**Longitudinal Study on Reading and Writing at the Word, Sentence, and Text Levels: Data Guide Book**

The purpose of this document is to serve as a guide for users looking to access and analyze the Longitudinal Study on Reading and Writing at the Word, Sentence, and Text Levels dataset. Within this document is a breakdown on variable naming, as well as some other information.

Variable naming: All variables in the document are broken down into 3-5 main components. May be some inconsistencies in terms of order of components and the presence (or lack thereof) of underscores between components.

**Variable name part 1 – Grade:** All variable names in the dataset begin with an indicator for the grade level at which the measure was assessed. There are four distinct options reported in the variable which can be viewed in the table below.

|  |  |
| --- | --- |
| **Name** | **Grade Level** |
| G1 | 1st Grade |
| G2 | 2nd Grade |
| G3 | 3rd Grade |
| G4 | 4th Grade |

Example: Any measure that was administered in grade 1 would begin with G1\_.... for grade 2 would being with G2\_..., etc.

**Variable name part 2 – Test:** The second aspect of each variables name is going to be the broad test or measure name that was used. For example, the Woodcock-Johnson III Tests of Achievement (WJAIII) is a set of tests for which many sub-measures were collected on the children in the sample. As such, for any test that was given as part of the larger WJAIII set of tests, the variable name will contain as the second part of the variable name the term “WJAIII”. The table below gives a breakdown of the broader tests that were given as a part of this study, as well as the variable naming component used.

|  |  |
| --- | --- |
| **Name** | **Test** |
| DIBELS | Dynamic Indicators of Basic Early Literacy Skills |
| StanBin | Stanford Binet Test |
| WRAT3 | Wide Range Achievement Test 3 |
| WJAIII | Woodcock-Johnson III Tests of Achievement |
| TOWRE2 | Test of Word Reading Efficiency (2nd Edition) |
| ExpPassRdg | Express Passage Reading |
| ExpPassListening | Express Passage Listening |
| WASI | Weschler Abbreviated Scale Intelligence |
| WIAT2 | Weschler Individual Achievement Test II |
| WRMT | Woodcock Reading Mastery Tests-Revised |
| MemSpan | Memory Span |
| TOSREC or ToSRE | Test of Silent Reading Efficiency and Comprehension |
| WUR |  |
| SAT10 | Stanford Achievement Test Series 10th Edition |

Example: Any measure that was given in grade 1 using a measure from the larger DIBELS series of tests would being with G1\_DIBELS\_... A measure at grade 2 using the WRAT3 would being with G2\_WRAT3\_..., etc.

**Variable name part 3-4 – Subtest/Form:** Following the broader test from which a measure was taken, the third (and sometimes fourth) component(s) of a variable name are going to specify the specific subtest, portion of a test, or testing form/version that was used. For example, within the WJAIII several subtests were given, including the Word Identification and Word Attack subtests. Given the different subtests applied, we must distinguish which subtest is reflected in each variable nam/column. For the Word Identification subtest, this is going to be designated by the term “Wid” and for the Word Attack subtest it will be designated with “Watt”. These terms will follow the broader test name.

Some variables however will have both a form/version and subtest, and will have two distinct codes within the variable name reflecting both of these aspects. This is however rare within the dataset, and intuitive in the variable naming where it takes place. As an example, the TOWRE contains a subtest focused on non-word reading which also has multiple forms. For all measure using the non-word reading subtest, the third component of the variable name (following the grade first and the broader test (TOWRE) second) will be the term “NonWd” to distinguish the non-word subtest. However, since there are multiple forms the name must further identify this. As such, in addition to specifying that subtest, the specific form used is designated following this. This then becomes a term specifying both the subtest and form, or something like “NonWd\_FormC”. Note that not all measures will have a subtest and/or form, so for some the form may be the third component of a subtest is not specified, for others the form will be the fourth term and will follow the subtest term, and for others there simply will be no form and/or subtest codes. See the table below for a breakdown of the subtests used.

