

La déontologie SIGMUND des tests d'évaluations

Les tests d'évaluations SIGMUND répondent à des exigences déontologiques rigoureuses.

Des exigences de compétences

Une équipe pluridisciplinaire

- . Dirigeants et hommes de la fonction personnelle
- . Consultants en ressources Humaines et formation
- . Psychologues, sociologues, chercheurs
- . Informaticiens et statisticiens.

1. Des exigences méthodologiques

La construction d'un test s'effectue en plusieurs étapes toutes rigoureusement menées par SIGMUND:

a) L'existence d'hypothèses de recherches

b) Conception et définition des critères que l'on souhaite évaluer en fonction des besoins des hypothèses de départ

Construction du test et des questions (items)

Ces items sont construits de façon hypothétique, il conviendra de vérifier qu'ils évaluent bien ce qu'on souhaite mesurer. On construit donc entre 3 et 4 fois plus d'items que ce qui est nécessaire.

c) Expérimentation sur un échantillon représentatif.

Cette expérimentation a pour but de vérifier la qualité des items par différentes méthodes statistiques.

La plus couramment utilisée est **l'analyse factorielle**.

A partir d'une étude des inter-corrélations entre les réponses de tous les sujets à tous les items des tests, l'analyse factorielle permet de vérifier que les items hypothétiquement prévus pour mesurer une variable, mesurent bien cette variable.

On analysera aussi le **potentiel discriminant des** items. Si plus de 80 % des personnes, ou moins de 20 %, choisissent la même réponse à un item, cela signifie que ce dernier n'est pas discriminant : si l'on évalue les aptitudes intellectuelles, c'est que la question est trop facile ou trop difficile, quand on évalue la personne dite ou les valeurs, on peut penser que la réponse correspond à une image sociale valorisante ou dévalorisante.

C'est aussi pendant la phase d'expérimentation que l'on vérifiera la compréhension des consignes de passation.

d) Une validation externe de ces hypothèses

Comparaison du test à des référentiels extérieurs au test

e) Une double validation interne

Vérification de la consistance du test

f) Un test de fiabilité

Vérification de la fiabilité du test

g) Un remaniement :

Sur la base des données collectées, des questions nécessitent d'être ajustées, voire réécrites. En fonction des statistiques obtenues les cotations sont définitivement élaborées.

h) Un étalonnage de plus de 1000 personnes :

L'étalonnage permet la construction des normes et de référentiels métiers précis

On vérifiera que la distribution des résultats des différentes personnes constituant l'échantillon répond bien à la loi normale théorique qui est **représentée par la Courbe de Gauss**.

Pour ce faire, on construira l'histogramme des résultats, qui peut, **par exemple prendre** la forme suivante **pour un critère** évalué par 15 **questions**.

i) Révision du test

En général, les notes des sujets se distribuent selon la loi normale théorique (courbe de Gauss). Plus la taille de l'échantillon est grande, plus la courbe sera proche de courbe normale parfaite. La courbe de Gauss est caractérisée par sa symétrie et sa forme en cloche (en effet un petit nombre de personnes ont des notes très basses ou très élevées, le plus grand nombre ayant une note moyenne).

La fiabilité d'un test dépend de ses 9 éléments, toujours réunis par les tests SIGMUND

2. Des exigences d'expérience « Terrain »

La « maturité terrain » de SIGMUND est aussi un élément de notre savoir faire :

1. Prés de 800 000 d'évaluations réalisées sur le terrain, plus de 20 ans d'expérience
2. Adaptation à différents métiers, cultures d'entreprise, pays (9 langues)
3. Présence dans plus de 200 entreprises

Des exigences éthiques pour les candidats

Avant d'utiliser un test, il conviendra de vérifier sa construction, les qualités métrologiques (validité, fiabilité) et la qualité de l'échantillon (taille, représentativité, ancienneté ...).

Avant d'interpréter un test, il conviendra de choisir le « bon » étalonnage. Pour les tests de personnalité, il est conseillé de commencer par l'étalonnage général puis d'affirmer l'interprétation en utilisant un (ou des) étalonnage (s) spécifiques.

Pour l'interprétation d'un test, on se réfère toujours aux notes standards (redressées) et jamais aux notes brutes.

Cela permet de positionner une personne par rapport à la population de référence.

Par exemple, si une personne obtient une note standard de 1 en sociabilité, cela signifie que, pour cette dimension, elle fait partie des 8,1 % des personnes qui obtiennent la note la plus basse et que 3,6 % des personnes obtiennent une note encore plus basse en sociabilité.

