

COMPARAISONS DES PROFILS D'INTELLIGENCE VERBALE ET
NON-VERBALE AU WISC-III ET AU STANFORD-BINET IV
CHEZ UN ÉCHANTILLON D'ENFANTS FRANCOPHONES
DU NOUVEAU-BRUNSWICK

THÈSE PRÉSENTÉE À LA FACULTÉ DES SCIENCES SOCIALES
EN VUE DE L'OBTENTION
DE LA MAÎTRISE ÈS ARTS EN PSYCHOLOGIE (M.A.Ps.)

PAR

ANICK MCGRATH, B.Ps.

UNIVERSITÉ DE MONCTON

JUIN 1997



**National Library
of Canada**

**Acquisitions and
Bibliographic Services**

**395 Wellington Street
Ottawa ON K1A 0N4
Canada**

**Bibliothèque nationale
du Canada**

**Acquisitions et
services bibliographiques**

**395, rue Wellington
Ottawa ON K1A 0N4
Canada**

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-23731-1

REMERCIEMENTS

La présente étude a été réalisée grâce à la participation de plusieurs personnes. Je dois d'abord remercier mon directeur de thèse, M. Pierre Cormier, Ph.D., qui m'a étroitement guidée et appuyée durant l'acheminement de cette thèse. J'en profite aussi pour remercier les autres membres de mon comité de thèse, M. Gaëtan Losier, Ph.D., de l'Université de Moncton et M. Serge Larivée, Ph.D., de l'Université de Montréal, pour leurs conseils et leurs encouragements. Un remerciement spécial va de plus à Nada Savoie, étudiante au Baccalauréat en psychologie à l'Université de Moncton, sans qui la cueillette des données aurait été plus ardue.

Mille mercis aux directeurs, enseignant-e-s et secrétaires des écoles suivantes: École Camille-Vautour de St-Antoine; École Carrefour Beausoleil de Miramichi; École Champlain de Moncton; École Mont-Carmel de Ste-Marie; École Notre-Dame de Notre-Dame; École Régionale de Baie Ste-Anne; École St-Paul de St-Paul; École W.F.Boisvert de Rogersville.

Aux 100 enfants de l'étude et à leurs parents, des remerciements les plus sincères. Leur motivation et leur enthousiasme ont permis de recueillir des données très importantes dans le domaine de l'évaluation cognitive.

Finalement, à tous les membres de ma famille et à mon époux, Larry, un remerciement très spécial. Leur amour, leur compréhension et leur appui m'ont permis d'atteindre plusieurs de mes objectifs de vie, dont ce travail.

CURRICULUM STUDIORUM

Anick (Turner) McGrath est née à Montréal, Québec, le 26 août 1971. Elle termina ses études secondaires à l'École Carrefour Beausoleil de Miramichi, Nouveau-Brunswick, où elle y reçut son diplôme de douzième année en juin 1989. En mai 1993, elle obtint son diplôme de Baccalauréat en psychologie (B.Ps.) de l'Université de Moncton.

SOMMAIRE

Dans cette étude, les profils d'intelligence verbale et non-verbale d'enfants francophones du Nouveau-Brunswick sont examinés dans deux tests d'intelligence. L'échantillon de l'étude regroupe 96 élèves (42 garçons et 54 filles) de classes régulières de la deuxième année (n = 27), de la quatrième année (n = 39) et de la sixième année (n = 30). Les administrations d'adaptation française du Wechsler Intelligence Scale for Children-Third Edition (WISC-III; Wechsler, 1991) et du Stanford-Binet Intelligence Scale: Fourth Edition (Stanford-Binet IV; Thorndike, Hagen, & Sattler, 1986a, 1986b) sont contrebalancées. L'analyse des données indique principalement que la performance des enfants de cet échantillon est significativement meilleure à l'échelle non-verbale par rapport à l'échelle verbale, tant au WISC-III, $F(1, 90) = 150,87, p < 0,001$, qu'au Stanford-Binet IV, $F(1, 90) = 208,33, p < 0,001$. Les écarts moyens entre ces échelles sont d'environ 16 points au WISC-III et 18 points au Stanford-Binet IV. L'examen des profils favorisant l'échelle d'intelligence non-verbale révèle une différence de 10 points et plus entre les cotes verbales et non-verbales chez 74% des enfants de l'échantillon au WISC-III et chez 78% d'entre eux au Stanford-Binet IV. Étant donné que la fréquence des écarts entre les échelles verbales et non-verbales est élevée et que la direction de ces écarts est surtout en faveur de la cote d'intelligence non-verbale, une interprétation quantitative des résultats

globaux au WISC-III et au Stanford-Binet IV s'avère inadéquate auprès des enfants de cette population. Les conclusions présentent entre autres des suggestions aux psychologues scolaires francophones de la province dans le but d'adapter leurs évaluations à la lumière des présents résultats.

TABLE DES MATIÈRES

REMERCIEMENTS.....	ii
CURRICULUM STUDIORUM.....	iii
SOMMAIRE.....	iv
TABLE DES MATIÈRES.....	vi
LISTE DES TABLEAUX.....	ix
LISTE DES FIGURES.....	x
INTRODUCTION.....	1
CHAPITRE I - INTELLIGENCE ET ÉVALUATION COGNITIVE.....	4
Concept d'intelligence.....	4
Définitions de l'intelligence.....	4
Approche psychométrique.....	6
Évaluation cognitive.....	8
WISC-III.....	10
Avantages et désavantages du WISC-III.....	12
Études comparatives avec le WISC-R.....	14
Stanford-Binet IV.....	18
Avantages et désavantages du Stanford-Binet IV.....	20
Études comparatives avec le Stanford-Binet L-M.....	23
Études comparatives entre les échelles d'intelligence de Wechsler et Stanford-Binet.....	27
Comparaisons entre le WISC-R et le Stanford-Binet IV.....	27
Comparaisons entre le WISC-III et le Stanford-Binet IV.....	32
Études des profils d'intelligence verbale et non-verbale aux échelles de Wechsler.....	36
Profils des populations minoritaires.....	39
Synthèse.....	46

Objectifs et hypothèses.....	49
CHAPITRE II - DESCRIPTION DE L'ÉTUDE.....	53
Méthode.....	53
Échantillon.....	53
Instruments de mesure.....	53
WISC-III.....	53
Stanford-Binet IV.....	54
Procédure.....	55
Analyses statistiques.....	56
CHAPITRE III - ANALYSES DES DONNÉES.....	59
Résultats.....	59
Statistiques descriptives.....	59
Analyse de l'échantillon de l'étude.....	66
Vérification des hypothèses.....	71
Corrélations inter-tests.....	71
Comparaisons des cotes de QI.....	73
Profils de QI verbal et de QI non-verbal.....	78
CHAPITRE IV - DISCUSSION ET CONCLUSIONS.....	82
Discussion.....	82
Comparaisons des tests.....	83
Profils de QI.....	85
Application des règles de la norme américaine.....	89
Conclusions.....	90
RÉFÉRENCES.....	97
ANNEXE A - Lettre de consentement parental.....	107

**ANNEXE B - Rapport global remis aux Districts scolaires
et aux écoles.....111**

ANNEXE C - Rapport global remis aux parents.....118

**ANNEXE D - Formule pour le rééchelonnement des QI du
Stanford-Binet IV.....122**

LISTE DES TABLEAUX

Tableau 1	Répartition de l'échantillon.....	61
Tableau 2	Statistiques descriptives du WISC-III.....	64
Tableau 3	Statistiques descriptives du Stanford-Binet IV.....	65
Tableau 4	Moyennes d'intelligence globale au WISC-III et au Stanford-Binet IV selon le sexe et l'année scolaire des sujets.....	68
Tableau 5	Moyennes d'intelligence globale au WISC-III et au Stanford-Binet IV selon le sexe et l'année scolaire des sujets du District scolaire no. 11.....	69
Tableau 6	Corrélations entre cotes de QI du WISC-III et du Stanford-Binet IV.....	72
Tableau 7	Moyennes d'intelligence globale, verbale et non-verbale au WISC-III et au Stanford-Binet IV.....	74
Tableau 8	Moyennes d'intelligence globale, verbale et non-verbale au WISC-III et au Stanford-Binet IV selon l'année scolaire des sujets.....	77
Tableau 9	Direction, fréquence et pourcentage cumulatif des écarts QIV-QIP au WISC-III.....	79
Tableau 10	Direction, fréquence et pourcentage cumulatif des écarts RV-RA/V au Stanford-Binet IV...	80

LISTE DES FIGURES

Figure 1	Distribution des cotes d'intelligence globale au WISC-III.....	62
Figure 2	Distribution des cotes d'intelligence globale au Stanford-Binet IV.....	63

INTRODUCTION

La présente étude s'intéresse à l'utilisation de tests d'intelligence américains auprès d'individus non représentés à l'intérieur du groupe normatif de ces instruments. Les résultats obtenus par un échantillon d'enfants francophones du Nouveau-Brunswick à des adaptations françaises du Wechsler Intelligence Scale for Children-Third Edition (WISC-III; Wechsler, 1991) et du Stanford-Binet Intelligence Scale: Fourth Edition (Stanford-Binet IV; Thorndike, Hagen, & Sattler, 1986a, 1986b) sont étudiés et comparés. L'objet principal de l'étude porte sur la comparaison des profils d'intelligence verbale et d'intelligence non-verbale des enfants de l'étude avec ceux de la population normative américaine.

La littérature scientifique américaine s'est intéressée à l'étude des profils d'intelligence produits par diverses populations d'enfants aux échelles de Wechsler. Entre autres, des études ont été menées auprès de populations américaines de différentes ethnies afin de pouvoir vérifier si les normes américaines des échelles de Wechsler rendent justice aux enfants de ces groupes minoritaires (Christiansen & Livermore, 1970; Clarizio, 1982; Dean, 1979, 1980; Figueroa, 1989; McCullough, Walker, & Diessner, 1985; McShane & Plas, 1984; Murphy, 1990; Swanson & Deblassie, 1979; Swerdlik, 1978; Tanner-Halverson, Burden, & Sabers, 1993; Taylor & Richards, 1991; Taylor, Ziegler, & Partenio, 1984). En général, ces études ont rapporté que, chez les

enfants qui diffèrent de la norme américaine, la performance obtenue au WISC-R ou au WISC-III ne concorde pas avec les règles de la norme établie dans les manuels de ces tests. Entre autres, les cotes d'intelligence verbale et non-verbale n'ont pas une prépondérance équivalente auprès de certains de ces groupes d'enfants, comme chez les Hispaniques et les Amérindiens. La tendance générale pour ces populations d'enfants est d'obtenir des cotes d'intelligence non-verbale qui sont significativement supérieures aux cotes d'intelligence verbale.

Au Nouveau-Brunswick, la population francophone représente une population qui est, elle aussi, différente de la norme américaine, tant à des niveaux social, linguistique, économique ou autre. Nous pouvons supposer que les profils d'intelligence des jeunes de cette population sont eux aussi différents de ceux du groupe normatif américain. Toutefois, aucune donnée empirique n'existe à ce sujet et les psychologues du Nouveau-Brunswick doivent présumer que le guide d'interprétation qui accompagne le WISC-III, s'utilise tel quel auprès des enfants de ce groupe. La présente étude va donc tenter de combler cette lacune en fournissant des données empiriques à ce sujet.

En première partie du chapitre 1, le concept d'intelligence est abordé avec une présentation de quelques définitions rattachées à cette notion, suivie d'une discussion de l'intelligence selon l'approche

psychométrique. En deuxième partie de ce chapitre, les deux mesures d'intelligence, soit le WISC-III et le Stanford-Binet IV, sont d'abord présentées et ensuite révisées à l'aide d'études portant sur la validité de ces deux tests. Cette section est suivie d'une revue des profils d'habiletés cognitives et des écarts entre les mesures verbale et non-verbale des échelles de Wechsler auprès de populations minoritaires diverses. C'est à la suite de cette revue de la littérature que les objectifs et les hypothèses d'étude sont formulés.

CHAPITRE PREMIER
INTELLIGENCE ET ÉVALUATION COGNITIVE

Concept d'intelligence

Définitions de l'intelligence

Définir l'intelligence est une tâche qui a suscité plusieurs débats scientifiques. Il n'y a pas de définition universellement acceptée de ce concept. D'ailleurs, il existe autant de définitions de l'intelligence qu'il y a d'experts l'ayant étudiée (Neisser et al., 1996; Sternberg, Conway, Ketron, & Bernstein, 1981; Wechsler, 1975; Weinberg, 1989). La divergence d'opinions à propos de la définition de ce concept ne suggère cependant pas son inexistence et ne représente pas un manque de connaissances à son sujet. Au contraire, les controverses appuient la pertinence de ce concept et suggèrent qu'il s'agit d'un concept complexe et multidimensionnel (Wechsler, 1975).

À des niveaux implicites et informels, l'intelligence est reconnue comme une forme de compétence mentale (Strommen, McKinney, & Fitzgerald, 1983). Le concept d'intelligence est reconnu par diverses cultures (Weinberg, 1989). Il est souvent associé à la qualité des réponses ou des comportements émis par un individu. Sternberg et al. (1981) concluent que les gens ordinaires définissent les comportements intelligents selon trois catégories principales: la capacité à résoudre des problèmes pratiques, la capacité verbale (lire, écrire, converser, etc.) et la

compétence sociale. Les experts du domaine sont généralement en accord avec ces catégories de définitions (Sternberg et al., 1981) mais ils vont davantage élaborer sur ces thèmes. Groth-Marnat (1984) note d'ailleurs que les définitions explicites de l'intelligence élaborent généralement sur l'un ou plusieurs des thèmes suivants: la pensée abstraite, l'apprentissage par expérience, la résolution de problèmes, l'ajustement à de nouvelles situations et la réalisation d'un but ou d'un objectif.

Selon Weinberg (1989), plusieurs avenues théoriques fournissent des définitions de l'intelligence. Entre autres, les chercheurs contemporains privilégient l'étude de l'intelligence selon des approches de traitement de l'information. Ils étudient et définissent l'intelligence à partir des processus cognitifs impliqués dans la résolution de problèmes. D'autres chercheurs définissent cependant l'intelligence selon un paradigme développemental. Ils recherchent alors la description des composantes qualitatives du développement cognitif, non la description des processus impliqués lors d'un travail cognitif. Par ailleurs, l'approche psychométrique tente plutôt l'analyse quantitative des habiletés mentales, où la performance des gens dans la résolution de problèmes est mesurée. La présente étude s'intéresse particulièrement à cette dernière approche.

Approche psychométrique. L'approche psychométrique offre trois définitions théoriques principales du concept d'intelligence (Weinberg, 1989). Une première définition indique que l'intelligence est une capacité cognitive globale, unifiée et commune à toutes sortes d'activités intellectuelles. Selon cette définition, l'intelligence consiste en un facteur général, "g", décrit par Spearman.

Une deuxième définition de l'intelligence s'appuie sur plusieurs habiletés cognitives séparées. Thurstone a vérifié la thèse de Spearman à l'aide d'une série d'analyses factorielles sur plusieurs tests d'intelligence. Ses analyses montrent que l'intelligence ne consiste pas en un facteur "g" mais plutôt en plusieurs habiletés cognitives spécifiques et séparées (Groth-Marnat, 1984; Sattler, 1992). Thurstone a identifié sept facteurs distincts de l'intelligence: la fluidité verbale, la compréhension verbale, l'aptitude spatiale, la vitesse perceptive, le raisonnement, l'aptitude numérique et la mémoire. Cette définition psychométrique de l'intelligence suppose l'indépendance des divers facteurs cognitifs et donne lieu à des théories multifactorielles de l'intelligence comme celle de Guilford par exemple.

Des recherches subséquentes effectuées à l'aide des facteurs identifiés par Thurstone ont cependant montré qu'ils sont corrélés les uns aux autres (Sattler, 1992). Ces recherches ont donné lieu à une troisième série de

définitions de l'intelligence décrite sous forme hiérarchique (Weinberg, 1989). Selon cette conception, l'intelligence est décrite comme une organisation hiérarchique consistant d'un à deux facteurs généraux qui se subdivisent en habiletés plus spécifiques, corrélées et donnant lieu au facteur "g" (Groth-Marnat, 1984; Sattler, 1992; Weinberg, 1989).

Wechsler (1974a, 1974b, 1975) a mis cette conceptualisation hiérarchique en théorie et en pratique. Au point de vue théorique, il reconnaît la présence de composantes intellectuelles verbales et non-verbales corrélées et rattachées à un facteur d'intelligence globale ("g"). En pratique, cette notion hiérarchique des habiletés intellectuelles a servi au développement de plusieurs échelles d'intelligence associées au nom de Wechsler. En plus de définir l'intelligence à l'aide de facteurs intellectuels (global, verbal et non-verbal), Wechsler reconnaît aussi que l'intelligence se compose de facteurs non-intellectuels (personnalité) et de facteurs permettant le respect des valeurs morales, esthétiques et sociales. L'intelligence est donc une manifestation de la personnalité des individus, en plus d'être une manifestation de leurs habiletés intellectuelles.

Horn et Cattell (1966) ont également défini l'intelligence comme une organisation hiérarchique de plusieurs facteurs. Leurs travaux ont mené au développement

de la théorie de l'intelligence fluide (Gf) et de l'intelligence cristallisée (Gc), deux facteurs dominants qui découlent directement de "g" (Grégoire, 1992; Papalia & Olds, 1985; Sattler, 1992). L'intelligence fluide est signe de flexibilité cognitive puisqu'elle permet une adaptation à de nouvelles situations. Elle dépend du développement neurophysiologique des individus et n'est pas assujettie aux influences de l'environnement. Par contre, l'intelligence cristallisée est la faculté de résoudre des problèmes en fonction d'un ensemble d'informations générales déjà accumulées. Cet aspect de l'intelligence est relié à l'éducation et à la culture des individus. Ces concepts de l'intelligence fluide et de l'intelligence cristallisée ont servi au développement théorique de la dernière révision des échelles d'intelligence Stanford-Binet (Thorndike et al., 1986a).

Évaluation cognitive

En psychométrie, l'intelligence est quantifiée. Binet a été le premier à proposer une représentation quantitative de l'intelligence, où la différence entre l'âge chronologique et l'âge mental d'un individu (mesuré à partir d'un test) servait d'indice quantitatif de son niveau d'intelligence (Groth-Marnat, 1984). Aujourd'hui, la notion de quotient intellectuel de déviation (QI) est couramment utilisée. Cette représentation quantitative de l'intelligence implique que le fonctionnement cognitif des individus se distribue

normalement autour d'une moyenne et selon un écart-type qui demeurent les mêmes pour chaque groupe d'âge. À l'aide du QI de déviation, les tests d'intelligence permettent de comparer la performance cognitive d'un individu à celle d'un barème normatif. L'étude pratique de l'intelligence en psychométrie favorise donc le développement de normes pour l'évaluation du fonctionnement cognitif des individus.

Aux États-Unis, les premières mesures d'habiletés intellectuelles ont été développées en milieu scolaire (Messadié, 1995; Sattler, 1992). En 1905, Alfred Binet et Théodore Simon mettaient à l'épreuve un premier test d'intelligence servant à repérer les écoliers qui échouaient en raison de difficultés scolaires. Depuis l'échelle Binet-Simon, plusieurs autres tests d'intelligence ont été développés, entre autres les échelles pour enfants de Wechsler qui sont couramment utilisées (Grégoire, 1992; Groth-Marnat, 1984; Kaufman, 1994; Sattler, 1992).

La plus récente révision du test de Wechsler, le WISC-III (Wechsler, 1991), comporte de nombreuses améliorations par rapport aux versions antérieures de ce test. Entre autres, des améliorations sont notées auprès des normes du test révisé. Les normes de sa version précédente, le WISC-R (Wechsler, 1974a), dataient de plus de 10 ans et étaient périmées. Une mise à jour des normes du test s'imposait puisqu'une amélioration d'environ 3 points par décennie a été notée dans le calcul du QI (Flynn, 1987;

Kaufman, 1994; Wechsler, 1991). La dernière révision des échelles Stanford-Binet est aussi utilisée en psychologie scolaire. En plus de bénéficier de normes plus récentes, ces échelles ont connu une amélioration théorique importante avec la venue du Stanford-Binet IV (Thorndike et al., 1986a, 1986b).

WISC-III

Le WISC-III est une version contemporaine du WISC-R (Wechsler, 1974a); 73% de la composition de la nouvelle échelle provient de la version précédente (Wechsler, 1991). Le WISC-III comporte 10 sous-tests standard également répartis en une échelle de quotient intellectuel verbal (QIV; Connaissances, Similitudes, Arithmétique, Vocabulaire et Jugement) et en une échelle de quotient intellectuel non-verbal ou de performance (QIP; Images à compléter, Codage, Arrangement d'images, Dessins avec blocs et Assemblages). La combinaison des résultats de ces deux échelles donne un indice d'intelligence globale ou quotient intellectuel total (QIT). Trois sous-tests supplémentaires (Empan numérique, Recherche de symboles et Labyrinthes) sont disponibles lors d'une substitution nécessaire des résultats d'un sous-test standard dans le calcul du QIT. L'ensemble de ces sous-tests permettent l'interprétation de quatre index factoriels optionnels. Ces derniers se composent de l'index de Compréhension verbale, celui de l'Organisation perceptuelle, celui d'Attention/Concentration et celui de Vitesse

psychomotrice. Par rapport au WISC-R, le WISC-III permet l'évaluation d'un nouveau facteur (Vitesse psychomotrice) défini à l'aide des résultats du sous-test Codage et du nouveau sous-test Recherche de symboles.

