



مجلة الآداب والعلوم الاجتماعية
ISSN : 2352-9849 ISBN : 2013-4803
المجلد الثاني عشر عدد خاص



Bilan psychologique, la conception des outils de mesure et leur utilité

Georges Cognet

Pr Honoraire Ecole de psychologues praticiens de Paris

تاريخ التقييم: 2020/07/15

تاريخ الإرسال: 2020/ 02 /27

Le concept même d'intelligence fait l'objet, depuis toujours serait-on tenté d'écrire, de critiques culturelles et sociales. Le QI est souvent remis en cause, parfois de façon virulente, et à travers lui « les tests sont accusés de légitimer et perpétuer les inégalités sociales et de ne pas apporter d'informations réellement intéressantes pour le diagnostic clinique. » (Grégoire, 2010).

La notion même d'intelligence fait aussi l'objet de critiques scientifiques. Ainsi, comme noté dans l'argumentaire de ce colloque, « les scientifiques ne sont pas parvenus à un consensus concernant la définition de l'intelligence [...] ». Effectivement, il n'y a jamais eue d'unanimité sur une définition précise de l'intelligence, si ce n'est une conception large, voire un peu vague, qui ramène l'intelligence à la capacité d'adaptation à des situations nouvelles, c'est-à-dire la capacité à comprendre, mais aussi à apprendre afin d'ajuster l'être au monde.

Georges Cognet

En 1997, cinquante-deux universitaires, scientifiques(dont Carroll, Cattell, Horn, Kaufman etVernon) publient une tribune (Gottfredson, 1997) afin de faire le point sur les connaissances partagées sur l'intelligence. Là encore, la définition produite par ces spécialistes mondialement reconnus, pour être consensuelle, n'en est pas moins généraliste, peu efficiente sur le plan scientifique. L'intelligence, écrivent-ils« est une capacité mentale très générale qui implique, entre autres,la capacité à raisonner, planifier, résoudre des problèmes,penser de manière abstraite, comprendre des idées complexes,apprendre rapidement et apprendre de l'expérience. » Paradoxalement, ils affirment qu'il existe, malgré l'impossibilité d'une définition consensuelle opérante et malgré la méconnaissance des processus cérébraux sous-jacents, la possibilité d'une mesure de l'intelligence, voire d'une mesure de qualité et, plus étonnant encore, ils écrivent que « bien qu'il existe différents types de tests d'intelligence, ils mesurent tous la même intelligence. »

Déjà, en 1944, David Wechsler l'indiquait clairement : « Lors de l'examen d'un très grand nombre d'individus, en utilisant des tests variés d'intelligence, on remarque que ceux qui ont obtenu une note élevée sur l'un de ces tests ont tendance à obtenir une bonne note sur les autres ; il en est de même pour ceux qui obtiennent une note basse ou intermédiaire. Depuis longtemps, les psychologues ont reconnu cette double caractéristique des aptitudes humaines : leur spécificité d'une part, et leur interdépendance d'autre part. »

Épreuves	Corrélations
WISC-IV / WISC-V	.81
WISC-V / WPPSI-IV	.82
WISC-IV / KABC-II	.78
NEMI-2 / WISC-III	.80

Tableau 1 : Corrélations entre différents tests d'intelligence

Les résultats présentés dans ce tableau montrent une corrélation très élevée et stable entre les différents tests d'intelligence adaptés en France.

,n

A. Les tests d'intelligence et leur implication

L'intelligence demeure un objet très investi par la recherche sur plusieurs aspects, citons en trois : l'exploration des processus neurologiques sous-jacents, son développement chez l'humain, de la naissance à sa mort, ou encore l'intérêt pour les différences interindividuelles. Cette dernière perspective est celle d'Alfred Binet, de David Wechsler et, toujours actuellement, des concepteurs de tests d'intelligence en lien avec les besoins de la société.

1. La nécessité du Binet-Simon

L'échelle métrique d'intelligence de Binet advient à l'exact entrecroisement des idées novatrices de la psychologie (Watson, Janet) du début du Xxe siècle, des propres recherches et observations de l'auteur et des besoins de la société, particulièrement ceux de l'instruction publique. Ainsi, le contexte scientifique de l'époque est à l'abandon de la psychologie des états de conscience pour se consacrer celle de l'activité, motrice ou verbale (Huteau, 2006) ; dans les mêmes temps, les expérimentations de Binet le portent vers l'étude des différences individuelles et enfin se posent, depuis l'instauration de l'instruction publique obligatoire, les questions de l'éducation des enfants anormaux d'intelligence. Avant d'éduquer, faut-il encore repérer, détecter les enfants dits arriérés ; distinguer le plus finement possible les sujets dont le retard est lié à un environnement ou à une scolarité défavorables de ceux qui présentent un déficit d'intelligence. La mise en place en 1904 en France d'une commission ministérielle pour l'enseignement spécial, ayant pour objet l'examen de deux questions importantes touchant à l'instruction de masse : le diagnostic des états d'arriération mentale et l'éducation des enfants anormaux, donna à Binet l'occasion de mettre en application concrète ses conceptions de la mesure des différences individuelles. « Il n'y a rien de tel que la nécessité pour faire surgir des méthodes nouvelles », écrit-il vers la fin de sa vie. Six mois plus tard, en 1905, dans *l'Année*

psychologique il publie sa méthode pour le diagnostic intellectuel des anormaux.

a) L'échelle métrique de 1905

L'enjeu social ainsi que l'intérêt de l'enfant sont au premier plan de cette échelle. Une évaluation objective de l'arriération, ne s'appuyant pas exclusivement sur un diagnostic médical ou pédagogique, est indispensable avec l'ouverture prochaine des écoles de perfectionnement.

