

# Exploring the Validity and Reliability of the WISC-IV: A Review of the Literature

Vasilios Andrikopoulos (Corresponding author)

University of Patras, Greece

Received: January 10, 2021    Accepted: November 2, 2021    Published: November 17, 2021

doi: 10.5296/jsss.v8i2.18166

URL: <https://doi.org/10.5296/jsss.v8i2.18166>

## Abstract

This study seeks to examine the psychometric properties, namely the validity and reliability as well as the overall psychometric quality of the WISC-IV. For this purpose, a systematic review of the literature was carried out. Data analysis revealed that the fourth edition of the Wechsler scale for children is more sophisticated in form and content, in line with the modern approaches and familiar models of intelligence and the measurement of mental abilities. However, research in the field of psychometric quality of this test does not give a clear picture of its interpretive power or its contribution to diagnostic evaluation. Despite its relative utilization in differential diagnostic and diagnostic assessment procedures, there is a strong criticism regarding its structural validity and the models on the basis of which it is explained, as well as the dominant structure emerging from WISC-IV. Over time, the four-factor model seems to be abandoned and its analysis oriented towards a five-factor model, in line with the CHC theory of intelligence and cognitive abilities. All in all, this study enriches our theoretical and practical understanding about the WISC-IV giving rise to other studies in this field.

**Keywords:** validity, reliability, psychometric quality, WISC-IV

## 1. Introduction

In the late 1930s, Wechsler began to have an indelible effect on the field of psychological evaluation by creating tests that measured psychological characteristics, such as intelligence. The purpose of creating a test battery was to extract clinical information from a series of cognitive projects, with the idea that the use of tests should serve something more than psychometric purposes (Flanagan & Kaufman, 2004).

The first edition of the Wechsler scale (Wechsler, 1949) was an adaptation of some of the tests that made up the Wechsler-Bellevue intelligence scale (Wechsler, 1939), which

included several new tests designed specifically for the purpose and objectives of the new scale. The test was organized into two subscales: The Verbal Scale and the Performance Scale providing scores, which concern three indicators, the Verbal IQ (VIQ), the performance IQ (PIQ) and the Full-Scale IQ (FSIQ). Each successive version adjusts the test to compensate for the Flynn effect (Flynn, 1984, 1987, 1999; Matarazzo, 1972; O'Keefe & Rodgers, 2020).

Therefore, subsequent revisions and improvements include changes to the questions to reduce bias against minorities and women, and up-to-date material to make the test more user-friendly and easier to administer. A revised version was published in 1974 as WISC-R (Wechsler, 1974) including the same tests. However, the age range of the test changed from 5-15 to 6-16 years. The third edition was published in 1991 (WISC-III, Wechsler, 1991) and brought with it a new measurement for Processing Speed. In addition to the traditional VIQ, PIQ and FSIQ scores, four new scores have been introduced representing more specific areas of cognitive function, such as the ones following: The Verbal Comprehension Index (VCI), the Perceptual Organization Index (POI) the Freedom from Distractibility Index (FDI) and the Processing Speed Index (PSI) (Koulakoglou, 2017).

The fourth edition of WISC, namely the WISC-IV, was published in 2003, and included 15 subtests, 10 of which were the main array. Thus, the main ten tests were organized in such a way, that arise four indicators (Verbal Comprehension, Perceptual Reasoning, Working Memory, Speed Processing), which in turn made up the final IQ (FS-IQ) (Flanagan & Kaufman, 2004; Wechsler, 2003). Since WISC-IV has not only been used in the American clinical research and practice but also in other cultural environments, apart from the American one, this study creates special interest, as it highlights the continuity in clinical research and forms a favorable framework for conducting wider research efforts aimed at improving the terms and conditions of diagnostic evaluation.

