

Indice d'aptitude général et indice de compétence cognitive pour le
WISC-IV : normes empiriques versus normes statistiques

LECERF, Thierry, *et al.*

Reference

LECERF, Thierry, *et al.* Indice d'aptitude général et indice de compétence cognitive pour le WISC-IV : normes empiriques versus normes statistiques. *Revue européenne de psychologie appliquée*, 2011, vol. 61, p. 115-122

DOI : 10.1016/j.erap.2011.01.001

Available at:

<http://archive-ouverte.unige.ch/unige:18881>

Disclaimer: layout of this document may differ from the published version.



**UNIVERSITÉ
DE GENÈVE**

Indice d'Aptitude Général et Indice de Compétence Cognitive
pour le WISC-IV : Normes empiriques vs. Normes statistiques

*General ability index and Cognitive proficiency index for the WISC-IV: Empirical norms vs.
statistical norms*

Résumé

L'*Indice d'Aptitude Général* (IAG) a été développé dans le WISC-III pour estimer le niveau cognitif général. Il s'agissait de proposer une alternative au QI Total. Plus récemment, l'*Indice de Compétence Cognitive* (ICC) a été proposé. Jusqu'à il y a peu, seules les normes américaines de ces indices étaient disponibles ; les normes françaises IAG et ICC pour le WISC-IV ont été récemment développées à partir d'une procédure d'approximation statistique. Toutefois, les normes développées de cette manière seraient moins appropriées que les normes construites sur la base d'un échantillon. Cette étude vise à comparer les normes IAG et ICC développées à partir d'une procédure d'approximation statistique à celles créées à partir d'un échantillon de 182 enfants âgés de 8 à 12 ans. Les corrélations entre les deux types de normes sont de 0,997 pour IAG et de 0,999 pour ICC ; les différences absolues moyennes sont respectivement de 5,38 et de 2,24 points pour IAG et ICC. Enfin, pour les enfants présentant un score IAG_S inférieur à 89, les normes IAG_E semblent indiquer que leurs compétences seraient inférieures d'environ 8 – 10 points à celui indiqué par le score IAG_S. Les résultats de cette étude indiquent que les scores IAG et ICC peuvent constituer des informations utiles sur les capacités cognitives de l'enfant.

Keywords: WISC-IV, Indice d'aptitude général, Indice de compétence cognitive, normes, CHC

Abstract

Introduction

The *General Ability Index* (GAI) was developed with the WISC-III to measure general cognitive ability. The GAI can be used as a substitute for the FSIQ. More recently, the *Cognitive Proficiency Index* (CPI) was developed. The GAI and CPI tables for the French WISC-IV were recently created using statistical approximation. However, it has been suggested that tables created with statistical approximation were less appropriate than tables created using sample.

Objective

The objective of this study was to compare GAI and CPI normative tables for the French WISC-IV created using a statistical approximation procedure and tables based on a sample.

Method

To further explore the validity of the GAI and ICC scores, we studied 182 children aged from 8 to 12 years old.

Results

Correlations between both types of tables were 0,997 for GAI and 0,999 for CPI, and the absolute mean difference is 5,38 points and 2,24, respectively. For the children presenting a GAI_S lower than 89, GAI_E suggested that it is likely that their scores are lower about 8 - 10 points than that indicated by GAI_S.

Conclusion

Results of this study support the use of GAI and CPI to provide different views into children's cognitive abilities.

Keywords: WISC-IV, General Ability Index, Cognitive Proficiency index, norms, CHC

Introduction

Lorsqu'en 1939 David Wechsler propose la première version de son test, le *Wechsler-Bellevue I (W-B I)*, il partage la conception théorique de Binet et Simon (1905), et considère que l'intelligence est une caractéristique globale, qui doit être évaluée à travers un ensemble d'activités. En outre, bien que proposant le calcul du QI Verbal et du QI Performance, Wechsler s'inscrit dans la continuité des propositions de Stern (1912) et adopte le Quotient Intellectuel Total (QIT) comme indicateur du fonctionnement intellectuel de l'individu. Au fait des théories de son époque, Wechsler s'appuie alors sur le modèle de Spearman (1927) pour légitimer le calcul du QIT. Rappelons que Spearman suggère à partir d'analyses factorielles, l'existence d'un facteur commun à toutes les épreuves : le facteur général d'intelligence (facteur *g*). Par la suite, les travaux réalisés par Thurstone (1938) vont remettre en question cette conception unidimensionnelle, puisque ce dernier aboutit à l'hypothèse de l'existence de 7 aptitudes mentales primaires ; Thurstone n'observe pas de facteur général. Par la suite, les propositions contradictoires de Spearman et de Thurstone sur la structure de l'intelligence vont être réconciliées en adoptant une perspective hiérarchique. Par exemple, Horn et Cattell (1967) distinguent 2 facteurs généraux : l'Intelligence Fluide (*Gf*) et l'Intelligence Cristallisée (*Gc*). Ces 2 facteurs sont ensuite subdivisés en facteurs de premier ordre. Carroll (1993), quant à lui, parvient à un modèle hiérarchique comprenant 3 niveaux : un facteur *g* de 3^{ème} ordre, 8 facteurs de 2^{ème} ordre (*Gf*, *Gc*, etc.), à leurs tours subdivisés en facteurs de 1^{er} ordre. Récemment, les propositions de Cattell-Horn et de Carroll ont été fusionnées au sein du modèle appelé *Cattell-Horn-Carroll (CHC)*, dans lequel on admet l'existence d'un facteur général d'intelligence, d'une dizaine de facteurs secondaires (intelligence fluide, etc.) et d'environ soixante dix facteurs primaires (relations spatiales, etc. ; McGrew, 2009). Ce modèle représente le développement le plus complet de cette

perspective multidimensionnelle des aptitudes cognitives (Grégoire, 2009; Lecerf, Rossier, Favez, Reverte, & Coleaux, 2010).

Cette évolution conceptuelle sur la structure de (des) l'intelligence(s) a des conséquences sur les aspects méthodologiques liés à la construction des batteries d'intelligence et à la mesure de l'intelligence (Huteau & Lautrey, 1999). Ainsi, la mesure de l'intelligence dans les échelles de Wechsler a suivi partiellement cette évolution d'une conception unidimensionnelle vers une conception multidimensionnelle. Dans la dernière édition de l'échelle de Wechsler pour enfants et adolescents (WISC-IV, Wechsler, 2005), le QI Verbal et le QI Performance ont été supprimés et l'approche multidimensionnelle a été renforcée par l'introduction de 4 indices factoriels : Compréhension Verbale (ICV), Raisonnement Perceptif (IRP), Vitesse de Traitement (IVT), Mémoire de Travail (IMT) ¹. L'introduction de ces 4 indices factoriels doit permettre une évaluation des compétences cognitives plus cohérente avec la théorie CHC des aptitudes cognitives. Cependant, le WISC-IV permet toujours le calcul et l'interprétation du QIT comme unique indicateur du fonctionnement intellectuel de l'individu. Or, le changement paradigmatique d'une conception unidimensionnelle vers une conception multidimensionnelle de la structure de l'intelligence a comme conséquence de questionner la pertinence et l'utilité du QI Total dans le cadre du bilan cognitif. L'utilisation et la restitution d'un score unique paraît *a priori* en opposition avec l'approche multidimensionnelle et pose donc le problème de l'évaluation du fonctionnement intellectuel général (Grégoire, 2009; Lautrey, 2005; Lecerf, Reverte, Coleaux, Favez, Rossier, 2010 ; Rozencwajg, 2006).

