



Validité prédictive des méthodes d'évaluation.

Analyse comparée des différentes méthodes d'évaluation en sélection.

Depuis maintenant plus de 100 ans, la psychologie a cherché à identifier quelles étaient les méthodes d'évaluation qui étaient les plus à même de prédire la capacité des personnes à réussir en contexte professionnel.

La méta analyse de Frank L. Schmidt

L'ensemble des travaux réalisés en la matière a récemment été synthétisé par *Franck L. Schmidt* de l'Université de l'Iowa en 2016 au travers d'une méta analyse intitulée : **The Validity and Utility of Selection Methods in Personnel Psychology: Practical and Theoretical Implications of 100 Years...**

Dans cette méta analyse, 31 méthodes d'évaluation ont été étudiées et comparées rigoureusement afin de distinguer celles qui apportent un réel incrément aux processus de présélection de celles qui ne sont pas ou peu utiles... voire néfastes en terme de capacité à distinguer les candidats du point de vue de leur faculté à réussir en poste.

Vous trouverez dans les pages qui suivent un condensé des résultats.



La validité prédictive, qu'est-ce que c'est ?

La validité prédictive d'une méthode d'évaluation représente la mesure selon laquelle les pronostics qu'elle permet de formuler sont exacts.

Elle est souvent appréciée au travers de la corrélation des éléments qui ressortent de l'application de cette méthode (ex. scores obtenus à un questionnaire de personnalité, à un test de connaissances techniques ou encore recommandations établies à la suite d'un entretien) avec le ou les critères que l'on souhaite évaluer (ex. la performance délivrée ultérieurement par les individus en poste une fois ceux-ci recrutés).

Cette corrélation peut théoriquement varier de -1 à +1.

Le cas $r = +1$

Une corrélation de +1 signifierait que la méthode d'évaluation prédit à 100% la performance des individus.

En gros, vous pourriez avoir confiance "les yeux fermés" en cette méthode. Elle permettrait de prendre les bonnes décisions 100% du temps et rendrait possible l'anticipation parfaite de la future performance en poste de chaque candidat potentiel.

Il va de soi qu'aucune méthode d'évaluation ne permet d'atteindre de tels résultats. La complexité des phénomènes à prédire et l'intrication des causes qui les déterminent ou les influencent rendent en effet impossible l'atteinte d'un tel pouvoir prédictif.

On peut en revanche - comme nous allons le voir - tenter de s'en rapprocher.

Le cas $r = -1$

Une corrélation de -1 signifierait que les résultats délivrés par les personnes en poste sont systématiquement inverses par rapport aux pronostics réalisés par cette méthode.



En gros, à chaque fois que la méthode anticipe une réussite du candidat, celui-ci échoue... et réciproquement. Il n'y aurait donc qu'à délibérer en sens inverse des préconisations fournies par l'application de la méthode en question. A chaque fois qu'elle dit noir, je choisis blanc. A chaque fois qu'elle dit blanc, je choisis noir.

Ce que l'on observe dans la pratique

Dans la pratique, la plupart des méthodes d'évaluation possèdent une corrélation positive allant de .01 à .70.

Dès .30, une corrélation peut être considérée comme étant acceptable.

En dessous de cette valeur ($r < .30$), les prédictions que la méthode d'évaluation permet de formuler doivent être considérées comme étant peu fiables.

Comment savoir la part de performance que l'on peut effectivement prédire avec une méthode d'évaluation en particulier ?

Le r est une corrélation entre les outputs d'une méthode d'évaluation et le critère que l'on cherche à prédire.

Pour bien se représenter l'utilité d'une méthode d'évaluation, il est conseillé d'utiliser le r^2 ("r carré"). Ce r^2 correspond à la part de variance de performance prédite par la méthode d'évaluation en question.

Exemple :

Imaginez que l'on considère 100 personnes qui occupent un poste de commercial et qui présentent toutes des résultats différents. Il y a dans l'équipe des bons, des très bons et des moins bons.

Les différences de performance constatées entre les différents commerciaux est ce que l'on appelle la variance de la performance. Cette variance pourra être très élevée si les différences entre les meilleurs et les moins bons commerciaux



sont très grandes. Elle pourra également être plutôt faible si tous les commerciaux obtiennent à peu près les mêmes résultats.

Si une méthode employée pour évaluer les commerciaux au moment de les présélectionner présente une validité prédictive de $r = .50$, cela signifie que cette méthode permet de prédire 25% des différences que l'on observera entre les meilleurs et les moins bon commerciaux.

