

# Is There a Universal Cognitive Diagnosis Related to the Tails of Intelligence Distribution?

Carmen Flores-Mendoza\*, Gabriela A. Pereira, Gabriela N. Nascimento, Mônica F. F. Novaes

Department of Psychology and Neuroscience Program, Federal University of Minas Gerais, Belo Horizonte, Brazil  
Email: \*carmenflor@uol.com.br

**How to cite this paper:** Flores-Mendoza, C., Pereira, G. A., Nascimento, G. N., & Novaes, M. F. F. (2020). Is There a Universal Cognitive Diagnosis Related to the Tails of Intelligence Distribution? *Psychology*, 11, 1661-1674.  
<https://doi.org/10.4236/psych.2020.1111105>

**Received:** October 14, 2020

**Accepted:** November 14, 2020

**Published:** November 17, 2020

Copyright © 2020 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).  
<http://creativecommons.org/licenses/by/4.0/>



Open Access

## Abstract

The diagnosis of giftedness or intellectual deficit is related to the extremes of IQ distribution. The universality of these diagnoses will depend on whether linguistic/cultural differences or Greenwich IQ (mean differences between nations) affect individual IQ rankings when different norms for the same cognitive measure are used. We verified both perspectives through two studies of the Wechsler Intelligence for Children—Fourth Edition, the most traditional intelligence scale. **Methods:** Study 1 analyzed American and Brazilian norms from the WISC-IV, and study 2 analyzed forty clinical protocols of children referred for cognitive assessment with the same scale. **Results:** Differences between WISC-IV American and Brazilian norms (study 1) impacted the IQ classification in the analyzed protocols (study 2). Using Brazilian norms, eight cases of high IQ (giftedness), and two cases of low IQ (intellectual deficit) were identified. Using American norms, one and six cases of high and low IQ respectively were estimated. The cognitive differences were more intense on the Verbal Comprehension Index. **Conclusion:** A broad academic discussion regarding the universality of cognitive diagnoses is urgent, especially in a time of increasing individuals' geographical mobility, and increasing demand for cross-cultural cognitive assessment.

## Keywords

Diagnosis, Universal Scale, Greenwich IQ, WISC-IV, Brazil

## 1. Introduction

The international Latin American student mobility has increased from 0.223% to an average of 0.297% (Colombia, Brazil, and Chile) from 2010 and 2018 (OECD, 2020). United States of America (USA) is one of the top ten countries that receives a largest number of international migrants, and presents significant increasing

of international students (Curtis, 2020). Thus, it is not surprising that when considering this student mobility, a reliable assessment of the student's cognitive potential is important for effective education, and supplemental education (if applicable).

Two kinds of clinical diagnosis related to the IQ scores distribution, both with considerable effects on education, are well known in applied psychology. These clinical diagnoses are intellectual development disorder ( $IQ < 70$ ) and intellectual giftedness ( $IQ > 130$ ). IQ scores below 70 or above 130 represent the extreme ends of the intelligence spectrum, when an intelligence test is administered to a nationally representative sample. Statistically, 2.27% of the population is considered "gifted", while 2.27% as "intellectually delayed". The 2.27% of people at each end of the intelligence spectrum would lead us to think that the larger the size of a country's population, the greater the number of people with intellectual delay or giftedness. This reasoning would be correct if the IQ average of 100 is equivalent for all nations. However, if an IQ score of 100 on a test implies a raw score of 50 points for country 1 and 60 points for country 2, there would be differences in the classification of giftedness and intellectual deficit in both countries. In other words, people that are considered gifted in country 1 would not be measured as gifted in country 2. Furthermore, people considered intellectually delayed in country 2 would not be classified as intellectually delayed in country 1. This distortion would affect not only diagnoses made by psychologists and psychiatrists, but it would also affect educational and psychotherapeutic treatments.

The International Test Commission (2018), and the American Educational Research Association, American Psychological Association, and National Council on Measurement in Education (2014), organizations that promotes policies and guidelines for effective testing and assessment, have provided guidelines to ensure test fairness and score comparability among culturally diverse populations. They recommend that if score differences across diverse linguistic groups are non-comparable, i.e., there are differences between two or more representative groups of their populations for the same test, the adapting test procedures should be revised or a linguistic/cultural investigation should be conducted. Therefore, such organizations consider an equivalence of psychological traits between populations, and biases such as no familiarity with test-taking skills, lack of motivation, little or no test practice, and linguistic peculiarities are underlying the differences between groups. It is a position defended for several researchers (Mushquash & Bova, 2007; Helms-Lorenz, Van de Vijver, & Poortinga, 2003), especially for the Wechsler scales (Funes, Roriguez, & Lopez, 2016; Sánchez-Escobedo, Hollingworth, & Fina, 2011).

