

# UnderStandingAmericaStudy

COGNITIVE COMPREHENSIVE FILE: DATA DESCRIPTION



USC Dornsife Center for Economic and Social Research  
Andrew Becker  
Michael Moldoff  
Bart Orriens  
Document Updated: 3/6/2025

# Contents

<b>1. Introduction .....</b>	<b>3</b>
<b>2. Overview of the Cognitive Comprehensive File .....</b>	<b>4</b>
2.1 Input Files and Survey Fielding Periods .....	4
2.2 Variable Name Prefixes.....	6
2.3 File Structure .....	7
2.4 CogCF Production Cycle .....	7
2.5 Obtaining the CogCF and Merging with Other UAS Data Files .....	7
<b>3. Variables.....</b>	<b>8</b>
3.1 Cognitive Assessment Variables .....	8
3.1.1 In the UAS HRS: Section D .....	8
3.1.2 In Study P: Financial Literacy, Numeracy.....	10
3.1.3 In Woodcock-Johnson Studies (Studies N, V, and A) .....	11
3.1.4 In the Stop and Go Switch (Study G) .....	11
3.1.5 In the Figure Identification Task (Study U) .....	12
3.1.6 In the Mental Health core surveys: PROMIS® Perceived Cognitive Function (Study C) .....	14
3.2 Survey Participation Flags .....	14
3.3 Demographic and Default Variables .....	15
3.3.1 Default Survey Variables.....	15
3.3.2 Demographics from My Household.....	15
<b>4. probability of cognitive impairment (Pci) score .....</b>	<b>17</b>
<b>5. Data Issues .....</b>	<b>19</b>
<b>Appendix A. Variable Lists.....</b>	<b>20</b>
A.1 Cognitive Comprehensive File (CogCF) - Complete Variable List.....	20
A.2 CogCF Demographic & Default Variables File - Summary of Variables Included .....	27



---

## 1. INTRODUCTION

---

This document describes the UAS Cognitive Comprehensive File (referred to as the CogCF in this document), which was created at the University of Southern California's Center for Economic and Social Research (CESR) with data collected in the Understanding America Study (UAS). Support for this dataset is provided by the Social Security Administration and the National Institute on Aging.

The UAS is a nationally representative panel of American households randomly recruited through address-based sampling. Respondents are aged 18 and up, and they complete surveys, on average, once or twice monthly via an online interface that is technologically powerful, user-friendly, and quick to deliver results.

The CogCF aggregates the cognitive assessment results calculated within 25 UAS surveys--all waves (currently a maximum of five) of seven distinct UAS studies. It also includes respondent and household identifiers, survey completion dates, and variables indicating whether the respondent completed each of the 25 surveys.

Accompanying the CogCF is the CogCF Demographic & Default Variables File, which contains the UAS standard variables (which are collected for every UAS survey) from each survey that contributed data to the CogCF.

Both data files, as well as this Data Description, can be found here. Please send all questions about the UAS, the CogCF, or this data description to [uas-l@mymailists.usc.edu](mailto:uas-l@mymailists.usc.edu).

---

## 2. OVERVIEW OF THE COGNITIVE COMPREHENSIVE FILE

---

### 2.1 Input Files and Survey Fielding Periods

The cognitive assessment variables contained in the CogCF are derived from eight UAS studies. Each study has from one to five waves of data collection, with (mostly) a single UAS survey for each wave. These surveys and their fielding periods are listed in Table 1 at the end of this subsection.

Waves for the four Woodcock-Johnson studies (N, V, and A) have a rolling fielding period and never close, with every respondent first completing Wave 1 (e.g., UAS 42 for Study N) of a given study, followed by Wave 2 (e.g., UAS 83), etc. Thus, new respondents always begin with the Wave 1 survey, no matter when they entered the UAS sample. Prior to Oct. 1, 2022 invites to each new wave were sent approximately two years after the previous wave was completed. From Oct. 1, 2022 to the present, invites to each new wave are sent approximately two years after the date on which respondents were invited to the previous wave--contingent on them having completed the previous wave. The minimum time between any two waves is one year.

The UAS HRS waves had a similar rolling fielding period sequence at first, and then at Wave 3 in 2019 shifted to a fully sequential fielding where respondents (whether new to the sample or reinterviews invited two years after completing the previous HRS wave) are only allowed to start a survey in that study's newest wave. Thus, each HRS wave is currently fielded within a particular two-year range (approximately). Waves of Study P also have a fully sequential fielding, with invites sent only for the newest wave<sup>[1]</sup>.