La standardisation du WISC-III est des plus remarquables (Edwards & Edwards, 1993; Kaufman, 1993; Little, 1992; Sattler, 1992). L'échantillon américain de normalisation regroupe 2200 enfants divisés en 11 groupes d'âges de 6 à 16 ans. Chaque groupe d'âge est composé de 200 sujets (100 filles et 100 garçons). Leur répartition selon la race/ethnie, la région géographique et le niveau d'éducation parentale, correspond à celle du recensement de 1988 aux États-Unis. Récemment, les résultats d'une étude canadienne (n = 1100) ont permis la publication de normes canadiennes pour l'administration standardisée du WISC-III auprès des enfants anglophones (Wechsler, 1996). Dans cette étude, les enfants canadiens-anglais obtiennent des résultats au QIT qui sont d'environ 3,34 points supérieurs au QIT standard américain. De nouvelles normes ont été définies pour les enfants anglo-canadiens afin de permettre des comparaisons plus justes de leurs résultats. Toutefois, l'utilisation des normes canadiennes-anglaises n'est pas recommandée lorsque les résultats du WISC-III sont comparés à ceux d'un autre test dont les normes n'ont pas été adaptées à la population cible (comme se sera le cas au Stanford-Binet IV dans la présente étude).

Avantages et désavantages du WISC-III. La modernisation du matériel psychométrique du WISC-III a été l'une des améliorations les plus évidentes depuis le WISC-R (Edwards & Edwards, 1993; Kaufman, 1993, 1994; Little, 1992; Sattler, 1992; Wechsler, 1991). En effet, le WISC-III contient des images de couleurs attrayantes pour les enfants. Le contenu des items a aussi été adapté pour permettre une représentation plus équitable des deux sexes et des différents groupes ethniques à travers les items des sous-tests.

Certaines directives et procédures d'administration ont aussi été révisées afin de faciliter l'utilisation du test et d'en améliorer la structure (Wechsler, 1991). Entre autres, l'ajout de nombreux items au contenu de plusieurs sous-tests tente de réduire les effets de plafonnement et de plancher (Bracken, 1988). De plus, l'ordre d'administration des sous-tests a été légèrement modifié depuis le WISC-R en vue de promouvoir l'établissement d'un rapport avec les enfants dès le début du test. L'administration du WISC-III débute avec un sous-test de nature non-verbale (Images à compléter), et non plus comme au WISC-R avec le sous-test Connaissances, maintenant administré en second lieu.

Un nouveau test de nature non-verbale, le sous-test de Recherche de symboles, constitue l'apport original du WISC-III (Little, 1992). Les raisons cliniques et empiriques de ce sous-test supplémentaire sont cependant mal définies

et sa validité n'a toujours pas été démontrée (Kaufman, 1993; Sattler, 1992; Sternberg, 1993). Jusqu'à présent, le seul indice d'utilité de ce nouveau sous-test est associé au développement du quatrième facteur, Vitesse psychomotrice (Little, 1992; Reynolds & Ford, 1994).

Kaufman (1993) remet en question l'importance accordée à la vitesse de réponse aux sous-tests non-verbaux: des points supplémentaires sont accordés pour de bonnes réponses émises en un temps limite. Cependant, pour les enfants de 12 ans et plus, la cote standard moyenne de 10 points est parfois inaccessible sans l'obtention de points supplémentaires. Les profils de QI peuvent donc varier en fonction de la vitesse de réponse des individus, remettant donc en question les construits mesurés à l'intérieur des sous-tests affectés. Une autre critique importante de Kaufman (1993) se rapporte aux sous-tests verbaux du WISC-III. À son avis, les sous-tests Connaissances, Arithmétique, Vocabulaire et Jugement requièrent un niveau d'apprentissage scolaire et culturel qui mettrait en désavantage les enfants de milieux minoritaires, les enfants bilingues (hispanophones et immigrants aux États-Unis) et les enfants ayant des troubles d'apprentissage. L'évaluation de ces populations, à l'aide de l'échelle verbale du WISC-III, serait donc plus ou moins juste.

Malgré ces critiques, le WISC-III demeure une révision importante des échelles pour enfants de Wechsler. Un

parcours des études comparatives entre le WISC-III et le WISC-R va permettre d'apprécier la validité des résultats obtenus au WISC-III.

Études comparatives avec le WISC-R. Wechsler (1991)

présente les résultats de plusieurs études corrélationnelles entre le WISC-III et le WISC-R. Puisqu'un faible pourcentage de l'échelle d'intelligence fut modifié depuis le WISC-R, les corrélations entre les résultats des deux tests devraient être élevées.

Dans une première étude, le WISC-R et le WISC-III sont administrés à un échantillon de 206 élèves non exceptionnels âgés de 6 à 16 ans (Wechsler, 1991). La présentation des tests est contrebalancée, avec un intervalle médian de 21 jours. La comparaison des résultats offre des corrélations relativement élevées entre les trois échelles principales de QI: des corrélations de .89 au QIT, de .90 au QIV et de .81 au QIP. L'étendue des corrélations observées entre les sous-tests du WISC-R et du WISC-III varie entre .42 (Arrangement d'images) et .80 (Connaissances). De façon générale, les sous-tests des échelles verbales sont plus hautement corrélés d'un test à l'autre que les sous-tests des échelles non-verbales. Ces résultats pourraient entre autres s'expliquer par l'importance accordée à la vitesse de réponse à plusieurs sous-tests de l'échelle QIP du WISC-III (Kaufman, 1993).

L'analyse des résultats obtenus aux échelles de QI aux deux tests révèle des différences significatives entre leurs moyennes d'intelligence (Wechsler, 1991). Au WISC-III, les moyennes d'intelligence sont de 101,5 (QIV), 104,2 (QIP) et 102,9 (QIT); elles s'élèvent à 103,9 (QIV), 111,6 (QIP) et 108,2 (QIT) au WISC-R. Les différences significatives de 2,7 points au QIV, 7,4 points au QIP et 5,3 points au QIT indiquent que le WISC-III donne des moyennes d'intelligence inférieures à celles du WISC-R. Les modifications apportées depuis le WISC-R semblent plus prononcées à l'échelle d'intelligence non-verbale.

La comparaison des résultats d'un groupe d'enfants exceptionnels (trouble d'apprentissage, trouble de lecture ou trouble d'hyperactivité avec déficit de l'attention) au WISC-III et au WISC-R rapporte des corrélations plus faibles que celles retrouvées auprès du groupe précédent (Wechsler, 1991). Toutefois, les corrélations s'élèvent à 0,86, 0,86 et 0,73 respectivement entre les échelles QIT, QIV et QIP. Les différences calculées entre les moyennes d'intelligence des deux tests sont significatives. Les résultats du WISC-III sont en moyenne inférieurs à ceux du WISC-R d'environ 6 points au QIT et 5 points au QIV et au QIP.

Des différences significatives semblables entre les trois échelles du WISC-R et du WISC ont déjà été observées en faveur du WISC (Larrabee & Holroyd, 1976). Lorsqu'une édition révisée d'un test est comparée à une version

précédente, les données recueillies sont souvent inférieures à celles de son prédécesseur (Bracken, 1988; Flynn, 1987; Kaufman, 1994; Larrabee & Holroyd, 1976). En effet, les tests publiés depuis au moins une dizaine d'années ont tendance à surestimer les moyennes de QI des individus parce que leurs normes sont vieilles. Les instruments révisés accompagnés de nouvelles normes donnent ainsi l'impression que leurs résultats sont inférieurs à ceux des versions précédentes.

D'après les corrélations obtenues entre le WISC-III et le WISC-R, Wechsler (1991) conclut que les construits mesurés à l'aide des tests sont similaires. Par contre, l'examen du fonctionnement cognitif à l'aide du WISC-III donne des résultats d'environ 5 points inférieurs à ceux du WISC-R. De plus, Wechsler indique que ces différences de QI se retrouvent davantage entre les échelles non-verbales des deux tests.

Plusieurs études subséquentes menées auprès d'enfants de diverses populations ont confirmé que le WISC-III donne des QI inférieurs à ceux du WISC-R, tout en y étant hautement corrélé (Bolen, Aichinger, Hall, & Webster, 1995; Doll & Boren, 1993; Graf & Hinton, 1994; Newby, Recht, Caldwell, & Schaefer, 1993; Slate, 1995a; Slate & Jones, 1995; Vance, Maddux, Fuller, & Awadh, 1996). La plupart de ces études rapportent que les corrélations entre les échelles non-verbales des tests sont les plus faibles (Bolen

et al., 1995; Doll & Boren, 1993; Graf & Hinton, 1994; Slate, 1995a; Slate & Jones, 1995; Vance et al., 1996), tout comme Wechsler (1991) l'a aussi observé. Par contre, de ces études, seulement trois d'entre elles ont obtenu des différences significatives plus élevées entre les cotes des échelles non-verbales (Bolen et al., 1995; Graf & Hinton, 1994; Vance et al., 1996). Il est à noter que les échantillons de ces trois études comptent plus de sujets que ceux des autres études, améliorant ainsi leur correspondance aux études de Wechsler et la puissance de leur test statistique.

Graf et Hinton (1994) ainsi que Slate et Jones (1995) ont obtenu des cotes moyennes globales à peu près équivalentes au WISC-III et au WISC-R. Quoique les différences obtenues soient de moins de 5 points, elles demeurent significatives, à l'exception de la différence calculée entre les échelles verbales de l'étude de Graf et Hinton. Sabatino, Spangler et Vance (1995), pour leur part, n'ont obtenu aucune différence statistique entre le QI du WISC-III et celui du WISC-R. Leurs résultats suggèrent que les profils d'habiletés cognitives d'enfants surdoués ($n = 51$) connaissent très peu de variations. D'ailleurs, les auteurs de cette étude expliquent la présence de faibles intercorrélations (0,36, 0,37 et 0,33 respectivement au QIT, QIV et QIP) par l'étendue restreinte des cotes standard chez ces enfants, et non par la mesure de construits différents

d'un test à l'autre. Par contre, il semble théoriquement possible d'observer des différences entre les cotes de QI des deux tests (Bracken, 1988; Flynn, 1987). L'absence d'écart auprès de l'échantillon d'étude de Sabatino et al. pourrait être signe d'effets de plafonnement non-correctés au WISC-III.

Des moyennes de QI significativement inférieures au WISC-III par rapport au WISC-R chez la plupart des populations sont rapportées dans l'ensemble de la littérature. Ces différences peuvent être attribuées à la modernisation des normes psychométriques (Bracken, 1988). Il semble donc important de comparer les performances du WISC-III à un autre instrument de mesure qui aborde des construits semblables à ceux des échelles de Wechsler et dont les normes seraient plus récentes que celles du WISC-R. La comparaison du WISC-III au Stanford-Binet IV semble pertinente puisque ces deux échelles d'intelligence partagent plusieurs ressemblances et que les normes du Stanford-Binet IV ont été publiées en 1986. La section suivante introduit le Stanford-Binet IV (Thorndike et al., 1986a, 1986b) en le comparant d'abord à sa version précédente, le Stanford-Binet L-M (Terman & Merrill, 1960). Par la suite, des études comparatives entre les échelles de Wechsler et Stanford-Binet sont recensées.

Stanford-Binet IV

L'utilisation des échelles Stanford-Binet est courante

dans l'évaluation de l'intelligence des enfants et des jeunes adultes, surtout pour les populations d'individus surdoués (Glutting, 1989; Murphy & Davidshofer, 1991; Robinson, 1992; Silverman & Kearney, 1992). Le Stanford-Binet IV (Thorndike et al., 1986a, 1986b) constitue la première révision du test basée sur un modèle hiérarchique défini, puisé entre autres dans la théorie de l'intelligence fluide et de l'intelligence cristallisée de Horn et Cattell (1966). Le premier niveau de la hiérarchie consiste en un indice d'habileté cognitive générale (Test composite ou TC; Thorndike et al., 1986a). Seul cet indice était obtenu à l'intérieur des versions anciennes des échelles de Stanford-Binet. Avec le Stanford-Binet IV, un deuxième et un troisième niveaux ont été ajoutés. Le deuxième niveau représente les facteurs de l'intelligence cristallisée, de l'intelligence fluide et de la mémoire à court terme (MCT). Le troisième niveau comprend le facteur de Raisonnement verbal (RV), le facteur de Raisonnement mathématique (RM) et le facteur de Raisonnement abstrait/visuel (RA/V). Ces facteurs du troisième niveau (RV, RM et RA/V), accompagnés du facteur MCT de deuxième niveau, représentent les quatre facteurs d'habiletés intellectuelles spécifiques mesurés au Stanford-Binet IV.

Chacun de ces quatre facteurs est mesuré à l'aide de différents sous-tests. Les sous-tests Vocabulaire, Compréhension, Absurdités et Relations verbales servent à

mesurer le RV; les sous-tests Analyse de patrons, Copie, Matrices ainsi que Pliage et découpage de papier servent à calculer l'indice du RA/V. Les sous-tests Quantités, Séries de nombres et Construction d'équations mesurent le RM, et les sous-tests de Mémorisation de chiffres, de perles, de phrases et d'objets donnent un indice de MCT. L'échelle entière compte au total 15 sous-tests qui ne sont pas tous administrés à un même individu. Le choix des sous-tests à administrer est basé sur l'âge du sujet et sur le niveau de ses capacités intellectuelles, défini à l'aide du sous-test Vocabulaire (Thorndike et al., 1986a).

L'échantillon de standardisation du Stanford-Binet IV compte 5013 individus répartis à l'intérieur de 17 groupes d'âge de 2 ans à 23 ans (Thorndike et al., 1986b). La sélection des individus de l'échantillon normatif a été basée sur la région géographique, la densité de population, le groupe ethnique, l'âge et le sexe des sujets, de sorte que la représentation de ces variables soit semblable à celle du recensement de 1980 aux États-Unis.

Avantages et désavantages du Stanford-Binet IV. Le Stanford-Binet IV est une révision du Stanford Binet, forme L-M (Stanford-Binet L-M; Terman & Merrill, 1960). Avec cette nouvelle version du test, Thorndike et al. (1986b) ont voulu créer une échelle plus apte à mesurer les habiletés cognitives des jeunes de 2 à 23 ans, tout en conservant les forces du Stanford-Binet L-M. Ainsi, les items faciles à

corriger, sans biais culturel, sans préjugé sexiste et applicables à plusieurs groupes d'âge, ont été préservés. Cependant, même si le nouveau test respecte une certaine continuité avec le Stanford-Binet L-M, son format en âge standard et son format théorique le distinguent de ses prédécesseurs (Glutting, 1989).

Thorndike et al. (1986b) indiquent que l'un des avantages du Stanford-Binet IV consiste en son administration adaptée selon l'âge des sujets et leur niveau de fonctionnement intellectuel. La combinaison de ces facteurs sert à déterminer le pallier de départ de chaque sous-test, en plus d'identifier les sous-tests à administrer. Le temps d'administration de l'échelle est relativement réduit, permettant d'éliminer la présentation d'items trop faciles ou trop difficiles pour un individu (Glutting, 1989).

Sattler (1992) donne l'avis que l'administration de ce test peut être de longue durée. En effet, le Stanford-Binet IV peut nécessiter jusqu'à 90 minutes d'administration. Des versions abrégées de ce test sont toutefois disponibles (Thorndike et al., 1986b). Cependant, malgré les bonnes corrélations retrouvées entre les versions abrégées et la batterie complète, les différences diagnostiques notées entre leurs administrations sont peu documentées (Glutting, 1989).

Une amélioration importante de la nouvelle échelle de Stanford-Binet consiste au regroupement par facteurs des cotes standard des sous-tests (Thorndike et al., 1986b). L'étendue des habiletés cognitives mesurées au Stanford-Binet IV est beaucoup plus large que celle rencontrée au Stanford-Binet L-M. L'indice du fonctionnement intellectuel global étant calculé au Stanford-Binet IV est plus exact. La structure factorielle du Stanford-Binet IV fait toutefois l'objet de critiques et elle demeure toujours mal définie (Sattler, 1992). Cependant, Robinson (1992) indique qu'elle représente une amélioration comparativement à la structure factorielle des échelles de Wechsler, puisqu'au Stanford-Binet IV, une distinction est faite entre le facteur de raisonnement mathématique et celui du raisonnement verbal.

Les niveaux de plafonnement du Stanford-Binet IV sont adéquats et permettent de déceler les enfants surdoués dès l'âge de 2 ans (avec un maximum possible d'un QI de 164 au TC) ainsi que les jeunes adultes surdoués de 20 ans (avec un maximum possible d'un QI de 149 au TC) (Glutting, 1989). Cependant, Silverman et Kearney (1992) ne recommandent pas l'utilisation du Stanford-Binet IV auprès des enfants surdoués. Même si les nouvelles normes du Stanford-Binet IV représentent une adaptation aux changements cognitifs de la population (Flynn, 1987), leur niveau de plafonnement est inférieur à celui du Stanford-Binet L-M. Les normes plus récentes discriminent donc moins bien les enfants surdoués

comparativement au Stanford-Binet L-M (Kitano & De Leon, 1988; McCall, Yates, Hendricks, Turner, & McNabb, 1989). Les niveaux plancher du Stanford-Binet IV sont aussi moins sensibles aux troubles cognitifs des jeunes enfants que ceux du Stanford-Binet L-M. Le Stanford-Binet IV permet de détecter le retard mental léger à partir de l'âge de 4 ans, tandis que le Stanford-Binet L-M le permet dès l'âge de 3 ans. Pour les très jeunes enfants, le Stanford-Binet IV n'est donc pas l'instrument de choix pour le dépistage du retard mental (Robinson, 1992).

Études comparatives avec le Stanford-Binet L-M. Le Stanford-Binet IV poursuit donc la tradition des échelles d'intelligence de Stanford-Binet. Plusieurs études corrélationnelles ont d'ailleurs tenté de comparer les résultats du Stanford-Binet IV à ceux du Stanford-Binet L-M afin de découvrir leur correspondance empirique et clinique.

Les auteurs du test ont effectué plusieurs comparaisons entre les deux dernières éditions de l'échelle (Thorndike et al., 1986b). Lorsque les résultats du Stanford-Binet IV et du Stanford-Binet L-M sont comparés chez des enfants non exceptionnels ($n = 139$), l'indice de corrélation calculé est élevé ($r = 0,81$). Les intercorrélations entre chaque secteur du Stanford-Binet IV et l'indice global du Stanford-Binet L-M varient de 0,56 (RA/V) à 0,76 (RV). Pour cet échantillon, la moyenne du TC au Stanford-Binet IV s'élève à 105,8. Au Stanford-Binet L-M par contre, la moyenne du TC

est supérieure à celle du Stanford-Binet IV, avec une moyenne de 108,1. Thorndike et al. (1986b) notent toutefois que cette différence pourrait être artificielle. D'abord parce que les années de standardisation des deux tests sont différentes. Ensuite parce que l'administration des deux tests n'a pas été contrebalancée de façon adéquate. En effet, 86% des enfants de leur étude ont été évalués premièrement à l'aide du Stanford-Binet IV. Une autre étude d'enfants non exceptionnels (n = 30) a d'ailleurs conclu que la classification des individus est la même à l'aide des deux tests (Hartwig, Sapp, & Clayton, 1987). Aucune différence significative n'a été obtenue entre la moyenne du TC au Stanford-Binet IV (114,43) et la moyenne du TC au Stanford-Binet L-M (113,07). Par contre, les conclusions de cette étude sont peu généralisables puisque l'échantillon d'étude compte peu de sujets.

Plusieurs recherches se sont davantage intéressées aux comparaisons des résultats du Stanford-Binet IV et du Stanford-Binet L-M auprès d'enfants surdoués (Kitano & De Leon, 1988; Kluever & Green, 1990; McCall et al., 1989; Thorndike et al., 1986b). De façon générale, la plupart de ces enfants obtiennent des résultats supérieurs au Stanford-Binet L-M, comme l'ont indiqué Silverman et Kearney (1992). Les écarts calculés entre la moyenne du TC au Stanford-Binet IV et la moyenne du TC au Stanford-Binet L-M se situent généralement entre 10 et 13,5 points (Kluever & Green, 1990;

McCall et al., 1989; Thorndike et al., 1986b). Par contre, les indices de corrélations calculés entre ces deux tests sont faibles (variant de $r = 0,09$ à $r = 0,40$) car l'étendue des résultats est restreinte chez ces échantillons d'enfants. Kitano et De Leon (1988) n'ont de leur côté décelé aucune différence significative entre les résultats des deux tests chez un groupe de 148 enfants surdoués. Cependant, les comparaisons entre les données du Stanford-Binet IV et celles du Stanford-Binet L-M n'ont pas été effectuées auprès des mêmes enfants. En fait, le Stanford-Binet L-M a été présenté à 104 enfants et le Stanford-Binet IV à 44 autres enfants. Leurs conclusions sont donc basées sur des comparaisons inter-groupes et non inter-sujets.

