JIDR Journal of Intellectual Disability Research
Published on behalf of mencap and in association with IASSID

 doi: 10.1111/jir.12499

Beyond the floor effect on the WISC-IV in individuals with Down syndrome: are there cognitive strengths and weaknesses?

L. Pezzuti,[1] R. Nacinovich,[2] S. Oggiano,[3] M. Bomba,[4] R. Ferri,[1] A. La Stella,[5] S. Rossetti[1] & A. Orsini[1]

1 *Department of Dynamic and Clinical Psychology, Sapienza University of Rome (Italy), Rome, Italy*
2 *School of Medicine and Surgery, Universita degli Studi di Milano-Bicocca, Milan, Lombardia, Italy*
3 *Child and Adolescent Mental Health Department, Universita degli Studi di Milano-Bicocca, Milan, Lombardia, Italy*
4 *Child and Adolescent Mental Health Department, San Gerardo Hospital, Monza, Italy*
5 *Service and Counseling of the People Down Italian Association, Rome, Italy*

Abstract

Background Individuals with Down syndrome generally show a floor effect on Wechsler Scales that is manifested by flat profiles and with many or all of the weighted scores on the subtests equal to 1.
Method The main aim of the present paper is to use the statistical Hessl method and the extended statistical method of Orsini, Pezzuti and Hulbert with a sample of individuals with Down syndrome (n = 128; 72 boys and 56 girls), to underline the variability of performance on Wechsler Intelligence Scale for Children-Fourth Edition subtests and indices, highlighting any strengths and weaknesses of this population that otherwise appear to be flattened.
Results Based on results using traditional transformation of raw scores into weighted scores, a very high percentage of subtests with weighted score of 1 occurred in the Down syndrome sample, with a floor effect and without any statistically significant difference between four core Wechsler Intelligence Scale for Children-Fourth Edition indices. The results, using traditional transformation, confirm a deep cognitive impairment of those with Down syndrome. Conversely, using the new statistical method, it is immediately apparent that the variability of the scores, both on subtests and indices, is wider with respect to the traditional method.
Conclusion Children with Down syndrome show a greater ability in the Verbal Comprehension Index than in the Working Memory Index.

Keywords Down syndrome, floor effect, intelligence, WISC-IV

Correspondence: Prof. Lina Pezzuti, Department of Dynamic and Clinical Psychology, Sapienza University of Rome (Italy), Via dei Marsi 78, 00195 Rome, Italy (e-mail: lina.pezzuti@uniroma1.it).

Introduction

Down syndrome, also known as trisomy 21, is a well-known genetic disorder. Typically, the nucleus of each cell of the human body contains 23 pairs of chromosomes, half of which are inherited from each parent. Down syndrome occurs when an individual has a full or partial extra copy of chromosome 21. There are three forms of Down syndrome, depending on the degree of the chromosomal abnormality. The standard trisomy 21 is the most prevalent and severe type of Down syndrome, accounting for some 95% of cases, where the whole of chromosome 21 is triplicated, whilst mosaicism and translocations involve limited triplication of only parts of chromosome 21 and account for 5% of cases of Down syndrome (Nadel 1999).

Over the past decade, significant advances have been made in understanding the specific differences in developmental trajectories of children and adolescents with Down syndrome, when compared with children and adolescents with typical development (Vicari & Carlesimo 2006). These differences are attributable to a number of factors: a high degree of variability in cognitive and social development (Patterson *et al.* 2013); the use and interpretation of assessment measures that are not always suitable (Pennington 2003); there is a difficulty in identifying how certain psychological factors such as motivation or anxiety can affect performance on development testing; children with Down syndrome show a variability in the pace of development (as shown by longitudinal studies) (Dykens *et al.* 2006).

The studies that have analysed intelligence quotients (IQs) with different evaluation measures have confirmed a slower rate of development in children with Down syndrome, compared with typically developing children, with a gradual increase in scoring disparities in relation to increased chronological age (Glue & Patterson 2009). However, we have no clear information in the literature that can tell us which instruments are most suitable for the measurement of IQ, nor whether differences in performance, using one tool rather than another, are found. Again, the most current lines of research are oriented to study comparisons with children with intellectual disability (ID), and some authors therefore propose that the normative groups must be with ID.

Through in literature, there are many studies that use other measures to estimate IQ; in our work, we use a Wechsler scale [Wechsler Intelligence Scale for Children-Fourth Edition (WISC-IV)], which is one of the most used tests for measuring IQ in individuals with typical and atypical development.

Clinical and research experience with intelligence testing in children with neurodevelopmental disorders shows that meaningful variation in performance is often obscured by floor effects when raw scores are converted to weighted scores (also known as scaled scores ranged 1 to 19) based on the normative data in test manuals. Whitaker and Gordon (2012) showed that the WISC-IV may overestimate some low IQs by several IQ points because of this floor effect. Their work focused on the WISC-IV and the Wechsler Adult Intelligence Scale-Third Edition and extrapolated the relationship between raw scores and scaled scores. According to their results in people with ID, the Wechsler Adult Intelligence Scale-Third Edition produces higher scores than the WISC-IV.