Finally, waves in Studies G, U and C also have a never-closing, rolling fielding period. Invites to each new wave are sent approximately two years after the date on which respondents were invited to the previous wave (contingent on them having completed the previous wave).

---

<sup>1</sup> Overlap between the launch date of a new wave and the final start date of the previous wave is due to a small number of respondents who received an invite during the previous wave's fielding period and were not reassigned to the new wave's list.

Table 1. Surveys in the Cognitive Comprehensive File and their Fielding Periods

UAS Study	UAS Survey	Wave	Launch Date	Final Start Date	Closing Date <sup>[2]</sup>
UAS HRS <sup>[3]</sup> , Sections A - D: Health and Retirement Study (includes serial sevens, general knowledge, --starting with UAS396--word recall, a box clicking speed score, and a Probability of Cognitive Impairment score)	UAS 20	1	5/20/2015	6/11/2019	10/30/2020
	UAS 95	2	6/1/2017	6/12/2019	10/30/2020
	UAS 185	3	6/7/2019	6/23/2021	8/22/2022
	UAS 396	4	7/2/2021	7/7/2023	still open
	UAS 530	5	7/14/2023	still inviting	still open
Study P <sup>[4]</sup> : includes Financial Literacy; Numeracy	UAS 1	1	5/31/2014	6/6/2018	7/12/2018
	UAS 121	2	1/15/2018	8/6/2020	8/25/2020
	UAS 237	3	4/3/2020	8/8/2022	10/18/2022
	UAS 458	4	8/8/2022	8/8/2022	8/9/2024
	UAS 593	5	8/9/2024	still inviting	still open
	UAS 594	5	8/9/2024	still inviting	still open
Study N: Woodcock-Johnson 1 – Numbers	UAS 42	1	6/13/2016	still inviting	still open
	UAS 83	2	6/22/2018	still inviting	still open
	UAS 292	3	6/26/2020	still inviting	still open
	UAS 483	4	8/19/2022	still inviting	still open
	UAS 642	5	8/19/2024	still inviting	still open
Study V: Woodcock-Johnson 2 – Picture Vocabulary	UAS 43	1	6/17/2016	still inviting	still open
	UAS 84	2	6/29/2018	still inviting	still open
	UAS 293	3	7/3/2020	still inviting	still open
	UAS 484	4	8/26/2022	still inviting	still open
	UAS 643	5	8/26/2024	still inviting	still open
Study A: Woodcock-Johnson 3 – Verbal Analogies	UAS 44	1	7/1/2016	still inviting	still open
	UAS 85	2	7/6/2018	still inviting	still open
	UAS 294	3	7/10/2020	still inviting	still open
	UAS 485	4	9/2/2022	still inviting	still open
	UAS 644	5	9/2/2024	still inviting	still open
Study G <sup>[5]</sup> : Stop and Go Switch	UAS 324	1	11/18/2020	still inviting	still open
	UAS 488	2	11/28/2022	still inviting	still open
	UAS 668	3	11/29/2024	still inviting	still open
Study U <sup>[5]</sup> : Figure Identification	UAS 327	1	11/18/2020	still inviting	still open
	UAS 489	2	11/28/2022	still inviting	still open
	UAS 669	3	11/29/2024	still inviting	still open
Study C: Cognitive Functioning	UAS 604	1	9/16/2024	still inviting	still open

<sup>2</sup> Access to a survey remains open during an extended period for those respondents who started the survey but did not complete before the wave stopped accepting additional respondents.

<sup>3</sup> Each wave of this study fields multiple surveys. The CogCF is only including variables from one of these surveys.

<sup>4</sup> Not all modules in Study P, which are listed here, are included in the CogCF.

<sup>5</sup> Wave 1 of this study fields multiple surveys. The CogCF is only including variables from the first of these surveys.

---

## 2.2 Variable Name Prefixes

All cognitive assessment variables in the CogCF are preceded by a prefix denoting the study and the wave in which they originated (e.g., the "n1" in variable "n1nsa\_score"). The first character in each cognitive assessment variable name, which is a letter, denotes the particular study, while the next character, a number, references the wave. Table 2 below displays the one-letter study-denoting prefixes along with a suggestion for how to remember the particular study that each letter prefix denotes.

