



Master

2011

Open Access

This version of the publication is provided by the author(s) and made available in accordance with the copyright holder(s).

L'évaluation des performances des enfants lors du bilan psychologique :
congruence des résultats à diverses épreuves censées évaluer les mêmes
aptitudes cognitives

Pellegrini, Fabien

How to cite

PELLEGRINI, Fabien. L'évaluation des performances des enfants lors du bilan psychologique : congruence des résultats à diverses épreuves censées évaluer les mêmes aptitudes cognitives. Master, 2011.

This publication URL: <https://archive-ouverte.unige.ch/unige:18138>



**UNIVERSITÉ
DE GENÈVE**

**FACULTÉ DE PSYCHOLOGIE
ET DES SCIENCES DE L'ÉDUCATION**

Mémoire de Master en Psychologie du Développement

**L'évaluation des performances des enfants lors du bilan
psychologique : congruence des résultats à diverses
épreuves censées évaluer les mêmes aptitudes cognitives**

Présenté par : Fabien Pellegrini

Dirigé par : Thierry Lecerf et Nicolas Favez

Membres du jury : Thierry Lecerf, Nicolas Favez et Isabelle Reverte

Année universitaire : 2010 - 2011

Email : pellegr0@etu.unige.ch

Téléphone : 022.301.46.69

Numéro d'étudiant : 02-305-860

Table des matières :

RESUME.....	2
1. INTRODUCTION GENERALE	3
2. INTRODUCTION THEORIQUE	4
2.1 L'ECHELLE D'INTELLIGENCE DE WECHSLER POUR ENFANTS (WISC).....	5
2.1.1 <i>Fondements</i>	5
2.1.2 <i>La quatrième version - Le WISC-IV</i>	7
2.2 STABILITE ET FLUCTUATION A COURT TERME DES PERFORMANCES	11
2.3 LE MODELE CATTELL-HORN-CARROLL (CHC).....	15
2.3.1 <i>Le modèle de Cattell et Horn</i>	15
2.3.2 <i>Le modèle de Carroll</i>	16
2.3.3 <i>Le modèle CHC des aptitudes cognitives</i>	17
2.4 LE WISC-IV SELON LE MODELE CHC - UNE STRUCTURE ALTERNATIVE	21
2.5 PROBLEMATIQUE.....	24
3. METHODE	28
3.1 PARTICIPANTS	28
3.2 MATERIEL	28
3.2.1 <i>L'échelle d'intelligence de Wechsler pour enfants - quatrième édition</i>	28
3.2.2 <i>La batterie d'évaluation de Kaufman pour enfants (K-ABC)</i>	36
3.2.3 <i>The Korkman Neuropsychological Scale (NEPSY)</i>	38
3.2.4 <i>The Woodcock-Johnson Psycho-Educational Battery Test – Revised (WJ-R)</i> ...	39
3.3 PROCEDURE.....	41
3.4 SYNTHÈSE	42
4. RESULTATS	43
4.1 STATISTIQUES DESCRIPTIVES.....	43
4.2 MESURES CENSEES EVALUER LA MEMOIRE A COURT TERME	46
4.3 MESURES CENSEES EVALUER LA VITESSE DE TRAITEMENT	47
4.4 MESURES CENSEES EVALUER LE TRAITEMENT VISUEL	49
4.5 MESURES CENSEES EVALUER L'INTELLIGENCE FLUIDE	51
4.6 MESURES CENSEES EVALUER L'INTELLIGENCE CRISTALLISEE	52
5. DISCUSSION	55
6. CONCLUSION.....	60
REFERENCES BIBLIOGRAPHIQUES :.....	62
ANNEXES :	65

Résumé

Cette recherche porte sur l'évaluation des performances cognitives à l'aide de la quatrième version française de l'*Echelle d'Intelligence de Wechsler pour enfants* (WISC-IV) et a pour origine le constat qu'elle est souvent administrée en plusieurs séances au lieu d'une seule, ce qui pose la question de la congruence des résultats aux épreuves censées évaluer les mêmes aptitudes cognitives. Ainsi, nous avons cherché à vérifier la stabilité des performances des enfants suisses-romands d'une séance à l'autre. Pour ce faire, nous avons analysé les performances d'un sous-échantillon de 105 enfants, âgés entre 8 et 12 ans, suite à l'administration du WISC-IV ainsi que de 5 autres tests, lors de trois séances. Nos interprétations sont basées sur le modèle le plus influent : la théorie de Cattell-Horn-Carroll (CHC). En effet, la structure factorielle du WISC-IV étant sujette à controverses, l'utilisation de ce modèle nous a aidé dans nos comparaisons et regroupements à effectuer.

Nos résultats ont montré que les épreuves qui ont été regroupées ensemble évaluaient bien les mêmes aptitudes cognitives. En effet, les corrélations étaient significatives, d'une intensité relativement élevée et proches de celles fournies par le manuel du WISC-IV. Enfin, au niveau des différences de performances, nous avons constaté que, pour la majorité des enfants, les performances restent stables. En effet, les variations moyennes absolues des performances sont inférieures à 4 points, seuil fourni par la littérature. Toutefois, on constate qu'un faible pourcentage d'enfants présente des fluctuations. Ces pourcentages étaient attendus et ne remettent pas en cause la stabilité. Ainsi, selon les paires d'épreuves, les différences concernaient 6,8% à 16,3% des enfants pour la *Mémoire à court terme* ; 12,6 à 30,1% pour la *Vitesse de traitement* ; 8,6 à 19,2% pour le *Traitement visuel* ; et 18,2 à 21,9% pour l'*Intelligence cristallisée*. Enfin, pour l'*Intelligence fluide*, nous avons constaté des fluctuations pour 5,7 à 53% des enfants. Néanmoins, il semble que cette différence soit due à une épreuve particulière qui posait problème.

En conclusion, nous considérons que les performances des enfants entre les séances restent stables et qu'administrer le WISC-IV sur plusieurs séances n'a pas d'influence.

1. Introduction générale

Cette recherche a pour intérêt la mesure de la stabilité des performances cognitives des enfants, à l'aide des échelles d'intelligence. Plus précisément, ce travail a pour origine le constat que, dans la pratique, l'évaluation cognitive se fait le plus souvent à l'aide de l'échelle d'intelligence de Wechsler pour les enfants et que son administration est souvent effectuée non pas en une seule séance, comme l'exige le manuel d'administration et de cotation, mais en plusieurs (deux au moins). En effet, la passation de l'ensemble de épreuves requiert beaucoup de temps (plus d'une heure pour les épreuves obligatoires), ce qui peut poser problème pour certains enfants s'ils présentent des difficultés majeures (ex : trouble de l'attention/hyperactivité, etc.). Ainsi, cela pose la question de l'influence que peut avoir l'étalement des passations dans le temps sur la congruence des résultats obtenus aux épreuves censées évaluer les mêmes aptitudes cognitives et, en définitive, sur les hypothèses et les prédictions des professionnels qui sont basées sur l'identification des forces et des faiblesses de l'enfant. En effet, la validité des hypothèses dépend de la stabilité des performances observées et le pronostic sera plus probable si les résultats sont similaires d'une session à l'autre.

Afin de tester cette stabilité des performances, nous avons administré l'intégralité des sous-tests du WISC-IV sur deux sessions. Ainsi que, lors d'une troisième session, 5 épreuves issues de trois autres batteries de tests (K-ABC, WJ-R, NEPSY). Par exemple, les épreuves *Cubes* (WISC-IV) de la première séance, *Complètement d'images* (WISC-IV) de la deuxième et *Triangles* (K-ABC) de la troisième, devraient indiquer des scores similaires. En effet, selon la littérature, les épreuve *Triangles* et *Cubes* évaluent les mêmes aptitudes cognitives. D'ailleurs, selon A. S. Kaufman et N. L. Kaufman (2004), on observe une corrélation de .66 entre ces deux épreuves. Ces observations nous amènent donc à penser que les différents sous-tests censés évaluer les mêmes aptitudes cognitives devraient être consistants et situer chaque enfant au même niveau de performances. Enfin, pour établir s'il existe bel est bien une différence de performances inter-séances, nous avons eu recours aux travaux de Grégoire (2006). En effet, selon cet auteur, on considère qu'une différence supérieure à 4 points entre deux épreuves regroupées sous un même facteur implique une fluctuation des performances.

Pour terminer, il est à souligner également que l'évaluation de cette stabilité ne pouvait se faire sans vérifier au préalable que les épreuves, regroupées sous chacun des facteurs de la batterie de tests, évaluent bien les mêmes aptitudes cognitives. Sur ce point, force et de constater que la quatrième version de cette batterie de tests (WISC-IV) demeure encore sujette

à controverses en ce qui concerne sa structure factorielle. Ainsi, les différents regroupement des épreuves, de même que leurs comparaisons et interprétations élaborées dans ce travail, ont été effectuées à la lumière du modèle théorique des aptitudes cognitives de Cattell-Horn-Carroll (CHC). Le choix de ce modèle repose sur le fait qu'il est devenu celui qui est le plus influent et qui a généré le plus de recherches dans le domaine des théories tentant d'expliquer le comportement cognitif humain (McGrew, 1997). D'ailleurs, selon la littérature, il a été montré qu'une structure alternative basée sur cette théorie est plus adaptée que la norme à quatre facteurs dont est actuellement composé le WISC-IV (Flanagan & Kaufman, 2004 ; Keith, Fine, Taub, Reynolds, & Kranzler, 2006 ; Lecerf, Rossier, Favez, Reverte, & Coleaux, 2010). Enfin, il est à noter également que les études menées sur le WISC-IV, et en particulier celles qui concernent la proposition d'une structure factorielle alternative selon le modèle CHC, se sont basées essentiellement sur la version américaine de la batterie de tests (Lecerf & al, 2010). Par conséquent, on ne peut être certains que les conclusions issues de ces évaluations soient valables pour des enfants suisses-romands.

2. Introduction théorique

Les définitions de l'intelligence sont nombreuses et ont beaucoup évolué jusqu'à nos jours. De manière générale, on la décrit comme étant « *la capacité d'adaptation à des situations nouvelles, capacité qui permet de connaître, de comprendre, d'apprendre* » (Huteau & Lautrey, 1999). Plus précisément, comme nous allons le voir dans ce travail, les premières théories de l'intelligence, comme celle de Spearman (1904), s'orientaient vers une approche unidimensionnelle. De nos jours, on postule plutôt une approche multidimensionnelle où l'intelligence serait décomposable en différentes aptitudes. On parle même de hiérarchie (Huteau et Lautrey, 1999). Cet aspect est important car les échelles de Wechsler ont également, en partie, suivi cette évolution conceptuelle (Lecerf, Reverte, Coleaux, Maillard, Favez, & Rossier, 2011).

Afin de répondre à la question de la stabilité des performances inter-séances des enfants, il est donc important, en premier lieu, de se pencher sur la description et les fondements de la quatrième version française de l'*Echelle d'intelligence de Wechsler pour enfant* (WISC-IV). En effet, on ne peut aborder notre problématique sans tenir compte de la théorie qui la sous-tend. Dans un deuxième temps, nous nous intéresserons à la stabilité des performances à proprement dite. Nous allons voir que plusieurs facteurs peuvent engendrer des fluctuations à court terme en ce qui concerne les performances des enfants. Enfin, nous présenterons une

description et les fondements du modèle CHC ainsi qu'une structure alternative du WISC-IV reposant sur cette théorie et sur laquelle nous nous sommes basé pour effectuer nos analyses.

2.1 L'échelle d'intelligence de Wechsler pour enfants (WISC)

Les échelles de David Wechsler (1896-1981), et notamment l'échelle d'intelligence pour enfants (WISC), sont les batteries cognitives les plus administrées au monde dans le cadre du bilan psychologique (Flanagan & Kaufman, 2004 ; Grégoire, 2006). Le WISC est, d'ailleurs, beaucoup utilisé en cas de difficultés majeures (déficience intellectuelle légère à moyenne, troubles du langage, déficience motrice, etc.) ou, au contraire, pour les enfants à haut potentiel. Wechsler lui-même considère qu'elle permet de fournir un point de départ essentiel dans l'approche psychologique globale de l'individu et ce, même si d'autres dimensions doivent être prises en considération par le psychologue comme la personnalité, les interactions avec l'environnement ou avec autrui, etc. En effet, elle permet de faire des hypothèses sur le fonctionnement cognitif de l'enfant âgé de 6 à 16 ans et 11 mois.

Avant d'aborder plus en profondeur les caractéristiques de la quatrième version, il est essentiel d'apporter en premier lieu un éclairage sur les fondements du WISC ainsi que sur son évolution. En effet, comme nous l'avons déjà mentionné, l'évaluation de la stabilité des performances dans notre étude est étroitement liée à la structure factorielle de cette batterie de tests ainsi qu'aux aspects théoriques sur lesquels elle repose.

2.1.1 Fondements

A l'origine, l'*Echelle d'Intelligence de Wechsler pour Enfants* (WISC) est une adaptation de l'*Echelle d'Intelligence de Wechsler-Bellevue* (1939) destinée aux adultes (Lecerf et al, 2011). Depuis sa version initiale, la structure du WISC a évolué au cours du temps jusqu'à sa quatrième version en 2003 (la version française date de 2005 et inclut des normes pour les enfants francophones du Canada). La première élaboration a pour origine le fait que l'auteur a été amené à être en désaccord avec la définition de l'intelligence de son époque. En effet, il considérait la théorie bifactorielle « g » et « s » de Charles Edward Spearman (1863-1945) comme étant trop réductrice. Pour Wechsler, la définition devait, au contraire, être abordée de manière plus étendue et être validée à travers des dimensions observables. Il est à noter qu'avant Spearman, la plupart des chercheurs ne pouvaient concevoir qu'il était possible de mesurer autre chose que la perception sensorielle. En effet, on considérait que les activités

mentales de haut niveau, comme le raisonnement ou encore le calcul, étaient des notions trop abstraites et que, par conséquent, leurs mesures étaient trop aléatoires (Huteau & Lautrey, 1999). Ainsi, Spearman fut le premier à proposer une théorie psychométrique, basée sur l'analyse factorielle, tentant d'expliquer l'organisation des aptitudes cognitives. Ses travaux l'on amené à développer une théorie bi-factorielle : un facteur général d'intelligence « g » et des facteurs spécifiques « s » (Grégoire, 2006 ; Huteau & Lautrey, 1999). Pour pouvoir valider cette théorie, il reprendra alors, en 1904, les travaux sur la mesure de l'intelligence de Sir Francis Galton¹ (1822-1911) et en particulier ses ébauches sur les principes de l'analyse factorielle et la primauté d'un facteur général d'intelligence sur des facteurs spécifiques. Sur le plan mathématique, si on peut s'accorder avec Spearman sur l'existence et la mesurabilité d'un score « g », en revanche, ce n'est pas le cas en ce qui concerne sa perspective unidimensionnelle de l'intelligence.

Ainsi, tout comme Alfred Binet (1857-1911) dont il est l'héritier, Wechsler suggère que l'intelligence devrait être abordée non plus à travers des processus primaires (temps de réactions, acuité visuelle, etc.) mais avec des processus mentaux supérieurs tels que la mémoire, le raisonnement ou encore l'attention. Il est à noter que Binet fut le premier à concevoir l'intelligence de cette manière (Huteau & Lautrey, 1999). Il a élaboré, en 1905 avec Théodore Simon, la première échelle métrique de l'intelligence. Puis, une nouvelle version, en 1908, qui avait déjà, plus ou moins, toutes les caractéristiques des tests actuels : le test d'intelligence Binet-Simon. Il a introduit également la notion de niveau mental, en situant les performances de l'enfant évalué en référence aux performances moyennes d'échantillons représentatifs d'âges différents (l'enfant effectuait les tâches jusqu'à ce qu'il n'y arrive plus, ce qui indiquait son niveau mental). Néanmoins, Wechsler se différencia de Binet, par le rejet de sa notion de niveau mental (Huteau & Lautrey, 1999), tout comme celle d'âge mental², qu'il substitue à une notion de rang, qui est la conception du QI actuelle, dans lequel la performance situe l'individu par rapport à son groupe d'âge (distribution gaussienne). En outre, il a procédé à l'intégration de tests de performance utilisant des objets et non des mots. En effet, il reprochait à Binet la présence trop importante de tests verbaux.

En définitive, Wechsler définit l'intelligence « *comme étant une aptitude globale qui permet d'agir intentionnellement en fonction d'un but, de penser de manière rationnelle et*

¹ Cousin de Charles Darwin et défenseur de la théorie de l'évolution, Galton chercha à faire le lien entre la théorie de la sélection naturelle et la recherche statistique afin de démontrer que des prévisions étaient possibles et vérifiables. C'est également le fondateur de la psychologie différentielle ainsi que l'inventeur de nombreuses notions et méthodes statistiques qui sont encore couramment employées aujourd'hui (Reuchlin, 2001).

² C'est Wilhelm Stern qui inventera, en 1912, le terme de « quotient intellectuel » et donnera naissance au test de QI classique, c'est-à-dire : $QI = (\hat{\text{Age}} \text{ mental} / \hat{\text{Age}} \text{ réel}) \times 100$.

d'interagir de manière efficace avec l'environnement » (Grégoire, 2006). Elle ne serait pas une entité unique, ni un simple ensemble d'aptitudes, mais une orchestration dont les caractéristiques peuvent être évaluées individuellement, à travers un ensemble de tests.

En ce qui concerne le WISC, Wechsler va réviser la batterie de tests (WISC-R) tout en gardant un regroupement des sous-tests selon deux catégories : un QI Verbal (QIV) et un QI Performance (QIP), qui forment eux-mêmes un QI total (QIT). Enfin, en ce qui concerne la troisième version de la batterie (WISC-III), on retrouve également le QIV, le QIP et le QIT (Huteau & Lautrey, 1999 ; Grégoire, 2006). Toutefois, la principale nouveauté consistait en la possibilité de calculer trois indices : l'*Indice de Compréhension verbale* (ICV), l'*Indice d'Organisation perceptive* (IOP) et l'*Indice de Vitesse de traitement* (IVT).

Abordons maintenant la quatrième version de cette batterie de test sur laquelle notre étude concernant la stabilité des performances a porté.

2.1.2 La quatrième version - Le WISC-IV

En ce qui concerne la dernière version, le WISC-IV, elle représente une révision importante (Grégoire, 2006). En effet, pour tenir compte des apports de l'approche factorielle de l'intelligence et de la psychologie cognitive, elle se base sur les modèles de l'intelligence de Cattell, de Horn et de Carroll (cf. chapitre 2.3). Le WISC-IV se différencie de part l'abandon du QIV et du QIP (suppression du modèle bifactoriel), au seul profit de quatre Indices (Grégoire, 2006) : l'*Indice Compréhension Verbale* (ICV), l'*Indice Raisonnement Perceptif* (IRP), l'*Indice Vitesse de Traitement* (IVT) et l'*Indice Mémoire de Travail* (IMT). Le QIT est resté mais devient secondaire, en comparaison de ces quatre Indices qui fournissent des repères quant aux forces et faiblesses des enfants (ipsatif) et qui permettent de situer leurs performances par rapport à leur groupe d'âge de référence (normatif). En ce qui concerne les autres changements apportés à cette quatrième version, il y a une diminution de l'impact du temps, la notion de retard mental a été retirée au profit d'une échelle allant de « très faible » (40) à « très supérieur » (160) et les normes ont été mises à jour. En outre, certains sous-tests disparaissent (Arrangement d'images, Labyrinthes et Assemblage d'objets), alors que d'autres font leur apparition (Raisonnement verbal, Identification de concepts, Matrices, Séquences lettres chiffres et Barrage). Au total, ce sont alors 15 épreuves,

dont 5 supplémentaires et non obligatoires³, qui se regroupent à travers les quatre Indices (Wechsler, 2005a). Le manuel en donne la description suivante :

- 1) Indice de compréhension verbale (ICV) : selon le manuel, cet indice mesure le facteur verbal, c'est-à-dire la compréhension générale de mots et la formation de concepts verbaux. Il implique l'application de compétences et de connaissances verbales dans la résolution de problèmes nouveaux, le traitement de l'information verbale ainsi que l'aptitude à penser avec des mots. Les tâches proposées permettent d'évaluer les enfants en fonction de leur capacité d'écoute et s'appuient sur les connaissances liées aux apprentissages. Il est à noter que les scores aux épreuves peuvent être influencés par la motivation, les intérêts, le milieu, l'éducation et la culture. Il est mesuré par trois épreuves et deux optionnelles :
 - **Similitude (SI)** : Selon le manuel, ce sous-test principal est censé évaluer le raisonnement verbal, la formation de concepts verbaux et la catégorisation. Il impliquerait la compréhension et l'expression verbale, la mémoire à long terme, la discrimination, l'aptitude à établir des relations (liens logiques), des connaissances générales des items ainsi que le concept de similitude et de différence. Ce serait une mesure du développement du langage.
 - **Vocabulaire (VO)** : également sous-test principal, il permettrait une mesure de la connaissance des mots et de leur sens ainsi que la formation de concepts verbaux. Il impliquerait le développement du langage, la mémoire à long-terme, la perception auditive, la compréhension verbale, le raisonnement abstrait et l'expression verbale.
 - **Compréhension (CO)** : cette épreuve est censée évaluer le raisonnement verbal, la formation de concepts verbaux, la compréhension et l'expression verbale ainsi que l'aptitude à évaluer, à utiliser l'expérience antérieure et à démontrer des connaissances pratiques. Il impliquerait également des connaissances de conventions, de normes sociales et morales. C'est un sous-test principal.
 - **Information (IN)** : cette épreuve supplémentaire est censée mesurer l'aptitude à appréhender, à retenir et à récupérer des connaissances générales à propos du monde. Ces connaissances sont acquises au cours des apprentissages. D'autres aptitudes seraient aussi impliquées comme la mémoire à long-terme, la perception auditive, la

³ Elles prolongent l'évaluation pour compléter les analyses et hypothèses. Elles peuvent remplacer des sous-tests principaux de manière justifiée : inadéquation avec les spécificités de l'enfant (limitations motrices, etc.), s'il y a eu interruption de la passation d'une épreuve obligatoire, etc. Un maximum de deux substitutions est toléré (Grégoire, 2006).

compréhension et l'expression verbale. Cette épreuve reflète l'influence du milieu socio-culturel et scolaire de l'enfant, la curiosité intellectuelle, l'ouverture au monde et l'adaptation scolaire.

