En conclusion, les tests de situation bien conduits apportent une contribution significative à l'évaluation de la personnalité et au pronostic du succès professionnel. La mise en œuvre conjointe des tests classiques, de tests de situation et d'entretiens représente sans aucun doute la méthode la plus valide et la plus valable pour l'évaluation et la sélection des cadres et de leur potentiel d'encadrement.

Chapitre X - Quel avenir pour l'évaluation ?

Dans un monde du travail en constante évolution, les activités d'évaluation vont-elles devenir plus importantes ou moins importantes ? Vont-elles plus vraisemblablement se diversifier et évoluer ?

La tension sur l'emploi explique l'intérêt accru pour l'évaluation car dans ces périodes, la mobilité des salariés est plus réduite. L'erreur de recrutement coûte cher car plus difficile à rattraper.

La globalisation de l'économie conduit également les sociétés à se restructurer et à devenir plus flexibles pour accroître leur compétitivité. Demande de compétences accrues et plus pointues, nécessité de travailler en équipe sont quelques exemples de facteurs qui appellent une gestion des ressources humaines efficace.

De nouvelles demandes venant des organisations

Le fait que le travail se fait de plus en plus souvent en groupes et par projets pose le problème de la constitution d'équipes aux compétences, voire aux personnalités complémentaires. Quelles sont alors les qualités individuelles qui fondent un succès d'équipe ? Cette question appelle une évaluation fondée sur la psychologie de groupe. Les qualités mises en lumière sont alors l'empathie et la compréhension des autres. Le travail de groupe pose également le problème difficile de l'évaluation de la contribution de chacun au succès du groupe.

L'accent est mis aujourd'hui sur la polyvalence plutôt que sur la spécialisation. Ce sont maintenant les qualités d'adaptation, la capacité à apprendre, et l'implication dans l'entreprise qui sont recherchées.

L'évaluation du personnel en poste est également de plus en plus importante et soulève des problèmes nouveaux. La notation peut se faire par plusieurs personnes - hiérarchiques dans les organisations matricielles, collègues et subordonnées dans le cas du 360° mais ces méthodes demandent des précautions pour y faire adhérer les cadres. L'évaluation des personnes en poste permet également d'influer de façon significative sur la performance, de manière bien plus corrélée que le salaire par exemple. Les analyses spécifiques deviennent dans ces conditions prépondérantes sur les analyses standardisées et identiques pour tout le monde et tous les postes.

Les nouveaux problèmes posés par le développement de l'informatique conduisent à des besoins nouveaux de spécialistes qualifiés pour gérer ces nouvelles techniques. L'apparition de nouvelles méthodes de travail à distance par exemple appelle une plus grande responsabilité des individus. Les tests doivent s'adapter et ne pas se contenter de mesurer les aptitudes existantes pour évaluer également l'adaptabilité des candidats aux nouvelles exigences des postes modernes.

Ces demandes vont dans le sens d'une évaluation de l'individu dans son ensemble et induisent la possibilité de faire des analyses précises des conduites mises en cause. Cela implique des variables complexes et le développement d'observations en situation.

De nouvelles demandes venant des individus

Le besoin des individus en termes d'informations sur soi a augmenté et conduit le psychologue à repenser sa pratique. Les centres d'intérêts se sont élargi pour prendre en compte l'ensemble de la situation dans son contexte ; cela signifie donner plus d'importance à la situation dans la sélection et donner plus d'importance aux aspects humains de l'évaluation c'est-à-dire à leurs effets sur l'image de soi et l'estime de soi.

De plus, il est admis que l'expérience apporte un élément indispensable à l'acquisition de compétences mais ne suffit pas, seule, à développer la compétence. Des qualités de fond individuelles sont nécessaires pour que l'expérience soit formatrice. Deux besoins en découlent de la part des individus en termes d'information sur soi:

- Une correspondance claire des compétences et des caractéristiques individuelles nécessaires à leur développement constituant un référentiel de compétences spécifiques.
- Plus d'information sur la capacité à apprendre en tirant parti de l'expérience.

Enfin, l'augmentation du nombre de femme souhaitant concilier vie professionnelle et vie familiale génère un nouveau besoin de gestion simultanée des carrières des deux membres d'un couple. L'équilibre professionnel et familial du couple en dépend et constitue une exigence importante pour les hommes comme pour les femmes.

De nouvelles ressources

La psychologie cognitive nous apprend que le fonctionnement cognitif se décompose en plusieurs étapes liées au traitement de l'information par l'individu (réception des informations, stockage, traitement, utilisation des résultats pour élaborer des comportements plus adaptés). Ces étapes doivent être diagnostiquées de manière indépendante, avec notamment des analyses de postes déclinées selon ces étapes. Cela nécessite des moyens importants.

Les logiciels informatiques peuvent faciliter la notation professionnelle selon un entretien structuré en organisant la réception de l'information de manière précise et en la plaçant dans un cadre pour la codifier et en en définissant sa fiabilité. Le logiciel peut par exemple suggérer les comportements, analyses de situations, etc... en fonction d'informations biographiques. L'apparition de nouvelles technologies (terminaux sans fils, etc) permet de faciliter grandement la mise en œuvre pratique des tests.

Actualité de la question et analyse personnelle

Dans le contexte actuel des évolutions technologiques très rapides, les entreprises sont obligées de revoir en permanence leurs priorités et leurs stratégies pour s'adapter à leur environnement et rester compétitives. De ce fait, elles doivent faire très attention à leurs recrutements et la question de l'évaluation du personnel prend donc de plus en plus d'importance dans l'entreprise.

