**Rivier University**

**Department of Education**

**Certification Program for Specialist in Assessment of Intellectual Functioning**

**Sample Explanations of Classification Labels for Test Scores**

"It is customary to break down the continuum of IQ test scores into categories. . . . other reasonable systems for dividing scores into qualitative levels do exist, and the choice of the dividing points between different categories is fairly arbitrary. It is also unreasonable to place too much importance on the particular label (e.g., 'borderline impaired') used by different tests that measure the same construct (intelligence, verbal ability, and so on)." [Roid, G. H. (2003). *Stanford-Binet Intelligence Scales, Fifth Edition, Examiner's* Manual. Itasca, IL: Riverside, p. 150.]

*A.  Use the various classification terms supplied by the test publishers and keep explaining why the same number gets different names. (I cannot find in any Wechsler manual any classification labels for subtest scaled scores, but everyone seems to use them anyhow.) Please see pages 4 and 8 for some examples.*

First The various tests that Ecomodine took use different classification schemes to describe

Mention: test scores. Therefore the same test score may be called different names on different tests. For example, a standard score of 110 is called "Average" on some tests, "High Average" on others, and "Above Average" on yet others. Verbal labels may seem to exaggerate small differences, as when 109 is "Average" and 110 is "High Average" or "Above Average" on some tests. Please see p. i of the Appendix for a complete description of the various sets of classification labels used with Ecomodine's tests.

Follow-up: Please remember that different tests use different verbal labels for the same scores. Please see p. i of the Appendix.

*B.  Pick one classification system and use it for all tests and keep explaining that these are not necessarily the same names used by the publishers of the various tests.*

*B.1. Stanines (page 5):*

First The various tests that Mordred took use different classification schemes to describe

Mention: test scores, so the same test score may confusingly be called different names on different tests. Therefore, I have taken the liberty of using stanine classification labels for all test scores in this report. Stanines 1, 2, and 3 (Very Low, Low, and Below Average) are the lowest 23% of students' test scores. Stanines 4, 5, and 6 (Low Average, Average, and High Average) are the middle 54%. The highest 23% of students' scores are called stanines 7, 8, and 9 (Above Average, High, and Very High). These are **not** the various classification labels provided with the different tests. Please see page i of the Appendix for a complete description of stanines and p. ii for a complete description of the various classification labels recommended by the publishers of the tests taken by Mordred.

Follow-up: Please remember that I am using stanine classification labels for Mordred's test scores. These are not the publishers' recommended labels. Please see pages i and ii of the Appendix. (*Use this parenthetical note or footnote frequently.*)

*Note*: *I usually provide in text or in an abbreviated table copied into the text from the appendix a standard score (Wechsler standard score, scaled score, T score, z-score, BOT-2 scaled score,* v*-scale score, or other), a percentile rank, or both along with the stanine. I always list the standard scores, confidence bands, and percentile ranks in the Appendix.* ***Be sure always to call stanines "stanines," not "ranges" or anything else!*** *"Mordred's score was in the Low Average stanine (4)." "Mordred's score was Below Average (stanine 3)." "Mordred's score was stanine 6 (High Average)."* ***Stanines badly upset some people.***

*B.2. Use one test's classification scheme (e.g., Woodcock-Johnson) for all of the tests.*

First The various tests that Quatherynne took use different classification schemes to describe

Mention: test scores, so the same test score may confusingly be called different names on different tests. Therefore, I have taken the liberty of using Woodcock-Johnson classification labels for all test scores in this report: Very Low, Low, and Low Average are the lowest 25% of students' test scores. The middle 50% of students' scores are called Average. The highest 25% are called High Average, Superior, and Very Superior. Please see page i of the Appendix for a complete description of Woodcock-Johnson classifications and the various other classification labels recommended by the publishers of the tests taken by Quatherynne.

Follow-up: Please remember that I am using Woodcock-Johnson classification labels for Quatherynne's test scores. These are not the labels recommended by the publishers of the other tests taken by Quatherynne. Please see page i of the Appendix.

*B.3. Use a classification scheme lifted and perhaps modified from a highly reputable source.*

First The various tests that Ralph took use different classification schemes to describe

Mention: test scores, so the same test score may confusingly be called different names on different tests. Therefore, I have taken the liberty of using classification labels recommended by Jerome M. Sattler (*Assessment of children: Cognitive foundations* (5th ed.). La Mesa, CA: Jerome M. Sattler, Publisher) for all test scores in this report. Please see page i of the Appendix for a complete description of these classifications and the various other classification labels recommended by the publishers of the tests taken by Ralph.

Follow-up: Please remember that I am using Professor Sattler's recommended classification labels for Ralph's test scores. These are not the labels recommended by the publishers of the other tests taken by Ralph. Please see page i of the Appendix.

*C.  Avoid names for scores altogether (perhaps simply reporting only percentiles)*.

First In this report, Quatherynne's various test scores (please see page i of the Appendix for a

Mention: complete description of standard, scaled, and other test scores) are also reported as percentile ranks. A percentile rank tells the percentage of students the same age (or in the same grade) who scored the same as Quatherynne or lower. For example a percentile rank of 37 would mean that Quatherynne scored as high as or higher than 37 percent of students his age and lower than the other 63 percent.