In particular, individuals with Down syndrome show a floor effect on the Wechsler scales that is manifested by flat profiles and with many or all of the weighted scores (ws) of the subtests equal to 1. This occurs despite the fact that there is wide variability of behaviour and performance (expressed by the range of raw scores corresponding to the weighted score of 1) among those given a weighted score of 1. All such differences are erased, and their strengths and weaknesses are hidden. This logically occurs with IQs equal to 40, where weighted scores of 1 are inevitable. The minimum in intelligence scales is usually 40, and in the clinical practice, a person with an IQ under 40 in these tests will have 40, but really, he does not have 40, he has less.

Clinicians, however, are aware of differences in the development of skills in the profile of children with Down syndrome because in the weighted score of 1, it brings together a large range of raw scores (Hessl *et al.* 2001; Dyer-Friedman *et al.* 2002; Glaser *et al.* 2003).

According to Davis and Escobar (2013), children with Down syndrome differ from many other children with ID in that Down syndrome subjects often have a variable neurocognitive profile with salient strengths and weaknesses as opposed to a flat globally depressed profile. Naturally, this variability is reduced among the individuals with Down syndrome with a severe ID, with the cognitive deficit deepens, and the cognitive profile flattens out with a floor effect for both subtests and indices.

The floor effect and other measurement problems in intelligence testing with children with IDs are very common; however, with a few exceptions such as those mentioned in the succeeding texts, they are not often recognised or discussed in the published studies.

In order to deal with the obstacle of this floor effect, Hessl *et al.* (2009) proposed a method that requires transforming WISC-III (Wechsler 1991) raw scores corresponding to a weighted score of 1 into z scores for any of the subtests of the scales. This method uses the mean and the standard deviation of raw scores in such subtests to perform the transformation using the following formula: $z = (x_{ij} - M_j)/S_j$, where x_{ij} is the raw score obtained by participant *i* in the age group *j* in any one of the subtests of the scale and where M_j and S_j are respectively the mean and the standard deviation of the jth age group in that subtest.

It is readily apparent that by using this method with the raw scores that correspond to a weighted score of 1, variation of the scores is recovered: raw scores (as opposed to a single weighted score of 1) are transformed into a plurality of z points.

Orsini *et al.* (2015) have extended this method, which so far has only been applied to subtest scores, to include IQs and Index scores. The main goal was to obtain IQs less than 40, which is the lower limit of indices and IQs for the WISC-IV Italian conversion tables (Orsini *et al.* 2012; Wechsler 2012). In subjects with a flat profile, i.e. with weighted scores of 1 in many subtests, the method consists of adding the z points of subtests attributable to Full Scale IQ and indices, and turning these sums into New Z (NZ) points using the test standardisation data, means and standard deviations of sums and distributions of identified z points. At this point, it is possible to reconvert the NZ points into new standard scores (IQs) with a mean of 100 and standard deviation of 15.

It might be helpful to point out that this method is essentially nothing but a double transformation of raw scores into *Z* scores where the first transformation is based on means and SD_s of the raw scores of the subtests, while the second transformation is based on means and SD_s of the sums of *Z* scores included in each composite score (for a more detailed statistical description of method, see Orsini *et al.* 2015).

The main aim of the present paper is to use the extended method of Orsini *et al.* (2015) with a large sample of Down syndrome subjects to underline the variability of the performance on subtests and indices that otherwise appear to be flattened, highlighting any strengths and weaknesses of this population.

This will be achieved by dividing the work into two steps. In the first step, we will be performing analyses and observations of the traditional scores of the Down syndrome sample, to have a cognitive efficiency framework of Down syndrome subjects according to the traditional manner of scoring the WISC-IV scale.

In the second step, we will apply the Orsini *et al.* method that takes into account the variability of the raw scores to the subtest scores corresponding to a standardised score of 1. So, we will compare this new method with the traditional method on our Down syndrome sample.

Method

Participants and instrument

The sample consisted of 128 participants (72 boys and 56 girls) aged between 7 and 16 years (mean age = 12.4, standard deviation = 2.59), with Down syndrome caused by Trisomy 21 confirmed by karyotype (including mosaicism and translocations).

All of these individuals have an ID according to Diagnostic and Statistical Manual of Mental Disorder-5 criteria, and thus under standard scoring, their scores would be subject to a floor effect on the WISC-IV (Wechsler 2008). The criteria of exclusion by the research were as follows: born preterm (gestational age at birth ≤34 weeks); severe uncorrected sensory alteration (auditory/visual); hypothyroidism; severe language disabilities; concurrent temporary illness in the last 2 weeks; altered consciousness (delirium); and not providing informed consent/assent to research.

