The significance of country norms in cognitive assessment: A WISC-V Canadian and U.S. normative comparison

Sarah E. Babcock[1], Jessie L. Miller[2], Donald H. Saklofske[1], Jianjun Zhu[2]

[1]University of Western Ontario; [2]Pearson Clinical Assessment Inc.

Abstract

In further support of the continued need for distinct sampling and independent normative data sets for Canadian ability measures, the goal of the current study is to evaluate and inform users of the differences that exist between the Canadian and U.S. WISC-V norms. Debate over the use of the WAIS-IV Canadian adult norms has been suitably addressed, therefore in this study the normative differences in the WISC-V child and adolescent data were examined and illustrated in three ways: 1) presenting a comparison of country normative data across subtest and composite scores for demographically matched and unmatched samples 2) a review of Canadian-U.S. standard FSIQ score differences across ability level and 3) comparison of the case detection accuracy of gifted and intellectually disabled children. Results demonstrate that there are small mean differences overall in the WISC-V norms when comparing the standardization samples in the two countries. Further, and importantly however, these results also show significant differences in the sensitivity of the Canadian norms compared to U.S. norms in identifying intellectually extreme cases (i.e., gifted and intellectual disability). These differences in sensitivity directly relate to the correct identification/diagnosis of children in these categories, and this highlights the importance of the use of Canadian norms in clinical decision making. These results support the use of Canadian norms when assessing these exceptional children following best clinical assessment practices and testing guidelines.

Introduction

- The development of specific Canadian normative data for American-based intelligence tests began more than two decades ago when Canadian practitioners argued that American normative information was not adequately representative of the Canadian population's performance or demographic characteristics[1]
- Despite the widespread acceptance of Canadian norms in standardized ability testing with the Wechsler scales, the Canadian normative data itself is not without criticism[2,3,4]
- Most recent findings showed that WAIS-IV Canadian norms produce lower standard scores compared to the U.S. norms and the number of individuals in their samples scoring below a Full Scale IQ (FSIQ) of 85 was far greater when using the Canadian norms compared to the American norms
- Follow up study shows major sampling issues were present in the studies claims[1]
- Criticisms have raised some concern and confusion among Canadian psychologists related to both understanding the need for and 'ethical' use of Canadian normative data

Purpose

- Given WAIS-IV concerns, address the normative differences that exist in the **WISC-V Canadian**
- Provide an analysis of the distributions by ability level in the Canadian sample, and within special group samples (gifted or intellectually disabled)
- Examine results for clinically significant differences in sensitivity which has implications for greater diagnostic accuracy.

Goal

Provide researchers & clinicians with an understanding of the differences that *do and do not* exist between the country samples.

Methods

Data: Canadian and U.S. **WISC-V** samples & special group samples (Intellectually Disabled and Gifted cases)
- Normative sample for both countries (Canadian N = 880; US N=2200)
- Standardization sample for Canadian study (N=920) *includes special group study cases*

Participants:
- Matched-sample (US-CDN) comparison group(N=605): M Age = 11.5 (SD 3.1), and 53.1% female; Ethnicity 92.4% Caucasian & 7.6% Asian. *Note: Comparisons of other ethnic groups was not possible because of the major country differences in primary ethnic composition; Parent education level: 28% no high school diploma 17.7% high school diploma 42% college or trades & 37.5% university
- Matched-control comparison group (ID and GT): Normative cases matched on age, parent education level, ethnicity, and sex.
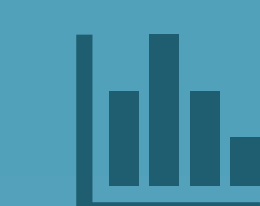
Analyses:
1. **Matched-sample comparison:** examine the mean standard score differences between the CDN and U.S. total and matched normative samples
2. **FSIQ Ability Level Differences:** Full CDN sample (including normative, clinical & exceptional cases) examined by ability level
3. **Special group comparisons:** Using CDN GT & ID data and matched-normative sample; chi-square analysis to determine the sensitivity and specificity

Results

Matched-Sample Comparison (Normative Sample)

- Canadian mean standard scores are slightly higher than U.S. mean standard scores across most subtests
- Few differences reach statistical significance and where they do, effect sizes are small (.20 or less).
- Processing Speed Index is the only composite mean standard score where CDN performance is very slightly lower than the U.S (non sig.)
- Largest composite difference is seen in the Visual Spatial Index; same index with the greatest performance difference in WPPSI–IV
- Overall, **mean FSIQ difference** is approximately **1.4 IQ points higher** for the Canadian sample (consistent with previous findings)

Standard Score Differences Across FSIQ Level

Frequency Distribution of Canadian-U.S. standard score differences, across ability level

FSIQ Difference	FSIQ RANGE 1 (130-160)	2 (120-129)	3 (110-119)	4 (90-109)	5 (80-89)	6 (70-79)	7 (40-69)	Total
-4	0	0	0	12	0	0	0	12
-3	0	0	0	83	8	4	1	96
-2	0	0	22	230	65	24	17	358
-1	7	23	74	229	67	31	9	440
1	12	0	0	0	0	0	0	12
2	2	0	0	0	0	0	0	2
Total	21	23	96	554	140	59	27	920

- The full CDN sample was used in this analysis (N = 920)
- Sample was scored using both CDN and U.S. norms
- Differences between FSIQ scores were calculated by subtracting the U.S. mean score from the Canadian mean score.
- Differences were recorded by frequency of the standard score differences (range: -4 to +2) for each ability level
- Majority of FSIQ differences are in the range of **1 or -2 point difference**
- Most differences fall in the **80 – 89, and 90 –109 FSIQ ranges**

