



Jacques Grégoire

Docteur en psychologie
Professeur à l'université
de Louvain

Comment interpréter

Le Wisc-v est un outil dont l'interprétation demeure riche et complexe. Jacques Grégoire revient ici sur les modèles théoriques qui l'ont fondé et sur les évolutions qui ont mené à cette nouvelle version. De manière claire et précise, il expose ce que mesurent les cinq indices qui le composent et nous donne quelques clés pour les interpréter.

La cinquième édition de l'échelle d'intelligence de Wechsler pour les enfants (Wisc-v) introduit un nouvel indice et modifie la mesure de plusieurs des indices existant dans la version précédente. Cette nouvelle échelle peut, dès lors, être déroutante pour des praticiens habitués à l'utilisation du Wisc-iv. Dans cet article, nous revenons sur les fondements de l'organisation de la nouvelle échelle dont nous présentons les bases théoriques, cliniques et statistiques. Nous passons ensuite en revue chacun des cinq indices et ce qu'ils mesurent. Nous exposons enfin les grandes lignes d'une méthode d'analyse des scores aux indices qui permet d'extraire un maximum d'informations pertinentes d'un protocole de Wisc-v.

DE LA MESURE DU QI À CELLE DES INDICES

David Wechsler a mené la plus grande partie de sa carrière professionnelle au Bellevue Psychiatric Hospital de New York. Se définissant avant tout comme psychologue clinicien, il a toujours défendu un usage des tests respectueux de la personne, qui ne confond pas le processus d'évaluation avec le *testing*. En 1975, à la fin de sa vie, il écrivait encore : « *Ce que nous mesurons avec les tests n'est pas ce que les tests mesurent – ni l'information,*

ni la perception spatiale, ni la capacité de raisonnement. Celles-ci ne sont que des moyens pour une fin. Ce que les tests d'intelligence mesurent, ce que nous espérons qu'ils mesurent, est quelque chose de beaucoup plus important : la capacité d'un individu de comprendre le monde qui l'entoure et ses ressources pour faire face aux défis qu'il lui présente. » (Wechsler, 1975). Pour atteindre cet objectif, David Wechsler a conçu son test dans le respect des principes énoncés par Alfred Binet au début du ^{xx}e siècle. Selon ce dernier, le niveau d'intelligence d'un individu ne peut être correctement estimé qu'à travers ses performances à un ensemble d'épreuves dont on identifie la tendance moyenne. Pour la première version de son test (1939), David Wechsler sélectionne onze épreuves variées dont ils avaient pu expérimenter l'intérêt clinique, lorsqu'il était jeune psychologue dans l'armée américaine. Ces épreuves, dont huit font encore partie de l'échelle actuelle, lui permettent de calculer un QI, dont une preuve de validité est son pouvoir de prédiction des performances scolaires et professionnelles.

De son expérience à l'armée, David Wechsler ne conserve pas seulement les épreuves qui constituent la base de son test. Il retient également l'intérêt de calculer, en plus d'un score global (QI Total), deux autres scores composites, l'un basé sur les épreuves verbales et l'autre sur les épreuves non verbales. Il nomme le premier « QI Verbal » et le second « QI de Performance ».

Ces deux mesures complémentaires au QIT vont connaître un grand succès auprès des praticiens qui y trouvent des informations utiles pour comprendre le fonctionnement intellectuel des personnes évaluées. L'organisation des échelles de Wechsler sur la base de trois scores composites (QIT, QIV et QIP) est restée inchangée jusqu'au décès de leur créateur, en 1981. Il faudra attendre le Wisc-III, publié aux États-Unis en 1991, pour que l'organisation du test soit (timidement) remise

Les indices du Wisc-v ?

en question. Pourtant, du vivant de Wechsler, plusieurs analyses factorielles des échantillons d'étalonnage avaient suggéré d'organiser les épreuves en trois sous-échelles plus cohérentes et plus pertinentes du point de vue clinique.