Forty-two children (26 boys and 16 girls) were follow-up outpatients from the Child and Adolescent Mental Health Clinic, San Gerardo Hospital, Monza, Italy; all participants were from the north of Italy, whereas 86 children (46 boys and 40 girls) were recruited from the Information Service and Counseling of the People Down Italian Association, Section of Rome in Italy.

All participants were administered the Italian version of the WISC-IV (Orsini *et al.* 2012; Wechsler 2012). For the purposes of the present study, we examined the scores obtained in the 10 core subtests of the WISC-IV, i.e. Block Design, Similarities, Digit Span, Picture Concepts, Coding, Vocabulary, Letter–Number Sequencing, Matrix Reasoning, Comprehension and Symbol Search. We calculated the Full Scale IQ from the sum of the 10 subtests, and the four factor indices: the Perceptual Reasoning

Index (PRI), which includes Block Design, Picture Concepts and Matrix Reasoning; the Verbal Comprehension Index (VCI), including Similarities, Vocabulary and Comprehension; the Working Memory Index (WMI) including Digit Span and Letter–Number Sequencing; and the Processing Speed Index (PSI) including Coding and Symbol Search.

Data analysis

First step

The traditional transformations of the Down syndrome sample were carried out: the raw scores of the subtests were converted into weighted scores (mean = 10 and standard deviation = 3, with a range from 1 to 19) according to the conversion tables of the WISC-IV Italian standardisation. So, the percentage of Down syndrome subjects who achieved a weighted score of 1 on 10 core WISC-IV subtests was calculated. Again, the weighted scores (ws) of the 10 WISC-IV subtests were converted into z points using the inverse formula of their transformation: $z_i = (ws - 10)/3$.

Second step

Using the method proposed by Hessl *et al.* (2009) and extended by Orsini *et al.* (2015), the raw scores of subtests that correspond to a weighted score (ws) of 1 were transformed into z points using means and standard deviations of the raw scores of the subtests of each age group. All other weighted scores different from 1 were transformed into z points using the inverse formula of their composition in the standardisation process: $z = (ws - 10)/3$. The sums of the z points components of the four indices, and the Full Scale IQ, were calculated, transforming these sums into new points z (NZ) and later into standard points with mean = 100 and standard deviation = 15. In this way, the resulting measures of IQ and indices are expressed in *Z* scores, and therefore, the floor effect presented by a minimum score of 40 was overcome. For more detailed information about this statistical method, see Orsini *et al.* (2015).

For example, one child of 11 years, 8 months and 15 days of age gets the raw scores reported in column a of Table 1 to the three subtests (Block Design, Picture Concepts and Matrix Reasoning) of PRI. Using the canonical method, such raw scores would be all converted to a weighted point of 1 (see column b) corresponding to a normalised z point of −3 (see column c), from which it would emerge (after appropriate psychometric transformations) an IQ of 40. Conversely, using the means and standard deviations of the age-relevant reference sample of the subject (see column d), the new z points are lower (see column e). Then, the new points z (column f) are converted to a new point Z by the mean and standard deviation of the weighted points of the entire calibration sample (column h) and finally converted to IQ by the formula Z (15) + 100 (see column i) to obtain PRI. We can see that PRI value drops to 22 IQ points, which is inferior to the 40 IQ points we would have obtained using the canonical method.

Results

First step: analysis and observations of the Wechsler Intelligence Scale for Children-Fourth Edition traditional scores on the Down syndrome sample

The raw scores of the subtests were converted into weighted scores (mean = 10 and standard deviation = 3, with a range from 1 to 19) and into four indices and Full Scale IQ (mean = 100 and standard deviation = 15, with a range from 40 to 160) according to the conversion tables of the WISC-IV Italian standardisation (Orsini *et al.* 2012; Wechsler 2012). This is the traditional transformation of the raw scores.

According to this traditional transformation, a very high percentage of subtests with weighted score of 1 occurred in the Down syndrome sample. Table 2 shows the percentage of Down syndrome subjects who achieved a weighted score of 1 at 10 core WISC-IV subtests.

As can be seen from the results, there are subtests that, more than others, have a weighted score of 1, in particular Symbol Search (84%), followed by Coding (81%), Digit Span (81%), Comprehension (72%) and Letter–Number Sequencing (70%). So the two subtests of PSI and the two subtests of the WMI are the ones that have a higher frequency of weighted scores of 1. The smallest proportion occurs in the Similarities subtest (27%).