Alternatively, differential psychology proposes another position. Mean IQs for all nations were provided by Lynn & Vanhanen (2002) and updated in subsequent publications, of which the most recent was published by Lynn & Becker (2019). In these studies, national IQs are calculated in relation to a British mean of 100 and a standard deviation of 15. This has been called the Greenwich IQ,

analogous to Greenwich Mean Time. The USA has a mean IQ of 98, and Latin America 86.4 (South America a mean IQ of 88—unweighted, based on means from samples of all studies found without consideration of data quality and sample size). Specifically, Brazil has a mean IQ of 87, mostly based on a non-verbal test, such as the Raven Matrices Progressives. The mean difference between Brazil and USA would be 11 points. The Greenwich IQ has been used in a considerable number of studies obtaining evidence of validity (Lynn & Becker, 2019; Rindermann, 2007).

Following the perspective of the Greenwich IQ, differences in average IQ between nations will be reflected in the standardization of intelligence tests, i.e., the IQ scale will have different psychological meaning for different countries. For instance, a Latin-American study (Duggan, Awakon, Loaiza, & Garcia-Barrera, 2019) conducted on a sample of 305 highly educated Colombian corporate executives, made a comparison in the WAIS-IV norms from Colombia, Chile, Mexico, Spain, United States, and Canada. One of their results is relevant for the present paper. The mean IQ of the Colombian sample was 120.3 (superior intelligence), but it was 112.7 (average intelligence) according to USA norms, a difference of 7.6 IQ points. Moreover, only 41.3% of the Colombian sample were within the same classification of USA norms for Full Scale IQ. The authors concluded there is an American-European assessment bias in cross-cultural test standardizations. Duggan et al. (2019) have the merit in presenting results, for the first time, regarding Latin American mean IQ differences through several norms of the Wechsler scales. However, is not clear as to why there was the highest agreement between the classifications Colombia/USA was observed in the Verbal Comprehension Index (52%), supposedly an index more sensitive to cultural/linguistic influence. Conversely, there was the lowest agreement was Working Memory Index (10%), supposedly less sensitive index to environment influence.

To examine the relevance of both perspectives (cultural/linguistic vs Greenwich IQ) for IQ classifications in children, we compared norms from the Wechsler Intelligence Scale for Children, fourth edition (WISC-IV) between the USA and Brazil (study 1) and, we analyzed the impact of these differences on Brazilian diagnoses of giftedness and intellectual delay (study 2).

Our hypotheses are:

- 1) If the Greenwich IQ perspective is valid, mean IQ differences will be noted between the WISC-IV's Brazilian and American norms, independent of the measured cognitive ability. There would be a universal IQ mean, and local cognitive diagnoses regarding giftedness or intellectual delay would not be reliable, even using appropriate standardizations;
- 2) If the cultural/linguistic perspective is valid, differences between Brazilian and American norms will specially be concentrated on the WISC-IV verbal subtests (VCI). There would not be a universal IQ mean and cognitive diagnoses of giftedness or intellectual delay would be reliable only in specific social contexts where standardizations of the test were conducted.

## 2. Methods

### 2.1. Study 1

#### 2.1.1. Brazilian and American WISC-IV Norms

The WISC-IV is the fourth version of the Wechsler Intelligence Scale for Children, an individually administered instrument for assessing a child's cognitive abilities. The USA norms (Wechsler, 2003), were based on a sample of 2200 children, aged between 6 years 0 months and 16 years 11 months from the Northeast (17%), South (34%), Midwest (24%) and West region (24%). The date of data collection was not provided, but it was evidently between 2000 (the 2000 American census was used for selecting the standardization sample) and 2002 (the test was published in 2003). The Brazilian norms (Wechsler, 2013) were based on a sample of 1860 schoolchildren enrolled in public and private schools from eight Brazilian states (69.2% Southeast region, 24.9% South region, and 5.9% North-east region). The data was collected between February and November 2010.

The WISC-IV consists of 16 subtests and four index IQs, namely 1) the Verbal Comprehension Index (VCI) composed of three core subtests: Similarities, Vocabulary and Comprehension (Information and Word Reasoning are supplementary subtests); 2) the Perceptual Reasoning Index (PRI), composed of three core subtests: Block Design, Picture Concepts and Matrix Reasoning (Picture Completion is a supplementary subtest); 3) the Processing Speed Index (PSI) composed of two core subtests: Coding and Symbol Search (Cancellation is the supplementary subtest); 4) and Working Memory Index (WMI) is composed of two core subtests: Digit Span and Letter-Number Sequences (Arithmetic is a supplementary test). Raw scores of each subtest are converted to scaled scores (1 to 19) that are based on the child's age. The sum of scaled scores of the core subtests is added, resulting in the index IQs. The index IQs are added to give the Full-Scale IQ. Several studies have showed the validity of WISC-IV for assessing gifted and intellectual delayed children, both in Brazil (Pereira, Araújo, Ciasca, & Rodrigues, 2015; Macedo, Mota, & Mettrau, 2017), and in USA (Koriakin, McCurdy, Papazoglou, Pritchard, Zabel, Mahone et al., 2013; McClain & Pfeiffer, 2012), and it is regarded as the gold standard in IQ testing by clinical and neuro-psychologists.