Comparison of Special Groups and Matched Controls

Gifted Sample Comparison

CANADIAN NORMS:
- GT Sample : M = 125.44 (SD =8.48; range = 105 – 145)
- Matched Controls: M= 107.76 (SD =11.58; range 83–132)

US NORMS:
- Mean FSIQ score was nearly identical to the CDN
- GT Sample: M = 125.55 (SD =8.22; range = 106 –145)
- Matched Controls: M=104.57 (SD =9.23; range = 85–123)

Chi-Square Analysis
- GT study inclusion criteria as the reference standard
- CDN : Sensitivity = 13/(13+16) 45%; Specificity= 28/(28 +1) 97%
- US : Sensitivity = 9/(9 + 20) 31%; Specificity 28/(28+ 1) 97%

Proportion of Cases Categorized as Gifted

		Canadian Norms Pre-assessment Score		U.S. Norms Pre-assessment Score	
		≥130	≤130	≥130	≤130
WISC-V Score	≥130	13	1	9	1
	≤130	16	28	20	28

Demographic data for the Canadian gifted and intellectual disability samples and matched-sample comparison groups

	Gifted	Matched-Sample	Intellectual Disability	Matched Sample
N	29	29	20	20
Age				
Mean	12.28	12.33	12.65	12.87
SD	2.91	2.83	2.78	2.79
Range	6-16	6-16	6-16	6-16
Sex				
Female	34.5	37.9	40.0	40.0
Male	65.5	62.1	60.0	60.0
Race/Ethnicity				
Asian	13.8	13.8	5.0	5.0
White	86.2	86.2	95.5	95.5
PED Level				
≤ 11 years	-	-	-	-
12 years	3.4	3.4	60.0	60.0
13–15 years	20.7	20.7	20.0	20.0
≥ 16 years	75.9	75.9	20.0	20.0

Intellectual Disability Sample Comparison

CANADIAN NORMS:
- ID Sample : M = 56.50 (SD=10.09; range = 41--76)
- Matched: M= 92.95 (SD =13.34; range 58-121)

US NORMS:
- ID Sample : M = 58.05 (SD=10.42; range = 42-77)
- Matched: M=93.47 (SD= 12.10 range = 60-116)

Chi-Square Analysis
- ID study inclusion criteria as the reference standard
- CDN: Sensitivity = 18/(18+2) 90%; Specificity = 19/(19 +1) 95%
- US : Sensitivity = 15/(15+5) 75%; Specificity = 19/(19 +1) 95%

Proportion of Cases Categorized as Intellectually Disabled

		Canadian Norms Pre-assessment Score		U.S. Norms Pre-assessment Score	
		≤70	≥70	≤70	≥70
WISC-V Score	≤70	18	1	15	1
	≥70	2	19	5	19

Discussion

- When scored using U.S. norms, CDN mean scores are slightly higher on most subtests
- Mean FSIQ difference is ~1.4 IQ points higher
- When demographically matched & scored using U.S. norms, score differences decrease & in most cases, are no longer statistically significant
- Two countries appear similar on most subtests
- Problem with only mean group differences in the normative samples is that it does not describe what is going on in the tails of the distribution
- **"Tails"** are where clinicians are making important recommendations around cognitive and psychoeducational functioning and service needs
- There is equally good specificity (i.e., correctly categorizing those who are **not** clinical or exceptional, as nonclinical/nonexceptional)
- However, the important differences lie in the **sensitivity** of the norms.
  - Gifted sample CDN norms show 45%, compared to 31% when using U.S. norms.
  - ID sample shows sensitivity when using CDN norms is 90%, compared to 75% for the U.S.
  - Critical finding - sensitivity refers to the power of the assessment to detect cases (gifted and intellectually disabled) when a disorder (or special group classification) is actually present
  - By using U.S. norms with the CDN sample, the chance of not detecting these children increases by 14% (GT) and 15% (ID)

Conclusion

- Consistent finding of a significantly higher FSIQ score in CDN samples when scored using U.S. norms justifies the need for distinct CDN norms.
- The clinical and exceptional group analyses clearly demonstrate different levels of accuracy in the diagnosis/identification
- Groups are most commonly the children being assessed for identification, support, and planning

Given availability of WISC-V[CDN] norms, clinicians can utilize the appropriate normative group that was rigorously collected and is the best available representation of Canadian children

References

1. Miller, J. L., Weiss, L. G., Beal, A. L., Saklofske, D. H., Zhu, J., & Holdnack, J. A. (2015). Intelligent Use of Intelligence Tests Empirical and Clinical Support for Canadian WAIS–IV Norms. *Journal of Psychoeducational Assessment*, 33(4), 312-328.
2. Harrison, A. G., Armstrong, I. T., Harrison, L. E., Lange, R. T., & Iverson, G. L. (2014). Comparing Canadian and American normative scores on the Wechsler Adult Intelligence Scale-Fourth Edition. *Archives of Clinical Neuropsychology*, 29, 737-746.
3. Harrison, A. G., Holmes, A., Silvestri, R., & Armstrong, I. T. (2015). Implications for educational classification and psychological diagnoses when using the Wechsler Adult Intelligence Scale-Fourth Edition with Canadian as opposed to American norms. *Journal of Psychoeducational Assessment*, 33, 299-311.
4. Chevalier, T., Stewart, G., Nelson, M., McInerney, R., & Brodie, N. (2016). Impaired or Not Impaired, That Is the Question: Navigating the Challenges Associated with Using Canadian Normative Data in a Comprehensive Test Battery That Contains American Tests. Archives of Clinical Neuropsychology 31, 446–455