Table 1 Example of computing new PRI on data of one child of 11 years, 8 months and 15 days of age

	a	b	c	d	e	f	g	h	i
Subtest	Raw score	Range of raw score corresponding to a weighted score (ws) of 1	Normalised z scores = (ws − 10)/3	Mean (SD) of the raw scores of the subtests on group of 11 years of age (*n* = 200)	Z new points with Orsini *et al.* (2015) method: $z = (x_{ij} - M_j)/S_j$	Sums of the z points about PRI	Mean (SD) of weighted scores about VCI (*n* = 2200)	$Z = (z_n - mean_w)/SD_{ws}$	IQ = Z(15) + 100
Block Design	2	0–4	−3.00	34.87 (11.20)	(2 − 34.87)/11.20 = −3.11	−11.94	0.00 (2.30)	Z = (−11.94 − 0)/2.30 = −5.2	PRI = −5.2(15) + 100 = 22
Picture Concepts	3	0–4	−3.00	16.73 (3.45)	(3 − 16.73)/3.45 = −4.85				
Matrix Reasoning	4	0–5	−3.00	20.55 (4.16)	(4 − 20.55)/4.16 = −3.98				

PRI, Perceptual Reasoning Index; SD, standard deviation; VCI, Verbal Comprehension Index.

Table 2 Percentage of DS subjects with weighted score of 1 to 10 core WISC-IV subtest

Subtest	% DS subjects with weighted score of 1 to 10 core subtest
Block Design	59
Similarities	27
Digit Span	81
Picture Concepts	61
Coding	81
Vocabulary	62
Letter–Number Sequencing	70
Matrix Reasoning	62
Comprehension	72
Symbol Search	84

DS, Down syndrome; WISC-IV, Wechsler Intelligence Scale for Children-Fourth Edition.

The number of Down syndrome subjects (and percentage) that have a weighted score of 1 in the different subtests is shown in Table 3: in particular, 18 Down syndrome subjects have a weighted score of 1 in all 10 subtests, 20 Down syndrome subjects have a weighted score of 1 in nine subtests and so on.

In the present Down syndrome sample of a total of 1280 scores on the subtests (128 subjects × 10 subtests = 1280 subtests scores), 845 scores, that is, 66%, correspond to a weighted score of 1 (Table 4). Table 4 shows the cumulative percentage of the

Table 3 Number of subjects (and percentage) with a different number of subtest with weighted score of 1

Number of subtest with weighted score of 1	Number of subjects (%)
10	18 (14.1)
9	20 (15.6)
8	14 (10.9)
7	18 (14.1)
6	13 (10.1)
5	14 (10.9)
4	15 (11.7)
3	10 (7.8)
2	3 (2.3)
1	3 (2.3)

Table 4 Percentage and cumulative (cum) percentage of weighted scores in the sample

Weighted score	Number of subtest	Percentage	Cum %
1	845	66.0	66.0
2	139	10.9	76.9
3	117	9.1	86.0
4	79	6.2	92.2
5	38	3.0	95.2
6	34	2.7	97.9
7	16	1.3	99.2
8	7	0.5	99.7
9	4	0.2	99.9
10	0	0.0	99.9
11	1	0.1	100.0

Note. 128 subjects × 10 subtest = 1280 subtest scores.

occurrence of weighted scores presented by the subjects in our Down syndrome sample. The data confirm and show the level of deep cognitive impairment of the Down syndrome subjects.

Table 5 shows the means and standard deviations of weighted scores (ws) of the 10 core WISC-IV subtests and the values expressed in the z points using the inverse formula of their transformation: $z_i = (ws - 10)/3$.

Table 5 Means and standard deviations (SDs) of the weighted scores of 10 subtests calculated according to traditional method and their transformation into z points

	Weighted scores		Normalised z scores	
Subtest	Mean	SD	Mean	SD
Block Design	1.91	1.50	−2.70	0.50
Similarities	3.16	1.97	−2.28	0.66
Digit Span	1.39	0.99	−2.87	0.33
Picture Concepts	1.92	1.43	−2.69	0.48
Coding	1.42	1.13	−2.86	0.38
Vocabulary	1.91	1.47	−2.70	0.49
Letter–Number Sequencing	1.77	1.45	−2.74	0.48
Matrix Reasoning	2.19	1.89	−2.60	0.63
Comprehension	1.64	1.44	−2.79	0.48
Symbol Search	1.43	1.19	−2.86	0.40

This transformation in z points is required to allow the comparison with the results of the method proposed by Hessl *et al.* (2009) and extended by Orsini *et al.* (2015), which will be performed in the second part of analysis of this paper. Now looking at Table 5, it can be said that the means of the subtests are usually between two and three standard deviations below the mean.

The conversion tables of sums of raw scores in IQ of WISC-IV Italian standardisation start at a sum of weighted score equal to 26 that corresponds to an IQ of 40. Sums of weighted scores <26 receive IQ <40. In the present sample, 105 of 128 subjects (82%) have an IQ <40 and are not otherwise assessable. Of the remaining subjects, the IQ ranged from 40 to 62.

The means of the four indices for the Down syndrome sample are 53.4 (VCI), 49.5 (WMI), 49.5 (PSI) and 47.5 (PRI), with a range between averages of 5.9 points and an apparent deficit of PRI compared with the other three indices. However, comparing the pairs of indices, no comparison was statistically significant. In summary, a fairly flat pattern of scores can be observed, and thus, it would seem difficult to speak of strength in VCI or weakness in the PRI.