1.1. QI total, facteur g : quelle mesure du fonctionnement intellectuel général dans le WISC-IV ?

Comme nous avons eu l'occasion de l'écrire précédemment (Lecerf, Reverte et al., 2010), tous les auteurs s'accordent à reconnaître que l'estimation du facteur g est nécessaire dans le cadre du bilan psychologique. C'est probablement la raison pour laquelle les concepteurs du WISC-IV ont maintenu le QIT, puisqu'ils indiquent que ce dernier peut être considéré comme une estimation du facteur g. Raiford, Weiss, Rolfhus, et Coalson (2005, p. 2) écrivent par exemple : « *le QI total est le plus fréquemment utilisé pour décrire l'intelligence générale, ou g* ». Pourtant, bien que le QIT et le facteur g soient tous deux destinés à évaluer le fonctionnement intellectuel général, ils sont conceptuellement différents (Grégoire, 2009 ; Lautrey, 2005). Le QIT est obtenu en moyennant divers sous-tests, lesquels contribuent avec la même importance à la moyenne (moyenne équipondérée). En revanche, le score en facteur g est obtenu à partir d'un ensemble de mesures, dont la contribution dépend de la saturation en facteur g de chacune des mesures. Autrement dit, un test ayant une forte saturation en facteur g aura un poids plus important qu'un sous-test ayant une faible saturation en facteur g. Prenons l'exemple des sous-tests Code et Similitudes : dans le calcul du QIT, ces 2 épreuves interviennent avec le même poids, soit 10 %. En revanche, dans le score en facteur g, la contribution du sous-test Similitudes est plus importante que celle de l'épreuve Code, car la saturation en facteur g de ce premier est plus élevée (respectivement 0,77 vs. 0,36 selon Grégoire, 2009).

1.2. Le QI : un concept inutile et dépassé ?

Dès l'apparition des premiers tests d'intelligence, les auteurs ont cherché à quantifier le niveau de fonctionnement intellectuel de l'individu ; Binet et Simon avaient proposé « le niveau mental », puis Stern (1912) le « QI développemental ». Depuis 1939, le « QI déviation » de Wechsler représente le score de référence (Grégoire & Wierzbicki, 2009). Le QIT fait cependant l'objet d'une controverse. En 2005, Lautrey écrit que « [...] le concept de

« QI n'est plus adapté pour rendre compte de cette richesse [de l'intelligence] et qu'il devrait donc être abandonné, y compris par les psychologues » (p. 150) ; il rajoute que « c'est aussi un concept qui n'est plus vraiment adapté aux connaissances scientifiques actuelles sur l'intelligence » (p. 150). La raison principale avancée par Lautrey est que le QI n'est plus adapté à la conception multidimensionnelle de l'intelligence. Si nous partageons l'argument avancé par Lautrey quant au décalage entre le QIT et l'approche multidimensionnelle, nous divergeons cependant quant à sa conclusion d'abandonner le QIT. Nous pensons pour notre part que la question de l'utilité et de la pertinence du QIT varie selon que l'on s'inscrit dans une démarche diagnostique ou dans une démarche pronostique. S'il s'agit d'émettre des hypothèses sur le fonctionnement cognitif d'un enfant (c.-à-d. l'approche diagnostic), de comprendre ses compétences et ses difficultés, le QIT ne présente que peu d'intérêt. En effet, un même QIT peut être obtenu par 2 enfants présentant des profils très différents et n'apporte donc aucune information relative au fonctionnement d'un enfant particulier. Rappelons toutefois, que le QIT est un des critères utilisés pour l'identification d'enfants atypiques, comme par exemple les enfants à Haut potentiel Intellectuel (HPI : $QIT \geq 125 - 130$) ou les enfants présentant des retard intellectuels (RI : $QIT \leq 70 - 75$; Grégoire, 2009 ; Rozencwajg, Aliamer, & Ombredane, 2009).

Si l'on se situe dans le cadre d'une approche pronostique, l'abandon du QIT nous paraît plus discutable. Tout d'abord, les travaux qui ont porté sur l'étude de la stabilité des mesures de l'intelligence à moyen et long terme ont montré que le QIT est la mesure la plus stable. Par exemple, Canivez et Watkins (2001) observent à partir d'une procédure Test - Retest (à 3 ans d'intervalle) une bonne stabilité du QIT dans différents groupes cliniques (r moyen = ,89). En ce qui concerne les indices factoriels, ces auteurs constatent que si les coefficients de fidélité sont acceptables pour la Compréhension Verbale et l'Organisation Perceptive (r moyen de 0,82 et de 0,83, respectivement), cela n'est pas le cas de la Vitesse de Traitement et

de la « Distractibilité » (0,58 à 0,74). On ne peut dès lors se baser sur ces deux derniers indices.

De même, les recherches portant sur la validité prédictive des mesures de l'intelligence témoignent de l'intérêt du QIT. En ce qui concerne le WISC-III, Glutting, Youngstrom, Ward, et Ward (1997), par exemple, montrent que le QIT est le meilleur prédicteur de la réussite académique (maths, langage, etc.), et qu'il explique environ 40 % de la variance des scores. En revanche, la part de variance additionnelle (validité incrémentale) expliquée par les indices factoriels est relativement faible, puisqu'elle varie entre 5 et 16 % lorsque l'on prend conjointement les 4 indices. Pris séparément, chaque facteur n'explique que de 0 à 5% de variance, lorsque l'on contrôle pour le QIT. Plus récemment, Watkins (2006) montre, sur les données de l'échantillon de standardisation Nord-américain du WISC-IV, que les pourcentages de variances expliqués par les indices factoriels sont relativement faibles, lorsque la variance expliquée par le facteur g est extraite en premier lieu (71,3 % de la variance partagée est expliquée par le facteur g, ce qui représente 38 % de la variance totale). Ainsi, l'indice de Compréhension Verbale (ICV) explique 12,1 % de la variance partagée (6,5 % de la variance totale) ; l'indice de Raisonnement Perceptif (IRP) explique 4,1 % ; l'indice de Mémoire de Travail (IMT) explique 4,3 % et enfin, l'indice de Vitesse de Traitement (IVT) 8,2 %.