En effet, un recrutement a un coût élevé pour l'entreprise et avec les exigences actuelles de compétitivité et de performance qui sont imposées aux sociétés, les responsables du recrutement comme les managers opérationnels ne peuvent se permettre de faire de mauvais recrutements.

Cette question est d'autant plus d'actualité quand le climat économique est difficile et donc peu propice à la mobilité. En effet, d'une part les salariés ne veulent pas prendre le risque de quitter un poste même s'ils ne s'y sentent pas à l'aise ou si ce poste ne correspond pas totalement à leur profil. D'autre part les entreprises ont un nombre limité d'ouvertures de postes et ont une sorte d'obligation de résultat pour ne pas commettre d'erreur de recrutement. L'entreprise doit donc plus que jamais faire attention à ses recrutements

De plus, dans un monde économique globalisé, l'heure est à la délocalisation de certaines compétences dans des régions du globe où la main d'œuvre est moins chère qu'en Europe. Dans un dossier spécial, l'Expansion de mai 2004 décrit ce phénomène en montrant que la délocalisation nécessaire pour les entreprises de certaines catégories d'emplois classe les compétences selon qu'elle seront recherchées demain ou non. On peut aller plus loin en remarquant que la délocalisation en elle même va générer des besoins en compétences nouvelles pour établir de nouvelles filiales ou gérer des collaborateurs à distance depuis le siège social d'une société par exemple. De nouveaux types de besoins voient le jour, d'anciens métiers s'en vont hors de nos frontières.

On peut en déduire que l'importance de l'évaluation du personnel va croître dans les entreprises au fur et à mesure que ces mouvements vont se généraliser et s'amplifier. Il sera en effet capital de bien veiller à conserver un portefeuille de compétences exploitables en Europe et de tenir compte de ces évolutions pour recruter les cadres qui organiseront ce mouvement.

Aussi les entreprises doivent-elles trouver des méthodes de recrutement qui soient fiables et, pour reprendre le terme de Claude Lévy-Leboyer, valide. En effet, il est important de ne pas « laisser passer » le bon candidat, tout comme il est pénalisant d'embaucher la mauvaise personne.

Il est dans ces conditions assez surprenant de constater que l'étude du CV et les entretiens sont utilisés dans une écrasante majorité des cas alors que leur validité et leur fiabilité est loin d'être démontrée. Les résultats de recherche ont mis en évidence les limites de ces outils et ce depuis longtemps. Y aurait-il donc un poids des habitudes dans les pratiques de recrutements actuelles ? Serait-ce une solution de facilité de la part des responsables des ressources humaines ?

Il y a certainement une tendance chez nos professionnels du recrutement à rester dans une sorte de complaisance qui les conduit à fermer les yeux et ne pas regarder s'il existe de meilleures méthodes que celles qu'ils utilisent habituellement. C'est ce que le livre de Claude LEVY-LEBOYER cherche à combattre en insistant sur la nécessité de choisir des méthodes adaptées à chaque situation, et en demandant aux professionnels de s'assurer du bien-fondé de leur méthode d'évaluation et en leur en donnant accès aux outils nécessaires. Idéalement, il faudrait que les recruteurs adaptent leur méthode d'évaluation à chaque poste à pourvoir.

Ces préconisations sont-elles réalisables en pratique ? Evaluer une méthode d'évaluation avant de l'utiliser peut se révéler très coûteux, lourd et fastidieux à mettre en œuvre. Il est vraisemblable que les responsables des ressources humaines n'y puissent y consacrer autant de temps et d'énergie que nécessaire. Dans ces conditions, le grand intérêt de l'ouvrage de Claude LEVY-LEBOYER est de présenter les résultats de recherches associées à chacune des méthodes d'évaluation qui existent (cv, entretiens, tests de situation, d'aptitudes...) en positionnant la validité de ces tests, en la plaçant dans le contexte des postes ou des types de postes à pourvoir et en donnant des éléments chiffrés sur leur efficacité.

Plus encore, l'auteur s'attache à décrire les conditions dans lesquelles certaines validités prouvées peuvent se généraliser à d'autres types de postes ou d'entreprises. Le recruteur n'est donc plus obligé de tester lui-même la méthode d'évaluation, mais doit se plonger dans l'analyse et des rapports de recherches existantes pour comparer les différentes méthodes. Le fera-t-il ? Il est permis d'en douter...

Enfin, et pour conclure, le développement de l'informatique permettra sans doute d'améliorer la qualité des instruments d'évaluation. Les tests adaptatifs sont des exemples d'amélioration de ces méthodes. De même, l'opérateur téléphonique américain ATT se penche sur la question de faire passer les tests par téléphone où les candidats répondront aux questions en appuyant sur les touches de leur clavier téléphonique. Des gains en efficacité existent certainement dans ces approches mais il est clair qu'elles ne peuvent compenser la perte de rapport humain tant appréciée par les recruteurs et par les candidats.

Reste un champs d'investigation important que Claude LEVY-LEBOYER ne manque pas de souligner à juste titre : les méthode d'évaluations et les études de validités associées viennent dans la très grande majorité des cas d'outre-atlantique. Les Etats-Unis ont en la matière une avance considérable sur l'Europe et sur la France en particulier. Il convient donc de développer ce secteur de recherche en France pour adapter les conclusions des études américaines au marché français et à ses spécificités, et de développer en France les méthodes les mieux adaptées. C'est probablement une des clés, à terme, de la compétitivité de la France dans le monde.