Glutting (1989) et Robinson (1992) indiquent que les niveaux plancher du Stanford-Binet IV sont moins sensibles au fonctionnement intellectuel sous la moyenne que ceux de son prédécesseur; les résultats du Stanford-Binet IV devraient donc être supérieurs à ceux du Stanford-Binet L-M. Thorndike et al. (1986b) ont obtenu des moyennes de 76,9 au Stanford-Binet L-M et de 79,9 au TC du Stanford-Binet IV auprès d'un groupe d'enfants avec troubles d'apprentissage ($n = 14$). La moyenne du TC au Stanford-Binet IV est donc supérieure au TC du Stanford-Binet L-M, mais ces indices demeurent hautement corrélés malgré l'écart de 3 points

($r = 0,79$). Par contre, le nombre de sujets de cette étude est faible et que la généralisation de ces résultats est douteuse.

Chez des jeunes souffrant d'un retard intellectuel de léger à modéré ($n = 31$), Lukens (1988) a décelé plusieurs patrons de résultats. L'examen des cotes individuelles a révélé que 19 sujets ont obtenu des résultats supérieurs au Stanford-Binet IV tandis que 11 sujets ont obtenu de meilleurs résultats à l'aide du Stanford-Binet L-M. Un seul des sujets a obtenu des résultats identiques aux deux tests. La différence absolue maximale entre les résultats du Stanford-Binet L-M et du Stanford-Binet IV s'est élevée à 12 points. Par contre, lorsque la moyenne du groupe est calculée, une faible différence de 2 points est notée en faveur du Stanford-Binet L-M. Après d'un autre groupe d'enfants de cette même population ($n = 22$), en moyenne, le Stanford-Binet IV donne des résultats supérieurs au Stanford-Binet L-M (Thorndike et al., 1986b). Toutefois, la faible différence de 1,4 point calculée entre les résultats globaux des deux tests est semblable à celle rapportée par Lukens.

La présence de différences significatives entre les moyennes du Stanford-Binet IV et du Stanford-Binet L-M peut s'expliquer entre autres par les différentes dates de publication des deux révisions et les changements cognitifs de la population (Bracken, 1988; Flynn, 1987). Par contre, à

l'intérieur de certaines études, notamment celle de Thorndike et al. (1986b) et celle de Lukens (1988) auprès d'enfants souffrant d'un retard intellectuel, les différences entre les deux tests peuvent être attribuées à l'administration non-contrebalancée des tests. Il peut sembler arbitraire de comparer le Stanford-Binet L-M au Stanford-Binet IV puisque ces deux échelles mesurent l'intelligence de façon différente (Kluever & Green, 1990). Notamment, le Stanford-Binet IV se base sur quatre différents secteurs d'habiletés cognitives pour calculer l'indice de TC, tandis que le Stanford-Binet L-M l'évalue à partir des habiletés cognitives verbales. Afin de mieux apprécier la plus récente version des échelles Stanford-Binet, il serait préférable de la comparer à un autre instrument qui mesure des construits semblables, par exemple le WISC-III.

Études comparatives entre les échelles d'intelligence de Wechsler et de Stanford-Binet

Comparaisons entre le WISC-R et le Stanford-Binet IV.

Plusieurs études corrélationnelles ont comparé le Stanford-Binet IV au WISC-R. Thorndike et al. (1986b) avaient formulés plusieurs hypothèses sur la correspondance des résultats aux deux tests. Entre autres, grâce à leur bagage élevé en habiletés verbales, les facteurs RV, RM et MCT du Stanford-Binet IV seraient plus hautement corrélés au QIV et au QIT qu'ils ne le seraient au QIP du WISC-R. Le facteur

RA/V serait plus fortement corrélé au QIP étant donné le contenu non-verbal de ces deux échelles.

Une étude réalisée auprès de 205 enfants non exceptionnels a confirmé en partie ces hypothèses (Thorndike et al., 1986b). En effet, le RV est hautement corrélé au QIV ($r = 0,72$) et avec le QIT ($r = 0,73$), mais moins corrélé au QIP ($r = 0,60$). Par contre, comparativement aux attentes de départ, les corrélations obtenues entre le facteur RM et les trois échelles de QI du WISC-R sont semblables ($r = 0,64$, $0,63$ et $0,69$ respectivement avec le QIV, le QIP et le QIT), tout comme les corrélations observées entre le facteur MCT et les trois échelles du WISC-R ($r = 0,64$, $0,63$ et $0,70$ respectivement avec le QIV, le QIP et le QIT). Les résultats précisent que le RA/V est autant corrélé au QIP ($r = 0,67$) qu'il ne l'est au QIV ($r = 0,68$). L'indice de corrélation calculé entre le TC et le QIT demeure néanmoins le plus élevé ($r = 0,83$).

Malgré les bonnes corrélations obtenues entre les deux mesures d'intelligence, la comparaison des moyennes indique toutefois la présence de différences significatives entre les résultats des échelles QIP ($M = 105,3$) et RA/V ($M = 98,9$) ainsi qu'entre le QIT ($M = 105,2$) et le TC ($M = 102,4$) (Thorndike et al., 1986b). D'après les résultats de cette étude, le WISC-R donne des indices de QIP et de QIT significativement plus élevés que les indices correspondants

du Stanford-Binet IV auprès d'un groupe d'enfants non exceptionnels.

Des études subséquentes ont comparé les résultats du WISC-R à ceux du Stanford-Binet IV. Rothlisberg (1987) a mené une étude auprès d'enfants de classes régulières ($n = 32$) et a obtenu des résultats similaires à ceux de Thorndike et al. (1986b). Les résultats de cette étude signalent que le RA/V donne des résultats significativement inférieurs à ceux du QIP, avec une différence de 11,4 points entre ces deux échelles. La moyenne des résultats du QIT ($\bar{M} = 112,53$) est supérieure à celle du TC ($\bar{M} = 105,53$). La différence de 1,1 point observée entre le QIV et le RV n'est cependant pas significative.

Une étude de Lukens (1990) effectuée auprès de 30 enfants avec retard mental léger révèle par ailleurs une différence significative entre le TC ($\bar{M} = 66,57$) et le QIT ($\bar{M} = 61,97$), en faveur du Stanford-Binet IV. Hollinger et Baldwin (1990) notent eux aussi que le Stanford-Binet IV donne des moyennes de QI plus élevées que le WISC-R chez un groupe d'enfants référés ($n = 19$), mais seulement lors de comparaisons entre le RV ($\bar{M} = 86,90$) et le QIV ($\bar{M} = 83,95$) ainsi qu'entre le RA/V ($\bar{M} = 89,42$) et le QIP ($\bar{M} = 85,84$). Les résultats obtenus aux indices d'intelligence globale des deux tests sont les mêmes (QIT = 83,63 et TC = 83,11).

Les résultats de ces deux dernières études indiquent que le Stanford-Binet IV donne des résultats supérieurs au

WISC-R chez des enfants référés. Les résultats de Prewett et Matavich (1992) pourraient servir à expliquer ces résultats. Auprès d'un groupe d'enfants référés ($n = 126$), ces auteurs montrent qu'en effet, le Stanford-Binet IV donne des résultats significativement plus élevés ($M = 71,54$) lorsque les enfants sont classifiés à des niveaux cognitifs de retard mental léger d'après le WISC-R ($M = 62,77$). Lorsque les résultats d'enfants avec intelligence moyenne (selon le WISC-R) sont comparés, le WISC-R ($M = 94,27$) donne des résultats significativement supérieurs au Stanford-Binet IV ($M = 87,86$). Les données de cette étude rejoignent donc celles de Rothlisberg (1987) et de Thorndike et al., où les enfants d'intelligence moyenne ont une meilleure performance au WISC-R comparativement au Stanford-Binet IV. Il semble donc que les différences significatives entre les résultats du WISC-R et du Stanford-Binet IV varient selon le niveau intellectuel des individus.

Par ailleurs, Carvajal et Weyand (1986), Phelps, Bell et Scott (1988) ne notent aucune différence entre les résultats des deux tests. L'examen de ces deux études révèle cependant des erreurs méthodologiques au niveau de l'administration et de l'analyse des données. Entre autres, des effets de pratique se sont probablement glissés dans les résultats de l'étude de Carvajal et Weyand puisque l'administration des tests n'a pas été contrebalancée. De plus, les conclusions proposées par Phelps et al. sont

basées sur des corrélations et non sur des comparaisons statistiques des moyennes des tests. Ainsi, leurs conclusions ne sont pas accompagnées de preuves statistiques adéquates.

Les études de comparaisons entre le WISC-R et le Stanford-Binet IV confirment donc en général la mesure de construits semblables à l'intérieur des deux tests. Les différences entre les cotes de QI du WISC-R et du Stanford-Binet IV peuvent s'expliquer en théorie par les dix années qui séparent leurs dates de publication (Bracken, 1988; Flynn, 1987). Il semble toutefois exister des différences entre les moyennes de QI de ces deux tests qui varient en fonction des niveaux intellectuels des sujets (Prewett & Matavich, 1992). Les enfants qui obtiennent des QI moyens atteignent de meilleurs résultats au WISC-R, tandis que ceux qui ont des habiletés cognitives plus faibles obtiennent de meilleurs résultats à l'aide du Stanford-Binet IV.

La variance des écarts attribuables au niveau intellectuel des individus suscite des hypothèses. Le Stanford-Binet IV pourrait être un instrument de mesure moins sensible aux troubles cognitifs, comparativement au WISC-R, puisqu'il donne des résultats supérieurs à celui-ci chez des enfants avec une intelligence sous la moyenne (Glutting, 1989; Robinson, 1992). Contrairement aux attentes théoriques (Bracken, 1988), le WISC-R pourrait sous-estimer (et non surestimer) le fonctionnement cognitif des individus

situés sous la moyenne d'intelligence, malgré ses normes désuètes. Le WISC-R pourrait donc présenter des tâches cognitives plus difficiles que le Stanford-Binet IV chez les enfants qui diffèrent de la norme.

Comparaisons entre le WISC-III et le Stanford-Binet IV.

De façon générale, le WISC-III et le Stanford-Binet IV mesurent des construits semblables, tout comme les études comparant le Stanford-Binet IV au WISC-R. En effet, les indices de corrélations calculés entre le QIT du WISC-III et le TC du Stanford-Binet IV varient entre 0,68 et 0,82 (Carvajal et al., 1993; Lavin, 1996; Lukens & Hurrell, 1996; Prewett & Matavich, 1994; Saklofske, Schwean, Yackulic, & Quinn, 1994). De plus, les corrélations entre leurs construits verbaux ($r = 0,66$ à $0,82$ entre le QIV et le RV) et leurs construits non-verbaux ($r = 0,63$ à $0,71$ entre le QIP et le RA/V) sont relativement élevées. Cependant, lorsque les moyennes des cotes du WISC-III et du Stanford-Binet IV sont comparées, deux situations différentes se présentent.

Notamment, Lavin (1990) ne rapporte aucune différence significative entre le TC (108,00) du Stanford-Binet IV et le QIT (106,95) du WISC-III, ainsi qu'entre les cotes de QI verbal (RV = 107,05 et QIV = 109,32) et les cotes de QI non-verbal (RA/V = 105,45 et QIP = 103,32) chez 40 enfants non exceptionnels. Carvajal et al. (1993) obtiennent des résultats semblables entre les cotes de QI global du

WISC-III (QIT = 112,6) et du Stanford-Binet IV, version abrégée (TC = 109,8) chez un groupe de 32 enfants non exceptionnels.

Par ailleurs, lorsque les résultats des tests sont comparés auprès d'enfants référés, les tests donnent des résultats différents. Une étude de Prewett et Matavich (1994) menée auprès de 73 enfants référés pour difficulté d'apprentissage, révèle des différences significatives entre les résultats du WISC-III et ceux du Stanford-Binet IV. Les écarts moyens entre le QIT (74,55) et le TC (82,92) sont de 9,4 points; les comparaisons des échelles verbales (QIV = 74,08 et RV = 86,37) et des échelles non-verbales (QIP = 79,07 et RA/V = 86,36) s'élèvent respectivement à 13,1 points et 8,1 points. De plus, les moyennes tirées du Stanford-Binet IV sont nettement plus élevées que celles obtenues au WISC-III.

Des études subséquentes ont reproduit les résultats de Prewett et Matavich (1994). Une étude effectuée auprès de 45 enfants avec un trouble d'hyperactivité avec déficit de l'attention (THADA) indique qu'une version abrégée du Stanford-Binet IV donne des résultats supérieurs au WISC-III (Saklofske et al., 1994). Dans cette étude, les écarts entre les QI s'élèvent à 4,0 points entre le QIT (97,96) et le TC (102,00), à 5,5 points entre le QIP (101,36) et le RA/V (106,87) et à 10,2 points entre le QIV (95,51) et le RV (105,73). Lukens et Hurrell (1996) ont eux aussi obtenu des

résultats supérieurs au Stanford-Binet IV. Leurs analyses des résultats du Stanford-Binet IV et du WISC-III chez des enfants avec retard mental léger ($n = 31$), révèlent des différences significatives de 7,5 points, 4,8 points et 10,5 points respectivement entre les cotes de QI global, non-verbal et verbal des deux instruments.

Prewett et Matavich (1994) ont particulièrement noté la différence élevée de 13,1 points calculée entre le QIV et le RV des enfants de leur étude. Cette différence est en effet la plus élevée des comparaisons effectuées et se situe à près d'un écart-type (15 points) entre les deux tests. Saklofske et al. (1994) et Lukens et Hurrell (1996) ont aussi noté une différence importante au niveau des comparaisons entre les échelles verbales. Le WISC-III semble donc donner des résultats inférieurs au Stanford-Binet IV chez des enfants exceptionnels, particulièrement au niveau des habiletés verbales. Cette particularité des écarts entre les échelles verbales n'est pas décelée lors de comparaisons établies entre le Stanford-Binet IV et le WISC-R auprès de populations semblables (Hollinger & Baldwin, 1990; Prewett & Matavich, 1992).

Un test publié depuis déjà quelques années surestime le fonctionnement cognitif des individus (Bracken, 1988; Flynn, 1987); la publication du Stanford-Binet IV précède celle du WISC-III de seulement cinq ans. Les différences entre leurs cotes de QI pourraient par conséquent être moindres puisque

leur date de publication ne sont pas séparées par une dizaine d'années. D'ailleurs, les études menées auprès d'enfants non exceptionnels démontrent que les moyennes de QI des deux tests sont en effet semblables (Carvajal et al., 1993; Lavin, 1990). Tout comme les études comparatives entre le WISC-R et le Stanford-Binet IV l'ont rapporté auprès d'enfants exceptionnels (Lukens, 1990; Hollinger & Baldwin, 1990; Prewett & Matavich, 1992), le Stanford-Binet IV donne des résultats supérieurs au WISC-III (Lukens & Hurrell, 1996; Prewett & Matavich, 1994; Saklofske et al., 1994).

Prewett et Matavich (1994) ont davantage poussé leurs analyses et ont observé la présence d'une différence significative entre les profils verbaux et non-verbaux des deux instruments de mesure. Tandis que le QIV de leur échantillon d'enfants référés est près de 5 points plus faible que le QIP au WISC-III, aucun écart significatif n'est calculé entre les échelles verbale et non-verbale au Stanford-Binet IV. Non seulement les cotes du WISC-III sont généralement inférieures aux cotes du Stanford-Binet IV, mais les impressions diagnostiques établies à l'aide des profils des deux tests sont différentes. Cette dernière constatation pourrait aussi s'appliquer auprès des moyennes de QI des études de Lukens et Hurrell (1996) et de Saklofske et al. (1994).

Le QIV du WISC-III semble jouer un rôle important auprès de ces résultats d'études, tant au niveau des

divergences inter-tests qu'au niveau des divergences intra-tests. D'ailleurs, la critique de Kaufman (1993) indiquant que l'échelle verbale du WISC-III pénalise particulièrement les groupes d'individus qui diffèrent de la norme américaine, appuie cette supposition. Des écarts significatifs entre le QIV et le QIP du WISC-III devraient être observés auprès d'enfants de milieux minoritaires et de milieux bilingues. Il devient donc important d'examiner les études empiriques traitant du sujet afin de mesurer la justesse de ces postulats.

Étude des profils d'intelligence verbale et non-verbale aux échelles de Wechsler

L'observation des écarts QIV-QIP dans l'interprétation des échelles de Wechsler est importante dans la formulation des impressions cliniques (Grégoire, 1992; Groth-Marnat, 1984; Kaufman, 1994; Sattler, 1992). Cette analyse doit d'abord établir la présence d'écarts statistiquement significatifs entre les échelles verbale et non-verbale. Suivant les règles statistiques, pour le WISC-III, des écarts entre les échelles verbale et non-verbale de 11 points et de 15 points sont respectivement significatifs à un seuil de décision statistique de 0,05 et de 0,01 (Wechsler, 1991). En théorie, des écarts de 11 points et plus sont donc non-aléatoires et diffèrent statistiquement d'une différence de 0 point. Néanmoins, une preuve statistique des différences QIV-QIP n'implique pas toujours

une importance au niveau clinique (Grégoire, 1992; Kaufman, 1994; Sattler, 1992). L'amplitude des écarts calculés entre les échelles verbale et non-verbale des échantillons de standardisation des tests s'avère importante.

Lorsque la différence entre les échelles verbale et non-verbale est calculée auprès des 2200 sujets de l'échantillon normatif du WISC-III, une différence moyenne absolue de 10 points est obtenue (Wechsler, 1991). Tandis que 45% de ces enfants ont des écarts absolus de 10 points et plus, 24% obtiennent des écarts significatifs de 15 points et plus. Des différences absolues moyennes semblables sont aussi notées auprès des échantillons normatifs du WISC-R. Les données d'étalonnage du WISC-R américain ont une moyenne absolue d'écarts QIV-QIP de 9,7 points (Kaufman, 1976). Des différences absolues de 9 points et plus sont présentes chez 50% de cet échantillon, tandis que des écarts significatifs de 15 points et plus se retrouvent chez 25% d'entre eux. Les résultats de l'échantillon de standardisation du WISC-R français ont donné une moyenne d'écarts absolus d'environ 10,6 points (Grégoire, 1992). Des écarts absolus de 9 points et plus apparaissent chez 51% des sujets de cet échantillon. Des différences absolues de 15 points et plus se retrouvent chez 27% d'entre eux. Près de la moitié des sujets des groupes de normalisation des tests obtiennent donc des moyennes d'écarts absolus d'environ 10 points. Ainsi, plusieurs individus du groupe normatif

présentent des différences assez importantes entre le QIV et le QIP.

Lorsque la direction des différences QIV-QIP est observée auprès des échantillons de normalisation, une prépondérance équivalente des habiletés verbales et des habiletés non-verbales est notée. Parmi les échantillons de standardisation du WISC-R américain et du WISC-R français, la fréquence des écarts $QIV > QIP$ est équivalente à la fréquence des écarts $QIP > QIV$ (Grégoire, 1992; Kaufman, 1976). Des patrons semblables d'écarts sont probablement présents dans les données normatives du WISC-III (le manuel ne fournit pas de précision à ce sujet). En revanche, les études menées auprès d'enfants qui diffèrent de la norme ne présentent pas ce même équilibre dans la direction des écarts QIV-QIP. En général, les cotes obtenues aux différentes échelles de Wechsler laissent entrevoir des patrons d'écarts favorisant l'échelle non-verbale chez les enfants bilingues (hispanophones ou autres aux États-Unis), les enfants de milieux minoritaires et les enfants exceptionnels (Altus, 1953; Christiansen & Livermore, 1970; Clarizio, 1982; Dean, 1979, 1980; Figueroa, 1989; McCullough et al., 1985; McShane & Plas, 1984; Murphy, 1990; Slate, 1995b; Swanson & Deblassie, 1979; Swerdlik, 1978; Tanner-Halverson et al., 1993; Taylor & Richards, 1991; Taylor et al., 1984).

Profils des populations minoritaires. Les données tirées de la norme américaine indiquent que la fréquence des patrons QIV > QIP est équivalente à la fréquence des patrons QIP > QIV et que l'amplitude des différences moyennes entre ces deux échelles est d'environ 10 points (Kaufman, 1976; Wechsler, 1991). Toutes les populations ne correspondant pas à ces caractéristiques sont probablement différentes de la norme.

Les profils d'intelligence des jeunes hispano-américains aux échelles de Wechsler ont été comparés à plusieurs reprises aux profils des enfants américains. Murphy (1990) compare les cotes de QI de 20 enfants bilingues (Hispano-américains) à ceux de 20 enfants unilingues (Américains) au WISC-R. Les cotes de QI des enfants bilingues s'élèvent à 119 au QIT, 107,8 au QIV et 126,6 au QIP. Chez les enfants unilingues, les moyennes obtenues sont de 129,6 au QIT, 125,9 au QIV et 126,2 au QIP. Les deux groupes d'enfants obtiennent des cotes situées autour ou au-dessus de la moyenne. Lorsque les moyennes de QI sont comparées, une différence importante est toutefois décelée au niveau de leurs habiletés verbales. En effet, la moyenne du QIV est significativement plus élevée chez les enfants américains que chez les enfants hispano-américains; aucune différence de ce genre n'est notée entre leurs cotes d'intelligence non-verbale. En plus de cette distinction entre leur performance verbale, les deux groupes obtiennent

des patrons d'écart QIV-QIP divergents. Tandis que les moyennes d'habiletés verbales et d'habiletés non-verbales des jeunes Américains sont identiques, chez les Hispano-américains, les résultats au QIP sont significativement supérieurs à ceux du QIV, avec une différence moyenne de 18,8 points.