*D. Completely eschew test scores and verbal labels and simply describe how the student functioned on each task.*

Calpurnia's test scores, explanations of the test statistics, and descriptions of the tests she took may be found in the Appendix.

*E. Do something different and explain that.*

**Remember to Explain Tests and Scores When They First Appear in the Report**

You do not want to bewilder and alienate your reader with unexplained scores in the History section at the beginning of your report.

Mortitia was tested when she was in third grade in 2011. Her scores included an NCE of 50 for Comprehension on the GMRT, a T score of 50 for Pattern Construction on the DAS-II, a PR of 50 on the ROCF, a score of 15 for Balance on the BOT-2, a score of 36 for Visual Association on the 1968 ITPA, a score of 10 for Atlantis on the KABC-II, a PRI score of 100 on the WISC-IV, and a score of 500 on the TTFC.

If that summary does not stop readers dead in their tracks, what will? What are the chances that most readers would know that all of those scores were in the 50th percentile? What are the chances that the reader will continue reading?

Footnotes can help. Sometimes. Maybe. I tend to include (for the benefit of advocates, attorneys, hearing officers, judges, and opposing expert witnesses) more technical information than Prof. Sattler, for example, advocates, but at least I relegate most of it to footnotes or parenthetical comments used once, the first time the term appears. From then on, I feel free to simply use the previously defined term (except for periodically reminding the reader about my use of Option A, B, C, D, or E above and perhaps parenthetically defining "percentile rank" every few pages, because that is such an essential concept).

 This type of standard score ranges from about 40 to about 160 with an average score of 100. The middle half of students' standard scores falls between 90 and 110. Please see pages i and ii.

 Scaled scores range from 1 to 19 or 20 with an average score of 10. The middle half of students' scaled scores falls between 8 and 12. Please see pages i and ii.

 T scores range from 20 to 80 with an average score of 50. The middle half of students' T scores fall between 43 and 57. Please see pages i and ii.

 *V-*scale scores range from 1 to 24 with an average score of 15. The middle half of students' scaled scores falls between 13 and 17. Please see pages i and ii.

1 Bruininks-Oseretsky subtest scores range from 1 to 30 with an average score of 15. The middle half of students' scores falls between 12 and 18.

 Normal Curve Equivalents (NCEs) range from 1 to 99 with an average NCE of 50. The middle half of students' NCEs falls between 36 and 64.

 Percentile ranks tell the percentage of students of the same age or grade whose scores Mortitia tied or exceeded. For example, a percentile rank of 36 would mean that Mortitia scored as high as or higher than 36 percent of peers and lower than the other 64 percent. Please see pages i and ii.

 Stanines range from 1 (Very Low) to 9 (Very High). The middle half (actually 54%) of students' scores falls in the broad average range or stanines 4, 5, and 6 (Low Average, Average, and High Average). These are **not** the publisher's classification labels. Please see pages i and ii.

 Test scores can never be perfectly accurate. This confidence band shows how much scores are likely to vary 90% of the time just by random variation.

 Test scores can never be perfectly accurate. Lucky and unlucky guesses, narrowly beating or exceeding time limits, and other random influences make scores on even the best tests less than totally reliable. A *significant difference* is one that is too large to occur by pure chance more than a certain percent of the time (in this report, 5% or p < .05) unless there were a real difference between the tested abilities. Even significant differences, however, are not necessarily uncommon. Human abilities normally vary. An *uncommon difference* is one that did not occur very frequently (in this report, less than 10 percent of the time) among the examinees whose scores were used to create the norms for the test.

 Most tests increase in difficulty from the lowest-numbered to the highest-numbered items and require that the student be given enough easier items ("basal") and enough difficult items ("ceiling") to cover the range of the student's strengths and weaknesses.

**Ecomodine's Scores for her Age on the Wechsler Intelligence Scale for Children Integrated, 4th ed. (WISC-IV) and Wechsler Individual Achievement Test, 3rd ed. (WIAT-III)**