Pris conjointement, ces deux courants de recherche indiquent clairement qu'un score global, et notamment le QIT, est le meilleur prédicteur de la réussite académique (entre 25 % et 50 % de la variance) tandis que les indices factoriels ne présentent à l'heure actuelle qu'un intérêt additionnel limité. On nous opposera peut-être que l'intérêt du QIT n'est vrai que si l'hétérogénéité entre les indices n'est pas trop importante. Rappelons, en effet, que les auteurs recommandent de n'interpréter le QIT que si la différence entre l'indice le plus haut et l'indice le plus bas est inférieure à 23 points (Sattler, 2001)². Or, les travaux conduits par

Freberg, Vandiver, Watkins, et Canivez (2008) montrent que le QIT est le meilleur prédicteur de la réussite scolaire, et cela indépendamment de la variabilité entre les 4 indices factoriels (ICV, IRP, IMT, IVT). En résumé, le QIT est un prédicteur robuste de la réussite académique (démarche pronostique), que l'hétérogénéité entre les indices factoriels soit importante ou non.

1.3.L'Indice d'Aptitude général (IAG) comme mesure du facteur g ?

C'est en 1998 que Prifitera, Weiss et Saklofske ont proposé de décrire le niveau de fonctionnement intellectuel à partir de l'Indice d'Aptitude Général. Il s'agissait notamment pour ces auteurs de répondre à des problèmes pratiques liés à l'utilisation du QIT. En effet, des enfants souffrant de troubles de l'attention ou des troubles des apprentissages présentaient des faiblesses dans les sous-tests Arithmétique et Code et voyaient leur QIT diminué. Prifitera et al. ont suggéré de retirer ces 2 épreuves, particulièrement sensibles à ces troubles. L'IAG devait alors permettre une évaluation plus correcte du fonctionnement intellectuel général de ces enfants. Avec le WISC-IV, l'intérêt de l'IAG a été renforcé dans la mesure où les troubles de l'attention ou des apprentissages vont apparaître principalement dans les indices de mémoire de travail (IMT) et de vitesse de traitement (IVT). Ainsi, l'IAG, basé uniquement sur les contenus de pensée (ICV et IRP) serait un indicateur plus adapté à l'estimation du fonctionnement cognitif des enfants à Haut Potentiel Intellectuel (HPI) ou des enfants souffrant de troubles de l'attention ou des apprentissages (Grégoire, 2009 ; Lecerf, Reverte et al. 2010).

Le calcul de l'IAG a récemment été complété par l'indice de Compétence Cognitive (ICC, *Cognitive Proficiency Index*) destiné à évaluer les processus psychologiques « de base », que sont la vitesse de traitement et la mémoire de travail (Weiss, Saklofske, Schwartz, Prifitera, & Courville, 2006). Les travaux réalisés en psychologie cognitive différentielle et

en psychologie développementale indiquent effectivement que la mémoire de travail et la vitesse de traitement seraient deux processus de base de la cognition (Ackerman, Beier, & Boyle, 2005 ; Kane, Hambrick, & Conway, 2005 ; Lecerf & Roulin, 2006 ; Oberauer, Schulze, Wihlem, & Süß, 2005). Plus précisément, on fait l'hypothèse que la vitesse de traitement influencerait les activités cognitives, et notamment l'intelligence fluide (Gf) via la mémoire de travail (de Ribaupierre & Lecerf, 2006). Les individus capables de traiter rapidement les stimuli maintiendraient simultanément activés en mémoire de travail davantage d'informations. A l'inverse, des individus dont la capacité à traiter des informations serait faible auraient moins d'informations stockées simultanément en mémoire de travail.

Au plan théorique, il est possible de positionner les indices IAG et ICC dans le cadre des modèles classiques du traitement de l'information et du modèle de Cattell-Horn-Carroll (CHC). Ainsi, Woodcock (1997 ; pour une synthèse : Floyd, 2005, p. 220) a proposé un modèle du traitement de l'information Gf-Gc (*Gf-Gc Information Processing model*). Woodcock intègre 10 aptitudes globales du modèle CHC (Gf, Gc, Gv, etc.) dans une perspective de traitement de l'information. En premier lieu, la mémoire de travail et la vitesse de traitement sont regroupées au sein de « l'efficacité cognitive » (*cognitive efficiency*). Ce regroupement correspond donc au calcul de l'indice de compétence cognitive (ICC). Quant à l'IAG, il regroupe d'une part les « aptitudes à la pensée » (*Thinking abilities* : Traitement visuel (Gv), intelligence fluide (Gf), traitement auditif (Ga), récupération et stockage à long terme (Glr)) et d'autre part, les connaissances acquises (*stores of acquired knowledge* : intelligence cristallisée (Gc), connaissances quantitatives (Gq) et lecture – écriture (Grw)). C'est la raison pour laquelle Sattler (2009) propose de renommer l'IAG et de parler d'Indice de Résolution de Problèmes et de Connaissances (IRPC) ; de même, l'ICC devrait être dénommé Indice de Vitesse Psychomotrice et d'Attention Soutenue (IVPAS).

Pour la version Nord-américaine du WISC-IV, il existe différentes sources pour les tables IAG, qui diffèrent par la méthode utilisée. Certaines ont été développées à partir de l'échantillon de standardisation (Saklofske, Prifitera, Weiss, Rolfhus, & Zhu, 2005), tandis que d'autres normes ont été créées en utilisant la procédure d'approximation statistique de Tellegen et Briggs (1967 ; Flanagan & Kaufman, 2004). En ce qui concerne la version Française, nous avons récemment développé les normes de l'IAG pour le WISC-IV en utilisant cette même procédure d'approximation statistique (Lecerf, Reverte et al. 2010). Toutefois, Saklofske, Weiss, Raiford, et Prifitera (2006) indiquent que les normes développées à partir de la formule de Tellegen et Briggs sous-estimeraient les performances dans la partie supérieure de la distribution et surestimeraient les performances dans la partie inférieure de la distribution. Selon Raiford et al. (2005), cette différence entre les normes « statistiques » et les normes « empiriques » serait en moyenne de 2 - 3 points, mais pourrait aller jusqu'à 6 points pour les scores très bas ou très élevés (c.-à-d. retard mental, haut potentiel). Bien que ces auteurs ne proposent aucune explication, on peut faire l'hypothèse que des « effets statistiques » pourraient être à l'origine de la surestimation et de la sous-estimation des scores statistiques relativement aux scores « empiriques ». Les notes ICV et IRP sont sommées et ré-étalonnées selon la loi normale en fonction de leur corrélation. Or, tous les enfants ne réussissent pas à la fois dans le domaine de la compréhension verbale et dans le domaine du raisonnement perceptif ; à l'inverse, tous les enfants n'échouent pas en même temps dans ces 2 domaines. On peut également faire l'hypothèse que la surestimation et la sous-estimation des scores statistiques sont liées à un phénomène de régression à la moyenne. Pourtant, Tellegen et Briggs (1967, p. 503) indiquent que leur procédure, contrairement à celle proposée par Doppelt (1956), permet d'éviter une sous-estimation ou une surestimation systématique selon que l'on se situe au-dessus ou au-dessous de la

moyenne. En d'autres termes, selon Tellegen et Briggs, ce phénomène de régression ne devrait pas intervenir dans leur procédure.