Une comparaison des écarts entre les échelles verbale et non-verbale du WISC-R chez trois groupes de différentes ethnies, indique que seuls les Hispano-américains (n = 184) obtiennent une fréquence plus élevée de patrons QIP > QIV que de patrons QIV > QIP (Taylor et al., 1984). Chez les jeunes Américains de race blanche (n = 184) et de race noire (n = 187), la prépondérance des échelles est équivalente. De plus, la moyenne des écarts QIV-QIP est davantage élevée auprès des Hispano-américains (\bar{M} = 13,72) qu'elle ne l'est auprès des Américains de race blanche (\bar{M} = 9,63) ou de race noire (\bar{M} = 8,91). Des résultats semblables sont obtenus chez un groupe d'enfants référés de cultures diverses. En effet, Dean (1979) a comparé les écarts QIV-QIP d'enfants référés de cultures américaine et hispano-américaine à l'aide du WISC-R (n = 120). Non seulement ses analyses des données indiquent que les moyennes des échelles QIV diffèrent entre les deux groupes (\bar{M} américaine = 91,78; \bar{M} hispano-américaine = 85,23), mais des différences significatives sont aussi présentes entre les échelles verbale et non-verbale des Hispano-américains seulement. En effet, chez ce groupe

d'enfants, le QIV donne des résultats inférieurs au QIP; aucune différence significative n'est signalée entre le QIV et le QIP auprès des Américains.

Des études additionnelles ont aussi signalé que les cotes d'intelligence non-verbale au WISC-R sont généralement supérieures aux cotes d'intelligence verbale auprès de divers échantillons d'Hispano-américains (Dean, 1980; Swanson & Deblassie, 1979; Swerdlik, 1978). L'observation des patrons des sous-tests au WISC-R appuie d'ailleurs ces faits auprès d'enfants de différentes ethnies. Taylor et Richards (1991) ont effectivement trouvé que les patrons des sous-tests au WISC-R varient en fonction de l'appartenance culturelle d'un individu. Dans une étude menée auprès de 300 enfants d'ethnies diverses, ces auteurs ont d'abord découvert que, lorsque la cote d'intelligence globale est maintenue constante parmi les sujets, les Hispano-américains obtiennent des cotes non-verbales supérieures à leurs cotes verbales. Les Américains de race noire obtiennent des résultats verbaux supérieurs comparativement à leurs cotes non-verbales. L'étude indique que les jeunes hispanophones sont ceux qui obtiennent la meilleure performance aux tâches visuo-spatiales du WISC-R, tandis que les Américains de race noire accusent une meilleure performance aux sous-tests verbaux. Par ailleurs, les Américains de race blanche parviennent à une meilleure performance auprès des tâches

mesurant la pensée abstraite ainsi que la connaissance des faits.

Les enfants hispano-américains ont généralement un QIV inférieur au QIP aux échelles de Wechsler. Chez des Hispanophones surdoués (n = 276), comparativement aux Américains surdoués de race blanche et de race noire, les résultats au QIV moyen du WISC-R sont équivalents à ceux du QIP moyen, tout comme auprès des Philippins surdoués (Saccuzzo, Johnson, & Russell, 1992). Chez les enfants hispano-américains de cette étude, la différence moyenne calculée entre l'échelle verbale et l'échelle non-verbale est relativement faible et est non significative. En effet, seulement 2,37 points séparent les deux échelles en faveur de l'échelle verbale. Toutefois, la faible étendue des cotes de QI des enfants surdoués pourrait expliquer la faible différence calculée entre le QIV et le QIP.

Le profil d'intelligence favorisant l'échelle non-verbale des tests de Wechsler n'est pas unique aux enfants de culture hispano-américaine. Des patrons d'écart semblables ont aussi été notés chez diverses populations amérindiennes aux États-Unis, en plus d'y retrouver des écarts importants entre les cotes d'habiletés verbales et d'habiletés non-verbales. Une étude de McCullough et al. (1985) soulève d'ailleurs l'importance de l'amplitude des écarts QIV-QIP chez des enfants amérindiens de classes régulières. Les différences moyennes calculées entre les

échelles d'intelligence verbales et non-verbales au WISC-R et au WAIS (échelle pour adultes de Wechsler) s'élève respectivement à 19 points et à 16 points auprès de cet échantillon. Quoique ces différences sont importantes, elles semblent typiques de l'échantillon étudié. En effet, McShane et Plas (1984) stipulent que le fonctionnement cognitif des amérindiens comporte normalement des écarts situés entre 8 et 19 points en faveur de l'échelle QIP au WISC-R, lorsque les normes américaines sont utilisées.

L'examen des profils d'intelligence d'enfants américains fait ressortir une fréquence élevée des patrons d'écarts favorisant l'échelle d'intelligence non-verbale (McCullough et al., 1985). Notamment, 72 des 75 enfants de cet échantillon obtiennent un patron d'écarts QIP > QIV au WISC-R ou au WAIS. Les indices de corrélations calculés entre le QIV et le QIP de ces tests sont non significatifs, $r = -0,17$ au WISC-R et $r = 0,02$ au WAIS. Puisque les écarts retrouvés entre les échelles des tests sont importants et que leurs corrélations sont faibles, les auteurs ont recommandé de ne pas interpréter le QIT auprès de cet échantillon d'étude. Au WISC-III, des patrons d'écarts semblables sont comparables à ceux trouvés par McCullough et al. chez des enfants amérindiens référés (Tanner-Halverson et al., 1993). L'écart moyen de 7,4 points calculé entre le QIP et le QIV des 110 sujets de l'étude est en faveur de l'échelle non-verbale. Des normes locales ont ici été

développées pour le WISC-III afin d'éviter l'utilisation des normes américaines auprès des jeunes amérindiens de cette population.

Des difficultés surgissent lors de l'application des règles de la norme américaine aux échelles de Wechsler pour interpréter les profils d'intelligence de populations minoritaires, particulièrement chez les enfants bilingues de culture hispanique ou de culture amérindienne aux États-Unis. Lorsque les patrons d'écart de ces groupes minoritaires sont examinés, une difficulté particulière est notée au niveau de l'échelle d'intelligence verbale. En effet, chez des enfants amérindiens entre autres, il a été noté que parfois, les cotes du QIV sont inférieures à la moyenne normative américaine ($M = 100$) pendant que celles du QIP y sont semblables (McCullough et al., 1985; McShane & Plas, 1984; Tanner-Halverson et al., 1993). Dean (1979), Naglieri (1982), Darou (1992), Flanagan (1995) et Kaufman (1994) estiment d'ailleurs que pour les enfants de culture amérindienne ou de culture hispanique, l'échelle verbale des tests de Wechsler évalue des habiletés linguistiques de langue anglaise, et non un indice d'intelligence verbale.

Les conclusions tirées des recherches abordant les écarts QIV-QIP des groupes minoritaires sont une preuve directe de la critique de Kaufman (1993) au sujet de l'échelle verbale du WISC-III. En effet, la mesure des tâches cognitives à l'échelle verbale semble pénaliser les

enfants bilingues et les enfants de milieux minoritaires. À ces groupes d'enfants américains ciblés par Kaufman pourrait de plus s'ajouter la population d'enfants francophones canadiens (Quintin, 1985).

Les enfants canadiens semblent aussi faire partie du groupe des populations minoritaires qui diffèrent de la norme américaine. D'ailleurs, des normes canadiennes-anglaises ont récemment été publiées pour le WISC-III (Wechsler, 1996). Cependant, les enfants francophones de cette population ne sont pas représentés à l'intérieur de l'échantillon normatif canadien. Les jeunes canadiens-français démontrent toutefois des profils distincts d'intelligence verbale et d'intelligence non-verbale. Quintin (1985) a déjà noté que les normes américaines des échelles de Wechsler ne rendent pas justice aux enfants canadien-français. Les données de son étude ont été recueillies à l'aide de l'échelle pour enfants d'âge pré-scolaire de Wechsler (WPPSI) chez 100 enfants québécois. Les indices moyens d'intelligence de cet échantillon sont de 102 au QIT, 98 au QIV et 106 au QIP. Un écart moyen de 8 points démarque les résultats de l'échelle verbale de ceux de l'échelle non-verbale, avec un patron d'écart $QIP > QIV$. Ce patron est d'ailleurs rencontré chez 81% des enfants de cet échantillon. Quoique 37% des sujets produisent ici des écarts $QIP > QIV$ de moins de 10 points, 44% obtiennent des écarts $QIP > QIV$ situés entre 10 et 34 points.

Quintin (1985) signale que l'interprétation clinique des écarts QIV-QIP effectuée auprès d'une population d'enfants francophones canadiens devrait être faite avec prudence. En effet, un patron d'écarts QIP > QIV semble "normal" auprès de cette population lorsque des adaptations françaises des échelles de Wechsler sont utilisées conjointement avec les normes américaines. Cette mise en garde de Quintin redouble d'importance sachant que les tests d'intelligence sont des outils diagnostiques grandement privilégiés pour l'évaluation de certaines populations canadiennes-françaises (Gauthier, 1988). De plus, Figueroa (1989) déconseille d'administrer des tests d'intelligence traduits littéralement. Il explique qu'une telle pratique peut occasionner des problèmes puisqu'il est difficile d'assurer une correspondance des propriétés psychométriques d'un test traduit comparativement à sa version originale. Par exemple, il indique que l'ordre de difficulté des items peut changer parce que l'usage de certains mots varie en fonction de la langue.

Synthèse

La présente étude tente d'évaluer si les tests d'intelligence normalisés aux États-Unis rendent justice aux enfants qui ne sont pas représentés à l'intérieur des échantillons de standardisation. La comparaison des profils d'intelligence verbale et non-verbale au WISC-III et au

Stanford-Binet IV est effectuée auprès d'enfants francophones du Nouveau-Brunswick.

La performance de certaines populations minoritaires aux États-Unis à des tests d'intelligence diffère de la performance des individus de la norme. D'abord, tandis que la comparaison du WISC-III et du Stanford-Binet IV chez des enfants non exceptionnels indique la présence de cotes semblables d'un test à l'autre (Carvajal et al., 1993; Lavin, 1996), auprès de divers échantillons d'enfants exceptionnels, les cotes du Stanford-Binet IV sont supérieures à celles du WISC-III (Lukens & Hurrell, 1996; Prewett & Matavich, 1994; Saklofske et al., 1994). L'examen des différences calculées entre les QI du WISC-III et du Stanford-Binet IV rend compte d'un écart particulièrement important entre les échelles verbales des deux tests.

Une autre divergence importante notée aux échelles de Wechsler entre les cotes de la norme américaine et celles de certaines populations minoritaires, se trouve au niveau des profils d'intelligence verbale et non-verbale de chaque population. Auprès de la norme américaine, un écart d'environ 10 points est fréquemment retrouvé entre les cotes d'intelligence verbale et non-verbale. De plus, il y a autant de profils où l'échelle non-verbale est supérieure à l'échelle verbale que de profils où l'échelle non-verbale est inférieure à l'échelle verbale (Kaufman, 1976; Wechsler, 1991). Quoique les écarts moyens soient aussi élevés auprès

des individus de populations minoritaires, aucune équivalence de la prépondérance des échelles n'est trouvée entre autres chez les groupes d'enfants hispanophones et d'enfants amérindiens aux États-Unis. D'ailleurs, les profils d'intelligence de ces groupes sont généralement unidirectionnels (non-verbal > verbal) (Altus, 1953; Christiansen & Livermore, 1970; Clarizio, 1982; Dean, 1979, 1980; Figueroa, 1989; McCullough et al., 1985; McShane & Plas, 1984; Murphy, 1990; Slate, 1995b; Swanson & Deblassie, 1979; Swerdlik, 1978; Tanner-Halverson et al., 1993; Taylor et al., 1984).

La règle normative prévoyant des écarts verbal-non-verbal relativement élevés et retrouvés également dans les deux directions (autant d'écarts QIP > QIV que d'écarts QIV > QIP) semble davantage inapplicable chez les enfants québécois. Quintin (1985) a en effet noté qu'au WPPSI, les jeunes canadiens-français obtiennent des patrons d'écarts QIP > QIV à une fréquence élevée.

L'interprétation des écarts significatifs calculés entre le QIV et le QIP ne peut suivre les règles de la norme, suite à l'évaluation cognitive d'enfants de groupes minoritaires, puisque leur profil d'écarts ne correspond pas à celui de la norme. Certains auteurs avisent d'ailleurs de ne pas procéder à l'interprétation du QIT chez ces populations (McCullough et al., 1985; Tanner-Halverson et al., 1993; Quintin, 1985), en particulier au niveau de

l'échelle d'intelligence verbale (Dean, 1979; Naglieri, 1982; Flanagan, 1995; Kaufman, 1994). À la lumière des postulats présentés dans la littérature, il semble donc nécessaire de pondérer l'utilisation des règles de la norme américaine auprès d'une population d'enfants francophones du Nouveau-Brunswick, en comparant les profils d'intelligence du WISC-III à ceux du Stanford-Binet IV.

Objectifs et hypothèses

Plusieurs objectifs guident la présente étude. D'abord, un premier objectif consiste en l'établissement de comparaisons empiriques entre les cotes standard et les QI du WISC-III et du Stanford-Binet IV auprès d'enfants francophones du Nouveau-Brunswick. Étant donné la popularité de ces deux instruments de mesure, il semble important de découvrir empiriquement les différences et les concordances retrouvées entre le WISC-III et le Stanford-Binet IV chez cet échantillon.

Le deuxième objectif d'étude vise à produire des données empiriques sur les écarts entre les mesures d'intelligence verbale et non-verbale au WISC-III et au Stanford-Binet IV. Par conséquent, un examen de la direction et de la fréquence des écarts auprès des enfants de l'étude est poursuivi. La cueillette de données est particulièrement importante, compte tenu de la valeur clinique accordée à l'interprétation des écarts QIV-QIP au WISC-III et de la carence de recherches auprès de la population cible.

Le troisième objectif est de nature pratique. La présente étude vise la communication des données aux psychologues scolaires francophones de la province en guise d'informations utiles et nécessaires à leur pratique.

Quatre hypothèses principales sont émises. Parmi ces hypothèses, les deux premières découlent des énoncés théoriques des tests (Bracken, 1988; Flynn, 1987; Wechsler, 1991). Les deux dernières hypothèses sont suggérées par les conclusions d'études tirées de la littérature. Elles concernent les écarts entre les moyennes des cotes verbales et non-verbales des deux tests.

Une première hypothèse stipule que des corrélations positives et significatives seront retrouvées entre les cotes du WISC-III et du Stanford-Binet IV. Sachant que les deux tests sont une mesure quantitative de l'intelligence, il est donc prévu de retrouver des corrélations positives. De plus, la corrélation retrouvée entre les indices d'intelligence globale des deux tests (QIT et TC) devrait s'avérer la plus élevée. En effet, parce que ces derniers indices représentent tous deux l'ensemble des tests, ils présenteront la corrélation la plus élevée de toutes. Par ailleurs, les échelles verbales des deux tests seront plus hautement corrélées l'une à l'autre qu'elles ne le seront avec l'échelle non-verbale de l'autre test, et vice-versa.

La deuxième hypothèse prévoit que les résultats moyens du WISC-III soient significativement inférieurs à ceux du

Stanford-Binet IV, puisque ce dernier test comprend des normes plus anciennes que le précédent. De plus, étant donné que l'échantillon de l'étude diffère de la population normative américaine (Kaufman, 1993), ces différences seront particulièrement prononcées entre les moyennes des échelles d'habiletés verbales (Lukens & Hurrell, 1996; Prewett & Matavich, 1994; Saklofske et al., 1994).

La troisième hypothèse propose que des différences significatives seront retrouvées entre les résultats des échelles verbale et non-verbale du WISC-III, mais pas au Stanford-Binet IV. En effet, chez des enfants qui se distinguent de la norme américaine, le profil des habiletés cognitives offert par le WISC-III est différent de celui du Stanford-Binet IV (Prewett & Matavich, 1994).

La quatrième et dernière hypothèse définit pour sa part la direction et la fréquence de ces écarts significatifs. Elle stipule qu'au WISC-III, la fréquence du patron d'écarts QIP > QIV sera beaucoup plus élevée que celle du patron d'écarts QIV > QIP auprès de l'échantillon étudié. Il a été démontré à plusieurs reprises que chez des populations minoritaires, les résultats obtenus à l'échelle non-verbale sont significativement plus élevés que ceux retrouvés à l'échelle verbale (Altus, 1953; Christiansen & Livermore, 1970; Clarizio, 1982; Dean, 1979, 1980; Figueroa, 1989; McCullough et al., 1985; McShane & Plas, 1984; Murphy, 1990; Slate, 1995b; Swanson & Deblasse, 1979; Swerdlik, 1978;

Tanner-Halverson et al., 1993; Taylor et al., 1984). Étant donné l'étude de Quintin (1985), il faudrait s'attendre à observer une majorité d'individus ayant un profil QIP > QIV dans notre échantillon d'enfants francophones.

CHAPITRE DEUXIÈME
DESCRIPTION DE L'ÉTUDE

Méthode

Échantillon

L'échantillon compte un total de 96 enfants francophones recrutés à l'intérieur de huit écoles élémentaires des Districts scolaires no.1 (région Moncton) et no.11 (région Bouctouche-Miramichi) du Nouveau-Brunswick. Parmi les 100 enfants rencontrés, quatre d'entre eux ont été éliminés de l'échantillon, suite à des évaluations incomplètes. L'échantillon d'analyse regroupe 42 garçons et 54 filles âgés entre 7,3 ans et 13,2 ans. Lors de l'administration du premier test, la moyenne d'âge des sujets était de 9,9 ans. Aucune information n'a été recueillie sur les enfants présentant des troubles quelconques du développement ou de l'apprentissage.

Les participants à l'étude sont pigés au hasard parmi les volontaires ayant fourni l'approbation écrite de leurs parents (n = 289) (voir copie de la lettre en annexe A). L'échantillon de l'étude se compose d'élèves suivant les classes régulières de la deuxième année (28,21%), de la quatrième année (40,6%) et de la sixième année (31,3%).

Instruments de mesure

WISC-III. Le Wechsler Intelligence Scale for Children-Third Edition (Wechsler, 1991) est un instrument de mesure valide et fidèle qui s'administre auprès des enfants de 6 à

16 ans. Son administration est d'une durée variant de 60 à 85 minutes. L'échelle compte au total treize sous-tests qui ont une cote standard moyenne de 10 avec un écart-type de 3. Le regroupement de ces sous-tests donne accès à trois échelles principales de QI et à quatre index factoriels optionnels. Les échelles et les facteurs du WISC-III offrent des indices de QI dont la moyenne est de 100 et l'écart-type de 15.

La consistance interne moyenne des échelles de QIT, QIV et QIP pour tous les groupes d'âges, s'élève respectivement à 0,96, 0,95 et 0,91. Les indices médians de consistance interne aux sous-tests du WISC-III sont par contre plus faibles et se situent entre 0,76 et 0,87. Par ailleurs, la moyenne des indices de consistance interne des quatre facteurs optionnels se retrouve entre 0,85 et 0,94.

Pour la présente étude, une adaptation canadienne-française du WISC-III (Maurice, 1992) est utilisée. Tous les dix sous-tests standard de l'échelle sont administrés, en plus de deux sous-tests supplémentaires, soit Recherche de symboles et Empan numérique. Le sous-test Labyrinthes n'est pas administré puisqu'il fournirait de l'information superflue, non comptabilisée dans le calcul des échelles d'intelligence ou des facteurs optionnels.

Stanford-Binet IV. Le Stanford-Binet Intelligence Scale: Fourth Edition (Thorndike et al., 1986a, 1986b) a été normalisé auprès d'un échantillon comptant 5013 sujets âgés

entre 2 et 23 ans. L'administration adaptative de l'échelle est d'une durée approximative de 60 à 90 minutes et comporte différents sous-tests choisis selon l'âge et les capacités cognitives des sujets. Tandis que les quinze sous-tests du Stanford-Binet IV donnent une cote standard moyenne de 50 avec un écart-type de 8, les cotes en âge standard (ou QI) des quatre secteurs et du Test composite de cette échelle, ont une moyenne de 100 avec un écart-type de 16.

En plus d'être un instrument de mesure valide, le Stanford-Binet IV est une mesure fiable de l'intelligence. La consistance interne médiane du TC est de 0,97, avec une étendue située entre 0,95 et 0,99 à travers les différents groupes d'âges. La consistance interne des sous-tests est par contre plus faible, variant entre des indices médians de 0,73 et 0,94 pour tous les groupes d'âges.

Une adaptation canadienne-française du Stanford-Binet IV (Chevrier, 1988) est utilisée lors de la présente étude. Parmi les quinze sous-tests de l'échelle, neuf d'entre eux sont sélectionnés comme administration standard pour tous les sujets de l'étude: Vocabulaire, Compréhension, Absurdités, Analyse de modèles, Matrices, Quantités, Série de nombres, Mémorisation de perles ainsi que Mémorisation de phrases.