The subtests (and items) are the same in the Brazilian and American versions except for a slight modification in one item, found in "Similarities". Item 10 was modified ("frown and smile" in the USA version were replaced by "laughing and crying" in the Brazilian version). In the present study only the core 10 subtests were analyzed (Block Design, Similarities, Digit Span, Picture Concept, Coding, Vocabulary, Letter-Number Sequences, Matrix Reasoning, Comprehension, and Symbol Search). The time to administer the 10 core subtests that yield the Full-Scale IQ and index factor scores average between 70 to 80 minutes.

#### 2.1.2. Analysis

The differences in the scores for the Brazilian and American samples were calculated at four ages (6, 10, 14, and 16 year olds). For each age, there are data for

every three months of age (e.g., 6.0 - 6.3, 6.4 - 6.7, and 6.8 - 6.11). A total of 24 age-norms were analyzed, 12 per country.

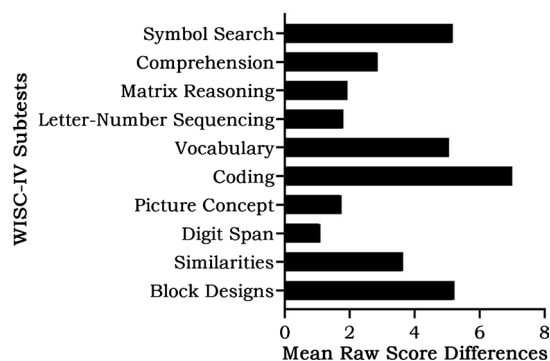
A positive difference between the raw scores from USA and Brazil indicates that a higher raw score was required by Americans than by Brazilians to achieve the same scaled score. A negative difference stated the contrary, i.e., a higher raw score was required by Brazilians than by Americans to achieve the same scaled score. For instance, considering norms for the age of 6.0 - 6.3 years, a scaled score of 4 in Block Design require 3 points for Americans while 1 point is required for Brazilians. In this case, the difference is +2 points. In Matrix Reasoning, a scaled score of 17 requires 33 points for Americans while 34 points are required for Brazilians. In this case, the average difference would be -1 point.

However, most of the scaled scores are derived from a range of raw scores. In these cases, the mean difference was calculated only for cases where the lowest raw score for country 1 was higher to the highest raw score for country 2, or vice-versa. For example, if the raw score range for country 1 was 28 - 30, and the raw score range for country 2 was 25 - 27, the mean difference was calculated ( $29 \text{ country 1} - 26 \text{ country 2} = 3 \text{ points}$ ). In this case, note that the lowest score required in country 1 was 28. It is higher than the highest score required in country 2, which was 27. If the range of scores was the contrary ( $25 - 27 \text{ country 1}, 28 - 30 \text{ country 2}$ ), the difference would be the same, but negative ( $-3 \text{ points}$ ). Otherwise, differences were not calculated (e.g.  $28 - 30 \text{ country 1}, 26 - 30 \text{ country 2}$  or  $14 - 16 \text{ country 1}, 14 - 15 \text{ country 2}$ ).

Finally, the Brazilian raw score equivalent to a scaled score of 10 for each subtest was converted to USA scaled scores. Note that a scaled score of 10 represents the average cognitive performance. Thus, a summed scaled score of 30 for VCI and for PRI, and 20 for WMI and for PSI is necessary to achieve an IQ of 100. Subsequently, the “new” scaled scores (in fact, American scaled scores) were summed for each index and converted to IQ according to the USA norms.

### 2.1.3. Results

Differences in average raw score from each WISC-IV subset based on the American and Brazilian norms are shown in **Figure 1**. Note that differences were

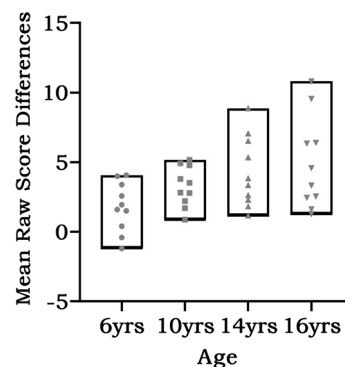


**Figure 1.** Raw score mean differences between the USA and Brazil for the scaled score of the WISC-IV.

obtained based on the scaled score (1 to 19) for each subtest and each of the 24 norms-age.