La version de 1905 comporte ainsi trente items, il ne s'agit cependant pas à proprement parler d'une échelle d'intelligence avec des niveaux d'âges, mais un outil qui vise le diagnostic rapide de l'arriération : « Notre but est, lorsqu'un enfant sera mis en notre présence, de faire la mesure de ses capacités intellectuelles, afin de savoir s'il est normal ou si c'est un arriéré. » (Binet & Simon, 1905 : p. 191).

L'échelle ne comporte aucune indication d'âge, mais les auteurs proposent, à travers le choix des items « un classement, une hiérarchie entre les intelligences diverses » (Binet & Simon, 1905) en partant des niveaux les plus bas. Ainsi l'item 1, intitulé : Le regard ; mouvement de la tête ou des yeux pour suivre le déplacement lent d'une allumette enflammée, teste-t-il la poursuite oculaire dont on sait qu'il s'agit d'une compétence présente à la naissance et qui est accompagnée d'un mouvement de tête dès l'âge de 2 mois. Il s'agit là d'une forme de réflexe qui permet une réponse du sujet aux sollicitations extérieures et que Piaget inclut dans le stade de l'intelligence sensorimotrice « au sens de réaction totale de l'individu, et non pas seulement mise en jeu d'automatismes particuliers ou locaux reliés entre eux du dedans seulement » (Piaget, 1936).

À l'autre extrémité du test, Binet et Simon proposent aux sujets de comparer et définir quelques termes tels qu'estime et amitié, ennui et chagrin, etc. Il s'agit ici de mots abstraits qui pour Binet établissent la frontière très nette entre les états d'arriération, d'imbécilité, et d'intelligence des enfants normaux.

b) L'échelle métrique de 1908

Dans l'article de l'année psychologique datée de 1908, intitulé *Le développement de l'intelligence chez les enfants*, Binet et Simon exposent de façon très détaillée leur échelle de mesure de l'intelligence infantine. Comme toujours, leur écrit dépasse la simple

technique pour envisager aussi « la loi du développement intellectuel des enfants » et s'intéresser aux aspects pédagogiques à la lumière de leurs recherches : « En nous occupant à tracer la ligne d'évolution de l'intelligence chez l'enfant, nous avons été poussés tout naturellement à jeter un coup d'œil sur les programmes d'enseignement, et à constater que certains de ces enseignements sont trop précoces, c'est-à-dire mal adaptés à la réceptivité mentale des jeunes. » (Binet et Simon, 1908, p. 2).

Il y a chez les auteurs un réel souci de standardisation : les consignes sont clairement explicitées, des exemples de bonnes et mauvaises réponses sont aussi proposés afin de faciliter la correction et la cotation. De plus, une note de bas de page donne des indications concernant l'étalonnage de référence (lieu de recueil, catégorie socioprofessionnelle des sujets).

Ainsi, les préalables méthodologiques énoncés en introduction, les recommandations aux utilisateurs, en d'autres termes la standardisation, font entrer l'échelle métrique de 1908 dans la modernité de ce que seront les tests psychologiques à venir.

c) L'échelle métrique de 1911

L'échelle métrique de 1911 est une amélioration nette sur les plans formels et théoriques de la précédente datant de 1908 : « on a fait à notre échelle quelques objections qui nous ont paru justes ; nous-mêmes, en l'employant, nous avons constaté ses défauts, et nous avons cherché à les réparer. » (Binet, 1911 : p. 146). Suivent un certain nombre de corrections, de suppressions d'épreuves faisant double emploi ou trop exclusivement scolaires et une nouvelle hiérarchisation des items.

De plus, cette échelle de 1911 inclut, par rapport à celle de 1908, des séries d'items pour les 15 ans et pour les adultes. Cette évolution vers l'évaluation de sujets plus âgés et, par conséquent, l'introduction de situations tests plus complexes est le signe tangible d'un changement d'objectif. Il n'est plus simplement question de dépister et de caractériser des niveaux de déficience, mais d'évaluer l'intelligence chez des sujets tout-venant.

Cependant, la forme définitive du Binet-Simon est plus due au décès brutal de son auteur, en 1911, qu'à un accomplissement. Il n'est pas

dans notre intention de réécrire l'histoire, mais plutôt de faire quelques conjectures : Binet, vivant, aurait peut-être eu le temps nécessaire, avant la première Grande Guerre mondiale, de proposer une quatrième version de son échelle métrique. Il aurait, bien entendu, réagi au calcul de W. Stern (1912) qui conduit à la notion de quotient intellectuel. Il aurait aussi, c'est certain, condamné les errements menés aux États-Unis par H., H. Goddard qui à partir de 1912, traque les « faibles d'esprit » parmi les immigrants débarqués à Ellis Island (voir plus loin).

Georges Cognet

TESTS BINET-SIMON	
Feuille de niveau (1)	
NOM ET PRÉNOMS	
Date de naissance	Date de l'examen
Age à la date de l'examen	Age d'intelligence
3 ans * Donner nom de famille. * Répéter 2 chiffres. * Enumérer une gravure. * Répéter phrase 6 syllabes. Montrer nez, œil, bouche.	9 ans * Reconnaître pièces de monnaie. * Rendre sur 1 fr. * Définir supérieurement à usage. * Questions faciles. Enumérer les mois.
4 ans * Répéter 3 chiffres. Donner son sexe. Nommer clé, couteau, sou. Comparer 2 lignes.	10 ans * Questions difficiles. Ordonner 5 poids. Deux dessins de mémoire. Critiquer phrases absurdes. 3 mots en 2 phrases.
5 ans * Répéter phrase 10 syllabes. * Compter 4 sous simples. Comparer 2 poids. Copier un carré. Jeu de patience.	12 ans 3 mots en une phrase. Résister à suggestion de lignes. Plus de 60 mots en 3 minutes. Définir mots abstraits. Deviner le sens de phr. en désordre.
6 ans * Compter 13 sous simples. * Définir par l'usage. Distinguer matin et soir. Copier un losange. Comparaisons esthétiques.	15 ans * Répéter 7 chiffres. * Répéter phrase de 26 syllabes. * Interpréter une gravure. Trouver 3 rimes. Problèmes de faits divers.
7 ans * Décrire une gravure. * 9 sous, dont 3 doubles. Main droite, oreille gauche. Nommer 4 couleurs. Exécuter 3 commissions.	Adultes Epreuve de découpage. Reconstruire un triangle. Différences de mots abstraits. Roi et Président. Pensée d'Hervieu.
8 ans * Répéter 5 chiffres. 2 objets de souvenir. Compter de 20 à 0. Lacunes de figures. Date du jour.	