Procédure

Suite à l'obtention du consentement écrit de leurs parents, les élèves sont choisis et rencontrés

individuellement à leur école respective. L'administration des tests est contrebalancée, avec 51% des sujets débutant l'évaluation à l'aide du Stanford-Binet IV et 49% d'entre eux débutant avec le WISC-III. Les tests sont administrés par trois examinateurs: l'auteure de la présente étude (n = 83), une étudiante du B.Ps. à l'Université de Moncton (n = 10) et un professeur du département de psychologie à l'Université de Moncton (n = 3). Tous les enfants rencontrent la même personne à deux reprises pour des périodes de 60 à 90 minutes chacune, durant lesquelles les tests sont administrés à tour de rôle. Un intervalle médian de 42 jours sépare l'administration des deux tests, l'étendue inter-test variant de 1 à 57 jours. Une récompense est remise à chaque enfant suite à l'administration du deuxième test. Les rencontres ont lieu entre mars et juin 1996. Des rapports globaux sur la présente étude sont acheminés aux Districts scolaires et aux écoles participantes (voir annexe B) ainsi qu'aux parents qui en ont manifesté le désir (voir annexe C).

Analyses statistiques

Les analyses statistiques de la présente étude sont effectuées à l'aide du programme informatique SPSS pour Unix (version 5,0). Précédant l'analyse des données, une transformation des cotes pour les quatre secteurs et le TC du Stanford-Binet IV, est réalisée dans le but de ramener ces résultats sur une échelle comparable à celle du

WISC-III. La moyenne de 100 avec un écart-type de 16 est transformée en moyenne de 100 avec un écart-type de 15, selon la formule figurant en l'annexe D.

Suivant cette transformation des données du Stanford-Binet IV, une analyse exploratoire définit les statistiques descriptives des données. Les indices de corrélations de Pearson entre les sous-tests et les échelles des deux tests sont calculés pour vérifier l'hypothèse de corrélations inter-tests.

Plusieurs analyses de la variance (ANOVAS) sont ensuite effectuées afin de pouvoir mesurer les effets de plusieurs variables sur les cotes principales des deux tests. D'abord, des ANOVAS à deux facteurs 2 (sexe: féminin, masculin) x 3 (année scolaire: deuxième année, quatrième année, sixième année) examinent les effets du sexe et de l'année scolaire sur les indices d'intelligence globale des deux tests (QIT et TC). Cette analyse tente d'établir la présence d'un biais de l'échantillon selon l'âge, le sexe ou l'interaction âge x sexe des sujets. Une ANOVA à mesures répétées à deux facteurs 2 (test: WISC-III, Stanford-Binet IV) x 3 (échelle: verbale, non-verbale, globale) examine ensuite les différences des moyennes des trois échelles principales (globale, non-verbale et verbale) selon le test utilisé. L'hypothèse de différences significatives entre les cotes moyennes des deux tests est vérifiée. Cette analyse aborde de plus l'hypothèse d'écart significatifs entre les cotes

d'habiletés verbale et d'habiletés non-verbales au WISC-III. Des analyses post hoc (HSD de Tukey) sont effectuées au besoin à la suite de ces ANOVAS, afin de situer les comparaisons individuelles significatives. Les écarts entre les cotes des échelles verbales et non-verbales des deux tests sont observés à l'aide de tableaux de fréquence et de statistiques descriptives. À partir de ceux-ci, la direction et la fréquence des profils d'intelligence sont observées.

CHAPITRE TROISIÈME
ANALYSE DES DONNÉES

Résultats

L'objectif global de la présente étude est de comparer les profils d'intelligence verbale et d'intelligence non-verbale des enfants francophones du Nouveau-Brunswick à ceux des enfants du groupe normatif aux États-Unis. Une cueillette de données est réalisée auprès d'un échantillon de la population cible afin de produire des données empiriques à ce sujet. Les données de l'étude sont analysées dans le présent chapitre.

En premier lieu, les statistiques descriptives des deux tests sont présentées, suivies d'une analyse de l'homogénéité de l'échantillon de l'étude selon le sexe et l'année scolaire des sujets. Ensuite, un bref examen des indices de corrélations obtenus entre les cotes des tests est effectué. Les analyses suivantes abordent plus directement les objectifs de l'étude. D'abord, l'observation empirique des différences entre les cotes de QI des deux tests (différences inter-tests) ainsi qu'entre les cotes de QI d'un même test (différences intra-tests) est réalisée. Cette analyse est suivie de l'examen des patrons d'écarts entre les cotes d'habiletés verbales et d'habiletés non-verbales des deux tests.

Statistiques descriptives

La répartition des sujets de l'échantillon de l'étude

selon le niveau scolaire, le sexe et l'école est présentée au tableau 1. De façon générale, la distribution des sujets varie d'une école à l'autre et d'un niveau scolaire à l'autre. Les élèves d'une école du District scolaire no. 11 (Carrefour Beausoleil) représentent à eux seuls plus du tiers des sujets de l'étude. Le groupe d'élèves de la quatrième année est sur-représenté comparativement aux deux autres sous-groupes scolaires. Néanmoins, malgré cette répartition inégale des sujets, la distribution des moyennes de QI s'apparente à celle d'une distribution normale, tant au WISC-III (voir figure 1), \underline{z} de Kolmogorov-Smirnov = 0,94, $p > 0,25$, qu'au Stanford-Binet IV (voir figure 2), \underline{z} de Kolmogorov-Smirnov = 0,67, $p > 0,25$.

Le tableau 2 présente les statistiques descriptives des cotes standard moyennes et des QI moyens pour le WISC-III. Au tableau 3 figurent les cotes standard moyennes et les QI moyens obtenus au Stanford-Binet IV. De façon générale, les moyennes de QI du WISC-III et du Stanford-Binet IV sont comparables et se situent à l'intérieur des limites de l'intelligence moyenne aux deux tests (100 ± 15).

Les cotes standard moyennes des 12 sous-tests du WISC-III sont contenues à l'intérieur des limites de la norme de ce test (10 ± 3). Il est en de même pour les résultats des 9 sous-tests du Stanford-Binet IV, se retrouvant eux aussi à l'intérieur des limites de la norme (50 ± 8). Au WISC-III, les sous-tests Vocabulaire et

Tableau 1
Répartition de l'échantillon

École	<u>n</u>	Année scolaire					
		2		4		6	
		sexe		sexe		sexe	
		F	M	F	M	F	M
Baie Ste-Anne	5	1	1	1	0	1	1
Camille Vautour	9	2	2	1	2	1	1
Carrefour Beausoleil	38	8	3	6	4	11	6
Champlain	13	0	0	6	7	0	0
Mont-Carmel	9	1	1	2	2	1	2
Notre-Dame	6	1	1	1	1	1	1
Saint-Paul	6	1	1	1	1	1	1
W.F. Boisvert	10	2	2	2	2	2	0
Total	96	16	11	20	19	18	12
	96	27		39		30	

Note. Toutes ces écoles sont situées dans le District scolaire no. 11, sauf l'école Champlain, dans le District scolaire no. 1.

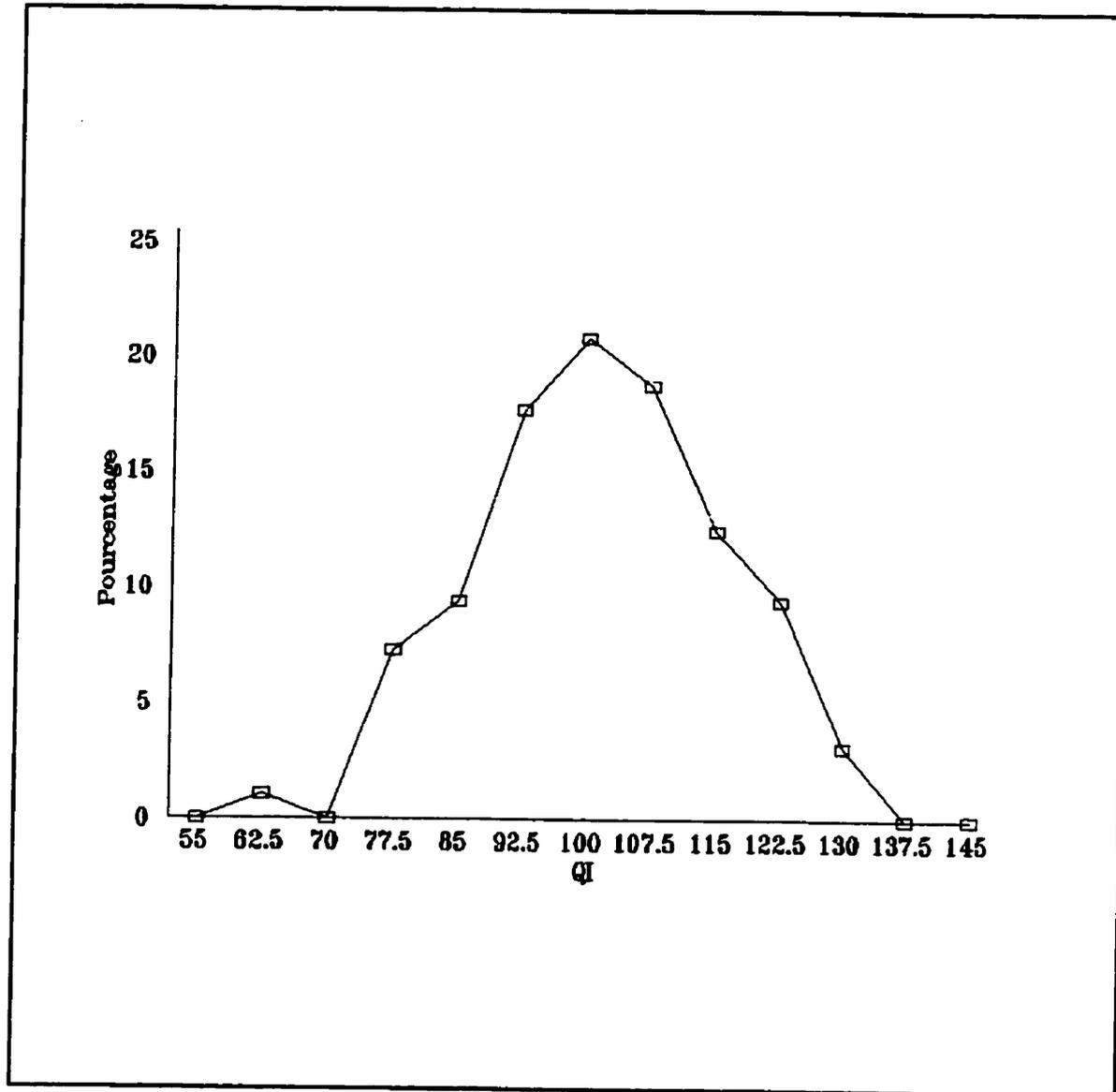


Figure 1. Distribution des cotes d'intelligence globale au WISC-III.

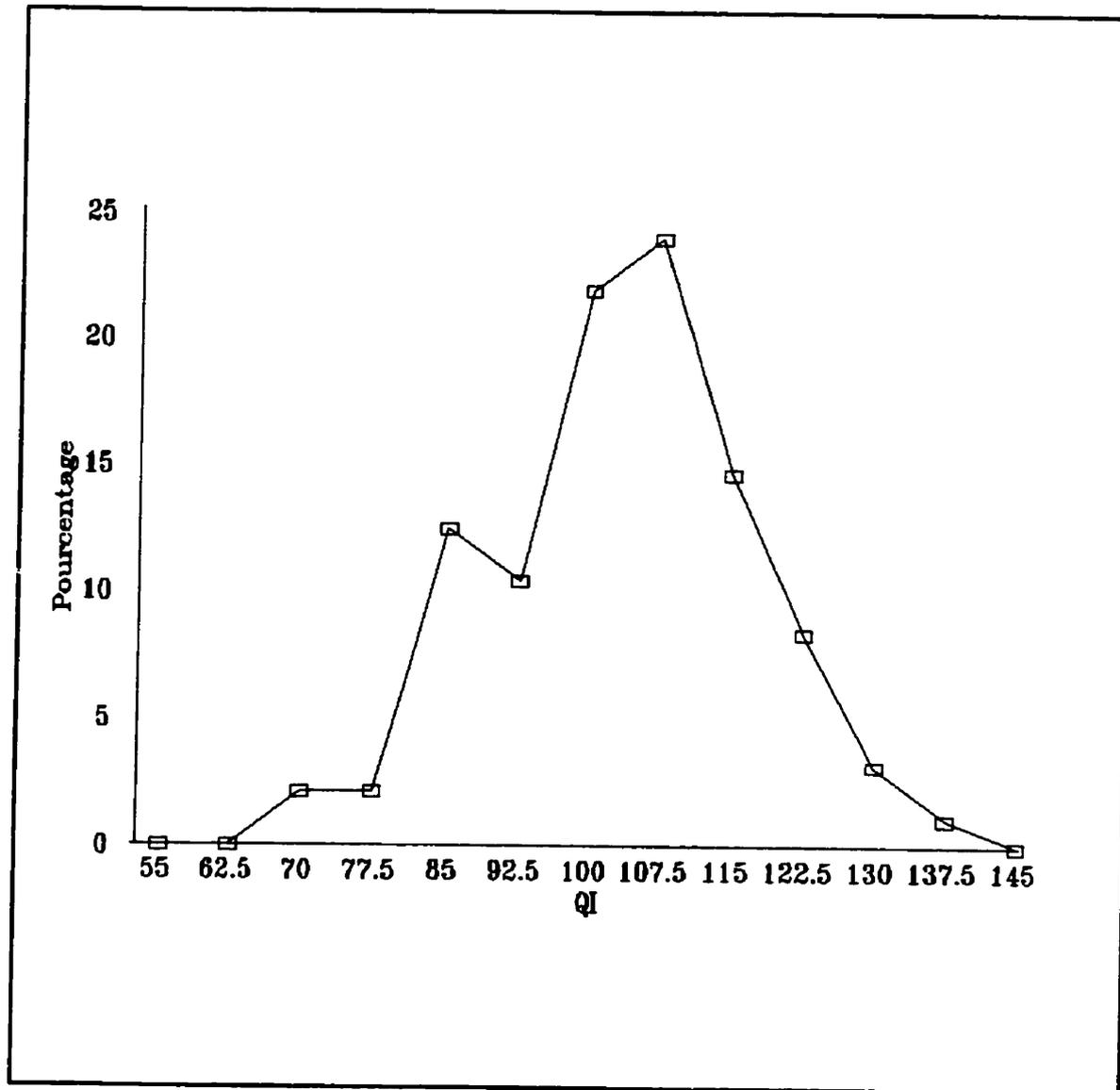


Figure 2. Distribution des cotes d'intelligence globale au Stanford-Binet IV (moyennes transformées).

Tableau 2
Statistiques descriptives du WISC-III

Test/échelle	<u>M</u>	<u>É.T.</u>	<u>Mdn.</u>	min.	max.
Ima	9,56	2,59	10,00	1	16
Con	8,33	2,22	8,00	3	14
Cod	11,58	3,06	12,00	4	18
Sim	8,42	3,29	9,00	1	18
Arr	9,83	3,61	10,00	1	19
Ari	9,69	3,03	9,00	2	17
Des	13,64	3,67	13,00	3	19
Voc	7,91	2,67	8,00	1	16
Ass	10,20	3,24	11,00	1	19
Jug	7,62	3,08	8,00	1	15
Rec	12,26	3,62	12,00	2	19
Emp	8,98	2,69	9,00	3	17
QIV	91,03	12,92	92,00	57	129
QIP	106,87	15,24	106,00	70	139
QIT	98,20	13,69	97,50	60	128
CV	89,93	12,45	91,00	60	125
OP	105,43	14,69	105,00	73	135
A/C	97,21	13,33	96,00	64	126
VP	110,92	15,02	112,50	70	143

Note. min.= valeur minimum; max.= valeur maximale; Ima= Images à compléter; Con= Connaissances; Cod= Codage; Sim= Similitudes; Arr= Arrangement d'images; Ari= Arithmétique; Des= Dessin avec blocs; Voc= Vocabulaire; Ass= Assemblages; Jug= Jugement; Rec= Recherche de symboles; Emp= Empan numérique; QIV= QI verbal; QIP= QI de performance; QIT= QI total; CV= Compréhension verbale; OP= Organisation perceptuelle; A/C= Attention-concentration; VP= Vitesse psychomotrice.

Tableau 3
Statistiques descriptives du Stanford-Binet IV^a

Test/échelle	<u>M</u>	<u>É.T.</u>	<u>Mdn.</u>	min.	max.
Voc	42,50	6,02	43,00	31	57
Com	45,89	7,95	47,00	30	66
Abs	50,16	6,66	50,00	33	67
Ana	55,96	7,99	57,00	30	72
Mat	54,21	6,74	54,00	32	73
Qua	49,47	7,58	50,00	30	76
Sér	52,84	12,08	54,00	35	69
Per	50,52	9,27	51,00	27	72
Phr	42,27	4,69	42,00	33	64
RV	92,57	11,71	93,50	65	124
RA/V	110,93	14,15	113,00	67	146
RM	104,06	14,63	102,50	63	145
MCT	92,35	13,70	92,00	57	132
TC	99,97	12,95	101,50	63	131

Note. min.= valeur minimum; max.= valeur maximale; Voc= Vocabulaire; Com= Compréhension; Abs= Absurdités; Ana= Analyse de modèles; Mat= Matrices; Qua= Quantités; Sér= Séries de nombres; Per= Mémorisation de perles; Phr= Mémorisation de phrases; RV= Raisonnement verbal; RA/V= Raisonnement abstrait-visuel; RM= Raisonnement mathématique; MCT= Mémoire à court terme; TC= Test composite.

^a moyennes transformées

Jugement, tous deux impliqués dans le calcul du QIV, atteignent les moyennes les plus faibles. De leur côté, Dessin avec blocs et Recherche de symboles ont les cotes standard les plus élevées. Les cotes standard moyennes les plus élevées au Stanford-Binet IV se situent aux sous-tests Analyse de modèles et Matrices, tous deux comptabilisés dans le calcul du RA/V. Par ailleurs, les résultats moyens des sous-tests Vocabulaire et Mémorisation de phrases sont les plus faibles. Ainsi, pour les deux instruments de mesure, les cotes standard les plus faibles sont calculées à l'intérieur de sous-tests à contenu verbal et les cotes standard les plus élevées, à l'intérieur de sous-tests mesurant entre autres l'intelligence non-verbale.

Analyse de l'échantillon de l'étude

Avant de procéder à des analyses subséquentes à l'aide de ces données, l'homogénéité de l'échantillon selon le sexe et l'année scolaire des sujets doit être établie. Les résultats d'une ANOVA à deux facteurs 2 (sexe: féminin, masculin) x 3 (année scolaire: deuxième année, quatrième année, sixième année) comparant les moyennes du QIT au WISC-III, indiquent la présence d'effets significatifs de l'année scolaire sur ces moyennes, $F(2, 90) = 4,21$, $p < 0,05$. Aucun effet de la variable sexe ni de l'interaction des variables sexe x année scolaire sur les moyennes de QIT n'est rapporté, respectivement

$F(1, 90) = 0,002, p > 0,05$ et $F(2, 90) = 0,52, p > 0,05$.
Lorsqu'une ANOVA similaire est effectuée avec les moyennes du TC au Stanford-Binet IV, des effets significatifs de l'année scolaire sont aussi notés, $F(2, 90) = 9,45, p < 0,001$. Tout comme au WISC-III, la variable sexe et l'interaction sexe x année scolaire n'ont pas d'effet significatif sur les moyennes du TC, respectivement $F(1, 90) = 1,21, p > 0,05$ et $F(2, 90) = 0,69, p > 0,05$.

Un examen des moyennes de QIT et de TC révèle que les élèves de la quatrième année obtiennent des résultats supérieurs à ceux de deuxième et de sixième années (voir tableau 4). Des analyses post hoc (HSD de Tukey) indiquent d'ailleurs que les enfants du groupe de la quatrième année obtiennent des moyennes supérieures aux deux autres sous-groupes d'élèves. Le tableau 1 ci-haut montre qu'il y a sur-représentation des élèves de la quatrième année dans l'échantillon total, avec une sur-représentation particulière des élèves du District scolaire no. 1.