**Table 2. Letter Prefixes within Variable Names and the Studies they Denote**

UAS study	Letter Prefix	Memory Suggestion
UAS HRS, Sections A - D: Health and Retirement Study	r	the <i>r</i> in <i>HRS</i>
Study P: Financial Literacy; Understanding Probabilities (A-DMC); Numeracy	p	the <i>p</i> in psychological scales or in probability
Study N: Woodcock-Johnson 1 – Number Series	n	the <i>n</i> in <i>numbers</i>
Study V: Woodcock-Johnson 2 – Picture Vocabulary	v	the <i>v</i> in <i>vocabulary</i>
Study A: Woodcock-Johnson 3 – Verbal Analogies	a	the <i>a</i> in <i>analogies</i>
Study G: Stop and Go Switch	g	the <i>g</i> in <i>go</i>
Study U: Figure Identification	u	the <i>u</i> in <i>figure</i>
Study C: Cognitive Functioning	c	the <i>c</i> in <i>cognitive</i>

The wave numbering scheme used in the CogCF (which numbers Waves 1, 2, and so on) is different than that found in the [UAS Comprehensive File](#) (which begins numbering at Wave 12, then Wave 13, etc. based on the [RAND HRS](#) wave numbering scheme). The Comprehensive File uses this naming structure to promote familiarity with its data for users who have previously worked with the RAND HRS data, as the Comprehensive File includes similarly derived variables from most sections of the core HRS questionnaire.

A crosswalk between CogCF variable names and their Comprehensive File names can be found in [Appendix A.1](#). This is a full list of the CogCF variables and their labels, and includes a column containing each CogCF variable's name within the Comprehensive File (when applicable).

---

## 2.3 File Structure

The CogCF has a respondent-level structure, with a single record per respondent. Individual respondents are uniquely identified by their UAS identification code, *uasid*. New waves of data are added to the file as additional columns.

---

## 2.4 CogCF Production Cycle

The CogCF is auto-generated nightly from a Stata do-file. Each nightly update contains the most recent data from all included UAS surveys. The variable *date\_cogcfCreated* contains the date when the CogCF was created, as well as when all included surveys were downloaded. Previous days' versions of the CogCF are archived. Due to the CogCF's auto-generation, sample weights are not calculated.

---

## 2.5 Obtaining the CogCF and Merging with Other UAS Data Files

The CogCF is available for download in Stata and .csv formats from the UAS website. If you would like the data in a different format, such as SAS, please email your request to [uas-l@mymaillists.usc.edu](mailto:uas-l@mymaillists.usc.edu). All data in the CogCF have their origin in UAS publicly available survey files, which are accessible through the [UAS All Surveys page](#).

Before accessing the CogCF or other UAS data, one must first obtain permission by [registering on the UAS site](#) to download the UAS public release files. Registering with UAS is considered agreeing to the “conditions of use” governing access and usage of the data.

The CogCF can be merged with other UAS special datasets and all publicly released UAS surveys using the unique person identifier *uasid*, which remains fixed across every survey a respondent takes. Other data products can be found on the [UAS Data Files And Products page](#), and a listing of all available UAS surveys is found on the [UAS All Surveys page](#).



---

### 3. VARIABLES

---

All variables in the CogCF are listed in [Appendix A.1](#) along with their labels. The variables found in the CogCF Demographic & Default Variables File are summarized in [Appendix A.2](#)

In the below variable descriptions, the string "<w>" in a variable name represents the set of all pertinent wave numbers. For the studies with multiple waves, this is the set of wave numbers 1-5. For example, *p<w>finlitscore* represents the set of finlitscores variables from each wave of Study P, which is currently *p1finlisscore*, *p2finlitscore*, *p3finlitscore*, and *p4finlitscore*. For the single wave studies, "<w>" simply gets replaced with "1".

---

#### 3.1 Cognitive Assessment Variables

The principal content in the CogCF is the set of cognitive assessment variables listed in this section. Unless otherwise stated, each of these variables is found in every wave of the pertinent study.

##### 3.1.1 In the UAS HRS: Section D

Each wave of the UAS HRS is divided into six UAS surveys in order to limit respondent burden. The CogCF incorporates a subset of variables collected in the first survey of the series, which is the survey that collects data from sections A - D of the HRS questionnaires.

Variables *r<w>d101* and *r<w>d102* variables contain self-reported answers for current memory and how the current memory compares to the memory at the time of the previous wave of the UAS HRS (or compared to two years ago if the first wave of the UAS HRS).

