**Differences Required for Significance When Each WISC-V Subtest**

**Scaled Score is Compared to the Mean Subtest Scaled Score**

 When considering strengths and weaknesses among WISC-V scaled scores, examiners may wish to compare subtest scores to the mean of the subtest scores in a group including the subtest of interest. The tables on the following pages, prepared by Ron Dumont, provide the critical values for significance at the .05 and .01 levels.

 For example, consider these WISC-V Verbal Comprehension subtest scaled scores obtained by Kate. Depending on the examiner's attention span, Kate might have taken 2, 3, or 4 Verbal Comprehension subtests, but in any event, these would have been her scores, the sums of her scores, and the means for each group of 2, 3, or 4 subtests.

|  |  |
| --- | --- |
|  | Scaled Scores and Differences from Means |
| Subtest | ss | diff | ss | diff | ss | diff | ss | diff |
| Similarities | 12 |  | 12 |  | 12 |  | 12 |  |
| Vocabulary  | 11 |  | 11 |  | 11 |  | 11 |  |
| Information |  |  |  6 |  |  |  |  6 |  |
| Comprehension |  |  |  |  | 11 |  | 11 |  |
| SUM | 23 |  | 29 |  | 34 |  | 40 |  |
| MEAN | 11.50 |  | 9.67 |  | 11.33 |  | 10.00 |  |

Now the examiner compares each subtest to the mean of all of the Verbal Comprehension subtests taken by Kate. We have shown 2, 3, and 4 subtests, but in reality, the examiner would be dealing with only one column because the examiner would have administered only 2, or 3, or 4 specific subtests.

Kate's Similarities score of 12 was 0.55 scaled score points above her mean of 11.50 for Similarities and Vocabulary. If the examiner had also administered Information, Kate's Similarities score of 12 would have been 2.33 points above her mean of 9.67 for Similarities, Vocabulary, and Information. If the examiner had also administered Comprehension but not Information, Kate's Similarities score of 12 would have been 0.67 points above her mean of 11.33 for Similarities, Vocabulary, and Comprehension. If the examiner had administered all four verbal subtests, Kate's Similarities score of 12 would have been 2.00 points above her mean of 9.67 for Similarities, Vocabulary, Information, and Comprehension. Again, only one of these groups would actually have been used; the other groups would be ignored.

|  |  |
| --- | --- |
|  | Scaled Scores and Differences from Means |
| Subtest | ss | diff | ss | diff | ss | diff | ss | diff |
| Similarities | 12 | +0.55 | 12 | +2.33 | 12 | +0.67 | 12 | +2.00 |
| Vocabulary  | 11 | –0.55 | 11 | +1.33 | 11 | –0.33 | 11 | +1.00 |
| Information |  |  |  6 | –3.67 |  |  |  6 | –4.00 |
| Comprehension |  |  |  |  | 11 | –0.33 | 11 | +1.00 |
| SUM | 23 | (0) | 29 | (0) | 34 | (0) | 40 | (0) |
| MEAN | 11.50 |  | 9.67 |  | 11.33 |  | 10.00 |  |

 The following table is copied from the upper, left corner of Ron Dumont's table for "All Ages." There are also separate tables for each year of age (e.g., 6:0:0 through 6:11:30). These are the critical values for the differences between subtest scores and the means of their group of 2, 3, or 4 subtests.

**Differences Required for Significance**

|  |  |
| --- | --- |
|   | VCI, VSI, FRI, WMI, PSI (2, 3, and 4 subtests) |
| Subtest | .05 | .01 | .05 | .01 | .05 | .01 | .05 | .01 |
| Similarities | 1.33 | 1.66 | 1.95 | 2.39 | 2.01 | 2.47 | 2.33 | 2.82 |
| Vocabulary  | 1.33 | 1.66 | 1.91 | 2.35 | 1.98 | 2.43 | 2.28 | 2.75 |
| Information |   |  | 2.04 | 2.50 |   |  | 2.48 | 2.99 |
| Comprehension |   |   |   |   | 2.29 | 2.82 | 2.78 | 3.35 |

 The critical values are given for significance at the .05 level (a difference too great to occur just by random variation more than 5 times in 100) and the .01 level (a difference too great to occur just by random variation more than 1 time in 100). If the difference between a subtest scaled score and its group mean is at least as large as the critical value, the difference is statistically significant (unlikely to occur just by chance) at the .05 or .01 level.

 The table below shows the differences that were statistically significant. Differences that are not significant could have occurred just by random variation and should never be considered differences at all. We should never, ever report that, "There was a non-significant difference between *this* and *that*." If the difference was not significant it should not be considered a difference.

 If Kate took only Similarities and Vocabulary, neither score differed significantly from her mean score for those two subtests. If Kate took Similarities, Vocabulary, and Information, her Similarities score was significantly higher (p < .05) and her Information score was significantly lower (p < .01) than her mean score for the three subtests. If, instead, Kate took Similarities, Vocabulary, and Comprehension, none of her subtest scores differed significantly from her mean score for those three subtests. However, if Kate took Similarities, Vocabulary, Information, and Comprehension, her Information score was significantly lower (p < .01) than her mean score for the four subtests and none of the other three subtests differed significantly from her four-subtest mean. The examiner would have an indication that Kate might, for some reason, have a limited fund of general information despite otherwise adequate verbal abilities. The next step would be to seek other information to refute or support this hypothesis. If the hypothesis turned out to be supported and not refuted by other data, then the examiner could explore possible causes for, consequences of, and interventions to help with that limitation.

|  |  |
| --- | --- |
|  | Scaled Scores and Differences from Means |
| Subtest | ss | diff | ss | diff | ss | diff | ss | diff |
| Similarities | 12 | +0.55 | 12 | +2.33 | 12 | +0.67 | 12 | +2.00 |
|  |  |  |  | p < .05 |  |  |  |  |
| Vocabulary  | 11 | –0.55 | 11 | +1.33 | 11 | –0.33 | 11 | +1.00 |
|  |  |  |  |  |  |  |  |  |
| Information |  |  |  6 | –3.67 |  |  |  6 | –4.00 |
|  |  |  |  | p < .01 |  |  |  | p < .01 |
| Comprehension |  |  |  |  | 11 | –0.33 | 11 | +1.00 |
|  |  |  |  |  |  |  |  |  |

 In real life, there would be only one set of scores since the examiner would have given only one specific set of 2, 3, or 4 subtests. If, for example, the examiner had given all four verbal subtests (which would have been a good idea for Kate's evaluation), the table would look like this.

|  |
| --- |
| Verbal Scaled Scores and Differences from Mean |
| Subtest | ScaledScore | Differencefrom Mean | Significance |
| Similarities | 12 | +2.00 |  |
| Vocabulary  | 11 | +1.00 |  |
| Information |  6 | –4.00 | p < .01 |
| Comprehension | 11 | +1.00 |  |
| Sum | 40 |  |  |
| Mean | 10.00 |  |  |