**in Scaled and Standard Scores, Percentile Ranks, and Wechsler Classifications4**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Subtests** | TestScore[[1]](#footnote-1) | 90% Confi- dence[[2]](#footnote-2) | Per-cen-tile[[3]](#footnote-3) | Classification (for Index scores; I have also used it for scaled scores)[[4]](#footnote-4) |
| **WISC-IV** |  |  |  |  |
| **Verbal Subtests** |  |  |  |  |
| explaining how two different things could be similar (SI) |  7 |  5 – 9  | 16 | Low Average |
| defining vocabulary words (VC) |  11 |  9 – 13  | 63 | Average  |
| answering questions of social and practical comprehension (CO) |  9 |  7 – 11  | 37 | Average |
| **Verbal Comprehension total score (SI VC CO)** |  **95** |  **90 – 101** | **37** | **Average** |
| **Perceptual Reasoning Subtests** |  |  |  |  |
| copying geometric designs with patterned cubes (BD) |  8 |  6 – 10  | 25 | Average |
| choosing one picture from each group to illustrate a concept (PCn) |  10 |  8 – 12  | 50 | Average |
| completing multiple-choice, logical matrix puzzles (MR) |  9 |  7 – 11  | 37 | Average |
| **Perceptual Reasoning total score (BD PCn MR)** |  **94** |  **88 – 101** | **34** | **Average** |
| **General Ability Index (SI, VS, CO, BD, PCn, MR)** |  **94** |  **90 – 99**  | **34** | **Average** |
| **Working Memory Subtests** |  |  |  |  |
| repeating series of dictated digits forward and backward (DS) |  5 |  3 – 8  | 05 | Borderline |
| repeating digits and letters digits first, then letters (LN) |  6 |  4 – 8  | 09 | Low Average |
| **Working Memory total score (DS LN)** |  **74** |  **70 – 83**  | **04** | **Borderline** |
| **Processing Speed Subtests** |  |  |  |  |
| speed of transcribing a digit-symbol code on paper (CD) |  6 |  4 – 8  | 09 | Low Average |
| speed of finding matching symbols in rows of symbols.(SS) |  4 |  2 – 6  | 02 | Borderline |
| **Processing Speed total score (CD SS)** |  **73** |  **67 – 79**  | **03** | **Borderline** |
| **Cognitive Proficiency Index (DS LN CD SS)** |  **71** |  **67 – 79**  | **03** | **Borderline** |
| **Full Scale total (SI VC CO BD PCn MR DS LN CD SS)** |  **81** |  **77 – 86**  | **10** | **Borderline** |
| **WIAT-III** |  |  |  |  |
| Reading Comprehension | 108 |  98 – 118  | 70 | Average |
| Word Reading |  89 |  85 – 93  | 23 | Average |
| Pseudoword Decoding  |  90 |  86 – 94  | 25 | Average |
| Oral Reading Fluency |  85 |  79 – 91  | 16 | Average |
| Total Reading Composite |  **89** |  **85 - 93** | **23** | **Average** |
| Sentence Composition | 109 | 100 – 118  | 73 | Average |
| Essay Composition | 107 |  98 – 116  | 68 | Average |
| Spelling |  85 |  80 – 90  | 16 | Average |
| **Written Expression Composite** |  **99** |  **93 – 105**  | **47** | **Average** |
| Math Problem Solving | 108 | 101 – 115 | 70 | Average |
| Numerical Operations |  89 |  84 – 94  | 23 | Average |
| Mathematics Composite |  **98** |  **93 – 103**  | **45** | **Average** |
| **Math Fluency Composite** |  **85** |  **79 – 91**  | **16** | **Average** |
| Listening Comprehension | 106 |  96 – 116 | 66 | Average |
| Oral Expression |  96 |  87 – 105  | 39 | Average |
| **Oral Language Composite** | **101** |  **94 – 108** | **53** | **Average** |

**Ecomodine's Scores for her Age on the Wechsler Intelligence Scale for Children Integrated, 4th ed. (WISC-IV)**

**and Wechsler Individual Achievement Test, 3rd ed. (WIAT-III)**

**in Scaled and Standard Scores, Percentile Ranks, and Stanine Classifications8**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Subtests** | TestScore[[5]](#footnote-5) | 90% Confi- dence[[6]](#footnote-6) | Per-cen-tile[[7]](#footnote-7) | Stanine[[8]](#footnote-8)123456789 |
| **WISC-IV** |  |  |  |  |
| **Verbal Subtests** |  |  |  |  |
| explaining how two different things could be similar (SI) |  7 |  5 – 9  | 16 |  3 Below Average  |
| defining vocabulary words (VC) |  11 |  9 – 13  | 63 |  6 High Average  |
| answering questions of social and practical comprehension (CO) |  9 |  7 – 11  | 37 |  4 Low Average  |
| **Verbal Comprehension total score (SI VC CO)** |  **95** |  **90 – 101** | **37** |  **4 Low Average**  |
| **Perceptual Reasoning Subtests** |  |  |  |  |
| copying geometric designs with patterned cubes (BD) |  8 |  6 – 10  | 25 |  4 Low Average  |
| choosing one picture from each group to illustrate a concept (PCn) |  10 |  8 – 12  | 50 |  5 Average  |
| completing multiple-choice, logical matrix puzzles (MR) |  9 |  7 – 11  | 37 |  4 Low Average  |
| **Perceptual Reasoning total score (BD PCn MR)** |  **94** |  **88 – 101** | **34** |  **4 Low Average**  |
| **General Ability Index (SI, VS, CO, BD, PCn, MR)** |  **94** |  **90 – 99**  | **34** |  **4 Low Average**  |
| **Working Memory Subtests** |  |  |  |  |
| repeating series of dictated digits forward and backward (DS) |  5 |  3 – 8  | 05 |  2 Low |
| repeating digits and letters digits first, then letters (LN) |  6 |  4 – 8  | 09 |  2 Low |
| **Working Memory total score (DS LN)** |  **74** |  **70 – 83**  | **04** |  **2 Low** |
| **Processing Speed Subtests** |  |  |  |  |
| speed of transcribing a digit-symbol code on paper (CD) |  6 |  4 – 8  | 09 |  2 Low |
| speed of finding matching symbols in rows of symbols.(SS) |  4 |  2 – 6  | 02 | 1 Very Low |
| **Processing Speed total score (CD SS)** |  **73** |  **67 – 79**  | **03** | **1 Very Low** |
| **Cognitive Proficiency Index (DS LN CD SS)** |  **71** |  **67 – 79**  | **03** | **1 Very Low** |
| **Full Scale total (SI VC CO BD PCn MR DS LN CD SS)** |  **81** |  **77 – 86**  | **10** |  **2 Low** |
| **WIAT-III** |  |  |  |  |
| Reading Comprehension | 108 |  98 – 118  | 70 |  6 High Average  |
| Word Reading |  89 |  85 – 93  | 23 |  4 Low Average  |
| Pseudoword Decoding  |  90 |  86 – 94  | 25 |  4 Low Average  |
| Oral Reading Fluency |  85 |  79 – 91  | 16 |  3 Below Average |
| **Total Reading Composite** |  **89** |  **85 - 93** | **23** |  **4 Low Average**  |
| Sentence Composition | 109 | 100 – 118  | 73 |  6 High Average  |
| Essay Composition | 107 |  98 – 116  | 68 |  6 High Average  |
| Spelling |  85 |  80 – 90  | 16 |  3 Below Average |
| **Written Expression Composite** |  **99** |  **93 – 105**  | **47** |  **4 Low Average**  |
| Math Problem Solving | 108 | 101 – 115 | 70 |  6 High Average  |
| Numerical Operations |  89 |  84 – 94  | 23 |  4 Low Average  |
| **Mathematics Composite** |  **98** |  **93 – 103**  | **45** |  **5 Average**  |
| **Math Fluency Composite** |  **85** |  **79 – 91**  | **16** |  **3 Below Average** |
| Listening Comprehension | 106 |  96 – 116 | 66 |  6 High Average  |
| Oral Expression |  96 |  87 – 105  | 39 |  4 Low Average  |
| **Oral Language Composite** | **101** |  **94 – 108** | **53** |  **5 Average**  |