(1) Les épreuves marquées d'un astérisque sont celles de la première feuille d'interrogatoire.

Figure 1 :Le Binet Simon de 1911

Les premiers utilisateurs, formés par Binet lui-même, sont des instituteurs, des directeurs d'école situés dans des quartiers pauvres de Paris. Puis, très vite, quelques autres utilisateurs, tous pédagogues, formés par la lecture de l'article de 1911, réalisent quelques passations dans d'autres milieux sociaux. Les premiers résultats montrent une influence du niveau social et économique sur les résultats aux tests ; « suivant la richesse de la population, le niveau intellectuel des enfants

se modifie » (Binet, 1911). D'autres expérimentations s'intéressent aux corrélations entre le niveau d'intelligence et le niveau scolaire : « Il existe une corrélation remarquable entre ces deux niveaux ; et nous pouvons l'exprimer de la manière suivante : quand les enfants sont en retard dans leurs études, ils ont une chance pour avoir une intelligence moyenne, contre deux chances pour avoir une intelligence inférieure à la moyenne, et aucune chance pour être brillants ; s'ils sont réguliers dans leurs études, ils ont une chance pour être brillants d'intelligence, une chance pour être lourds d'intelligence, et deux chances pour être moyens d'intelligence ; s'ils sont en avance dans leurs études, ils ont autant de chances à peu près pour être réguliers ou brillants d'intelligence. » (Binet, 1911 : p. 162).

2. La NEMI-2, une rupture avec le test de Binet et la NEMI de Zazzo

À la fin du Xxème siècle, alors que l'examen psychologique — questionné dans son déroulement, ses objectifs, ses comptes-rendus par des praticiens et des universitaires — retrouve une légitimité et une vigueur nouvelle dans la clinique des psychologues, la NEMI de René Zazzo, qui date de 1966, apparaît comme une épreuve dépassée qui n'a pas été révisée ni n'a bénéficié d'aucun réétalonnage depuis plus de trente ans. Les étalonnages de 1949 ou de 1966 ne rendent plus compte du développement des enfants d'aujourd'hui. Les praticiens en délaissent donc l'usage au profit des échelles de Wechsler.

a) Modèle théorique et structure

La NEMI-2 est un test qui propose un ensemble composite d'épreuves qui toutes font appel aux *processus mentaux supérieurs* tels que les définissait Binet (1909).

En revanche, l'organisation de la NEMI-2 s'affranchit, et c'est une rupture avec les échelles de Binet et Zazzo, de la double contrainte liée à l'échelle d'âge qui nécessite, d'une part, de créer des items spécifiques d'un âge donné, c'est-à-dire des items échoués avant l'âge seuil puis largement réussis, celui-ci dépassé, et d'autre part de sélectionner, parmi ceux-ci, les meilleurs items qui évaluent l'intelligence. De plus, une échelle d'âge ne peut pas, par construction, proposer les mêmes items ou les mêmes types d'items à tous les

sujets, jeunes ou âgés. La NEMI-2 qui abandonne ce principe devient ainsi un test beaucoup plus homogène : ce sont les mêmes épreuves qui sont présentées à tous les niveaux d'âge (de 4 ans et demi à 12 ans et demi).

Le modèle de l'intelligence retenu pour cette révision est celui de Carroll (1993), qui réalise une synthèse entre les conceptions unitaires et multifactorielles de l'intelligence.

En prenant en compte une partie seulement de ce modèle, La NEMI-2 se compose de sept épreuves dont quatre sont obligatoires (*Connaissances, Comparaisons, Matrices analogiques et Vocabulaire*) et trois facultatives (*Adaptation sociale, Répétition de chiffres, Copie de figures géométriques* ou *Comptage de cubes*) laissées à l'appréciation du clinicien.

De nombreux items du Binet-Simon et de la NEMI, qui ont fait leurs preuves, ont été retenus pour l'élaboration de la NEMI-2.

b) L'expression des résultats

La NEMI-2 permet, comme sa devancière, le calcul d'une note globale. Cependant, le terme QI n'a pas été conservé puisque le résultat n'est plus « un véritable quotient d'âges [...mais] un indice qui traduit la distance, l'écart à la moyenne » (Zazzo, Dague, Schmelk, Rossi, 1975). Aussi, a été créé un *Indice d'efficacité cognitive* (IEC) issu de la somme des notes standard aux quatre épreuves obligatoires. Cet indice, basé sur la distribution des scores selon la loi normale (moyenne = 100, écart-type = 15), est exclusivement exprimé sous la forme d'un intervalle de confiance. Cette forme d'expression a l'avantage de relativiser la note obtenue en ne réduisant pas la performance de l'enfant à un seul nombre.

Malgré l'abandon de la perspective génétique dans la construction de l'échelle, la référence à l'âge a été maintenue pour chacune des épreuves sous la forme d'un tableau convertissant les notes obtenues en âges de développement.

c) De l'utilité de la NEMI-2

La NEMI-2 est particulièrement utilisée par les psychologues de l'Éducation nationale en France pour ses qualités de pronostic de la réussite scolaire.