Les objectifs de cette recherche sont de proposer des normes empiriques des indices IAG et ICC pour une population d'enfants suisse francophone, et d'évaluer les différences quantitatives et qualitatives d'avec les normes « statistiques ». Il s'agit notamment de tester l'hypothèse selon laquelle les différences entre les deux types de tables IAG sont plus importantes dans les parties extrêmes de la distribution. On cherchera également à déterminer si cette hypothèse s'applique aux tables ICC. Autrement dit, cet article se donne pour objectif de proposer des normes « empiriques » des scores IAG et ICC, et de comparer ces tables empiriques avec les tables statistiques existantes. En cela, il fait suite à notre précédente étude qui portait uniquement sur l'IAG (Lecerf, Reverte et al., 2010). Ainsi, la première caractéristique de cette étude, par rapport à la recherche précédent, est qu'elle inclue l'analyse de l'ICC et porte sur un échantillon plus conséquent (182 au lieu de 60). Cette étude se démarque également de la précédente, car il s'agit de déterminer pour l'IAG et l'ICC si les différences entre les 2 types de tables sont plus importantes dans les parties extrêmes de la distribution.

Méthode

Participants et épreuves³

Tous les sous-tests du WISC-IV (principaux et optionnels) ont été administrés à cent quatre-vingt deux enfants âgés de 8 à 12 ans (âge moyen = 10,12 ans ; écart-type = 1,17). L'échantillon comporte 89 filles et 93 garçons. Tous les participants se trouvent dans le degré scolaire correspondant à leur âge chronologique. La sélection d'enfants entre 8 et 12 ans s'explique par le fait que cette étude s'inscrit dans le cadre d'un projet de recherche plus large (financé par le Fonds National Suisse de la Recherche Scientifique), dans lequel des analyses

factorielles confirmatoires multigroupes seront réalisées. Il s'agira de comparer la structure factorielle des enfants de 8-9 ans à celle d'enfants de 10-11 ans selon le modèle CHC.

*Procédure*⁴

Les passations des sous-tests du WISC-IV se sont déroulées de manière individuelle durant les heures scolaires dans les locaux de différentes écoles du Canton de Genève. La passation et la cotation ont été réalisées selon la procédure du manuel d'administration et de cotation (Wechsler, 2005).

Résultats

Indiquons tout d'abord que les performances moyennes observées pour cet échantillon sont proches des moyennes (100) et des écart-types attendus (15). Le score moyen du QIT est de 105 et les scores des indices varient entre 96 (IMT) et 108 (ICV). Les écart-types varient entre 13 et 14 points pour les 4 indices factoriels et le QIT.

Les scores « statistiques » IAG_S et ICC_S de chaque participant ont été identifiés à partir de la procédure et des tables que nous avons développées (voir Lecerf, Reverte et al, 2010, pour les tables IAG). Rappelons brièvement que la procédure d'approximation statistique pour IAG utilise la formule suivante : $IAG_S = 0,575(ICV + IRP) - 15,09$. Pour le score ICC, la formule est la suivante : $ICC_S = 0,63(IMT + IVT) - 25,49$.

En ce qui concerne le développement des tables empiriques IAG (IAG_E), la première étape a consisté à calculer la somme des notes standards des sous-tests de Compréhension Verbale et de Raisonnement Perceptif (Similitudes, Vocabulaire, Compréhension, Cubes, Matrices, et Identification de Concepts). Les tables empiriques ICC ont été calculées en sommant les notes standards des sous-tests de Mémoire de travail et de Vitesse de traitement (Mémoire de chiffres, Séquence lettres-chiffres, Code, et Symboles). Ces sommes des notes standards (IAG et ICC) ont ensuite été centrées sur une moyenne de 100 et un écart-type de

15. Cette procédure permet ainsi d’obtenir un score IAG et ICC « empirique » pour chaque participant.

Résultats de l’Indice d’Aptitude Général

Le tableau 1 rapporte les moyennes (écart-types) et les corrélations entre les normes statistiques (IAG_S) et empiriques (IAG_E) pour l’échantillon total. Les différences moyennes absolues ($|IAG_S - IAG_E|$) et relatives ($IAG_S - IAG_E$) entre les normes statistiques et empiriques sont équivalentes et sont de 5,38 (e-t = 2,04). On constate que la moyenne du score IAG_S est proche de celle du QIT. Dans l’ensemble, ces résultats indiquent que le score « statistique » (IAG_S) est supérieur au score « empirique » (IAG_E), et que les trois scores corrélaient fortement.

— Insérer Tableau 1 —

Dans un deuxième temps, nous avons testé l’hypothèse selon laquelle les normes statistiques sous-estimeraient les scores dans la partie supérieure de la distribution, tandis qu’elles les surestimeraient dans la partie inférieure de la distribution. Pour tester cette hypothèse, sept groupes de participants ont été constitués sur la base de leurs scores IAG empiriques (IAG_E), et cela conformément à la classification proposée dans le WISC-IV :

Groupe 1 : $IAG \leq 69$ (N= 5) ; groupe 2 : IAG entre 70 - 79 (N = 12) ; groupe 3 : IAG entre 80 - 89 (N= 24) ; Groupe 4 : IAG entre 90 - 109 (N = 90) ; groupe 5 : IAG entre 110 - 119 (N = 37) ; Groupe 6 : IAG entre 120 - 129 (N = 6) ; groupe 7 : IAG : 130 et plus (N = 8). Le tableau 2 rapporte les moyennes (écart-types) des normes statistiques (IAG_S) et empiriques (IAG_E) pour les 7 groupes. Les différences absolues entre les normes statistiques et empiriques sont également présentées. Il faut souligner que les différences relatives sont strictement identiques aux différences absolues ; en conséquence, elles n’ont pas été rapportées pour des raisons de simplification. Les différences absolues ont l’avantage de

sommer les différences indépendamment du sens de la différence. Autrement dit, les différences en faveur de IAG_S ne sont pas « contrebalancées » par les différences en faveur de IAG_E. Ces résultats indiquent que globalement le score « statistique » est supérieur au score « empirique », et cela quel que soit le groupe. Une analyse de la variance a été réalisée sur les facteurs Groupes (facteur inter-sujet : 7 groupes) et Scores (facteur intra-sujet : IAG_S – IAG_E)⁵. L'interaction, significative, $F(6, 175) = 119,96, p < ,01, \eta_p^2 = ,80$, a été décomposée à l'aide de la correction de Bonferroni, et montre que toutes les comparaisons sont significatives. Ainsi, pour tous les groupes, l'IAG_S est supérieur à l'IAG_E. De même, les différences entre les groupes sont toutes significatives, et cela aussi bien pour IAG_S que pour IAG_E (G7 > G6 > ... > G2 > G1). Pour terminer, les comparaisons entre IAG_S et IAG_E au sein de chaque groupe permettent de constater que les tailles d'effets (η_p^2) varient entre ,92 et ,99. Ces résultats sont à prendre avec beaucoup de précaution compte tenu de la taille réduite des effectifs de certains groupes et de la très forte hétérogénéité de ces mêmes effectifs entre les groupes.