- **Raisonnement verbal (RV)** : comme son nom l'indique, ce sous-test permettrait de mesurer le raisonnement verbal (général et analogique), mais également la compréhension verbale, l'abstraction verbale, l'aptitude à synthétiser et à intégrer l'information et la formation de concepts. Il impliquerait aussi la mémoire à long terme, la flexibilité cognitive, l'apprentissage par tâtonnement et l'attention. Il s'agit là aussi d'une épreuve supplémentaire.

2) Indice de raisonnement perceptif (IRP) : cet indice concerne le raisonnement basé sur du matériel nécessitant la perception visuelle, la formation de concepts non-verbaux ainsi que sur l'organisation et la planification visuelle. Il concerne l'aptitude à réaliser des opérations mentales de manière rapide comme la manipulation de symboles abstraits. Il impliquerait la coordination visuo-motrice et visuo-spatiale. Il est mesuré par trois épreuves et une optionnelle :

- **Cubes (CU)** : ce sous-test principal est censé permettre d'évaluer l'aptitude à analyser et à synthétiser des stimuli visuels abstraits. Il impliquerait le traitement simultané, la formation de concepts non-verbaux, l'organisation spatiale, la coordination visuo-motrice et visuo-spatiale, ainsi que la dépendance-indépendance à l'égard du champ⁴.
- **Identification de concepts (IDC)** : ce sous-test principal est censé mesurer le raisonnement non-verbal et en particulier l'aptitude de l'enfant à découvrir les caractéristiques sous-jacentes (règles, concepts, etc.) qui régissent un ensemble d'éléments. Il évaluerait le raisonnement abstrait, la catégorisation, l'organisation perceptive visuelle et la flexibilité mentale.
- **Matrices (MA)** : ce dernier sous-test principal est censé permettre d'évaluer l'aptitude à percevoir, à analyser et à synthétiser l'information visuelle (traitement simultané) ainsi que le raisonnement abstrait et analogique, la classification et la catégorisation. Elle impliquerait la mémoire de travail et le raisonnement inductif et déductif. Difficulté.

⁴ Capacité à isoler un élément de son contexte et à le réutiliser dans un contexte différent. La *dépendance* se rapporte à une perception dominée par l'organisation générale du champ. Ainsi, contrairement aux *indépendants*, certains enfants ont des difficultés à se décentrer de la forme globale d'une image, à déstructurer et à isoler les parties du champ.

- **Complètement d'images (CI)** : cet unique sous-test supplémentaire est censé évaluer la perception et l'organisation visuelle ainsi que la reconnaissance et la discrimination visuelle. Il impliquerait également la mémoire visuelle à long terme, l'attention, la flexibilité mentale et la dépendance/indépendance à l'égard du champ.
- 3) Indice de mémoire de travail (IMT) : il permet une mesure du maintien à court terme d'un ensemble d'informations et de leur traitement simultané dans la résolution de problèmes (manipulations mentales) ainsi que la flexibilité cognitive et l'aptitude à faire des séquences. Il est sensible à la fatigue et au stress. Il est mesuré par deux épreuves et une optionnelle :
- **Mémoire des chiffres (MC)** : ce sous-test principal est censé évaluer l'encodage et le traitement auditif, la mémoire auditive à court terme, ainsi que l'aptitude à faire des séquences. Elle impliquerait également les représentations visuo-spatiales, la manipulation et la transformation de l'information, la flexibilité cognitive et l'attention.
 - **Séquences lettres-chiffres (SLC)** : ce sous-test principal est censé mesurer la mémoire auditive à court terme et l'aptitude à faire des séquences. Il impliquerait aussi l'attention, la manipulation mentale, les représentations visuo-spatiales et la vitesse de traitement.
 - **Arithmétique (AR)** : ce sous-test supplémentaire permet une mesure de la mémoire à court terme, de l'aptitude à manipuler l'information, à faire des séquences, des opérations et des calculs mentaux (additions, soustractions, multiplications, divisions). Il implique l'attention, le raisonnement numérique et la vigilance.
- 4) Indice de vitesse de traitement (IVT) : ce dernier Indice permet une mesure de l'aptitude de l'enfant à analyser, à discriminer et à sélectionner des informations visuelles simples rapidement. Il implique aussi la vitesse grapho-motrice, l'organisation, la planification, la flexibilité cognitive, l'attention visuelle sélective et la vigilance. Ainsi, dans les épreuves, c'est le nombre d'items traités en un temps limité (vitesse d'exécution) qui permet de calculer le score. Il est mesuré par deux épreuves et une optionnelle :
- **Code (CD)** : c'est un sous-test principal qui a pour but la mesure de la vitesse de traitement, la mémoire à court terme visuelle, la perception visuelle et le traitement visuel séquentiel. Il implique la coordination visuo-motrice, la flexibilité cognitive et l'attention. Il est sensible au stress et à la méticulosité.

- **Symboles (SY)** : également sous-test principal, il permet une mesure de la vitesse de traitement sur une tâche cognitive simple mais également de la mémoire visuelle à court terme. Il implique la coordination visuo-motrice, l'organisation perceptive ainsi que la planification, la flexibilité cognitive, la discrimination visuelle, l'attention et l'aisance à manipuler des symboles abstraits.
- **Barrage (BA)** : c'est un sous-test supplémentaire qui permet de mesurer la vitesse de traitement, la discrimination visuelle, l'attention visuelle sélective, la vigilance et la négligence visuelle. Il implique la coordination visuo-motrice.

Actuellement, cette description des épreuves est remise en question. C'est également le cas en ce qui concerne celle de la structure factorielle du WISC-IV, et du nombre de facteurs qu'elle distingue. En effet, elle fait l'objet de nombreuses controverses (Keith et al., 2006). Comme nous allons le voir par la suite (cf. chapitre 2.4), plusieurs auteurs ont alors fait appel à la théorie de Cattell-Horn-Carroll (CHC) pour rejeter cette description donnée par le manuel et pour déterminer plus adéquatement une structure factorielle, de manière théorique et empirique.

Avant de décrire plus en détail le WISC-IV à la lumière de ce modèle CHC, théorie sur laquelle nous nous sommes basé pour effectuer nos analyses, abordons le sujet principal de cette étude. C'est-à-dire : la stabilité des performances aux sous-tests administrés lors de différentes séances.

2.2 Stabilité et fluctuation à court terme des performances

Selon le manuel d'administration et de cotation du WISC-IV, les sous-tests doivent être effectués en une seule fois. Or, dans la pratique, on recourt souvent à plusieurs séances. De ce fait, cela pose la question de l'influence que peut avoir cet étalement de l'évaluation dans le temps sur la congruence des résultats obtenus aux tests censés évaluer les mêmes aptitudes cognitives. A travers nos recherches littéraires, nous n'avons pas trouvé d'étude spécifique portant sur cette problématique. Ainsi, à notre connaissance, il ne semble pas y avoir de travaux portant sur l'influence de ce découpage, d'où l'intérêt de cette étude. Toutefois, différentes recherches, que nous allons présenter ici, nous ont permis de nous interroger sur la stabilité des performances et de nous guider dans nos interprétations.

Bien que l'intelligence soit considérée comme relativement stable au cours du temps, des fluctuations de performances à court terme et à long terme, plus ou moins importantes,

peuvent être observées (Grégoire, 2006). En effet, selon Floyd, Bergeron, McCormack, Anderson et Hargrove-Owens (2005), les aspects temporels des séances d'évaluation peuvent contribuer à l'erreur de mesure, en particulier en ce qui concerne l'intervalle de temps entre les séances. Grégoire (2006) a d'ailleurs observé ce phénomène à travers la méthode test-retest pour l'ensemble des épreuves du WISC-IV, sur un intervalle de temps relativement court (délais moyen de 27 jours), pour 93 enfants âgés de 6 à 15 ans. Il a constaté que le fait d'avoir passé un test une première fois peut aider à mieux réussir la fois suivante, par l'activation de connaissances en mémoire à long terme et l'utilisation de stratégies de résolution de problèmes acquises précédemment. C'est l'effet d'apprentissage. Ainsi, alors que la corrélation pour le QI Total entre les deux passations est très élevée (.91), indiquant que les différences inter-individuelles restent stables dès l'âge de 6-7 ans, ce n'est pas le cas, en ce qui concerne les résultats individuels. Si on se réfère aux Indices, on constate une augmentation de 2,1 points pour l'*Indice de Compréhension verbale* (ICV), de 8,2 points pour l'*Indice de raisonnement perceptif* (IRP), de 3,6 points pour l'*Indice de Mémoire de travail* (IMT) et de 12,0 points pour l'*Indice de Vitesse de traitement* (IVT). En ce qui concerne le QI Total, l'augmentation était de 8,3 points. De plus, les scores aux sous-tests sont moins stables dans le temps que ceux des indices ou de l'échelle totale. Ainsi, une nette amélioration est constatée, en moyenne, d'une passation à l'autre, en particulier en ce qui concerne les sous-tests de l'IVT et de l'IRP. En effet, les gains sont plus importants en ce qui concerne les sous-tests non-verbaux (ex : Complètement d'images, Code, etc.) que les verbaux (ex : Information, Compréhension, etc.). Toutefois, ces effets d'apprentissage ne seraient pas permanents et tendraient à disparaître progressivement avec le temps. On considère, d'ailleurs, qu'un intervalle de temps d'au moins une année est nécessaire (Canivez & Watkins, 1998 ; Grégoire, 2006). Ainsi, la réévaluation de l'intelligence semble poser problème surtout quand l'intervalle de temps est court. De ce fait, on peut se poser la question au sujet de la validité de la mesure qui a été effectuée à un moment donné et, plus précisément, si les performances obtenues par les enfants ainsi que leur position par rapport à leur groupe d'âge correspondent réellement à leur niveau de compétence.

Dans notre étude, nous n'avons pas procédé à une procédure test-retest. On ne peut donc pas être certain que ces effets d'apprentissage soient valables lorsque les sous-tests sont administrés une seule fois mais sur plusieurs séances. En revanche, étant donné que Grégoire (2006) stipule qu'une différence supérieure à 4 points implique une différence de performance, nous pouvons nous baser sur cette affirmation pour vérifier la stabilité des

scores aux sous-tests censés évaluer les mêmes aptitudes cognitives lorsqu'ils sont administrés sur différentes sessions d'évaluation.

Toujours selon Grégoire (2006), la présence d'une différence de performance entre des épreuves censées mesurer les mêmes construits cognitifs n'est pas rare chez les individus tout-venant au sein d'une seule et même passation. Elle serait même relativement fréquente. Ainsi, même si les capacités d'une personne sont homogènes, une certaine variabilité des scores est généralement observée. Par exemple, pour l'échantillon d'étalonnage, une différence de 5 points et plus est constatée chez 6,1% des individus entre les épreuves *Similitudes* et *Vocabulaire* ; chez 10,9% entre *Vocabulaire* et *Compréhension*, chez 11,7% entre *Compréhension* et *Similitudes*, chez 16,8% entre *Matrices* et *Identification de concepts*, chez 12,8% entre *Mémoire des chiffres* et *Séquence lettres-chiffres* et chez 10,3% entre les sous-tests *Code* et *Symboles*. Par ailleurs, plus les écarts de performances entre paires de sous-tests sont élevés et plus le pourcentage d'individus qui les présentent diminue. Il est à noter, toutefois, que si les fluctuations des performances sont très importantes, la probabilité qu'elles soient le reflet de différences réelles augmente, ce qui engendre un affaiblissement de la valeur clinique des indices concernés (Grégoire, 2006). Ainsi, au regard de ce qui vient d'être présenté, nous pouvons émettre l'hypothèse que si l'étalement de l'administration du WISC-IV sur plusieurs séances n'a pas d'influence, nous devrions retrouver de faibles différences entre les scores pour la majorité des enfants de même que des fluctuations plus importantes pour un petit pourcentage d'individus. Comme nous allons le voir par la suite, ces fluctuations peuvent être attribuées à des erreurs de mesures, à un profil particulier de l'individu évalué ou à des aspects développementaux et non pas, automatiquement, à des caractéristiques pathologiques.

Selon Grégoire (2006), il ne faut pas oublier qu'une mesure sans erreur n'est pas possible (on ne peut supprimer l'erreur aléatoire). Les variations des performances peuvent alors être liées à la fidélité imparfaite des instruments de mesure et à un ensemble de facteurs qu'on ne peut jamais parfaitement contrôler et qui peuvent influencer les résultats aux tests tels que : la compréhension des consignes, la motivation (certaines tâches peuvent être plus attrayantes ou familières que d'autres), la distractibilité, le niveau de stress, la fatigue, etc. Par exemple, pour le sous-test *Code* du WISC-IV, la pression du temps limité peut faire qu'un enfant trop anxieux est susceptible de perdre ses moyens et, de ce fait, l'amener à fournir des performances plus faibles. De plus, une méticulosité excessive peut aussi ralentir l'enfant et faire baisser son score. Enfin, selon Floyd et al. (2005), les exigences des tâches administrées peuvent également contribuer à la fluctuation des performances. En effet, les sous-tests

classés selon les aptitudes générales peuvent impliquer des aptitudes spécifiques différentes et plus ou moins développées chez l'individu, ce qui peut affecter les scores. Ainsi, un enfant peut présenter un score élevé à une épreuve et un score plus faible à une autre, alors qu'elles évaluent la même aptitude générale.

Toutefois, on ne peut considérer la variabilité intra-individuelle des performances comme résultant uniquement de biais méthodologiques et des propriétés psychométriques des épreuves utilisées. En effet, ces variations peuvent aussi provenir de l'individu lui-même (Grégoire, 2006). Par exemple, la stabilité des scores peut également être abordée sur le plan développemental. En effet, selon l'hypothèse de la *différentiation – dédifférentiation*, les aptitudes cognitives partagent de nombreux traits communs chez les enfants mais, en ce qui concerne les jeunes adultes, il y aurait, au contraire, une différenciation. Par exemple, dans une étude de Ghisletta, Lecerf, et de Ribaupierre (2003), les corrélations inter-tests pour des épreuves d'inhibition, de mémoire de travail et de vitesse de traitement sont plus élevées chez les enfants que chez les jeunes adultes. Ainsi, et en accord avec Grégoire (2006), nous pouvons nous attendre à ce que les épreuves censées évaluer les mêmes aptitudes cognitives soient effectivement corrélées entre elles de manière significative et avec des intensités élevées mais qu'elles corrèleront également, de manière plus modérée, avec d'autres sous-tests.

De plus, des fluctuations à court terme, réversibles et représentant un état stable, peuvent mettre en évidence des changements sur le plan du fonctionnement cognitif (caractère adaptatif de l'individu). D'ailleurs, ces changements développementaux, qui eux sont relativement permanents au cours du cycle de vie, peuvent être observés à travers l'administration d'un ensemble d'épreuves (Ghisletta & al., 2003). Ainsi, la présence de fluctuations de performances chez les enfants n'est pas forcément liée à des erreurs de mesures mais peut représenter des changements réels. Il est à noter que les concepts de changement et de variabilité intra-individuelle sont essentiels dans l'évaluation psychologique, d'autant que les interprétations ainsi que l'établissement ou les modifications d'une intervention découlent de l'observation d'une amélioration, d'un déclin ou d'un dysfonctionnement (Salthouse, 2007). Cette appréciation étant, en partie, tributaire des scores obtenus par l'individu, la présence d'une fluctuation à court terme des performances cognitives peut être utile pour le diagnostic et le pronostic mais également rendre difficile l'interprétation des résultats par le professionnel.

Enfin, selon Salthouse (2007), l'évaluation unique n'est jamais suffisante. En effet, cet auteur a montré, après avoir administré différents tests cognitifs sur trois séances et dans un

intervalle de temps d'environ deux semaines, que les performances cognitives ne seraient pas stables à court terme et suivraient une loi normale qui est propre à chaque individu et dans laquelle la dispersion autour du niveau moyen de performances est différente pour chacun. De ce fait, bien que l'évaluation multiple sur une période de temps relativement courte soit coûteuse, elle n'en demeure par moins nécessaire pour obtenir des niveaux de performances plus vraisemblables. L'auteur suggère également le développement de nouvelles normes et de nouvelles épreuves parallèles en ce qui concerne les outils psychométriques afin de prendre en compte cette dimension. Il est à souligner, néanmoins, que l'étude de Salthouse porte sur une population d'individus âgés de 18 à 97 ans. C'est pourquoi, nous ne pouvons en conclure que cela est valable pour les enfants.

2.3 Le modèle Cattell-Horn-Carroll (CHC)

Comme convenu, nous allons maintenant présenter le modèle Cattell-Horn-Carroll (CHC). En effet, il est à rappeler que l'évaluation de la stabilité des performances aux sous-tests du WISC-IV, administrés lors de différentes séances, s'est faite à la lumière de cette théorie. Plus précisément, elle nous a permis d'élaborer, dans notre étude, les différents regroupements des épreuves de même que les comparaisons et interprétations à effectuer. Commençons, tout d'abord, par les fondements, l'évolution et la description de ce modèle pour ensuite aborder plus en détail le WISC-IV au regard du CHC.

2.3.1 *Le modèle de Cattell et Horn*

C'est Raymond Bernard Cattell (1905-1998), qui, inspiré des travaux de Louis Léon Thurstone⁵ (1887-1955), a posé la première pierre dans l'élaboration de cette théorie (Huteau & Lautrey, 1999). En outre, il a également été, après Spearman, l'un des premiers à introduire l'utilisation de méthodes d'analyse factorielle pour ses travaux sur les aptitudes cognitives. Dans sa conception initiale, Cattell a considéré la structure de l'organisation des aptitudes cognitives selon deux facteurs (Grégoire, 2006) : l'intelligence fluide (Gf) et l'intelligence cristallisée (Gc). Il élaborera alors un modèle d'intelligence dichotomique : le modèle Gf-Gc (McGrew, 2005). En ce qui concerne John L. Horn (1928-2006), il est l'élève de Cattell (Huteau et Lautrey, 1999). Inspiré par les travaux de Spearman et de Thurstone, Horn est

⁵ Thurstone a repris les analyses de Spearman, qu'il remet en cause, et introduit une conception multifactorielle de l'intelligence. Selon lui, il n'y aurait pas de facteur général, mais des secteurs de compétences. Il conclut à l'existence de sept facteurs principaux représentant autant de types d'intelligence et n'ayant pas de lien entre eux.

surtout connu pour ses travaux dans le domaine de la psychométrie (analyses multivariées) et des performances cognitives dans une perspective lifespan (tout au long de la vie). Selon lui, ces deux formes d'intelligence (Gf et Gc) ont des trajectoires très différentes au cours du développement de l'enfance à l'âge adulte. Ainsi, l'intelligence fluide augmenterait jusqu'à l'adolescence et déclinerait à partir de l'âge adulte, surtout quand l'individu atteint un âge avancé. En revanche, l'intelligence cristallisée reste relativement stable.

Pour terminer, en ce qui concerne ce modèle initial, ces deux auteurs ont élargi le nombre des facteurs primaires et le nombre des facteurs généraux de second ordre, tout en conservant Gf et Gc (McGrew, 2005). Il est à noter que la théorie a continué d'être appelée théorie Gf-Gc, alors que ces deux aptitudes ont été intégrées aux autres de manière égale, c'est-à-dire sans hiérarchie (Schrank, McGrew & Woodcock, 2001). Selon Horn et Noll (1997), elle est composée de 9 facteurs: Intelligence fluide, Intelligence cristallisée, Traitement visuel, Traitement auditif, Vitesse de traitement, Vitesse de réaction, Mémoire à long terme, Mémoire à court terme et Connaissances quantitatives. Ces aptitudes sont indépendantes et ne se regroupent pas autour d'un facteur général (Grégoire, 2006).

2.3.2 *Le modèle de Carroll*

Par la suite, John Bissell Carroll (1916-2003), également influencé par Thurstone et connu pour sa grande contribution provenant de ses nombreux travaux sur l'analyse factorielle des aptitudes cognitives, proposa en 1993 une théorie de l'intelligence avec une structure hiérarchique et multifactorielle sur trois niveaux différents : la *Three-Stratum Theory* (McGrew, 2005). Sa structure se définit ainsi : la strate III se compose d'un facteur « g » (comparable à celui de Spearman). Il est au sommet de la hiérarchie et est extrait par l'analyse factorielle des facteurs de la strate II. En ce qui concerne cette deuxième strate, elle contient 8 grandes aptitudes cognitives (intelligence fluide, intelligence cristallisée, mémoire et apprentissage, représentation visuo-spatiale, représentation auditive, récupération en mémoire à long terme, rapidité cognitive, vitesse de traitement) qui peuvent gouverner ou influencer une grande variété de comportements. Ces aptitudes, proches de celles de Horn et Cattell, saturent les facteurs primaires qui sont tous corrélés entre eux. Enfin, la strate I est le niveau le plus bas et contient 69 facteurs primaires (Grégoire, 2006).

Il est à souligner que le modèle de Cattell-Horn et celui de Carroll diffèrent en ce qui concerne l'existence du facteur d'intelligence générale (Alfonso, Flanagan, & Radwan, 2005).

En effet, contrairement à Carroll, Cattell et Horn n'ont jamais admis son existence (Huteau & Lautrey, 1999).

2.3.3 Le modèle CHC des aptitudes cognitives

Etant donné la très grande similarité entre le modèle de Cattell-Horn et celui de Carroll, Kevin S. McGrew (1997), et par la suite avec l'aide de Flanagan (2000), a proposé un nouveau modèle cognitif basé sur la synthèse des deux modèles précédents : le modèle Cattell-Horn-Carroll ou modèle CHC des aptitudes cognitives. Sur la base de ses propres analyses, Flanagan identifie 10 grandes aptitudes générales et plus de 70 aptitudes spécifiques (Flanagan, Ortiz & Alfonso, 2007).

Dans notre étude, nous nous sommes basés sur ce modèle pour effectuer nos regroupements et comparaisons, afin de tester la stabilité des performances. Avant de présenter sa structure, il est à noter que, de part leur grand nombre, seuls les facteurs impliqués dans les épreuves des différentes batteries de tests administrées dans notre recherche seront décrits ici :

- 1) L'intelligence fluide (Gf): cette aptitude générale concerne les opérations mentales utilisées lorsque l'individu est confronté à une nouvelle tâche qui ne peut être résolue automatiquement. Elle concerne l'aptitude générale de raisonnement (inductif et déductif), l'identification de relations (inférences), la compréhension d'incidences, la reconnaissance et la formation de concepts, la classification, la transformation de l'information et la génération d'hypothèses. L'intelligence fluide comporte, par exemple, les facteurs primaires suivants :
 - **Induction (I)** : elle concerne l'aptitude à découvrir les caractéristiques sous-jacentes (règles, principes, appartenance à une classe, etc.) qui sous-tendent une relation entre des éléments, ou à appliquer des règles apprises précédemment. Elle implique la capacité à combiner des éléments d'information séparés dans la formation d'inférences, de règles, d'hypothèses ou de conclusions.
 - **Raisonnement séquentiel général (RG)** : cette aptitude primaire correspond à la capacité à utiliser des règles, des conditions, etc. pour la résolution de problèmes par déduction et en plusieurs étapes.
 - **Raisonnement quantitatif (RQ)** : cette aptitude primaire correspond au raisonnement (par induction et/ou déduction) à travers des concepts, des relations et des propriétés mathématiques (stimuli quantitatifs).