Chaque année, les établissements scolaires proposent à l'ensemble des élèves des classes de CE2 et 6^e une évaluation diagnostique de début d'année. Cette procédure d'évaluation est coordonnée par la *Direction de l'évaluation et de la prospective* du ministère de l'Éducation nationale. Les protocoles d'évaluation constituent un outil de diagnostic individuel permettant, dès le début de l'année, d'identifier les acquis et les faiblesses de chacun des élèves. Une recherche a permis d'établir une corrélation entre les résultats à ces évaluations et le résultat global obtenu à la NEMI-2 pour deux groupes d'élèves.

Évaluations nationales (n= 96)	Corrélation
Mathématiques	0.65
Français	0.61
Mathématiques et français	0.68

Tableau 2 : Corrélation entre les performances scolaires mesurées aux Évaluations nationales de CE2 et 6^e et l'IEC de la NEMI-2

On observe des corrélations élevées entre l'IEC de la NEMI-2, issu des notes aux quatre épreuves obligatoires (*Connaissances, Comparaisons, Matrices analogiques, Vocabulaire*) et les résultats aux Évaluations nationales de CE2 et 6^e. La NEMI-2 semble ainsi être un bon prédicteur de la réussite scolaire en primaire et dans les premières années du collège.

La NEMI-2 apparaît aussi pertinente pour mettre en évidence les aptitudes visuospatiales et, le cas échéant, leur déficit, grâce à l'épreuve *Comptage de cubes*, et pour évaluer les compétences graphomotrices par l'épreuve de *Copies de figure géométrique*.

Une étude spécifique a été conduite auprès d'enfants présentant un trouble persistant de la coordination motrice globale non explicable par un retard mental ou une affection somatique identifiée. Les performances aux épreuves *Comptage de cubes* et *Copies de figures géométriques* sont nettement au-dessous du niveau escompté compte tenu de l'âge des sujets.

3. L'évolution des échelles de Wechsler

David Wechsler, après Alfred Binet, fut le plus illustre représentant de l'évaluation de l'intelligence. Pour lui, l'intelligence est conçue

comme une qualité émergente : « L'intelligence est la capacité globale ou complexe de l'individu d'agir dans un but déterminé, de penser d'une manière rationnelle et d'avoir des rapports utiles avec son milieu. Elle est globale parce qu'elle caractérise le comportement de l'individu dans son ensemble ; elle est complexe parce qu'elle est composée d'éléments, d'attitudes qui, sans être entièrement indépendants, sont qualitativement différenciables. En fin de compte, c'est par la mesure de ses aptitudes que nous évaluons l'intelligence. » (Wechsler, 1944)

Comme Binet, Wechsler apparaît comme un précurseur créatif voire pour certains un génie avec, par exemple, l'invention d'une note globale, nommée *QI* par Stern, issue de la distribution normale des scores, ou encore l'introduction de tâches nonverbales dans la mesure globale de l'intelligence et aussi par ce souhait de mesurer l'intelligence, sur un même modèle, tout au long de la vie des individus avec une panoplie d'outils bien connus : *WPPSI* pour les plus jeunes, *WISC* pour les enfants au-delà de 6 ans et les adolescents et *WAIS* pour les jeunes à partir de 16 ans et les adultes.

a) La rupture du WISC-IV

La rupture est intervenue avec le WISC-IV. Cette version de l'échelle d'intelligence de Wechsler pour enfants et adolescents apporte alors un changement en profondeur avec les conceptions anciennes de la mesure de l'intelligence par l'abandon des traditionnels *QI verbal* et *QI performance* qui structuraient les échelles de Wechsler, jusqu'au WISC-III, et en introduisant des indices factoriels inspirés de la théorie de l'intelligence dite CHC (Carroll, Horn, Cattell).

Par rapport à ses devancières, l'échelle WISC-IV représente une réelle avancée par son organisation en quatre domaines : la *compréhension verbale*, le *raisonnement perceptif*, la *mémoire de travail* et la *vitesse de traitement*, qui évaluent des fonctions cognitives qui « interagissent de manière très concrète », « fonctionnent ensemble et sont le plus souvent citées dans la littérature comme nécessaires aux apprentissages. » (Wechsler, 2005). À chacun des domaines cités, correspond un indice, comparable dans sa construction à un *QI* : *ICV* (*Indice de compréhension verbale*), *IRP* (*Indice de raisonnement perceptif*), *IMT* (*Indice de mémoire de travail*) et *IVT* (*Indice de vitesse de traitement*). Cette nouvelle organisation de l'échelle de

Wechsler enrichit l'évaluation en distinguant et en comparant des domaines essentiels du fonctionnement cognitif. Les psychologues de l'enfance et de l'adolescence ont apprécié l'approche analytique en quatre grands domaines que propose le WISC-IV. L'examen psychologique, grâce à cette échelle, ne se limite plus au calcul du QI total et à l'interprétation des réussites et des échecs aux subtests, mais permet, évolution majeure en 2005, le calcul de quatre indices fondamentaux et l'établissement d'un profil cognitif.

Cependant, à l'usage, les chercheurs et les praticiens ont pointé certaines limites propres à cet outil, nous en retenons particulièrement trois : l'hétérogénéité de l'échelle de *Raisonnement perceptif*, l'évaluation de la mémoire de travail qui demeure essentiellement sur un mode auditivoverbal et la perte de significativité du QI.

b) L'aboutissement représenté par le WISC-V

Si le WISC-IV a représenté une véritable rupture conceptuelle et pratique, le WISC-V se positionne comme une forme d'aboutissement de l'évolution des échelles de Wechsler pour enfants et adolescents. Certes, l'ensemble du modèle CHC (voir modèle plus loin) n'est pas repris dans cette échelle, mais son organisation même, au plus près du modèle théorique de référence, la proposition de subtests adaptés à la mesure de capacités précises et la redéfinition du QI total, qui devient un indice, donnent un outil extrêmement abouti. Ainsi, avec une certaine économie de moyens - seuls dix subtests sont nécessaires pour établir le profil cognitif et sept pour calculer le QIT - le WISC-V réussit à proposer aux praticiens une approche globale de l'intelligence et des principales fonctions cognitives tout en préservant les observations cliniques.

