

ICPSR 34437

**Gates Millennium Scholars  
Program (GMS) Cohort 2,  
2001-2006**

*Bill and Melinda Gates Foundation*

Codebook for Freshmen Cohort 2 (2001)  
Non-cognitive Scores - Restricted Use

Inter-university Consortium for  
Political and Social Research  
P.O. Box 1248  
Ann Arbor, Michigan 48106  
[www.icpsr.umich.edu](http://www.icpsr.umich.edu)

# Terms of Use

The terms of use for this study can be found at:  
<http://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/34437/terms>

## Information about Copyrighted Content

Some instruments administered as part of this study may contain in whole or substantially in part contents from copyrighted instruments. Reproductions of the instruments are provided as documentation for the analysis of the data associated with this collection. Restrictions on "fair use" apply to all copyrighted content. More information about the reproduction of copyrighted works by educators and librarians is available from the United States Copyright Office.

### NOTICE

#### WARNING CONCERNING COPYRIGHT RESTRICTIONS

The copyright law of the United States (Title 17, United States Code) governs the making of photocopies or other reproductions of copyrighted material. Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be "used for any purpose other than private study, scholarship, or research." If a user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of "fair use," that user may be liable for copyright infringement.

**ICPSR PROCESSING NOTE FOR # 34437**  
*Gates Millennium Scholars Program (GMS) Cohort 2, 2001-2006*  
***Restricted use data***

**1) Confidentiality:**

- In the restricted use data, ICPSR has blanked the all responses that contain respondent names and other identifiable characteristics to protect respondent confidentiality.
- These responses are shown in the data as "BLANKED FOR CONFIDENTIALITY - ICPSR".

**2) Question Text:**

- Question text was added to the data based on the original data collection instrument deposited by the Principal Investigator.

**3) Value labels:**

- Value labels were standardized based on the original data collection instrument deposited by the Principal Investigator.
- Any spelling errors found in the value labels were corrected during ICPSR processing.

**4) Missing Values:**

- Missing values were standardized as per the study's convention. Values "-4" and "-5" were recoded or labeled as "Logical Skip" and "Missing Data" respectively.

**5) Unknown code:**

- Any unknown codes were given the value label "Unknown Code – ICPSR".

**6) Special Considerations**

- Credits, grades and GPA were reported in a wide variety of ways in the data; no effort has been made to regularize or convert grades, GPA or credits. Users of this study should use caution when utilizing these variables in their research.

## Data Completeness Report

Notes: (1) Variables are individually listed only if they have greater than 5% missing data. These variables are listed under the appropriate percentage category in the order in which they appear in the data file. (2) The Data Completeness Report only captures information about system missing or other values that are declared missing. Codes that have a label implying that they are missing but that are not declared missing values are not reflected in this report. Data users should consult the codebook for more specific information about missing values. (3) Some variables that have 100% missing data may have been blanked by ICPSR to protect respondent confidentiality. Data users should consult the codebook for more specific information about blanked variables. (4) Data do not contain skip patterns or skip patterns are not reflected in the data as coded.

**Table 1: Distribution of Variables by Percentage of Missing Values--Freshmen Cohort 2 (2001)**

### Non-cognitive Scores - Restricted Use

Variable Name and Label (Total Cases = 2338 )	Percent of Cases with Missing Values
38.1% ( 8 of 21 variables)	have 0% Missing Values
61.9% ( 13 of 21 variables)	have 0% - 1% Missing Values
0.0% ( 0 of 21 variables)	have 1% - 3% Missing Values
0.0% ( 0 of 21 variables)	have 3% - 5% Missing Values
0.0% ( 0 of 21 variables)	have 5% - 10% Missing Values
0.0% ( 0 of 21 variables)	have 10% - 20% Missing Values
0.0% ( 0 of 21 variables)	have 20% - 40% Missing Values
0.0% ( 0 of 21 variables)	have 40% - 99% Missing Values
0.0% ( 0 of 21 variables)	have 100% missing values

