



Article professionnel

Article

2001

Published version

Open Access

This is the published version of the publication, made available in accordance with the publisher's policy.

Les tests d'intelligence

Lecerf, Thierry

How to cite

LECERF, Thierry. Les tests d'intelligence. In: Psychoscope, 2001, vol. 22, n° 6, p. 24–27.

This publication URL: <https://archive-ouverte.unige.ch/unige:18954>

Les tests d'intelligence

Des instruments qui mesurent les compétences acquises et non innées

Les tests d'intelligence sont ceux qui ont donné lieu aux controverses les plus virulentes. Le psychologue franco-suisse Thierry Lecerf, de l'Université de Genève, revient sur les principaux problèmes liés à ces mesures. Et il démontre leur apport dans le diagnostic, le pronostic ou l'orientation.

Ce fait est reconnu depuis des centaines d'années : les individus diffèrent les uns des autres – sur le plan cognitif, comportemental et de la personnalité – et ces différences peuvent être évaluées. Dès 2200 avant J.-C., les mandarins de la Chine impériale avaient introduit l'utilisation de tests pour sélectionner les « fonctionnaires ». De même, Platon et Aristote avaient proposé l'emploi de tests pour le recrutement professionnel.

L'utilisation de tests a ensuite disparu et il faudra attendre la fin du XIX^e siècle, avec notamment Galton et Binet, pour que l'idée de mesurer les processus mentaux ressurgisse. De nos jours, les tests engendrent à la fois fascination et inquiétude. Preuve en est le nombre de magazines proposant régulièrement d'évaluer tel aspect de la personnalité ou telle compétence cognitive.

Les psychologues francophones disposent de plus de 2000 tests : cognitifs (intelligence et aptitudes), conatifs (intérêts, motivation), affectifs (questionnaires de personnalité, tests projectifs), etc. La notion de test recouvre une grande diversité, si bien qu'il serait illusoire de les aborder tous. Aussi cet article sera-t-il centré sur une catégorie

particulière, les tests d'intelligence, qui sont utilisés dans trois grands secteurs de la psychologie : l'éducation, la santé et le travail.

La première partie est consacrée à l'ambiguïté de la notion d'intelligence et de sa mesure. Dans la deuxième partie, l'utilité des tests d'intelligence est abordée, avec notamment les problèmes de validité. Pouvons-nous, sur la base des performances observées à un test d'intelligence, prédire la réussite scolaire, professionnelle, et de quelles informations disposons-nous sur le fonctionnement cognitif de la personne ? Cette deuxième partie permettra de confronter les tests d'intelligence aux théories actuelles. La troisième partie porte sur le problème du vieillissement des tests, notamment en termes d'étalonnages. Enfin, la dernière partie aborde un des problèmes majeurs des tests d'intelligence, celui de l'origine des différences individuelles. Elle vise à répondre à la question suivante : les tests d'intelligence témoignent-ils des capacités innées des sujets ?

Qu'est-ce que l'intelligence ?

En premier lieu, la construction et l'utilisation des tests d'intelligence pose un problème conceptuel : qu'est-ce que l'intelligence ? Étymologiquement, le mot « intelligence » vient du latin *intellegere* qui signifie *comprendre*. On se souvient que, selon Wechsler (1944), l'intelligence est « la capacité d'un individu à agir en fonction d'un but, à penser rationnellement et à faire face de manière efficace à son environnement ». Les psychologues s'accordent sur cette définition générale. Mais les divergences surviennent si l'on cherche à donner une définition plus précise, et sur la structure de l'intelligence : existe-t-il une intelligence générale (facteur *g* de Spearman) ou des intelligences spécifiques à des domaines particuliers (théorie de Thurstone ; théorie de Gardner) ? Le modèle consensuel de Carroll (1993) propose une structure hiérarchique avec un facteur général d'intelligence et des aptitudes plus spécifiques. Les divergences sur la définition précise de l'intelligence ne repré-

sentent pas des obstacles à sa mesure, puisque notre culture a contextualisé la notion d'intelligence, en considérant qu'elle repose sur des aspects que l'on peut qualifier de verbo-conceptuels.

Comment la mesurer ?