Notons que les différences relatives entre les normes statistiques et les normes empiriques du score IAG ont également été testées pour des groupes constitués sur la base de leur QIT, conformément à la classification proposée dans le WISC-IV. Les résultats indiquent que le score « statistique » est supérieur au score « empirique », et cela quel que soit le groupe. L'amplitude des différences varie de 11 points pour le groupe 2 (QIT : 70 – 79) à 3 points pour le groupe 7 (QIT > 130).

— Insérer Tableau 2 —

Dans un troisième temps, le pourcentage d'accord de classification réalisé à partir du score IAG_S et IAG_E a été calculé. On constate que 63,2 % des enfants sont classés dans le même groupe pour les deux indices. Parmi les 36,8 % d'enfants restants, on remarque que dans tous les cas le score IAG_S est supérieur d'une catégorie au score IAG_E. Autrement dit,

115 enfants sont situés exactement dans le même groupe à partir des normes IAG_S et IAG_E tandis que 67 enfants présentent un IAG_S supérieur à IAG_E d'une catégorie. Sur le plan quantitatif, parmi les enfants présentant un désaccord dans la classification, 11 enfants (16,4 %) présentent une différence entre 3 et 4 points ; 50 enfants (74,63 %) présentent une différence entre 5 et 9 points ; enfin, 6 enfants (8,96 %) présentent une différence égale ou supérieure à 10 points (maximum de 14 points de différence pour 2 enfants).

Résultats de l'Indice de Compétence Cognitive

Des analyses similaires ont été réalisées pour les indices ICC. Le tableau 3 rapporte les moyennes (écart-types) et les corrélations entre les normes statistiques (ICC_S) et empiriques (ICC_E) pour l'échantillon total. Les différences absolues ($| ICC_S - ICC_E |$) entre les normes statistiques et empiriques sont de 2,24 points (e-t = 1,85), tandis que les différences relatives (ICC_S – ICC_E) sont de 1,77 points (e-t = 2,30). Globalement le score « statistique » est très légèrement supérieur au score « empirique », et les trois scores corrèlent fortement.

— Insérer Tableau 3 —

Dans un deuxième temps, nous avons cherché à déterminer si l'hypothèse selon laquelle les normes statistiques sous-estiment les scores dans la partie supérieure de la distribution, tandis qu'elles les surestiment dans la partie inférieure de la distribution, s'applique aux scores ICC. Sept groupes de participants ont été constitués sur la base de leurs scores ICC empiriques, et cela conformément à la classification proposée dans le WISC-IV : Groupe 1 : ICC ≤ 69 (N= 5) ; groupe 2 : ICC entre 70 - 79 (N = 9) ; groupe 3 : ICC entre 80 - 89 (N= 33) ; Groupe 4 : ICC entre 90 - 109 (N = 95) ; groupe 5 : ICC entre 110 - 119 (N = 25) ; Groupe 6 : ICC entre 120 - 129 (N = 8) ; groupe 7 : ICC : 130 et plus (N = 7). Le tableau 4 rapporte les moyennes (écart-types) des normes statistiques (ICC_S) et empiriques

(ICC_E) pour les 7 groupes. Les différences absolues et relatives entre les normes statistiques et empiriques sont également présentées.

Une analyse de la variance a été réalisée sur les facteurs Groupes (facteur inter-sujet : 7 groupes) et Scores (facteur intra-sujet : ICC_S – ICC_E)⁵. L'interaction, significative, $F(6, 175) = 187,52, p < .01, \eta_p^2 = ,87$, a été décomposée à l'aide de la correction de Bonferroni, et montre que toutes les comparaisons sont significatives, à l'exception de la différence entre ICC_S et ICC_E pour le groupe 5. Pour les groupes 1, 2, 3 et 4, ICC_S est supérieur à ICC_E. En revanche, pour les groupes 6 et 7, ICC_E est supérieur à ICC_S. Les différences entre les groupes sont toutes significatives, et cela aussi bien pour ICC_S que pour ICC_E (G7 > G6 > ... > G2 > G1). Pour terminer, les comparaisons entre ICC_S et ICC_E au sein de chaque groupe permettent de constater que les tailles d'effets (η_p^2) varient entre ,16 pour le groupe 5 à ,987 pour le groupe 2. Comme indiqué précédemment, ces résultats sont à prendre avec beaucoup de précaution.

Les différences relatives entre les normes statistiques et les normes empiriques du score ICC ont également été testées pour des groupes constitués sur la base de leur QIT, conformément à la classification proposée dans le WISC-IV. Les résultats indiquent que le score ICC « statistique » est supérieur au score « empirique » pour les groupes inférieurs et moyens (jusqu'au groupe 4), tandis que le score « empirique » est équivalent au score « statistique » pour les groupes supérieurs (groupes 5 à 7).

— Insérer Tableau 4 —

Dans un troisième temps, le pourcentage d'accord de classification réalisé à partir du score ICC_S et ICC_E a été calculé. On constate que 83 % des enfants (soit 151) sont classés dans le même groupe à partir des 2 tables. Parmi les 17 % d'enfants discordants, on constate que pour 87,10 % (N = 27) le score ICC_S est supérieur d'une catégorie au score ICC_E, tandis que pour les 12,90 % des enfants discordants restants le score ICC_E est supérieur d'une

catégorie au score ICC_S. Il faut noter que la supériorité du score ICC_S sur l'indice ICC_E concerne des enfants appartenant aux groupes 1 à 4, tandis que la supériorité de l'indice ICC_E sur ICC_S s'observe pour des enfants appartenant aux groupes 6 et 7. Sur le plan quantitatif, parmi les enfants présentant un désaccord dans la classification (31 enfants), 21 présentent une différence absolue entre 1 et 4 points (67,7 %), 9 enfants présentent une différence comprise entre 5 et 9 points (30 %), et 1 enfant présente une différence de 10 points entre les 2 scores ICC. En ce qui concerne les différences relatives, 4 enfants présentent une différence entre -3 et -1 points, 17 enfants présentent une différence entre 1 et 4 points, 9 enfants présentent une différence entre 5 et 9 points et enfin, 1 enfant présente une différence de 10 points.

Discussion

La recherche présentée avait pour objectif de proposer des normes « empiriques » pour les Indices d'aptitude Général (IAG) et de Compétence Cognitive (ICC), et de comparer ces normes empiriques (IAG_E – ICC_E) aux normes « statistiques » (IAG_S – ICC_S). Les normes empiriques ont été établies à partir d'un échantillon de 182 enfants Genevois « tout venant » âgés de 8 à 12 ans.