- 2) L'intelligence cristallisée (Gc): cette aptitude générale concerne l'étendue et la profondeur des connaissances en ce qui concerne le langage, les concepts verbaux et les informations verbales. En d'autres termes, elle concerne la capacité à utiliser notre base de connaissances essentiellement verbales développée au cours des apprentissages et des expériences éducatives. Elle est liée à la culture et est fondée sur la connaissance du langage déclaratif et procédural. De plus, elle est associée à la compréhension verbale, à la connaissance lexicale, à la capacité à communiquer des informations. Elle comporte, par exemple, les facteurs primaires suivants :
- **Développement du langage (LD)** : cette aptitude primaire correspond au développement général du langage ainsi qu'à la compréhension et à l'application des mots, des phrases, etc. Elle concerne les compétences linguistiques dans l'expression et la communication.
 - **Connaissances lexicales (VL)** : cette aptitude primaire correspond à la capacité à utiliser et à comprendre correctement le sens des mots (noms, verbes ou adjectifs). Elle est fortement liée au développement du langage et à l'étendue du vocabulaire.
 - **Informations générales verbales (KO)** : elle concerne la base de connaissances et l'aptitude à utiliser les connaissances générales (essentiellement verbales).
 - **Aptitude d'écoute (LS)** : elle correspond à l'aptitude primaire à écouter et à comprendre le sens de la communication orale (mots, expressions, phrases, etc.). Cette aptitude primaire n'est pas évaluée dans la WISC-IV mais intervient dans une autre épreuve administrée dans notre étude (*Compréhension de consignes* du NEPSY).
- 3) La mémoire à court terme (Gsm): elle concerne l'aptitude à appréhender, à maintenir et à rappeler des informations à court terme et à les utiliser après quelques secondes. De plus, c'est un système limité en ce qui concerne la conservation de la quantité d'informations. En effet, il y a une perte, sauf si une stratégie cognitive est utilisée pour conserver les informations. Elle se compose, par exemple, des aptitudes primaires suivantes :
- **Empan mnésique (MS)** : cette aptitude primaire correspond à la capacité à se rappeler une série d'éléments liés entre eux de manière aléatoire et à les reproduire dans l'ordre correct, après quelques secondes et une seule présentation (ex : des lettres, des chiffres).
 - **Mémoire de travail (MW)** : elle correspond à l'aptitude à stocker temporairement des informations sur lesquelles un ensemble d'opérations cognitives sont effectuées.

Elle exige l'attention divisée et la gestion de la capacité limitée des ressources de la mémoire à court terme.

- 4) Les connaissances quantitatives (Gq): cette aptitude générale concerne le matériel numérique et consiste en la capacité à comprendre et à appliquer des concepts et des connaissances mathématiques pour résoudre des problèmes quantitatifs (relations quantitatives, manipulation de symboles numériques). Comme Gc, Gq est essentiellement acquis au cours des expériences éducatives (connaissances quantitatives, déclaratives, procédurales). Il faut savoir que Gq n'équivaut pas à l'aptitude spécifique RQ (présent dans Gf), qui est l'aptitude à raisonner, par induction et déduction, pour résoudre des problèmes quantitatifs. Gq comporte, par exemple, le facteur primaire suivant :
 - **Performances mathématiques (A3)** : cette aptitude primaire correspond à la capacité à démontrer des habiletés mathématiques.

- 5) La vitesse de traitement (Gs) : cette aptitude concerne l'exécution automatique, rapide et efficace de tâches cognitives et en particulier quand elles exigent un maintien de l'attention et de la concentration. Plus précisément, c'est un indicateur de l'aptitude à scanner et à comparer rapidement des stimuli visuels simples. Gs est associé à des processus cognitifs élémentaires et est habituellement évalué par des tâches qui nécessitent un traitement cognitif rapide mais peu de raisonnement. Il comprend, par exemple, les facteurs primaires suivants :
 - **Vitesse perceptive (P)** : elle correspond à l'aptitude à rechercher rapidement, de manière précise et correctement des similitudes ou des différences sur des éléments visuels présentés côte à côte ou séparés dans le champ visuel (comparaisons).
 - **Taux de réponse à un test (R9)** : cette aptitude primaire se rapporte à des éléments cognitifs sur lesquels sont effectuées des opérations rapides (décisions simples). Cette aptitude n'est pas associée à un type particulier de contenu ou de stimuli.

- 6) Le traitement visuel (Gv) : il concerne l'aptitude à manipuler mentalement des stimuli visuels de différentes manières (faire pivoter, tourner, basculer, etc.). Cette aptitude implique également la perception, l'analyse, le stockage, la synthèse, la génération, la récupération et la transformation de stimuli visuels (figuratifs ou géométriques). Il comprend, par exemple, les facteurs primaires suivants :

- **Visualisation (VZ)** : elle correspond à l'aptitude primaire à appréhender, à manipuler et à transformer mentalement des stimuli visuels en deux ou trois dimensions (faire pivoter, etc.) ainsi qu'à les faire correspondre à d'autres stimuli.
- **Relations spatiales (SR)** : cette aptitude primaire correspond à la capacité à résoudre des problèmes impliquant des relations spatiales. En d'autres termes, elle implique l'identification et la manipulation de stimuli visuels à travers différentes perspectives (angles ou positions).
- **Flexibilité de fermeture (CF)** : elle concerne l'aptitude à identifier une figure ou un motif visuel particulier intégré dans un schéma complexe de distraction visuelle.

Actuellement, il existe un certain consensus en ce qui concerne le modèle CHC, bien qu'il continue d'être revu et augmenté par divers théoriciens. En effet, certains auteurs, comme McGrew (2005 ; Flanagan & McGrew, 1998), plaident pour un certain nombre d'extensions taxonomiques. Cet auteur, par exemple, a proposé de séparer le facteur de vitesse de traitement en trois aptitudes ainsi que d'ajouter quatre grands facteurs qui sont : les aptitudes psychomotrices, les aptitudes kinesthésiques, les aptitudes tactiles et les aptitudes olfactives.

Ce modèle CHC n'en demeure pas moins le modèle psychométrique de l'intelligence le plus à jour théoriquement et empiriquement (Flanagan et McGrew, 1998). De plus, étant donné que la plupart des batteries de tests actuelles sont basées sur le modèle CHC, cette théorie offre un cadre unificateur, un langage commun pour les professionnels (McGrew, 1997) et est, d'ailleurs, beaucoup utilisée pour guider l'interprétation des scores, lors du bilan cognitif. Enfin, il est aussi utilisé comme fondement théorique pour la construction et la révision de batteries d'évaluation cherchant à représenter plus pleinement l'ensemble des aptitudes cognitives humaines. Ainsi, afin de fournir un certain nombre de mesures et de directives pour réduire l'écart entre la théorie et la pratique, Flanagan et collègues (Flanagan & McGrew, 1998 ; Flanagan, McGrew & Ortiz, 2000 ; Flanagan, Ortiz & Alfonso, 2007) ont élaboré une méthode de batterie-croisée⁶ (cross-battery) fondée sur la théorie CHC. Cette méthode consiste à combiner différentes batteries afin d'évaluer un plus grand nombre d'aptitudes générales et spécifiques et de tester des hypothèses. De cette façon, pour chacune des épreuves concernées, une hypothèse est formulée en ce qui concerne le(s) aptitude(s) générale(s) et spécifique(s) impliquée(s), ce qui permet d'effectuer une classification, à l'aide de techniques d'analyses factorielles confirmatoires, de l'ensemble des sous-tests dans les

⁶ C'est Richard Woodcock qui, en 1990, a proposé cette méthode consistant à combiner les batteries de tests. Il révisera, par ailleurs sa batterie de test (WJ-III) sur la base du modèle CHC.

aptitudes générales et spécifiques décrites par la théorie CHC. L'administration d'une épreuve d'une autre batterie évaluant les mêmes aptitudes permettra alors de rejeter ou non cette hypothèse. Au cours des dernières années, la révision de plusieurs tests d'intelligence importants, avec pour base la théorie CHC, ont été publiés comme le WJ-III, le KABC-II ou le NEPSY-II. En ce qui concerne le WISC-IV, et comme nous allons le voir dans le chapitre qui suit, il fait également l'objet d'études en lien avec cette théorie. D'ailleurs, comme nous l'avons déjà évoqué à plusieurs reprises, notre recherche s'est basée sur les travaux en lien avec ce modèle, afin d'établir les regroupement et les comparaisons des épreuves à effectuer pour évaluer la stabilité des performances entre les séances.

2.4 Le WISC-IV selon le modèle CHC - Une structure alternative

Le WISC a été longtemps considéré comme athéorique. Bien que les changements apportés à la quatrième version soient notables, il lui manque encore un cadre théorique explicite à jour, en ce qui concerne son organisation et sa structure. En effet, bien qu'il s'inspire des modèles de Cattell-Horn et de Carroll, le WISC-IV n'a pas été explicitement développé sur la base du modèle CHC. Il est à noter aussi que les épreuves du WISC-IV ne couvrent qu'une partie des grands facteurs décrits par la théorie CHC (Flanagan & al., 2000). Par conséquent, cette batterie de tests ne reflète pas la structure théorique et empirique actuelle de l'intelligence (Horn & Blankson, 2005).

Différentes études américaines ont montré que le modèle théorique CHC conduit à une meilleure interprétation des résultats au WISC-IV et qu'une structure alternative est plus adaptée. Ainsi, Keith et ses collègues (2006), ont suggéré l'existence non plus de quatre, mais de cinq facteurs : l'Intelligence fluide (Gf), l'Intelligence cristallisée (Gc), la Mémoire à court terme (Gsm), la Vitesse de traitement (Gs) et le Traitement visuel (Gv). Flanagan et Kaufman (2004), quant à eux, proposent même un sixième facteur : les Connaissances quantitatives (Gq). Il faut savoir que le sous-test *Arithmétique* est la seule épreuve qui sature sur Gq dans le modèle à 6 facteurs et que, selon Flanagan et Kaufman (2004) et Keith et al. (2006), elle serait également une mesure de Gf. En revanche, selon Grégoire (2006), elle serait une mesure de Gsm et de Gc. Enfin, selon The Psychological Corporation, elle serait une mesure de Gq et de Gsm. En ce qui concerne les autres sous-tests, The Psychological Corporation a classé les épreuves *Similitudes* et *Raisonnement verbal* en tant que mesure de Gf alors que Keith et al. (2006), ou encore Grégoire (2006), les ont classées en tant que mesure de Gc. En ce qui concerne l'épreuve *Cubes*, Kaufman (1994) l'a classée en tant que mesure de Gf, alors que

d'autres l'ont classée en tant que mesure de Gv (Grégoire, 2006). Ces différences amènent donc à penser que la construction de chaque épreuve n'est pas encore bien définie actuellement et n'est pas équivalente aux différentes versions alternatives.

En ce qui concerne la version française du WISC-IV, il n'est pas possible d'être certain que les conclusions proposées concernant la construction et la classification des sous-tests par les études américaines soient valables ou équivalentes pour la version francophone et, en particulier, pour un échantillon d'enfants suisses-romands. En effet, une adaptation ne se limite pas à une simple traduction des items et des consignes. Ainsi, la plupart des épreuves font l'objet de modifications plus ou moins importantes (Grégoire, 2006). D'ailleurs, selon Grégoire (2006), si l'on compare la version américaine à la française à travers des analyses factorielles exploratoires, on constate certaines divergences. Par exemple, dans la version américaine, le sous-test *Cubes* a une saturation de .78 sur l'Indice de Raisonnement perceptif (IRP), alors que la saturation est de seulement .52 dans la version française. Autre exemple, dans la version américaine, le sous-test *Barrage* a une saturation de .65 sur l'Indice de Vitesse traitement (IVT), alors que la saturation n'est que de .42 dans la version française. Enfin, le sous-test *Matrices* a une saturation de .64 dans la version anglaise sur l'Indice de Raisonnement perceptif (IRP), alors qu'elle est de .51 dans la version française.

Ainsi, Lecerf et collaborateurs (2010) ont élaboré une étude portant sur les données françaises du WISC-IV. Leurs résultats ont montré également que la batterie de tests est mieux décrite avec un modèle alternatif, basé sur la théorie CHC, qu'avec une structure en quatre facteurs. Comme Flanagan et Kaufman (2004), ce modèle alternatif comporte six facteurs (Gf, Gc, Gsm, Gs, Gv et Gq). Au regard de cette recherche, la présente étude prendra comme référence leur modèle alternatif basé sur la théorie CHC afin de nous aider à choisir les comparaisons et les regroupements à effectuer dans notre analyse de la stabilité des performances. Cette construction se définit de la manière suivante : les épreuves regroupées sous l'Indice de Compréhension verbale (ICV) ainsi que celles de l'Indice de Vitesse de traitement (IVT) ont été réparties respectivement de la même manière sous le facteur Gc et le facteur Gs. En revanche, les épreuves regroupées sous l'Indice de raisonnement perceptif (IRP) ont été réparties entre les facteurs Gf et Gv. En ce qui concerne les épreuves regroupées sous l'Indice de Mémoire de travail (IMT), elles ont été réparties sous le facteur Gsm à l'exception de l'épreuve *Arithmétique* (AR). En effet, étant donné que cette épreuve semble mesurer Gq, Gsm et, dans une moindre mesure, Gc (Lecerf & al., 2010), nous avons décidé de ne pas l'intégrer directement dans notre modèle. Enfin, il est à remarquer que l'épreuve *Complètement d'images* (CI), censée mesurer Gv, mesure principalement Gc (Lecerf & al.,

2010). Toutefois, afin de pouvoir effectuer des comparaisons, un minimum de deux sous-tests du WISC-IV par facteurs étaient nécessaires dans notre étude. Ainsi, l'épreuve *Complètement d'images* (CI) a été regroupée sous Gv avec le sous-test *Cubes* (CU).

En définitive, dans notre étude, le modèle CHC nous permettrait l'interprétation de 5 grandes aptitudes cognitives et de 12 aptitudes spécifiques mesurées par le WISC-IV. Les hypothèses sont les suivantes :

1) Intelligence fluide (Gf) :

- **Matrices (MA) :** ce sous-test serait un indicateur de l'*Induction* (I) et du *Raisonnement séquentiel général* (RG).
- **Identification de concepts (IDC) :** il serait aussi un indicateur de l'*Induction* (I).

2) Intelligence cristallisée (Gc) :

- **Similitudes (SI) :** cette épreuve serait un indicateur du *Développement du langage* (LD) et des *Connaissances lexicales* (VL).
- **Vocabulaire (VO) :** ce sous-test serait une mesure des *Connaissances lexicales* (VL).
- **Compréhension (CO) :** cette épreuve serait une mesure des *Informations générales verbales* (K0) et du *Développement du langage* (LD).
- **Information (IN) :** ce sous-test serait une mesure des *Informations générales verbales* (K0).
- **Raisonnement verbal (RV) :** cette épreuve serait une mesure des *Connaissances lexicales* (VL).

3) Mémoire à court terme (Gsm) :

- **Mémoire des chiffres (MC) :** ce sous-test serait une mesure de l'*Empan mnésique* (MS) et en partie de la *Mémoire de travail* (MW).
- **Séquences lettre-chiffres (SLC) :** cette épreuve serait un indicateur de la *Mémoire de travail* (MW).

4) Traitement visuel (Gv) :

- **Cubes (CU) :** ce sous-test serait une mesure des *Relations spatiales* (SR) et de la *Visualisation* (VZ).
- **Complètement d'images (CI) :** cette épreuve serait une mesure de *Flexibilité de fermeture* (CF) et des *Informations générales verbales* (Gc-K0).

5) Vitesse de traitement (Gs) :

- **Code (CD)** : cette épreuve serait un indicateur du *Taux de réponse à un test* (R9).
- **Symboles (SY)** : ce sous-test mesurerait la *Vitesse perceptive* (P) et, en partie, le *Taux de réponse à un test* (R9).
- **Barrage (BA)** : cette épreuve mesurerait également la *Vitesse perceptive* (P) et, en partie, le *Taux de réponse à un test* (R9).

2.5 Problématique

Dans cette recherche, et au regard du modèle CHC, nous avons voulu vérifier si le fait d'effectuer les épreuves du WISC-IV sur plusieurs séances peut avoir une influence sur la stabilité des performances des enfants suisses-romands. En effet, dans la pratique, la batterie de tests est souvent administrée en ayant recours à plus d'une séance, alors qu'à l'origine elle a été conçue pour être réalisée au travers d'une seule et même session d'évaluation. Ce découpage tient généralement du fait que certains profils d'enfants ne permettent pas toujours une évaluation de longue durée (ex : trouble de l'attention/hyperactivité, etc.). Or, ne serait-ce que la passation des dix sous-tests obligatoires, sur les quinze qui composent le WISC-IV, requiert environ 60 à 80 minutes.

En nous référant à la littérature, et comme nous l'avons vu précédemment, les aspects temporels des sessions d'évaluation peuvent engendrer des erreurs de mesures. C'est le cas, notamment, lorsqu'on administre une seconde fois la batterie de tests sur un intervalle de temps relativement court. En effet, les performances ont tendance à augmenter suite à un effet d'apprentissage. Toutefois, étant donné que dans notre étude nous ne sommes pas dans une situation test-retest, nous ne pouvons être certains que cet effet soit également présent lorsque l'on administre le WISC-IV une seule fois mais sur plusieurs séances.

Par ailleurs, lors d'une même séance d'évaluation, il a été constaté que des différences de performances entre des épreuves censées évaluer les mêmes aptitudes cognitives pouvaient apparaître fréquemment chez les individus tout-venants alors que leurs capacités sont homogènes. Généralement, pour la grande majorité des individus, ces fluctuations sont relativement faibles et ne sont pas le reflet de différences réelles. D'ailleurs, plus les écarts entre les scores sont importants et plus le pourcentage d'individus qui les présentent diminue. Ces différences sont souvent attribuées à des erreurs de mesures mais peuvent également être liées à un profil particulier de l'individu évalué ou encore à des aspects développementaux.

Ainsi, au regard de ces différents éléments, nous pouvons donc nous poser la question si le découpage de la batterie de tests pour une administration sur plusieurs séances a un effet sur les résultats des mesures dont dépendent les interprétations de l'évaluateur. Est-ce que cela contribue à l'erreur de mesure ? Est-ce que le pourcentage d'enfants présentant des différences de performances va augmenter ? Ces fluctuations seront-elles très importantes ? C'est ce que nous allons essayer de vérifier.

Tout d'abord, afin d'évaluer cette stabilité, il a été primordial de s'assurer au préalable que les différents sous-tests évaluent bien les mêmes aptitudes cognitives, en fonction du facteur sous lequel chacun d'eux a été classé. Ainsi, nous avons procédé à des analyses, via des corrélations et leur comparaisons à celle du manuel du WISC-IV, afin de vérifier s'il existe bien un lien, pour notre échantillon d'enfants suisses-romands, entre les sous-tests censés mesurer les mêmes construits cognitifs et entre les performances inter-séances. Il est à noter que nous pouvons également nous attendre à des corrélations plus modérées entre des épreuves mesurant des construits cognitifs différents. En effet, selon l'hypothèse de la *différentiation – dédifférentiation*, les aptitudes cognitives partagent de nombreux traits communs chez les enfants.

Par ailleurs, étant donné que la structure actuelle du WISC-IV est sujette à controverses, nous avons eu recours au modèle CHC, afin de nous aider à choisir les regroupements et les comparaisons à effectuer. Le choix de ce modèle repose sur le fait qu'il est devenu celui qui est le plus influent et qui a généré le plus de recherches. De plus, au regard de cette théorie, différentes études ont montré qu'une structure alternative était plus adéquate que la structure actuelle en quatre facteurs de la batterie de tests. En définitive, et comme nous l'avons vu précédemment, nous avons classé les épreuves selon 5 facteurs : l'intelligence fluide, l'intelligence cristallisée, la mémoire à court terme, le traitement visuel et la vitesse de traitement.

En ce qui concerne la vérification de la stabilité des performances inter-séances aux épreuves censées évaluer les mêmes aptitudes cognitives, certaines « contraintes » se sont posées. En effet, il était nécessaire, pour chaque séance, d'avoir au minimum une épreuve pour chacun des cinq facteurs. Ainsi, étant donné que certains de ces facteurs étaient composés de deux épreuves (Gf, Gsm et Gv), les comparaisons ne pouvaient être effectuées que sur deux sessions. Afin de pouvoir disposer d'une séance supplémentaire, nous avons eu alors recours à cinq autres sous-tests issus du K-ABC (*Triangles*), du NEPSY⁷

⁷ Il est à noter que les normes du NEPSY sont problématiques pour établir un pronostic mais, cette étude s'intéressant à la comparaison des performances des enfants entre eux, ce problème est moindre.

(*Compréhension de consignes et Répétition de phrases*) et du WJ-R (*Formation de concepts et Cross-out*). En ce qui concerne le K-ABC, l'épreuve des *Triangles* a été choisie car, selon le modèle CHC, elle évalue les mêmes aptitudes cognitives générales (Gv) et spécifiques (SR et VZ) que l'épreuve *Cubes* dans le WISC-IV. D'ailleurs, selon A. S. Kaufman et N. L. Kaufman (2004), on observe une corrélation de .66 entre ces deux épreuves. Ces observations nous amènent donc à penser que les différents sous-tests censés évaluer les mêmes aptitudes cognitives, comme par exemple les *Triangles* (K-ABC) et les *Cubes* (WISC-IV), devraient être consistants et situer chaque enfant au même niveau de performances. Ces épreuves, issues d'autres batteries de tests se répartissent et se définissent de la manière suivante :

1) Intelligence fluide (Gf) :

- **Formation de concepts (FC) :** ce sous-test issu du WJ-R serait un indicateur de l'*Induction* (I). Cette épreuve est fréquemment utilisée et est considérée comme l'un des meilleurs indicateurs de l'intelligence fluide.