# **ICPSR 34437**

## **Gates Millennium Scholars Program (GMS) Cohort 2, 2001-2006**

### **Variable Description and Frequencies**

**Note: Frequencies displayed for the variables are not weighted. They are purely descriptive and may not be representative of the study population. Please review any sampling or weighting information available with the study.**

**Summary statistics (minimum, maximum, mean, median, and standard deviation) may not be available for every variable in the codebook. Conversely, a listing of frequencies in table format may not be present for every variable in the codebook either. However, all variables in the dataset are present and display sufficient information about each variable. These decisions are made intentionally and are at the discretion of the archive producing this codebook.**

## Gates Millennium Scholars Program (GMS) Cohort 2, 2001-2006

### Freshmen Cohort 2 (2001) Non-cognitive Scores - Restricted Use

SU_ID	Respondent ID
-------	---------------

Location: 1-8 (width: 8; decimal: 0)  
 Variable Type: character  
 Question: Respondent ID  
 Based upon 2338 valid cases out of 2338 total cases.

PL_2	Preload Population Type
------	-------------------------

Location: 9-10 (width: 2; decimal: 0)  
 Variable Type: numeric  
 Question: Preload Population Type

Value	Label	Unweighted Frequency	%
1	2000 Freshmen Recipients	0	0.0 %
2	2000 Continuing Undergraduate Recipients	0	0.0 %
3	2000 Graduate Recipients	0	0.0 %
4	2001 Freshmen Recipients	1000	42.8 %
5	2000 Freshmen Non-recipients	0	0.0 %
6	2000 Continuing Undergraduate Non-recipients	0	0.0 %
7	2000 Graduate Non-recipients	0	0.0 %
8	2001 Freshmen Non-recipients	1338	57.2 %
9	2002 Freshmen Recipients	0	0.0 %
10	2002 Freshmen Non-recipients	0	0.0 %
11	2004 Freshmen Recipients	0	0.0 %
12	2004 Freshmen Non-recipients	0	0.0 %

Based upon 2338 valid cases out of 2338 total cases.

- Mean: 6.29
- Median: 8.00
- Mode: 8.00
- Minimum: 4.00
- Maximum: 8.00
- Standard Deviation: 1.98

RACE_DB	Preload Race
---------	--------------

Location: 11-12 (width: 2; decimal: 0)  
 Variable Type: numeric  
 Question: Preload Race

- Study 34437 -

<i>Value</i>	<i>Label</i>	<i>Unweighted Frequency</i>	<i>%</i>
1	African Americans	853	36.5 %
2	American Indians	329	14.1 %
3	Asian/Pacific Islanders	425	18.2 %
4	Hispanic Americans	731	31.3 %

Based upon 2338 valid cases out of 2338 total cases.

- Mean: 2.44
- Median: 2.00
- Mode: 1.00
- Minimum: 1.00
- Maximum: 4.00
- Standard Deviation: 1.27

**SEX\_DB Preload Sex/Gender**

Location: 13-14 (width: 2; decimal: 0)  
 Variable Type: numeric  
 Question: Preload Sex/Gender

<i>Value</i>	<i>Label</i>	<i>Unweighted Frequency</i>	<i>%</i>
1	Male	754	32.2 %
2	Female	1584	67.8 %
3	Unknown	0	0.0 %

Based upon 2338 valid cases out of 2338 total cases.

- Mean: 1.68
- Median: 2.00
- Mode: 2.00
- Minimum: 1.00
- Maximum: 2.00
- Standard Deviation: 0.47

**PELL\_DB Preload Pell Grant Eligibility**

Location: 15-16 (width: 2; decimal: 0)  
 Variable Type: numeric  
 Question: Preload Pell Grant Eligibility

<i>Value</i>	<i>Label</i>	<i>Unweighted Frequency</i>	<i>%</i>
1	Pell-Eligible	1001	42.8 %
2	Pell-Ineligible	424	18.1 %
3	Pell Eligible Unknown	913	39.1 %

- Study 34437 -

<i>Value</i>	<i>Label</i>	<i>Unweighted Frequency</i>	<i>%</i>
4	Not Applicable (i.e Graduate student)	0	0.0 %

Based upon 2338 valid cases out of 2338 total cases.

