Overview

This technical report is the second in a series intended to introduce the Wechsler Intelligence Scale for Children—Fourth Edition (WISC–IV). Technical Report #1 presents the theoretical structure and test blueprint for the WISC–IV, as well as subtest changes from Wechsler Intelligence Scale for Children—Third Edition (WISC–III); Technical Report #2 presents the psychometric properties of WISC–IV; Technical Report #3 addresses the instrument’s clinical validity.

Neurocognitive models of information processing provides the basis for the new structure of the WISC–IV, which replaces the traditional Verbal IQ (VIQ)/Performance IQ (PIQ) dichotomy. The index scores that were supplemental in WISC–III are now primary, and each has been enhanced according to contemporary research. The names of two indices have been updated in order to more accurately reflect the content measured by subtests introduced in this revision. Detailed information is provided in the WISC–IV Technical and Interpretive Manual.

WISC–IV Standardization Sample

The WISC–IV standardization sample was representative of the U.S. population of children age 6–16. The stratified random sampling plan utilized the following variables, based on U.S. Bureau of the Census data from the March 2000 census: age, sex, race, parent education level, and geographic region. The standardization sample for the WISC–IV included 2,200 children who were divided by age into eleven groups, each consisting of 200 children. The figure below shows the demographic characteristics of the total sample.

Figure 1. Demographic Characteristics of the Standardization Sample Compared to the U.S. Population

- Parent Education Level
- Race/Ethnicity
- Geographic Region

![Graphs showing demographic characteristics of the standardization sample compared to the U.S. population.]

Evidence of Reliability: Internal Consistency

The evidence of internal consistency reliability for the normative sample was obtained using the split-half method for all subtests except the speeded tasks (Coding, Symbol Search, and Cancellation), for which the test-retest coefficients were used. Table 1 compares the average internal consistency reliability coefficients across ages for the WISC–III and the WISC–IV subtest and composite scales. The reliability coefficients for the WISC–IV composite scales range from .88 (Processing Speed) to .97 (Full Scale). The reliability coefficients of the WISC–IV composite scales are identical to or slightly better than corresponding composite scales in the WISC–III.

That these results can be appropriately generalized is supported by information obtained from the special and clinical samples. Evidence of reliability was obtained utilizing the split-half method from a sample of 661 children from 16 special and clinical groups. Table 1 provides average internal consistency reliability coefficients of subtests for these groups. The majority of the subtest reliability coefficients across special groups are similar to or higher than those coefficients reported for the normative sample, suggesting that the WISC–IV is equally reliable instrument for assessing children who are developing typically and children with clinical diagnoses.

Evidence of Reliability: Test-Retest Stability

The evidence of the WISC–IV’s test-retest stability for subtest and composite scales was evaluated with information obtained from a sample of 243 children. Participants were administered the WISC–IV on two separate occasions, with a mean test-retest interval of 32 days. Table 2 presents the mean subtest scaled scores and composite scores and their standard deviations. As can be seen, the data indicate that the WISC–IV scores are stable across time. The mean retest scores for all subtests are higher than the mean test scores from the first administration, with effect sizes ranging from .08 (Comprehension) to .60 (Picture Completion). In general, the test-retest gains are less pronounced for the verbal subtests than for other subtests. This analysis is presented in five separate age bands that show that the pattern of results is similar.
Factor-Analytic Studies

Exploratory and confirmatory factor analyses were conducted to evaluate the internal structure of the WISC–IV. Because the revision retained ten subtests from the WISC–III, as well as including five new subtests designed to measure similar constructs, WISC–IV was expected to measure the same four cognitive domains as the WISC–III (i.e., Verbal Comprehension, Perceptual Organization, Freedom from Distractibility, and Processing Speed; Wechsler, 1991).

The initial step in examining the factor structure of the WISC–IV was to determine if the pattern of obtained results matched the hypothesized four factor structure. The stability of the factor structure was then examined across different age groups. Finally, the predicted model was tested against alternative models using confirmatory factor analytic methods.
Table 3. Exploratory Factor Pattern Loadings for Core and Supplemental Subtests