La construction des tests d'intelligence pose également des problèmes méthodologiques : comment mesurer l'intelligence ? C'est ici que survient la confusion entre concept et mesure, dont les psychologues sont en partie responsables. On se souvient de la boutade de Boring (1923) : « L'intelligence, c'est ce que mesure mon test. » Boring semble indiquer que l'intelligence peut se réduire à sa mesure, par exemple au QI. Cette assertion jette un trouble entre une dimension théorique, l'intelligence, et une dimension opérationnelle, le QI. En fait, ces deux termes ne sont pas synonymes. Le QI est un indice de mesure de l'intelligence et non l'intelligence. En conséquence, les tests fournissent des données quantitatives qui ne mesurent pas directement l'intelligence. Le psychologue infère alors le niveau d'intelligence à partir des performances observées au test. Dans l'optique de supprimer cette confusion, on pourrait abolir le terme d'intelligence et le remplacer par celui d'aptitudes scolaires, de potentiel académique, etc. En résumé, les tests d'intelligence ne doivent pas être rejetés sous prétexte qu'ils ne mesurent pas l'intelligence, même si cette dernière n'est pas clairement définie. Les tests mesurent des comportements cognitifs susceptibles de faire appel à cette compétence générale, l'intelligence, relativement à notre contexte socioculturel.

Diagnostic et pronostic

A des fins pratiques, les tests d'intelligence sont utilisés comme des instruments de diagnostic ou de pronostic. Ils pourront être considérés comme valides s'ils contribuent à de bons diagnostics ou pronostics. Dans ce sens, les tests permettent une meilleure compréhension de l'origine de certaines difficultés rencontrées par une personne. En outre, ce pronostic sera d'autant plus utile

Bibliographie

- Anastasi, A., & Urbina, S. (1997). *Psychological testing* (7th ed.). New Jersey : Prentice Hall.
- Carroll, J.B. (1993). *Human cognitive abilities*. Cambridge : Cambridge University Press.
- Grégoire, J. (2000). *L'évaluation clinique de l'intelligence de l'enfant. Théorie et pratique du WISC-III*. Sprimont (Belgique) : Mardaga.
- Huteau, M., & Lautrey, J. (1999). *Évaluer l'intelligence. Psychométrie cognitive*. Paris : Presses Universitaires de France.

Zusammenfassung

Psychodiagnostische Verfahren zur Messung menschlicher Intelligenz gehören zu den umstrittensten Methoden. Der französisch-schweizerische Psychologe Thierry Lecerf von der Universität Genf zeichnet die Streitpunkte nach und skizziert die grundlegenden Probleme bei der Intelligenzmessung, stellt aber auch die Beiträge der Intelligenztests bei Diagnose und Prognose dar.

qu'il prend place dans un modèle du fonctionnement cognitif. Or, la validité prédictive des tests d'intelligence s'est révélée excellente. En effet, contrairement à ce que l'on peut entendre dans la rue, le QI est un des meilleurs prédicteurs de l'apprentissage, de la réussite scolaire ou professionnelle. On peut néanmoins se demander si la réussite scolaire est un bon critère d'intelligence. Quoi qu'il en soit, la relation entre réussite scolaire et QI signifie que les capacités cognitives mesurées dans les tests d'intelligence sont en partie utiles pour réussir dans les apprentissages scolaires. Par exemple, les corrélations sont de l'ordre de .55 à .65 entre le WISC-III et divers tests d'acquis scolaires (lecture, mathématiques). Ces niveaux de corrélation permettent d'accepter la capacité du WISC-III à mesurer l'intelligence, et son intérêt pour évaluer par exemple les enfants en difficulté scolaire.

Trois ou quatre facteurs

L'intérêt clinique du WISC-III découle également des études de validité conceptuelle qui visent à identifier les compétences mises en jeu dans les subtests. Les données américaines montrent que la meilleure solution comprend 4 facteurs : Compréhension ver-

bale, Organisation perceptive, Attention/Concentration et Vitesse de traitement.