In addition, a number of changes, such as the reconsideration of the distinction between practical and verbal intelligence, the addition of new subtests as well as its application for the differential diagnosis of students with disorders requiring early intervention, models of interpretation and evaluation of intellectual ability-functionality in combination with similar theories of intelligence, attract the interest of researchers in the latest versions of the Wechsler scale for children leading to the study and examination of WISC-IV and WISC-V. Given that WISC-V is the latest version of this scale, our study seeks to present key issues of validity and reliability of WISC-IV through a systematic review of the existing literature, in order to show to what extent the weak and strong points led to the revision and further modification of this scale and its succession by the WISC-V.

Thus, the structure of the present study is formulated as follows: The next (second) section summarizes the structure and content of WISC-IV. The third section is dedicated to the standardization process, norms, and reliability of the WISC-IV. In the fourth section the validity of this scale is examined, as shown by the results of studies used in the diagnostic evaluation of children with specific disorders, in order to determine its applicability and usefulness. Then, research results are presented for the use of this scale in cultural contexts other than the American one, in which it was constructed and developed as well as applied for the

differential diagnosis and diagnosis of difficulties and disorders. Finally, a summary of the main findings of this study is attempted in the form of findings from the literature research.

## **2. WISC-IV Standardization, Norms and Reliability**

The WISC-IV was standardized using a sample of 2,200 children selected to match the results of a 2002 U.S. census taking into account variables such as age, gender, geographical origin, ethnicity and socioeconomic status (parents' education). The standardization sample was divided into 11 age groups, each of which included 200 children, and into 2 groups based on gender. Standardized data were collected between August 2001 and October 2002. The norm tables were divided into 4-month age intervals. Thus, it is argued that the correlation between standardization data and the US population was exemplary, as the differences between stratification variables and the population were less than two percentage points (Flanagan, Alfonso, Mascolo, & Hale, 2011).

In terms of reliability, the internal consistency scores were .94 for the VCI index, .92 for the PRI index, .92 for the WMI index, .88 for the PSI index, and .97 for the final IQ index (FS-IQ). Internal consistency values for all age-based subtests ranged between .72 for Coding (6-7 years) and .94 for Vocabulary (for 15 years). The average internal consistency for all subtests is between .79 (Symbols and Cancellation) and .90 (Letter-Number Sequence). The overall score for WISC-IV (FS-IQ) and the synthesis of lower-order indicators in terms of reliability is particularly high (> .90+) based on age range, while the reliability of each subtest is moderate (.80 -.89) (Kaufman, Flanagan, Alfonso, & Mascolo, 2006).

The test-retest reliability (mean = 32 days) for the overall index and the sub-indices in a sample of 243 children aged 6-16 years was high to moderate for the five age groups. Also, the WISC-IV is considered a stable tool with a test-retest reliability coefficient of .93, .89, .86, and .93 for each of the following indicators VCI, PRI, WMI, PSI, and FS-IQ, respectively. In general, the effect of the review practice is stronger for ages 6-7 and decreases with age (Flanagan & Kaufman, 2004).

## **3. Validity of WISC-IV**

The validity of the WISC-IV (relative to a criterion) is also supported by correlations with other already tested, valid measurement tools. WISC-IV FS-IQ was correlated with the corresponding WISC-III, WPPSI-III and WAIS-III FS- IQs (.89) as well as FS-IQ-4 by WASI. In this regard, it is noted that WISC-IV also shows highly satisfactory convergent and divergent validity, as shown by the correlation with previous indicators of the Wechsler scale. In this way, mean correlation of VCI was found to be .83. Similarly, the PRI index has an average correlation of 0.76 with the other visual-perceptual measures of the previous Wechsler scales. The validity of WISC-IV was also further investigated by examining its relationship to WIAT-II. The correlations between the components of FS-IQ and WIAT-II ranged from 0.75 (Spoken Speech) to 0.78 (Reading and Mathematics), indicating that WISC-IV FS-IQ explains 56% to 60% of variation in these areas of achievement. The correlation between the overall grade of FS-IQ and WIAT-II is .87

(explaining 76% of the variance), and is about as high as the correlation between WISC-IV FS-IQ and FS- IQs of other Wechsler scales (.89) (Canivez , Watkins, James, Good, & James, 2014; Giudicessi, Ibarra, Visconti, Zenit, & Pelayo, 2019; Styck & Watkins, 2014).