Création de deux nouveaux indices

Afin de répondre à l'hétérogénéité de l'indice de raisonnement perceptif, le WISC-V propose deux nouveaux des indices très spécifiques : *l'Indice visuospatial* et *l'Indice de raisonnement fluide*. Chacun étant dédié à l'évaluation de fonctions précises qui auparavant étaient confondues dans une seule échelle.

Ainsi, *l'Indice visuospatial* est composé de l'épreuve des *Cubes*, subtest traditionnel des épreuves de Wechsler, auquel vient s'adjoindre une nouvelle épreuve, *Puzzles visuels*. Ces deux subtests sont très

différents dans leurs modalités d'administration, le premier propose la reconstruction en trois dimensions de modèles présentés sur une image et le second demande aux sujets de recomposer virtuellement, sans manipulation et sur un plan en deux dimensions, un puzzle géométrique à partir de trois images. L'un des subtests est très concret l'autre beaucoup plus abstrait, mais les deux mettent en jeu le facteur visualisation spatiale (Gv) et la rotation mentale.

Le second indice nouveau est *l'Indice de raisonnement fluide*. Il regroupe le subtest classique de *Matrices* et le subtest *Balances*, déjà connu des utilisateurs de la WAIS-IV. Cette échelle évalue le raisonnement logique à partir d'un support perceptif, sans solliciter des connaissances structurées préalablement, et par conséquent la capacité à donner du sens à une confusion d'éléments et à établir des systèmes de pensée non verbaux.

Le subtest *Matrices* est composé de deux types d'items : des matrices orthogonales, mais aussi des suites. Dans les deux cas, il est demandé au sujet de compléter logiquement la matrice ou la suite en pointant une réponse parmi plusieurs propositions. Le subtest *Balances* est un nouveau subtest qui consiste, à partir de la représentation de deux ou trois balances, de réaliser des équivalences de poids. Cette épreuve mobilise, comme celle de *Matrices*, la pensée inductive et déductive, la mémoire de travail, la concentration, mais aussi le raisonnement numérique et quantitatif.

Renouvellement de l'échelle de Mémoire de travail

Afin d'optimiser l'évaluation de la mémoire de travail, le WISC-V complète la traditionnelle épreuve de *Mémoire des chiffres* par une épreuve de *Mémoire des images*. Ainsi, *l'Indice de mémoire de travail* combine une évaluation auditivoverbale avec une autre visuomotrice. Le subtest *Séquence Lettres-Chiffres* devient supplémentaire.

Un QI rénové

Les échelles de Wechsler proposent une approche globale du sujet placé au centre de l'évaluation, qui prend en compte la cognition, mais aussi les facteurs non intellectifs. Dans les versions précédentes du WISC-V, le calcul du QI provenait de l'ensemble des subtests utilisés pour calculer les différents indices.

Dans le calcul du QI, qui est une somme équipondérée des performances, chaque épreuve pèse le même poids. L'épreuve de *Code* par exemple a le même poids que l'épreuve de *Similitudes*, alors que tous les praticiens savent bien que les deux ne sollicitent pas des processus de même niveau et qu'ils accorderont bien évidemment beaucoup plus d'importance à une épreuve de conceptualisation verbale plutôt qu'à une épreuve de rapidité de traitement et de données perceptives.

Avec le WISC-V, le mode de calcul du QI reste le même que dans les versions précédentes, mais la conception et le poids des différents domaines évoluent. Tous les subtests ne participent pas au calcul du QI, seuls sept d'entre eux sont retenus : *Similitudes*, *Vocabulaire*, *Cubes*, *Matrices*, *Balances*, *Mémoire des chiffres* et *Code*. La nature même du QI a changé, il ne s'agit plus de représenter l'échelle totale, mais grâce à une sélection raisonnée, d'apporter une information supplémentaire sous la forme d'un indice rendant compte d'une grande diversité d'aptitudes.

c) De l'utilité du WISC-V

Là encore avec le WISC-V, on note une très bonne corrélation entre le QIT et les résultats aux tests scolaires. Le facteur *g* joue un rôle majeur dans les apprentissages scolaires (Cammerrer, Maddocks, Keith et Reynold, 2018) et le WISC-V est en mesure de rendre compte de ce rôle éminent.

Épreuves	Corrélations
KTEA-3 Comp. écrit / WISC-V	.74
KTEA-3 mathématiques / WISC-V	.79

Tableau 3 : Corrélations entre les scores composites du KTEA-3 (*Kaufman Test of Educational Achievement*) et du WISC-V (Wechsler, 2014)

Les troubles d'apprentissage établissent un lien fort entre les différentes fonctions cognitives et leurs dysfonctionnements. C'est parce qu'une fonction cognitive, participant bien entendu aux opérations essentielles liées à l'information – versus encodage, traitement, production – vient à manquer, à se montrer défaillante, qu'apparaissent des incapacités à mettre en place, par exemple, des systèmes automatisés d'attention focalisée, d'analyse perceptive des

données, de communication par le langage ou encore d'accès à la lecture ou au calcul.

Ces troubles, pour qu'ils soient dits d'apprentissage, ne doivent pas affecter le fonctionnement cognitif global comme peut le faire une déficience intellectuelle, mais une ou plusieurs fonctions cognitives particulières. Ils sont dits spécifiques en rapport à leur étiologie qui n'est pas une déficience intellectuelle, par essence globale, ni une lésion neurologique, ni une pathologie psychique, ni un trouble du spectre autistique. Le retentissement de ces troubles sur les apprentissages s'avère fort, durable et résistant aux remédiations habituellement mises en œuvre dans le cadre scolaire. Ils affectent sévèrement le parcours scolaire des enfants qui en sont atteints. Ceux-ci, d'ailleurs, expriment souvent leur fatigue d'apprendre, voire leur épuisement, et la perte de la confiance consubstantielle au développement et à l'enfance. Le risque de décrochage demeure toujours présent.