En premier lieu, on observe de très fortes corrélations entre les deux types de tables, puisque la corrélation est de 0,997 entre IAG_S et IAG_E, et de 0,999 entre ICC_S et ICC_E. De même, la corrélation entre IAG et QIT est très élevée ($r = 0,92$ pour IAG_S et IAG_E). Au passage, notons que cette très forte corrélation ne signifie pas que ces deux scores sont strictement équivalents. Sur le plan théorique, et selon le modèle CHC, l'IAG inclut l'intelligence cristallisée (Gc), l'intelligence fluide (Gf) et le traitement visuel (Gv) ; les composantes de mémoire à court terme (Gsm) et de vitesse de traitement sont donc exclus (Gs). Or, ces deux dernières composantes représentent 40 % du calcul du QIT. En outre, rappelons que les corrélations traduisent la stabilité des différences interindividuelles, mais ne

fournissent aucune information sur les variations des niveaux de performance. Sur le plan clinique, et comme indiqué précédemment, Prifitera et al. (1998) ont proposé l'IAG pour répondre à des problèmes pratiques liés à l'utilisation du QIT. Par exemple, les enfants souffrant de troubles des apprentissages ou de troubles de l'attention peuvent voir leur QIT diminué à cause de leurs faibles performances en Mémoire de travail (IMT) et/ou en Vitesse de traitement (IVT). De même, les enfants à Haut potentiel intellectuel présentent un QIT déprimé en raison de ces mêmes indices de Mémoire de travail et de Vitesse de traitement. Dès lors, l'IAG permettrait une évaluation plus correcte de leur niveau de fonctionnement intellectuel et sera plus élevé que le QIT. A l'inverse, les enfants présentant un QIT faible (70-80) tendront à obtenir un IAG inférieur au QIT. Ces enfants ont tendance à présenter relativement de meilleures performances pour les indices de Mémoire de travail et de Vitesse de traitement, ce qui relève leur QIT.

En ce qui concerne les différences absolues entre les 2 types de tables, elles sont en moyenne de 5,38 pour l'IAG et de 2,24 pour l'ICC. Les résultats indiquent également que les enfants sont catégorisés de manière relativement identique à partir des 2 types de tables ICC. En effet, 83 % des enfants sont classés dans la même catégorie. Pour les enfants discordants, si le score ICC_E est supérieur à 120 (groupes 6 et 7), la classification réalisée à partir des normes ICC_E est supérieure à celle obtenue à partir des tables ICC_S ; l'inverse est vrai si le score ICC_E est compris entre 69 et 109 (groupes 1 à 4). Parmi les enfants présentant un désaccord, la différence quantitative entre ICC_S et ICC_E varie entre 1 et 9 points pour 96,77 % de ces enfants. Notons que des analyses complémentaires montrent que 48,35% des enfants sont catégorisés de manière identique à partir du score ICC_E et du score QIT. Pour les enfants discordants, la classification réalisée à partir du score ICC_E est supérieure à celle réalisée à partir du QIT pour 11,54 % ; pour les 40,11 % d'enfants restants la classification réalisée à partir du QIT est donc supérieure à celle effectuée à partir du score ICC_E .

En ce qui concerne l'IAG, le pourcentage d'accord entre les deux types de table IAG est plus faible, puisqu'il est de 63,2 %. On constate que pour tous les enfants discordants (36,8 %), le score IAG_S est supérieur d'une catégorie au score IAG_E. En outre, les résultats montrent que seulement 8,96 % des enfants discordants présentent une différence égale ou supérieure à 10 points. La comparaison entre les classifications réalisées à partir du score IAG_E et du QIT indique que 54,4 % sont positionnés dans la même catégorie. Pour 35,71 % des enfants, la catégorie donnée à partir du QIT est supérieure à celle du score IAG_E ; pour les 9,89 % d'enfants restant, la catégorie IAG est donc supérieure à celle indiquée par le QIT.

Rappelons qu'en ce qui concerne le score IAG, Saklofske et al. (2006) ont indiqué que les tables statistiques (IAG_S) sous-estimeraient les performances dans la partie supérieure de la distribution et surestimeraient les performances dans la partie inférieure de la distribution (et cela contrairement aux écrits de Tellegen et Briggs, 1967). L'objectif principal de cette étude était donc de tester cette hypothèse pour le score IAG, mais également pour le score ICC. Pour cela, sept groupes d'enfants ont été créés à partir de leurs scores IAG_E et ICC_E, et cela conformément à la classification des QIT proposée dans le WISC-IV. En ce qui concerne le score IAG, nos résultats valident partiellement l'hypothèse. Si les normes statistiques surestiment les performances dans la partie inférieure de la distribution (12,40 points pour le groupe 1 et 8,42 points pour le groupe 2), on n'observe pas, en revanche, de sous-estimation des performances dans la partie supérieure de la distribution. Pour les groupes 6 et 7 (IAG_E > 120), les normes IAG_S continuent de donner un score plus élevé que les normes IAG_E, même si la différence est plus petite (2 – 3 points). Ainsi, l'IAG_S surestime toujours les performances par rapport au score IAG_E. Aussi, dans le cas d'enfants présentant un score IAG_S inférieur à 89, nous recommandons de considérer que le niveau cognitif des enfants est probablement inférieur d'environ 8 – 10 points par rapport à celui indiqué par IAG_S. En ce qui concerne le score ICC, l'hypothèse initiale est davantage validée.

En effet, on observe bien une surestimation du score statistique par rapport au score empirique dans la partie inférieure de la distribution, et une sous-estimation des normes statistiques sur les normes empiriques dans la partie supérieure de la distribution. En conséquence, dans le cas d'enfants présentant un score ICC_S inférieur à 89, nous recommandons de considérer que le niveau cognitif des enfants est probablement inférieur d'environ 5 – 6 points par rapport à celui indiqué par ICC_S. A l'inverse, dans le cas d'enfants présentant un score ICC_S supérieur à 120, nous recommandons de considérer que le niveau cognitif des enfants est probablement supérieur d'environ 2 – 3 points par rapport à celui indiqué par ICC_S. Soulignons que la limite principale de cette recherche, tient à la taille de l'échantillon, qui est relativement modeste (N = 182), et notamment dans les groupes extrêmes. Il est indiscutable que des normes empiriques établies sur un échantillon de taille plus conséquente seraient souhaitables pour l'ICC ⁶.

Ainsi, l'hypothèse générale postulait que les tables statistiques sous-estimeraient les performances dans la partie supérieure de la distribution et surestimeraient les performances dans la partie inférieure de la distribution. Les résultats montrent que cette hypothèse est davantage validée pour l'indice de Compétence Cognitive (ICC) que pour l'indice d'Aptitude Général (IAG). Selon nous, cette différence entre IAG et ICC pourrait s'expliquer par les caractéristiques de notre échantillon. En effet, si la corrélation entre ICV et IRP est significative est modérée ($r = ,48$), en revanche, la corrélation entre IVT et IMT n'est pas significative ($r = ,125$ $p = ,09$). Plus important, la corrélation ICV – IRP est proche de celle rapportée pour les données d'étalonnage du WISC-IV ($r = ,51$; Wechsler, 2005), données que nous avons utilisées pour la création des tables IAG statistique. Par contre, la corrélation entre IMT et IVT est plus importante dans les données d'étalonnages françaises ($r = ,25$; Wechsler, 2005) que dans notre étude ($r = ,125$). Rappelons que les tables ICC statistiques ont été créées à partir de la corrélation entre IVT et IMT des données d'étalonnages ($r = ,25$).