The structural validity of the test is presumed by the factor analysis presented in the Technical and Interpretive Manual of WISC-IV, which provides evidence of reliability and structural validity for the structure of the four factors resulting from exploratory factor analysis. The analyses support the structure of the basic subtests and a combination of basic and complementary subtests. Picture Concepts alone cannot charge in the defined factor (Kaufman et al., 2006).

The structural validity of the four indicators also appears to be supported by confirmatory factor analysis (CFA), although validity is increased when only the basic subtests are included in the analysis. The results that show that there is a good fit to the data (goodness-of-fit) range from 0.96 to 0.98, while the addition of optional subtests reduces the indications of good fit by values of 0.90 to 0.95 (Flanagan et al., 2011).

It is worth noting that the results of the confirmatory factor analysis are not reported in the WISC-IV Technical and Interpretive Manual. Thus, no information is provided on the stability or variability of the factors based on age nor the charges and correlations of the confirmatory factor analysis, making unclear the nature of the cognitive structures under study. Predominant can be considered the model of the four factors in the interpretation of data from normal samples but also with children with learning difficulties or those fall in neuro-psychological disorders (Bodin, Pardini, Burns, & Stevens, 2009; Flanagan et al., 2011; Prifitera, Saklofske, Weiss, & Rolfhus, 2005).

However, a five-factor model based on the Cattell-Horn-Carroll theory (CHC) and the homonymous model concerning intelligence and its theoretical framework is suggested by Keith, Fine, Taub, Reynolds, & Kranzler (2006). According to this model, WISC-IV measures crystallized intelligence (Gc), visual processing (Gv), flow reasoning (Gf), short-term memory (Gsm), and processing speed (Gs) (Alfonso, Flanagan, & Radwan, 2005; Keith et al., 2006). Similar findings are reported in content validity studies by Flanagan & Kaufman (2004) and Flanagan, Ortiz, Alfonso, & Dynda (2006). Therefore, exploratory factor analysis yields four factors, while confirmatory factor analysis by Keith et al. (2006) supports the existence of five factors.

The standard scores listed in the WISC-IV Technical and Interpretive Manual for all subtests extend **two** standard deviations above and below the average. Thus, it is argued that it can be used in the evaluation of individuals belonging to the category of the gifted but also the mentally retarded (Chen & Zhu, 2012; Kaufman et al., 2006). This manual also includes scores of the test for specific groups, in order to provide information on its specialization and clinical utility in the diagnostic evaluation process. The special groups studied include children with autism spectrum disorder, Asperger syndrome, specific language disorder and mixed receptive and expressive language disorder, ADHD and children with learning disabilities. Samples were also used with children with moderate or mild mental retardation, gifted children, children with brain damage and mobility problems (Hebben,

2004).

Fiorello et al. (2007) used this data to fragment the FSIQ and examine the unique and common variance of the four WISC-IV indicators in predicting reading performance and mathematics from the WIAT-II and found that the FSIQ consisted mainly of the unique and not a common factor variation in children with language-specific disorder (SLD), ADHD and brain damage (TBI). Because the unique factor *g* found during exploratory factor analysis appears to be composed of different cognitive constructs, and because these dimensions were able to predict reading and mathematical performance differently, the FSIQ interpretation seems to be of little use for diagnosis and intervention (Fiorello et al., 2007; Flanagan et al., 2006).

Flanagan et al. (2011) in turn indicated that the data presented in the manual of the WISC-IV cannot be used in the diagnostic evaluation of learning difficulties, because all indicators are near average and no indicator differs substantially from the other. Such a pattern is incompatible with the learning disabilities and deficits with which it is associated. The researchers interpret these results by identifying the cause in the heterogeneous nature of the participating group, which therefore deprives them of the ability to distinguish between the mean performance of the VCI, PRI, WMI and PSI scores for all groups of learning disabilities.