**Figure 1** shows the positive differences from all WISC-IV subsets, which means that a higher raw score is required for Americans than for Brazilians to reach the same scaled score. As a consequence, for any raw score, a higher scaled score is expected if Brazilian norms are used, and conversely, a lower scaled score is expected if American norms are used.

On the other hand, we plotted the average differences obtained in each age for each subtest in order to observe the trajectory of these differences. **Figure 2** shows that differences increase with age. The older age, the higher the raw score required for Americans.



**Figure 2.** Mean raw score differences for the ten subtests (inside of rectangle) in each age.

More specifically, the differences that led to more increases with age were related to: Coding (3.8 points differences at 6 years old increase to 10.7 points at 16 years old;  $\rho = 0.951$ ,  $p = 0.000$ ), Similarities (2.45 points at 6 yrs to 4.64 at 16.8 yrs;  $\rho = 0.909$ ,  $p = 0.000$ ) and Vocabulary (0.19 point at 6 yrs to 9.35 points at 16 yrs;  $\rho = 0.902$ ,  $p = 0.000$ ). The differences not related to age were Digit Span ( $\rho = -0.249$ ,  $p = 0.436$ ), Letter-Number Sequence ( $\rho = 0.228$ ,  $p = 0.476$ ) and Matrix Reasoning ( $\rho = 0.399$ ;  $p = 0.199$ ). The Spearman  $\rho$  between total mean differences from the WISC-IV subtests and age was 0.972.

Regarding the Indexes, the PSI (IQ 89) and the VCI (IQ 91.7) were the more distant indexes from the USA mean IQ of 100, while the WMI (IQ 95) was closer. Considering the Full-Scale IQ, the Brazilian mean IQ would be 89, relative to the USA norms, a difference of 11 IQ points between USA and Brazil (**Table 1**).

**Table 1.** Estimates of Brazilian index score according to USA norms.

Ages	VCI	PRI	WMI	PSI	Full Scale IQ
6.0 - 6.3	91	96	91	100	92
6.4 - 6.7	95	96	91	97	92
6.8 - 6.11	95	96	97	94	93
10.0 - 10.3	91	86	94	88	87
10.4 - 10.7	89	92	94	91	88

**Continued**

10.8 - 10.11	93	94	97	88	90
14.0 - 14.3	91	92	97	85	88
14.4 - 14.7	93	92	97	85	89
14.8 - 14.11	89	94	97	88	89
16.0 - 16.3	91	90	97	85	88
16.4 - 16.7	91	90	97	85	88
16.8 - 16.11	91	90	91	85	87
Mean	91.7	92.3	95.0	89.3	89.3

Note: VCI = Verbal Comprehension Index; PRI = Perceptual Reasoning Index; WMI = Working Memory Index; PSI = Processing Speed Index.

## 2.2. Study 2

### 2.2.1. Effects on Clinical Diagnosis

In order to verify the effect of WISC-IV norms differences (Study 1) on Brazilian clinical diagnoses, 40 WISC-IV protocols of children assessed in psychological clinics from two universities were analyzed. Twenty protocols were for children suspected of being gifted by their parents and teachers, and twenty protocols were for children with learning disabilities according to their teachers' opinions. All families signed a consent term for the use of the results for subsequent investigation, regarding giftedness or learning difficulties. The mean age of the analyzed group was 9.7 (SD = 2.6), with 80% male.

### 2.2.2. Analysis

Raw scores from each subtest and age were converted to scaled scores according to Brazilian and USA norms. In this analysis, positive differences between the scaled scores were due to the effect of the Brazilian norms. Negative differences between the scaled scores were due to the effect of the USA norms.

Regarding the indexes VCI, PRI, WMI and PSI, differences were calculated using Brazilian and American norms. However, the variations were classified into 3 types:

*Type 1* = Brazilian index is higher than the USA index, but falls within the USA confidence interval of 95%. In this case, the difference was not significant. For example, if a Brazilian WMI index was 115 with a 95% confidence interval between 107 - 121 and a USA WMI index was 110 with a 95% confidence interval between 102 - 117, the difference between populations parameters ( $WMI_{Br}$  115 vs.  $WMI_{USA}$  110) would not be significant. Note that the 115  $WMI_{Br}$  falls within the 95% USA confidence interval (102 - 117), and the 110  $WMI_{USA}$  falls within the 95% Brazilian confidence interval (107 - 121).