Second step: analysis of the raw scores with new method and comparison with the traditional method

A second transformation of the raw scores was carried out using the method proposed by Hessl *et al.* (2009) and extended by Orsini *et al.* (2015) to include the calculation of Full Scale IQ and four indices.

In this case, the raw scores that correspond to a weighted score of 1 were transformed into z points using means and standard deviations of the raw scores on the subtests of each age group of Italian standardisation sample.

All other weighted scores different from 1 were transformed into z points using the inverse formula of their composition in the standardisation process: $z = (ws - 10)/3$.

So all the scores of all the subtests, of all subjects, were expressed in z points, and, always following the method described by Orsini *et al.* (2015), the sums of the z points components – the four indices and the Full Scale IQ – were calculated, transforming these sums into new points z (NZ) and latter in standard points with mean = 100 and standard deviation = 15.

For comparison with the traditional transformation method of the first step analysis, also weighted score of traditional method was transformed into z points with the formula z = (ws − 10)/3. One analysis of variance was performed for each subtest, and an index between the two z scores obtained by each subject with the two methods (traditional vs. Orsini *et al.* method) was calculated.

Table 6 shows the means and standard deviation of z points of the 10 WISC-IV subtests calculated by the two methods. Compared with the traditional method, using the new method results in a decrease in means, a significant increase in standard deviations and essentially a much larger inter-individual variability.

An identical result emerges for the indices (Table 7): with the second method, a wider variability of scores is highlighted.

However, the more simple case to show the major variability of scores going from the first to the second method is to observe the change of the Indices in the 18 Down syndrome subjects with a weighted score of 1 in all 10 subtests of the WISC-IV (cf. Table 2). All 18 subjects have a VCI = 46, PRI = 41, WMI = 46, PSI = 47, then a flattening of the profile between the indices.

Using the method proposed by Orsini *et al.* (2015), the VCI of 18 subjects ranging from −1 to 51, the PRI from 7 to 37, the WMI between −10 and 31 and the PSI between 10 and 42.

Table 6 Means and standard deviation (SD) of z points to subtest calculated with the traditional and the Orsini *et al.* (2015) methods

	Z points with traditional transformation (normalised z scores)		Z points with the Orsini *et al.* (2015) method					
Subtest	**Mean**	**SD**	**Mean**	**SD**	$F_{1,127}$	*P*	**Eta squared**	**Power**
Block Design	−2.70	0.50	−2.78	0.70	3.93	0.049	0.03	0.503
Similarities	−2.28	0.66	−2.43	0.88	24.53	0.000	0.16	0.998
Digit Span	−2.87	0.33	−3.53	0.91	102.9	0.000	0.45	1.000
Picture Concepts	−2.69	0.48	−3.54	1.28	103.2	0.000	0.45	1.000
Coding	−2.86	0.38	−3.52	0.88	124.3	0.000	0.50	1.000
Vocabulary	−2.70	0.49	−3.30	1.10	77.82	0.000	0.38	1.000
Letter-Number Sequencing	−2.74	0.48	−4.08	1.64	126.9	0.000	0.50	1.000
Matrix Reasoning	−2.60	0.63	−3.06	1.06	73.60	0.000	0.37	1.000
Comprehension	−2.79	0.48	−3.92	1.39	135.8	0.000	0.52	1.000
Symbol Search	−2.86	0.40	−3.57	0.84	189.4	0.000	0.60	1.000

Table 7 Means and standard deviation of z points to indices calculated with the traditional and the Orsini *et al.* (2015) methods

	Traditional method				Orsini *et al.* (2015) method							
Indices	**Min**	**Max**	**Mean**	**SD**	**Min**	**Max**	**Mean**	**SD**	$F_{1,127}$	*P*	**Eta squared**	**Power**
VCI	−3.56	−0.79	−3.07	0.56	−17.05	−2.00	−9.65	2.99	866.22	0.000	0.87	1.000
PRI	−3.91	−1.31	−3.48	0.51	−15.40	−3.00	−9.37	2.54	974.65	0.000	0.88	1.000
WMI	−3.62	−2.01	−3.38	0.34	−12.16	−3.33	−7.60	1.90	808.89	0.000	0.86	1.000
PSI	−3.53	−2.16	−3.36	0.33	−10.23	−3.48	−7.09	1.34	1447.00	0.000	0.92	1.000

VCI, Verbal Comprehension Index; PRI, Perceptual Reasoning Index; WMI, Working Memory Index; PSI, Processing Speed Index; SD, standard deviation.

Using the method proposed by Orsini *et al.* (2015), the means of the z points decrease compared with the traditional method, while standard deviations increased and in some cases, doubled. This would indicate that there is a large increase in performance variability.

The same result emerges for the indices. However, what is the effect of this 'compression' of raw scores in the weighted scores 1, on four indices? We can see it well both in Table 7 of the means and standard deviations of z points and in Table 8 with the means and standard deviations of IQ points calculated with the two methods and for all four indices.