Les étalonnages spécifiques :

Pour un test donné, on pourra construire un étalonnage général, sur l'ensemble de la population mais aussi des étalonnages spécifiques sur une partie de la population : **par exemple** des étalonnages en fonction des métiers ou de la fonction occupée.

L'intérêt des étalonnages spécifiques est d'affiner l'interprétation en ciblant davantage la population d'appartenance.

Pendant l'interprétation, on prendra en compte, non seulement la définition des dimensions, mais aussi toutes les données statistiques descriptives du mode de fonctionnement moyen.

Ne jamais oublier qu'un test est **un outil psychométrique**, qu'il permet d'évaluer les dimensions, en fonction d'un mode de comportement moyen (la norme).

Tout test doit respecter les règles éthiques de respect de la personne évaluée. Pour SIGMUND:

- Comparaison du candidat sur chacun des critères évalués, avec une moyenne fondée sur les résultats d'au moins 1000 personnes.
- Restitution des résultats
- Confidentialité de la modalité de réponse choisie, pour chaque question
- Transparence des critères évalués

La méthodologie SIGMUND

Un outil informatique

La société *SIGMUND* a mis au point une approche nouvelle de l'évaluation du facteur humain en entreprise. Cette approche nouvelle est à l'origine d'une gamme d'outils informatisés élaborés par le laboratoire de psychologie de la société.

Ces outils se proposent d'établir un diagnostic quantifié sur un certain nombre de critères, à partir de la passation d'une évaluation informatisée par le biais d'un questionnaire.

3. Différents types de tests

SIGMUND propose plusieurs catégories de tests.

- **Les tests d'intelligences**

Le mesure de l'intelligence reste un sujet grandement controversé dans la mesure où il reste difficile de donner une seule et simple définition de l'intelligence. Les tests d'intelligences sont donc aussi nombreux qu'il existe de définition différente de l'intelligence.

Schématiquement, les tests d'intelligence sont ceux qui identifient les conduites intellectuelles par la capacité à conceptualiser, à raisonner, à résoudre des problèmes, et à s'adapter à des situations nouvelles. Ces capacités mettant en œuvre les processus cognitifs les plus hiérarchiquement élevés et complexes de la pensée humaine.

Cela inclus la capacité à acquérir des connaissances et à les appliquer, à la capacité à raisonner et à produire un jugement, la capacité à faire preuve d'intuition, à trouver le bon mot au bon moment, la capacité à s'adapter à des situations nouvelles ou de nouveaux type de problème.

Ces tests d'intelligence ont pour finalité de situer le sujet par rapport à ses pairs.

- **Les tests d'attitudes professionnelles**

Les aptitudes des tests SIGMUND sont des dispositions à effectuer des tâches professionnelle, et a avoir les comportement adaptés avec une plus ou moins grande efficacité. Ils ont pour objectifs, à travers un grand nombre de questions, de dresser

un profil des habiletés professionnelles. Ils sont construits sur la base d'analyses factorielles en fonction des critères explicatifs de l'échec ou du succès professionnel et fournissent un spectre très large de facteurs.

L'interprétation de ces tests est particulièrement délicate et requiert une formation ainsi qu'une solide expérience. Des recherches ont montré que les tests d'aptitudes professionnelles fournissent des informations plus objectives que l'entretien ; ces tests permettent d'éviter la subjectivité et fournissent une appréciation fine et pertinente. Leur utilisation dans le cadre professionnel est garantie par une déontologie très stricte.

Les tests d'aptitudes de SIGMUND sont rassemblés en batteries en fonction des spécificités professionnelles.

- **Les tests de motivation**

Ces tests évaluent les motifs de motivation et de démotivation en situation professionnelle.

- **Les tests d'intérêt**

Ces tests appréhendent les préférences qui vont motiver la personne, qui vont lui permettre de se réaliser professionnellement. Les tests d'intérêts de SIGMUND portent sur le choix professionnel : intérêt pour les nombres, pour les objets, pour les personnes, intérêt scientifique, commercial ...

Ils sont particulièrement utiles pour la personne elle-même qui doit choisir un métier ou une voie de formation.

Les questionnaires d'intérêts ou de valeurs, comme d'ailleurs les tests d'attitudes, sont un excellent support de dialogue et d'échanges entre le conseiller et la personne, qui a ainsi l'occasion de parler d'elle, de mieux se connaître et de mieux comprendre l'image qu'elle s'est forgée d'elle-même.

Ce type de tests enrichit particulièrement une situation de bilan professionnel et personnel, où la personne est acteur et au centre des investigations.