Considérant que le maximum atteignable se situe autour de $.70$, cela signifie que l'on peut - tout au plus - expliquer environ 50% des différences entre les meilleurs et les moins bons candidats au travers des méthodes d'évaluation qui existent aujourd'hui ($.70^2 = 0,49 = 49\%$).

Est-ce beaucoup... ou pas ?

On peut en effet se poser la question. Prédire 50% de la capacité d'une personne à réussir, c'est également ne rien pouvoir dire des 50% restants...

Une analyse critique simple de la situation permet cependant de prendre conscience de l'énorme pouvoir prédictif d'une méthode qui permettrait d'atteindre les $.70$ en validité opérationnelle.

Pensez à ça une seconde : bien évidemment les caractéristiques de la personne vont avoir un impact fort sur sa capacité à réussir et à s'intégrer... mais il y a également beaucoup d'autres facteurs qui vont rentrer en jeu et qui sont susceptibles d'impacter le niveau de performance d'un individu en poste :

- La cohérence entre la façon dont le poste a été présenté et la réalité de ce poste.
- La façon dont la personne va être intégrée.
- Le qualité et le style de management de son manager.
- La culture de l'entreprise et/ou de l'équipe en place.
- L'affinité de la personne recrutée avec les membres de l'équipe.



- Les conflits qui peuvent survenir au sein de l'équipe ou avec certains interlocuteurs (en interne ou en externe).
- Les méthodes et outils de travail qui permettent d'optimiser ou non le travail des individus.
- Le contexte économique dans lequel se situe l'entreprise au moment où la personne prend le poste.
- La concurrence plus ou moins rude à laquelle l'entreprise est confrontée.
- Les changements de stratégie commerciale éventuellement décidés une fois la personne recrutées ou 6 mois après son arrivée.
- La modification des réglementations en lien avec l'exercice de la profession ou avec l'exécution des transactions dont la personne recrutée a la responsabilité.
- ...

Autant de facteurs (on pourrait en citer des dizaines et des dizaines) qui sont susceptibles de moduler la performance que les personnes recrutées délivreront une fois dans l'entreprise.

Par conséquent, être en mesure de prédire 10, 15, 20, 30 voire 40% de la performance qu'un individu sera en mesure de délivrer - avant même qu'il n'ait mis un pied dans l'entreprise - ce n'est pas si anodin que cela. C'est même plutôt conséquent si l'on y réfléchit à 2 fois !



**Ce qui ne
marche
pas.**



Les critères à bannir de votre processus de présélection.

L'un des intérêts de la méta analyse conduite par Franck L. Schmidt est de mettre en évidence certains critères parfois utilisés en sélection et qui pourtant n'ont quasiment aucun pouvoir prédictif.

Non seulement leur prise en compte n'apportera rien en terme d'information, mais il se peut surtout qu'ils puissent avoir un effet contre productif sur les décisions qui reposeront sur ces critères.

Parmi ces critères :

L'âge des candidats.

La corrélation observée entre l'âge des individus et la performance qu'ils délivrent est égale à $r = .00$.

Cela signifie qu'il y absolument aucun lien entre l'âge des individus et leur capacité à réussir en poste. On peut être très jeune et réussir au plus haut niveau tout comme on peut être beaucoup plus âgé et réussir au même niveau.

Par conséquent - d'un point de vue scientifique - sélectionner les personnes sur la base de leur âge est une aberration totale.

La formation des candidats.

La corrélation observée entre les années de formation effectuées par les candidats et la performance qu'ils délivrent est égale à $r = .10$.

Le r^2 étant de 0,01 cela signifie que cette approche permet de prédire 1% des différences qui seront présentées entre les meilleurs et le moins bons éléments recrutés.

Par conséquent, continuer à exiger des personnes que l'on recrute qu'elles aient suivi au moins 3, 4 ou 5 années d'enseignement supérieur n'a aucun sens.



Cela peut rassurer un recruteur ou un manager qui recrute (parce qu'ils sont "convaincus" que ça joue...) mais d'un point de vue strictement rationnel, cela est un non sens total.

Les années d'expérience en poste.

La corrélation observée entre les années d'expérience des candidats sur un poste donné et la performance qu'ils délivrent est égale à $r = .16$.

Le r^2 étant de 0,03 cela signifie que cette approche permet de prédire 3% des différences qui seront présentées entre les meilleurs et le moins bons éléments recrutés.