- Mean: 1.96
- Median: 2.00
- Mode: 1.00
- Minimum: 1.00
- Maximum: 3.00
- Standard Deviation: 0.90

**SEL\_NSEL Preload Select or No-select**

Location: 17-18 (width: 2; decimal: 0)  
 Variable Type: numeric  
 Question: Preload Select or No-select

<i>Value</i>	<i>Label</i>	<i>Unweighted Frequency</i>	<i>%</i>
1	Select	2010	86.0 %
2	Non-Select	328	14.0 %

Based upon 2338 valid cases out of 2338 total cases.

- Mean: 1.14
- Median: 1.00
- Mode: 1.00
- Minimum: 1.00
- Maximum: 2.00
- Standard Deviation: 0.35

**REC\_NREC Preload Scholar or Non-Recipients**

Location: 19-20 (width: 2; decimal: 0)  
 Variable Type: numeric  
 Question: Preload Scholar or Non-Recipient

<i>Value</i>	<i>Label</i>	<i>Unweighted Frequency</i>	<i>%</i>
1	Scholar	1000	42.8 %
2	Non-Recipient	1338	57.2 %

Based upon 2338 valid cases out of 2338 total cases.

- Mean: 1.57
- Median: 2.00
- Mode: 2.00



- Study 34437 -

- Minimum: 1.00
- Maximum: 2.00
- Standard Deviation: 0.49

**SCH\_YR**                      **Preload School Year**

Location:                      21-22 (width: 2; decimal: 0)  
 Variable Type:                numeric  
 Question:                      School Year

<i>Value</i>	<i>Label</i>	<i>Unweighted Frequency</i>	<i>%</i>
1	2000-2001	0	0.0 %
2	2001-2002	2338	100.0 %
3	2002-2003	0	0.0 %
4	2003-2004	0	0.0 %
5	2004-2005	0	0.0 %

Based upon 2338 valid cases out of 2338 total cases.

- Mean: 2.00
- Median: 2.00
- Mode: 2.00
- Minimum: 2.00
- Maximum: 2.00
- Standard Deviation: 0.00

**NCTOTAL**                      **Non-cognitive: Composite score**

Location:                      23-24 (width: 2; decimal: 0)  
 Variable Type:                numeric  
 Range of Missing Values (M): -5  
 Question:                      Non-cognitive: Composite score

<i>Value</i>	<i>Label</i>	<i>Unweighted Frequency</i>	<i>%</i>
36	-	1	0.0 %
39	-	1	0.0 %
43	-	2	0.1 %
44	-	1	0.0 %
45	-	2	0.1 %
46	-	2	0.1 %
47	-	7	0.3 %
48	-	8	0.3 %
49	-	7	0.3 %
50	-	9	0.4 %
51	-	15	0.6 %

- Study 34437 -

<i>Value</i>	<i>Label</i>	<i>Unweighted Frequency</i>	<i>%</i>
52	-	16	0.7 %
53	-	13	0.6 %
54	-	18	0.8 %
55	-	18	0.8 %
56	-	27	1.2 %
57	-	26	1.1 %
58	-	17	0.7 %
59	-	27	1.2 %
60	-	34	1.5 %
61	-	32	1.4 %
62	-	40	1.7 %
63	-	31	1.3 %
64	-	34	1.5 %
65	-	28	1.2 %
66	-	43	1.8 %
67	-	50	2.1 %
68	-	86	3.7 %
69	-	69	3.0 %
70	-	72	3.1 %
71	-	125	5.3 %
72	-	179	7.7 %
73	-	158	6.8 %
74	-	141	6.0 %
75	-	132	5.6 %
76	-	159	6.8 %
77	-	128	5.5 %
78	-	116	5.0 %
79	-	117	5.0 %
80	-	74	3.2 %
81	-	57	2.4 %
82	-	69	3.0 %
83	-	52	2.2 %
84	-	36	1.5 %
85	-	34	1.5 %
86	-	22	0.9 %
-5 (M)	Missing Data	3	0.1 %

Based upon 2335 valid cases out of 2338 total cases.