2) Intelligence cristallisée (Gc) :

- **Compréhension de consignes (CC) :** ce sous-test issu du NEPSY serait un indicateur de l'*Aptitude d'écoute* (LS).

3) Mémoire à court terme (Gsm) :

- **Répétition de phrases (RP) :** issu du NEPSY, cette épreuve serait une mesure de l'*Empan mnésique* (MS).

4) Traitement visuel (Gv) :

- **Triangles (TG) :** cette épreuve, issue du K-ABC, serait une mesure des *Relations spatiales* (SR) et serait un très bon indicateur de *Visualisation* (VZ).

5) Vitesse de traitement (Gs) :

- **Cross-Out (Xout) :** issu du WJ-R, ce sous-test serait une mesure de la *Vitesse perceptive* (P)

Ensuite, afin d'établir s'il existe bel est bien une différence de performances inter-séances, nous avons eu recours aux travaux de Grégoire (2006). En effet, selon cet auteur, on considère qu'une différence supérieure à 4 points, entre deux épreuves censées évaluer les mêmes aptitudes cognitive, implique une fluctuation des performances. Il est à noter que lorsque différentes épreuves regroupées sous le même facteur ont été administrées lors de la même session, une analyse intra-séance a été effectuée. Si aucune différence de performance

n'a été observée, alors une épreuve était sélectionnée afin d'effectuer nos comparaisons inter-séances.

Enfin, si la stabilité a été vérifiée, nous avons calculé le pourcentage d'enfants qui pourraient néanmoins présenter des différences de performances. En effet, selon Grégoire (2006), la présence de fluctuations chez des individus tout-venants est relativement fréquente. Ainsi, on peut s'attendre à une différence de 4 points et plus sur l'ensemble de chaque paires d'épreuves administrées pour une proportion d'enfants allant jusqu'à environ 30% des sujets. En revanche, si une fluctuation importante est observée pour une très grande proportion d'enfant, alors nous pourrions en conclure que les performances ne sont pas restées stables et nous chercherons à déterminer si c'est l'étalement de l'administration des sous-tests sur plusieurs séances qui est en cause ou si d'autres facteurs peuvent être incriminés.

En définitive, nos hypothèses théoriques sont donc les suivantes : selon notre première hypothèse, nous nous attendons donc à ce que les performances aux sous-tests censés évaluer les mêmes aptitudes cognitives restent stables d'une épreuve et d'une séance à l'autre. Ainsi, cela devrait être le cas entre les épreuves *Matrices* (MA), *Identification de concepts* (IDC) et *Formation de concepts* (FC) pour le facteur *Intelligence fluide* (Gf) ; entre les sous-tests *Similitudes* (SI), *Raisonnement verbal* (RV) et *Compréhension de consignes* (CC) pour le facteur *Intelligence cristallisée* (Gc) ; entre les épreuves *Mémoire des chiffres* (MC), *Séquences lettres-chiffres* (SLC) et *Répétition de phrases* (RP) pour le facteur *Mémoire à court terme* (Gsm) ; entre les sous-tests *Cubes* (CU), *Complètement d'images* (CI) et *Triangles* (TG) pour le facteurs *Traitement visuel* (Gv) ; et entre les épreuves *Code* (CD), *Symboles* (SY) et *Cross-out* (Xout) pour le facteur *Vitesse de traitement* (Gs).

Selon notre seconde hypothèse, nous nous attendons à ce que les épreuves, regroupées selon le modèle CHC, évaluent les mêmes aptitudes cognitives pour notre échantillon d'enfants suisses-romands. De ce fait, les épreuves suivantes devraient mesurer les mêmes construits cognitifs : *Matrices* (MA), *Identification de concepts* (IDC) et *Formation de concepts* (FC) pour le facteur *Intelligence fluide* (Gf) ; *Similitudes* (SI), *Vocabulaire* (VO), *Compréhension* (CO), *Information* (IN), *Raisonnement verbal* (RV) et *Compréhension de consignes* (CC) pour le facteur *Intelligence cristallisée* (Gc) ; *Mémoire des chiffres* (MC), *Séquences lettres-chiffres* (SLC) et *Répétition de phrases* (RP) pour le facteur *Mémoire à court terme* (Gsm) ; *Cubes* (CU), *Complètement d'images* (CI) et *Triangles* (TG) pour le facteur *Traitement visuel* (Gv) ; *Code* (CD), *Symboles* (SY), *Barrage* (BA) et *Cross-out* (Xout) pour le facteur *Vitesse de traitement* (Gs).

3. Méthode

3.1 Participants

Nous avons recruté, dans des classes de 3^{ème} à 6^{ème} primaire, 250 enfants (125 garçons et 125 filles) âgés entre 8 et 12 ans. Ils étaient tous francophones, issus de milieux socio-économiques hétérogènes et se trouvaient dans le degré scolaire correspondant à leur âge chronologique. Ainsi, aucun d'entre eux n'avait doublé ni sauté de classe. Afin de pouvoir procéder à nos comparaisons pour évaluer les différences de performances entre les sessions aux épreuves censées évaluer les mêmes aptitudes cognitives, il était indispensable d'avoir un minimum d'une épreuve issue de chaque facteur pour chacune des trois séances d'évaluation. Les enfants qui ne respectaient pas cette condition ont alors été éliminés de notre échantillon. De plus, il était également important d'avoir des séances homogènes et contenant les mêmes épreuves. De ce fait, nous avons cherché à garder la plus grande proportion d'enfants qui ont effectués les mêmes sous-tests aux mêmes sessions (les deux premières séances étaient consacrées au WISC-IV et la troisième aux cinq épreuves des autres batteries de tests). Concernant les enfants qui ont eu besoin d'une séance supplémentaire pour achever l'ensemble des épreuves, ils ont aussi été écartés. Finalement, le sous-échantillon que nous avons constitué était composé de 105 enfants de 8 à 12 ans ($M = 9,75$, $SD = 1,30$).

3.2 Matériel

3.2.1 *L'échelle d'intelligence de Wechsler pour enfants - quatrième édition*

Le WISC-IV est destiné aux enfants de 6 ans à 16 ans et 11 mois. La durée de passation de la totalité du WISC est d'environ 60 à 90 minutes (pour les sous-tests principaux) selon l'âge et les aptitudes de l'enfant. Enfin, l'ordre de passation est prédéterminé.

1) Intelligence fluide (Gf) :

- **Matrices (MA) :** ce sous-test principal est censé permettre d'évaluer l'aptitude à découvrir les caractéristiques sous-jacentes qui sous-tendent une relation entre des éléments. Il implique la capacité à combiner des éléments d'informations séparés dans la formation d'inférences, de règles, d'hypothèses et de conclusions pour la résolution de problèmes par déduction et en plusieurs étapes. On présente à l'enfant une série ou une séquence d'images incomplètes (35 items) et on lui demande d'identifier la partie manquante, parmi cinq choix de réponses possibles. Par exemple,

pour l’item présenté ci-dessous (cf. *Figure 1*), la bonne réponse est le papillon bleu (proposition 1) car sa forme et sa couleur correspondent aux trois papillons de l’item. En ce qui concerne la cotation, on donne 1 point si la réponse est correcte et 0 point si elle est incorrecte. L’épreuve est arrêtée après 4 notes consécutives de 0 point. Le score maximum est de 35 points.

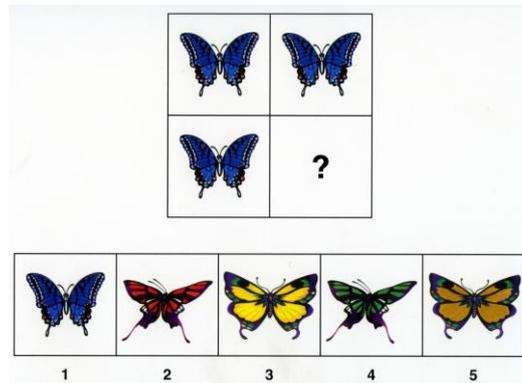


Figure 1. Illustration de la tâche *Matrices* (MA)

- **Identification de concepts (IDC) :** ce sous-test principal est censé mesurer l’aptitude à découvrir les caractéristiques sous-jacentes qui sous-tendent une relation entre des éléments. Il implique la capacité à combiner des éléments d’informations séparés dans la formation d’inférences, de règles, d’hypothèses et de conclusions. Ainsi, il est demandé à l’enfant d’identifier, parmi deux ou trois rangées d’images, celles qui peuvent être regroupées en fonction d’un concept commun (une image par rangée). Par exemple, pour l’item présenté ci-dessous (cf. *Figure 2*), la bonne réponse est l’image 3 de la première rangée et l’image 2 de la deuxième car ils correspondent à la même catégorie, c’est-à-dire aux poissons. Ce sous-test comporte 28 items (ex : *Figure 2*). Au niveau de la cotation, pour chaque item, on donne 1 point si l’image désignée dans chaque rangée est correcte et 0 point si les réponses sont incorrectes ou si l’enfant ne donne pas de réponse. On s’arrête suite à 5 notes consécutives de 0 point. Le score maximum est de 28 points.

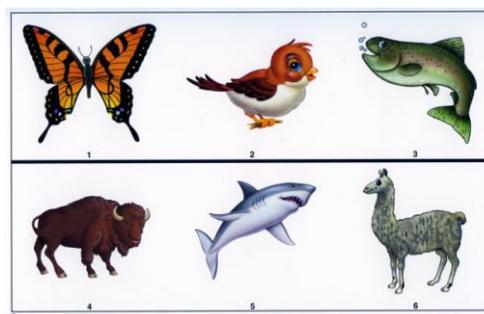


Figure 2. Illustration de la tâche *Identification de concepts* (IDC)

2) Intelligence cristallisée (Gc) :

- **Similitude (SI) :** selon le le modèle CHC, ce sous-test principal est censé évaluer le développement du langage, la compréhension et l'application de mots, de phrases, etc. de même que la capacité à utiliser et à comprendre correctement le sens des mots (noms, verbes, adjectifs). L'enfant doit décrire en quoi sont similaires des paires de mots présentés oralement et représentant un même objet ou concept (23 items). Exemple : « en quoi le bleu et le rouge se ressemblent ? Ou encore : « en quoi le chat et la souris se ressemblent ? ». En ce qui concerne la cotation, 2 points sont donnés si la réponse est pertinente, 1 point s'il existe un point commun avec la bonne réponse et 0 point si la réponse est incorrecte. Si l'enfant échoue cinq items à la suite, on s'arrête. Le score maximum est de 44 points.
- **Vocabulaire (VO) :** ce sous-test principal permettrait une mesure de la capacité à utiliser et à comprendre correctement le sens des mots (noms, verbes, adjectifs). Pour les enfants âgés jusqu'à 9 ans, on leur demande de nommer des images qui leurs sont présentées. Pour les plus âgés, c'est la définition de mots, présentés avec un support visuel, qui est demandée. La difficulté est d'ordre croissante (36 items). Exemples : « qu'est qu'un nénuphar ? », « que veut dire *transparent* ? ». En ce qui concerne la cotation, pour les items 1 à 4, on donne 1 point si la réponse est correcte et 0 si elle est incorrecte, si l'enfant ne répond pas ou si les réponses sont personnalisées (ex : « maman l'utilise », etc.). Pour les items 5 à 36, on accorde 2 points pour une description, un synonyme ou exemple pertinent du mot ou s'il donne la classe générale à laquelle il appartient (ex : « une *vache*, c'est un mammifère »). On accorde 1 point pour une réponse correcte mais pauvre (ex : un *parapluie*, c'est ce qu'on met au-dessus de la tête »), pour une utilisation secondaire de l'objet (ex : « un *gant*, c'est pour faire la vaisselle ») ou pour une définition d'un mot de la même famille (ex : si l'enfant définit « obéissant » au lieu d'*obéir* »). Enfin, 0 point est donné si la réponse est incorrecte, trop pauvre, montre l'incompréhension de l'enfant ou si des gestes démonstratifs sont utilisés. L'administration de l'épreuve s'arrête après 5 notes consécutives de 0 point. Le score maximum est de 65 points.
- **Compréhension (CO) :** cette épreuve principale est censée évaluer le développement général du langage, la compréhension et l'application de mots, phrases, etc. ainsi que la base de connaissance et l'aptitude à utiliser des connaissances générales. L'enfant doit répondre à des questions présentées oralement à partir de sa compréhension de problèmes de la vie quotidienne, de règles ou de concepts relatifs à la vie sociale, au

monde, etc. (21 items). Par exemple, il s'agira d'expliquer des règles sociales comme : « Que dois-tu faire, si tu vois de la fumée sortir de la fenêtre de la maison de ton voisin ? Que veut dire : quand le chat n'est pas là, les souris dansent ? ». Au niveau de la cotation, pour les items contenant une seule idée générale, on donne 2 ou 1 points si l'enfant évoque cette idée. Pour les items contenant plusieurs idées générales, l'enfant doit en évoquer au moins deux différentes pour obtenir 2 points. S'il n'évoque qu'une seule idée, on donne 1 point. Si la réponse est incorrecte, on donne 0 point. L'épreuve est arrêtée suite à 4 notes consécutives de 0 point. Le score maximum est de 42 points.

- **Information (IN)** : cette épreuve supplémentaire est censée mesurer la base de connaissances et l'aptitude à utiliser des connaissances. On demande à l'enfant de répondre à des questions portant sur des connaissances générales (en science, histoire, géographie, etc.). Les questions sont de difficulté croissante (33 items). Exemple : « Qu'est ce qu'un fossile ? A quoi sert l'estomac ? Qui était Charles Darwin ? ». Au niveau de la cotation, on donne 1 point si la réponse est correcte et 0 point si la réponse est incorrecte. L'épreuve est arrêtée suite à 5 notes consécutives de 0 point. Le score maximum est de 33 points.
- **Raisonnement verbal (RV)** : ce sous-test supplémentaire permettrait de mesurer la capacité à utiliser et à comprendre correctement le sens des mots (noms, verbes, adjectifs). A partir d'une série d'indices, l'enfant doit identifier un concept commun ou sous-jacent. En d'autres termes, les items prennent la forme de devinettes (24 items). Exemple : « C'est haut dans le ciel et on les voit seulement la nuit », « c'est un bâtiment où l'on fait du sport ». Pour la cotation : on donne 1 point pour une réponse correcte et 0 point pour une réponse incorrecte. L'épreuve est arrêtée suite à 5 notes consécutives de 0 point. Le score maximum est de 24 points.

3) Mémoire à court terme (Gsm) :

- **Mémoire des chiffres (MC)** : ce sous-test principal est censé évaluer la capacité à se rappeler une série d'éléments liés entre eux de manière aléatoire et à les reproduire dans l'ordre correct, après quelques secondes et en une seule présentation,. Il évalue également l'aptitude à stocker temporairement des informations sur lesquelles un ensemble d'opérations cognitives sont effectuées. Il est demandé à l'enfant de répéter des séquences de chiffres (de 2 à 9 chiffres), présentées oralement, dans un même ordre et dans un ordre inverse, dans une deuxième partie de la tâche. Ce sous-tests

comporte 8 items pour l'ordre direct et 8 items pour l'ordre inverse. Exemple : pour l'ordre direct, si l'énoncé est « 3-4-1-7 », l'enfant devra répondre « 3-4-1-7 ». Pour l'ordre inversé, lors de la présentation de l'énoncé « 2-4-3-7-4 », l'enfant devra répondre : 4-7-3-4-2. Au niveau de la cotation, 1 point est donné si la réponse est correcte ou 0 si elle est incorrecte. La note brute totale correspond à la somme des notes aux items des deux parties de l'épreuves (ordre direct et ordre inverse). Des notes additionnelles peuvent être calculées⁸. On arrête l'administration du test après une note de 0 point aux deux essais d'un même item, que ce soit pour l'ordre direct ou inverse. Le score maximum est de 32 points.

- **Séquences lettres-chiffres (SLC) :** ce sous-test principal est censé mesurer l'aptitude à stocker temporairement des informations sur lesquelles un ensemble d'opérations cognitives sont effectuées. L'enfant doit répéter, pour chaque séquence entendue, des chiffres et des lettres. En premier les chiffres par ordre croissant, puis les lettres par ordre alphabétique (10 items comprenant chacun trois essais). Par exemple, pour l'énoncé 7-B-P-M-4, l'enfant devra répondre : 4-7-B-M-P. En ce qui concerne la cotation, on donne 1 point si l'enfant restitue tous les chiffres et toutes les lettres dans l'ordre correct de la séquence et 0 point dans le cas contraire. La note à un item consiste en la somme des notes des trois essais (de 0 à 3 points). L'épreuve est arrêtée suite à 3 notes de 0 point aux trois essais d'un même item. Le score maximum est de 30 points.

4) Traitement visuel (Gv) :

- **Cubes (CU) :** ce sous-test principal est censé permettre d'évaluer la capacité à résoudre des problèmes impliquant des relations spatiales à travers l'identification et la manipulation de stimuli visuels selon différentes perspectives (angles ou positions). Il évalue également l'aptitude à appréhender, à manipuler, à transformer mentalement des stimuli visuels en deux et en trois dimensions et à les faire correspondre. On demande alors à l'enfant de regarder des motifs géométriques bidimensionnels et de les reproduire à l'aide de cubes (tridimensionnel) tricolores (rouge, blanc, rouge et blanc), en un temps limité. La difficulté est croissante (14 items). Par exemple, pour l'item présenté ci-dessous (cf. *Figure 3*), l'enfant devra utiliser quatre cubes pour reproduire le motif : deux faces blanches et deux faces bicolores. Le temps de

⁸ La note additionnelle pour l'ordre direct correspond au nombre de chiffres restitué correctement lors du dernier essai (9 points maximum). La note additionnelle pour l'ordre inverse correspond au nombre de chiffres restitués correctement lors du dernier essai (8 points maximum).

réalisation est de 30 secondes pour le premier item, de 45 secondes pour les items 2 à 5, de 75 secondes pour les items 6 à 10 et de 120 secondes pour les items 11 à 14. Le temps nécessaire à la réalisation ainsi que l'échec ou la réussite doivent être indiqués. En ce qui concerne la cotation, pour les items 1 à 3, on donne 2 points si la construction est réalisée correctement, sans rotation (< 30 degrés) et dans les temps ; on accorde 1 point si la construction est réalisée correctement, sans rotation, dans le temps imparti lors du deuxième essais ; et on donne 0 point si la réalisation est incorrecte, comporte une rotation ou si elle est inachevée dans le temps imparti. Pour les items 4 à 8, on accorde 4 points si la construction est réalisée correctement, sans rotation et dans les temps ou 0 point si la construction est incorrecte, comporte une rotation ou est inachevée. Enfin, pour les items 9 à 14, on donne 4, 5, 6 ou 7 points selon le temps utilisé par l'enfant⁹, si la construction est correcte, sans rotation et réalisée dans les temps. Une note additionnelle peut être calculée à partir de la note brute¹⁰. On arrête la passation après 3 notes consécutives de 0 point. Le score maximum est de 68 points.

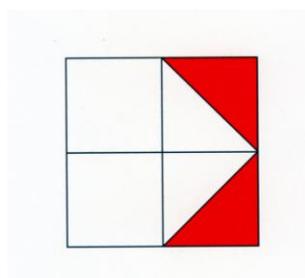


Figure 3. Illustration de la tâche Cubes (CU)

- **Complètement d'images (CI) :** ce sous-test supplémentaire est censé évaluer l'aptitude à identifier une caractéristique visuelle particulière intégrée dans un schéma complexe de distraction visuelle. Il est censé évaluer la base de connaissances et l'aptitude à utiliser des connaissances générales. On présente à l'enfant des images (38 items) pour lesquelles il doit identifier, dans un temps limité (20 secondes par image), le détail manquant le plus important. Par exemple, pour l'item présenté ci-dessous (cf. Figure 4), la partie manquante est l'oreille du renard. Au niveau de la cotation, on accorde 1 point si l'enfant identifie la partie manquante dans les délais et

⁹ Pour les items 9 et 10, on donne 4 points si le temps est compris entre 31 et 75 secondes, 5 points entre 21 et 30 secondes, 6 points entre 11 et 20 secondes et 7 points entre 1 et 10 secondes. Pour les items 11 à 14, on donne 4 points si le temps est compris entre 71 et 120 secondes, 5 points entre 51 et 70 secondes, 6 points entre 31 et 50 secondes et 7 points entre 1 et 10 secondes.

¹⁰ On note 2, 1 ou 0 points pour les items 1 à 3 et on donne 4 points, pour les items 4 à 14, si la construction est réalisée correctement, sans rotation et dans les temps. On donne 0 points, si la construction est incorrecte, comporte une rotation ou n'est pas réalisée dans les temps. La note maximale est de 50 points.

0 point, si la réponse est incorrecte ou s'il n'a pas répondu à temps. L'épreuve est arrêtée suite à 6 notes consécutives de 0 point. Le score maximum est de 38 points.



Figure 4. Illustration de la tâche *Complètement d'images* (CI)

5) Vitesse de traitement (Gs) :

- **Code (CD) :** ce sous-test principal aurait pour but la mesure de l'aptitude à effectuer des opérations rapides (décisions simples). L'enfant doit, dans un temps limité de 120 secondes, copier des symboles appariés à des chiffres dans la case correspondante. Le Code A est destiné aux enfants entre 6 et 7 ans et le Code B à ceux entre 8 et 16 ans (cf. *Figure 5*). En ce qui concerne la cotation, par exemple pour le Code A, 59 points sont donnés si l'épreuve est réalisée entre 116 et 120 secondes, 60 points entre 111 et 115 secondes (1 point de bonification), 61 points entre 106 et 110 secondes (2 points de bonification), 62 points entre 101 et 105 secondes (3 points de bonification), 63 points entre 96 et 100 secondes (4 points de bonification), 64 points entre 86 et 95 secondes (5 points de bonification) et 65 points si le temps est égal ou inférieur à 85 secondes (6 points de bonification). Le score maximum est de 65 points (points de bonification inclus) pour le Code A et de 119 points pour le Code B (pas de points de bonification).