En revanche, les données françaises montrent qu'une solution comprenant trois facteurs est plus adéquate : Compréhension verbale (Information, Similitudes, Vocabulaire, Compréhension), Organisation perceptive (Complètement d'images, Arrangement d'images, Cubes, Assemblage d'objets) et Vitesse de traitement (Code, Symboles). Ainsi, l'administration du WISC-III permet non seulement d'obtenir trois QI (QIV, QIP, QIT), mais également trois indices factoriels (ICV, IOP, IVT). Ces indices semblent plus intéressants pour le clinicien que les QI, car il s'agit d'indicateurs plus homogènes. La compréhension verbale est un bon indicateur de l'intelligence verbale. On sait que des enfants issus de milieux socio-économiques défavorisés ou des enfants dyslexiques peuvent présenter de faibles performances sur cet indice. Le facteur d'Organisation perceptive est un bon indice de l'intelligence visuo-spatiale, englobant par exemple la mémoire visuelle à long terme (CI). L'intérêt du facteur Vitesse de traitement est plus discutable.

WISC-III et hyperactivité ?

Sur le plan clinique, on montre que des enfants dyslexiques ou souffrant de troubles attentionnels peuvent présenter de faibles performances dans l'épreuve Code. En revanche, l'utilité clinique du subtest Symboles n'est pas encore établie. Il impliquerait des aspects de concentration et de contrôle visuo-moteur. Les indices ACID (Arith., Code, Info., Mém.) et SCAD (Symb., etc.) semblent permettre une bonne identification des enfants dyslexiques. En ce qui concerne les troubles du déficit d'attention avec hyperactivité, on a également tenté de mettre en évi-

dence un profil spécifique dans le WISC-III. Ainsi, quelques auteurs américains utilisent le facteur Attention/Concentration comme indicateur. Pourtant, on a observé à plusieurs reprises que des enfants souffrant de troubles attentionnels ne présentent pas de faibles performances à cet indice. Cela remet donc en question l'utilité clinique de l'indice Attention/Concentration dans le diagnostic des troubles attentionnels et de l'hyperactivité. Cette remise en question est d'autant plus vraie pour les francophones que le facteur Attention/Concentration n'a pas été observé. Il est donc recommandé la plus grande prudence dans l'interprétation du facteur Vitesse de traitement (Grégoire, 2000). De faibles performances pourraient avoir d'autres origines (par exemple, problème de contrôle graphomoteur). Le psychologue souhaitant tester l'hypothèse de déficits attentionnels devrait donc utiliser d'autres tests (Stroop, appariement d'images, etc.). Enfin, soulignons que ces facteurs s'intègrent parfaitement dans le modèle de l'intelligence de Carroll, dans lequel on trouve un facteur général et des facteurs plus spécifiques. Tous les subtests du WISC-III évaluent à la fois une même capacité et des compétences spécifiques, justifiant l'utilisation du QI Total ainsi que des autres QI ou indices. Toutefois, rien ne garantit que cette capacité générale soit bien l'Intelligence.

Plutôt décevant

Désirant remplacer les échelles de Wechsler, jugées faibles sur le plan théorique, Kaufmann a créé le K-ABC. Pour cela, il s'est inspiré des théories plus récentes de l'intelligence et du modèle neuropsychologique de Luria (1973). Le K-ABC permettrait d'évaluer les processus Simultanés,

L'auteur

Thierry Lecerf a effectué ses études de psychologie à l'Université Pierre Mendès France de Grenoble et son doctorat à l'Université de Genève. Il est maître d'enseignement et de recherche suppléant à la Faculté de psychologie et des sciences de l'éducation (FPSE) de Genève, où il enseigne actuellement la psychologie différentielle.

Adresse

Thierry Lecerf, FPSE, 40, Bd du Pont-d'Arve, 1205 Genève. E-mail : thierry.lecerf@pse.unige.ch

Séquentiels, Composites et les Connaissances. L'échelle Simultanée apprécierait les capacités d'intégration et de synthèse et l'échelle Séquentielle serait un indice des processus impliqués lorsque les informations sont traitées successivement. L'échelle Composite englobe les processus Simultanés et Séquentiels. Enfin, l'échelle des Connaissances mesurerait l'intelligence cristallisée.