Afin de savoir si les enfants du District scolaire no. 1 sont la source de ce biais, la comparaison des moyennes des deux tests est effectuée à l'aide d'une ANOVA à mesures répétées 2 (sexe: féminin, masculin) x 3 (année scolaire: deuxième année, quatrième année, sixième année) x 2 (test: WISC-III, Stanford-Binet IV) auprès des élèves du District scolaire no. 11 seulement ($n = 83$). Le tableau 5 présente les moyennes de ces élèves au QIT et au TC. Aucun effet

Tableau 4

Moyennes d'intelligence globale au WISC-III et au Stanford-Binet IV selon le sexe et l'année scolaire des sujets

Groupes	<u>M</u>	<u>É.T.</u>	<u>n</u>
WISC-III			
Garçons			
Deuxième année	94,45	12,25	11
Quatrième année	104,37	13,08	19
Sixième année	93,50	17,58	12
Filles			
Deuxième année	93,94	10,63	16
Quatrième année	101,45	14,14	20
Sixième année	97,28	12,53	18
Stanford-Binet IV ^a			
Garçons			
Deuxième année	96,08	8,28	11
Quatrième année	104,74	10,80	19
Sixième année	92,03	13,78	12
Filles			
Deuxième année	94,90	12,12	16
Quatrième année	107,78	12,31	20
Sixième année	98,44	13,18	18

^a moyennes transformées

Tableau 5

Moyennes d'intelligence globale au WISC-III et au Stanford-Binet IV selon le sexe et l'année scolaire des sujets du District scolaire no. 11

Groupes	<u>M</u>	<u>É.T.</u>	<u>n</u>
WISC-III			
Garçons			
Deuxième année	94,45	12,25	11
Quatrième année	99,97	10,16	12
Sixième année	93,50	17,58	12
Filles			
Deuxième année	93,94	10,63	16
Quatrième année	96,79	13,53	14
Sixième année	97,28	12,53	18
Stanford-Binet IV ^a			
Garçons			
Deuxième année	96,08	8,28	11
Quatrième année	102,73	8,30	12
Sixième année	92,03	13,78	12
Filles			
Deuxième année	94,90	12,12	16
Quatrième année	104,49	12,92	14
Sixième année	98,44	13,18	18

^a moyennes transformées

significatif de l'année scolaire n'est retrouvé sur les moyennes d'intelligence globale auprès de ce sous-groupe, $F(2, 77) = 2,14, p > 0,05$, la même chose étant noté selon le sexe des sujets, $F(1, 77) = 0,16, p > 0,05$. Cependant, la variable test a des effets significatifs sur les résultats obtenus, $F(1, 77) = 6,76, p < 0,05$. De plus, il y a une interaction significative des variables année scolaire x test, $F(2, 77) = 4,00, p < 0,05$, où les élèves de la quatrième année obtiennent des cotes globales plus élevées au Stanford-Binet IV (voir tableau 5). Aucun effet significatif des interactions entre les variables sexe x test, $F(1, 77) = 1,47, p > 0,05$, les variables sexe x année scolaire, $F(2, 77) = 0,57, p > 0,05$, ainsi que les variables sexe x année scolaire x test, $F(2, 77) = 0,89, p > 0,05$, n'est trouvé.

Quoique les moyennes de QI des élèves de la quatrième année du District scolaire no. 1 ont tendance à faire augmenter la moyenne globale de l'échantillon total, cette tendance semble déjà présente en l'absence de ces élèves. Les effets de l'interaction année scolaire x test auprès des élèves du District scolaire no. 11 sont d'ailleurs un exemple de cette tendance. Les analyses subséquentes sont donc effectuées auprès de l'échantillon total ($N = 96$), tout en tenant compte du biais d'échantillonnage rencontré à l'intérieur du groupe de quatrième année.

Vérification des hypothèses

Corrélations inter-tests. Les corrélations calculées entre les moyennes de QI du WISC-III et du Stanford-Binet IV sont rapportées au tableau 6. Toutes les corrélations calculées sont positives et statistiquement significatives. L'indice de corrélation le plus élevé entre les deux tests est rapporté entre le QIT et le TC. Le facteur Vitesse psychomotrice (VP) du WISC-III, de son côté, est le facteur qui correspond le moins aux autres facteurs des deux tests, particulièrement avec les composantes verbales QIV, RV et Compréhension verbale (CV) où les corrélations sont les plus faibles.

L'indice de corrélation calculé entre le QIT et le TC appuie la mesure de construits globaux semblables au WISC-III et au Stanford-Binet IV auprès d'un échantillon francophone néo-brunswickois. Par ailleurs, les échelles verbales des tests sont plus hautement corrélées les unes aux autres qu'elles ne le sont avec les échelles non-verbales, et vice-versa. Ainsi le QIV du WISC-III est plus hautement corrélé avec le RV qu'avec le RA/V du Stanford-Binet IV, la même chose étant notée avec le CV. Le QIP de son côté est plus hautement corrélé avec le RA/V qu'avec le RV, tout comme l'est aussi l'index Organisation perceptuelle (OP).

Tableau 6

Corrélations entre cotes de QI du WISC-III et du Stanford-Binet IV*

Échelles	1	2	3	4	5	6	7	8	9	10	11	12
WISC-III												
1.QIV	---											
2.QIP	0,61	---										
3.QIT	0,89	0,90	---									
4.CV	0,97	0,57	0,86	---								
5.OP	0,61	0,97	0,88	0,57	---							
6.A/C	0,69	0,59	0,72	0,59	0,57	---						
7.VP	0,38	0,72	0,61	0,34	0,59	0,54	---					
Stanford-Binet IV												
8.RV	0,70	0,57	0,70	0,70	0,60	0,43	0,30	---				
9.RA/V	0,64	0,71	0,75	0,60	0,71	0,63	0,53	0,56	---			
10.RM	0,67	0,61	0,72	0,63	0,61	0,56	0,44	0,64	0,71	---		
11.MCT	0,60	0,53	0,62	0,56	0,53	0,60	0,46	0,54	0,65	0,62	---	
12.TC	0,77	0,71	0,82	0,73	0,72	0,66	0,52	0,79	0,87	0,88	0,83	---

Note. Toutes corrélations sont significatives, $r = .254$, $p < .01$. QIV= QI verbal; QIP= QI de performance; QIT= QI total; CV= Compréhension verbale; OP= Organisation perceptuelle; A/C= Attention-concentration; VP= Vitesse psychomotrice; RV= Raisonnement verbal; RA/V= Raisonnement abstrait-visuel; RM= Raisonnement mathématique; MCT= Mémoire à court terme; TC= Test composite.

* moyennes transformées

Ces corrélations supposent que le WISC-III et le Stanford-Binet IV possèdent des mesures d'intelligence verbale qui correspondent à un construit semblable. Cette même constatation s'applique aussi à leurs mesures d'habiletés non-verbales de l'intelligence. De plus, les indices de corrélation calculés entre les échelles verbales des tests sont à peu près équivalents aux corrélations entre les échelles non-verbales des tests. Enfin, la corrélation de chaque échelle d'intelligence verbale et de chaque échelle d'intelligence non-verbale avec l'indice de QI global du test opposé est soit équivalente ou soit supérieure aux corrélations calculées avec leurs échelles correspondantes.

Comparaisons des cotes de QI. Une ANOVA à mesures répétées (variables intra-sujets seulement) 2 (test: WISC-III, Stanford-Binet IV) x 3 (échelle: verbale, non-verbale, globale) est calculée afin de pouvoir discerner si les variables test et échelle ont des effets sur les cotes moyennes des tests. Les résultats de cette analyse indiquent effectivement la présence d'effets du test, $F(1, 95) = 11,15, p = 0,001$, et de l'échelle, $F(2, 190) = 269,28, p < 0,001$, sur les moyennes des trois échelles principales des tests (verbale, non-verbale et globale) (voir tableau 7). Aucun effet de l'interaction des variables test x échelle n'est trouvé, $F(2, 190) = 2,94, p > 0,05$. Ces résultats indiquent qu'il existe des

Tableau 7

Moyennes d'intelligence globale, verbale et non-verbale au WISC-III et au Stanford-Binet IV*

Échelle	<u>M</u>	<u>É.T.</u>	Intervalle de confiance	
WISC-III				
QIT	98,20	13,69	95,42	100,97
QIV	91,03	12,92	88,41	93,65
QIP	106,87	15,24	103,79	109,96
Stanford-Binet IV				
TC	99,97	12,95	97,34	102,59
RV	92,57	11,71	90,20	94,94
RA/V	110,93	14,14	108,06	113,79

Note. L'intervalle de confiance est calculée à 95%.

• moyennes transformées

différences significatives entre les moyennes de QI des tests, tant au niveau des comparaisons inter-tests qu'au niveau intra-test.

L'hypothèse de différences significatives entre les moyennes d'intelligence des tests n'est que partiellement confirmée. Le Stanford-Binet IV donne généralement des résultats supérieurs au WISC-III mais cette tendance n'est significative qu'au niveau non-verbal. En effet, des analyses post hoc effectuées à l'aide du HSD de Tukey révèlent que des effets du test sont retrouvés auprès de l'échantillon étudié mais uniquement entre les QI non-verbaux des tests. L'écart de 4,06 points calculés entre le QIP et le RA/V des tests est donc le seul à être significatif. Les données du tableau 7 laissent entrevoir que la cote moyenne du RA/V au Stanford-Binet IV est supérieure au QIP moyen du WISC-III, tout comme l'hypothèse l'avait prédit. Par contre, les différences entre les moyennes du QIT et du TC (1,77 point) ainsi qu'entre celles du QIV et du RV (1,54 point) sont non significatives.

Les analyses post hoc (HSD de Tukey) indiquent également que les écarts calculés entre les échelles d'intelligence d'un même test sont tous significatifs, illustrant ainsi les effets de l'échelle sur les moyennes de QI des tests. Ces résultats confirment l'hypothèse d'écarts significatifs entre les cotes d'intelligence verbale et non-verbale au WISC-III. Contrairement aux attentes de départ

par contre, ces résultats présentent aussi des différences significatives calculées entre le RV et le RA/V du Stanford-Binet IV. Au WISC-III, un écart significatif élevé de 15,84 points sépare les moyennes du QIV et du QIP; au Stanford-Binet IV, un écart significatif de 18,36 points est noté entre les moyennes du RV et du RA/V. En plus de retrouver des écarts significatifs entre les cotes d'intelligence verbale et les cotes d'intelligence non-verbale aux deux tests, il y a aussi présence d'écarts significatifs entre les cotes d'intelligence globale et verbale ainsi qu'entre les cotes d'intelligence globale et non-verbale aux deux tests. Les écarts entre les composantes non-verbales et globales des tests se situent respectivement à 8,67 points au WISC-III (QIP et TC) et à 10,96 points au Stanford-Binet IV (RA/V et TC). De plus, les différences de 7,17 points et de 7,40 points calculées respectivement entre le QIV et le QIT ainsi qu'entre le RV et le TC des tests, sont aussi toutes deux significatives.

Le tableau 8 présente les moyennes de QI obtenues aux deux tests selon l'année scolaire des sujets. L'examen de ce tableau révèle que malgré la présence du biais d'échantillonnage, les comparaisons établies entre les cotes de QI de l'échantillon total demeurent adéquates. En effet, les tendances notées ci-haut entre les cotes de QI du WISC-III et du Stanford-Binet IV ainsi qu'entre les cotes

Tableau 8

Moyennes d'intelligence globale, verbale et non-verbale au WISC-III et au Stanford-Binet IV^a selon l'année scolaire des sujets

Année scolaire	QI		
	global	verbal	non-verbal
	WISC-III		
Deuxième année	94,19	88,90	101,85
Quatrième année	102,91	95,83	110,68
Sixième année	95,39	86,70	106,26
	Stanford-Binet IV		
Deuxième année	95,49	89,71	107,52
Quatrième année	106,26	97,34	116,19
Sixième année	95,23	88,92	106,46

^a moyennes transformées

d'un même test demeurent les mêmes, outre les résultats élevés des élèves de la quatrième année.

Profils de QI verbal et de QI non-verbal. Un examen du tableau 7 révèle que les différences significatives calculées entre les moyennes des cotes d'intelligence verbale et non-verbale au WISC-III et au Stanford-Binet IV, sont en faveur de l'échelle non-verbale. En effet, les moyennes de QI non-verbal sont supérieures aux moyennes de QI verbal pour l'ensemble des sujets aux deux tests. Cela signifie qu'en moyenne, les enfants de l'échantillon ont obtenu des cotes supérieures au QI non-verbal de façon consistante. Ces résultats confirment l'hypothèse d'une certaine "normalité" d'un profil non-verbal > verbal auprès des jeunes de l'étude. Afin d'observer la fréquence de ce patron d'écart chez les enfants de l'étude, l'examen des données individuelles est nécessaire.

Les tableaux 9 et 10 présentent la direction, la fréquence et le pourcentage cumulatif des écarts calculés entre les échelles verbales et non-verbales du WISC-III et du Stanford-Binet IV. De façon générale, les écarts entre les moyennes verbales et non-verbales des deux tests offrent principalement des patrons non-verbal > verbal. Au WISC-III, 88,5% des enfants de l'étude ont des patrons d'écart QIP > QIV. Au Stanford-Binet IV, des patrons d'écart RA/V > RV sont présents chez 89,6% des enfants de l'échantillon. Quoique 14,6% des enfants obtiennent des écarts QIP > QIV de

Tableau 9
Direction, fréquence et pourcentage cumulatif des écarts
QIV-QIP au WISC-III

Écarts cumulatif	Fréquence	Pourcentage
-50 à -54	1	1,0
-45 à -49	1	2,1
-40 à -44	2	4,2
-35 à -39	3	7,3
-30 à -34	8	15,6
-25 à -29	3	18,8
-20 à -24	15	34,4
-15 à -19	19	54,2
-10 à -14	19	74,0
-5 à -9	7	81,3
-1 à -4	7	88,5
0	1	89,6
1 à 4	5	94,8
5 à 9	4	99,0
10 à 14	1	100,0

Note. Les écarts négatifs représentent des patrons QIP > QIV
et les écarts positifs représentent des patrons QIV > QIP.

Tableau 10

Direction, fréquence et pourcentage cumulatif des écarts RV-
RA/V au Stanford-Binet IV^a

Écarts cumulatif	Fréquence	Pourcentage
-45 à -49	1	1,0
-40 à -44	3	4,2
-35 à -39	4	8,3
-30 à -34	7	15,6
-25 à -29	17	33,3
-20 à -24	13	46,9
-15 à -19	18	65,6
-10 à -14	12	78,1
-5 à -9	8	86,5
-1 à -4	3	89,6
0	4	93,8
1 à 4	2	95,8
5 à 9	2	97,9
10 à 14	1	99,0
15 à 19	1	100,0

Note. Les écarts négatifs représentent des patrons RA/V > RV
et les écarts positifs représentent des patrons RV > RA/V.

^a moyennes transformées

10 points et moins au WISC-III, 73,96% d'entre eux obtiennent ce même patron avec des écarts de 10 points et plus. Au Stanford-Binet IV, tandis que des écarts RA/V > RV de 10 points et moins se retrouvent auprès de 11,5% des sujets, 78,1% obtiennent un patron d'écart RA/V > RV de 10 points et plus. De telles données appuient l'hypothèse d'une fréquence de patrons non-verbal > verbal beaucoup plus élevée que de patrons verbal > non-verbal au WISC-III, en plus de retracer la même chose au Stanford-Binet IV.

Lorsqu'un coefficient de corrélation est calculé entre les différences d'échelles verbales et non-verbales au WISC-III et au Stanford-Binet IV, une faible corrélation significative est trouvée, $r = 0,26$, $p < 0,05$. L'examen du graphique de cette relation et des corrélations entre les variables de la présente étude (âge, sexe, dates d'administration des tests, cote totale au WISC-III et au Stanford-Binet IV, etc.) en plus des indices de différence verbal-non-verbal, ne fait rien voir d'anormal par rapport à cette relation. Par conséquent, cet indice de corrélation indique que l'écart calculé entre les composantes verbales et non-verbales n'est pas nécessairement le même (quantitativement à tout le moins) pour un même sujet aux deux tests.

CHAPITRE QUATRIÈME
DISCUSSION ET CONCLUSIONS

Discussion

La présente étude met à l'épreuve l'application des tests d'intelligence américains auprès d'individus qui ne sont pas représentés dans les échantillons de normalisation. Les profils d'intelligence verbale et non-verbale obtenus à l'aide de versions traduites et adaptées du WISC-III et du Stanford-Binet IV sont examinés auprès d'un groupe d'enfants francophones du Nouveau-Brunswick.

En général, les analyses des données établissent que la population cible est différente de la population normative américaine. Par contre, les profils de QI des enfants de l'étude correspondent à ceux des groupes minoritaires et des groupes bilingues aux États-Unis. Toutefois, la fréquence du profil d'intelligence non-verbal > verbal et les écarts moyens entre les cotes d'habiletés verbales et d'habiletés non-verbales des enfants du présent échantillon sont plus prononcés que chez ces groupes d'enfants américains.

Le présent chapitre aborde la discussion des résultats obtenus. En premier lieu, les données de comparaisons entre les deux tests sont examinées. Comme hypothèses d'étude, il était prévu de trouver des corrélations positives significatives entre les cotes de QI du WISC-III et du Stanford-Binet IV, tout en mesurant des différences significatives entre les cotes des deux tests. En second

lieu, l'analyse des profils d'intelligence verbale et d'intelligence non-verbale aux deux tests est approfondie. D'après les données empiriques recueillies lors de la recension des écrits, les hypothèses de départ prévoyaient que chez des enfants qui diffèrent de la norme américaine, des écarts significatifs soient calculés entre les indices d'intelligence verbale et d'intelligence non-verbale au WISC-III, avec une fréquence élevée du patron d'écart QIP > QIV. L'application des règles de la norme américaine est examinée auprès de l'échantillon de l'étude. Les conclusions présentent en dernière partie des considérations pratiques à l'intention des psychologues scolaires de la province. Certaines suggestions en vue d'études futures sont présentées à la lumière des forces et des limites de la présente étude.

Comparaisons des tests

La théorie précise qu'il est possible pour deux tests d'intelligence semblables, publiés et normalisés à différentes époques, de produire des résultats différents (Bracken, 1988; Flynn, 1987). Plusieurs études empiriques ont d'ailleurs démontré ces faits en présence de corrélations élevées entre deux différentes révisions d'un même test (Bolen et al., 1995; Doll & Boren, 1993; Graf & Hinton, 1994; Newby et al., 1993; Slate, 1995a; Slate & Jones, 1995; Thorndike et al., 1986b; Vance et al., 1996; Wechsler, 1991).

Les corrélations élevées calculées entre les données moyennes de l'étude appuient la validité de l'adaptation canadienne-française du WISC-III lorsque comparée à celle du Stanford-Binet IV. La corrélation la plus élevée entre les résultats du WISC-III et du Stanford-Binet IV est notée entre leurs échelles d'intelligence globale. Cet indice de corrélation correspond étroitement à ceux rapportés en littérature entre le QIT du WISC-III et le TC du Stanford-Binet IV (Carvajal et al., 1993; Lavin, 1996; Lukens & Hurrell, 1996; Prewett & Matavich, 1994; Saklofske et al., 1994). De plus, l'affinité des échelles verbales et des échelles non-verbales de la présente étude supporte la mesure de construits semblables aux deux tests.

Les corrélations établies entre le WISC-III et le Stanford-Binet IV sont donc comparables aux résultats d'études antérieures. Par contre, contrairement aux attentes théoriques (Bracken, 1988; Flynn, 1987) et aux résultats d'études empiriques (Lukens & Hurrell, 1996; Prewett & Matavich, 1994; Saklofske et al., 1994), les moyennes calculées aux deux tests sont semblables. En effet, à l'exception des échelles non-verbales qui sont séparées par une différence significative de 4,06 points en faveur du Stanford-Binet IV, les moyennes obtenues aux deux tests sont identiques. L'hypothèse de différences significatives entre les moyennes des tests est ainsi rejetée en partie.

Des moyennes de QI semblables au WISC-III et au Stanford-Binet IV ont été obtenues chez des enfants non exceptionnels américains (Carvajal et al., 1993; Lavin, 1996). Il est donc possible de supposer que pour des enfants d'intelligence moyenne, le WISC-III et le Stanford-Binet IV auraient tendance à offrir des résultats semblables, même auprès des groupes d'enfants qui diffèrent de la norme américaine. Les résultats de la présente étude nécessiteraient alors des explications de la différence des moyennes notée entre le QIP et le RA/V des tests.

Il est possible entre autres d'envisager deux explications. Le biais d'échantillonnage de la présente étude fait que les enfants de quatrième année de la région de Bouctouche-Miramichi ont eu de meilleurs résultats à l'échelle RA/V du Stanford-Binet IV. De façon plus générale, le WISC-III a une exigence de plus que le Stanford-Binet IV au niveau de l'échelle non-verbale, la vitesse de réponse. Des études subséquentes pourraient mettre à l'épreuve ces explications ou d'autres.

Profils de QI

L'examen des profils d'intelligence des échelles de Wechsler auprès d'enfants de diverses populations minoritaires aux États-Unis, révèle une supériorité de l'échelle d'intelligence non-verbale (Altus, 1953; Christiansen & Livermore, 1970; Clarizio, 1982; Dean, 1979, 1980; Figueroa, 1989; McCullough et al., 1985; McShane &

Plas, 1984; Murphy, 1990; Slate, 1995b; Swanson & Deblassie, 1979; Swerdlik, 1978; Tanner-Halverson et al., 1993; Taylor & Richards, 1991; Taylor et al., 1984). Après d'un groupe d'enfants québécois, Quintin (1985) a noté que la fréquence du patron non-verbal > verbal est élevée et qu'une supériorité de la cote non-verbale semble être la norme chez une population canadienne-française.

Après de l'échantillon de l'étude, des différences significatives étaient prévues entre les résultats des échelles verbale et non-verbale au WISC-III. Cette hypothèse suppose que les enfants francophones du Nouveau-Brunswick sont différents de la norme américaine et que leurs profils de QI ressemblent à ceux des populations minoritaires aux États-Unis. Les profils d'habiletés verbales et non-verbales produits par l'échantillon de l'étude se distinguent nettement de ceux produits par les échantillons de normalisation aux États-Unis. En effet, la plupart des jeunes de l'étude présentent un profil de QI avec une supériorité de l'échelle non-verbale. Contrairement aux attentes (Prewett & Matavich, 1994), cette supériorité de l'échelle non-verbale est aussi notée au Stanford-Binet IV, non seulement au WISC-III.