4. Tests ipsatifs et tests normatifs

Test Ipsatif : c'est un test qui dans son approche des cotations et de l'interprétation évalue les forces et les qualités d'un individu au regard des autres qualités mesurés. Les tests dont le nom finit en Pro sont chez SIGMUND des tests ipsatifs

Test Normatif ; c'est un test dans le quel les réponses de l'interviewé le positionne par rapport à une catégorie particulière (ici l'étalonnage).

Ces tests SIGMUND se proposent d'établir un diagnostic quantifié sur un certain nombre de critères (20 à 50 critères selon le type d'outil), à partir de la passation d'une évaluation informatisée de 120 à 450 questions.

La fiabilité de ces outils, la qualité de l'information recueillie et traitée par ces systèmes repose en grande partie sur le support informatique qui permet à la fois :

5. La maîtrise des conditions expérimentales

- Standardisation des conditions de passation
- Grilles de correction automatisées
- Forte implication du candidat (test plus convivial et plus ludique, confidentialité des réponses)
- Pour certain test, l'intégration au logiciel d'une échelle d' « *adhésion au test* » pour contrôler la sincérité

6. La réalisation de gros échantillons de référence

Etalonnage des questionnaires sur d'importantes populations de référence.

Validations à grande échelle qui permettent de réaliser des études statistiques fiables.

Mises à jour périodique des tests et ré-étalonnages.

7. Deux hypothèses

La méthode qui sous-tend la création et la fiabilisation des systèmes *SIGMUND* est de type scientifique. Elle a consisté en la formulation d'une hypothèse générale, concrétisée par des hypothèses de recherche :

- Hypothèse générale :

Il est possible de mesurer les attitudes le comportement, le mode de fonctionnement d'une personne dans un cadre de travail

- Hypothèse de recherche :

On peut mesurer les valeurs de réussite professionnelle et les comportements et la motivation au travail, ainsi que le sens des responsabilités d'un individu en entreprise.

8. Des travaux de validation et d'analyse des données

Après construction et mise au point du questionnaire un premier étalonnage a été effectué. La construction de cet étalonnage a permis de mettre en évidence des courbes de Gauss pour chacun des critères et de mener les travaux suivants :

Opérations de validation

Il s'agit de vérifier les hypothèses de recherche qui fondent le test : pour vérifier que le système mesure bien des valeurs de réussite professionnelle et de motivation au travail

Détermination d'axes explicatifs des comportements

On utilise l'analyse factorielle en composantes principales pour situer les critères les uns par rapport aux autres, observer quels critères fonctionnent dans le même sens, quels groupe de critères s'opposent.

Cette analyse permet d'affiner la compréhension des valeurs véhiculées et mesurées par les systèmes et de fiabiliser l'évaluation.

Une matrice d'inter corrélation linéaire permet de connaître les liens existants entre les critères.

Double contrôle de fiabilité

Il s'agit de vérifier que la performance du test est la même dans une variété de circonstances différentes.

La démarche SIGMUND

Un questionnaire construit sur deux hypothèses de recherche

9. Des hypothèses de recherche sur les relations Homme/Entreprise

Avant d'élaborer un test d'évaluation, il convient de poser un certain nombre d'hypothèses de recherche. Nous avons donc posé les questions suivantes :

- Peut-on mesurer le potentiel humain en entreprise ?
- Peut-on mesurer les motivations de l'homme à travailler ?
- Peut-on mesurer la volonté de réussite en entreprise ?
- Peut-on mesurer les différentes attitudes et comportement en situation professionnel
- Peut-on mesurer le sens des responsabilités en entreprise ?

10. Une approche empirique

Les attitudes à mesurer son toujours issu des besoins des entreprises.

La démarche SIGMUND est conduite en plusieurs étapes :

- Interviews de responsables du recrutement et de chefs d'entreprise, afin de préciser leurs critères d'appréciation des performances.
- Compilation des principales grilles de tests utilisées par les services de recrutement, pour analyser les critères d'évaluation.
- Analyse bibliographique des principaux tests psychologiques utilisés en entreprise.
- Synthèse générale et élaboration d'hypothèses de recherche (sur l'évaluation du potentiel, des attitudes professionnels, de la motivation au travail, du sens des responsabilités).
- Elaboration d'une grille de critères en rapport avec ces hypothèses de recherche.
- Rédaction des questions pour les critères à mesurer

Cette démarche pluridisciplinaire, réunissant des psychologues, des sociologues, des consultants en ressources humaines, des D.R.H. et des managers, a permis de déterminer une grille d'évaluation, synthèse entre les dimensions évaluées par les méthodes psychologiques traditionnelles (questionnaires de personnalité), et les critères de performance et d'attitudes professionnels retenus comme indispensables par les entreprises.