Code B
Age 8-16 ans

1	2	3	4	5	6	7	8	9
÷)	+	⊥	⌊	∇	⊂	÷	⊖

ITEMS D'EXEMPLE

2	1	4	6	3	5	2	1	3	4	2	1	3	1	2	3	1	4	2	6	3
1	2	5	1	3	1	5	4	2	7	4	6	9	2	5	8	4	7	6	1	8
7	5	4	8	6	9	4	3	1	8	2	9	7	6	2	5	8	7	3	6	4
5	9	4	1	6	8	9	3	7	5	1	4	9	1	5	8	7	6	9	7	8
2	4	8	3	5	6	7	1	9	4	3	6	2	7	9	3	5	6	7	4	5
2	7	8	1	3	9	2	6	8	4	1	3	2	6	4	9	3	8	5	1	8

Figure 5. Illustration de la tâche *Code* (CD)

- **Symboles (SY)** : ce sous-test principal permettrait une mesure de l'aptitude à chercher rapidement, de manière précise et correctement des similitudes ou des différences sur des éléments présentés côte à côte (comparaisons) et à effectuer des opérations rapides (décisions simples). Il est demandé à l'enfant, dans un temps limité de 120 secondes, de regarder des séries de un ou deux symboles isolés et une série de trois ou cinq symboles et de décider, pour chaque série présentée, si oui ou non un symbole cible est présent. Par exemple, pour le premier item présenté ci-dessous (cf. *Figure 6*), la bonne réponse est « oui » car le symbole du cercle avec un « + » au milieu se retrouve dans la série juxtaposée. En ce qui concerne la cotation, le score correspond au nombre de réponses correctes moins le nombre de réponses incorrectes. Les omissions et les items non-complétés après la limite de temps ne sont pas pris en considération. Si l'enfant répond à la fois oui et non, on considère que l'item est incorrect. Le score maximum est de 45 points pour les Symboles A (pour les enfants de 6 à 7 ans) et de 60 points pour les Symboles B (pour les enfants de 8 à 16 ans).

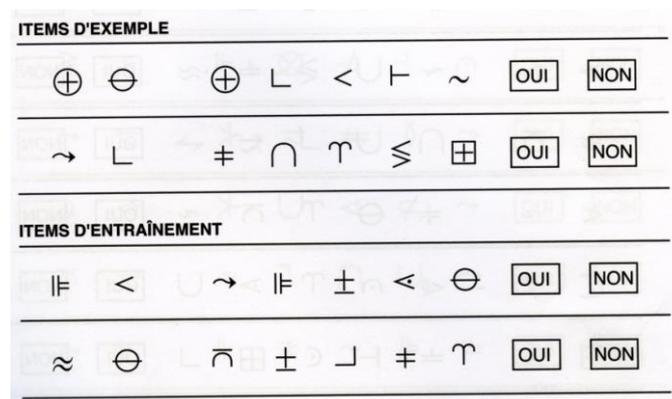


Figure 6. Illustration de la tâche *Symboles (SY)*

- **Barrage (BA)** : ce sous-test supplémentaire permet de mesurer de l'aptitude à chercher rapidement, de manière précise et correctement des similitudes ou des différences sur des éléments présentés côte à côte (comparaisons) et à effectuer des opérations rapides (décisions simples). Dans un temps limité de 45 secondes, il est demandé à l'enfant de repérer et de barrer les images d'animaux à travers un ensemble aléatoire d'images (ex : *Figure 7*). Ensuite, il doit répéter l'exercice mais avec un ensemble structuré d'images (ex : *Figure 8*). Pour la cotation, on comptabilise le nombre total d'images barrées correctement. Si l'enfant a donné des réponses incorrectes, il faut alors les soustraire du nombre de bonnes réponses. Si la différence est inférieure à 60, on enregistre alors cette différence comme note brute

a pour objectif d'évaluer le développement cognitif chez les enfants âgés de 2 ans et 6 mois à 12 ans et 11 mois. La batterie comporte 16 sous-tests dont l'administration se réalise dans un ordre constant. Toutefois, le nombre d'épreuves varie selon l'âge de l'enfant : un minimum de 7 épreuves pour les enfants de 7 ans et 6 mois et un maximum de 13 sous-tests pour ceux âgés de 7 à 12 ans et 6 mois. Enfin, il faut compter entre 35 et 85 minutes pour administrer le K-ABC (le temps de passation augmente avec l'âge).

Cette échelle reprend la dichotomie des processus de traitement de l'information décrits par les travaux en psychologie cognitive et des théories neuropsychologiques de Luria. De ce fait, cette batterie de tests se divise en deux composantes principales : (1) une échelle des processus mentaux composites qui regroupe deux sous-échelles (processus séquentiels et processus simultanés) et (2) une échelle qui évalue les connaissances acquises au cours des apprentissages. Il est à noter que la batterie comporte également une échelle supplémentaire non verbale (en cas de trouble auditif, du langage, etc.).

En ce qui concerne les épreuves, seule celle qui a été utilisée dans cette étude sera présentée ici. Il s'agit de l'épreuve *Triangles* qui évalue, selon le modèle CHC, le Traitement visuel (Gv) :

- Traitement visuel (Gv) :

- **Triangles (TG)** : cette épreuve est destinée aux enfants âgés de 4 à 12 ans et 6 mois. Selon le modèle CHC, elle évalue la capacité à résoudre des problèmes impliquant des relations spatiales à travers l'identification et la manipulation de stimuli visuels selon différentes perspectives (angles ou positions). Elle évalue également l'aptitude à appréhender, à manipuler, à transformer mentalement des stimuli visuels en deux et en trois dimensions et à les faire correspondre (24 items). Ce sous-test est très semblable à la tâche des *Cubes* du WISC-IV, à la différence que les triangles sont bicolores et comportent uniquement deux faces (l'une jaune et l'autre bleue). Elle évalue la capacité à résoudre des problèmes impliquant des relations spatiales à travers l'identification et la manipulation de stimuli visuels selon différentes perspectives (angles ou positions). Ainsi, on demande à l'enfant de reconstruire, à partir de triangles identiques, l'image de dessins abstraits. Par exemple, pour l'item présenté ci-dessous (cf. *Figure 9*), l'enfant devra utiliser quatre triangles pour reproduire le motif : trois faces bleues et une face jaune. L'enfant dispose de 30 secondes à partir de l'item 10 (le temps n'est pas pris en compte dans les précédents). Ensuite, il dispose de 45 secondes pour les items 15 à 19, de 90 secondes pour les items 20 à 21 et de 105 secondes pour les items 22 à 24. En ce qui concerne la

cotation, on donne, pour les deux premiers items, 2 points par réponse correcte au premier essai et 1 point pour une réponse correcte au deuxième essai. En ce qui concerne les items 3 à 14, on accorde 1 point par réponse correcte. Pour les items 15 à 24, 1 à 3 points sont accordés selon le temps mis par l'enfant dans le délai imparti. Enfin, on donne 0 point si l'enfant échoue ou dépasse le temps limite. L'épreuve est arrêtée suite à 3 notes consécutives de 0 point. Le score maximum est de 46 points.



Figure 9. Illustration de la tâche *Triangles* (TR)

3.2.3 *The Korkman Neuropsychological Scale (NEPSY)*

Le NEPSY est utilisé pour l'évaluation neuropsychologique du développement de l'enfant âgé de 3 à 12 ans. Il a été élaboré par Marit Korkman, Ursula Kirk et Sally Kemp en 1997. La version française date de 2003. Il est basé sur les méthodes cliniques de Luria ainsi que sur des théories neuropsychologiques de l'enfant. Cette batterie de 27 tests permet d'identifier les forces et les faiblesses de l'enfant et d'analyser d'éventuels déficits à travers cinq domaines : les fonctions attentionnelles et exécutives, le langage, les fonctions sensori-motrices, le traitement visuo-spatial ainsi que la mémoire et les apprentissages. La passation de l'ensemble des tâches requiert plus de deux heures.

En ce qui concerne les épreuves, et encore une fois, seules celles qui sont utilisées dans cette étude seront présentées ici. Il s'agit des épreuves *Répétition de phrases* et *Compréhension de consignes* qui évaluent respectivement, selon le modèle CHC, la Mémoire à court terme (Gsm) et l'Intelligence cristallisée (Gc) :

- Mémoire à court terme (Gsm) :
 - o **Répétition de phrases (RP) :** ce sous-test évalue l'aptitude à se rappeler une série d'éléments liés entre eux et à les reproduire dans l'ordre correct, après quelques secondes et une seule présentation. On va alors demander à l'enfant, âgé de 3 à 12 ans, de répéter des phrases de complexité et de longueur croissante. Exemples : « les enfants font la queue pour déjeuner », « lorsque le soleil s'est couché, nous avons monté nos tentes ». En ce qui concerne la cotation, 2 points sont donnés par items pour chaque répétition correcte, 1 point si l'enfant commet une ou deux erreurs et 0

point s'il commet plus de deux erreurs (omissions, ordre des mots, changement ou ajout d'un mot). L'épreuve est arrêtée après 4 notes consécutives de 0 point. La note maximum est de 34 points.

- Intelligence cristallisée (Gc) :

- **Compréhension de consignes (CC) :** dans ce sous-test, pour les enfants âgés de 3 à 12 ans, on évalue, selon le modèle CHC, l'aptitude à écouter et à comprendre le sens de la communication orale. Pour se faire, on montre à l'enfant des formes géométriques et on lui demande de désigner certaines d'entre elles, en fonction de leur couleur, de leur position et de leur relation à d'autres figures (28 items). Exemple : « Montre-moi une croix blanche après m'avoir montré une forme rouge sous une forme bleue » (ex : *Figure 10*). En ce qui concerne la notation, 1 point est donné par réponse correcte et 0 point pour une réponse incorrecte ou donnée dans le mauvais ordre. L'épreuve est arrêtée suite à 4 échecs consécutifs. Le score maximum est de 28 points.

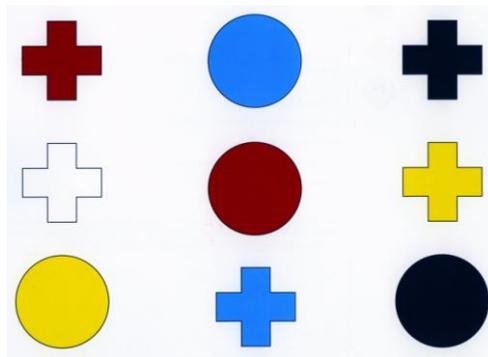


Figure 10. Illustration de la tâche *Compréhension de consignes (CC)*

3.2.4 *The Woodcock-Johnson Psycho-Educational Battery Test – Revised (WJ-R)*

La première version de cette batterie de test date de 1977. Elle a été élaborée par Richard Woodcock et révisée en 1989 sous le nom de WJ-R. Il est à noter qu'actuellement il n'existe pas de version française et que les épreuves utilisées ici ont été traduites afin de pouvoir les utiliser. Elle est administrée individuellement et a pour objectif d'évaluer les aptitudes cognitives mais également la réussite et l'intérêt scolaire, le potentiel par rapport au rendement. Elle est conçue pour des individus âgés de 2 à 90 ans et est organisée en items de difficulté croissante, en fonction de l'âge. Le WJ-R est basé sur la théorie Gf-Gc de Cattell-Horn. Il est composé de 35 sous-tests dont 21 épreuves évaluant les aptitudes cognitives (Raisonnement fluide, Compréhension/connaissances, Mémoire à court terme, Mémoire à

long terme, Vitesse de traitement, Traitement visuel et Traitement auditif) ainsi que 14 épreuves évaluant les performances (Mathématiques, Langage écrit et Lecture).

En ce qui concerne les épreuves, et de nouveau, seules celles qui sont utilisées dans cette étude seront présentées ici. Il s'agit des épreuves *Formation de concepts* et *Cross-out* qui évaluent respectivement, selon le modèle CHC, l'Intelligence fluide (Gf) et la Vitesse de traitement (Gv) :

- Intelligence fluide :

- **Formation de concepts (FC) :** cette tâche évalue l'aptitude à découvrir les caractéristiques sous-jacentes qui sous-tendent une relation entre des éléments. Il implique la capacité à combiner des éléments d'informations séparés dans la formation d'inférences, de règles, d'hypothèses et de conclusions. Ainsi, pour un ensemble de formes géométriques de couleurs alignées, on demande à l'enfant, pour chaque item, de trouver la règle, le concept sous-jacent et de dire pourquoi certaines formes sont encadrées (35 items). Il est à noter que cette tâche nécessite le concept de « ET » et « OU ». Par exemple, pour l'item 12 présenté ci-dessous (cf. *Figure 11*), la bonne réponse est « carré » ; pour l'item 22, la réponse est « jaune ET carré » ; et pour l'item 37, la réponse est « rond OU rouge OU double (par deux) ». En ce qui concerne la notation, 1 point est attribué par réponse correcte et 0 point par réponse incorrecte. Le score maximum est de 35 points.

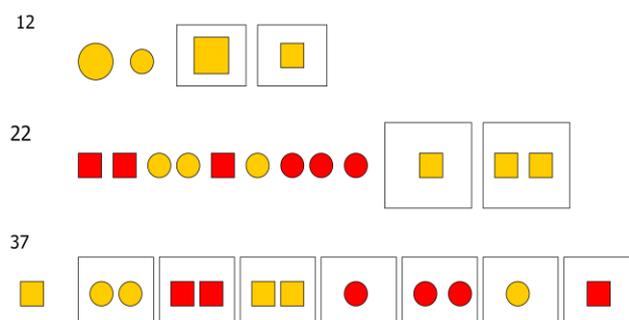


Figure 11. Illustration de la tâche *Formation de concepts* (FC)

- Vitesse de traitement (Gs) :

- **Cross-out (Xout) :** cette épreuve mesure l'aptitude à rechercher rapidement, de manière précise et correctement des similitudes ou des différences sur des éléments visuels présentés côte à côte (comparaisons). Pour se faire, on demande à l'enfant, dans un temps limité de 3 minutes, d'identifier et de barrer les dessins identiques sur des séries de 20 figures. Par exemple, pour le premier item de la figure présentée ci-dessous (cf. *Figure 12*), l'enfant devra repérer les triangles avec un « + » au milieu. Il

doit essayer de compléter autant de lignes que possible. Le score maximum est de 37 points.

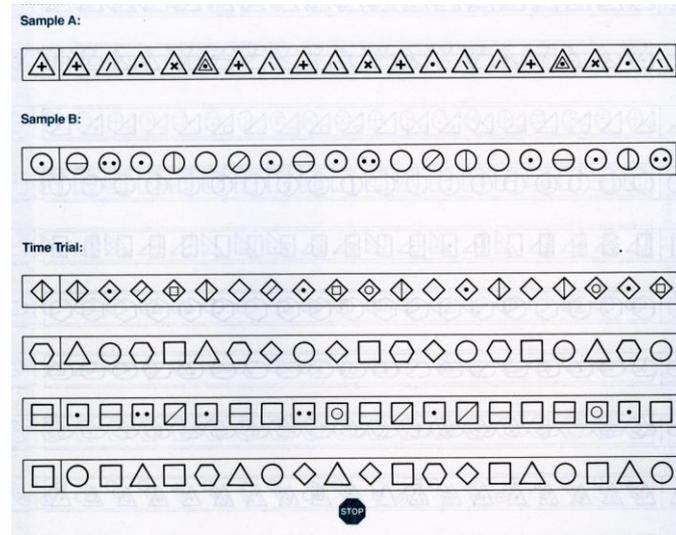


Figure 12. Illustration de la tâche Cross-out (Xout)

3.3 Procédure

Les enfants ont été sélectionnés par tirage au sort par leur enseignant(e) et avec l'autorisation du Département de l'instruction publique (DIP) du Canton de Genève et/ou des représentants légaux des enfants. Pour des raisons de limites pratiques, nous n'avons examiné qu'un petit échantillon d'enfants dans plusieurs écoles du canton de Genève, durant les heures scolaires et en présence de deux examinateurs.

Nous avons administré l'intégralité des sous-tests du WISC-IV (15 sous-tests) ainsi que cinq épreuves issus de trois autres batteries de tests : l'épreuve *Triangle* du K-ABC, les épreuves *Compréhension de consignes* et *Répétition de phrases* du NEPSY ainsi que les sous-tests *Formation de concepts* et *Cross-out* du WJ-R. Par ailleurs, les passations, ainsi que les cotations, ont été faites en suivant les procédures décrites dans les manuels d'administration et de cotation.

Pour l'ensemble de notre échantillon, nous avons eu recours à trois séances individuelles, à un intervalle de temps d'environ une semaine, dans les locaux des différentes écoles. Lors de la première séance, nous avons administré une première partie des épreuves du WISC-IV : *Cubes*, *Similitudes*, *Mémoire des chiffres*, *Identification de concepts*, *Code* et *Vocabulaire*. Lors d'une deuxième séance, nous avons administré les autres sous-tests restant du WISC-IV : *Séquences Lettres-Chiffres*, *Matrices*, *Compréhension*, *Symboles*, *Complètement d'images*, *Barrage*, *Information*, *Arithmétique* et *Raisonnement verbal*. Enfin, lors de la

troisième séance, nous avons administré les épreuves des autres batteries de tests : *Triangles*, *Compréhension de consignes*, *Répétition de phrases*, *Formation de concepts* et *Cross-out*. Il est à noter qu'une quatrième séance a parfois été nécessaire, dans le cas où l'enfant n'était pas parvenu à effectuer l'ensemble des tâches dans le temps imparti. Dans ce cas de figure, les enfants ont été éliminés de notre échantillon.

Enfin, cette recherche respecte le code déontologique de la Société Suisse de Psychologie (SSP) et de la Fédération Suisse des Psychologues (FSP). Bien entendu, toutes les informations obtenues auprès des enfants demeurent confidentielles et sont conservées de manière anonyme. Il n'y a pas eu de retour individuel.

3.4 Synthèse

En définitive et afin de procéder à un récapitulatif des comparaisons et des regroupements que nous avons effectués dans cette étude, le tableau présenté ci-dessous (cf. Tableau 1) résume l'ensemble des sous-tests sur lesquels nos analyses ont porté. Il est à noter que le sous-test *Arithmétique* (AR) n'y figure pas, étant donné qu'il est censé mesurer plusieurs aptitudes cognitives différentes (Gq, Gsm et Gc).

Tableau 1 : Récapitulatif du regroupement des épreuves selon un modèle CHC

Facteurs	WISC-IV	K-ABC	NEPSY	WJ-R
Intelligence fluide (Gf)	MA (<i>I et RG</i>) IDC (<i>I</i>) SI (<i>LD et VL</i>) VO (<i>VL</i>)			FC (<i>I</i>)
Intelligence cristallisée (Gc)	CO (<i>K0 et LD</i>) IN (<i>K0</i>) RV (<i>VL</i>)		CC (<i>LS</i>)	
Mémoire à court terme (Gsm)	MC (<i>MS et MW</i>) SLC (<i>MW</i>)		RP (<i>MS</i>)	
Traitement visuel (Gv)	CU (<i>SR et VZ</i>) CI (<i>CF et Gc-K0</i>) CD (<i>R9</i>)	TG (<i>SR et VZ</i>)		
Vitesse de traitement (Gs)	SY (<i>P et R9</i>) BA (<i>P et R9</i>)			Xout (<i>P</i>)

En ce qui concerne les hypothèses opérationnelles élaborées dans cette recherche, elles sont donc les suivantes : selon notre première hypothèse, nous nous attendons à ce que les différences de performances entre les épreuves censées évaluer les mêmes aptitudes

cognitives soient égales ou inférieures à 4 points entre les épreuves *Matrices* (MA), *Identification de concepts* (IDC) et *Formation de concepts* (FC) pour le facteur *Intelligence fluide* (Gf) ; entre les épreuves *Vocabulaire* (VO), *Raisonnement verbal* (RV) et *Compréhension de consignes* (CC) pour le facteur *Intelligence cristallisée* (Gc) ; entre les épreuves *Mémoire des chiffres* (MC), *Séquences lettres-chiffres* (SLC) et *Répétition de phrases* (RP) pour le facteur *Mémoire à court terme* (Gsm) ; entre les épreuves *Cubes* (CU), *Complètement d'images* (CI) et *Triangles* (TG) pour le facteur *Traitement visuel* (Gv) ; et entre les sous-tests *Code* (CD), *Symboles* (SY) et *Cross-out* (Xout) pour le facteur *Vitesse de traitement* (Gs).

Selon notre seconde hypothèse, nous nous attendons à ce que les épreuves censées évaluer les mêmes aptitudes cognitives soient corrélées entre elles avec des intensités élevées et proches de celles présentées dans le manuel du WISC-IV. Les épreuves sont les suivantes : *Matrices* (MA), *Identification de concepts* (IDC) et *Formation de concepts* (FC) pour le facteur *Intelligence fluide* (Gf) ; *Similitudes* (SI), *Vocabulaire* (VO), *Compréhension* (CO), *Information* (IN), *Raisonnement verbal* (RV) et *Compréhension de consignes* (CC) pour le facteur *Intelligence cristallisée* (Gc) ; *Mémoire des chiffres* (MC), *Séquences lettres-chiffres* (SLC) et *Répétition de phrases* (RP) pour le facteur *Mémoire à court terme* (Gsm) ; *Cubes* (CU), *Complètement d'images* (CI) et *Triangles* (TG) pour le facteur *Traitement visuel* (Gv) ; *Code* (CD), *Symboles* (SY), *Barrage* (BA) et *Cross-out* (Xout) pour le facteur *Vitesse de traitement* (Gs).

4. Résultats

4.1 Statistiques descriptives

D'un point de vue descriptif, en ce qui concerne les quatre indices du WISC-IV, nous constatons que les performances moyennes des enfants à l'ICV, l'IRP, l'IMT et l'IVT sont respectivement de 107,70 ($SD = 14,71$) ; de 100,81 ($SD = 13,31$) ; de 94,38 ($SD = 13,23$) ; et de 107,95 ($SD = 13,3$). Au niveau des écart-types, ils sont tous proches de 15. Au regard de ces résultats, on constate donc que les performances moyennes et leurs écart-types sont très proches de ceux de la population de référence, c'est-à-dire de 100 et d'un écart-type de 15. Il est à noter que ces moyennes et leur écart-type ne nous donnent qu'une vision globale des résultats. En effet, certains enfants s'écartent de ces moyennes avec une différence allant de deux, voire trois fois la valeur de l'écart-type, se situant ainsi dans les extrêmes supérieurs et

inférieurs. Enfin, les indices de tendance centrale ont des valeurs très similaires au sein des quatre indices, ce qui indique que les distributions des performances sont proches d'un échantillon de standardisation (cf. Tableau 2).

Pour les sous-tests du WISC-IV, les performances moyennes ainsi que les écarts-types sont très proches les uns des autres et également proches de 10 et d'un écart type de 3. Ainsi, ces valeurs correspondent à celles de l'échantillon d'étalonnage mentionnées dans le manuel du WISC-IV. Par ailleurs, les indices de tendance centrale ont des valeurs très similaires au sein des différentes épreuves, indiquant encore une fois que les distributions sont relativement symétriques et proche d'un échantillon de standardisation (cf. Tableau 2). On peut constater aussi que quelques enfants ont présenté des scores se situant jusqu'à 3 écarts-types en dessous ou en dessous de la moyenne (sujets extrêmes).