Là aussi, continuer à exiger des personnes que l'on recrute qu'elles justifient d'au moins 1, 2 voire 3 années d'expérience préalable dans le type de poste pour lequel on souhaite les recruter est une aberration.

Nota :

Notez bien que ces conclusions ne sont nullement influencées par quelque idéologie que ce soit. Elles sont la conséquence logique découlant d'une analyse méticuleuse des résultats obtenus au terme de 100 ans d'analyse des pratiques de recrutement, partout dans le monde.



**Ce qui
marche
plutôt
bien.**



Le facteur g explique 42% de la performance future.

Le meilleur prédicteur de la performance des individus en contexte professionnel est le facteur g, tel qu'évalué par les GMA Tests. ($r = .65$ soit un r^2 de 42%).

Le facteur g correspond à la capacité de la personne à raisonner de manière juste et à résoudre des problèmes de différents niveaux de complexité.

Ce résultat est parfaitement en ligne avec les conclusions auxquelles arrivent régulièrement les professionnels RH et certains managers. Ceux-ci décrivent en effet la résolution de problème comme étant la compétence la plus utile pour réussir dans l'environnement dans lequel nous évoluons aujourd'hui (plus changeant, plus incertain, plus complexe, plus ambiguë), ceci quel que soit le niveau de technicité auquel les individus ont affaire dans le cadre de leur activité professionnelle au quotidien.

BRAIN by AssessFirst

Le test BRAIN, intégré à AssessFirst a été spécialement conçu pour évaluer ce facteur g. La spécificité de BRAIN est qu'il repose sur une approche adaptative et entièrement gamifiée.

Dans les tests dits "adaptatifs", les questions posées aux individus le sont en fonction des réponses qu'ils apportent aux questions précédentes. Ceci permet de garantir une expérience plaisante : On ne va pas pousser des items démesurément compliqués à une personne qui ne sera manifestement pas en mesure de les résoudre.

Par ailleurs, l'utilisation des codes inspirés du Gaming permet d'engager les individus et de soutenir leur attention pendant le temps nécessaire à l'évaluation de leurs capacités cognitives.

A noter également que le format adaptatif permet de réduire de moitié le temps nécessaire à procéder à cette évaluation du facteur g.



Plus le poste est complexe, plus l'impact du facteur g est important.

Il est à noter que le caractère prédictif du facteur g augmente de manière exponentiel à mesure que la complexité associée au poste augmente également.

Pour un poste de complexité faible, une personne ayant un score élevé sur un test de facteur g (tel que BRAIN) délivrera un niveau de performance X fois plus important qu'une personne qui obtient un score faible sur ce même test.

Si le degré de complexité du poste est moyen, cette différence de performance constatée passe de 3 pour 1 (X3).

Pour les postes les plus complexes, la différence de performance constatée est 8 fois plus importante entre les personnes qui obtiennent un score élevé au test de facteur g et les personnes qui obtiennent des scores faibles.



La personnalité, un prédicteur puissant... dès lors qu'il est remis en contexte.

Si la validité prédictive de la personnalité (Modèle Big5 / OCEAN / FFM) varie en fonction des traits considérés, il est à noter toutefois que celle-ci présente un fort intérêt en sélection des personnes.

L'utilisation combinée d'un test de GMA (BRAIN) et d'un inventaire de personnalité (SHAPE) permet en effet d'atteindre des niveaux de prédictivité allant de $r = .65$ à $r = .78$ soit une variance de performance prédite allant de 42% à 61%.

SHAPE by AssessFirst

Le questionnaire de personnalité SHAPE développé par AssessFirst permet d'évaluer 20 dimensions de personnalité qui influencent et conditionnent les comportements que les individus adoptent en situations professionnelles.

Du point de vue de la personnalité, il est bien important de comprendre que les dimensions sont assez peu prédictives de la réussite au travail "en soi".

Chaque fonction, chaque métier, chaque poste comporte en effet des spécificités qui lui sont propres.

Par conséquent les dimensions de personnalité impliquées dans la prédiction de la réussite varient énormément d'une fonction, d'un métier et d'un poste à l'autre.

Il est donc indispensable, préalablement à l'utilisation de questionnaires de personnalité de bien définir les critères qui influent sur la réussite dans les postes considérés.



Les motivations, un complément utile.