**STANINE CLASSIFICATIONS OF SCORES ON TESTS TAKEN BY ECOMODINE**

When a new test is developed, it is *normed* on a *sample* of hundreds or thousands of people. The sample should be like that for a good opinion poll: female and male, urban and rural, different parts of the country, different income levels, etc. The scores from that norming sample are used as a yardstick for measuring the performance of people who then take the test. This human yardstick allows for the difficulty levels of different tests. The student is being compared to other students on both difficult and easy tasks. You can see from the illustration below that there are more scores in the middle than at the very high and low ends. Many different scoring systems are used, just as you can measure the same distance as 1 yard, 3 feet, 36 inches, 91.4 centimeters, 0.91 meter, or 1/1760 mile.

**PERCENTILE RANKS (PR)** simply state the percent of persons in the norming sample who scored the same as or lower than the student. A percentile rank of 63 would be high average – as high as or higher than 63% and lower than the other 37% of the norming sample. It would be in Stanine 6. The middle half of scores falls between percentile ranks of 25 and 75.

**STANDARD SCORES** ("quotients" on some tests) have an average (*mean)* of 100 and a *standard deviation* of 15. A standard score of 105 would also be at the 63rd percentile rank. Similarly, it would be in Stanine 6. The middle half of these standard scores falls between 90 and 110.

**SCALED SCORES** ("standard scores" on some tests) are standard scores with an average (*mean)* of 10 and a *standard deviation* of 3. A scaled score of 11 would also be at the 63rd percentile rank and in Stanine 6. The middle half of these standard scores falls between 8 and 12.

**STANINES** (standard nines) are a nine-point scoring system. Stanines 4, 5, and 6 are approximately the middle half of scores, or average range. Stanines 1, 2, and 3 are approximately the lowest one fourth. Stanines 7, 8, and 9 are approximately the highest one fourth. Throughout this report, for all of the tests, I am using the stanine labels shown below (Very Low, Low, Below Average, Low Average, Average, High Average, Above Average, High, and Very High), **even if the particular test may have a different labeling system in its manual**.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  There are  | 200 **&**s, so |  |  | **&&&&&** |  |  |  |  |
|  |  Each **&&** | = 1 % |  |  **&&&&&&**  | **&&&&&&&** |  **&&&&&&** |  |  |  |
|  |  |  |  **&&&** | **&&&&&&&** | **&&&&&&&** | **&&&&&&&** |  **&&&** |  |  |
|  |  |  | **&&&&&&&** | **&&&&&&&** | **&&&&&&&** | **&&&&&&&** | **&&&&&&&** |  |  |
|  |  **&** | **&&&&&&&** | **&&&&&&&** | **&&&&&&&** | **&&&&&&&** | **&&&&&&&** | **&&&&&&&** | **&&&&&&&** |  **&**  |
|  | **&&&&&&&** | **&&&&&&&** | **&&&&&&&** | **&&&&&&&** | **&&&&&&&** | **&&&&&&&** | **&&&&&&&** | **&&&&&&&** | **&&&&&&&** |
|  |  |  |  |  |  |  |  |  |  |
| **Stanine** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** |
|  | **Very** |  | **Below** | **Low** |  | **High** | **Above** |  | **Very** |
|  | **Low** | **Low** | **Average** | **Average** | **Average** | **Average** | **Average** | **High** | **High** |
|  | 4% | 7% | 12% | 17% | 20% | 17% | 12% | 7% | 4% |
| Percentile | 1 – 4 | 4 - 11 | 11 – 23 | 23 - 40 | 40 – 60 | 60 - 77 | 77 - 89 | 89 - 96 | 96 -99  |
| Standard Score |  - 73 | 74 - 81 | 82 – 88 | 89 - 96 | 97 – 103 | 104 - 111 | 112- 118 | 119 - 126 |  127 -  |
| Scaled Score | 1 - 4  |  5 6 | 7 |  8 9 | 10 |  11 12 | 13 |  14 15 |  16 - 19  |