Also, for the same reasons, there is difficulty in distinguishing children with ADHD and without ADHD using WISC-IV. In addition, WISC-IV cannot distinguish average children from those with a low rate of brain injury. Children with mobility problems are expected to have difficulties with projects that require visual-motor coordination (Drawing with Cubes) and that involve the element of time in the execution of the task (PSI Index). Finally, WISC-IV is not considered a particularly suitable test for people with language disorders (Fenollar-Cortés, Navarro-Soria, González-Gómez, & García-Sevilla, 2015; Gomez, Vance, & Watson, 2016).

#### **4. WISC-IV Survey Results in Other Countries (Except US)**

Gygi, Hagmann-von Arx, Schweizer, & Grob (2017) investigated the predictive validity of WISC-IV in a sample of German students attending regular school. The results of the research showed that this tool has a high predictive validity and can predict the students' language and mathematical performance of the next three grades. Likewise, in the research conducted by Golay, Reverte, Rossier, Favez, & Lecerf (2012) attempting to detect the structural validity of WISC-IV it was found out that the French version of WISC-IV is more accurately described by the five-factor model, as proposed by Keith et al. (2006). The researchers argue that factor analysis points to the CHC model, reinforcing the idea that the PRI index should be divided into two sub-indices (*Gf* and *Gv*). Also, the Arithmetic subtest charges only the *Gsm* factor and not the *Gf*.

Similarly, Martínez (2015) found that WISC-IV has structural validity, as evidenced by the factor analysis performed, which seems to support the five-factor model as the most appropriate. This finding seems to be consistent with the results of the factor analysis



concerning the French version of the test (Golay et al., 2013; Reverte et al., 2014), the British version (Canivez, James, James, & Good, 2013) and the corresponding American version that emerged from the administration of the test to people with neurological deficits (Bodin, 2009).

Gomez et al. (2016) argued that for people with ADHD it is preferable to take into account the overall IQ rather than the performance in each subtest, as the confirmatory factor analysis of the research found increased support for the two-factor model (overall index intelligence on the one hand and individual indicators on the other). In fact, it is pointed out that the final IQ can predict the performance in mathematics and reading of children with ADHD.

Similarly, in the study of Fenollar-Cortés et al. (2015) the usefulness of WISC-IV in the identification and diagnosis of children with ADHD found was confirmed. In particular, the clinical samples had significantly lower scores on the Working Memory and Processing Speed indices and therefore highlighted a characteristic cognitive profile of children with ADHD, which can be revealed by administering this test. Similar findings emerge in the research of Ünal et al. (2020), with an additional low scoring performance of children with ADHD in the Index of Verbal Comprehension (Similarities) and Perceptual Reasoning (Matrix Reasoning) compared to the group of children without ADHD.

A study by Watkins & Smith (2013) attempted to investigate the control-retest reliability of the WISC-IV using a sample of 344 children with disabilities and found that there were no significant differences in the scores regarding the subtests as opposed to those in level of indicators and the overall index, where differences of the order of 10 points were recorded. The researchers argue that this differentiation indicates a lack of consistency or consistency in children's WISC-IV scoring performance over long periods of time.

Other research points to the relative superiority of WISC-IV in recognizing gifted children, as it offers these children the opportunity to participate in tests of reasoning, abstract thinking, and verbal ability, while its visualized elements reduce the emphasis on response time, in contrast to WISC-III (Erden, Yiğit, Çelik, & Guzey, 2020; Molinero, Mata, Calero, Garc ía-Mart ín, & Araque-Cuenca, 2015; Rimm, Gilman, & Silverman, 2008).

## **5. Discussion-Conclusion**

The present study attempted to investigate the validity, reliability, and the overall psychometric quality of WISC-IV through a systematic review of the relevant literature. From the passage of the theoretical discourse but also of the research data and results, specific findings and remarks can emerge in relation to the fourth edition of the Wechsler scale for children. In particular, it can be asserted that the fourth edition of the Wechsler scale for children is more sophisticated in form and content, in line with modern approaches and familiar models of intelligence and the measurement of intellectual ability.