Quoique les profils de QI des enfants de l'étude et ceux de la norme américaine soient différents, lorsque leurs cotes de QI global sont comparées, cette différence n'est pas discernée. En effet, les indices d'intelligence globale

des enfants de l'étude sont contenus à l'intérieur des limites de l'erreur standard de mesure (environ 100 ± 3) (Wechsler, 1991); leurs cotes de QI global sont équivalentes à celles de la norme américaine. Par contre, les cotes d'intelligence verbale et non-verbale obtenues par l'échantillon de l'étude dépassent cette marge standard d'erreur aux deux tests. Les moyennes des échelles d'intelligence verbale sont inférieures à la moyenne de la norme, tandis que les cotes moyennes obtenues aux échelles d'intelligence non-verbale y sont supérieures. De tels résultats suggèrent que l'interprétation isolée des indices d'intelligence globale est trompeuse puisqu'elle dissimule des données importantes au sujet des profils d'intelligence au WISC-III et au Stanford-Binet IV.

Les écarts notés entre les cotes d'intelligence verbale et non-verbale au WISC-III et au Stanford-Binet IV sont plus importants auprès de l'échantillon de l'étude qu'auprès de la norme américaine. En effet, les écarts moyens calculés dans la présente étude sont de plus de 15 points. D'après les données de standardisation du WISC-III, des écarts de 15 points et plus sont statistiquement significatifs (Wechsler, 1991). Auprès de l'échantillon de l'étude, ces différences moyennes sont présentes chez plus de la moitié des enfants et leur fréquence est deux fois plus élevée que chez les enfants de la norme américaine. Ceci rend donc difficile l'interprétation clinique des écarts entre les échelles

verbale et non-verbale chez les enfants représentés par le présent échantillon.

Malgré le biais d'échantillonnage de la présente étude, des différences fondamentales entre la population étudiée et la population de référence semblent entraîner les divergences notées entre leurs profils de QI. Ces différences fondamentales peuvent provenir d'une nette distinction entre les habiletés cognitives des deux populations. Notamment, les Anglo-canadiens ont un profil d'intelligence différent de la norme américaine (Wechsler, 1996). Les Canadiens francophones pourraient eux aussi connaître des forces et des faiblesses cognitives particulières comparativement aux Anglo-américains, puisque des différences culturelles, linguistiques, sociales, économiques, etc., les distinguent de la norme américaine.

En revanche, les différences entre les profils de QI des jeunes de l'étude et ceux de la norme américaine peuvent provenir directement des tests administrés. Les adaptations canadiennes-françaises du WISC-III et du Stanford-Binet IV pourraient ne pas rendre justice aux enfants francophones du Nouveau-Brunswick lorsqu'utilisées conjointement avec les normes américaines. Notamment, la traduction de ces tests peut avoir modifié la difficulté des items originaux (Figueroa, 1989). De plus, un test standardisé aux États-Unis pourrait inadéquatement mesurer les capacités cognitives de jeunes non exceptionnels qui diffèrent de sa

population normative (Murphy, 1990; Taylor & Richards, 1991; Taylor et al., 1984); l'échelle verbale des tests a particulièrement été ciblée comme cause probable de ce problème (Dean, 1979; Naglieri, 1982; Kaufman, 1994).

Application des règles de la norme américaine

La distinction établie entre les résultats des individus de la norme américaine et des enfants francophones du Nouveau-Brunswick à des tests de QI occasionne certains problèmes. D'abord, sachant que leurs moyennes de QI et leurs profils d'intelligence sont différents, l'application des règles de la norme auprès des enfants francophones du Nouveau-Brunswick est inappropriée. Particulièrement, au WISC-III, le seuil de 11 points ($\alpha = 0,05$) comme écart statistiquement significatif entre les échelles d'intelligence verbale et non-verbale est inadéquat auprès des enfants de l'étude. Même le seuil de 15 points ($\alpha = 0,01$) représente peu la réalité des jeunes francophones du Nouveau-Brunswick puisqu'un tel écart semble être la norme pour eux. Par ailleurs, les règles de la norme américaine indiquent qu'en moyenne, les individus obtiennent un QI verbal équivalent au QI non-verbal. La présente étude décèle plutôt que chez la population cible, le potentiel intellectuel non-verbal est en moyenne supérieur au potentiel intellectuel verbal. De plus, leur niveau d'intelligence non-verbale dépasse en moyenne celui des

enfants de la norme américaine, tandis que le contraire est noté auprès de leur niveau d'intelligence verbale.

L'interprétation des profils d'intelligence des enfants de l'étude est inappropriée à l'aide des données normatives américaines. De plus, la pertinence des indices d'intelligence globale au WISC-III et au Stanford-Binet IV est remise en question puisque le calcul du QIT et celui du TC dépendent des résultats obtenus aux échelles verbales et non-verbales des tests. Aux deux tests, seul l'examen de la cote globale ne peut distinguer les différences des profils entre les enfants de l'étude et ceux de la norme américaine. L'amplitude et la direction des écarts moyens entre les échelles d'intelligence verbale et d'intelligence non-verbale suggèrent de plus qu'une interprétation des résultats à l'aide des cotes d'intelligence globale serait inadéquate. La valeur de l'indice de QIT aux échelles de Wechsler a déjà fait l'objet de critiques auprès d'enfants qui diffèrent de la norme américaine (McCullough et al., 1985; Quintin, 1985).

Conclusions

Parmi les objectifs ayant guidé la présente étude figurent d'abord le besoin de comparer empiriquement les cotes standard et les QI du WISC-III et du Stanford-Binet IV. L'examen des écarts entre les mesures d'intelligence verbale et non-verbale au WISC-III et au Stanford-Binet IV a particulièrement été visé. Comme objectif pratique, il est

attendu de communiquer les données empiriques de la présente étude aux psychologues scolaires francophones du Nouveau-Brunswick.

Trois résultats principaux ressortent des données de la présente étude. D'abord, les enfants de l'échantillon obtiennent des profils d'intelligence qui diffèrent de ceux de la norme américaine. De plus, la direction des écarts entre les indices d'intelligence verbale et non-verbale du WISC-III et du Stanford-Binet IV favorise surtout l'échelle d'intelligence non-verbale. Finalement, la fréquence élevée de ce patron d'écarts (non-verbal > verbal) suggère une certaine "normalité" de ce patron auprès des enfants de l'étude.

Il est difficile de cerner jusqu'à quel point les données de la présente étude sont généralisables auprès des autres populations canadiennes-françaises. D'ailleurs, il y a un biais dans l'échantillon de la présente étude où les enfants de quatrième année ont obtenu une performance supérieure à celle des enfants de deuxième et sixième année. Il s'agit probablement d'une différence fortuite. Un éventail plus diversifié d'enfants provenant de plusieurs communautés francophones du Nouveau-Brunswick est nécessaire avant de pouvoir généraliser les présents résultats d'étude à l'ensemble de la population acadienne de la province. De plus, l'échantillon correspond aux critères d'un échantillon de convenance plutôt que d'un échantillon aléatoire. Malgré

ces limites, les données de la présente étude sont d'une importance considérable. Les administrations contrebalancées des deux tests ont permis de réduire des effets possibles de pratique. De plus, l'échantillon compte un nombre de participants assez important, ce qui donne une valeur empirique à l'étude et augmente la puissance statistique des tests.

Puisque les conclusions de la présente étude correspondent à celles de Quintin (1985), il est possible de retrouver des résultats semblables en d'autres milieux canadiens-français. Cependant, avant que de telles conclusions soient émises, les résultats de la présente étude doivent être reproduits par d'autres chercheurs. En attendant, les données recueillies peuvent servir de lignes directrices pour l'interprétation des moyennes de QI auprès d'une population francophone au Nouveau-Brunswick. À la lumière des résultats de la présente étude, certaines suggestions peuvent être formulées sur l'usage des tests d'intelligence américains, traduits et accompagnés des normes américaines, auprès des enfants francophones du Nouveau-Brunswick.

L'adaptation canadienne-française du WISC-III (Maurice, 1992) donne des résultats semblables à celle du Stanford-Binet IV (Chevrier, 1988). Puisque le biais d'échantillonnage semble avoir joué un rôle auprès de la différence significative calculée entre les échelles non-

verbales des tests, il est difficile de jauger la pertinence de cet écart inter-test. Toutefois, les enfants de l'échantillon produisent des profils d'intelligence verbale et d'intelligence non-verbale semblables d'un test à l'autre. En pratique, les données de la présente étude suggèrent que la sélection d'un instrument de mesure plutôt que de l'autre, est laissée libre à la discrétion de l'examineur ou de l'examinatrice. Cependant, en théorie, l'utilisation du WISC-III est préconisée puisque ses normes sont plus récentes que celles du Stanford-Binet IV (Bracken, 1988; Flynn, 1987; Wechsler, 1991).

Deux considérations importantes concernent l'interprétation des moyennes de QI des enfants d'un milieu francophone au Nouveau-Brunswick. D'après les analyses des données, une prudence extrême est recommandée lors de l'interprétation des écarts entre les habiletés verbales et les habiletés non-verbales au WISC-III et au Stanford-Binet IV. Contrairement aux recommandations du manuel du WISC-III (Wechsler, 1991), l'interprétation clinique des écarts de 15 points et plus semble inadéquate auprès des enfants francophones du Nouveau-Brunswick. En effet, la fréquence d'écarts importants est élevée auprès des sujets de l'étude. Un nouveau seuil statistique sera nécessaire pour la population visée si les présents résultats sont reproduits. L'éventuel développement de normes canadiennes-françaises au WISC-III pourrait répondre à cette lacune.

L'interprétation isolée de la cote d'intelligence globale au WISC-III et au Stanford-Binet IV ne doit pas être effectuée chez les enfants francophones néo-brunswickois. Kaufman (1994) avertit d'ailleurs les professionnels travaillant en milieu multiculturel et bilingue de ne pas interpréter les résultats de la cote globale au WISC-III. D'après les résultats de la présente étude, l'indice d'intelligence globale est trompeur aux deux tests car il n'indique aucunement la présence d'un écart favorable au potentiel intellectuel non-verbal. Ainsi, l'interprétation des résultats doit plutôt se faire à l'aide des profils de QI, en considérant que pour les enfants francophones du Nouveau-Brunswick représentés par l'échantillon de la présente étude, il semble commun d'obtenir des cotes verbales inférieures aux cotes non-verbales lorsque les normes américaines des tests sont utilisées. Les résultats de l'étude appuient l'utilisation d'une batterie de tests plus complète afin de pouvoir mieux évaluer le fonctionnement cognitif des enfants de la population cible.

Comme projets d'études futures, il serait nécessaire de poursuivre la présente étude auprès de groupes d'enfants exceptionnels de langue française au Nouveau-Brunswick. Quoique les participants de la présente étude sont des élèves de classes régulières, aucune distinction n'a été faite au niveau des pathologies possiblement rencontrées (trouble d'apprentissage, trouble du comportement, trouble

expressif verbal, etc.); une vue d'ensemble de la population était plutôt recherchée. Des études subséquentes pourraient étudier les cotes de QI des enfants THADA ou connaissant divers troubles d'apprentissage, afin de pouvoir apprécier les divergences ou les ressemblances des profils des deux tests auprès de groupes diagnostics précis. De plus, la fréquence et la direction des écarts entre le calcul de l'intelligence verbale et non-verbale pourraient être identifiées auprès de ces enfants. De telles études pourraient tenter de cibler des enfants aux deux extrémités du continuum des habiletés cognitives afin de mieux décrire les résultats d'une population d'enfants francophones néo-brunswickois. Il serait de plus nécessaire d'examiner les cotes d'intelligence au WISC-III à l'aide des nouvelles normes canadiennes-anglaises, pour découvrir si ces normes sont plus aptes à représenter le vrai potentiel intellectuel des jeunes de la population cible.

Ceci soulève le besoin d'établir des normes pour la population visée par cette étude. Les Anglo-canadiens bénéficient maintenant de nouvelles normes au WISC-III (Wechsler, 1996). Il aurait été intéressant d'analyser les données de la présente étude à l'aide de ces nouvelles normes, puisque les normes canadiennes-anglaises seraient probablement plus appropriées que celles des États-Unis. Toutefois, l'utilisation de ces normes aurait été inadéquate puisque les comparaisons avec le Stanford-Binet IV

n'auraient pas été basées sur des normes correspondantes (Wechsler, 1996).

L'objectif principal de l'étude, qui était d'obtenir des données empiriques auprès des enfants francophones néo-brunswickois, a été atteint. Cependant, les données de la présente étude ne sont pas en mesure d'expliquer les résultats obtenus au WISC-III et au Stanford-Binet IV. Des études beaucoup plus vastes, c'est-à-dire abordant entre autres la culture, le niveau socio-économique, la scolarité des enfants, etc., sont nécessaires afin de découvrir le pourquoi de la tendance des résultats trouvés ici. La présente étude permet d'avoir accès à des données de comparaison, ce qui n'était pas disponible jusqu'à aujourd'hui. Pour l'instant, la présente étude fournit un appui empirique à l'argument de Kaufman (1993) concernant les populations d'enfants pénalisés par les éléments verbaux du WISC-III. Elle en étend l'application dans un milieu précis, soit celui des enfants francophones au Nouveau-Brunswick, en plus d'indiquer que l'argument de Kaufman s'applique aussi au Stanford-Binet IV pour les enfants visés par la présente étude.

RÉFÉRENCES

Altus, G. T. (1953). W.I.S.C. patterns of a selective sample of bilingual school children. Journal of Genetic Psychology, 83, 241-248.

Bolen, L. M., Aichinger, K. S., Hall, C. W., & Webster, R. E. (1995). A comparison of the performance of cognitively disabled children on the WISC-R and WISC-III. Journal of Clinical Psychology, 51, 89-94.

Bracken, B. A. (1988). Ten psychometric reasons why similar tests produce dissimilar results. Journal of School Psychology, 26, 155-166.

Carvajal, H. H., Hayes, J. E., Lackey, K. L., Rathke, M. L., Wiebe, D. A., & Weaver, K. A. (1993). Correlations between scores on the Wechsler Intelligence Scale for Children-III and the General Purpose Abbreviated Battery for the Stanford-Binet IV. Psychological Reports, 72, 1167-1170.

Carvajal, H., & Weyand, K. (1986). Relationships between scores on Stanford-Binet IV and Wechsler Intelligence Scale for Children-Revised. Psychological Reports, 59, 963-966.

Chevrier, J.-M. (1988). Échelle d'intelligence Stanford-Binet: Quatrième édition. Montréal, PQ: Institut de Recherches psychologiques.

Christiansen, T., & Livermore, G. (1970). A comparison of Anglo-American and Spanish-American children on the WISC. Journal of Social Psychology, 81, 9-14.

Clarizio, H. F. (1982). Intellectual assessment of Hispanic children. Psychology in the Schools, 19, 61-71.

Darou, W. G. (1992). Native Canadians and intelligence testing. Canadian Journal of Counselling, 26, 96-99.

Dean, R. S. (1979). Distinguishing patterns for Mexican-American children on the WISC-R. Journal of Clinical Psychology, 35, 790-794.

Dean, R. S. (1980). Factor structure of the WISC-R with Anglos and Mexican-American children. Journal of School Psychology, 18, 234-239.

Doll, B., & Boren, R. (1993). Performance of severely language-impaired students on the WISC-III, language scales, and academic achievement measures [Monograph]. Journal of Psychoeducational Assessment, 11, 77-86.

Edwards, R., & Edwards, J. L. (1993). The WISC-III: A practitioner perspective [Monograph]. Journal of Psychoeducational Assessment, 11, 144-150.

Figueroa, R. A. (1989). Psychological testing of linguistic-minority students: Knowledge gaps and regulations. Exceptional Children, 56, 145-152.

Flanagan, R. (1995). The utility of the Kaufman Assessment Battery for Children (K-ABC) and the Wechsler Intelligence Scales for linguistically different children: Clinical considerations. Psychology in the Schools, 32, 5-11.

Flynn, J. R. (1987). Massive IQ gains in 14 nations: What IQ tests really measure. Psychological Bulletin, 101, 171-191.

Gauthier, Y. (1988). L'influence des tests diagnostiques d'intelligence dans un milieu franco-ontarien. Apprentissage et Socialisation, 11, 111-114.

Glutting, J. J. (1989). Introduction to the structure and application of the Stanford-Binet Intelligence Scale-Fourth Edition. Journal of School Psychology, 27, 69-80.

Graf, M. H., & Hinton, R. N. (1994). A 3-year comparison study of WISC-R and WISC-III IQ scores for a sample of special education students. Educational and Psychological Measurement, 54, 128-133.

Grégoire, J. (1992). Évaluer l'intelligence de l'enfant: Échelle de Wechsler pour enfants. Liège, Belgique: Pierre Mardaga.

Groth-Marnat, G. (1984). Handbook of psychological assessment. New York: Van Nostrand.

Hartwig, S. S., Sapp, G. L., & Clayton, G. A. (1987). Comparison of the Stanford-Binet Intelligence Scale: Form L-M and the Stanford-Binet Intelligence Scale: Fourth Edition. Psychological Reports, 60, 1215-1218.

Hollinger, C. L., & Baldwin, C. (1990). Comparing scores on the Stanford-Binet, Fourth Edition with the WISC-R for exceptional children. Psychological Reports, 66, 979-984.

Horn, J. L., & Cattell, R. B. (1966). Refinement and test of the theory of fluid and crystallized general intelligences. Journal of Educational Psychology, 57, 253-270.

Kaufman, A. S. (1976). Verbal-performance IQ discrepancies on the WISC-R. Journal of Consulting and Clinical Psychology, 44, 739-744.

Kaufman, A. S. (1993). King WISC the Third assumes the throne. Journal of School Psychology, 31, 345-354.

Kaufman, A. S. (1994). Intelligent testing with the WISC-III. New York: Wiley.

Kitano, M. K., & De Leon, J. (1988). Use of the Stanford-Binet Fourth Edition in identifying young gifted children. Roeper Review, 10, 156-159.

Kluever, R. C., & Green, K. E. (1990). Identification of gifted children: A comparison of the scores on Stanford-Binet 4th Edition and Form L-M. Roeper Review, 13, 16-20.

Larrabee, G. J., & Holroyd, R. G. (1976). Comparison of WISC and WISC-R using a sample of highly intelligent children. Psychological Reports, 38, 1071-1074.

Lavin, C. (1996). The Wechsler Intelligence Scale for Children-Third Edition and the Stanford-Binet Intelligence Scale: Fourth Edition: A preliminary study of validity. Psychological Reports, 78, 491-496.

Little, S. G. (1992). The WISC-III: Everything old is new again. School Psychology Quarterly, 7, 148-154.

Lukens, J. (1988). Comparison of the Fourth Edition and the L-M Edition of the Stanford-Binet used with mentally retarded persons. Journal of School Psychology, 26, 87-89.

Lukens, J. (1990). Stanford-Binet, Fourth Edition and the WISC-R for children in the lower range of intelligence. Perceptual and Motor Skills, 70, 819-822.

Lukens, J., & Hurrell, R. M. (1996). A comparison of the Stanford-Binet IV and the WISC-III with mildly retarded children. Psychology in the Schools, 33, 24-27.

Maurice, P. (1992). Échelle Wechsler d'intelligence pour enfants, 3è édition. Consignes de passation et de notation (version préliminaire française-américaine). Montréal, PQ: UQAM.

McCall, V. W., Yates, B., Hendricks, S., Turner, K., & McNabb, B. (1989). Comparison of the Stanford-Binet: L-M and the Stanford-Binet: Fourth Edition with a group of gifted children. Contemporary Educational Psychology, 14, 93-96.

McCullough, C. S., Walker, J. L., & Diessner, R. (1985). The use of Wechsler Scales in the assessment of Native Americans of the Columbia River Basin. Psychology in the Schools, 22, 23-28.

McShane, D. A., & Plas, J. M. (1984). The cognitive functioning of American Indian children: Moving from the WISC to the WISC-R. School Psychology Review, 13, 61-73.

Messadié, G. (1995, Janvier). Intelligence et hérédité. Science & Vie, 928, 47-55.

Murphy, D. G. (1990). Comparison of scores of bilingual urban and monolingual suburban elementary school children

for two measures of intelligence. Psychological Reports, 67, 1375-1378.

Murphy, K. R., & Davidshofer, C. O. (1991). Psychological testing: Principles and applications (2è éd.). Englewood Cliffs, NJ: Prentice Hall.

Naglieri, J. A. (1982). Does the WISC-R measure verbal intelligence for nonenglish-speaking children? Psychology in the Schools, 19, 478-479.

Neisser, U., Boodoo, G., Bouchard, T. J., Jr., Boykin, A. W., Brody, N., Ceci, S. J., Halpern, D. F., Loehlin, J. C., Perloff, R., Sternberg, R. J., & Urbina, S. (1996). Intelligence: Knowns and unknowns. American Psychologist, 51, 77-101.

Newby, R. F., Recht, D. R., Caldwell, J., & Schaefer, J. (1993). Comparison of WISC-III and WISC-R IQ changes over a 2-year time span in a sample of children with dyslexia [Monograph]. Journal of Psychoeducational Assessment, 11, 87-93.

Papalia, D. E., & Olds, S. W. (1985). Introduction à la psychologie. Montréal, PQ: McGraw Hill.

Phelps, L., Bell, M. C., & Scott, M. J. (1988). Correlations between the Stanford-Binet: Fourth Edition and the WISC-R with a learning disabled population. Psychology in the Schools, 25, 380-382.

Prewett, P. N., & Matavich, M. A. (1992). Mean score differences between the WISC-R and the Stanford-Binet Intelligence Scale: Fourth Edition. Diagnostique, 17, 195-201.