Adapted from Willis, J. O. & Dumont, R. P., *Guide to identification of learning disabilities* (1998 New York State ed.) (Acton, MA: Copley Custom Publishing, 1998, p. 26). Also available at http://alpha.fdu.edu/psychology/test\_score\_descriptions.htm.

**PUBLISHERS' CLASSIFICATIONS OF SCORES ON TESTS TAKEN BY ECOMODINE**

When a new test is developed, it is *normed* on a *sample* of hundreds or thousands of people. The sample should be like that for a good opinion poll: female and male, urban and rural, different parts of the country, different income levels, etc. The scores from that norming sample are used as a yardstick for measuring the performance of people who then take the test. This human yardstick allows for the difficulty levels of different tests. The student is being compared to other students on both difficult and easy tasks. You can see from the illustration below that there are more scores in the middle than at the very high and low ends. Many different scoring systems are used, just as you can measure the same distance as 1 yard, 3, feet, 36 inches, 91.4 centimeters, 0.91 meter, or 1/1760 mile.

**PERCENTILE RANKS (PR)** simply state the percent of persons in the norming sample who scored the same as or lower than the student. A percentile rank of 50 would be Average – as high as or higher than 50% and lower than the other 50% of the norming sample. The middle half of scores falls between percentile ranks of 25 and 75.

**STANDARD SCORES** ("quotients" on some tests) have an average (*mean)* of 100 and a *standard deviation* of 15. A standard score of 100 would also be at the 50th percentile rank. The middle half of these standard scores falls between 90 and 110.

**SCALED SCORES** ("standard scores on some tests) are standard scores with an average (*mean)* of 10 and a *standard deviation* of 3. A scaled score of 10 would also be at the 50th percentile rank. The middle half of these standard scores falls between 8 and 12.

**T SCORES** have an average (*mean)* of 50 and a *standard deviation* of 10. A T score of 50 would be at the 50th percentile rank. The middle half of T scores falls between approximately 43 and 57.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  | **&& &&** |  |  |  |
|  |  There are  | 200 **&**s. |  | **&&&&&& &&&&&&** |  |  |  |
|  |  Each **&&** | = 1%. |  | **&&&&&& &&&&&&** |  |  |  |
|  |  |  |  **&&**  | **&&&&&& &&&&&&** |  **&&** |  |  |
|  |  |  | **&&&&&&** | **&&&&&& &&&&&&** | **&&&&&&** |  |  |
|  |  |  | **&&&&&&** | **&&&&&& &&&&&&** | **&&&&&&** |  |  |
|  |  |  **&** | **&&&&&&** | **&&&&&& &&&&&&** | **&&&&&&** |  **&** |  |
|  |  | **&&&&&&**  | **&&&&&&** | **&&&&&& &&&&&&** | **&&&&&&** | **&&&&&&** |  |
|  |  **& & & &**  | **&&&&&&** | **&&&&&&** | **&&&&&& &&&&&&** | **&&&&&&** | **&&&&&&** |  **& & & &** |
|  |  |  |  |  |  |  |  |
| Percent in each | 2.2% | 6.7% | 16.1% | 50% | 16.1% | 6.7% | 2.2% |
| Standard Scores | – 69 | 70 – 79 | 80 – 89 | 90 – 109 | 110 – 119 | 120 – 129 | 130 –  |
| Scaled Scores | 1 2 3 |  4 5  |  6 7 |  8 9 10 11 |  12 13 |  14 15 | 16 17 18 19 |
| Percentile Ranks | – 02 | 03 – 08 | 09 – 24 | 25 – 74 | 75 – 90 | 91 – 97 | 98 –  |
| Wechsler Classification  | Extremely Low | Borderline | LowAverage | Average | HighAverage | Superior | VerySuperior |
| WIAT-III Classification | Very Low<55 | Low55 – 69  | Below Average 70 – 84 | Average85 – 115 | Above Average116 – 130 | Super-ior131-145 | Very Super-ior 146 –  |
| Stanines | Very Low – 73  | Low 74 – 81  | Below Average 82 - 88 | Low Average89 – 96  | Average97 – 103  | High Average104 - 111 | Above Average 112 – 118 | High119 – 126  | Very High127 –  |

Adapted from Willis, J. O. & Dumont, R. P., *Guide to identification of learning disabilities* (1998 New York State ed.) (Acton, MA: Copley Custom Publishing, 1998, p. 27). Also available at http://alpha.fdu.edu/psychology/test\_score\_descriptions.htm.