In this way, the contribution of WISC-IV to clinical research and practice is considered particularly important, to the extent that as a psychometric tool it seems to allow the recognition of gifted children and the diagnostic evaluation of children with ADHD. It also has increased predictive validity, as it can predict the language and math performance of

children with and without ADHD in the American environment as well as in other cultural contexts. It also shows relatively high reliability both overall as a scale and in the individual subscales that make it up.

However, research in the field of psychometric quality of this test does not give a clear picture of its interpretive power or its usefulness in diagnostic evaluation. Despite its relative utilization in differential diagnostic and diagnostic evaluation procedures, there is a strong criticism regarding its structural validity and the models on the basis of which it is explained, as well as the dominant structure emerging from WISC-IV. Over time, the model of four factors seems essentially abandoned and analysis oriented towards a model of five factors being in line with the CHC theory of intelligence and cognitive ability.

As the CHC model is considered more appropriate for describing WISC-IV and interpreting the results obtained from its administration, this version and format of the test has been revised and replaced by the Fifth Edition of the Wechsler Scale for Children (WISC-V), which is more in line with the theory of intelligence mentioned above (CHC) (Canivez, Watkins, & Dombrowski, 2017; Dombrowski, Canivez, Watkins, & Beaujean, 2015).

In conclusion, our study through the presentation, analysis and discussion of issues related to the validity and reliability of WISC-IV can be a springboard for conducting larger-scale research in the field of psychometric and psychological evaluation giving further impetus to study of intelligence and mental capacity. Thus, through presenting the structure, composition, organization and content of this tool as well as highlighting the strengths and weaknesses of the scale in question and its usefulness and usability, which results from and ends with the basic characteristics of the scale WISC - IV, i.e. its validity, reliability and practicality in clinical research and practice, this work can contribute to the reframing, reconsideration of scales as psychodiagnostic tools providing a reflective approach to WISC-IV and the later version of that scale, namely the WISC-V as well.

## References

- Alfonso, V. C., Flanagan, D. P., & Radwan, S. (2005). The impact of Cattell-Horn-Carroll (CHC) theory on test development and the interpretation of cognitive and academic abilities. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp.185-202). New York: Guilford.
- Bodin, D., Pardini, D. A., Burns, T. G., & Stevens, A. B. (2009). Higher order factor structure of the WISC-IV in a clinical neuropsychological sample. *Child Neuropsychology*, *15*(5), 417-424. <https://doi.org/10.1080/09297040802603661>
- Canivez, G. L., Watkins, M. W., James, T., Good, R., & James, K. (2014). Incremental validity of WISC-IV UK factor index scores with a referred Irish sample: Predicting performance on the WIAT-II UK. *British Journal of Educational Psychology*, *84*(4), 667-684. <https://doi.org/10.1111/bjep.12056>
- Canivez, G. L., Watkins, M. W., & Dombrowski, S. C. (2017). Structural validity of the Wechsler Intelligence Scale for Children-Fifth Edition: Confirmatory factor analyses with the

16 primary and secondary subtests. *Psychological Assessment*, 29(4), 458-472. <https://doi.org/10.1037/pas0000358>

Chen, H., & Zhu, J. (2012). Measurement invariance of WISC-IV across normative and clinical samples. *Personality and Individual Differences*, 52(2), 161-166. <https://doi.org/10.1016/j.paid.2011.10.006>

Dombrowski, S. C., Canivez, G. L., Watkins, M. W., & Beaujean, A. A. (2015). Exploratory bifactor analysis of the Wechsler Intelligence Scale for Children-Fifth Edition with the 16 primary and secondary subtests. *Intelligence*, 53, 194-201. <https://doi.org/10.1016/j.intell.2015.10.009>

Erden, G., Yiğit, İ., Çelik, C., & Guzey, M. (2020). The diagnostic utility of the Wechsler Intelligence Scale for Children-(WISC-IV) in identification of gifted children. *The Journal of General Psychology*, 1-20. <https://doi.org/10.1080/00221309.2020.1862038>

Flanagan, D., Alfonso, V., Mascolo, J., & Hale, J. (2011). The Wechsler Intelligence Scale for Children-Fourth Edition in Neuropsychological Practice. In Davis, A. (Ed.), *Handbook of Pediatric Neuropsychology* (pp. 397-414). New York: Springer.