Aux côtés des QIV et QIP, une troisième sous-échelle regroupait les épreuves d'Arithmétique, de Mémoire de chiffres et de Code, qui étaient fortement corrélées entre elles et sensibles aux troubles de l'attention. Malgré les évidences empiriques, David Wechsler n'a toutefois jamais souhaité modifier l'organisation de son échelle, la jugeant pleinement satisfaisante du point de vue clinique.

Wechsler disparu, la composition et l'organisation des échelles ont progressivement évolué.

Dans le Wisc-III, l'introduction d'une nouvelle épreuve – Symboles – a permis de mettre en évidence une organisation des subtests en quatre sous-échelles : Compréhension verbale, Organisation perceptive, Attention / Concentration et Vitesse de traitement. En conséquence, les praticiens ont eu la possibilité de calculer quatre scores composites (indices) correspondant à ces quatre sous-échelles. Il ne s'agissait toutefois que d'une alternative possible au calcul des QIV et QIP qui restait privilégié. Cette cohabitation de deux systèmes

de calcul des scores composites a perduré dans la troisième édition de l'échelle d'intelligence de Wechsler pour adultes (WAIS-III).

Ce n'est qu'en 2003, avec la parution du Wisc-IV aux États-Unis, que les QIV et QIP ont été abandonnés au profit des seuls indices. Simultanément, les quatre indices développés sur des bases statistiques étaient mis en relation avec les composantes de deux modèles de la structure de l'intelligence, celui de John B. Carroll (1993) et celui de Raymond Cattell et John L. Horn (1966), rassemblés aujourd'hui sous l'acronyme CHC (Cattell-Horn-Cattell). Le modèle hiérarchique de John B. Carroll est illustré dans la figure 1. Il s'appuie sur l'analyse de plus de quatre cent soixante bases de données rassemblées lors de recherches sur les aptitudes intellectuelles réalisées entre 1930 et 1985. Il représente la structure des aptitudes humaines qui, globalement, constitue l'intelligence. À la base du modèle se trouvent les aptitudes élémentaires qui, chacune,

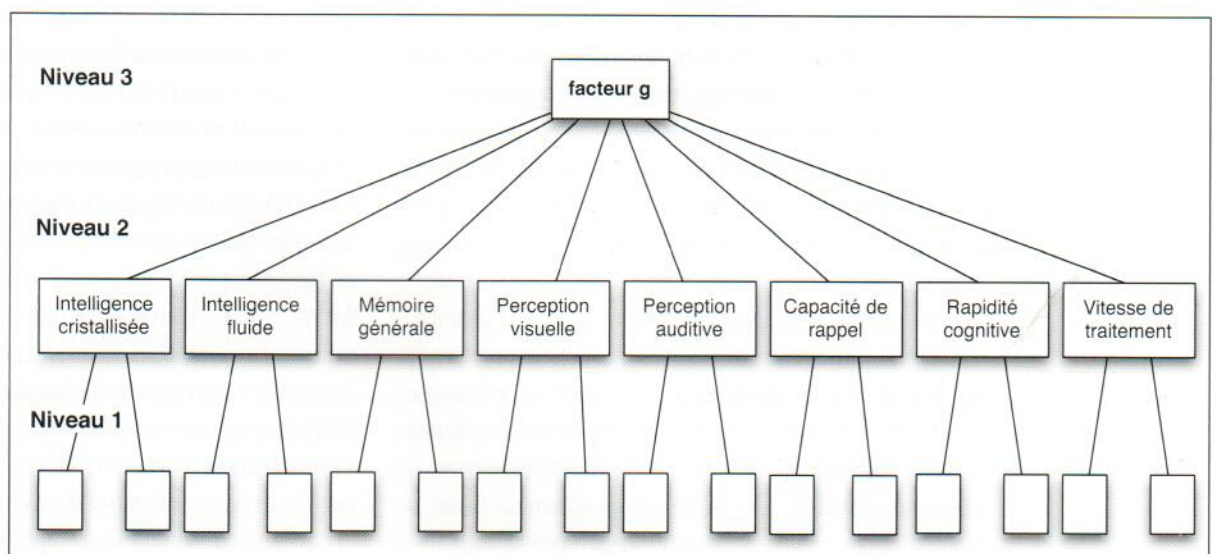


Figure 1 : Modèle hiérarchique de l'intelligence selon John B. Carroll.