**Option A**

**Ecomodine's Scores for her Age on Tests of Working Memory, Verbal Ability, and Nonverbal Ability**

**in Standard Scores, Percentile Ranks, and Test Publishers' Classifications10**

|  |  |  |  |
| --- | --- | --- | --- |
| **Test** | **Standard****Score** | **Percentile****Rank**[[9]](#footnote-9) | **Publisher's****Classification**[[10]](#footnote-10) |
| **Working Memory Tests** |  |  |  |
| **WISC-IV**: Working Memory Index | 110 | 75 | High Average |
| **DAS-II:** Working Memory Cluster | 110 | 75 | Above Average |
| **WJ III**: Working Memory Cluster | 110 | 75 | Average |
| **WRAML2**: Working Memory | 110 | 75 | (none provided in Manual) |
|  |  |  |  |
| **Verbal Ability Tests** |  |  |  |
| **WISC-IV**: Verbal Comprehension |  65 | 01 | Extremely Low |
| **DAS-II:** Verbal Ability Cluster |  65 | 01 | Very Low |
| **WJ III**: Verbal Ability Cluster |  65 | 01 | Very Low |
| **KABC-II**: Gc Scale |  65 | 01 | Lower Extreme |
| **WRAML2**: Verbal Memory |  65 | 01 | (none provided in Manual) |
| **RIAS**: Verbal Index (VIX) |  65 | 01 | Significantly Below Average |
| **KTEA-II**: Oral Language Composite |  65 | 01 | Lower Extreme |
| **WIAT-III**: Oral Language Composite |  65 | 01 | Low |
| **TOAL-4**: General Language Composite |  65 | 01 | Very Poor |
| **OWLS-II**: Listening Comprehension |  65 | 01 | Deficient |
| **SB5**: Verbal IQ |  65 | 01 | Mildly Impaired |
| **BBCS-3**: Receptive Composite |  65 | 01 | Very Delayed |
|  |  |  |  |
| **Nonverbal Ability Tests** |  |  |  |
| **WISC-IV**: Perceptual Reasoning | 133 | 99 | Very Superior |
| **DAS-II:** Spatial Ability Cluster | 133 | 99 | Very High |
| **WJ III**: Visual/Spatial Thinking (Gv) | 133 | 99 | Very Superior |
| **KABC-II**: Gv Scale | 133 | 99 | Upper Extreme |
| **WRAML2**: Visual Memory | 133 | 99 | (none provided in Manual) |
| **RIAS**: Nonverbal Index (NIX) | 133 | 99 | Significantly Above Average |
| **BOT-2**: Total Motor Composite | 133 | 99 | Well-Above Average |
| **SB5**: Nonverbal IQ | 133 | 99 | Gifted |
| **OWLS-II**: Written Expression | 133 | 99 | Exceptional |

**ADDENDUM: EXAMPLES OF VARIOUS PUBLISHERS' CLASSIFICATION SCHEMES**

When a new test is developed, it is *normed* on a *sample* of hundreds or thousands of people. The sample should be like that for a good opinion poll: female and male, urban and rural, different parts of the country, different income levels, etc. The scores from that norming sample are used as a yardstick for measuring the performance of people who then take the test. This human yardstick allows for the difficulty levels of different tests. The student is being compared to other students on both difficult and easy tasks. You can see from the illustration below that there are more scores in the middle than at the very high and low ends. Many different scoring systems are used, just as you can measure the same distance as 1 yard, 3, feet, 36 inches, 91.4 centimeters, 0.91 meter, or 1/1760 mile.

**PERCENTILE RANKS (PR)** simply state the percent of persons in the norming sample who scored the same as or lower than the student. A percentile rank of 50 would be Average – as high as or higher than 50% and lower than the other 50% of the norming sample. The middle half of scores falls between percentile ranks of 25 and 75.

**STANDARD SCORES** ("quotients" on some tests) have an average (*mean)* of 100 and a *standard deviation* of 15. A standard score of 100 would also be at the 50th percentile rank. The middle half of these standard scores falls between 90 and 110.

**SCALED SCORES** ("standard scores on some tests) are standard scores with an average (*mean)* of 10 and a *standard deviation* of 3. A scaled score of 10 would also be at the 50th percentile rank. The middle half of these standard scores falls between 8 and 12.