En définitive, et étant donné que la performance moyenne absolue des différents indices et sous-tests sont proches, respectivement, de 100 ($SD = 15$) et de 10 ($SD = 3$), nous sommes en droit d'affirmer ici que, globalement, les résultats correspondent à un échantillon « tout-venant ».

Tableau 2 : Résultats descriptifs pour les sous-tests, les indices et le QI

Sous-tests	N	Moyenne	Ecart type	Médiane	Mode	Minimum	Maximum
Compréhension Verbale	105	107,70	14,713	108	112	74	155
Raisonnement Perceptif	105	100,81	13,310	102	104	65	135
Mémoire de Travail	105	94,38	13,228	94	94	58	121
Vitesse de Traitement	105	107,95	13,300	109	109	81	148
Matrices	105	9,39	2,687	9	8	1	15
Identification de concepts	105	9,58	2,507	10	8	4	15
Similitudes	105	11,89	3,014	11	10	3	19
Vocabulaire	105	11,01	2,593	11	12	4	18
Compréhension	105	10,93	2,836	11	10	4	19
Information	105	10,66	3,382	11	12	3	19
Raisonnement verbal	105	8,5	3,083	9	10	1	16
Mémoires de chiffres	105	9,28	2,532	10	7	3	15
Séquence lettres-chiffres	105	8,84	2,697	10	10	1	14
Cubes	105	11,50	3,083	11	11	3	19
Complètement d'images	105	10,14	2,644	10	10	1	16
Code	105	10,97	2,712	11	10	3	19
Symboles	105	11,70	2,579	12	13	5	19
Barrages	105	11,35	2,852	12	12	1	19
Arithmétique	105	8,66	3,103	9	10	1	16
Cross-out	105	12,491	2,9148	12,8	10,4	5,8	18,2
Formation de concept	105	13,76	2,3722	13,6	12,2	6,8	19,8
Triangles	105	11,17	2,168	12	12	5	16
Compréhension consignes	105	9,77	2,956	10	9	2	15
Répétitions de phrases	105	9,84	3,154	10	13	2	16

En ce qui concerne les corrélations entre les indices, nos analyses ont indiqué les résultats suivants : à la différence des corrélations entre l'*Indice de Vitesse de traitement* (IVT) et l'*Indice de Mémoire de travail* (IMT) ainsi qu'entre l'*Indice de Compréhension verbale* (ICV) et l'*Indice de Vitesse de traitement* (IVT), qui ne sont pas significatives et qui indiquent qu'ils ne mesurent pas les mêmes construits cognitifs, les corrélations entre tous les autres indices sont significatives à un seuil de significativité 1% (cf. Tableau 3). En effet l'ICV du WISC-IV présente une corrélation de .43 ($p < 0,1$) avec l'*Indice de Raisonnement perceptif* (IRP), et de .27 ($p < .01$) avec l'IMT. De plus, la corrélation entre l'IRP et l'IMT est de .30 ($p < .01$) et celle entre l'IRP et l'IVT est de .31 ($p < .01$). Toutefois, en se référant aux valeurs mentionnées dans le manuel du WISC-IV, on constate que ces corrélations sont relativement proches, ce qui indique une concordance.

Tableau 3 : Corrélations pour les indices du WISC-IV

Variabes	1	2	3	4
1. Indice de Compréhension verbale		.51	.43	.29
2. Indice de Raisonnement perceptif	.430**		.42	.34
3. Indice de Mémoire de travail	.273**	.297**		.27
4. Indice de Vitesse de traitement	.150	.312**	.042	

** : $p < .01$; NB : en dessous de la diagonale = valeurs absolues ; au-dessus de la diagonale = valeurs rapportées dans le manuel du WISC-IV

De manière générale, au niveau des corrélations entre les différents sous-tests censées évaluer les mêmes aptitudes cognitives, nous constatons qu'elles sont significatives et relativement proches de celles du manuel du WISC-IV (cf. Tableau 4). Ainsi, par exemple, la corrélation entre les sous-tests *Cubes* (CU) et *Complètement d'images* (CI) de .38 ($p < .01$) dans notre étude est identique à celle mentionnée dans le manuel ($r = .38$). Par ailleurs, pour la corrélation entre les épreuves *Similitudes* (SI) et *Vocabulaire* (VO), elle est de .60 pour notre échantillon et est très proche de celle indiquée par le manuel ($r = .67$). Au regard de ces résultats, nous pouvons donc dire que nos résultats concordent de manière générale avec les données issues du manuel du WISC-IV. Toutefois, il est à constater que bien que les épreuves censées évaluer les mêmes aptitudes cognitives soient corrélées entre elles de manière significatives et avec des intensités relativement élevées, on peut observer que certaines d'entre elles corréles également avec d'autres sous-tests. Par exemple, l'épreuve *Identification de concept* (IDC) est corrélée conformément avec le sous-test *Matrices* (MA) mais également avec *Vocabulaire* (VO) et *Raisonnement verbal* (RV). En outre, l'épreuve *Matrices* (MA) est

également corrélée avec le sous-test *Cubes* (CU). Ainsi, il semble donc, et en accord avec la littérature, que certaines épreuves soient aussi sous-tendues par d'autres aptitudes cognitives.

Tableau 4 : Corrélations pour les sous-tests du WISC-IV

	CU	SI	MC	IDC	CD	VO	SLC	MA	CO	SY	CI	BA	IN	RV	AR
CU		.37	.26	.26	.24	.32	.26	.40	.21	.30	.38	.16	.35	.27	.37
SI	.302**		.33	.34	.21	.67	.34	.38	.54	.25	.38	.16	.62	.55	.48
MC	.105	.207*		.22	.12	.30	.47	.30	.25	.17	.20	.13	.31	.26	.39
IDC	.058	.317**	.237**		.15	.35	.25	.37	.29	.21	.36	.16	.26	.28	.32
CD	.103	.021	.077	.086		.22	.23	.20	.17	.51	.13	.33	.16	.16	.24
VO	.229**	.596**	.172*	.328**	-.018		.38	.39	.56	.25	.41	.15	.65	.54	.52
SLC	.186*	.187*	.445**	.131	-.165*	.260**		.33	.27	.27	.23	.17	.39	.31	.49
MA	.415**	.349**	.289**	.264**	.068	.397**	.187*		.28	.23	.33	.16	.36	.30	.42
CO	.118	.587**	.167*	.134	-.120	.583**	.209*	.186*		.20	.34	.16	.48	.48	.36
SY	.432**	.335**	.079	.189*	.367**	.281**	.140	.240**	.229**		.18	.33	.23	.26	.31
CI	.382**	.306**	.023	.063	.046	.233**	.099	.198*	.195*	.280**		.15	.36	.36	.28
BA	.270**	.128	-.002	.165*	.321**	.127	-.028	.287**	.093	.352**	.183*		.14	.13	.13
IN	.392**	.645**	.285**	.223*	.002	.617**	.297**	.374**	.544**	.319**	.372**	.097		.54	.52
RV	.296**	.615**	.199*	.341**	-.070	.564**	.264**	.377**	.532**	.223*	.368**	.081	.620**		.41
AR	.316**	.374**	.273**	.222*	.061	.407**	.394**	.306**	.390**	.322**	.236**	.050	.452**	.433**	

* : $p < .05$; ** : $p < .01$; NB : en dessous de la diagonale = valeurs absolues ; au-dessus de la diagonale = valeurs rapportées dans le manuel du WISC-IV

4.2 Mesures censées évaluer la Mémoire à court terme

Dans le modèle à 4 facteurs du WISC-IV, les épreuves censées mesurer l'*Indice de Mémoire de travail* (IMT) ont été regroupées sous le facteur *Mémoire à court terme* (Gsm) dans le modèle à 5 facteurs (CHC), à l'exception de l'épreuve *Arithmétique* (AR) dont nous n'avons pas tenu compte dans notre étude. Ainsi, seuls ont été conservés les sous-tests *Mémoire des chiffres* (MC) qui a été effectué lors de la première séance ainsi que *Séquences lettres-chiffres* (SLC) lors de la deuxième séance. Enfin, l'épreuve *Répétition de phrases* (RP) issue du NEPSY et évaluant également Gsm a été réalisée lors de la troisième séance.

L'analyse des corrélations nous montre que ces épreuves du WISC-IV censées évaluer les mêmes aptitudes cognitives sont corrélées très significativement entre elles avec une intensité relativement élevée. En effet, la corrélation entre MC et SLC est de .45 ($p < 0,1$). En se référant au manuel du WISC-IV, on constate que cette valeur est très proche ($r = .47$). De plus, avec l'épreuve RP du NEPSY, on observe également des corrélations très significatives : .47 ($p < .01$) entre SLC et RP et .46 ($p < .01$) entre RP et MC (cf. Tableau 5). Au regard de ces résultats, nous pouvons donc dire qu'il existe un lien entre les performances observées pour les trois séances et que ces épreuves évaluent bien Gsm.

Tableau 5 : Corrélations entre les épreuves du WISC-IV et du NEPSY supposées mesurer la Mémoire à court terme

Variables	1	2	3
1. Répétition de phrases (NEPSY)			
2. Mémoire des chiffres (WISC-IV)	.463**		.47
3. Séquence lettres-chiffres (WISC-IV)	.466**	.445**	

** : $p < .01$; NB : en dessous de la diagonale = valeurs absolues ; au-dessus de la diagonale = valeurs rapportées dans le manuel du WISC-IV

En ce qui concerne la stabilité des performances entre les trois séances, la variation moyenne absolue est de 2,11 ($SD = 1,81$) entre MC et SLC effectuées, respectivement, lors de la première et deuxième séance. De plus, en ce qui concerne la troisième séance, les variations moyennes absolues sont de 2,33 ($SD = 1,94$) entre MC et RP ainsi que de 2,52 ($SD = 1,97$) entre RP et SLC. Les valeurs étant toutes inférieures à 4 points, nous pouvons donc dire que les performances des enfants sont stables entre les séances.

Cependant, nous constatons que ce n'est pas le cas pour tous les enfants. En effet, 6,8% d'entre eux présentent des variations de performance supérieures à 4 points entre MC et SLC (cf. Annexe I). Par exemple, sur ces 6,8%, 2,9% des enfants ont présenté une différence de 5 points, 1% de 6 points, 1,9% de 8 points et 1% de 9 points. Ces pourcentages concordent avec les données de Grégoire où 12,8% des enfants de l'échantillon d'étalonnage présentaient une différence de performance de 5 points et plus entre ces deux épreuves. Quant aux sous-tests MC et RP, 16,2% des enfants ont une différence de performance allant de 5 à 7 points (cf. Annexe II). Plus précisément, 7,6% présentaient une différence de 5 points, 4,8% de 6 points et 3,8% de 9 points. Enfin, en ce qui concerne les épreuves SLC et RP, 16,3% présentent une différence de performance de 5 à 8 points (cf. Annexe III). En effet, 8,6% des enfants présentaient une différence de 5 points, 2,9% de 6 points, 1,9% de 7 points et 2,9% de 8 points. En conclusion, et étant donné que la proportion d'enfants présentant des différences de plus de 5 points est « négligeable », nous pouvons en conclure que les performances des enfants entre les séances restent stables.

4.3 Mesures censées évaluer la Vitesse de traitement

Dans le modèle à 5 facteurs, les trois sous-tests du WISC-IV regroupés sous le facteur *Vitesse de traitement* (Gs) sont les mêmes que pour l'*Indice de Vitesse de traitement* dans le

modèle à 4 facteurs. Le sous-test *Code* (CD) a été administré lors de la première séance et les sous-tests *Symboles* (SY) et *Barrage* (BA), lors de la deuxième séance. Enfin, l'épreuve *Cross-out* (Xout) issue du WJ-R et évaluant également Gs a été effectuée lors de la troisième séance.

L'analyse des corrélations nous montre que toutes les épreuves du WISC-IV évaluant Gs sont très significativement corrélées entre elles, signifiant l'existence d'une dépendance entre ces différentes épreuves évaluant Gs (cf. Tableau 6). En effet, l'épreuve *Code* (CD) est significativement corrélée aux épreuves SY ($r = .37, p < 0,1$) et BA ($r = .32, p < 0,1$). De plus, la corrélation est également significative entre les épreuves SY et BA ($r = .35, p < 0,1$) administrées lors de la deuxième séance. En ce qui concerne les données du manuel du WISC-IV, on peut observer que ces valeurs sont relativement proches de celles obtenues dans notre étude : la corrélation est de .33 entre les sous-tests CD et BA ; de .51 entre les épreuves CD et SY et de .33 entre SY et BA. Il y a donc une dépendance entre ces épreuves. Au niveau des corrélations entre l'épreuve Xout du WJ-R et les épreuves du WISC-IV mesurant Gs, nous pouvons déduire, avec un seuil de significativité de 5%, qu'il existe une corrélation entre l'épreuve Xout et les épreuves BA ($r = .49, p < .01$) ; SY ($r = .41, p < .01$) ; et CD ($r = .27, p < .01$) du WISC-IV. En conclusion, nous pouvons donc dire qu'il existe un lien entre les performances observées pour les trois séances et que ces épreuves évaluent bien Gs.

Tableau 6 : *Corrélations entre les épreuves du WISC-IV et du WJ-R supposées mesurer la Vitesse de traitement*

Variabiles	1	2	3	4
1. Cross-out (WJ-R)				
2. Code (WISC-IV)	.266**		.51	.33
3. Symboles (WISC-IV)	.407**	.367**		.33
4. Barrage (WISC-IV)	.493**	.321**	.352**	

** : $p < .01$; NB : en dessous de la diagonale = valeurs absolues ; au-dessus de la diagonale = valeurs rapportées dans le manuel du WISC-IV

En ce qui concerne l'analyse de la stabilité des performances intra-séance, on observe que les sous-tests SY et BA, administrées lors de deuxième séance, présentent une variation moyenne absolue des performances de 2,43 ($SD = 1,94$). Etant donné que cette différence est inférieure 4 points, nous pouvons donc dire que les performances des enfants restent stables. Toutefois, il est important de signaler que cela n'est valable que pour 84,6% des enfants. En effet, on remarque des variations allant de 5 à 9 points chez 15,4% d'entre eux (cf. Annexe

IV). Plus précisément, 6,7% des enfants présentaient une différence de 5 points, 4,8% de 6 points, 1,9% de 7 points, 1% de 8 points et 1% de 9 points. En conclusion, il ressort que la proportion d'enfants présentant des différences de plus de 5 points est « négligeable ». En conséquence, nous pouvons affirmer que les performances des enfants aux sous-tests SY et BA restent stables.

Au niveau des variations moyennes absolues des performances entre les trois séances, les analyses montrent une variation moyenne absolue de 3,04 ($SD = 2,15$) entre CD et Xout ; de 2,29 ($SD = 2,02$) entre CD et SY ; et de 2,40 ($SD = 1,96$) entre SY et Xout. Ainsi, étant donné que ces valeurs sont toutes inférieures à 4 points, on peut donc considérer que les performances sont restées stables. Toutefois, il est à signaler que certains enfants présentent des différences de performances supérieures à 4 points. En effet, on observe une fluctuation des performances de 5 à 10 points entre les sous-tests SY et CD pour 12,6% des participants (cf. Annexe V). Plus précisément, 4,8% des enfants présentent une différence de 5 points, 2,9% de 6 points, 1,9% de 7 points, 1% de 8 points, 1% de 9 points et 1% de 10 points. Ces pourcentages concordent avec les données de Grégoire où 10,3% des enfants de l'échantillon d'étalonnage présentaient une différence de performance de 5 points et plus entre ces deux épreuves.

En ce qui concerne les épreuves CD et Xout, on observe une fluctuation de performance de 4,20 à 9,20 points pour 30,1% des enfants qui se répartissent de la manière suivante : 11,5% des enfants présentent une différence de 4,2 à 4,8 points, 7,8% entre 5 et 5,6 points, 5,9% entre 6,2 et 6,8 points, 1% de 7,4 points, 1% de 8,4 points et 2,9% de 9,2 points (cf. Annexe VI).

Enfin, concernant les épreuves SY et Xout, on constate des différences de performance allant de 4,2 à 11,2 chez 16,6% des participants (cf. Annexe VII). Plus précisément, 5,8% présentent une différence entre 4,2 et 4,8 points, 5,8% entre 5 et 5,8 points, 2% de 6,2 points, 1% de 7 points, 1% de 9,2 points et 1% de 11,2 points.

Au regard de ces résultats, et étant donné que la proportion d'enfants présentant des différences de plus de 5 points est « négligeable », nous pouvons en conclure que les performances des enfants entre les séances restent stables.

4.4 Mesures censées évaluer le Traitement visuel

Dans le modèle à 4 facteurs, les épreuves censées mesurer l'*Indice de Raisonnement perceptif*, étaient les sous-tests *Cubes* (CU), *Identification de concepts* (IDC), *Matrices* (MA)

et *Complètement d'images* (CI). Dans le modèle à 5 facteurs, seules les épreuves CU et CI, ont été conservées sous le facteur *Traitement visuel* (Gv), les autres ayant été regroupées sous le facteur *Intelligence fluide* (Gf). Les sous-tests CU et CI ont été administrées respectivement à la première et deuxième séance. Enfin, l'épreuve *Triangles* (TG) issue du K-ABC et évaluant également Gv a été effectuée lors de la troisième séance.

En ce qui concerne les corrélations, les analyses nous montrent l'existence d'une dépendance entre chacun des sous-tests TG, CI et CU (cf. Tableau 7). En effet, nous remarquons que toutes les corrélations sont élevées et très significatives : .42 ($p < 0,1$) entre TG et CI ; .63 ($p < 0,1$) entre TG et CU; et .38 ($p < 0,1$) entre CI et CU. En se référant au manuel du WISC-IV on constate que la corrélation entre les épreuves CU et CI ($r = .38$) est similaire à celle observée dans notre étude ($r = .38, p < .01$). En conclusion, nous pouvons donc dire qu'il existe un lien entre les performances observées pour les trois séances et que ces épreuves évaluent bien Gv.

Tableau 7 : Corrélations entre les épreuves du WISC-IV et du K-ABC supposées mesurer le *Traitement visuel*

Variabiles	1	2	3
1. Triangles (K-ABC)			
2. Cubes (WISC-IV)	.627**		.38
3. Complètement d'images (WISC-IV)	.417**	.382**	

** : $p < .01$; NB : en dessous de la diagonale = valeurs absolues ; au-dessus de la diagonale = valeurs rapportées dans le manuel du WISC-IV

En ce qui concerne l'analyse de la stabilité des performances entre les trois séances, la variation moyenne absolue d'une épreuve à une autre est de 1,83 ($SD = 1,60$) entre CU et TG ; de 2,19 ($SD = 1,77$) entre CI et TG ; et de 2,81 ($SD = 2,02$) entre CI et CU. Au regard de ces résultats, nous pouvons donc dire que les performances sont restées stables, étant donné qu'elles sont inférieures à 4 points. Toutefois, il est important de signaler que cela n'était pas le cas pour 8,6% des enfants pour les épreuves CU et TG. En effet, certains enfants présentent une différence de performances allant de 5 à 6 points (cf. Annexe VIII). Plus précisément, 5,7% des enfants présentent une différence de 5 points et 2,9% une différence de 6 points. De plus, pour les épreuves CI et TG, 8,7% des enfants ont présenté une différence de performances allant de 5 à 10 points (cf. Annexe IX). En effet, 3,8% des enfants présentent une différence de 5 points, 1% de 6 points, 2,9% de 7 points et 1% de 10 points. Enfin, 19,2%

des enfants ont présentés une différence de performances entre 5 et 10 points entre les épreuves CI et CU (cf. Annexe X). Ainsi, 11,4% des enfants présentent une différence de 5 points, 1,9% de 6 points, 2,9% de 7 points et 1% pour respectivement 8, 9 et 10 points.

Au regard de ces résultats, et étant donné que la proportion d'enfants présentant des différences de plus de 5 points est « négligeable », nous pouvons en conclure que les performances des enfants entre les séances restent stables.

4.5 Mesures censées évaluer l'Intelligence fluide

Dans le modèle à 4 facteurs, les épreuves censées mesurer l'*Indice de Raisonnement perceptif*, étaient les sous-tests *Cubes* (CU), *Identification de concepts* (IDC), *Matrices* (MA) et *Complètement d'images* (CI). Dans le modèle à 5 facteurs, seules les épreuves IDC et MA, ont été conservées sous le facteur *Intelligence fluide* (Gf), les autres ayant été regroupées sous le facteur *Traitement visuel* (Gv). Les sous-tests IDC et MA ont été administrée respectivement à la première et deuxième séance. Enfin, l'épreuve *Formation de concepts* (FC) issue du WJ-R et évaluant également Gf a été effectuée lors de la troisième séance.

Au niveau des corrélations, et en prenant le seuil de significativité de 1%, on observe une corrélation entre l'épreuve MA et IDC de .26 ($p < .01$). En ce qui concerne l'épreuve *Formation de concepts* (FC), qui a été administrée lors de la troisième séance, on constate que les corrélations ne sont pas très élevées. En effet, sa corrélation est de .23 ($p < .05$) avec IDC et de .32 ($p < .01$) avec MA (cf. Tableau 8). En conclusion, la dépendance entre chacun des sous-tests est plus modérée mais indique, de manière plus nuancée, qu'elles évaluent bien Gf et qu'il existe un lien entre les performances observées pour les trois séances.

Tableau 8 : *Corrélations entre les épreuves du WISC-IV et du WJ-R supposées mesurer l'Intelligence fluide*

Variables	1	2	3
1. Formation de concepts (WJ-R)			
2. Identification de concepts (WISC-IV)	.230*		.37
3. Matrices (WISC-IV)	.316**	.264**	

* : $p < .05$; ** : $p < .01$; NB : en dessous de la diagonale = valeurs absolues ; au-dessus de la diagonale = valeurs rapportées dans le manuel du WISC-IV

En ce qui concerne la stabilité des performances entre les trois séances, les variations moyennes absolues des performances sont de 2,53 ($SD = 1,87$) entre MA et IDC ; de 4,50 ($SD = 2,75$) entre MA et FC ; et 4,35 ($SD = 2,77$) entre IDC et FC. Au regard de ces résultats, nous pouvons donc dire que les performances restent stables uniquement entre l'épreuve MA et IDC. Toutefois, 14,3% des enfants présentent une différence de performances allant de 5 à 8 points entre ces deux épreuves (cf. Annexe XI). En effet, 5,7% des enfants présentent une différence de performance de 5 points, 3,8% de 6 points, 3,8% de 7 points et 1% de 8 points. Ces pourcentages concordent avec les données de Grégoire où 16,8% des enfants de l'échantillon d'étalonnage présentaient une différence de performance de 5 points et plus entre ces deux épreuves. Ainsi, nous pouvons donc en conclure que les performances des enfants entre ces deux épreuves restent stables.