→ couvrent un domaine étroit de l'activité intellectuelle. Le second niveau du modèle comprend huit aptitudes générales qui couvrent des domaines étendus de l'activité intellectuelle. Enfin, au sommet du modèle, on trouve un facteur général (*g*) qui représente des caractéristiques partagées par toutes les aptitudes intellectuelles et qui explique les corrélations observées entre ces dernières. La nature du facteur *g* fait encore débat, certains l'associant avec des propriétés neuronales comme la vitesse de conductivité des neurones, alors que d'autres considèrent qu'il s'agit plutôt d'une capacité de contrôle de nos activités intellectuelles. Développé en suivant un chemin différent de celui de John B. Carroll, le modèle de Raymond Cattell et John L. Horn aboutit à une structure de l'intelligence quasi identique, ce qui justifie l'intégration des deux modèles au sein du modèle CHC.

Bien que le modèle CHC comprenne huit aptitudes au niveau 2, seulement cinq de celles-ci sont mesurées par le Wisc-v pour des raisons de valeur clinique et de temps de passation.

La mise en relation des indices du Wisc-IV avec certaines composantes de second niveau du modèle CHC a été faite *a posteriori*. Les correspondances sont, par conséquent, approximatives. C'est particulièrement le cas pour l'indice Raisonnement perceptif qui correspond à la fois à l'Intelligence cristallisée et à la Perception visuelle, ce qui complique son interprétation, car il n'est pas aisé de démêler l'influence des deux aptitudes intellectuelles dans les performances observées. Lors du développement du Wisc-v, la volonté a été de rechercher une correspondance terme à terme entre chaque indice et une seule aptitude de niveau 2. Bien que le modèle CHC comprenne huit aptitudes au niveau 2, seulement cinq de celles-ci sont mesurées par le Wisc-v pour des raisons de valeur clinique et de temps de passation. Il s'agit des aptitudes suivantes : Intelligence cristallisée, Intelligence fluide, Mémoire générale, Perception visuelle et Rapidité cognitive. Ces cinq facettes de l'intelligence se révèlent les plus intéressantes dans le cadre de l'examen diagnostique des troubles du raisonnement et des apprentissages. Par ailleurs, pour être mesurée de manière fiable, une aptitude nécessite d'être évaluée par plusieurs épreuves. Dans le cas du Wisc-v, la décision a été de mesurer chacune des cinq aptitudes sélectionnées

par deux épreuves afin de limiter la durée de l'examen intellectuel.

QUE MESURENT EFFECTIVEMENT LES CINQ INDICES ?