*Type 2* = Brazilian index is higher than the USA index, and is outside the USA confidence interval, but the USA index falls within the Brazilian confidence interval. In this case, the difference was not significant. For example, if a Brazilian PSI index was 121 with a 95% confidence interval between 110 - 127, while a



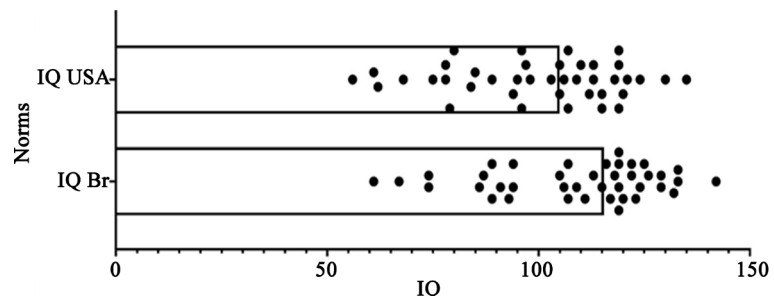
USA PSI index was 112 with a 95% confidence interval of 102 - 120, the difference between population parameters ( $PSI_{Br}$  121 and  $PSI_{USA}$  112) would not be significant. In this case, the 121  $PSI_{Br}$  falls outside of the 95% USA confidence interval (102 - 120), but the 112  $PSI_{USA}$  falls within the 95% Brazilian confidence interval (110 - 127).

*Type 3* = Brazilian index is higher than the USA index; it is outside of the USA confidence interval, and the USA index is below the Brazilian confidence interval. In this case, the difference is significant. For example, if a Brazilian VCI index is 115 with a 95% confidence interval between 107 - 121, while the USA VCI index is 102 with a 95% confidence interval between 95 - 109, differences between population parameters ( $VCI_{Br}$  115 vs.  $VCI_{USA}$  102) would be significant. The 115  $VCI_{Br}$  is outside of 95% USA confidence interval (95 - 109), as well, the 102  $VCI_{USA}$  is below to the 95% Brazilian confidence interval (107 - 121).

From 160 calculations (40 clinical protocols  $\times$  4 Brazil and USA indexes), only nine (5.6%) showed negative differences (the USA index was higher than the Brazilian index), and they were classified as type 1 (not significantly different). Therefore, the absolute majority (94.6%) of the differences were positive, that is, the Brazilian scaled score required a lower raw score than the American norms. These positive differences will affect the estimation of Full Scale IQ.

### 2.2.3. Results

The mean IQ of Brazilian protocols was 109.5 and 100.2, according to USA norms (Median Br = IQ 115.5, and Median USA = IQ 105). Thus, a difference of 9.3 IQ points. The distribution of the forty protocols along the IQ scale is shown in **Figure 3**.



**Figure 3.** Distribution of 40 protocols along the IQ scale according to the USA and the Brazilian WISC-IV norms.

**Figure 3** shows that more protocols with higher IQs and fewer protocols with lower IQs were observed when Brazilian norms are used. More considerable scaled score differences were observed in Symbol Search, Coding, Vocabulary and Comprehension. Conversely, fewer differences were observed in Letter-Number Sequence, Picture Concept, Digit Span and Matrix Reasoning.

Regarding the relationship between differences and age, Vocabulary ( $\rho = 0.766$ ,  $p = 0.000$ ) and Coding ( $\rho = 0.720$ ,  $p = 0.000$ ) data demonstrated the highest correlations with age, while Digit Span ( $\rho = -0.286$ ,  $p = 0.073$ ), Let-



ter-Number Sequence ( $\rho = -0.282$ ,  $p = 0.078$ ), Similarities ( $\rho = 0.134$ ,  $p = 0.410$ ) and Comprehension ( $\rho = 0.145$ ,  $p = 0.371$ ) did not show significant correlations. The Spearman correlation between total mean differences from the WISC-IV subtests and the 40 ages examined was 0.710. These results are similar to those obtained in study 1.

Regarding the indexes, the mean IQ difference between the use of the Brazilian and the American norms was 8.9 for PSI, 7.9 for VCI, 6.3 for PRI, and 5.6 for WMI. Thus, similar to study 1, skills such as velocity (PSI) and verbal comprehension (VCI) were the more distant, while memory (WMI) was less distant from the USA norms. The criteria for classifying differences in index factors (Analysis section) were applied to the analysis of the 40 clinical protocols (Table 2). Based on these criteria, the VCI showed a high percentage of cases classified as type 3 (57.5%), while the WMI was the index factor with a low rate of cases (15%) classified as Type 3. If the Full Scale IQ is considered, 75% of protocols presented difference Type 3; 7.5% presented difference Type 2, and 17.5% presented difference Type 1.

