



# Big Words Club Assessment: Codebook

## Table of Contents

1. <a href="#">Data Set Information</a> .....	2
2. <a href="#">Variable Descriptions</a> .....	4
3. <a href="#">Summary Statistics</a> .....	14
4. <a href="#">Online Appendix</a> .....	24



## Data Set Information

This project evaluated the Big Words Club (BWC) program by conducting an RCT during the 2017-2018 school year. 47 schools participated across Arizona, Colorado, and Texas. Randomization was done at the school level (21 control and 26 treatment). The unit of observation in this data set is an individual student (818 total).

The primary outcome variable in this study was student test scores. While there was no assessment at the beginning of program usage, students were assessed at two points in time: 17 weeks after beginning the program and 25 weeks after beginning the program. These correspond to March and May of 2018 respectively.

During Assessment 1 (week 17), students took a 38-question assessment (a copy of this test is available in the online appendix), of which 19 words are commonly used words and the other 19 were ones specifically taught as a part of the BWC program. During Assessment 2 (week 25), students took this same assessment again, and additionally took a standardized PPVT-4 exam (a link to the description of this exam is in the online appendix). 818 students participated in Assessment 1, and 603 in Assessment 2. Some test scores were deemed invalid, and these are noted in the dataset via indicator variables for each of these 3 exams (A1, A2, PPVT).

As described in the next section, this data set also contains demographic variables for each student (gender, age, etc.), as well as school-level variables (Title 1 status, percentage on free or reduced lunch, etc.), and teacher level variables (teacher's years of experience, teacher's class usage of the BWC program, etc.).

[Back to Top](#)



BEHAVIORAL INSIGHTS  
AND PARENTING LAB  
THE UNIVERSITY OF CHICAGO

## Variable Descriptions

**NOTE:** For **every** variable, code of “999” indicates a **missing** value.

**participant\_id** – *Student Identification Number*

Unique ID number given to each student in the study

**teacher\_id** – *Teacher Identification Number*

Unique ID number given to each teacher in the study

**school\_id** – *School's Identification Number*

Unique ID number given to each school in the study

**treat** – *Treatment Status*

0 = Control Group

1 = Treatment Group

**age\_months** – *Student's Age in Months*

Age is as of May 2018 (Assessment 2)

**female** – *Student's Gender*

0 = Male

1 = Female

**kindergarten** – *Student's Grade Level*

0 = Pre-School

1 = Kindergarten

**private\_school** – *School Type*

0 = Public School

1 = Private School



BEHAVIORAL INSIGHTS  
AND PARENTING LAB  
THE UNIVERSITY OF CHICAGO

**title1** – *School's Title 1 Status*

- 0 = Non-Title 1 School
- 1 = Title 1 School

**free\_reduced\_lunch** – *School's Free Reduced Lunch Proportion*

This gives the percent of students in the student's school that receive free or reduced lunch. Administrative data received from district website.

**state** – *Student's State of Residence*

- 1 = Arizona
- 2 = Colorado
- 3 = Texas

**esl\_observed** – *Student's English as a Second Language Status*

Note: student's ESL status was determined based on an observation by our assessment administrator, not based on data from the school.

- 0 = Not ESL student
- 1 = ESL student

**special\_ed\_observed** – *Student's Special Education Status*

Note: student's SPED status was determined based on an observation by our assessment administrator, not based on data from the school.

- 0 = Not Special Education student
- 1 = Special Education student

**new\_student** – *New Student*

Indicates whether the student enrolled in the school after our study began.

- 0 = Not a New Student
- 1 = New Student

**distracted\_a1** – *Distracted During Assessment 1*

Our Assessment Administrator's determination of how distracted a student was during the week 17 BWC assessment.

- 0 = Not Distracted
- 1 = Mildly Distracted
- 2 = Moderately Distracted
- 3 = Extremely Distracted



**distracted\_a2** – *Distracted During Assessment 2*

Our Assessment Administrator's determination of how distracted a student was during the week 25 BWC assessment.

0 = Not Distracted

1 = Mildly Distracted

2 = Moderately Distracted

3 = Extremely Distracted

**distracted\_ppvt** – *Distracted During PPVT Assessment*

Our Assessment Administrator's determination of how distracted a student was during the week 25 PPVT assessment.

0 = Not Distracted

1 = Mildly Distracted

2 = Moderately Distracted

3 = Extremely Distracted

**score\_a1** – *Score on Assessment 1*

Student Score (out of 38 points) on BWC Assessment 1 (Week 17).

**invalid\_a1** – *Invalid Assessment 1 Score*

Our Assessor's determination of whether a score was invalid, based in part on potentially incomplete or random answers given by students.

0 = Valid Score

1 = Invalid Score

**score\_a2** – *Score on Assessment 2*

Student Score (out of 38 points) on BWC Assessment 2 (Week 25).