Les intérêts et motivations (telles qu'évaluées par le questionnaire DRIVE) possèdent en soi un pouvoir prédictif moyen de $r = .31$. Leur prise en compte conjointe avec le facteur g des individus permet d'obtenir un validité opérationnelle $r = .71$ (soit un r^2 de 50%).

Tout comme pour la personnalité, les motivations nécessitent d'être contextualisées si l'on souhaite en augmenter le pouvoir prédictif en sélection professionnelle.

Il va de soi en effet que des postes différents font appel à des motivations différentes.

DRIVE by AssessFirst

Le questionnaire de motivations DRIVE développé par AssessFirst a été conçu tout spécialement pour appréhender un spectre le plus large possible des moteurs individuels.

Ainsi, ce sont 20 besoins fondamentaux qui sont évalués au travers du seul questionnaire DRIVE.

Il peut donc être utilisé pour prédire la réussite professionnelle des individus dans un spectre très large de fonctions, de métiers et de postes.

Ces mêmes besoins se sont par ailleurs révélés particulièrement utiles pour participer à définir ce qui fonde les spécificités culturelles d'une entreprise.



Maximiser le pouvoir prédictif de votre process de (pré)sélection...

La maximisation du pouvoir prédictif d'un dispositif d'évaluation - que celui-ci soit utilisé en sélection ou en présélection - est directement lié à la capacité du recruteur à identifier les bons prédicteurs de la réussite.

Pour identifier les meilleurs prédicteurs de la réussite pour un poste donné, AssessFirst recommande très fortement de créer des modèles prédictifs sur la base de vos populations en poste en utilisant notre module d'Intelligence Artificielle (I.A.).

Ceci vous permettra de disposer d'algorithmes personnalisés présentant un haut pouvoir prédictif (validité opérationnelle $r = .65$ à $.70$ en moyenne).

Pourquoi la plupart des recruteurs obtiennent des résultats mitigés ?

On sait aujourd'hui qu'une grande partie des recrutements effectués par les entreprises se soldent par un échec à court terme.

Au bout de 6 mois à peine, c'est en effet 19% des candidats recrutés qui ne sont plus en poste. Au bout d'un an, 36% des candidats ont quitté l'entreprise ou se sont fait licencier. A 18 mois, c'est 46% des personnes recrutées - soit près d'une sur 2 - qui ne font plus partie des effectifs.

Et dans 89% des cas, ceci est dû à une mauvaise prise en compte des facteurs comportementaux.

Donc d'une part l'accent est encore trop souvent porté sur les compétences techniques (hard skills) alors que les soft skills ont un pouvoir prédictif très fort, ceci dans de nombreux postes.

Par ailleurs, la plupart des recruteurs ne sont pas suffisamment rigoureux dans la définition des facteurs qui leur servent de critères de présélection.



La plupart du temps, ces critères sont en effet définis en demandant tout simplement au manager ses préférences concernant le profil à recruter.

La limite énorme de cette méthode, c'est qu'elle conduit souvent à ne recueillir que les représentations du manager opérationnel concernant ce qu'il pense être prédictif de la réussite. On est souvent bien loin de la réalité.

Observer la réalité plutôt que de projeter ses propres représentations.

Par conséquent, lorsque vous souhaitez procéder à l'identification de ces critères prédictifs de la réussite en poste - préalablement à la présélection des candidats - nous vous recommandons très fortement d'observer la réalité plutôt que de vous baser sur les impressions de telle ou telle personne, peu importe qu'elle occupe elle-même ou pas le poste pour lequel vous avez à recruter.

La meilleure façon de le faire est d'évaluer les personnes occupant effectivement ce poste et de recueillir pour chacune d'entre elles les indicateurs qui servent à évaluer leur performance au quotidien (ou chaque semaine, chaque mois...). C'est ensuite l'analyse croisée rigoureuse des ces différents types de données qui vous permettra de mettre en lumière les critères sur lesquels vous devriez opérer votre présélection.

Notez qu'AssessFirst a développé une Intelligence Artificielle qui vous permet d'automatiser une très grande partie du travail nécessaire à l'identification de ces prédictifs.

About AssessFirst

AssessFirst has developed a predictive recruitment solution allowing companies to predict how well candidates and employees will succeed and thrive in their job. The AssessFirst solution analyses data on over 5,000,000 profiles, whether candidates, employees or recruitment professionals.

Today, over 3,500 companies use the AssessFirst solution to raise their performance by up to 25%, drive down their recruitment costs by 20% and reduce their employee turnover rate by 50%.

Find out more: www.assessfirst.com