Autrement dit, la corrélation modérée entre ICV et IRP indique que les enfants présentant de bonnes performances en compréhension verbale (ICV) tendent à avoir de bonnes performances en raisonnement perceptif (IRP) ; à l'inverse, les enfants présentant de faibles performances en ICV tendent également à avoir de faibles performances en IRP. En ce qui concerne l'ICC, l'absence de corrélations entre IMT et IVT ($r = ,125$), ainsi que les faibles performances des enfants de notre échantillon sur l'indice de mémoire de travail (IMT = 96) associées à de relativement bonnes performances sur l'indice de vitesse de traitement (IVT = 107) seraient à l'origine de ces phénomènes de régression à la moyenne et donc de la sous-estimation et de la surestimation systématique selon la position sur la distribution normale.

Conclusion

En résumé, si l'utilisation d'un score unique est en contradiction avec la conception multidimensionnelle de l'intelligence, il est en revanche légitime selon l'organisation hiérarchique proposée dans le modèle CHC. En outre, tous les auteurs s'accordent à reconnaître qu'une estimation globale basée sur le facteur g est nécessaire, et cela notamment dans le cadre d'une perspective pronostique. Aussi, nous recommandons d'utiliser le score IAG en complément du QIT dans le cadre du bilan psychologique. Pour cela, les tables IAG que nous avons proposées (Lecerf et al. 2010) peuvent être utilisées tant que celles développées à partir des données d'étalonnage ne sont pas disponibles ⁷. Comme le QIT, le score IAG est une moyenne non pondérée des scores des différents sous-tests, mais contrairement au QIT, l'IAG ne comprend pas les sous-tests dont les saturations en facteur g sont les plus faibles (Code et Symboles) ; il présente donc une saturation moyenne en facteur g légèrement plus élevée que celle du QIT. On pourra nous opposer que l'IAG n'est pas non plus la meilleure estimation du facteur g , car il n'est pas calculé à partir des 6 épreuves les plus saturées en facteur g . C'est la raison pour laquelle, nous avons calculé 2 autres indices statistiques : un QIT - pondéré (QIT-P) et un indice de facteur g (IFG). Le QIT-P est obtenu à

partir des sous-tests compris dans le calcul du QIT, mais qui ont été pondérés par leur saturation en facteur g (sur la base des valeurs rapportées par Grégoire, 2009). Il s'agit donc d'une moyenne pondérée. La corrélation entre le QIT et le QIT-P est de 0,989, tandis que la corrélation entre IAG et QIT-P est de 0,95 ; ces 2 corrélations diffèrent significativement l'une de l'autre. Cette dernière corrélation est donc légèrement supérieure à celle rapportée dans cette étude entre IAG et QIT (0,91). Le calcul de l'IAG se rapproche donc du calcul d'un QIT dans lequel la saturation en facteur g serait prise en compte (QIT-P). Enfin, nous avons calculé un indice de facteur g (IFG) : ce score moyen non pondéré a été calculé à partir des 6 sous-tests principaux les plus saturés en facteur g (Vocabulaire, Similitudes, Compréhension, Séquence Lettres - Chiffres, Matrices et Cubes). Les corrélations entre cet indice IFG et les scores IAG et QIT-P sont de 0,96, tandis que la corrélation avec le QIT est de 0,92. Pris dans leur ensemble, ces résultats indiquent que l'indice IAG représente la réponse la plus simple à mettre en œuvre pour tenir compte des saturations en facteur g . En outre, l'utilisation de l'indice complémentaire de Compétence Cognitive (ICC) permet d'estimer la contribution des mécanismes de base de la cognition que sont la mémoire de travail et la vitesse de traitement. Les futures recherches doivent maintenant porter sur l'étude de la validité prédictive des scores IAG et ICC.

Références

- Ackerman, P. L., Beier, M. E., & Boyle, M. O. (2005). Working memory and intelligence: The same or different constructs? Psychological Bulletin, *131*, 30-60.
- Binet, A., & Simon, T. (1905). Méthodes nouvelles pour le diagnostic du niveau intellectuel des anormaux. L'Année Psychologique, *11*, 191-244.
- Canivez, G. L., & Watkins, M. W. (2001). Long-term stability of the Wechsler Intelligence Scale for Children-Third Edition among students with disabilities. School Psychology Review, *30*, 438-453.
- Carroll, J. B. (1993). Human cognitive abilities: A survey of factor-analytic studies. New York: Cambridge University Press.
- Doppelt, J. E. (1956). Estimating the Full Scale score on the Wechsler Adult Intelligence Scale from scores on four subtests. Journal of Consulting Psychology, *20*, 63-66.
- de Ribaupierre, A., & Lecerf, T. (2006). Relationships between working memory and intelligence from a developmental perspective: Convergent evidence from a neo-Piagetian and a psychometric approach. European Journal of cognitive Psychology, *18*, 109-137.
- Flanagan, D. P., & Kaufman, A. S. (2004). Essentials of WISC-IV assessment. Hoboken, New Jersey: John Wiley, & Sons, Inc.
- Floyd, R. G. (2005). Information-processing approaches to interpretation of contemporary intellectual assessment instruments. In D. P. Flanagan & P. L. Harrison (Eds.), Contemporary intellectual assessment. Theories, tests, and issues (pp. 203-233). New York: The Guilford Press.
- Freberg, M. E., Vandiver, B. J., Watkins, M. W., & Canivez, G. L. (2008). Significant Factor Score Variability and the Validity of the WISC-III Full Scale IQ in Predicting Later Academic Achievement. Applied Neuropsychology, *15*, 131-139.
- Glutting, J. J., Youngstrom, E. A., Ward, T., & Ward, S. (1997). Incremental Efficacy of WISC-III Factor Scores in Predicting Achievement: What Do They Tell Us? Psychological Assessment, *9*, 295-301.
- Grégoire, J. (2009). L'examen clinique de l'intelligence de l'enfant. Fondements et pratique du WISC-IV. 2ème édition revue et complétée. Wavre, Belgique: Mardaga.
- Grégoire, J., & Wierzbicki, C. (2009). Comparaison de quatre formes abrégées de l'échelle d'intelligence de Wechsler pour adultes – troisième édition (WAIS-III). Revue Européenne de Psychologie appliquée, *59*, 17-24.
- Horn, J. L., & Cattell, R. B. (1967). Age differences in fluid and crystallized intelligence. Acta Psychologica, *26*, 107-129.
- Huteau, M., & Lautrey, J. (1999). Evaluer l'intelligence. Psychométrie cognitive. Paris: PUF.
- Kane, M. J., Hambrick, D. Z., & Conway, A. R. A. (2005). Working memory capacity and