|  |  |  |
| --- | --- | --- |
| **Larger Test** | **Subtest/Form Code** | **Definition** |
| DIBELS | Spring |  |
|  | IceCream |  |
|  | CheckUp |  |
| StanBin |  | This test contains no subtests |
| WRAT3 | Spell | The spelling subtest of the WRAT3 was used. \*This is the only WRAT3 Measure in the dataset. |
| WJAIII | Wid | Word Identification subtest |
|  | PComp | Passage Comprehension subtest |
|  | WAttack | Word Attack subtest |
|  | OComp | Oral Comprehension subtest |
| TOWRE2 | NWdA or NWdB | Non-word reading (Forms A or B) |
|  | WdA or WdB or WdC or WdD  SightWdA or SightWdB or SightWdC or SightWdD | Word reading (Forms A , B, C, or D). Wd used in grades 1-2, SightWd used in grades 3-4 |
| ExpPassRdg | 1Windows | Passage about windows (Grades 1 and 2) |
|  | 2Earth | Passage about Earth (Grades 1 and 2) |
|  | 1NewHome | Passage about a new home (Grade 3) |
|  | 2Oceans | Passage about oceans (Grade 3) |
|  | AndrewsFath | Passage about Andrew’s father (Grade 4) |
|  | MYTHS | Passage about myths (Grade 4) |
| ExpPassListening | 1Pierre | Passage about Pierre (Grades 1 and 2) |
|  | 2Tree | Passage about trees (Grades 1 and 2) |
|  | TheOwl | Passage about an owl (Grade 3) |
|  | Bats | Passage about bats (Grade 3) |
|  | SummerVac1 | Passage about summer vacation (Grade 4) |
|  | RAINBOWS1 | Passage about rainbows (Grade 4) |
| WASI | Vocab | Vocabulary subtest. \*Only subtest used in dataset |
| WIAT2 | WF | Fluency |
|  | Sentences |  |
|  | Para |  |
|  | WE |  |
|  | SP |  |
|  | WL |  |
|  | Alpha |  |
| WRMT | PComp or PSGComp | Passage comprehension (PComp in grades 1-2; PSGComp in grades 3-4). \*Only subtest used in dataset |
| MemSpan | A or B | Designating forms A or B were used |
| TOSREC or ToSRE | FormA or FormD | Specifies form A or D |
| WUR | 1N, 2E, 3N, or 4E | Specifies forms used |
|  | 0708, 0809, or 0910 | Specifies school year tests were given (0708 in grade 1, 0809 in grade 2, 0910 in grade 3). \*No year is listed for grade 4 measures, but this would relate to the 2010-2011 school year |
| SAT10 | Rd or \*NONE | Designates that reading was assessed. \*Only subtest given, but some variable names may have no subtest code listed |

Example: A measure taken in grade 1, from the broader WJAII test using Passage comprehension subtest will have the name G1\_WJAIII\_PComp… A measure taken from TOWRE in grade 3, using the non-word subtest, with form C, will have the name G3\_TOWRE2\_NonWd\_FormC…

Variable name final part – Measure value/score-type: Up to this point in the variable naming scheme, the variable name tells us the grade the test was administered, the broader test used, and the specific subtest and form used when applicable. The final part as such is going to specify the value that the score actually represents. That is, different tests and subtest within the dataset measure their respective constructs of interests in different ways. For some, number of words attempted/correct/incorrect is important, versus in others a numeric value will be computed. Alternatively, for measure of fluency or other timed measures, the time elapsed may be of interest. For example, the WJAII PComp subtest has a raw score, a standardized score, or a weighted score. For the variable reporting the PComp raw score, this term will end in “RS”. For one specifying standardized score, this will end in “SS”. See the table below for a breakdown of different variable endings used and what they mean. Note that there is quite a bit of overlap across the measures in the type of data reported. As such, this table does not break down the terms used here by specific measure.

|  |  |
| --- | --- |
| **Term** | **Definition** |
| TTLattempted | Total words attempted (TOWRE) |
| INCR\_words or NumINCORRECT or NUM)INCR | Number of Incorrect Words. (TOWRE) |
| Correct or corr or NumCorrect or corre | Number of words correct |
| time | Time elapsed |
| RS | Raw score |
| SS | Scaled score |
| WS | Weighted Score |
| IS | Incorrect score |
| AgeBasedRS | Age based raw score |
| AgeBasedSS | Age based scale score |
| GradeBasedRS | Grade based raw score |
| GradeBasedSS | Grade based scale score |
| AgeBased\_Q | Age adjusted score (specific to WIAT2\_WF) |
| GradeBased\_Q | Grade adjusted score (specific to WIAT2\_WF) |
| AgeBasedsumSS | Age based sum score |
| GradeBasedSumSS | Grade based sum score |

**Other Variables:** In addition to this variable naming scheme, there are also two additional variables included in the dataset. These variables are described below:

|  |  |  |
| --- | --- | --- |
| **Variable** | **Definition** | **Values** |
| Race | Participant selected race | 1-Black  2-White  3-Hispanic  4-Asian  5-Mixed/Multiple Races  6-Unknown |
| Sex | Participant selected biological sex | 1-Male  2-Female |