Flanagan, D. P., Ortiz, S. O., Alfonso, V. C., & Dynda, A. (2006). Integration of response to intervention and norm-referenced tests in learning disability identification: Learning from the tower of Babel. *Psychology in the Schools*, 43, 807-825. <https://doi.org/10.1002/pits.20190>

Fenollar-Cortés, J., Navarro-Soria, I., González-Gómez, C., & García-Sevilla, J. (2015). Cognitive profile for children with ADHD by using WISC-IV: Subtype differences?. *Revista de Psicodidáctica*, 20(1), 157-176. <https://doi.org/10.1387/RevPsicodidact.12531>

Fiorello, C. A., Hale, J. B., Holdnack, J. A., Kavanagh, J. A., Terrell, J., & Long, L. (2007). Interpreting intelligence test results for children with disabilities: Is global intelligence relevant? *Applied Neuropsychology*, 14, 2-12. <https://doi.org/10.1080/09084280701280338>

Flanagan, D. P., & Kaufman, S. (2004). *Essentials of assessment with WISC-IV*. New York, NY: Wiley.

Flynn, J. R. (1984). The mean IQ of Americans: Massive gains 1932 to 1978. *Psychological Bulletin*, 95(1), 29-51. <https://doi.org/10.1037/0033-2909.95.1.29>

Flynn, J. R. (1987). Massive IQ gains in 14 nations: What IQ tests really measure. *Psychological Bulletin*, 101(2), 171-191. <https://doi.org/10.1037/0033-2909.101.2.171>

Flynn, J. R. (1999). Searching for justice: The discovery of IQ gains over time. *American Psychologist*, 54(1), 5-20. <https://doi.org/10.1037/0003-066X.54.1.5>

Gygi, J. T., Hagmann-von Arx, P., Schweizer, F., & Grob, A. (2017). The predictive validity of four intelligence tests for school grades: A small sample longitudinal study. *Frontiers in Psychology*, 8, 375-364. <https://doi.org/10.3389/fpsyg.2017.00375>

Golay, P., Reverte, I., Rossier, J., Favez, N., & Lecerf, T. (2013). Further insights on the French WISC-IV factor structure through Bayesian structural equation modeling.



*Psychological Assessment*, 25(2), 496-508. <https://doi.org/10.1037/a0030676>

Gomez, R., Vance, A., & Watson, S. D. (2016). Structure of the Wechsler Intelligence Scale for Children-Fourth Edition in a group of children with ADHD. *Frontiers in Psychology*, 7(737), 1-11. <https://doi.org/10.3389/fpsyg.2016.00737>

Giudicessi, A., Ibarra, C., Visconti, V., Zenit, A., & Pelayo, H. (2019). C-44 Exploring the Criterion Validity of the WISC-IV in a Clinically Referred Sample of Children in Mexico. *Archives of Clinical Neuropsychology*, 34(6), 1073-1073. <https://doi.org/10.1093/arclin/acz034.206>

Hebben, N. (2004). Review of special group studies and utility of the process approach with the WISC-IV. In D. P. Flanagan & A. S. Kaufman (Eds.), *Essentials of WISC-IV assessment* (pp. 183-199). New York: Wiley.

Kaufman, A. S., Flanagan, D. P., Alfonso, V. C., & Mascolo, J. T. (2006). Test review: Wechsler intelligence scale for children, (WISC-IV). *Journal of Psychoeducational Assessment*, 24(3), 278-295. <https://doi.org/10.1177/0734282906288389>

Keith, T. Z., Fine, J. G., Taub, G. E., Reynolds, M. R., & Kranzler, J. H. (2006). Higher order, multi-sample, confirmatory factor analysis of the Wechsler Intelligence Scale for Children-Fourth Edition: What does it measure. *School Psychology Review*, 35(1), 108-127. <https://doi.org/10.1080/02796015.2006.12088005>

Koulakoglou, K. (2017). *Psychometric and psychological evaluation*. Athens: Patakis (In Greek).