**Table 2.** Index factor scores according to the Brazilian and American norms, correlation between mean differences and age, and distribution of type of differences in each index factor and for the total IQ using the 40 protocols analyzed.

Index Factor	Brazil norms	USA norms	Diff	$r$ diff x age
VCI	109.30	101.40	7.90	0.47
PRI	109.00	102.70	6.30	0.54
WMI	105.70	100.10	5.60	-0.47
PSI	101.80	92.85	8.95	0.58
IQ Total	109.48	100.18	9.30	0.49
Type differences	VCI	PRI	WMI	PSI
1	12 (30.0%)	22 (55.0%)	32 (80.0%)	25 (62.5%)
2	3 (7.5%)	7 (17.5%)	2 (5.0%)	5 (12.5%)
3	23 (57.5%)	11 (27.5%)	6 (15.0%)	10 (25.0%)

Subsequently, we analyzed 30 clinical protocols that presented Type 3 differences. Note that Type 3 was the criteria for identifying how many protocols were significantly affected by the previously detected differences. We observed that 24 of clinical protocols (80%) changed its IQ classification.

**Table 3.** Changes in IQ classification of clinical protocols that presented Type 3 differences.

Protocol	IQ Br	IQ USA	Classification Brazil	Classification USA
1	142	135	Extremely High	Extremely High
2	132	118	Extremely High	High Average
3	129	119	Very High	High Average
4	129	119	Very High	High Average
5	126	115	Very High	High Average

## Continued

6	122	110	Very High	High Average
7	122	113	Very High	High Average
8	120	106	Very High	Average
9	119	107	High Average	Average
10	119	109	High Average	Average
11	119	105	High Average	Average
12	118	107	High Average	Average
13	116	97	High Average	Average
14	113	103	High Average	Average
15	111	102	High Average	Average
16	111	105	High Average	Average
17	109	94	Average	Average
18	107	95	Average	Average
19	107	96	Average	Average
20	105	96	Average	Average
21	98	89	Average	Low Average
22	94	84	Average	Low Average
23	93	85	Average	Low Average
24	91	78	Average	Very Low
25	89	79	Low Average	Very Low
26	89	75	Low Average	Very Low
27	86	75	Low Average	Very Low
28	84	70	Low Average	Very Low
29	74	68	Very Low	Extremely Low
30	69	62	Extremely Low	Extremely Low

*Note:* The Wechsler IQ classification is Extremely Low (69 and below); Very Low (70 - 79); Low Average (80 - 89); Average (90 - 109); High Average (110 - 119); Very High (120 - 129); and, Extremely High (130 and above).

The range of IQs associated with special education is above 120 (giftedness) and below 79 (intellectually delayed). According to Brazilian norms, **Table 3** shows eight protocols with IQ above 120 and two protocols with IQ below 79. However, according to American norms, there was just one protocol with IQ above 120 and seven protocols below IQ 79.

### 3. Discussion

In this paper, we analyzed the possibility of the universality of the diagnosis of giftedness or intellectual delay through two hypotheses. First, the Greenwich IQ perspective would be valid, and it affects the diagnosis linked to the extreme of the intelligence distribution. The second hypothesis, the cultural/linguistic pers-

pective would be correct, and the diagnoses made using local standardized tests are reliable. The first hypothesis implies that differences will be noted in all cognitive tasks, especially in those with more cognitive demand, favoring the country with higher mean IQ. The second hypothesis implies that differences will be more significant in verbal tasks than non-verbal tasks. For testing these two hypotheses, we conducted two studies. In the first, norms for two countries (Brazil and USA), using the Wechsler Intelligence Scale for Children-IV, were compared (Wechsler, 2003; Wechsler, 2013). Before the comparison, we verified the content, artwork, and scoring processing. From the 10 WISC-IV subtests analyzed, only Similarities had 1 item modified, out of 23. The remaining subtests and items, were the same for the USA and Brazilian standardizations. In the second study, the impact of the results of study 1 on forty Brazilian clinical protocols that used the WISC-IV was analyzed. Our results indicated four points of interest, which are discussed as follows:

*Positive differences.* The USA and Brazilian scaled scores of the 10 WISC-IV subtests showed that American children are required to obtain a higher raw score than Brazilian children to achieve the same scaled score. As a consequence, if USA norms are used instead of Brazilian norms, the final index and the Full Scale IQ would be lower (Study 1). Additionally, these differences were noted in the scoring of clinical protocols (Study 2). The difference was 11 IQ points, the same value estimated by the Greenwich IQ between the two countries. This result renders support to the first hypothesis, i.e., there would be a universal IQ mean, and local cognitive diagnoses would not be reliable, even using appropriate standardizations.