En revanche ce n'est pas le cas pour les épreuves MA et FC. En effet, on constate que 53% des enfants présentent une différence de 4,2 à 10,8 points entre ces épreuves, ce qui indique qu'il existe bien une différence de performances (cf. Annexe XII). En effet, 9,6% des enfants présentent une différence de performance entre 4,2 et 4,8 points, 12,4% entre 5 et 5,8 points, 9,6% entre 6 et 6,8 points, 10,6% entre 7 et 7,8 points, 2% entre 8 et 8,4 points, 4,9% entre 9,2 et 9,8 points et 3,9% entre 10,2 et 10,8 points.

Pour terminer, en ce qui concerne les épreuves IDC et FC, on constate également qu'il existe une fluctuation des performances car 50,2% des enfants présentent une différence entre 4,2 et 10,4 points (cf. Annexe XIII). Plus précisément, 10,6% des enfants présentent une différence de performance entre 4,2 et 4,8 points, 10,5% entre 5 et 5,8 points, 6,7% entre 6 et 6,8 points, 10,7% entre 7 et 7,8 points, 4,9% entre 8,6 et 8,8 points, 3,9% entre 9 et 9,6 points et 2,9% entre 10,2 et 10,4 points. En conclusion, ces pourcentages nous montrent qu'il existe bien une différence de performances très importante entre les séances avec l'épreuve FC du WJ-R, effectuée lors de la troisième séance, mais pas entre les épreuves IDC et MA du WISC-IV effectuées respectivement à la première et deuxième séance.

4.6 Mesures censées évaluer l'Intelligence cristallisée

Dans le modèle à 5 facteurs, les cinq sous-tests du WISC-IV regroupés sous le facteur *Intelligence cristallisée* (Gc) sont les mêmes que pour *l'Indice de Compréhension verbale* (ICV) dans le modèle à 4 facteurs. En d'autres termes, Gc est donc équivalent à l'ICV. Les sous-tests *Similitudes* (SI) et *Vocabulaire* (VO) ont été administrés lors de la première séance et les sous-tests *Compréhension* (CO), *Information* (IN) et *Raisonnement verbal* (RV), lors de

la deuxième séance. Enfin, l'épreuve *Compréhension de consignes* (CC) issue du NEPSY et évaluant également Gc a été effectuée lors de la troisième séance.

Tout d'abord, l'analyse des corrélations nous montrent que toutes les épreuves du WISC-IV évaluant Gc présentent des corrélations significatives avec une intensité élevée (cf. Tableau 9) : les épreuves SI et VO sont fortement corrélées entre elles ($r = .60, p < 0,1$) et le sous-test RV est fortement corrélé aux épreuves SI ($r = .62, p < .01$), VO ($r = .57, p < .01$) ; CO ($r = .53, p < .01$) et IN ($r = .62, p < .01$). De plus, l'épreuve SI est fortement corrélée aux épreuves CO ($r = .59, p < .01$) et IN ($r = .65, p < .01$). En ce qui concerne le sous-test VO, il est fortement corrélé avec les épreuves CO ($r = .58, p < .01$) et IN ($r = .62, p < .01$). Enfin, l'épreuve IN est fortement corrélée à l'épreuve CO ($r = .54, p < .01$). En se référant au manuel du WISC-IV, on constate que ces résultats concordent et sont même très proches. Ainsi, il existe bien une dépendance entre chacun de ces sous tests. Enfin, en ce qui concerne les corrélations avec l'épreuve CC du NEPSY, les valeurs de p sont toutes inférieures à .01 et l'intensité des corrélations sont plus modérées en comparaison des autres corrélations (cf. Tableau IX) : l'épreuve CC a une corrélation de .28 ($p < .01$) avec SI ; de .40 ($p < .01$) avec VO ; de .25 ($p < .01$) avec CO ; de .36 ($p < .01$) avec IN ; et de .36 ($p < .01$) avec RV. Au regard de ces résultats, nous pouvons conclure qu'il existe un lien entre les performances observées pour les trois séances et que ces épreuves évaluent bien Gc.

Tableau 9 : *Corrélations entre les épreuves du WISC-IV et du NEPSY supposées mesurer l'Intelligence cristallisée*

Variables	1	2	3	4	5	6
1. Compréhension de consignes (NEPSY)						
2. Similitude (WISC-IV)	.281**		.67	.54	.62	.55
3. Vocabulaire (WISC-IV)	.399**	.596**		.56	.65	.54
4. Compréhension (WISC-IV)	.251**	.587**	.583**		.48	.48
5. Information (WISC-IV)	.360**	.645**	.617**	.544**		.54
6. Raisonnement verbal (WISC-IV)	.361**	.621**	.570**	.534**	.622**	

** : $p < .01$; NB : en dessous de la diagonale = valeurs absolues ; au-dessus de la diagonale = valeurs rapportées dans le manuel du WISC-IV

En ce qui concerne la variation moyenne absolue des performances des enfants lors de la première séance entre les épreuves SI et VO, elle est de 2,04 ($SD = 1,75$). Etant donné que cette valeur est inférieure à 4 points, nous pouvons en conclure que les performances restent stables. Toutefois, on observe que cette stabilité n'est vérifiée que pour 91,4% des participants et que 8,6% des enfants présentent des fluctuations allant de 5 à 8 points (cf. Annexe XIV).

En effet, 3,8% des enfants présentent une différence de performance de 5 points, 1,9% de 6 points et 2,9% de 8 points. Ces pourcentages concordent avec les données de Grégoire où 6,1% des enfants de l'échantillon d'étalonnage présentaient une différence de performance de 5 points et plus entre ces deux épreuves. En conclusion, il ressort que la proportion d'enfants présentant des différences de plus de 5 points est « négligeable ». Par conséquent, nous pouvons affirmer que les performances pour la première séance restent stables.

Au niveau de la deuxième séance, les variations moyennes absolues sont de 3,06 ($SD = 2,17$) entre CO et RV ; de 2,31 ($SD = 1,93$) entre CO et IN et de 2,88 ($SD = 2,07$) entre RV et IN. Les performances restent donc stables étant donné que les valeurs sont toutes inférieures à 4 points. Néanmoins, on constate que ce n'est pas le cas pour tous les enfants. En effet, 21% d'entre eux présentent une différence de performance entre 8 et 10 points entre les épreuves CO et RV (cf. Annexe XV). Plus précisément, 7,6% des enfants présentent une différence de performance de 5 points, 5,7% de 6 points, 1% de 7 points, 5,7% de 8 points et 1% de 10 points. Concernant les épreuves CO et IN, 12,4% des enfants présentent une différence de performance entre 5 et 10 points (cf. Annexe XVI). En effet, 5,7% des enfants présentent une différence de performance de 5 points, 3,8% de 6 points, 1,9% de 7 points et 1% de 10 points. Enfin, pour les sous-tests RV et IN, on constate que 19% des enfants présentent une différence de performance entre 5 et 9 points (cf. Annexe XVII). En effet, 9,5% des enfants présente une différence de performance de 5 points, 3,8% de 6 points, 1,9% de 7 points, 1,9% de 8 points et 1,9% de 9 points. Étant donné que la proportion d'enfants présentant des différences de plus de 5 points est « négligeable », nous pouvons en conclure que les performances des enfants pour la deuxième séance restent stables.

Concernant les variations moyennes absolues des performances entre les trois séances, elles sont de 2,98 ($SD = 2,11$) entre VO et RV, de 2,65 ($SD = 1,95$) entre VO et CC et de 2,75 ($SD = 2,37$) entre CC et RV. Les performances sont donc restées stables entre les séances, étant donné que ces valeurs ne dépassent pas 4 points. Toutefois, on constate encore une fois que cette stabilité n'est valable que pour une majorité d'enfants. En effet, 24% des enfants présentent une différence de performances allant de 5 à 10 points entre les épreuves VO et RV (cf. Annexe XVIII). Plus précisément, 10,5% des enfants présentent une différence de performance de 5 points, 6,7% de 6 points, 4,8% de 7 points, 1% de 8 points et 1% de 10 points. Par ailleurs, 21,9% des enfants présentent une différence de performance allant de 5 à 8 points entre les épreuves VO et CC (cf. Annexe IX). En effet, 13,3% des enfants présentent une différence de performance de 5 points, 7,6% de 6 points et 1% de 8 points. Enfin, pour les épreuves CC et RV, les données nous montrent une différence de performance allant de 5 à 10

points chez 18,2% des enfants (cf. Annexe XX). En effet, 4,8% des enfants présentent une différence de performance de 5 points, 2,9% de 6 points, 3,8% de 7 points, 3,8% de 8 points, 1,9% de 9 points et 1% de 10 points.

En conclusion, nous pouvons dire que la proportion d'enfants présentant des différences de plus de 5 points est « négligeable » et que les performances des enfants entre les séances restent stables.

5. Discussion

Dans cette étude, nous nous sommes intéressés à la stabilité des performances des enfants aux épreuves censées évaluer les mêmes aptitudes cognitives lorsque l'on administre le WISC-IV sur plusieurs séances au lieu d'une. De manière générale, nos résultats ont confirmé nos deux hypothèses. En effet, les épreuves regroupées pour chaque facteur étaient corrélées entre elles de manière significative et avec des intensités relativement élevées indiquant ainsi qu'elles évaluent bien les mêmes aptitudes cognitives. Par ailleurs, la stabilité a été confirmée mais n'est valable que pour la majorité des enfants. En effet, pour une petite partie d'entre eux, des fluctuations ont été observées. Toutefois, cela concorde avec les données de la littérature. En effet, nous pouvions nous attendre à ce que cela soit le cas étant donné que selon Grégoire (2006), il n'est pas rare d'observer des différences de performances supérieures à 4 points chez certains enfants au sein d'un échantillon tout-venant et ce, même si leurs capacités sont homogènes.

Nos résultats mettent en évidence que la performance moyenne absolue des différents indices et sous-tests sont proches, respectivement, de 100 ($SD = 15$) et de 10 ($SD = 3$), ce qui indique que, globalement, les résultats correspondent à un échantillon « tout-venant ». De plus, en ce qui concerne les corrélations entre les indices, nos analyses ont indiqué qu'à la différence des corrélations entre l'*Indice de Vitesse de traitement* (IVT) et l'*Indice de Mémoire de travail* (IMT) ainsi qu'entre l'*Indice de Compréhension verbale* (ICV) et l'*Indice de Vitesse de traitement* (IVT), qui ne sont pas significatives et qui indiquent qu'ils ne mesurent pas les mêmes construits cognitifs, les corrélations entre tous les autres indices sont significatives à un seuil de significativité de 1%. Toutefois, en se référant aux valeurs mentionnées dans le manuel du WISC-IV, on constate que ces corrélations sont relativement proches, ce qui indique une concordance. Enfin, au niveau des corrélations entre les différents sous-tests censées évaluer les mêmes aptitudes cognitives, nous constatons qu'elles sont

significatives et relativement proches de celles du manuel du WISC-IV. Toutefois, il est à constater que la plupart des épreuves corrélaient également avec d'autres sous-tests. De ce fait, il semble donc, et en accord avec la littérature, que la réalisation de certaines épreuves soient aussi sous-tendues par d'autres aptitudes cognitives.

Tout d'abord, pour la *Mémoire à court terme* (Gsm), les épreuves concernées étaient *Mémoire des chiffres* (MC) et *Séquences lettres-chiffres* (SLC) du WISC-IV ainsi que *Répétition de phrases* (RP) du NEPSY. Nos analyses nous ont montré que ces épreuves évaluent bien Gsm, confirmant de cette façon notre seconde hypothèse, étant donné qu'elles sont corrélées très significativement entre elles et avec des intensités relativement élevée. Toutefois, il a été constaté que l'épreuve SLC présente également des corrélations significatives avec des épreuves regroupées sous Gc (ex : Information et Raisonnement verbal). C'est également le cas pour MC avec des épreuves de Gf et Gc (cf. Tableau 4). Au niveau de la stabilité des performances entre les trois séances, notre première hypothèse a également été confirmée étant donné que les variations moyennes absolues des performances entre les épreuves sont toutes inférieures à 4 points entre MC et SLC ; MC et RP ; ainsi que RP et SLC. Toutefois, nos analyses nous ont montré que certains enfants ont présenté des fluctuations supérieures à 4 points. C'est le cas entre les épreuves MC et SLC (6,8%), MC et RP (16,2%) ; ainsi que SLC et RP (16,3%). Il est à noter également que, pour ces pourcentages, la majorité de ces enfants présentaient les différences les moins élevées et que, globalement, plus la différence augmentait et plus le pourcentage d'individus concernés diminuait.

Une des explications qui peut être donnée à ces pourcentages d'enfants présentant des fluctuations peut être attribuée aux aptitudes impliquées pour chacun de ces sous-tests. En effet, nous avons vu qu'elles évaluent également d'autres facteurs généraux. De plus, au niveau des aptitudes spécifiques, MC et SLC mesurent la *Mémoire de travail* (MW) mais MC serait également un indicateur de l'*Empan mnésique* (MS). Ces différences peuvent également être observées avec RP. En effet, il serait aussi un indicateur de l'*Empan mnésique* (MS). Ainsi, il aurait en commun cette aptitude spécifique avec MC et aucune avec SLC.

Concernant la *Vitesse de traitement* (Gs), ce facteur regroupait les épreuves *Code* (CD), *Symboles* (SY) et *Barrage* (BA) du WISC-IV de même que *Cross-out* (Xout) du WJ-R. Nos analyses nous ont montré que toutes ces épreuves sont corrélées significativement entre elles. De ce fait, nous pouvons donc en conclure que notre seconde hypothèse est vérifiée étant donné que les sous-tests évaluent bien Gs. Toutefois, il est à constater que SY est aussi corrélé avec des épreuves regroupées sous Gv, Gc et Gf (cf. Tableau 4). Par ailleurs, tant au niveau

intra- qu'inter-séances, les performances des enfants sont restées stables, ce qui confirme notre première hypothèse. En effet, les variations moyennes absolues des performances entre les trois séances sont toutes inférieures à 4 points entre CD et SY ; CD et Xout ; ainsi que SY et Xout. Néanmoins, il est à signaler que certains enfants ont également présenté ici des différences de performances supérieures à 4 points. En effet, c'est le cas entre SY et CD (12,6%) ; CD et Xout (30,1%) ; ainsi que SY et Xout (16,6%). De plus, pour ces pourcentages, la majorité des enfants présentaient les différences les moins élevées et plus la différence augmentait et plus le pourcentage d'individus concernés diminuait.

Il est à noter que le pourcentage le plus important d'enfants qui présentaient des différences se situait entre les épreuves CD et Xout. Une des explications possible est que bien que ces deux épreuves évaluent la même aptitude générale (Gs), elles mesurent également des aptitudes générales et spécifiques différentes. En effet, en ce qui concerne l'épreuve Xout, elle présente des corrélations avec des sous-tests regroupés sous Gf, Gs ou encore Gv. Par ailleurs, concernant les aptitudes spécifiques, CD serait un indicateur du *Taux de réponse à un test* (R9) alors que Xout mesurerait la *Vitesse perceptive* (P). En revanche, bien que SY mesure aussi la *Vitesse perceptive* (P), elle a en commun avec CD l'aptitude spécifique *Taux de réponse à un test* (R9), ce qui pourrait expliquer le plus faible pourcentage d'enfants présentant des fluctuations entre ces deux épreuves. Ce serait également le cas entre les épreuves SY et Xout qui ont en commun P mais pas R9.

En ce qui concerne le *Traitement visuel* (Gv), il regroupait les sous-tests *Cubes* (CU) et *Complètement d'images* (CI) du WISC-IV ainsi que *Triangles* (TR) issue du K-ABC. Encore une fois, nos analyses ont montré qu'elles évaluaient bien Gv et ont confirmé notre seconde hypothèse. En effet, les corrélations étaient élevées et très significatives entre les épreuves CI et CU ; TG et CI ; ainsi que TG et CU. Toutefois, il est à noter que CU présente également des corrélations significatives avec des épreuves regroupées sous d'autres aptitudes générales (Gc, Gf et Gs), ce qui est également le cas pour CI (Gc et Gs). Au niveau de la stabilité des performances entre les trois séances, nos résultats ont également confirmé notre première hypothèse. En effet, les variations moyennes absolues étaient toutes inférieures à 4 points entre CI et CU ; CU et TG ; et CI et TG. Toutefois, nos résultats nous ont montré que ce n'est pas le cas pour une minorité d'enfants. Ainsi, c'est le cas entre les sous-tests CI et CU (19,2%) ; CU et TG (8,6%) ; et CI et TG (8,7%). Par ailleurs, il a été observé que, pour ces pourcentages, la majorité de ces enfants présentaient les différences les moins élevées et que, globalement, plus la différence augmentait et plus le pourcentage d'individus concernés diminuait.

En ce qui concerne les explications qui pourraient être données au sujet de ces pourcentages d'enfants présentant des différences de performances supérieures à 4 points, l'implication d'aptitudes générales et spécifiques différentes semble pertinente ici aussi. En effet, nous avons vu que les épreuves du WISC-IV étaient corrélées avec des épreuves regroupées sous d'autres facteurs que Gv. Par ailleurs, en ce qui concerne les aptitudes spécifiques, on constate que CU et TG mesurent exactement la même chose : les *Relations spatiales* (SR) et la *Visualisation* (VZ). Ainsi, il est possible que ces enfants aient activé des connaissances en mémoire à long terme et utilisé des stratégies de résolution de problèmes acquises précédemment. En d'autres termes, il y aurait eu un effet d'apprentissage entre ces deux sous-tests. Par ailleurs, en ce qui concerne CI, il n'aurait aucune aptitude spécifique en commun avec CU et avec TG. De plus, il mesurerait également l'Intelligence cristallisée car il serait un indicateur des *Information générales verbales* (K0). De ce fait, il est tout à fait envisageable que les différences de performances avec CI soient liées à ces aptitudes spécifiques.

D'autres interprétations peuvent également être envisagées. Par exemple, en ce qui concerne l'épreuve CU, les différences de performances avec les autres sous-tests pourraient être liées à la pression du temps limité. En effet, si un enfant est trop anxieux, il est susceptible de perdre ses moyens et ainsi fournir des performances plus faibles. De plus, une méticulosité excessive peut aussi ralentir l'enfant et faire baisser son score. Un autre exemple qui peut expliquer les différences entre les scores peut concerner la dépendance/indépendance à l'égard du champ, c'est-à-dire à la capacité à isoler un élément de son contexte et à le réutiliser dans un contexte différent. Ainsi, contrairement aux indépendants, certains enfants ont des difficultés à se décentrer de la forme globale d'une image, à déstructurer et à isoler les parties du champ.

Concernant, le facteur *Intelligence fluide* (Gf), les épreuves qui ont été administrées étaient *Matrices* (MA) et *Identification de concepts* (IDC) du WISC-IV ainsi que *Formation de concepts* (FC) du WJ-R. Les résultats de nos analyses ont montré que les corrélations étaient plutôt modérées entre MA et IDC ; IDC et FC ; et entre MA et FC. De ce fait, notre seconde hypothèse a été confirmée, car ces sous-tests évaluent bien Gf, mais est plus nuancée. Toutefois, si l'on se réfère au tableau de l'ensemble des corrélations pour le WISC-IV (CF. Tableau 4), on constate que MA est également corrélée avec des épreuves regroupées sous Gs, Gc, Gv et Gsm et que IDC est corrélé avec des épreuves classées sous Gc et Gsm. Au niveau de notre première hypothèse, elle n'a été que partiellement confirmée. Les variations moyennes absolues des performances étaient inférieures à 4 points entre les sous-tests MA et

IDC et seulement une faible proportion d'enfants ont présenté une différence de performances au-dessus de 4 points (14,3%). De plus, il est à noter que, pour ce pourcentage, la majorité de ces enfants présentaient les différences les moins élevées et que, globalement, plus la différence augmentait et plus le pourcentage d'individus concernés diminuait.

En revanche, entre MA et FC et entre IDC et FC, il existe bel et bien une différence de performance. D'ailleurs, en nous référant aux pourcentages d'enfants qui présentent des différences supérieures à 4 points, on constate que plus de la moitié des enfants ont présenté une fluctuation entre ces épreuves. Toutefois, on peut relativiser ces résultats du fait que le sous-test FC est issu du WJ-R. En effet, l'explication la plus probable tient du fait qu'il n'existe pas à ce jour de version française pour cette batterie de tests et qu'une traduction a donc été effectuée afin de pouvoir l'administrer. Ainsi, il est tout à fait possible que cela ait posé problème dans notre étude. D'ailleurs, le pourcentage d'enfants présentant des différences de performances supérieures à 4 points entre les sous-tests MA et IDC est largement inférieur.

En ce qui concerne les fluctuations entre MA et IDC, une des explications qui peut être donnée concerne les aptitudes générales et spécifiques qu'elles impliquent. En effet, nous avons vu que ces deux sous-tests corrélaient avec des aptitudes générales différentes. De plus, en ce qui concerne les aptitudes spécifiques, on constate que ces deux épreuves mesurent l'*Induction* (I) mais que MA serait également un indicateur du *Raisonnement séquentiel général* (RG). Ainsi, il est possible que les fluctuations des performances soient liées à ces aspects.

Enfin, pour le facteur *Intelligence cristallisée* (Gc), il regroupe les sous-tests *Similitudes* (SI), *Vocabulaire* (VO), *Compréhension* (CO), *Information* (IN) *Raisonnement verbal* (RV) du WISC-IV de même que l'épreuve *Compréhension de consignes* (CC) du NEPSY. En ce qui concerne les corrélations entre les épreuves du WISC-IV, nous avons constaté qu'elles sont toutes significatives et avec des intensités très élevées. En revanche, elles sont plus modérées avec le sous-test CC. Ainsi, tous ces sous-tests évaluent bien Gc, ce qui confirme notre première hypothèse. Toutefois, les épreuves RV et VO, que nous avons utilisé pour évaluer les différences inter-séances, corrélaient également avec des épreuves regroupées sous Gv, Gf, Gsm et Gs. En ce qui concerne la stabilité des performances, nos résultats ont montré qu'elle est présente tant au niveau intra- qu'inter-séance. De ce fait, notre première hypothèse a été confirmée. En effet, les variations moyennes absolues des performances entre les trois séances sont toutes inférieures à 4 points. Néanmoins, et encore une fois, on constate que cette stabilité n'est valable que pour une majorité d'enfants. En effet,

certaines d'entre eux ont présenté une différence de performances supérieure à 4 points entre VO et RV (24%) ; entre VO et CC (21,9%) ; et entre RV et CC (18,2%). De plus, pour ces pourcentages, la majorité de ces enfants présentaient les différences les moins élevées et, globalement, plus la différence augmentait et plus le pourcentage d'individus concernés diminuait.