Les épreuves sélectionnées pour mesurer les Indices et leurs correspondances avec les aptitudes du modèle CHC sont présentes dans le tableau 1. Le choix des épreuves s'est fait sur la base de leur intérêt clinique et de leurs corrélations avec les facteurs correspondants. Les praticiens doivent avoir à l'esprit que les épreuves choisies ne permettent qu'une estimation des aptitudes visées. Ces dernières ne peuvent, en effet, pas être mesurées directement, mais seulement à travers des performances à des tâches utilisées comme révélateurs. Or, ces tâches font toutes appel à des processus mentaux relativement complexes qui débordent la seule aptitude visée. Pour interpréter correctement les performances à une tâche donnée, il est, dès lors, essentiel de bien connaître les processus impliqués dans cette dernière, afin de pouvoir identifier les sources d'éventuelles difficultés. Un même score peut, en effet, être la résultante des processus et/ou de dysfonctionnements variés. L'indice Compréhension verbale (icv) est une mesure de ce que Raymond Cattell appelle l'« Intelligence cristallisée ». Il s'agit de l'ensemble des connaissances et des compétences acquises tout au long de notre existence qui nous permettent de faire face à la majorité des problèmes que nous rencontrons dans la vie quotidienne. Pour une large part, il s'agit d'aptitudes verbales qui nous permettent de penser et de communiquer. Il est donc logique que les deux épreuves choisies pour mesurer cet indice fassent toutes les deux appel à l'élaboration de concepts verbaux. Il ne faut toutefois pas en conclure que l'intelligence cristallisée est uniquement verbale. Effectuer une addition de nombres décimaux, comprendre une séquence d'images dans une bande dessinée ou utiliser un téléphone portable sont également des compétences acquises qui font partie de l'Intelligence cristallisée, bien qu'elles soient non verbales. Du fait de sa composition, il n'est pas étonnant que l'icv soit l'indice le mieux corrélé avec les performances scolaires qui font largement appel aux compétences intellectuelles verbales. Selon les disciplines scolaires, les corrélations observées s'étendent de 0,50 à 0,70. Logiquement, les opportunités d'apprentissage favorisent le développement de l'intelligence cristallisée, dont l'icv est un bon reflet. Cet indice est d'ailleurs le plus lié au niveau éducatif des parents.

Ainsi, selon que les parents n'ont aucun diplôme ou qu'ils ont un diplôme bac + 3, on observe un écart de performances de leurs enfants de 15 points en moyenne à l'icv. L'indice Visuo-spatial (ivs) est une mesure de ce que John B. Carroll appelle la « Perception visuelle » et que nous appelons généralement « l'intelligence visuo-spatiale ». Il s'agit de la capacité d'analyser, d'encoder et de manipuler mentalement des formes spatiales. C'est ce que mesurent les épreuves de Cubes et de Puzzles visuels. Afin que le score à l'icv soit le plus focalisé sur l'aptitude visée, l'impact de la vitesse a été réduit au maximum ; les bonus de points associés au temps de résolution n'apparaissent, en effet, plus que dans les quatre derniers items de l'épreuve des Cubes.

Indices du Wisc-v	Épreuves	Modèle CHC	Sigle
Compréhension verbale	Similitudes	Intelligence cristallisée	Gc
	Vocabulaire		
Visuo-spatial	Cubes	Perception visuelle	Gv
	Puzzles visuels		
Raisonnement fluide	Matrices	Intelligence fluide	Gf
	Balances		
Mémoire de travail	Mémoire de chiffres	Mémoire générale	Gsm
	Mémoire d'images		
Vitesse de traitement	Code	Rapidité cognitive	Gs
	Symboles		

Tableau 1. Correspondance entre les indices du Wisc-v et les aptitudes du modèle CHC.

Cette dernière est la seule

épreuve du Wisc-v qui demande

de manipuler des objets. Cette externalisation de la démarche de résolution de problèmes permet d'intéressantes observations des procédures utilisées par les enfants (démarche systématique ou par essais et erreurs ; réponses gâchées...). L'épreuve de Puzzles visuels remplace l'ancienne épreuve du même nom qui demandait d'assembler des pièces pour construire une figure. À présent, les enfants ne doivent plus manipuler du matériel, mais résoudre les items de manière purement mentale, ce qui est une source de complexité. Par ailleurs, le plus grand nombre d'items de la nouvelle épreuve et la disparition de la mesure du temps ont permis une amélioration notable de la fidélité et de la validité.

L'indice Raisonnement fluide (IRF) correspond à l'aptitude du même nom dans le modèle CHC. Il mesure une capacité générale de résolution de problème qui consiste à extraire des règles qui sous-tendent une situation problématique, puis à les utiliser pour résoudre le problème posé.