- Mean: 72.01
- Median: 73.00

- Study 34437 -

- Mode: 72.00
- Minimum: 36.00
- Maximum: 86.00
- Standard Deviation: 8.11

**NUEQNCOG**      **Num. nominees in pool with same noncog composite score**

Location: 25-27 (width: 3; decimal: 0)  
 Variable Type: numeric  
 Range of Missing Values (M): -5  
 Question: Num. nominees in pool with same noncog composite score

<i>Value</i>	<i>Label</i>	<i>Unweighted Frequency</i>	<i>%</i>
1	-	2	0.1 %
2	-	3	0.1 %
9	-	2	0.1 %
10	-	2	0.1 %
17	-	7	0.3 %
20	-	15	0.6 %
22	-	31	1.3 %
30	-	15	0.6 %
32	-	16	0.7 %
34	-	34	1.5 %
36	-	36	1.5 %
49	-	13	0.6 %
53	-	70	3.0 %
57	-	57	2.4 %
62	-	18	0.8 %
69	-	69	3.0 %
76	-	74	3.2 %
79	-	17	0.7 %
87	-	26	1.1 %
89	-	54	2.3 %
116	-	150	6.4 %
117	-	149	6.4 %
125	-	34	1.5 %
128	-	128	5.5 %
132	-	132	5.6 %
135	-	28	1.2 %
141	-	141	6.0 %
142	-	40	1.7 %
159	-	317	13.6 %
160	-	31	1.3 %

- Study 34437 -

<i>Value</i>	<i>Label</i>	<i>Unweighted Frequency</i>	<i>%</i>
166	-	125	5.3 %
177	-	86	3.7 %
180	-	179	7.7 %
186	-	119	5.1 %
196	-	72	3.1 %
203	-	43	1.8 %
-5 (M)	Missing Data	3	0.1 %

Based upon 2335 valid cases out of 2338 total cases.

- Mean: 129.69
- Median: 135.00
- Mode: 159.00
- Minimum: 1.00
- Maximum: 203.00
- Standard Deviation: 47.84

**NCSUBSR1 Non-cognitive: Positive self-concept**

Location: 28-29 (width: 2; decimal: 0)  
 Variable Type: numeric  
 Range of Missing Values (M): -5  
 Question: Non-cognitive: Positive self-concept

<i>Value</i>	<i>Label</i>	<i>Unweighted Frequency</i>	<i>%</i>
3	-	3	0.1 %
4	-	65	2.8 %
5	-	119	5.1 %
6	-	510	21.8 %
7	-	953	40.8 %
8	-	685	29.3 %
-5 (M)	Missing Data	3	0.1 %

Based upon 2335 valid cases out of 2338 total cases.

- Mean: 6.88
- Median: 7.00
- Mode: 7.00
- Minimum: 3.00
- Maximum: 8.00
- Standard Deviation: 0.99

**NCSUBSR2 Non-cognitive: Realistic self-appraisal**

- Study 34437 -

Location: 30-31 (width: 2; decimal: 0)  
 Variable Type: numeric  
 Range of Missing Values (M): -5  
 Question: Non-cognitive: Realistic self-appraisal

<i>Value</i>	<i>Label</i>	<i>Unweighted Frequency</i>	<i>%</i>
3	-	4	0.2 %
4	-	87	3.7 %
5	-	180	7.7 %
6	-	525	22.5 %
7	-	993	42.5 %
8	-	546	23.4 %
-5 (M)	Missing Data	3	0.1 %

Based upon 2335 valid cases out of 2338 total cases.