Si l'on se réfère aux aptitudes générales et spécifiques impliquées par ces sous-tests, une explication à ces fluctuations peut être donnée. En effet, nous avons vu qu'elles impliquent également la mesure d'autres aptitudes générales. De plus, en ce qui concerne les aptitudes spécifiques, on constate que VO et RV ont la même en commun : les *connaissances lexicales* (VL). Ainsi, il est possible qu'un effet d'apprentissage ait pu contribuer à la fluctuation des performances. En revanche, ce n'est pas le cas avec CC. En effet, étant donné que cette épreuve mesure l'*Aptitude d'écoute* (LS), contrairement à VO et RV, les fluctuations pourraient être expliquées par le fait qu'il n'y a pas d'aptitudes spécifiques en commun entre ces sous-tests.

6. Conclusion

En définitive, procéder à un découpage des épreuves du WISC-IV, afin de les administrer sur plusieurs séances, semble ne pas poser de problèmes. Toutefois, il est important que les examinateurs prennent en considération que certains enfants peuvent tout de même présenter des fluctuations au niveau de leurs performances.

Pour terminer, nous ne saurions insister suffisamment sur certains aspects de l'évaluation. Ainsi, au regard de ce qui a été présenté dans cette étude, nous recommandons vivement que l'administration du WISC-IV soit effectuée par des personnes bien formées car, malheureusement, force est de constater que c'est loin d'être toujours le cas dans la pratique clinique de manière générale.

De plus, il ne faut pas oublier non plus que l'intelligence ne peut s'évaluer de manière directe et pure. Ainsi, bien que le WISC-IV fournisse des indications précieuses, il ne permet que l'observation d'une situation d'évaluation, c'est-à-dire les facettes de l'activité de l'intelligence, son expression durant l'évaluation. En d'autres termes, ce que l'on mesure n'est donc pas l'intelligence en elle-même mais l'efficacité avec laquelle l'enfant utilise ses potentialités intellectuelles. Par ailleurs, les tests psychométriques n'en demeurent pas moins

des outils qui nécessitent une interprétation et c'est donc au psychologue de donner du sens aux mesures.

Nous soulignons également que le WISC-IV ne permet pas une mesure de toutes les facettes de l'intelligence. En effet, d'autres aspects peuvent être considérés comme la mémoire à long terme, l'intelligence sociale, la créativité, l'interaction émotion-cognition, etc. Il faudrait donc chercher à les intégrer, afin que cette batterie de tests soit réellement considérée comme une échelle d'intelligence générale.

Enfin, lors du bilan psychologique, il est également nécessaire de considérer d'autres tests ou questionnaires ainsi que d'autres instruments d'évaluation tels que l'entretien, l'évaluation informelle ou encore les observations directes, avant de tirer certaines conclusions. Les interprétations doivent également être faites à la lumière de ce qui est observé durant la passation comme la réussite partielle aux épreuves, les stratégies de résolution de problèmes, les réactions, les verbalisations, les attitudes envers les tâches et l'examineur, etc. En effet, les informations obtenues en dehors de celles fournies par la batterie de tests apportent des éléments cliniques qui, au-delà des scores, permettent une analyse plus qualitative et une formulation des hypothèses plus fiable. Ainsi, le contexte de passation est considéré comme un moment central pour appréhender l'individu et ne se limite donc pas au cadre de la mesure des aptitudes de l'enfant.

Références bibliographiques :

- Alfonso, V.C., Flanagan, D.P., & Radwan, S. (2005). The impact of the Cattell-Horn-Carroll theory on test development and interpretation of cognitive and academic abilities. In Flanagan, D.P., & Harrison, P.L. (Eds.), *Contemporary intellectual assessment. Theories, tests, and issues* (pp. 185-202). New York: The Guilford Press.
- Canivez, G.L., & Watkins, M.W. (1998). Long-term stability of the Wechsler Intelligence Scale for Children-Third Edition. *Psychological Assessment, 10*, 285-291.
- Flanagan, D.P., & Kaufman, A.S. (2004). *Essentials of WISC-IV assessment*. Hoboken, New Jersey: John Wiley & Sons, Inc.
- Flanagan, D.P., & McGrew, K.S. (1998). Interpreting Intelligence Tests from Contemporary Gf-Gc Theory: Joint Confirmatory Factor Analysis of the WJ-R and KAIT in a Non-White Sample. *Journal of School Psychology, 36*(2), 151-182.
- Flanagan, D.P., McGrew, K.S., & Ortiz, S.O. (2000). *The Wechsler Intelligence Scales and Gf-Gc theory: A contemporary approach to interpretation*. Boston: Allyn and Bacon.
- Flanagan, D.P., Ortiz, S.O., & Alfonso, V.C. (2007). *Essentials of cross-battery assessment* (2nd Ed.). Hoboken, New Jersey: John Wiley & Sons. Inc.
- Floyd, R.G., Bergeron, R., McCormack, A.C., Anderson, J.L., & Hargrove-Owens, G.L. (2005). Are Cattell-Horn-Carroll Broad Ability Composite Scores Exchangeable Across Batteries? *School Psychology Review, 34*(3), 329-357.
- Ghisletta, P., Lecerf, T., & de Ribaupierre, A. (2003). Etude de la variabilité inter- et intra-individuelle dans des épreuves d'inhibition, de mémoire de travail et de vitesse de traitement. In vom Hofe, A., Charvin, H., Bernaud, J.L., & Guédon, D. (Eds). *Psychologie différentielle: recherches et réflexions*. (pp. 217-221). Rennes: Presses Universitaires de Rennes.
- Grégoire, J. (2006). *L'examen clinique de l'intelligence de l'enfant. Fondements et pratique du WISC-IV*. Sprimont: Mardaga.

- Horn, J.L., & Noll, J. (1997). Human cognitive capabilities: Gf-Gc theory. In Flanagan, D.P., Genshaft, J.L., & Harrison, P.L. (Eds.), *Contemporary intellectual assessment*. New York: Guilford Press.
- Huteau, M., & Lautrey, J. (1999). *Évaluer l'intelligence. Psychométrie cognitive*. Paris : PUF.
- Kaufman, A.S. (1994). *Intelligent testing with the WISC-III*. New York: John Wiley & Sons, Inc.
- Kaufman, A.S., & Kaufman, N.L. (2004). *Kaufman assessment battery for children* (2nd Ed.). Circle Pines, MN: AGS Publishing.
- Kaufman, A.S., Lichtenberger, E.O., Fletcher-Janzen, E., & Kaufman, N.L. (2005). *Essential of KABC II assessment*. Hoboken, New Jersey: John Wiley & Sons. Inc.
- Keith, T.Z., Fine, J.G., Taub, G.E., Reynolds, M.R., & Kranzler, J.H. (2006). Higher Order, Multisample, Confirmatory Factor Analysis of the Wechsler Intelligence Scale for Children - Fourth Edition: What Does It Measure? *School Psychology Review*, 35, 108–127.
- Lecerf, T., Reverte, I., Coleaux, L., Maillard, F., Favez, N., & Rossier, J. (2011). Indice d'aptitude général et indice de compétence cognitive pour le WISC-IV : normes empiriques vs. normes statistiques. *Revue Européenne de Psychologie appliquée / European Journal of Applied Psychology*, 61, 115-122.
- Lecerf, T., Rossier, J., Favez, N., Reverte, I., & Coleaux, L. (2010). Comparing the usual four- and an alternative six-factor structure of the French WISC-IV using Confirmatory Factor Analyses. *Swiss Journal of Psychology*, 69, 221-232.
- McGrew, K.S. (1997). Analysis of the major intelligence batteries according to a proposed comprehensive Gf-Gc framework. In Flanagan, D.P., Genshaft, J.L., & Harrison P.L. (Eds.). *Contemporary intellectual assessment. Theories, tests, and issues* (pp. 151-180). New York: The Guilford Press.

- McGrew, K.S. (2005). The Cattell-Horn-Carroll (CHC) theory of cognitive abilities: Past, present and future. In Flanagan, D.P., & Harrison P.L. (2nd Eds.), *Contemporary intellectual assessment. Theories, tests, and issues* (pp. 136-202). New York: The Guilford Press.
- Reuchlin, M. (2001). *La Psychologie différentielle*. Paris: PUF.
- Salthouse, T.A. (2007). Implications of Within-Person Variability in Cognitive and Neuropsychological Functioning for the Interpretation of Change. *Neuropsychology*, 21(4), 401-411.
- Schrank, F.A., McGrew, K.S., & Woodcock, R.W. (2001). *Technical Abstract* (Woodcock-Johnson III Assessment Service Bulletin No. 2). Itasca, IL: Riverside Publishing.
- Spearman, C. (1904). General intelligence objectively determined and measured. *American Journal of Psychology*, 15(2), 201–293.
- Wechsler, D. (2005a). *Manuel de l’Echelle d’Intelligence de Wechsler pour Enfants-4e édition* (Manual for the Wechsler Intelligence Scale for Children - Fourth Edition). Paris : Editions du Centre de Psychologie Appliquée.

Annexes :

Annexe I : Différences de performances entre les épreuves Mémoire des chiffres (MC) et Séquences lettres-chiffres (SLC)

Différences	Fréquences	Pourcentages	Pourcentages cumulés
0	21	20.0	20.0
1	24	22.9	42.9
2	20	19.0	61.9
3	20	19.0	81.0
4	13	12.4	93.3
5	3	2.9	96.2
6	1	1.0	97.1
8	2	1.9	99.0
9	1	1.0	100
Total	105	100	

Annexe II : Différences de performances entre les épreuves Mémoire des chiffres et Répétition de phrases

Différences	Fréquences	Pourcentages	Pourcentages cumulés
0	19	18.1	18.1
1	26	24.8	42.9
2	16	15.2	58.1
3	19	18.1	76.2
4	8	7.6	83.8
5	8	7.6	91.4
6	5	4.8	96.2
7	4	3.8	100
Total	105	100	

Annexe III : Différences de performances entre les épreuves Séquences lettres-chiffres et Répétition de phrases

Différences	Fréquences	Pourcentages	Pourcentages cumulés
0	14	13.3	13.3
1	24	22.9	36.2
2	22	21.0	57.1
3	16	15.2	72.4
4	12	11.4	83.8
5	9	8.6	92.4
6	3	2.9	95.2
7	2	1.9	97.1
8	3	2.9	100
Total	105	100	

Annexe IV : Différences de performances entre les épreuves Symboles et Barrage

Différences	Fréquences	Pourcentages	Pourcentages cumulés
0	17	16.2	16.2
1	18	17.1	33.3
2	29	27.6	61.0
3	17	16.2	77.1
4	8	7.6	84.8
5	7	6.7	91.4
6	5	4.8	96.2
7	2	1.9	98.1
8	1	1.0	99.0
9	1	1.0	100
Total	105	100	

Annexe V : Différences de performances entre les épreuves Code et Symboles

Différences	Fréquences	Pourcentages	Pourcentages cumulés
0	18	17.1	17.1
1	24	22.9	40.0
2	24	22.9	62.9
3	19	18.1	81.0
4	7	6.7	87.6
5	5	4.8	92.4
6	3	2.9	95.2
7	2	1.9	97.1
8	1	1.0	98.1
9	1	1.0	99.0
10	1	1.0	100
Total	105	100	

Annexe VI : Différences de performances entre les épreuves Code et Cross-out

Différences	Fréquences	Pourcentages	Pourcentages cumulés
0.0	1	1.0	1.0
0.2	6	5.7	6.7
0.4	4	3.8	10.5
0.6	3	2.9	13.3
0.8	3	2.9	16.2
1.0	4	3.8	20.0
1.2	3	2.9	22.9
1.4	4	3.8	26.7
1.6	3	2.9	29.5
1.8	4	3.8	33.3
2.0	1	1.0	34.3
2.2	7	6.7	41.0
2.4	7	6.7	47.6
2.6	5	4.8	52.4
2.8	6	5.7	58.1
3.0	1	1.0	59.0
3.2	2	1.9	61.0
3.4	2	2.0	62.9
3.6	4	3.8	66.7
3.8	2	1.9	68.6
4.0	2	1.9	70.5
4.2	6	5.7	76.2
4.4	3	2.9	79.0
4.6	1	1.0	80.0
4.8	2	1.9	81.9
5.0	3	2.9	84.8
5.2	3	2.9	87.6
5.4	1	1.0	88.6
5.6	1	1.0	89.5
6.2	1	1.0	90.5
6.4	1	1.0	91.4
6.6	1	1.0	92.4
6.8	3	2.9	95.2
7.4	1	1.0	96.2
8.4	1	1.0	97.1
9.2	3	2.9	100
Total	105	100	

Annexe VII : Différences de performances entre les épreuves Symboles et Cross-out

Différences	Fréquences	Pourcentages	Pourcentages cumulés
0.0	1	1.0	1.0
0.2	6	5.8	6.7
0.4	4	3.8	10.5
0.6	4	3.8	14.3
0.8	10	9.5	23.8
1.0	3	2.9	26.7
1.2	7	6.7	33.3
1.4	5	4.8	38.1
1.6	11	10.5	48.6
1.8	7	6.7	55.2
2.0	2	1.9	57.1
2.2	2	2.0	59.0
2.4	2	1.9	61.0
2.6	5	4.8	65.7
2.8	1	1.0	66.7
3.0	2	1.9	68.6
3.2	5	4.8	73.3
3.4	4	3.8	77.1
3.6	2	2.0	79.0
3.8	4	3.8	82.9
4.0	1	1.0	83.8
4.2	1	1.0	84.8
4.4	3	2.9	87.6
4.8	2	1.9	89.5
5.0	2	1.9	91.4
5.4	1	1.0	92.4
5.6	1	1.0	93.3
5.8	2	1.9	95.2
6.2	2	2.0	97.1
7.0	1	1.0	98.1
9.2	1	1.0	99.0
11.2	1	1.0	100
Total	105	100	

Annexe VIII : Différences de performances entre les épreuves Cubes et Triangles

Différences	Fréquences	Pourcentages	Pourcentages cumulés
0	20	19.0	19.0
1	38	36.2	55.2
2	17	16.2	71.4
3	12	11.4	82.9
4	9	8.6	91.4
5	6	5.7	97.1
6	3	2.9	100
Total	105	100	

Annexe IX : Différences de performances entre les épreuves Complètement d'images et Triangles

Différences	Fréquences	Pourcentages	Pourcentages cumulés
0	15	14.3	14.3
1	24	22.9	37.1
2	32	30.5	67.6
3	15	14.3	81.9
4	10	9.5	91.4
5	4	3.8	95.2
6	1	1.0	96.2
7	3	2.9	99.0
10	1	1.0	100
Total	105	100	

Annexe X : Différences de performances entre les épreuves Complètement d'images et Cubes

Différences	Fréquences	Pourcentages	Pourcentages cumulés
0	10	9.5	9.5
1	22	21.0	30.5
2	16	15.2	45.7
3	26	24.8	70.5
4	11	10.5	81.0
5	12	11.4	92.4
6	2	1.9	94.3
7	3	2.9	97.1
8	1	1.0	98.1
9	1	1.0	99.0
10	1	1.0	100
Total	105	100	

Annexe XI : Différences de performances entre les épreuves Matrices et Identification de concepts

Différences	Fréquences	Pourcentages	Pourcentages cumulés
0	9	8.6	8.6
1	31	29.5	38.1
2	19	18.1	56.2
3	17	16.2	72.4
4	14	13.3	85.7
5	6	5.7	91.4
6	4	3.8	95.2
7	4	3.8	99.0
8	1	1.0	100
Total	105	100	

Annexe XII : Différences de performances entre les épreuves Matrices et Formation de concepts

Différences	Fréquences	Pourcentage	Pourcentage cumulé
0.0	1	1.0	1.0
0.2	3	2.9	3.8
0.4	2	1.9	5.7
0.8	2	1.9	7.6
1.0	5	4.8	12.4
1.2	1	1.0	13.3
1.4	4	3.8	17.1
1.6	1	1.0	18.1
1.8	3	2.9	21.0
2.0	1	1.0	21.9
2.2	3	2.9	24.8
2.6	3	2.9	27.6
2.8	4	3.8	31.4
3.0	5	4.8	36.2
3.2	3	2.9	39.0
3.4	1	1.0	40.0
3.6	1	1.0	41.0
3.8	3	2.9	43.8
4.0	3	2.9	46.7
4.2	2	1.9	48.6
4.4	1	1.0	49.5
4.6	5	4.8	54.3
4.8	2	1.9	56.2
5.0	2	1.9	58.1
5.2	2	1.9	60.0
5.4	1	1.0	61.0
5.6	4	3.8	64.8
5.8	4	3.8	68.6
6.0	2	1.9	70.5
6.2	2	1.9	72.4
6.4	1	1.0	73.3
6.6	3	2.9	76.2
6.8	2	1.9	78.1
7.0	3	2.9	81.0
7.2	2	1.9	82.9
7.4	1	1.0	83.8
7.6	2	1.9	85.7
7.8	3	2.9	88.6
8.0	1	1.0	89.5
8.4	1	1.0	90.5
9.2	2	1.9	93.3
9.4	1	1.0	94.3
9.6	1	1.0	95.2
9.8	1	1.0	96.2
10.2	3	2.9	99.0
10.8	1	1.0	100
Total	105	100	

Annexe XIII : Différences de performances entre les épreuves Identification de concepts et Formation de concepts

Différences	Fréquences	Pourcentage	Pourcentage cumulé
0.0	3	2.9	2.9
0.2	3	2.9	5.7
0.4	2	1.9	7.6
0.8	2	1.9	9.5
1.0	2	1.9	11.4
1.2	3	2.9	14.3
1.4	4	3.8	18.1
1.6	1	1.0	19.0
1.8	3	2.9	21.9
2.0	4	3.8	25.7
2.2	6	5.7	31.4
2.4	1	1.0	32.4
2.6	2	1.9	34.3
2.8	2	1.9	36.2
3.0	3	2.9	39.0
3.2	2	1.9	41.0
3.4	2	1.9	42.9
3.6	1	1.0	43.8
3.8	3	2.9	46.7
4.0	3	2.9	49.5
4.2	1	1.0	50.5
4.4	1	1.0	51.4
4.6	4	3.8	55.2
4.8	5	4.8	60.0
5.0	2	1.9	61.9
5.2	2	1.9	62.9
5.6	5	4.8	67.6
5.8	2	1.9	70.5
6.0	2	1.9	72.4
6.2	2	1.9	74.3
6.8	3	2.9	77.1
7.0	2	1.9	80.0
7.2	3	2.9	82.9
7.4	1	1.0	83.8
7.6	4	3.9	87.6
7.8	1	1.0	88.6
8.6	2	1.9	90.5
8.8	3	2.9	93.3
9.0	2	1.9	95.2
9.4	1	1.0	96.2
9.6	1	1.0	97.1
10.2	2	1.9	99.0
10.4	1	1.0	100
Total	52	100	

Annexe XIV : Différences de performances entre les épreuves Similitudes et Vocabulaire

Différences	Fréquences	Pourcentages	Pourcentages cumulés
0	16	15.2	15.2
1	33	31.4	46.7
2	24	22.9	69.5
3	15	14.3	83.8
4	8	7.6	91.4
5	4	3.8	95.2
6	2	1.9	97.1
8	3	2.9	100
Total	105	100	

Annexe XV : Différences de performances entre les épreuves Compréhension et Raisonnement verbal

Différences	Fréquences	Pourcentages	Pourcentages cumulés
0	8	7.6	7.6
1	20	19.0	26.7
2	22	21.0	47.6
3	16	15.2	62.9
4	17	16.2	79.0
5	8	7.6	86.7
6	6	5.7	92.4
7	1	1.0	93.3
8	6	5.7	99.0
10	1	1.0	100
Total	105	100	

Annexe XVI : Différences de performances entre les épreuves Compréhension et Information

Différences	Fréquences	Pourcentages	Pourcentages cumulés
0	18	17.1	17.1
1	26	24.8	41.9
2	17	16.2	58.1
3	19	18.1	76.2
4	12	11.4	87.6
5	6	5.7	93.3
6	4	3.8	97.1
7	2	1.9	99.0
10	1	1.0	100
Total	105	100	

Annexe XVII : Différences de performances entre les épreuves Information et Raisonnement verbal

Différences	Fréquences	Pourcentages	Pourcentages cumulés
0	13	12.4	12.4
1	16	15.2	27.6
2	21	20.0	47.6
3	18	17.1	64.8
4	17	16.2	81.0
5	10	9.5	90.5
6	4	3.8	94.3
7	2	1.9	96.2
8	2	1.9	98.1
9	2	1.9	100
Total	105	100	

Annexe XVIII : Différences de performances entre les épreuves Vocabulaire et Raisonnement verbal

Différences	Fréquences	Pourcentages	Pourcentages cumulés
0	9	8.6	8.6
1	23	21.9	30.5
2	18	17.1	47.6
3	16	15.2	62.9
4	14	13.3	76.2
5	11	10.5	86.7
6	7	6.7	93.3
7	5	4.8	98.1
8	1	1.0	99.0
10	1	1.0	100
Total	105	100	

Annexe XIX : Différences de performances entre les épreuves Vocabulaire et Compréhension de consignes

Différences	Fréquences	Pourcentages	Pourcentages cumulés
0	15	14.3	14.3
1	23	21.9	36.2
2	16	15.2	51.4
3	15	14.3	65.7
4	13	12.4	78.1
5	14	13.3	91.4
6	8	7.6	99.0
8	1	1.0	100
Total	105	100	

Annexe XX : Différences de performances entre les épreuves Compréhension de consignes et Raisonement verbal

Différences	Fréquences	Pourcentages	Pourcentages cumulés
0	14	13.3	13.3
1	25	23.8	37.1
2	21	20.0	57.1
3	13	12.4	69.5
4	13	12.4	81.9
5	5	4.8	86.7
6	3	2.9	89.5
7	4	3.8	93.3
8	4	3.8	97.1
9	2	1.9	99.0
10	1	1.0	100
Total	105	100	