- fluid intelligence are strongly related constructs: Comments on Ackerman, Beier, and Boyle (2005). Psychological Bulletin, 131, 66-71.
- Lautrey, J. (2005). Le Q. I : concept mal compris ou concept dépassé ?, A. N. A. E. 17, pp. 146–149.
- Lecerf, T., Reverte, I., Coleaux, L., Favez, N., & Rossier, J. (2010). Indice d'aptitude général pour le WISC-IV : normes françaises. Pratiques psychologiques, 16, 109-121.
- Lecerf, T., Rossier, J., Favez, N., Reverte, I., & Coleaux, L. (2010). The four- vs. alternative six-factor structure of the French WISC-IV: comparison using Confirmatory Factor Analyses. Swiss Journal of Psychology, 69, 221-232.
- Lecerf, T., & Roulin, J.-L. (2006). Distinction between visuo-spatial short-term-memory and working memory span tasks. Swiss Journal of Psychology, 65, 37-54.
- McGrew, K. S. (2009). CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. Intelligence, 37, 1-10.
- Oberauer, K., Schulze, R., Wihlem, O., & Süß, H.-M. (2005). Working memory and intelligence - Their correlation and their relation: Comment on Ackerman, Beier, and Boyle (2005). Psychological Bulletin, 131, 61-65.
- Prifitera, A., Weiss, L. G., & Saklofske, D. (1998). The WISC-III in the context. In A. Prifitera & D. Saklofske (Eds.), WISC-III. Clinical use and interpretation: Scientist-practitioner perspective (pp. 1-38). New York: Academic Press.
- Raiford, S. E., Weiss, L. G., Rolfhus, E., & Coalson, D. (2005). General aptitude index (WISC-IV Technical report #4). Retrieved from <http://www.ksde.org/Portals/0/Special%20Education%20Services/gifted/WISCIVTechReport4.pdf>
- Rozencwajg, P. (2006). Quelques réflexions sur l'évaluation de l'intelligence générale : un retour à Binet ? Some questions on the evaluation of general intelligence: a flashback to Binet? Pratiques Psychologiques, 12, 395-410.
- Rozencwajg, P., Aliamer, V., & Ombredane, E. (2009). Le fonctionnement cognitif d'enfants atypiques à travers leur QI. Pratiques Psychologiques, 15, 343-365.
- Saklofske, D. H., Prifitera, A., Weiss, L. G., Rolfhus, E., & Zhu, J. (2005). Clinical interpretation of the WISC-IV FSIQ and GAI. In A. Prifitera, D. H. Saklofske & L. G. Weiss (Eds.), WISC-IV. Clinical use and interpretation. Scientist-practitioner perspectives (pp. 33-69). San Diego, USA: Elsevier, Inc.
- Saklofske, D., Weiss, L. G., Raiford, S. E., & Prifitera, A. (2006). Advanced interpretive issues with the WISC-IV full-scale IQ and General ability index scores. In L. G. Weiss, D. Saklofske, A. Prifitera & J. Holdnack (Eds.), WISC-IV. Advanced clinical interpretation (pp. 99-138). San Diego: Elsevier Inc.
- Sattler, J. M. (2001). Assessment of children. Cognitive applications (Fourth Edition ed.): Jerome Sattler, Publishers, Inc.

- Sattler, J. M., & Ryan, J. J. (2009). Assessment with the WAIS-IV. La Mesa, California: Jerome M. Sattler, Publisher, Inc.
- Spearman, C. (1927). The abilities of man: Their nature and measurement. New York: Mac Millan.
- Stern, W. (1912). Die Psychologische methoden der intelligenzprufung. Barth, Leipzig.
- Tellegen, A., & Briggs, P. F. (1967). Old wine in new skins: Grouping Wechsler subtests into new scales. Journal of Consulting Psychology, 31, 499-506.
- Thurstone, L. L. (1938). Primary mental abilities. University of Chicago Press, Chicago.
- Watkins, M. W. (2006). Orthogonal higher order structure of the Wechsler Intelligence Scale for Children - Fourth edition. Psychological Assessment, 18, 123-125.
- Wechsler, D. (1939). The measurement of adult intelligence. Baltimore: Williams & Wilkins. p. 229.
- Wechsler, D. (2005). Manuel de l'Echelle d'Intelligence de Wechsler pour Enfants - 4e édition. Paris : Editions du Centre de Psychologie Appliquée.
- Weiss, L. G., Saklofske, D. H., Schwartz, D. M., Prifitera, A., & Courville, T. (2006). Advanced clinical interpretation of WISC-IV index scores. In L. G. Weiss, D. Saklofske, A. Prifitera & J. Holdnack (Eds.), WISC-IV. Advanced clinical interpretation (pp. 139-179). San Diego: Elsevier Inc.
- Woodcock, R. W. (1997). The Woodcock-Johnson tests of cognitive ability - Revised. In D. P. Flanagan, J. L. Genshaft & P. L. Harrison (Eds.), Contemporary intellectual assessment. Theories, tests, and issues (pp. 230-246). New York: The Guilford Press.

Notes de bas de page

¹ La distinction entre échelle Verbale (QIV) et échelle Performance (QIP) proposée par Wechsler, sur la base du modèle d'Alexander, peut être considérée comme la 1^{ère} étape vers une conception multidimensionnelle. Toutefois, la suppression du QIV et du QIP est liée à leur faible pertinence clinique.

² Plus précisément, 4 critères sont actuellement proposés pour évaluer la pertinence du QIT.

³ Cette recherche est soutenue par le Fonds National Suisse de la Recherche Scientifique dans le cadre d'un subside alloué à T. Lecerf, N. Favez, & J. Rossier (requête N° 100011-107764 : *Analysis of the french WISC-IV structure according to the Cattell-Horn-Carroll (CHC) narrow ability classification*).

⁴ Cette recherche respecte le code déontologique de la Société Suisse de psychologie (SSP) et de la Fédération Suisse des psychologues (FSP). L'autorisation d'administrer le WISC-IV aux enfants a été donnée par le Département de l'Instruction Public (DIP) du Canton de Genève et/ou par les représentants légaux des enfants.

⁵ Cette suggestion a été faite par l'un des experts.

⁶ Jacques Grégoire (2009), dans la révision de son livre sur le WISC-IV, vient de proposer les tables IAG sur la base des données d'étalonnages françaises ; mais les tables ICC ne sont pas proposées.

⁷ Notons que 96,2% des enfants sont classés dans la même catégorie à partir de l'IAG_S et de l'IAG proposé par Grégoire (2009) ; en terme de points, cela représente des différences entre -2 et +3 points.

Conflit d'intérêt : aucun