Le WISC-V avec ses cinq échelles s'avère être un outil extrêmement pertinent, dans une première approche, pour repérer d'éventuels troubles d'apprentissages.

B. Le paradoxe des tests d'intelligence

1. Des fondements théoriques solides

Les grands modèles actuels de l'intelligence sont des modèles issus de l'analyse factorielle et de l'avancée des connaissances en psychologie cognitive et en neuropsychologie. Nous présentons succinctement les principaux modèles :

a) Le premier modèle de Cattell et Horn

Cattell et Horn reprennent la démarche de Spearman, mais effectuent une analyse factorielle dite de second ordre entre les facteurs déjà extraits.

Ils mettent en évidence non pas un facteur général, mais deux facteurs qu'ils nomment :

- Gc, intelligence cristallisée qui sature principalement les tests verbaux et les tests de connaissance. Gc est plus efficace si la base de connaissance du sujet est développée et structurée. D'ailleurs, le

Georges Cognet

terme « cristallisé » renvoie à l'organisation sous-jacente des connaissances mises en mémoire à long terme. Gc est très corrélée à la réussite scolaire.

- Gf, intelligence fluide, qui sature des tests de raisonnement où les connaissances requises sont minimales. Gf représente le potentiel biologique de l'intelligence qui s'actualise dans Gc par l'expérience et avec l'acquisition incidente (non délibérée) de connaissances. Gf fait peu appel aux connaissances stockées en mémoire à long terme.

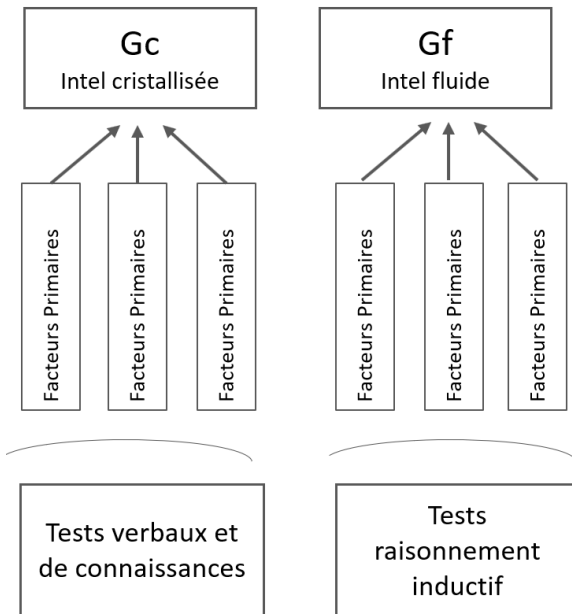


Figure 2 : le premier modèle Horn-Cattell

b) La structure hiérarchique à trois niveaux de J., B. Carroll, 1993

Par un travail gigantesque de compilation des données de 460 analyses factorielles de tests d'intelligence, John B. Carroll a mis en évidence, en 1993, une structure factorielle commune à toutes les données des tests d'intelligence.

Il produit ainsi une structure factorielle qui obtint, rapidement, un consensus. Cette structure hiérarchique se dispose sur trois niveaux :

Georges Cognet

- Le premier niveau (Stratum I) est composé de plus de 60 facteurs primaires qui correspondent peu ou prou aux différents subtests des épreuves d'intelligence. Certains sont corrélés entre eux et donnent naissance aux facteurs du 2^e niveau.
- Les facteurs de second ordre (Stratum II) représentent 8 grands domaines de l'intelligence (intelligence cristallisée, fluide, etc.) et sont aussi corrélés entre eux, ce qui indique l'existence d'un facteur général d'intelligence.
- Le troisième niveau (Stratum III) représente donc le facteur général d'intelligence.

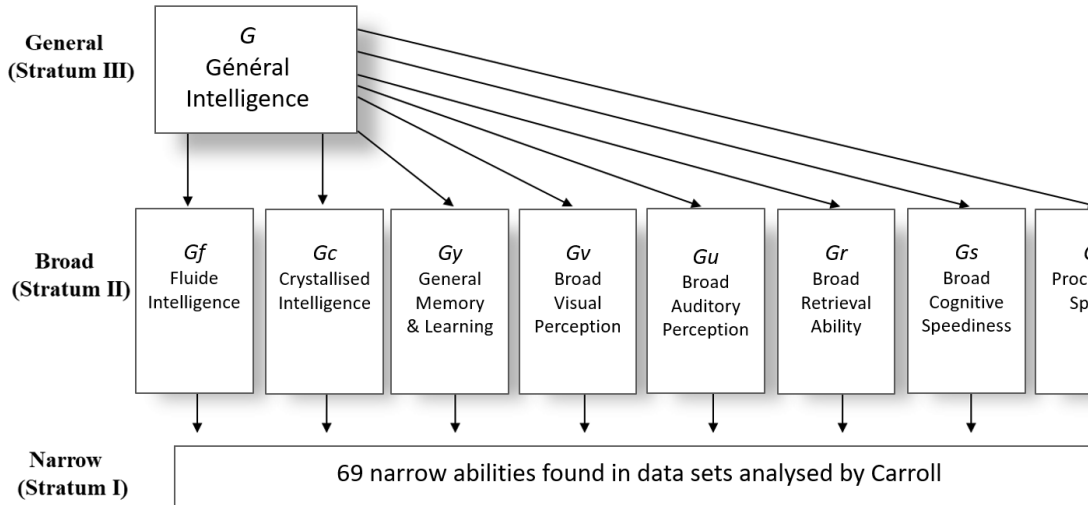


Figure 3 : Le modèle de Carroll (1993)

c) Le modèle CHC (Cattell, Horn, Carroll)

Le modèle dit CHC est une synthèse des travaux de John B. Carroll, de ceux de Raymond B. Cattell et de ceux de John L. Horn. Les modèles de l'intelligence de Cattell – Horn et de Carroll, bien que différents, montrent des points communs importants qui ont permis de les réunir dans un modèle synthétique nommé modèle CHC.