L'illustration typique de cette démarche intellectuelle est donnée par l'épreuve Matrices, inspirée du test créé par John C. Raven en 1938. Le sujet doit analyser une séquence de figures géométriques, afin d'identifier les règles qui sous-tendent cette suite. Il doit ensuite appliquer ces règles pour compléter la série. Il s'agit donc de mettre en œuvre une suite de raisonnements inductifs et déductifs. Cette démarche générale permettant à l'être humain de faire

face aux problèmes rencontrés lorsqu'il ne possède pas de procédure spécifique acquise par l'expérience ou l'éducation. Certains auteurs ont identifié l'intelligence fluide à un noyau central de l'intelligence d'origine innée. Ils en ont trop rapidement conclu que les tests de suites logiques fournissent une bonne estimation de l'intelligence dégagée des influences culturelles. Cette conception ne résiste pas à l'examen des données empiriques. De toutes les épreuves intellectuelles, les séries logiques sont, en effet, les plus sensibles à l'effet Flynn, c'est-à-dire à l'évolution, depuis plus de cinquante ans, des performances intellectuelles moyennes au sein des populations des pays développés. Or, une telle évolution rapide n'est possible que si les aptitudes impliquées dans les épreuves sont influencées par l'éducation. Par ailleurs, les données d'étalonnage du Wisc-v montrent que les performances des enfants à l'IRF sont presque aussi sensibles au niveau d'études de leurs parents que leurs performances à l'icv. Enfin, il ne faut pas perdre de vue que les raisonnements inductifs et déductifs ne se déroulent pas dans le vide, mais portent toujours sur des contenus spécifiques avec lesquels le sujet examiné est plus ou moins à l'aise en fonction de ses intérêts et de ses expériences. Dans l'épreuve Matrices, ces contenus sont essentiellement géométriques, alors que, dans l'épreuve Balances, ils sont principalement numériques. L'indice Mémoire de travail (IMT) permet une évaluation de l'aptitude Mémoire générale au sein du modèle CHC. ➔

DOSSIER

Le Wisc-v : du nouveau pour l'observation clinique

→ Dans le Wisc-v, la mesure de la mémoire a été élargie avec l'introduction de la nouvelle épreuve Mémoire des images qui fournit une mesure de la mémoire de travail visuelle en plus de la mémoire de travail auditive déjà évaluée par l'épreuve Mémoire de chiffres. La mémoire de travail occupe une place privilégiée au sein de l'intelligence. Elle est la composante de l'intelligence où nous traitons l'information de manière consciente. Sa contenance est limitée, ce qui en fait le goulot d'étranglement du système cognitif. Plus nous sommes capables de traiter d'informations en même temps et à une plus grande vitesse, plus nous pouvons résoudre des problèmes complexes. C'est sur cette base que sont construites les suites logiques de l'IRF. Leur complexité augmente en fonction du nombre de règles que nous devons gérer simultanément. Bien qu'il permette une mesure élargie de la mémoire de travail, l'IMT ne couvre toutefois pas toute l'étendue de cette dernière. Il ne peut pas, dès lors, être considéré comme une mesure générale de la mémoire de travail. Lorsqu'un score faible est observé à l'IMT, il peut, par conséquent, être utile de compléter la mesure de la mémoire de travail par des tests qui en évaluent d'autres facettes. Par ailleurs, il est souvent intéressant d'examiner les scores complémentaires à cette épreuve. C'est, par exemple, le cas de la mesure de l'empan de chiffres en ordre direct et en ordre inversé. Les deux mesures peuvent, en effet, être très différentes, car le rappel en ordre direct est une mesure de la seule mémoire immédiate, alors que le rappel en ordre inversé est également une mesure de l'efficacité de l'administrateur central (gestion de la double tâche de mémorisation de la série et de sa réorganisation en ordre inversé). L'IMT demande un réel effort de concentration, ce qui le rend

très sensible aux troubles de l'attention. Il est aussi sensible à la dyslexie où le traitement des séquences d'information peut être perturbé.