Mart ínez, A. B. (2015). Identification of two SLI profiles through WISC-IV, CELF-4 and FON. *Journal of Educational Psychology*, 3(2), 45-79. <https://doi.org/10.20511/pyr2015.v3n2.78>

Matarazzo, J. D. (1972). *Wechsler's measurement and appraisal of adult intelligence*. Baltimore: Williams & Wilkins.

Molinero, C., Mata, S., Calero, M. D., Garc ía-Mart ín, M. B., & Araque-Cuenca, A. (2015). Usefulness of WISC-IV in determining intellectual giftedness. *The Spanish Journal of Psychology*, 18, 1-10. <https://doi.org/10.1017/sjp.2015.63>

O'Keefe, P., & Rodgers, J. L. (2020). The Flynn effect can become embedded in tests: How cross-sectional age norms can corrupt longitudinal research. *Intelligence*, 82, 1-9. <https://doi.org/10.1016/j.intell.2020.101481>

Prifitera, A., Saklofske, D. H., Weiss, L. G., & Rolfhus, E. (2005). The WISC-IV in the clinical assessment context. In A. Prifitera, D. H. Saklofske, & L. G. Weiss (Eds.), *WISC-IV clinical use and interpretation: Scientist-practitioner perspectives* (pp. 3-32). New York: Elsevier. <https://doi.org/10.1016/B978-012564931-5/50002-5>

Reverte, I., Golay, P., Favez, N., Rossier, J., & Lecerf, T. (2014). Structural validity of the Wechsler Intelligence Scale for Children (WISC-IV) in a French-speaking Swiss sample.

*Learning and Individual Differences*, 29, 114-119.

<https://doi.org/10.1016/j.lindif.2013.10.013>

Rimm, S., Gilman, B. J., & Silverman, L. K. (2008). Non-traditional applications of traditional testing. In J. Van Tassel-Baska (Ed.), *Alternative assessments with gifted and talented students* (pp. 175-202). Waco, Texas: Prufrock Press. <https://doi.org/10.4324/9781003232988-10>

Styck, K. M., & Watkins, M. W. (2014). Discriminant validity of the WISC-IV culture-language interpretive matrix. *Contemporary School Psychology*, 18(3), 168-177. <https://doi.org/10.1007/s40688-014-0021-y>

Ünal, D., Çiçek, N. M., Çak, T., Sakarya, G., Artik, A., Karaboncuk, Y., ... Kültür, E. Ç. (2020). Comparative analysis of the WISC-IV in a clinical setting: ADHD vs. non-ADHD. *Archives de P édiatrie*, 1-7. <https://doi.org/10.1016/j.arcped.2020.11.001>

Watkins, M. W., Canivez, G. L., James, T., James, K., & Good, R. (2013). Construct validity of the WISC-IVUK with a large referred Irish sample. *International Journal of School & Educational Psychology*, 1(2), 102-111. <https://doi.org/10.1080/21683603.2013.794439>

Watkins, M. W., & Smith, L. G. (2013). Long-term stability of the Wechsler Intelligence Scale for Children-Fourth Edition. *Psychological Assessment*, 25(2), 477-483. <https://doi.org/10.1037/a0031653>

Wechsler, D. (1939). *Wechsler-Bellevue intelligence scale*. New York: The Psychological Corporation. <https://doi.org/10.1037/t06871-000>

Wechsler, D. (1949). *Wechsler intelligence scale for children*. New York: The Psychological Corporation.

Wechsler, D. (1974). *Wechsler intelligence scale for children-revised*. New York: The Psychological Corporation.

Wechsler, D. (1991). *The Wechsler intelligence scale for children-third edition*. San Antonio, Texas: The Psychological Corporation.

Wechsler, D. (2003). *The Wechsler intelligence scale for children-fourth edition*. London: Pearson. <https://doi.org/10.1037/t15174-000>

### Copyright Disclaimer

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).