- Mean: 6.74
- Median: 7.00
- Mode: 7.00
- Minimum: 3.00
- Maximum: 8.00
- Standard Deviation: 1.03

**NCSUBSR3 Non-cognitive: Understand and navigate social system**

Location: 32-33 (width: 2; decimal: 0)  
 Variable Type: numeric  
 Range of Missing Values (M): -5  
 Question: Non-cognitive: Understand and navigate social system

<i>Value</i>	<i>Label</i>	<i>Unweighted Frequency</i>	<i>%</i>
3	-	6	0.3 %
4	-	157	6.7 %
5	-	307	13.1 %
6	-	650	27.8 %
7	-	798	34.1 %
8	-	417	17.8 %
-5 (M)	Missing Data	3	0.1 %

Based upon 2335 valid cases out of 2338 total cases.

- Mean: 6.43
- Median: 7.00
- Mode: 7.00
- Minimum: 3.00
- Maximum: 8.00

- Study 34437 -

- Standard Deviation: 1.14

**NCSUBSR4 Non-cognitive: Pref long-range goals over short-term/immed. needs**

Location: 34-35 (width: 2; decimal: 0)  
 Variable Type: numeric  
 Range of Missing Values (M): -5  
 Question: Non-cognitive: Pref long-range goals over short-term/immed. needs

<i>Value</i>	<i>Label</i>	<i>Unweighted Frequency</i>	<i>%</i>
3	-	4	0.2 %
4	-	85	3.6 %
5	-	179	7.7 %
6	-	557	23.8 %
7	-	923	39.5 %
8	-	587	25.1 %
-5 (M)	Missing Data	3	0.1 %

Based upon 2335 valid cases out of 2338 total cases.

- Mean: 6.74
- Median: 7.00
- Mode: 7.00
- Minimum: 3.00
- Maximum: 8.00
- Standard Deviation: 1.04

**NCSUBSR5 Non-cognitive: Availability of strong support person**

Location: 36-37 (width: 2; decimal: 0)  
 Variable Type: numeric  
 Range of Missing Values (M): -5  
 Question: Non-cognitive: Availability of strong support person

<i>Value</i>	<i>Label</i>	<i>Unweighted Frequency</i>	<i>%</i>
4	-	305	13.0 %
5	-	831	35.5 %
6	-	1199	51.3 %
-5 (M)	Missing Data	3	0.1 %

Based upon 2335 valid cases out of 2338 total cases.

- Mean: 5.38
- Median: 6.00
- Mode: 6.00
- Minimum: 4.00

- Study 34437 -

- Maximum: 6.00
- Standard Deviation: 0.71

**NCSUBSR6 Non-cognitive: Leadership**

Location: 38-39 (width: 2; decimal: 0)  
 Variable Type: numeric  
 Range of Missing Values (M): -5  
 Question: Non-cognitive: Leadership

<i>Value</i>	<i>Label</i>	<i>Unweighted Frequency</i>	<i>%</i>
3	-	7	0.3 %
4	-	68	2.9 %
5	-	236	10.1 %
6	-	626	26.8 %
7	-	883	37.8 %
8	-	515	22.0 %
-5 (M)	Missing Data	3	0.1 %

Based upon 2335 valid cases out of 2338 total cases.

- Mean: 6.65
- Median: 7.00
- Mode: 7.00
- Minimum: 3.00
- Maximum: 8.00
- Standard Deviation: 1.04

**NCSUBSR7 Non-cognitive: Community service/involvement**

Location: 40-41 (width: 2; decimal: 0)  
 Variable Type: numeric  
 Range of Missing Values (M): -5  
 Question: Non-cognitive: Community service/involvement

<i>Value</i>	<i>Label</i>	<i>Unweighted Frequency</i>	<i>%</i>
3	-	15	0.6 %
4	-	111	4.7 %
5	-	282	12.1 %
6	-	742	31.7 %
7	-	794	34.0 %
8	-	391	16.7 %
-5 (M)	Missing Data	3	0.1 %

Based upon 2335 valid cases out of 2338 total cases.