11. Un questionnaire d'un nouveau type

La détermination de traits mesurables

Pour chacun de ces critères de personnalité ont été définis un certain nombre de traits de caractère, d'attitudes et de comportements, ainsi qu'un ensemble de situations dans lesquelles le trait de personnalité décrit ou le comportement attendu est censé intervenir. Là encore une démarche pluridisciplinaire a été mise en œuvre, lors de séances de créativité réunissant des psychologues, des chargés de recrutement et des responsables d'entreprise.

Dans un second temps ont été imaginées des situations courantes de la vie quotidienne, professionnelle et sociale, mettant en scène d'autres personnes (collègues, supérieurs hiérarchiques etc...) pour y intégrer les traits, attitudes et comportements ainsi définis. Ces situations à mesurer qui n'étaient pas directement observable.

La maquette expérimentale du test comporte deux fois plus de questions que ce que l'on souhaite avoir dans la version finale. Les questions sont multidimensionnelles, c'est-à-dire que pour chaque situation le choix d'une réponse entraîne un marquage quantitatif en plus ou en moins dans plusieurs critères à la fois.

12. Un test plus impliquant

Les questionnaires SIGMUND présentent un certain nombre de particularités, indépendamment de leur présentation informatisée.

La plupart des questions évitent le choix alternatif « Oui/Non » dans leurs modalités de réponse, le comportement humain n'étant pas binaire.

Le principe du test informatisé entraîne une implication importante de la part du passateur du test, liée à l'interactivité du système. Par ailleurs les mécanismes défensifs générés par la situation d'évaluation sont moindres puisque les réponses restent confidentielles : en effet, l'évaluateur n'a pas accès aux réponses particulières faites à chaque question, mais uniquement au résultat global du critère.

Vérification des qualités métrologiques

La validation d'un test consiste à vérifier les qualités métrologiques du test, c'est à dire ses qualités de mesure.

En plus des informations statistiques de base sur le test, les deux qualités métrologiques les plus importantes sont **la validité** et **la fiabilité**.

13. La validité

La **validité** permet de vérifier que le test évalue bien ce qu'il est censé évaluer, c'est à dire le degré de pertinence des résultats du test par rapport à ce qu'il prédit.

On vérifiera plusieurs types de validités :

La validité **interne**, qui est liée aux études statistiques effectuées pendant la construction du test (analyse factorielle **par exemple**)

La validité **externe** qui est liée à la comparaison du test à des critères externes (autres tests, critères de réussite, jury, ...)

14. La fiabilité

La **fiabilité** concerne les prévisions de la mesure, la confiance que l'on peut avoir dans la note de la personne.

Les études de **fiabilité** permettent de quantifier la marge d'erreur.

Il y a plusieurs façon d'estimer la fiabilité, la plus fréquente est le test — Re-test : un même groupe de personnes passe le test à un intervalle de temps de 1 à 2 mois, dans les mêmes conditions. On vérifie alors la proximité des résultats.

Ce type de démarche permet de calculer **l'erreur type de mesure**, qui, exprimée en notes brutes, nous indique la zone de confiance que nous pouvons accorder à une note.

Par exemple, si pour une dimension donnée l'erreur type de mesure est de 4, cela signifie que pour une personne obtenant 10 à cette dimension, sa note « réelle » se situe entre 8 et 12.

15. Les étalonnages

Le test a-t-il été appliqué à suffisamment de personnes pour pouvoir légitimer sa valeur normative ?

« On appelle mesure l'attribution de nombres à des objets ou à des événements selon certaines règles ».

Les résultats d'un test sont généralement exprimés sous forme de score. Les statistiques qui accompagnent les scores sont des outils servant à décrire les scores, établir des relations entre les scores et exprimer des conclusions sur le sens des scores.

La mesure en psychométrie consiste à assigner des nombres à des caractéristiques de certain objet (comportement, aptitudes, attitudes, etc.).

Les échelles utilisées peuvent être soit nominales soit ordinales. L'échelle nominale fait référence à un groupe (homme / femme). L'échelle ordinale permet en plus de classification selon les propriétés accordées aux nombres. (Modalités de réponses en 1, 2, 3, 4, 5 où l'on sait que 5 est une catégorie supérieure à 1)

Les chiffres et les graphes utilisés pour décrire et représenter des objets sont regroupés dans la catégorie des statistiques descriptives.