More particularly, the results of Table 8 show that the new IQs are lower than the means calculated with the traditional method, and the differences between the pairs of means are always statistically significant. The same results emerge for the z points data shown in Table 7.

In Table 9, the four new indices were compared with each other to identify any strengths or

Table 8 Means and standard deviation of IQ points to indices calculated with the traditional and the Orsini *et al.* (2015) methods

	Traditional method				Method of Orsini *et al.* (2015)							
Indices	Min	Max	Mean	SD	Min	Max	Mean	SD	$F_{1,127}$	*P*	Eta squared	Power
VCI	46	88	53.42	8.51	−1	88	42.70	17.71	113.86	0.000	0.47	1.000
PRI	41	80	47.54	7.63	0	80	38.87	16.57	79.33	0.000	0.38	1.000
WMI	46	70	49.47	5.07	−10	70	31.31	17.21	212.26	0.000	0.63	1.000
PSI	47	68	49.55	5.12	10	60	37.53	11.91	249.06	0.000	0.66	1.000

IQ, intelligence quotient; VCI, Verbal Comprehension Index; PRI, Perceptual Reasoning Index; WMI, Working Memory Index; PSI, Processing Speed Index; SD, standard deviation.

Table 9 Comparisons between pairs of subtest and indices using the Orsini *et al.* (2015) method with Down syndrome sample

Indices		Mean	SD	Mean	SD	$F_{1,127}$	*P*	Eta squared*	Power
	Comparison between subtest (in z means points)	Subtest 1		Subtest 2					
VCI	Similarities vs. Vocabulary	−2.43	0.87	−3.30	1.10	130.48	*0.000*	0.51	1.000
	Similarities vs. Comprehension	−2.43	0.87	−3.92	1.39	228.92	*0.000*	0.64	1.000
	Vocabulary vs. Comprehension	−3.30	1.10	−3.92	1.39	63.52	*0.000*	0.33	1.000
PRI	Block Design vs. Picture Concepts	−2.78	0.70	−3.54	1.28	61.93	*0.000*	0.33	1.000
	Block Design vs. Matrix Reasoning	−2.78	0.70	−3.05	1.06	11.23	*0.001*	0.08	0.914
	Picture Concepts vs. Matrix Reasoning	−3.54	1.28	−3.05	1.06	25.41	*0.000*	0.17	0.999
WMI	Digit Span vs. Letter–Number Sequencing	−2.87	0.33	−2.74	0.48	5.45	*0.021*	0.04	0.639
PSI	Coding vs. Symbol Search	−3.52	0.88	−3.57	0.84	0.34	*0.563*	0.00	0.089
	Comparison between indices (in IQ points)	Index 1		Index 2					
	VCI–PRI	42.70	17.72	38.87	16.57	7.51	0.007	0.06	0.776
	VCI–WMI	42.70	17.72	31.31	17.21	67.76	0.000	0.35	1.000
	VCI–PSI	42.70	17.72	37.53	11.91	15.10	0.000	0.11	0.971
	PRI–WMI	38.87	16.57	31.31	17.21	23.18	0.000	0.15	0.998
	PRI–PSI	38.87	16.57	37.53	11.91	1.06	0.304	0.01	0.176
	WMI–PSI	31.31	17.21	37.53	11.91	21.47	0.000	0.15	0.996

*For the interpretation of Eta squared, ≥0.01 is a small effect, ≥0.06 is a medium effect and ≥0.14 is a large effect.
SD, standard deviation; VCI, Verbal Comprehension Index; PRI, Perceptual Reasoning Index; WMI, Working Memory Index; PSI, Processing Speed Index; IQ, intelligence quotient.

weaknesses of the Down syndrome sample examined. Looking at Table 9, it can be seen that the Down syndrome sample has a strength (with significantly higher mean) in VCI and weakness in the WMI.

Comparisons between the indices show statistically significant differences with a large effect (see Eta squared) for the comparisons between VCI and WMI, between PRI and WMI and between WMI and PSI, where the WMI always presents the lowest mean. The comparison between PRI and PSI is not statistically significant.

Finally, if we analyse the Full Scale IQs calculated with the new method, they range between 1 and 62 (mean = 23, standard deviation = 14), and there are 21 IQs off the scale, i.e. with $z < -6.66$ points (i.e. 21 Full Scale IQs are below 1).

Discussion and conclusions

The analysis of scores according to traditional transformation showed that the Down syndrome subjects in this sample showed a greater frequency of weighted scores of 1, mainly for two subtests of PSI and the two subtests of WMI. While the lowest percentage of weighted scores of 1 is in the subtest of Similarities that belongs to the VCI. In addition, 66% of all scores available for the sample (data from the file of 128 subjects for 10 subtests) correspond to a weighted score of 1. Then looking at the z points means of the sample for the subtests, they are between two and three standard deviations below average.