L'indice Vitesse de traitement (IVT) permet une estimation de l'aptitude Rapidité cognitive au sein du modèle CHC. Il ne fournit toutefois qu'une mesure très étroite de cette dernière, car les deux épreuves qui le composent n'évaluent que la vitesse de traitement de *stimuli* visuels non sémantiques et de réponse motrice manuelle. De nombreux aspects de la vitesse de traitement ne sont pas reflétés dans cet indice. Les praticiens doivent, dès lors, se garder de toute généralisation abusive des résultats observés. Comme les réponses impliquent d'écrire, l'IVT est sensible au degré d'habileté graphomotrice et aux dyspraxies. La composante motrice des tâches de l'IVT peut aussi être pénalisante pour les enfants surdoués chez qui les traitements mentaux sont très rapides, mais dont le développement moteur correspond souvent à celui des enfants du même âge. Un score élevé à l'IVT suppose une grande concentration dans les deux épreuves proposées. Il en découle que l'IVT est très sensible aux troubles de l'attention et à tout ce qui peut ralentir le traitement cognitif, comme la dépression et les troubles d'apprentissage.

QUELLE MÉTHODE D'INTERPRÉTATION DES INDICES ?

Interpréter les indices implique d'examiner d'abord leur dispersion et leur homogénéité. Dans l'hypothèse où un enfant obtiendrait une note standard de 10 à tous les subtests et, par conséquent, un score de 100 à chacun des cinq indices, nous pourrions simplement conclure que

toutes ses performances se situent dans la moyenne, sans pouvoir dire grand-chose d'autre. Dans une telle situation, le QIT serait le meilleur indicateur de l'ensemble des capacités intellectuelles de l'enfant. Toutefois, dans la grande majorité des cas, nous n'observons pas un tel tableau de performances homogènes. Des écarts plus ou moins grands existent entre les scores aux indices et entre les subtests qui composent ces derniers. La question est, dès lors, de savoir comment évaluer l'importance de ces écarts afin d'identifier ceux qui méritent de retenir notre attention et d'être interprétés. La méthode la plus simple pour analyser la dispersion des indices est illustrée par la figure 2.

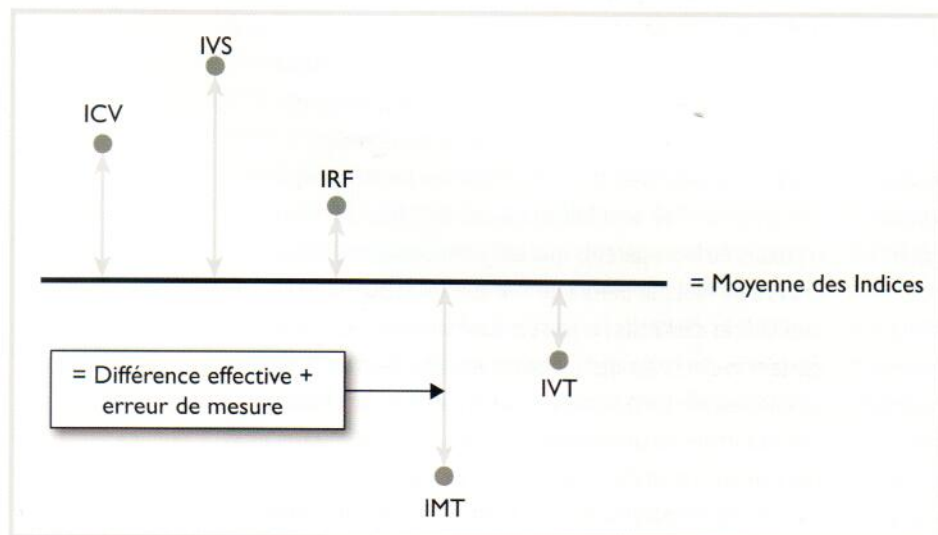


Figure 2. Analyse de la dispersion des scores aux indices.

Elle consiste à calculer la moyenne des scores aux cinq indices, puis la différence entre les scores à chaque indice et cette moyenne.