**invalid\_a2** – *Invalid Assessment 2 Score*

Our Assessor's determination of whether a score was invalid, based in part on potentially incomplete or random answers given by students.

0 = Valid Score

1 = Invalid Score



**score\_ppvt** – *Score on PPVT-4 Assessment*

PPVT-4 Assessment Score, given in week 25. This test is a standardized vocabulary test created by Pearson. See Online Appendix for more details.

**score\_ppvt\_age** – *Score on PPVT-4 Assessment with Age Adjustment*

PPVT-4 Assessment Score, given in week 25. This standard measure adjusts the score based on the student's age, to make scores more comparable.

**invalid\_ppvt** – *Invalid PPVT-4 Score*

Our Assessor's determination of whether a score was invalid, based in part on potentially incomplete or random answers given by students.

0 = Valid Score

1 = Invalid Score

**t\_logins\_april** – *Teacher Logins: April*

The number of times a teacher logged into the BWC program website during the month of April (in between the two Assessments).

**t\_logins\_total** – *Teacher Logins: Total*

The total number of times a teacher logged into the BWC program website during the entire course of our study.

**t\_weeks\_used** – *Weeks of Teacher BWC Use*

This was an open-ended question asked to teachers in the survey (Q6) given in May. The exact phrasing of the question:

How many weeks of the Big Word Club program have you completed with your classroom?

**t\_words\_learned** – *Words Learned*

This was a multiple-choice question asked to teachers in the survey (Q16) given in May. The exact phrasing of the question:

About how many words would you say your students learned from the Big Word Club?

0 = almost none

1 = 1 to 5

2 = 6 to 10



3 = 11 to 15

4 = 16 to 20

5 = 21 or more

**t\_financial\_struggle** – *Student's Financial Struggle*

An open-ended question asked to teachers in the survey (Q30) given in May was: How many children in each classroom you teach have families in these situations? If you do not know, take your best guess. Leave blank if classroom does not apply.

\_\_\_ Struggle to make ends meet

\_\_\_ Have just enough to get by

\_\_\_ Have enough to comfortably get by

\_\_\_ Have money for many extras

This variable is the number that the teacher answered for the **first category: Struggle to Make Ends Meet**.

**t\_misbehavior** – *Student Misbehavior*

A multiple-choice question asked to teachers in the survey (Q31) given in May was: How often does the misbehavior of children in your classroom(s) interfere with your ability to teach?

0 = Never

1 = Rarely

2 = Occasionally

3 = Frequently

Teachers gave multiple answers in case they taught 2 classes, so this variable is the **average of their two answers**.

**t\_years\_experience** – *Teacher Experience*

This was an open-ended question asked to teachers in the survey (Q36) given in May. The exact phrasing of the question:

Counting this school year, how many years have you been a school teacher (including as a lead or assistant teacher, full- or part-time)?

[Back to Top](#)



## Summary Statistics

The following table shows descriptive statistics for continuous variables in the dataset.

**Note:** that the total sample size is 818, so the number of missing observations for each variable can be inferred from its N provided in the table (first column).

	N	Mean	SD	Min	Max
age_months	818	67.77	9.121	41	103
free_reduced_lunch	818	55.57	35.48	0	100
score_a1	818	23.64	5.760	0	37
score_a2	603	25.36	5.480	7	38
score_ppvt	591	89.44	25.75	20	149
score_ppvt_age	595	99.33	15.86	27	137
t_logins_april	332	10.02	12.70	0	40
t_logins_total	332	59.92	60.60	1	196
t_weeks_used	259	19.68	6.830	3	35
t_words_learned	260	3.642	1.201	2	5
t_financial_struggle	524	29.31	28.02	0	100
t_misbehavior	556	1.820	0.755	0	3
t_years_experience	556	18.39	9.981	1	45

[Back to Top](#)





BEHAVIORAL INSIGHTS  
AND PARENTING LAB  
THE UNIVERSITY OF CHICAGO

## Online Appendix

Link to Online Appendix: <https://sites.google.com/view/bip-bwc-online-appendix/home>

The **Online Appendix** contains the following 4 items:

### 1. BWC Assessment

This document is the test that was given to the student participants in our study during both Assessment 1 and Assessment 2.

### 2. BWC Score Sheet

This document is the rubric used to score the assessments given to the student participants during Assessment 1 and Assessment 2.

### 3. Teacher Survey

This document is the exact survey given to each teacher. Responses to questions used in our analysis are included in the main dataset.

### 4. PPVT Assessment

PPVT-4 was one of the assessments we used during Assessment 2, created by Pearson. We provide a link to Pearson website's description of the PPVT-4 Test.

If there is any information or data about this project you would like but could not find in this codebook or on the online appendix please email us at [biplab@uchicago.edu](mailto:biplab@uchicago.edu).

[Back to Top](#)