Ce modèle s'avère être un cadre théorique qui permet la construction d'outils évaluant les capacités cognitives humaines. Certaines

dénominations de facteurs de niveau deux sont différentes dans le modèle de Carroll et dans le modèle CHC.

Le modèle CHC s'organise sur 3 niveaux :

- le niveau 1 étant celui des facteurs dits primaires ou spécifiques, représentés par exemple dans le WISC-V par les subtests (*Matrices* et *Balances* pour le raisonnement fluide, *Similitudes*, *Vocabulaire*, *Compréhension* et *Information* pour l'intelligence cristallisée, etc.).
- Au niveau 2 se situent les grandes capacités cognitives ou domaines de l'intelligence. Pour le WISC-V, il s'agit de la *Compréhension verbale*, du *Raisonnement fluide*, des aptitudes *Visuospatiales*, de la *Mémoire de travail* et de la *Vitesse de traitement*.
- Au niveau 3 se situe le facteur *g*.

Avec l'évolution des méthodes d'analyse factorielle, le modèle CHC connaît une évolution permanente. C'est ce qu'indique très directement McGrew en 2009 : « Il est important que les chercheurs et les utilisateurs des tests d'intelligence ne soient pas aveuglés par le point de repère que représente le modèle CHC actuel. » De nombreux documents, régulièrement mis à jour, sont disponibles sur le site themindhub.com

2. Les contraintes de construction des tests d'intelligence

On observe cependant que le WISC-V n'évalue pas la totalité des domaines du deuxième niveau du modèle CHC. Pourquoi une telle restriction, pourquoi le modèle CHC, auquel le WISC-V fait référence, n'est pas intégralement repris ?

La réponse tient dans les contraintes de la construction d'un test d'intelligence qui réunit deux principes opposés : proposer une approche aussi complète que possible des processus cognitifs en jeu, des contenus, mais aussi tenir compte de la temporalité propre à l'examen psychologique qui « est une observation limitée dans le temps, brève, une méthode rapide pour connaître le sujet examiné » (Emmanuelli, 2004). On le sait, l'outil le plus complet qui proposerait une palette exhaustive de subtests et d'items, dont la passation durerait plusieurs heures, ne trouverait pas, quelles que soient ses qualités

scientifiques, sa place dans l'examen psychologique de l'enfant tel que le pratiquent les cliniciens. C'est cette double contrainte, à la fois une diversité des processus, des supports et une limitation dans le temps, qui structure, depuis le Binet-Simon, les tests d'intelligence. Les concepteurs doivent faire des choix et par conséquent exclure de leurs échelles certaines dimensions décrites dans les modèles du fonctionnement cognitif auxquels ils se réfèrent.

C. Pour une utilisation intelligente des tests d'intelligence

Plusieurs écueils guettent le praticien qui évalue l'intelligence d'un sujet, enfant, adolescent ou adulte. Nous évoquerons deux types d'écueils bien distincts. Le premier, évidemment le plus grave, consiste à se servir des tests d'intelligence à des fins qui ne sont pas compatibles avec les codes de déontologie des psychologues ni même avec le respect des droits humains, le second type d'écueil concerne les différents biais qui peuvent affecter le jugement du professionnel utilisant des outils de mesure.

1. La mal-mesure de l'homme

Rappelons-nous les heures noires de l'utilisation des tests, déplaçons-nous donc sur l'île d'Ellis Island, en 1912, alors qu'arrivent les bateaux transportant les émigrants européens harassés par un très long voyage dans l'entrepont. À ces Italiens, ces Hongrois, ces Russes qui n'avaient, pour la plupart, jamais tenu un crayon, H.H. Goddard (le premier traducteur du *Binet-Simon* aux États-Unis) et ses collaborateurs sélectionnent au faciès les « simples d'esprit » et soumettent ceux-ci à un rapide test d'efficacité. Il s'agit d'une figure montrée quelques secondes puis cachée à la vue ou des épreuves non verbales proposées collectivement. Les biais culturels sont criants. Il faut par exemple dessiner la trajectoire que prend la balle de tennis lors d'un échange, situer la boule sur la piste d'un bowling ou compléter la représentation d'une ampoule à incandescence à laquelle manque le filament. Les immigrants, fatigués, apeurés, n'ont évidemment jamais joué au tennis, ni au bowling et pour nombre d'entre eux, n'ont pas encore rencontré la *fée électricité*, mais ils

comprennent rapidement l'enjeu de la réussite ou non de ces tests : s'installer sur le sol du *Nouveau Monde* ou retourner d'où ils viennent. Avec l'utilisation de normes sévères, non adaptées aux populations évaluées, Goddard et les innéistes, s'appuyant sur les tests d'intelligence, ont pour projet de refouler « le plus médiocre de chaque race » (Gould, 1997).

2. L'humain est affecté par des biais de raisonnement

Le cerveau est extraordinairement complexe, quatre-vingt-six milliards de neurones, un million de milliards de connexions. Pour comprendre le monde, nous nous retrouvons dans l'obligation de simplifier le réel. Pour cela, l'humain a inventé des stratégies, des heuristiques, rapides et efficaces, qui sont cependant affectées par les biais cognitifs. Ceux-ci représentent une tendance humaine qui amène à prendre en compte des facteurs non pertinents pour raisonner. Envisageons deux biais cognitifs assez fréquents.

a) Les biais de confirmation

Il s'agit là d'une des erreurs épistémologiques parmi les plus fréquentes. C'est par exemple, face à un enfant qui présente des difficultés que l'on pense avoir identifiées, trouver des faits qui vont dans le sens de ce que nous pensons. C'est-à-dire que toutes les observations supplémentaires que le professionnel peut développer risquent de confirmer son idée princeps.