Using the traditional transformation, 82% of the Down syndrome subjects have an IQ <40 and otherwise not measurable, while 18% have an IQ ranging from 40 to 62. Again, the means of the four indices oscillate between 47.5 for the PRI and 53.4 for the VCI with an apparent lack of PRI compared with other indices. In summary, there is a fairly flat average trend and a profound ID of this sample.

However, this is a very particular situation in which there is also wide variability in behaviours and performance (that for this type of subject can be expressed through a range of raw scores corresponding to the minimum weighted score of 1), and under the standard scoring procedure that variability is reduced to a weighted score of 1, which erases all the differences. Many authors have highlighted a possible floor effect with the WISC-IV (Orsini *et al.* 2015; Whitaker 2005; Whitaker & Gordon 2012; Whitaker & Wood 2008) because of weighted scores of 1 being given for very low raw scores and the non-measurement of functioning below 4 standard deviations below average (IQ = 40).

However, according to Hessl *et al.* (2009, p. 34),

> the lack of sensitivity of intelligence tests in this range of functioning is typically due to relative dearth of children with intellectual disabilities of varying levels of severity in the standardization samples, and limitations in the range of difficulty of test items and tasks that prevent measurement of lower levels of ability.

Using the new method expanded by Orsini *et al.* (2015), it was possible to draw all the information from the performance to the subtest considering all the raw scores corresponding to the weighted score of 1. So, we showed the distribution of the usual standard scores in this sample of Down syndrome subjects in comparison with that for scores derived from the new method for calculating new normalised scores representing each child's actual deviation from the standardisation sample, based on the raw score descriptive statistics of the latter sample.

Using this method, it is immediately apparent that the variability of the scores for the subtest indices is wider compared with scores obtained through the traditional method. The same Full Scale IQs, with the new method, vary between 1 and 62, and there are 21 IQs out of scale that are less than 1.

Then, using this method, the variability in performance is recovered, and it is possible to observe any strengths and weaknesses both in the subtests and the indices. In particular, as regards the subtest of VCI, a point of strength for the subtest of similarities and a point of weakness in the other two subtests of vocabulary and comprehension emerge.

For the subtests that relate to the PRI, a point of strength for the Block Design subtest emerges. Relative to the subtests that belong to the other two indices (WMI and PSI), strengths and weaknesses are not apparent: the performances on the subtests are similar.

Finally, the new four indices were compared for four traditional indices showing a greater ability of the Down syndrome subject in VCI than in WMI. These results seem to find confirmation in the literature where the Down syndrome subject seems to highlight criticality in verbal short-term memory (Lanfranchi *et al.* 2004; Lanfranchi *et al.* 2012; Stavroussi *et al.* 2016).

In summary, the results of the present study confirm firstly the need to identify appropriate measures for an intellectual development assessment of individuals with Down syndrome.

It is true that the use of standardised measures enables comparisons both between groups of subjects evaluated in different socio-cultural contexts, both within the same group longitudinally evaluated. The use of standard scores is therefore preferable, but it can become problematic if in the studied population the standard scores do not prove to be adequate assessments to identify the intra-individual variability.

The use of Orsini, Pezzuti and Hulbert's method for scoring the Wechsler scales can provide a measurement that does not flatten individual differences, both on the IQ scores and both on subtests scores; on the contrary, the traditional method of computing IQ in Wechsler scales seems not valid per this type of population. Although other intelligence assessment tests (e.g. Raven Progressive Matrices, Leiter-III, KABC-II and CTONI-II) can be used, Wechsler scales are the most widely used in the context of the Italian clinical experience. Then, the use of one scale (WISC-IV), both for children with typical and atypical development, could help plan better educational programs, considering that in Italy, children with Down syndrome always attend regular classes.

However, it is important to highlight that in clinical interventions with population that present an ID, the classic intelligence tests should be complemented by other neuropsychological tests (e.g. the Tower of London, Weigl's Colour Form Sorting Test, Behaviour Rating Inventory of Executive Function, Cambridge Examination for Mental Disorders of Older People with Down's Syndrome and others with IDs), to obtain a cognitive profile (e.g. Esteba-Castillo *et al.* 2013; Garcia-Alba *et al.* 2017).

The repercussions of this are very important, because we can improve our understanding of the cognitive phenotype of Down syndrome, and we can also schedule clinical interventions that respond to the needs of individual children and adolescents with Down syndrome.

References

Davis A. S. & Escobar L. F. (2013) Early childhood cognitive disorders: Down syndrome. In: *Psychopathology of Childhood and Adolescence: A Neuropsychological Approach* (ed. A. S. Davis), pp. 569–80. Springer, New York.

Dyer-Friedman J., Glaser B., Hessl D., Johnston C., Huffman L. C., Taylor A. *et al.* (2002) Genetic and environmental influences on the cognitive outcomes of children with fragile X syndrome. *Journal American Academy Child Adolescent Psychiatry* **41**, 237–44.