Un bon moyen pour comprendre le sens donné au chiffre et de les représenter graphiquement. Une autre solution est de condensé l'information sous forme de statistiques.

Avant d'appliquer des formules statistiques, il peut être intéressant d'organiser les données pour mieux les comprendre. La distribution en fréquence permet cela.

Les tests **ne sont que des instruments qui fournissent des informations quantifiées permettant d'élaborer des hypothèses**. Comme tout instrument de mesure, les tests doivent prouver la qualité de leur conception. Un test doit donner des garanties aux utilisateurs. Il est tout à fait possible de connaître le niveau de fiabilité d'un test.

Une notion importante en psychométrie : la corrélation

Il y a corrélation quand 2 variables évoluent de façon simultanée. On prendra en compte 2 facteurs : le sens et la force de corrélation. Le **sens** : positif ou négatif

On parle de corrélation positive quand les deux variables évoluent dans le même sens : l'une augmente, l'autre augmente; l'une diminue, l'autre diminue.

Par exemple, on constate qu'il existe une corrélation positive entre la longueur des jambes et la taille des pas.

On constate aussi qu'il y a une corrélation négative, en voiture, entre la vitesse et la sécurité.

La force : la liaison entre les deux variables est-elle assez forte pour qu'elle ait une signification et ne soit pas due au hasard?

On calcul donc **la significativité** d'une corrélation.

L'analyse de la corrélation entre deux variables permet de calculer un **Coefficient de corrélation = degré de liaison entre les deux variables**

Les coefficients de corrélations se situent entre -1 et + 1 et s'expriment de la façon suivante : .80, .20, -.70, -.30

On constatera que l'on indique ni le 0, ni le signe +.

Dans la **construction** d'un test, les études de corrélations sont à la base du choix des items, des études de validité, de fiabilité, des études sur le fonctionnement d'une population donnée par rapport aux différentes dimensions évaluées.

Dans **l'interprétation** d'un test, il est important de tenir compte des constats faits sur le fonctionnement des différentes populations.

Cela permet d'affiner ou de pondérer l'interprétation et d'éviter les projections personnelles.

Vérification de la Validité

Définition

Jusqu'à quel point l'instrument mesure ce qu'il est supposé mesurer.

Exemple:

Si j'utilise une série de questions censées mesurer le travail en équipe et que je me rends compte que j'ai en fait mesuré l'intérêt pour les autres, ma mesure est une mesure non valide du travail en équipe.

Il existe deux types de validité:

Validité Externe

La **validité** permet de vérifier que le test évalue bien ce qu'il est censé évaluer, c'est à dire le degré de pertinence des résultats du test par rapport à ce qu'il prédit.

La validité **externe** qui permet la comparaison du test à des critères externes (autres tests, critères de réussite, jury, ...)

16. Corrélacion scores du test par rapport à l'Auto-évaluation

Les principes de l'auto-évaluation

Une question d'auto-évaluation est posée pour chacun des critères de personnalité. L'idée est de respecter le principe de contradiction du candidat : lui fournir la possibilité de donner son avis sur les qualités qu'il se reconnaît à lui même, en contredisant éventuellement l'évaluation opérée par le système.

Validation Interne

La validité **interne**, indique jusqu'à quel point chaque indicateur constitue une mesure d'un et d'un seul concept. On utilisera plusieurs outils statistiques pour vérifier les relations des critères entre eux, la constitution des critères, les relations entre les critères et les questions, les relations des questions entre elles.

17. L'analyse factorielle

C'est avec l'application de traitement statistique que cette hypothèse a été confirmée. Vers 1904, le Britannique SPEARMAN, inventeur de l'analyse factorielle, montre qu'en effet il existe des relations entre les résultats obtenus à différentes épreuves, et que cette corrélation s'explique par la mise en évidence d'un facteur général appelé « facteur g », qui conditionne la réussite à de nombreuses tâches.

L'analyse factorielle est une méthode mathématique dont le but est de regrouper un grand nombre de données en plus petit nombre de catégories ou de facteurs distincts. La base de l'analyse factorielle est le coefficient de corrélation. Celui-ci est un index numérique indiquant le degré de liaison entre 2 données.

L'analyse factorielle consiste en l'application du test à un large échantillon de la population concernée et à vérifier la structuration des réponses en fonction d'un certain nombre de facteurs. Les résultats obtenus par un grand nombre de sujets permettent de construire les normes d'étalonnage. Si l'échantillon de sujets a été construit sans biaisage, alors leurs résultats se distribuent suivant une courbe « normale », qui permet des découpages en classes ou notes standardisées.