Prewett, P. N., & Matavich, M. A. (1994). A comparison of referred students performance on the WISC-III and the Stanford-Binet Intelligence Scale: Fourth Edition. Journal of Psychoeducational Assessment, 12, 42-48.

Quintin, E. P. (1985). Note sur l'usage clinique du WPPSI. Psychologie Canadienne, 26, 214-217.

Reynolds, C. R., & Ford, L. (1994). Comparative three-factor solutions of the WISC-III and WISC-R at 11 age levels between 6-½ and 16-½ years. Archives of Clinical Neuropsychology, 9, 553-570.

Robinson, N. M. (1992). Stanford-Binet IV, of course! Time marches on! Roeper Review, 15, 33-34.

Rothlisberg, B. A. (1987). Comparing the Stanford-Binet, Fourth Edition to the WISC-R: A concurrent validity study. Journal of School Psychology, 25, 193-196.

Sabatino, D. A., Spangler, R. S., & Vance, H. B. (1995). The relationship between the Wechsler Intelligence Scale for Children-Revised and the Wechsler Intelligence Scale for Children-III scales and subtests with gifted children. Psychology in the Schools, 32, 18-23.

Saccuzzo, D. P., Johnson, N. E., & Russell, G. (1992). Verbal versus performance IQs for gifted African-American,

Caucasian, Filipino, and Hispanic children. Psychological Assessment, 4, 239-244.

Saklofske, D. H., Schwean, V. L., Yackulic, R. A., & Quinn, D. (1994). WISC-III and SB:FE performance of children with attention deficit hyperactivity disorder. Canadian Journal of School Psychology, 10, 167-171.

Sattler, J. M. (1992). Assessment of children (3è éd. rév.). San Diego, CA: Auteur.

Silverman, L. K., & Kearney, K. (1992). The case for the Stanford-Binet L-M as a supplemental test. Roeper Review, 15, 34-37.

Slate, J. R. (1995a). Two investigations of the validity of the WISC-III. Psychological Reports, 76, 299-306.

Slate, J. R. (1995b). Discrepancies between IQ and index scores for a clinical sample of students: Useful diagnostic indicators? Psychology in the Schools, 32, 103-108.

Slate, J. R., & Jones, C. H. (1995). Preliminary evidence of the validity of the WISC-III for African American students undergoing special education evaluation. Educational and Psychological Measurement, 55, 1039-1046.

Sternberg, R. J. (1993). Rocky is back again: A review of the WISC-III [Monograph]. Journal of Psychoeducational Assessment, 11, 161-164.

Sternberg, R. J., Conway, B. E., Ketron, J. L.,

& Bernstein, M. (1981). People's conceptions of intelligence. Journal of Personality and Social Psychology, 41, 37-55.

Strommen, G. A., McKinney, J. P., & Fitzgerald, H. E. (1983). Developmental psychology: The school-aged child (éd. rév.). Homewood, IL: Dorsey.

Swanson, E., & Deblassie, R. (1979). Interpreter and spanish administration effects on the WISC performances of Mexican-American children. Journal of School Psychology, 17, 231-236.

Swerdlik, M. (1978). Comparison of WISC and WISC-R scores of referred Black, White, and Latino children. Journal of School Psychology, 16, 110-125.

Tanner-Halverson, P., Burden, T., & Sabers, D. (1993). WISC-III normative data for Tohono O'odham Native-American children [Monograph]. Journal of Psychoeducational Assessment, 11, 125-133.

Taylor, R. L., & Richards, S. B. (1991). Patterns of intellectual differences of Black, Hispanic, and White children. Psychology in the Schools, 28, 5-9.

Taylor, R. L., Ziegler, E. W., & Partenio, I. (1984). An investigation of WISC-R verbal-performance differences as a function of ethnic status. Psychology in the Schools, 21, 437-441.

Terman, L. M., & Merrill, M. A. (1960). Stanford-Binet Intelligence Scale. Boston: Houghton Mifflin.

Thorndike, R. L., Hagen, E. P., & Sattler, J. M. (1986a). The Stanford-Binet Intelligence Scale: Fourth Edition. Guide for administering and scoring. Chicago: Riverside.

Thorndike, R. L., Hagen, E. P., & Sattler, J. M. (1986b). The Stanford-Binet Intelligence Scale: Fourth Edition. Technical manual. Chicago: Riverside.

Vance, H., Maddux, C. D., Fuller, G. B., & Awadh, A. M. (1996). A longitudinal comparison of WISC-III and WISC-R scores of special education students. Psychology in the Schools, 33, 113-118.

Wechsler, D. (1974a). Manual for the Wechsler Intelligence Scale for Children-Revised. San Antonio, TX: Psychological Corporation.

Wechsler, D. (1974b). Selected papers of David Wechsler. New York: Academic.

Wechsler, D. (1975). Intelligence defined and undefined: A relativistic appraisal. American Psychologist, 30, 135-139.

Wechsler, D. (1991). Manual for the Wechsler Intelligence Scale for Children-Third Edition. San Antonio, TX: Psychological Corporation.

Wechsler, D. (1996). Manual for the Wechsler Intelligence Scale for Children-Third Edition: Canadian supplement. Toronto, ON: Harcourt.

Weinberg, R. A. (1989). Intelligence and IQ: Landmark issues and great debates. American Psychologist, 44, 98-104.

ANNEXE A

Lettre de consentement parental

Chers parents,

En tant qu'étudiante de maîtrise en psychologie à l'Université de Moncton, je m'intéresse à la question de la mesure d'intelligence chez les jeunes francophones. La littérature scientifique récente nous indique que certaines difficultés sont rencontrées lorsque l'administration de tests d'intelligence normalisés aux États-Unis est effectuée auprès de populations qui diffèrent de la population normative américaine. Entre autres, des écarts entre les capacités verbales et non-verbales sont plus fréquents chez les enfants bilingues et de milieux minoritaires que ce que nous indique la norme américaine.

Monsieur Pierre Cormier, professeur de psychologie à l'Université de Moncton, et moi-même avons mis sur pied un projet de recherche portant sur l'utilisation des tests d'intelligence américains pour mesurer les capacités verbales et non-verbales de jeunes francophones du Nouveau-Brunswick. Nous nous proposons de rencontrer une centaine d'enfants à qui deux tests d'intelligence seront administrés à l'école même. Je rencontrerai chaque enfant individuellement à deux reprises en dehors de la classe, pour des périodes approximatives de 90 minutes chaque. Les horaires d'administration des tests seront organisés de façon à nuire le moins possible au progrès de l'enfant en classe. De plus, chaque enfant recevra un cadeau de remerciement suivant l'administration du deuxième test.

Notre expérience antérieure dans d'autres recherches nous indique que les enfants aiment bien participer à ce genre de tâche. Les enfants participant à l'étude seront choisis parmi ceux qui ramèneront l'approbation écrite de leurs parents. Nous vous prions d'indiquer la date de naissance de votre enfant puisqu'elle nous sera utile pour définir l'échantillon de cette étude.

La participation de votre enfant va nous permettre de former un échantillon normatif francophone dans la province du Nouveau-Brunswick. À partir de cet échantillon, les psychologues seront en mesure de mieux situer les capacités verbales et non-verbales des enfants francophones de la province en comparant leurs résultats à des données normatives qui se rapprochent beaucoup plus d'eux, tant au niveau de la langue qu'au niveau de la culture.

Je tiens à vous assurer que toute information concernant votre enfant demeurera confidentielle. Aucun nom ne sera dévoilé. Seuls Monsieur Cormier et moi auront accès aux données individuelles de votre enfant.

Vous pouvez décider de retirer votre enfant du projet à n'importe quel moment, en envoyant une note écrite à l'éducateur(trice) de votre enfant. Un rapport écrit décrivant les données globales de cette recherche sera disponible à l'école à l'automne prochain. Si vous désirez une copie de ce rapport, veuillez inscrire votre adresse postale à l'endroit approprié ci-bas afin que nous puissions vous en faire parvenir une copie.

N'hésitez pas à me contacter au (506) 773-6588 pour de plus amples renseignements sur ce projet. Je vous remercie de votre bienveillante attention.

Anick McGrath, B.Ps.
Étudiante de maîtrise
Département de psychologie
Université de Moncton

P.S. Prière de retourner la demi-page ci-jointe à l'école.

_____ J'ai lu la lettre décrivant le projet de Madame Anick McGrath sur la mesure de l'intelligence et je refuse que mon enfant y participe.

_____ J'ai lu la lettre décrivant le projet de Madame Anick McGrath sur la mesure de l'intelligence et j'accepte que mon enfant y participe.

Signature du parent

Date

Nom et date de naissance de l'enfant

Nom de l'éducateur(trice)

P.S. Je désire obtenir une copie du rapport décrivant les données globales de l'étude. Voici mon adresse postale:

ANNEXE B

Rapport global remis aux Districts scolaires et aux écoles

Comparaisons des profils d'intelligence verbale et
non-verbale au WISC-III et au Stanford-Binet IV
chez un échantillon d'enfants francophones
du Nouveau-Brunswick

par

Anick McGrath et Pierre Cormier
Département de psychologie, Université de Moncton

Rapport de recherche intérimaire présenté à l'administration
du District scolaire no.11 et aux écoles participantes

Février 1997

Comparaisons des profils d'intelligence verbale et
non-verbale au WISC-III et au Stanford-Binet IV
chez un échantillon d'enfants francophones
du Nouveau-Brunswick

L'évaluation du potentiel intellectuel d'un enfant est une des composantes majeures de l'intervention psychologique auprès de ces enfants. Une difficulté particulière de cette évaluation réside dans le fait que les normes de nombreux tests d'intelligence ne correspondent pas à une population spécifique, comme celle des enfants francophones du Nouveau-Brunswick. Le ou la psychologue scolaire faisant une évaluation doit alors supposer qu'il n'y a pas de différence entre la population spécifique à laquelle appartient l'enfant et la population normative. Une telle supposition n'a toutefois pas d'appui empirique.

De très nombreuses variables peuvent produire des différences entre deux populations. Dans le cas des Acadiens du Nouveau-Brunswick, il suffit de songer à la langue et au contexte social, économique et historique de cette population comme source possible de différences avec le groupe normatif habituel des tests d'intelligence, soit les enfants anglophones des États-Unis. Cette situation rend plus incertaine l'évaluation que fait le ou la psychologue scolaire. Peut-il ou peut-elle se fier aux mêmes directives et critères de décision avec de petits francophones néo-brunswickois qu'avec les anglophones des États-Unis? Cette question est étudiée sous deux angles dans la présente étude.

Le fait que les populations minoritaires n'enregistrent pas exactement le même niveau de performance que la population normative a été largement étudié aux États-Unis. Il s'agit de savoir dans ces recherches si les mêmes normes

rendent justice à tous les enfants de quelque groupe que ce soit.

Parmi les critères fréquemment utilisés en psychologie scolaire se trouve celui de l'écart entre le potentiel intellectuel verbal et le potentiel intellectuel non-verbal. La règle suivie est qu'un écart de plus de 12 points entre le potentiel verbal et le potentiel non-verbal indique la présence d'un patron anormal d'habiletés verbales et non-verbales. L'utilisation de cette règle suppose qu'une différence verbale-non-verbale est relativement rare et relativement balancée dans sa direction (autant d'enfants avec des habiletés verbales supérieures aux habiletés non-verbales que d'enfants avec des habiletés non-verbales supérieures aux habiletés verbales). L'étude de Quintin (1985) renseigne sur la possibilité que les conditions préalables à cette règle ne se rencontrent pas dans les populations francophones minoritaires. Elle avait trouvé avec une échelle d'intelligence pré-scolaire, le WPPSI, une prépondérance d'enfants avec des habiletés non-verbales supérieures aux habiletés verbales chez des petits francophones du Québec.

Ces deux constatations issues de la littérature scientifique sur la mesure de l'intelligence, soit les différences entre les groupes minoritaires et le groupe normatif de même que la possibilité qu'il y ait un écart entre le potentiel verbal et le potentiel non-verbal différent chez les francophones par rapport aux anglophones, ont motivé la présente recherche. Une autre source de motivation a été l'arrivée sur le marché d'une version remise à jour de l'échelle d'intelligence pour enfants de Wechsler, le WISC-III. Ceci donnait l'occasion de comparer cette échelle révisée à une autre échelle d'intelligence relativement populaire en psychologie scolaire, le Stanford-Binet IV, dont la dernière révision date d'à peu près dix ans, et de constater par la même occasion si la révision du

WISC réduisait les problèmes d'évaluation des enfants francophones.

Objectifs spécifiques

Deux objectifs spécifiques sont poursuivis dans cette recherche.

Premièrement, il s'agit de voir comment se comparent les enfants francophones du Nouveau-Brunswick par rapport aux normes des enfants des États-Unis sur chacun des deux tests d'intelligence utilisés.

Deuxièmement, il s'agit d'observer la fréquence et la direction des écarts entre les habiletés verbales et les habiletés non-verbales chez les enfants francophones du Nouveau-Brunswick sur chacun des deux tests d'intelligence utilisés.

Procédure

Les adaptations françaises du Wechsler Intelligence Scale for Children-Third Edition (WISC-III) et du Stanford-Binet-4^e édition ont été administrées à 96 enfants francophones du Nouveau-Brunswick. Les enfants provenaient de classes de 2^e, 4^e et 6^e année scolaire d'écoles de la région de Bouctouche-Miramichi et de Moncton. L'échantillon de sujets n'est pas nécessairement représentatif de la population puisqu'il est composé d'enfants qui ont été portés volontaires pour l'étude par leurs parents. Une seule école de Moncton a pu participer à l'étude parce que les deux étudiantes chargées de la collecte de données à Moncton ont eu des problèmes personnels qui les ont empêchées de mener à bon port leur tâche. Bénéficiant d'un fonds de recherche supplémentaire, nous espérons compléter l'échantillonnage des données de la région de Moncton.

Chaque enfant était rencontré individuellement lors de deux sessions d'une durée d'approximativement une heure et demie. L'ordre d'administration des tests était contrebalancée aléatoirement parmi les enfants pour éviter qu'un effet d'entraînement ou de fatigue n'affecte la

performance à l'un des deux tests (comme ça pourrait être le cas si le même ordre d'administration des tests était utilisé chez tous les enfants). Chaque enfant recevait une récompense pour sa participation.

Résultats

Les enfants de cet échantillon ont obtenu des résultats globaux aux deux tests d'intelligence dans la normale. Leurs données sont comparables à celles de l'échantillon de normalisation américain. Il y avait une petite différence significative d'un point en faveur du Stanford-Binet par rapport au WISC-III. Il y avait une différence de trois à quatre points en faveur des enfants de 4^e année par rapport aux autres enfants. Cette différence se retrouvait aussi bien chez les enfants de Moncton que ceux de la région de Bouctouche-Miramichi.

La très grande majorité des enfants ont présenté un écart entre leur potentiel intellectuel verbal et non-verbal en faveur du non-verbal et ce, aussi bien au WISC-III qu'au Stanford-Binet IV. La grande majorité de ces écarts étaient de plus de dix points.

Conclusions (intérimaires)

"Si la tendance se maintient" après la collecte de données dans la région de Moncton, les résultats globaux des enfants francophones du Nouveau-Brunswick devraient être équivalents à ceux de l'échantillon normatifs des enfants des États-Unis. Les différences notées au niveau des résultats globaux sont à l'intérieur de l'erreur de mesure des tests d'intelligence (trois points). Dans ce contexte, ces différences doivent être considérées comme petites et aléatoires à moins qu'elles ne soient systématiquement reproduites dans d'autres études.

Toutefois, l'examen des écarts entre le potentiel intellectuel verbal et le potentiel intellectuel non-verbal suggère que juger ces enfants avec leurs résultats globaux est plutôt trompeur car ils présentent pour la plupart un

écart favorable aux habiletés non-verbales. Cet écart est de plus non-négligeable et se reproduit sur les deux tests. Il y a là une différence fondamentale entre cet échantillon et l'échantillon normatif des États-Unis. Cette observation suggère que les conditions pour trouver d'intérêt clinique toute différence entre les potentiels intellectuels de plus de douze points peut importe sa direction, doit être modifiée pour tenir compte de la réalité néo-brunswickoise (si ce n'est pas de la réalité canadienne-française compte tenu des résultats de Quintin, 1985).

La nature de cet écart favorable au potentiel intellectuel non-verbal reflète sans doute les particularités sociales, économiques et historiques d'un échantillon comme celui de la présente étude. Cet écart amène la question à savoir où se situe la vraie nature du potentiel intellectuel des francophones du Nouveau-Brunswick, du côté de leur force ou de leur faiblesse? Si l'on considère que la vraie nature de ce potentiel de cet échantillon se trouve du côté de sa force, il faut alors conclure que les adaptations françaises des tests d'intelligence comme le WISC-III et le Stanford-Binet IV ne suffisent pas à rendre compte équitablement des habiletés verbales des enfants testés. Une normalisation canadienne-française de l'un ou l'autre de ces tests d'intelligence est encouragée.

Des tableaux présentant les statistiques descriptives complètes de l'étude seront fournis aux District scolaires participantes lorsque l'étude sera complète. Nous espérons que les présents résultats feront l'objet d'une publication scientifique.

ANNEXE C

Rapport global remis aux parents

Cher parent,

vous avez accepté que votre enfant participe à une étude sur l'intelligence lors de la session d'hiver 1996. La présente est pour vous informer des résultats obtenus.

Que voulions-nous faire?

Nous cherchions à voir dans cette étude si les traductions françaises des tests d'intelligence sont adéquates pour des enfants francophones.

Pourquoi?

Les tests d'intelligence sont construits aux États-Unis. Il est possible que leurs résultats reflètent plus les qualités de la population de ce pays que de celle du Canada. Ceci serait d'autant plus vrai pour les Acadiens puisqu'ils parlent français.

Qu'avons-nous fait?

Nous avons administré deux tests d'intelligence fréquemment utilisés par les psychologues scolaires à chacun des enfants. Les enfants étaient recrutés en 2^e, 4^e et 6^e année. Chaque enfant était vu deux fois, une fois pour chacun des tests.

Les résultats

Les tests d'intelligence donnent trois notes. Il y a une note pour l'intelligence verbale (comment résoudre des problèmes à l'aide de mots, ex.: faire des mots croisés). Il y a aussi une note pour l'intelligence visuo-spatiale (comment résoudre des problèmes en utilisant ses habiletés manuelles et visuelles, ex. faire des casse-tête). Et, en faisant l'addition des deux notes, on obtient une autre note pour l'intelligence globale.

Au point de vue de la note d'intelligence globale, il n'y avait pas de différence notable entre les enfants acadiens et américains.

Toutefois, chez la grande majorité des enfants acadiens, les notes d'intelligence visuo-spatiale étaient nettement supérieures à celles de l'intelligence verbale.

Que penser de ces résultats?

Cette étude montre que, chez les enfants acadiens, l'intelligence visuo-spatiale est plus forte que l'intelligence verbale. Ceci est différent des enfants des États-Unis où ces deux formes d'intelligence sont égales. Cette caractéristique des enfants acadiens peut être produite par un talent naturel propre à cette population ou par une erreur incorporée aux tests d'intelligence faits en anglais et traduits en français. Des recherches ultérieures (avec d'autres enfants) permettront de trancher cette question.

Ces résultats sont importants au point de vue de l'évaluation des enfants acadiens. Ils indiquent qu'il faut être très prudent dans l'utilisation de la note globale d'intelligence. Cette note est faite à partir de l'addition des notes d'intelligence verbale et d'intelligence visuo-spatiale. Cependant, les résultats de notre étude montrent que, chez les enfants acadiens, les deux notes d'intelligence verbale et d'intelligence visuo-spatiale sont différentes. Puisqu'il est difficile d'additionner des choses différentes, il serait préférable d'utiliser les deux notes d'intelligence verbale et d'intelligence visuo-spatiale séparément plutôt que seulement la note globale.

Aux États-Unis, les enfants ayant des problèmes d'apprentissage ont souvent des profils d'intelligence différents de la population normale. Les psychologues scolaires se servent de la présence d'une supériorité verbale ou d'une supériorité visuo-spatiale pour intervenir

auprès de ces jeunes en utilisant leurs forces mentales. Il faudra ajuster les règles d'intervention puisque, chez les enfants acadiens, il y a supériorité de l'intelligence visuo-spatiale par rapport à l'intelligence verbale.

En guise de conclusion, les résultats des enfants ayant participé à cette recherche seront présentés dans une thèse et une publication scientifique. Ils serviront aussi à faire des recommandations auprès des psychologues scolaires de la région sur la façon d'utiliser les tests d'intelligence dans leurs évaluations.

En terminant, je tiens à vous remercier, vous et votre enfant. N'hésitez pas à me contacter si vous avez des questions. Vous pouvez me rejoindre en soirée au (506) 545-6862, ou pouvez m'écrire à l'adresse suivante: 828 rue Poirier, Bathurst, Nouveau-Brunswick, E2A 2T8.

Anick McGrath, B.Ps.
Étudiante en maîtrise
Université de Moncton

ANNEXE D

Formule pour le rééchelonnement des QI du Stanford-Binet IV

$$(((X - 100) / 16) * 15) + 100$$

X = QI du Test composite; QI de Raisonnement verbal; QI de Raisonnement abstrait-visuel; QI de Raisonnement mathématique; QI de Mémoire à court terme