Dykens E., Hodapp R. & Evans D. (2006) Profiles and development of adaptive behavior in children with Down syndrome. *Down Syndrome Research and Practice* **9**, 45–50.

Esteba-Castillo S., Dalmau-Bueno A., Ribas-Vidal N., Vila-Alsina M., Novell-Alsina R. & Garcia-Alba J. (2013) Adaptation and validation of CAMDEX-DS (Cambridge Examination for Mental Disorders of Older People with Down Syndrome and others with intellectual disabilities) in Spanish population with intellectual disabilities. *Revista de Neurologia* **57**, 337–46.

Garcia-Alba J., Esteba-Castillo S., Castellanos Lopez M. A., Rodriguez Hidalgo E., Ribas Vidal N., Moldenhauer Diaz F. *et al.* (2017) Validation and normalization of the Tower of London-Drexel University 2nd Edition in adult population with intellectual disability. *The Spanish Journal of Psychology* **20**, 1–4. https://doi.org/10.1017/sjp.2017.30.

Glaser B., Hessl D., Dyer-Friedman J., Johnston C., Wisbeck J., Taylor A. *et al.* (2003) Biological and environmental contributions to adaptive behavior in fragile X syndrome. *American Journal Medicine Genetic* **117**, 21–9.

Glue P. & Patterson T. (2009) Can drug treatments enhance learning in subjects with intellectual disability?. *Australian & New Zealand Journal of Psychiatry* **43**, 899–904.

Hessl D., Dyer-Friedman J., Glaser B., Wisbeck J., Barajas R. G., Taylor A. *et?al.* (2001) The influence of environmental and genetic factors on behavior problems and autistic symptoms in boys and girls with fragile X syndrome. *Pediatrics* **108**, e88–e88.

Hessl D., Nguyen D. V., Green C., Chavez A., Tassone F., Hagerman R. J. *et al.* (2009) A solution to limitations of cognitive testing in children with intellectual disabilities: the case of fragile X syndrome. *Journal of Neurodevelopment Disorders* **1**, 33–45.

Lanfranchi S., Baddeley A., Gathercole S. & Vianello R. (2012) Working memory in Down syndrome: is there a dual task deficit? *Journal of Intellectual Disability Research* **56**, 157–66.

Lanfranchi S., Cornoldi C. & Vianello R. (2004) Verbal and visuo-spatial working memory deficits in children with Down syndrome. *American Journal on Mental Retardation* **109**, 456–66.

Nadel L. (1999) Down syndrome in cognitive neuroscience perspective. *Neurodevelopmental Disorders*, 197–222.

Orsini A., Pezzuti L. & Hulbert S. (2015) Beyond the floor effect on the Wechsler Intelligence Scale for Children -4th Ed.(WISC-IV): calculating IQ and Indexes of subjects presenting a floored pattern of results. *Journal of Intellectual Disability Research* **59**, 468–73.

Orsini A., Pezzuti L. & Picone L. (2012) *WISC-IV. Contributo alla taratura italiana*. Giunti OS, Firenze.

Patterson T., Rapsey C. M. & Glue P. (2013) Systematic review of cognitive development across childhood in Down syndrome: implications for treatment interventions. *Journal of Intellectual Disability Research* **57**, 306–18.

Pennington B. F., Moon J., Edgin J., Stedron J. & Nadel L. (2003) The neuropsychology of Down syndrome: evidence for hippocampal dysfunction. *Child development* **74**, 75–93.

Stavroussi P., Andreou G. & Karagiannopoulou D. (2016) Verbal fluency and verbal short-term memory in adults with Down syndrome and unspecified intellectual disability. *International Journal of Disability, Development and Education* **63**, 122–39.

Vicari S. & Carlesimo G. A. (2006) Short-term memory deficits are not uniform in Down and Williams syndromes. *Neuropsychology review* **16**, 87–94.

Wechsler D. (1991) *Wechsler Intelligence Scale for Children, Third Edition*. The Psychological Corporation, San Antonio, TX.

Wechsler, D. (2008). *Wechsler Intelligence Scale for Children, Forth Edition*. [Administration and Scoring Manual]. Pearson, San Antonio, TX.

Wechsler, D. (2012). *WISC-IV Wechsler Intelligence Scale for Children Quarta Edizione*. [Manuale di somministrazione e scoring]. Firenze. Giunti OS.

Whitaker S. (2005) The uses of the WISC-III and the WAIS-III with people with a learning disability: Three concerns. In: *Clinical Psychology Forum* (Vol. 50, pp. 37–40). The British Psychological Society.

Whitaker S. & Wood C. (2008) The Distribution of Scaled Scores and Possible Floor Effects on the WISC-III and WAIS-III. *Journal of Applied Research in Intellectual Disabilities* **21**, 136–41.

Whitaker S. & Gordon S. (2012) Floor effects on the WISC-IV. *International Journal of Developmental Disabilities* **58**, 111–19.

Accepted 1 April 2018