*Major and minor differences.* In both studies, differences were more intense in VCI (verbal comprehension), than WMI (memory). At this point, higher differences in VCI which resulted in changes in IQ classifications, as observed in Study 2, rendered support to the second hypothesis, i.e., there would not be a universal IQ mean, and cognitive diagnoses related to the tail of intelligence distribution would be valid only in the context where the standardized IQ test was conducted.

*Differences and age.* Measured differences increased with age (studies 1 and 2). This result can be interpreted as evidence of the influence that the environment had on intellectual differences, i.e., the earliest ages, Brazilian and American children would present little cognitive differences. As they grow older, the environmental context tended to increase the differences. Nevertheless, research in behavioral genetics has indicated the reverse, i.e., an increase of genetic influences and decrease of shared environmental influences over the years (Bartels, Rietveld, van Baal, & Boomsma, 2002; Deary, Spinath, & Bates, 2006).

*Cognitive diagnosis differences.* Considering clinical protocols that showed differences in Type 3 (IQ differences outside confidence intervals), just one, instead of eight cases of high IQ (>120), and seven, instead two cases of low IQ (<79), were observed. Note that the children with low cognitive performance could receive specialized education if they were living in the USA. However, differences were more concentrated in VCI (57.5% of protocols), which could have

affected the Full Scale IQ score.

Both studies' results render support to the second hypothesis for explaining the differences noted between the Brazilian and the American norms. Differences were more significant in verbal tasks. However, it is unclear why there were also substantial differences in PSI or PRI (considering Type 3 differences); the indexes cultural/linguistically little influenced.

Perhaps the fairest form of comparison between American and Brazilian norms would be the use of WMI, the index IQ that showed the fewest Type 3 differences (15% of protocols), and it was not influenced by age. Note that [Sánchez-Escobedo et al. \(2011\)](#) showed the same result with the WISC-IV Mexican version. Working Memory is considered a factor that is strongly related to psychometric intelligence ([Colom, Flores-Mendoza & Rebollo, 2003](#); [Salthouse & Pink, 2008](#)). Unfortunately, it would be difficult to expect an academic and professional consensus into considering the WMI as the primary reference for the diagnosis of giftedness and intellectual deficit.

#### 4. Conclusion

This paper analyzed how universal is the cognitive diagnoses related to the tail of intelligence distribution. The question merits a scientific answer considering the increase in students and families' geographical mobility and, consequently, the search for special education. Here we presented differences in IQ diagnosis in the most commonly used scale—the WISC-IV—for assessing intelligence. To our knowledge, there is no doubt this scale is valid in capturing performance at the extreme end of the intelligence spectrum. However, there is a higher probability of diagnosing giftedness and a lower probability of diagnosing intellectual delaying, if the Brazilian norms, instead USA norms are considered, which affect international psychological assessment practices. It is the same conclusion that other researchers reached with Spanish versions of the Wechsler scales ([Sánchez-Escobedo, Hollingworth, & Fina, 2011](#); [Fuentes et al., 2016](#)). On the other hand, differences between Brazil/USA norms were mostly influenced by differences in the VCI, which indicate that a cognitive diagnosis would be valid only for the social context where the standardization was conducted. Perhaps, the WMI index is the fairest cognitive measure when people culturally diverse are assessed through traditionally measures such the Wechsler scale. Nevertheless, the universality of cognitive diagnosis is still an open question. An accurate answer could be reached using different cognitive measures administered on representative samples from developing countries. Our study is a preliminary step in that direction. It is a challenge and a duty for behavioral science to investigate whether local standardizations hide actual group differences or flattening group linguistic-cultural differences.

#### Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

## References

- American Educational Research Association, American Psychological Association, National Council on Measurement in Education, Joint Committee on Standards for Educational and Psychological Testing (U.S.) (2014). *Standards for Educational and Psychological Testing*. Washington DC: AERA.
- Bartels, M., Rietveld, M. J. H., van Baal, G. C., & Boomsma, D. I. (2002). Genetic and Environmental Influences on the Development of Intelligence. *Behavior Genetics*, *32*, 237-249. <https://doi.org/10.1023/A:1019772628912>
- Colom, R., Flores-Mendoza, C. E., & Rebollo, I. (2003). Working Memory and Intelligence. *Personality and Individual Differences*, *34*, 33-39. [https://doi.org/10.1016/S0191-8869\(02\)00023-5](https://doi.org/10.1016/S0191-8869(02)00023-5)
- Curtis, C. J. (2020). Predictors of Cognitive Skills Development among International Students: Background Characteristics, Precollege Experiences, and Current College Experiences. *Journal of International Students*, *10*, 501-526. <https://doi.org/10.32674/jis.v10i2.253>
- Deary, I. J., Spinath, F. M., & Bates, T. C. (2006). Genetics of Intelligence. *European Journal of Human Genetics*, *14*, 690-700. <https://doi.org/10.1038/sj.ejhg.5201588>
- Duggan, E. C., Awakon, L. M., Loaiza, C. C., & Garcia-Barrera, M. A. (2019). Contributing towards a Cultural Neuropsychology Assessment Decision-Making Framework: Comparison of WAIS-IV Norms from Colombia, Chile, Mexico, Spain, United States, and Canada. *Archives of Clinical Neuropsychology*, *34*, 657-681. <https://doi.org/10.1093/arclin/acy074>
- Funes, C. M., Rodriguez, J. H., & Lopez, S. R. (2016). Norm Comparisons of the Spanish-Language and English-Language WAIS-III: Implications for Clinical Assessment and Test Adaptation. *Psychological Assessment*, *12*, 1709-1715. <https://doi.org/10.1037/pas0000302>
- Helms-Lorenz, M., Van de Vijver, F. J. R., & Poortinga, Y. H. (2003). Cross-Cultural Differences in Cognitive Performance and Spearman's Hypothesis. *Intelligence*, *31*, 9-29. [https://doi.org/10.1016/S0160-2896\(02\)00111-3](https://doi.org/10.1016/S0160-2896(02)00111-3)
- International Test Commission (2018). *The ITC Guidelines for the Large-Scale Assessment of Linguistically and Culturally Diverse Populations*. <http://www.InTestCom.org>
- Koriakin, T. A., McCurdy, M. D., Papazoglou, A., Pritchard, A. E., Zabel, T. A., Mahone, E. M., & Jacobson, L. A. (2013). Classification of Intellectual Disability Using the Wechsler Intelligence Scale for Children: Full Scale IQ or General Abilities Index? *Developmental Medicine and Child Neurology*, *55*, 840-845. <https://doi.org/10.1111/dmcn.12201>
- Lynn, R., & Becker, D. (2019). *The Intelligence of Nations*. London: Ulster Institute for Social Research.
- Lynn, R., & Vanhanen, T. (2002). *IQ and the Wealth of Nations*. Westport, CT: Praeger.
- Macedo, M. M. F., Mota, M. E. da., & Mettrau, M. B. (2017). WISC-IV: Evidências de Validade para Grupos Especiais de Superdotados WISC-IV. *Psicologia em Pesquisa*, *11*, 1-2. <https://doi.org/10.24879/2017001100100213>
- McClain, M. C., & Pfeiffer, S. (2012). Identification of Gifted Students in the United States Today: A Look at State Definitions, Policies, and Practices. *Journal of Applied School Psychology*, *28*, 59-88. <https://doi.org/10.1080/15377903.2012.643757>
- Mushquash, C. J., & Bova, D. L. (2007). Cross-Cultural Assessment and Measurement Issues. *Journal on Developmental Disabilities*, *13*, 53-65.
- OECD (2020). *International Student Mobility*.

- Pereira, A. M., Araújo, C. R., Ciasca, S. M., & Rodrigues, S. D. (2015). Avaliação da memória em crianças e adolescentes com capacidade intelectual limítrofe e deficiência intelectual leve. *Revista Psicopedagogia*, *32*, 302-313.
- Rindermann, H. (2007). The g-Factor of International Cognitive Ability Comparisons: The Homogeneity of Results in PISA, TIMSS, PIRLS and IQ-Tests across Nations. *European Journal of Personality*, *21*, 667-706. <https://doi.org/10.1002/per.634>
- Salthouse, T. A., & Pink, J. E. (2008). Why Is Working Memory Related to Fluid Intelligence? *Psychonomic Bulletin & Review*, *15*, 364-371. <https://doi.org/10.3758/PBR.15.2.364>
- Sánchez-Escobedo, P., Hollingworth, L., & Fina, A. D. (2011). A Cross-Cultural, Comparative Study of the American, Spanish, and Mexican Version of the WISC-IV. *TESOL Quarterly*, *45*, 781-792. <https://doi.org/10.5054/tq.2011.268057>
- Wechsler, D. (2003). *The WISC-IV Technical and Interpretive Manual*. San Antonio, TX: Psychological Corporation.
- Wechsler, D. (2013). *Escala de Inteligência Wechsler para Crianças: WISC-IV. Manual de instruções para aplicação e avaliação*. Adaptação e Padronização Brasileira: Rueda, F. J. M., Noronha, A. P. P., Sisto, F. F., Santos, A. A. A., & Castro, N. R. C. 4ed. São Paulo: Casa do Psicólogo.