**T SCORES** have an average (*mean)* of 50 and a *standard deviation* of 10. A T score of 50 would be at the 50th percentile rank. The middle half of T scores falls between approximately 43 and 57.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  | **&& &&** |  |  |  |
|  |  There are  | 200 **&**s. |  | **&&&&&& &&&&&&** |  |  |  |
|  |  Each **&&** | = 1%. |  | **&&&&&& &&&&&&** |  |  |  |
|  |  |  |  **&&**  | **&&&&&& &&&&&&** |  **&&** |  |  |
|  |  |  | **&&&&&&** | **&&&&&& &&&&&&** | **&&&&&&** |  |  |
|  |  |  | **&&&&&&** | **&&&&&& &&&&&&** | **&&&&&&** |  |  |
|  |  |  **&** | **&&&&&&** | **&&&&&& &&&&&&** | **&&&&&&** |  **&** |  |
|  |  | **&&&&&&**  | **&&&&&&** | **&&&&&& &&&&&&** | **&&&&&&** | **&&&&&&** |  |
|  |  **& & & &**  | **&&&&&&** | **&&&&&&** | **&&&&&& &&&&&&** | **&&&&&&** | **&&&&&&** |  **& & & &** |
|  |  |  |  |  |  |  |  |
| Percent in each | 2.2% | 6.7% | 16.1% | 50% | 16.1% | 6.7% | 2.2% |
| Standard Scores | – 69 | 70 – 79 | 80 – 89 | 90 – 109 | 110 – 119 | 120 – 129 | 130 –  |
| Scaled Scores | 1 2 3 |  4 5  |  6 7 |  8 9 10 11 |  12 13 |  14 15 | 16 17 18 19 |
| T Scores | – 29 | 30 – 36 | 37 – 42 | 43 – 56 | 57 – 62 | 63 – 69  | 70 – |
| Percentile Ranks | – 02 | 03 – 08 | 09 – 24 | 25 – 74 | 75 – 90 | 91 – 97 | 98 –  |
| Wechsler Classification  | Extremely Low | Borderline | LowAverage | Average | HighAverage | Superior | VerySuperior |
| DAS-IIClassification | VeryLow | Low | BelowAverage | Average | AboveAverage | High | VeryHigh |
| Woodcock-Johnson Classif. | VeryLow | Low | LowAverage | Average(90 – 110) | High Average (111 – 120) | Superior(121 – 130) | Very Superior(131 – ) |
| Pro-EdClassification | VeryPoor | Poor | BelowAverage | Average | Above Average  | Superior | Very Superior |
| WIAT-III Classification | Very Low<55 | Low55 – 69  | Below Average 70 – 84 | Average85 – 115 | Above Average116 – 130 | Super-ior131-145 | Very Super-ior 146 –  |
| PPVT-4 Classifications | Extremely Low | Moderately Low | Low | High | Moderately High | Extremely High |
| Average |
| Stanines | Very Low – 73  | Low 74 – 81  | Below Average 82 - 88 | Low Average89 – 96  | Average97 – 103  | High Average104 - 111 | Above Average 112 – 118 | High119 – 126  | Very High127 –  |

Adapted from Willis, J. O. & Dumont, R. P., *Guide to identification of learning disabilities* (1998 New York State ed.) (Acton, MA: Copley Custom Publishing, 1998, p. 27). Also available at http://alpha.fdu.edu/psychology/test\_score\_descriptions.htm.

**ADDENDUM: MORE EXAMPLES OF VARIOUS PUBLISHERS' CLASSIFICATION SCHEMES**

*V-***SCALE SCORES** have a *mean* of 15 and *standard deviation* of 3. A *v-*scale score of 15 would also be at the 50th percentile rank and in Stanine 5. The middle half of *v-*scale scores falls between 13 and 17.