Sur le plan clinique, Kaufmann a constaté que son test évalue des compétences similaires à celles du WISC-III ($r=.70$) et a observé des corrélations de l'ordre de .30 à .50 entre le K-ABC et divers tests d'acquisitions scolaires. Globalement, ces corrélations sont inférieures à celles qui sont observées avec le WISC-III. La distinction originale Simultané/Séquentiel doit être considérée avec prudence, car elle fait l'objet de nombreuses critiques théoriques. Le K-ABC, construit sur une base théorique, est jugé comme relativement décevant et n'a pas révolutionné le diagnostic des troubles cognitifs, comme attendu. Les capacités évaluées par le K-ABC correspondent à celles qui le sont également par le WISC-III, et les qualités métrologiques du K-ABC sont inférieures à celles du WISC-III. Ainsi, les facteurs évalués par le K-ABC sont très proches des facteurs observés au WISC-III. Les processus Simultanés correspondent au facteur d'Organisation perceptive ; le facteur Connaissances est proche du facteur de la Compréhension verbale ; et les processus Séquentiels correspondent au facteur Attention/Concentration (version américaine). En résumé, le choix du K-ABC ne devrait pas se faire dans l'optique d'apprécier les processus Simultané et Séquentiel, puisque ces mêmes compétences sont évaluées dans le WISC-III.

Utilité des tests

L'utilité clinique d'un test peut reposer sur d'autres aspects que la validité prédictive. Par exemple, dans la nouvelle version du WAIS, l'étendue des scores possibles est élargie vers le bas, et l'importance accordée aux performances

chronométrées est moindre. En outre, l'introduction de l'évaluation de l'apprentissage incident rend le WAIS-III plus adapté au champ de la neuropsychologie. Le WAIS-III semble également présenter un intérêt clinique pour des personnes manifestant un retard mental, des troubles psychiatriques, une précocité intellectuelle (ou WISC-III), etc. Dans ce dernier cas, aucune mesure spécifique ne permet une aussi bonne prédiction que les tests d'intelligence. Enfin, les bases théoriques sont renforcées, puisque le WAIS-III permet d'évaluer trois QI et quatre indices factoriels. On mesure la Compréhension verbale (Info., Sim., Voc.), l'Organisation perceptive (CI, Cub., Matrices), la Mémoire de travail (Arith., Mém. et Séquence Lettres-Chiffres) et la Vitesse de traitement (Code, Symb.). Les données recueillies à ce jour semblent démontrer l'intérêt de ces indices dans le diagnostic de divers troubles cognitifs.

Comme énoncé précédemment, le concept d'intelligence est défini en référence à une époque, un contexte socioculturel. C'est la raison pour laquelle les items peuvent devenir obsolètes avec le temps et doivent être réanalysés régulièrement. Cela concerne plus particulièrement les subtests verbaux. Dans la première version du WISC, on posait par exemple la question suivante dans le subtest Information : « Qui a inventé la T.S.F ? » Il est probable que peu d'enfants (voire d'adultes) pourraient de nos jours répondre à cette question. On voit ici la nécessité non seulement d'actualiser régulièrement le matériel sous peine d'administrer des items désuets, mais également que le psychologue utilise la version la plus récente d'un test.

Réactualiser les étalonnages

Cependant, l'élément le plus important de la réactualisation des tests concerne les étalonnages disponibles. L'utilité d'un test réside dans sa capacité à mettre en évidence les compétences et les difficultés d'une personne, relativement à un groupe de référence. En ce qui concerne l'intelligence, on a obser-

vé une évolution régulière des QI, baptisée « effet Flynn ». Flynn (1987) a analysé 73 études réalisées entre 1932 et 1978 et a observé un gain moyen de 3 points de QI par décennie. Cette croissance a été observée pour plusieurs nations et pour différents groupes d'âges. L'existence de cet « effet Flynn » a des conséquences importantes sur l'évaluation d'une personne et justifie les nécessaires réactualisations des étalonnages. Utiliser de nos jours le WISC-R ou le WAIS-R conduit à une surestimation des performances, qui ne témoigne pas des véritables compétences d'une personne. Par exemple, un enfant obtenant 100 au WISC-R, aussi bien pour le QIV, le QIP ou le QIT, présenterait probablement au WISC-III un QIV compris entre 97 et 100, un QIP compris entre 93 et 96 et un QIT compris entre 94 et 97.