Il s'agit évidemment d'un biais qui peut avoir des conséquences graves, particulièrement dans le cadre de l'examen psychologique, où le professionnel doit pouvoir faire preuve d'ouverture d'esprit, de flexibilité, de résistance à la demande initiale et à l'attente des tiers demandeurs, dans le cas d'un enfant, et face aux symptômes que présente le sujet.

b) Le biais de l'illusion de la fin de l'histoire

Le jugement, particulièrement celui des professionnels qui s'occupent de sujets en développement, peut être obscurci par ce que l'on pourrait nommer l'illusion de la fin de l'histoire, en référence à une étude publiée en 2013 dans la revue *Science* (Quoidbach, Gilbert, Wilson), qui interroge 19 000 personnes âgées de 18 ans à 68 ans. Des questions sont posées aux sujets : on leur demande d'évaluer leur

« changement » au cours des dix dernières années et de prédire leur « changement » pour les prochaines dix années à venir.

Les sujets, en grande majorité, insistent sur combien ils ont changé lors de la décennie passée et combien ils pensent ne pas beaucoup changer dans la période des dix années à venir. Il y a, chez les sujets de l'étude, le sentiment que le présent est un moment charnière, ce qui les amène à penser qu'ils vont demeurer ainsi tout le restant de leur vie.

On retrouve ce sentiment chez les éducateurs, les psychologues qui croient que ce qu'ils observent *hic et nunc* est ce qui demeurera à l'avenir. Comme si l'enfant ou l'adolescent était devenu la personne qu'il sera toute son existence ; comme si ses « caractéristiques » scolaires, cognitives ou affectives ne devaient pas évoluer. La plupart des humains sous-estiment l'importance des changements à venir, particulièrement chez les jeunes en difficulté personnelle ou scolaire.

Il s'agit là d'un biais particulièrement dommageable pour le sujet de l'examen psychologique qui, dans ses fondements, se positionne comme une évaluation dont l'objet est, toujours, directement ou secondairement, le changement.

D. Science avec conscience

Pour conclure, rappelons qu'il convient toujours d'effectuer un bilan cognitif dans le cadre d'un examen psychologique.

Si comme nous le pensons, l'examen psychologique fait appel à l'art du psychologue, nous n'oublions pas qu'il tire ses fondements de la clinique bien entendu, mais aussi des sciences cognitives et de la neuropsychologie, qu'il prend appui, voire s'arc-boute, sur les règles éthiques développées par notre profession afin de protéger les consultants et les psychologues des mésusages de cet acte si particulier qui engage un sujet, ses proches, ainsi que les autres professionnels qui travaillent avec lui. Cette rencontre féconde n'advient que dans un cadre original, comme il n'en existe nulle part ailleurs, qui allie un espace contenant, un temps limité, l'utilisation d'épreuves psychologiques scientifiquement validées et dont l'accent est porté sur le sujet, sa singularité, et non sur les caractéristiques de ses troubles. Cette approche respectueuse du sujet prend bien évidemment en compte les difficultés, la symptomatologie et le

handicap qu'elle intègre dans le processus de compréhension et dans l'objectif ultime d'aide au développement.

Clinique, sciences cognitives et éthiques, trois dimensions essentielles qui organisent notre façon de penser et notre activité professionnelle.

Bibliographie

Bachelier, D., Cognet, G. (2017). *Le bilan avec l'le WISC-V*. Paris : Dunod.

Binet, A. (1909). *Les idées modernes sur les enfants*. Paris : Ernest Flammarion.

Carroll, J. B. (1993). *Human cognitive abilities. A survey of factor-analytic studies*. New York : Cambridge University press.

Cattell, R., B. (1963). Theory of fluid and crystallized intelligence, a critical experiment. *Journal of Educational Psychology*, 54, 1-22.

Cognet, G., Bachelier, D. (2016). *Clinique de l'examen psychologique de l'enfant et de l'adolescent*. Paris : Dunod.

Cognet, G. (2011). Les trois vies du Binet-Simon, Binet-Simon, Nemi, Nemi-2. *Recherches & Educations*, 5, 165-179.

Cognet, G. (2006). *NEMI-2*. Paris : Editions du Centre de Psychologie Appliquée.

Flanagan, D. P., Alfonso, V. C. (2016). *Essentials of WISC-V assessment*. Hoboken. New Jersey : Wiley.

Gottfredson, L. S. (1997). Mainstream science on intelligence : an editorial with 52 signatories, history and bibliography. *Intelligence*, vol 24, 1, 13-23.

Gould, S. (1997). *La mal-mesure de l'homme*. Paris : Ed Odile Jacob.

Grégoire, J. (2010). Qu'est ce que l'intelligence ? Est-elle quantifiable ? *ANAE*, 110, 353-359.

Grégoire, J. (2019). *WISC-V, L'examen clinique de l'intelligence de l'enfant*. Wavre : Mardaga.

Houdé, O. (2019). *L'intelligence humaine n'est pas un algorithme*. Paris : Odile Jacob.

Huteau, M., Lautrey J. (1999). *Évaluer l'intelligence*. Paris : Puf.

Mackintosh, N. J. (2004). *QI et intelligence humaine*. Bruxelles : De Boeck.

- McGrew, K., S. (2003). *Cattell-Horn-Carroll, definition project*. Institute for applied psychometrics.<http://themindhub.com/>
- Schneider, W. J., & McGrew, K. S. (2017). The Cattell-Horn-Carroll Theory of Cognitive Abilities. In D. P. Flanagan & Erin M .McDonough Eds.), *Contemporary intellectual assessment: Theories, tests and issues* (4th ed.,). New York: Guilford Press.
- Spearman C. (1904). General intelligence, objectively determined and measured.*American Journal of Psychology*, 15, 201-293.
- Zazzo, R. (1979). *Les déficits mentales*. Paris : Armand-Colin, 3e ed.