Comme les scores au test sont inévitablement entachés d'erreurs de mesure, toutes les différences observées ne sont pas nécessairement le reflet d'une différence effective entre certaines aptitudes. La probabilité est très élevée que les petites différences ne soient que le fruit d'erreurs de mesure aléatoires. Inversement, plus les différences sont grandes, plus il est probable qu'elles traduisent de véritables différences entre certaines aptitudes. Le manuel fournit les valeurs seuils à partir desquelles il est raisonnable de prendre en considération les différences observées et de les interpréter en termes de forces ou de faiblesses. Ces différences bien réelles ne doivent toutefois pas être systématiquement considérées comme les manifestations de troubles. En effet, une certaine dispersion des performances est courante chez les sujets tout-venant, sans pathologie avérée. Cette variabilité est normale et ne devrait pas nous étonner. Elle est le fruit de différences d'opportunités et d'intérêts au cours du développement intellectuel. Ainsi, certains enfants développent plus la facette verbale de leur intelligence, alors que d'autres enfants investissent plus la facette visuo-spatiale. Le tableau 2 présente, pour chaque indice, le pourcentage d'enfants de l'échantillon d'étalonnage dont le score s'écarte significativement de la moyenne des cinq indices. On peut constater que cela représente entre un cinquième et un quart des enfants de l'échantillon d'étalonnage. Parmi ceux-ci, seulement 34,8 % ne présentent aucun indice s'écartant significativement de la moyenne. Un certain degré de dispersion des indices est donc un phénomène courant qui doit être évalué avec discernement en évitant toute interprétation systématique en termes pathologiques.

La dispersion des indices peut avoir des répercussions sur l'interprétation du QIT. En effet, plus la dispersion est importante, moins le QIT est un indicateur pertinent du fonctionnement intellectuel global. Le même raisonnement doit être tenu à propos des indices eux-mêmes. Lorsque les scores aux deux subtests qui composent un indice s'écartent significativement l'un de l'autre, la validité de cet indice comme mesure d'une aptitude de grande étendue peut être mise en question. Le manuel du Wisc-v fournit les valeurs seuils qui permettent de déterminer quand une différence entre deux subtests peut être considérée comme statistiquement significative.

Sur la base de l'analyse de la dispersion des indices et de leur homogénéité, il est possible d'identifier

	%
I. Compréhension verbale	19,0
I. Visuo-spatial	22,0
I. Raisonnement fluide	25,1
I. Mémoire de travail	17,6
I. Vitesse de traitement	25,4

Tableau 2. Fréquence chez les enfants tout-venant des écarts significatifs par rapport à la moyenne pour chaque indice.

les scores qui méritent de retenir notre attention et faire l'objet d'une interprétation. Dans certains cas, le profil sera suffisamment homogène pour que nous considérions le QIT comme l'indicateur le plus représentatif du fonctionnement intellectuel du sujet examiné. Dans d'autres cas, face à un profil de scores plus contrasté, nous pourrions parler de « forces » et de « faiblesses » au sein des composantes de l'intelligence. Il faudra alors être attentif à ne pas qualifier de « pathologiques » toutes les faiblesses. Celles-ci sont, en effet, relatives. Par exemple, chez un enfant surdoué dont quatre indices sont supérieurs à 135, un indice à 118 sera une faiblesse relative, mais ne pourra pas être considéré comme le signe d'un trouble cognitif. Pour parler de « troubles », il est nécessaire qu'un score soit nettement inférieur à la moyenne des sujets de l'échantillon d'étalonnage et qu'il soit associé à des difficultés d'adaptation ou d'apprentissage. Dans ce cas, pour identifier la signification d'un score apparemment pathologique, il sera important de croiser cette donnée avec d'autres informations quantitatives et qualitatives issues de l'ensemble de l'examen diagnostique.

EN CONCLUSION

À travers ce bref article d'introduction à l'interprétation des indices du Wisc-v, le lecteur peut se faire une idée de la richesse, mais aussi de la complexité de cette interprétation. Le Wisc-v est un excellent outil de mesure, mais sa véritable valeur ne peut s'exprimer que dans les mains d'un praticien compétent, bien informé des forces et des faiblesses de l'instrument, et respectueux des règles méthodologiques et éthiques qui devraient présider à toute interprétation des résultats. ■