<table>
<thead>
<tr>
<th>Ages 6:0–16:11 (N=1525)</th>
<th>Four Factor Model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Verbal Comprehension</td>
</tr>
<tr>
<td>Similarities</td>
<td>.71</td>
</tr>
<tr>
<td>Vocabulary</td>
<td>.87</td>
</tr>
<tr>
<td>Comprehension</td>
<td>.78</td>
</tr>
<tr>
<td>Information</td>
<td>.71</td>
</tr>
<tr>
<td>Word Reasoning</td>
<td>.73</td>
</tr>
<tr>
<td>Block Design</td>
<td>-.06</td>
</tr>
<tr>
<td>Picture Concepts</td>
<td>.16</td>
</tr>
<tr>
<td>Matrix Reasoning</td>
<td>-.03</td>
</tr>
<tr>
<td>Picture Completion</td>
<td>.32</td>
</tr>
<tr>
<td>Digit Span</td>
<td>.00</td>
</tr>
<tr>
<td>Letter-Number Seq.</td>
<td>.11</td>
</tr>
<tr>
<td>Arithmetic</td>
<td>.14</td>
</tr>
<tr>
<td>Coding</td>
<td>-.02</td>
</tr>
<tr>
<td>Symbol Search</td>
<td>.01</td>
</tr>
<tr>
<td>Cancellation</td>
<td>.01</td>
</tr>
</tbody>
</table>

Note: See WISC–IV Technical and Interpretive Manual for complete table and further discussion.

Exploratory Factor Analysis

Table 3 presents the factor analysis results for the core and supplemental subtests for all ages. The four factor structure is clearly observed and the primary loading of each subtest is found on the expected factor. On the Verbal Comprehension factor, a small secondary loading across ages is observed for Picture Completion. To further examine the stability of the factor structure, the data in four age bands (6–7, 8–9, 10–11, 12–13, 14–16) were then analyzed separately. With minor variations, each age band supported the overall four factor structure. The factor loading for Picture Concepts for ages 6–7 was evenly split between Verbal Comprehension and Perceptual Reasoning. The response processes of younger children on this task may require more verbal mediation than in older children. Above age 11, Arithmetic—clearly a working memory subtest at all age bands—has a small secondary loading on Verbal Comprehension and on Perceptual Reasoning. At age 10 and below, Information exhibited a small secondary loading on the Working Memory factor.

The results from the Confirmatory Factory Analysis of the total sample verified that, as compared to reasonable alternative models, the four-factor model is clearly fits the data best; this finding was also consistent across the four age bands.

Evidence of Validity: Relationships to Other Measures

Both the WISC–IV and the WISC–III were administered in counterbalanced order to 244 children from ages 6–16; the test-retest interval was 5 to 42 days. Table 4 presents the means, standard deviations, corrected and uncorrected correlations, and standard differences.

The corrected correlation between the WISC–III VIQ and WISC–IV VCI is .87 and .74 between the WISC–III PIQ and the WISC–IV PRI. The lower correlation between PIQ and PRI reflects important changes made to this composite in WISC–IV. Tasks that were primarily visual and spatial (Object Assembly, Picture Completion, Picture Arrangement, and Coding) were replaced with fluid reasoning tasks (Matrix Reasoning, Picture Concepts), making the PRI a stronger measure of fluid reasoning than the PIQ; for this reason, a moderate correlation was expected. The WISC–III FSIQ and the WISC–IV FSIQ correlate highly (r = .89).

As anticipated, the older WISC–III norms provided slightly inflated estimates for today’s children. The overall difference between the WISC–III and WISC–IV FSIQ scores is 2.5 points, with WISC–III scores the higher of the two. Of the WISC–III scores, processing speed tasks showed the most inflation; the least inflation was observed on the working memory tasks. As Table 4 shows, the WISC–III PSI mean is 5.5 points higher then the WISC–IV PSI mean and the WISC–III FDI mean is 1.5 points higher than the WISC–IV WMI mean. These results are consistent with the Flynn Effect (Flynn, 1984, 1987).
WISC–IV in Comparison to WISC–III

Table 5 provides the expected ranges of the WISC–IV Composite scores for selected WISC–III IQ and Index scores. These ranges are relatively narrow near the middle of the IQ score distribution (i.e., 100) and wider at the upper and lower score levels. The WISC–IV and WISC–III PRI and PIQ scores can be expected to differ more than VCI and VIQ due to the changes in the construct measured by PRI as compared to PIQ. Note that these are 95% confidence intervals based on a non-clinical sample; special education and other clinical children may fall outside of these ranges when they are retested.

Practitioners should keep in mind that these are average differences, and that an individual child who has been administered the WISC–III and is retested with the WISC–IV may score more or less than 2.5 points lower on the WISC–IV FSIQ, as compared to his or her previous WISC–III scores. When retesting clinical or special education students, many factors can contribute to score differences (for example, the compound effects of the disorder or disability with increased educational and environmental demands as the child ages).

Summary

This technical report presents some of the basic psychometric properties of the WISC–IV, including information about the demographically-representative standardization sample, internal consistency reliability, test-retest stability, factor structure, and the correlations with and score differences from the WISC–III. The interested reader is referred to the WISC–IV Technical and Interpretive Manual for further information.

References