L'inné et l'acquis

Un reproche fréquemment adressé aux tests d'intelligence est de perpétuer la domination de certains individus et, par conséquent, de maintenir les injustices socioculturelles. Ainsi, dans le débat inné-acquis sur l'origine des différences individuelles, les tests d'intelligence sont accusés de mettre en évidence une compétence innée. Le darwinisme social considère alors qu'il est inutile de chercher à réformer la société. Cette position idéologique repose sur deux postulats : l'intelligence est unidimensionnelle et les aspects génétiques sont centraux.

Depuis les premiers travaux de Galton (1884), de nombreuses recherches quantifient la part respective de l'inné et de l'acquis et estiment l'héritabilité de l'intelligence à environ 60%. L'héritabilité représente un indice statistique qui évalue l'importance de l'action des gènes. Une héritabilité d'environ 60% signifie que plus de la moitié de la variance observée dans les tests peut s'expliquer par des différences génétiques entre individus. S'il est raisonnable de penser que les facteurs héréditaires jouent un rôle, il est probable que leur part relative est sur-

estimée en raison notamment d'une incapacité à tenir compte des interactions entre facteurs génétiques et environnementaux. Mais, quelle que soit la part des déterminants héréditaires dans l'intelligence, les résultats d'un test administré à un individu ne peuvent en aucun cas être considérés comme les indicateurs d'aptitudes innées. Cette lecture innéiste ne repose sur aucun fondement et doit être combattue par les psychologues. Il est de leur devoir d'expliquer au *quidam* que les tests d'intelligence ne témoignent pas du caractère inné de l'intelligence. En d'autres termes, les performances d'une personne testée traduisent ses compétences modelées par son éducation et ses apprentissages au moment de l'administration du test, c'est-à-dire ses compétences acquises. Les différences d'intelligence observées à un test reflètent donc en partie les inégalités du milieu. Le psychologue qui observe de faibles performances à un test peut alors proposer des méthodes pédagogiques compensatoires et des procédures de remédiation.

En résumé, les tests ne sont pas des instruments d'injustice sociale mesurant des compétences innées et destinées à perpétuer une domination d'un groupe social sur un autre. Il est indispensable que les psychologues diffusent l'information selon laquelle les tests d'intelligence ne mesurent pas une aptitude innée et stable, mais des compétences acquises. Observer des différences individuelles sur le plan cognitif est inévitable : il faut dès lors utiliser cette information pour définir un cadre pédagogique compensatoire, et non pour exclure une personne.

Les autres facteurs

En conclusion, si nous savons mesurer une dimension nommée « Intelligence », nous ne savons pas exactement ce qui est mesuré. Les résultats empiriques montrent que, quelle que soit cette dimension, elle permet de bonnes pré-



Photo: Annette Rutsch

dictions pour les réussites académique et professionnelle, si importantes dans notre société. Les tests d'intelligence sont valides, puisqu'ils permettent de formuler de relativement bons diagnostics ou pronostics. Enfin, les tests d'intelligence constituent une situation d'observation clinique très riche. Ils fournissent non seulement des indications sur l'intelligence de la personne, mais aussi sur sa personnalité, par exemple la gestion de ses émotions, sa motivation, etc. Les tests de QI sont de bons instruments pour se faire une première idée du niveau et du mode de fonctionnement d'un individu. Mais ils ne représentent qu'un des moyens permettant l'évaluation du fonctionnement d'une personne et leur utilisation doit s'inscrire dans une démarche générale

dans laquelle d'autres aspects doivent être intégrés (entretien clinique). Le psychologue doit toujours prendre en compte les facteurs non intellectuels et les informations de l'histoire du sujet (sociale, médicale, culturelle, etc.) lorsqu'il interprète les résultats. Les tests d'intelligence ne valent que par l'utilisation que le psychologue en fait. Une connaissance des qualités et des limites de chaque test est nécessaire pour interpréter les résultats de manière adéquate et ne pas leur faire dire plus qu'ils ne le peuvent.