**BRUININKS-OSERETSKY (BOT-2)** subtest scores have a *mean* of 15 and *standard deviation* of 5. A BOT-2 subtest score of 15 would be at the 50th percentile rank and in Stanine 5. The middle half of BOT-2 subtest scores falls between 12 and 18.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  | **&& &&** |  |  |  |
|  |  There are  | 200 **&**s. |  | **&&&&&& &&&&&&** |  |  |  |
|  |  Each **&&** | = 1%. |  | **&&&&&& &&&&&&** |  |  |  |
|  |  |  |  **&&**  | **&&&&&& &&&&&&** |  **&&** |  |  |
|  |  |  | **&&&&&&** | **&&&&&& &&&&&&** | **&&&&&&** |  |  |
|  |  |  | **&&&&&&** | **&&&&&& &&&&&&** | **&&&&&&** |  |  |
|  |  |  **&** | **&&&&&&** | **&&&&&& &&&&&&** | **&&&&&&** |  **&** |  |
|  |  | **&&&&&&**  | **&&&&&&** | **&&&&&& &&&&&&** | **&&&&&&** | **&&&&&&** |  |
|  |  **& & & &**  | **&&&&&&** | **&&&&&&** | **&&&&&& &&&&&&** | **&&&&&&** | **&&&&&&** |  **& & & &** |
|  |  |  |  |  |  |  |  |
| Percent in each | 2.2% | 6.7% | 16.1% | 50% | 16.1% | 6.7% | 2.2% |
| Standard Scores | – 69 | 70 – 79 | 80 – 89 | 90 – 109 | 110 – 119 | 120 – 129 | 130 –  |
| Scaled Scores | 1 2 3 |  4 5  |  6 7 |  8 9 10 11 |  12 13 |  14 15 | 16 17 18 19 |
| *V-*Scale Scores |  1 – 8  |  9 10 | 11 12 | 13 14 15 16 |  17 18 |  19 20 | 21 – 24  |
| T Scores | – 29 | 30 – 36 | 37 – 42 | 43 – 56 | 57 – 62 | 63 – 69  | 70 – |
| z-scores | < –2.00  |  –2.00 – –1.34  | –1.33 – –0.68 | –0.67 – 0.66  | 0.67 – 1.32 | 1.33 – 1.99 | 2.00 – |
| Bruininks-Oseretsky |  0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28  |
| Percentile Ranks | – 02 | 03 – 08 | 09 – 24 | 25 – 74 | 75 – 90 | 91 – 97 | 98 –  |
| WRAT4 Classification | Lower Extreme | Low | Below Average | Average | Above Average | Superior | Upper Extreme |
| VMIClassification | VeryLow | Low | BelowAverage | Average | AboveAverage | High | VeryHigh |
| RIASClassification | Significantly Below Av. | Moderately Below Av. | BelowAverage | Average | AboveAverage | Moderately Above Av. | Significantly Above Av. |
| Stanford-Binet Classification | Moder-ately Impaired40-54 | Mildly Impaired55-69 | Borderline | LowAverage | Average | High Average | Superior | Gifted130-144 | Very Gifted 145-160 |
| Leiter Classification | Mod- -erateDelay40-54 | Very Low/ Mild Delay55-69 | Low | BelowAverage | Average | Above Average | High | Very High/Gifted |
|  Severe Delay =  30 – 39  |
| OWLS-II Classification | Deficient– 69  | Below Average 70 – 84 | Average85 – 115 | Above Average116 – 130 | Exceptional 131 –  |
| KTEA-II Classification | Lower Extreme | Below Average 70 – 84 | Average85 – 115 | Above Average116 – 130 | Upper Extreme |
| Vineland Adaptive Levels | Low– 70 | Moderately Low71 – 85  | Adequate or Average86 – 114 | Moderately High115 – 129  | High130 – |
| CELF-4 Classifications | Very Low– 70 | Low71 – 77  | Borderline78 – 85  | Average86 – 114  | Above Average115 – |
| Stanines | Very Low – 73  | Low 74 – 81  | Below Average 82 - 88 | Low Average89 – 96  | Average97 – 103  | High Average104 - 111 | Above Average 112 – 118 | High119 – 126  | Very High127 –  |

Adapted from Willis, J. O. & Dumont, R. P., *Guide to identification of learning disabilities* (1998 New York State ed.) (Acton, MA: Copley Custom Publishing, 1998, p. 27). Also available at http://alpha.fdu.edu/psychology/test\_score\_descriptions.htm.

John Willis, 2/20/14

1. These are scaled scores for subtests and standard scores for totals and factors. Please see page 7. [↑](#footnote-ref-1)
2. Test scores can never be perfectly reliable, even on the very best tests. Lucky and unlucky guesses, barely beating or missing time limits, and other random influences inevitably alter scores. This score interval shows how much scores are likely to vary 90% of the time just by pure chance. [↑](#footnote-ref-2)
3. Percentile ranks tell the percentage of students of the same age or grade whose scores Ecomodine tied or exceeded. For example, a percentile rank of 36 would mean that Ecomodine scored as high as or higher than 36 percent of peers and lower than the other 64 percent. [↑](#footnote-ref-3)
4. The WISC-IV and WIAT-III use different classification terms for standard scores. Please see page 7. The WISC-IV does not provide classification terms for scaled scores, so I used the terms for statistically equivalent standard scores. [↑](#footnote-ref-4)
5. These are scaled scores for subtests and standard scores for totals and factors. Please see pages 6 and 7. [↑](#footnote-ref-5)
6. Test scores can never be perfectly reliable, even on the very best tests. Lucky and unlucky guesses, barely beating or missing time limits, and other random influences inevitably alter scores. This score interval shows how much scores are likely to vary 90% of the time just by pure chance. [↑](#footnote-ref-6)
7. Percentile ranks tell the percentage of students of the same age or grade whose scores Ecomodine tied or exceeded. For example, a percentile rank of 36 would mean that Ecomodine scored as high as or higher than 36 percent of peers and lower than the other 64 percent. [↑](#footnote-ref-7)
8. Stanines range from 1 (Very Low) to 9 (Very High). The middle half (actually 54%) of scores falls in the broad average range or stanines 4, 5, and 6 (Low Average, Average, and High Average). These are **not** the classifications used by the publishers. Please see pages 6 and 7. [↑](#footnote-ref-8)
9. Percentile ranks tell the percentage of students of the same age or grade whose scores Ecomodine tied or exceeded. For example, a percentile rank of 36 would mean that Ecomodine scored as high as or higher than 36 percent of peers and lower than the other 64 percent. [↑](#footnote-ref-9)
10. The various tests that Ecomodine took use different classification schemes to describe test scores. Therefore the same test score may be called different names on different tests. For example, a standard score of 110 is called "Average" on some tests, "High Average" on others, and "Above Average" on yet others. Verbal labels may seem to exaggerate small differences, as when 109 is "Average" and 110 is "High Average" or "Above Average" on some tests. Please see page 7 for a complete description of the various sets of classification labels used with Ecomodine's tests. [↑](#footnote-ref-10)