- Study 34437 -

- Mean: 6.44
- Median: 7.00
- Mode: 7.00
- Minimum: 3.00
- Maximum: 8.00
- Standard Deviation: 1.09

**NCSUBSR8 Non-cognitive: Abil. to acquire knowledge in non-tradit. ways**

Location: 42-43 (width: 2; decimal: 0)  
 Variable Type: numeric  
 Range of Missing Values (M): -5  
 Question: Non-cognitive: Abil. to acquire knowledge in non-tradit. ways

<i>Value</i>	<i>Label</i>	<i>Unweighted Frequency</i>	<i>%</i>
3	-	3	0.1 %
4	-	133	5.7 %
5	-	322	13.8 %
6	-	696	29.8 %
7	-	816	34.9 %
8	-	365	15.6 %
-5 (M)	Missing Data	3	0.1 %

Based upon 2335 valid cases out of 2338 total cases.

- Mean: 6.41
- Median: 7.00
- Mode: 7.00
- Minimum: 3.00
- Maximum: 8.00
- Standard Deviation: 1.09

**NCSUBSR9 Non-cognitive: Rigor of course work**

Location: 44-45 (width: 2; decimal: 0)  
 Variable Type: numeric  
 Range of Missing Values (M): -5  
 Question: Non-cognitive: Rigor of course work

<i>Value</i>	<i>Label</i>	<i>Unweighted Frequency</i>	<i>%</i>
3	-	2	0.1 %
4	-	32	1.4 %
5	-	133	5.7 %
6	-	448	19.2 %
7	-	722	30.9 %



- Study 34437 -

<i>Value</i>	<i>Label</i>	<i>Unweighted Frequency</i>	<i>%</i>
8	-	998	42.7 %
-5 (M)	Missing Data	3	0.1 %

Based upon 2335 valid cases out of 2338 total cases.

- Mean: 7.08
- Median: 7.00
- Mode: 8.00
- Minimum: 3.00
- Maximum: 8.00
- Standard Deviation: 0.99

**NCSUBSR11 Non-cognitive: Math/science/foreign language courses**

Location: 46-47 (width: 2; decimal: 0)  
 Variable Type: numeric  
 Range of Missing Values (M): -5  
 Question: Non-cognitive: Math/science/foreign language courses

<i>Value</i>	<i>Label</i>	<i>Unweighted Frequency</i>	<i>%</i>
3	-	2	0.1 %
4	-	39	1.7 %
5	-	135	5.8 %
6	-	525	22.5 %
7	-	620	26.5 %
8	-	1014	43.4 %
-5 (M)	Missing Data	3	0.1 %

Based upon 2335 valid cases out of 2338 total cases.

- Mean: 7.04
- Median: 7.00
- Mode: 8.00
- Minimum: 3.00
- Maximum: 8.00
- Standard Deviation: 1.03

**NCSUBSR12 Non-cognitive: Scholarly nature of essay**

Location: 48-49 (width: 2; decimal: 0)  
 Variable Type: numeric  
 Range of Missing Values (M): -5  
 Question: Non-cognitive: Scholarly nature of essay

- Study 34437 -

<i>Value</i>	<i>Label</i>	<i>Unweighted Frequency</i>	<i>%</i>
2	-	1	0.0 %
3	-	46	2.0 %
4	-	142	6.1 %
5	-	319	13.6 %
6	-	841	36.0 %
7	-	698	29.9 %
8	-	288	12.3 %
-5 (M)	Missing Data	3	0.1 %

Based upon 2335 valid cases out of 2338 total cases.

- Mean: 6.23
- Median: 6.00
- Mode: 6.00
- Minimum: 2.00
- Maximum: 8.00
- Standard Deviation: 1.14