The *r<w>serialseven* variables (e.g., *r1serialseven*, *r2serialseven*, etc.) contain the number of correct answers in the HRS serial sevens cognitive test. More specifically, they contain the number of correct subtractions within a series of five subtraction questions. The first subtraction question asks respondents to subtract 7 from 100, and then each of the four subsequent questions ask respondents to subtract 7 from their preceding answer.

The variables *r<w>d155\_ok*, *r<w>d156\_ok*, *r<w>d157\_ok*, and *r<w>d158\_ok* are the results of the HRS general knowledge test (sometimes referred to as the "naming" test). They contain the dichotomous scoring (correct vs. not correct) of respondent answers to questions on who is the president, who is the vice president, and to two questions asking the respondent to name common objects. Starting in Wave 3, questions *d155-d158* are no longer asked in the UAS HRS, so these four variables are only available for Waves 1 and 2.

The word recall test variables (*r<w>wordlistvalue*, *r<w>recall1\_score*, *r<w>recall1\_nowords*, *r<w>recall1\_refuse*, *r<w>recall2\_score*, *r<w>recall2\_nowords*, and *r<w>recall2\_refuse*) were collected beginning with Wave 4. In this test respondents are shown a list of ten words and later asked to recall the list at two separate points within the survey. The total number of words remembered at the first and second points are contained in the two *score* variables, while the *nowords* and *refuse* variables store whether respondents selected "don't remember any words" or "don't want to try to remember any words," instead of attempting to type out the word list. It is left to each CogCF user to decide whether to code one, both, or neither of these as equivalent to a score of zero<sup>[5]</sup>. Those respondents who selected one of these two options during the first recall question were not presented with the second recall question (in which case the second question's variables were assigned a value of .a), and it is again left to the user to decide how to code these .a values, that is, whether to code the second recall as equivalent to a score of zero when the first recall response was *nowords*, when it was *refuse*, both, or neither.

Also starting in Wave 4, the box clicking test was adopted from the Health and Retirement Study on-line TICS. Respondents were presented with 4 boxes and asked to click inside each box as quickly as they could. The score in variable *r5\_box\_click\_score* captures the number of seconds from when the box clicking test page was loaded to when the final box was clicked. Scores were log-transformed to normalize their distribution. Box clicking times above the 99th percentile (90 seconds) were set to missing as they were considered outliers, consistent with prior treatment of response time data. Additionally, scores where boxes were clicked more than seven times were set to missing. Information about the validity of the box clicking test score is provided here:

"Hernandez, R., Gatz, M., Schneider, S., Finkel, D., Darling, J. E., Orriens, B., Liu, Y., & Kapteyn, A. (2024). Visual–Motor Integration (VMI) Is Also Relevant for Computer, Smartphone, and Tablet Use by Adults: Introducing the Brief Box Clicking Test. *The American Journal of Occupational Therapy*, 78(5), 7805205010. <https://doi.org/10.5014/ajot.2024.050680>"

Finally, in Wave 4 and following waves a typing speed measure was created from the Health and Retirement Study on-line TICS. Respondents were asked to type in a pre-specified sentence as fast as possible ("The quick brown fox jumps over the lazy dog"; for Spanish participants "Cada vez que trabajo, Felix me paga un whisky"). Two variables are included to signify typing speed with adjustment for accuracy: *r<w>awpm* and its log-transformation *r<w>lawpm*. These variables

---

<sup>6</sup> It is typical to code "don't remember any words" as equivalent to a recall score of 0.

capture the number of words typed per minute as a product of typing speed and accuracy, as it was operationalized in prior work (Pinet et. al., 2022). Typing speed was computed as words typed per minute; accuracy is calculated as the Levenshtein distance between the pre-specified target sentence and the typed sentence. Typing speed for observations with less than 50% accuracy were set to missing for being over erroneous. A higher value for awpm signifies a higher typing speed, with adjustment for level of accuracy (e.g., lower accuracy makes awpm lower). Further details about the calculations behind the typing speed scores can be found here:

“Pinet, S., Zielinski, C., Alario, F.-X., & Longcamp, M. (2022). Typing expertise in a large student population. *Cognitive Research: Principles and Implications*, 7(1), 77. <https://doi.org/10.1186/s41235-022-00424-3>”

More about the methodology behind these UAS HRS cognitive assessment variables can be found in the pdf documents "UAS Cognitive Scoring Wave <w>", currently available for download on the Wave 1 - Wave 5 survey webpages ([UAS 20](#), [UAS 95](#), [UAS 185](#), [UAS 396](#), [UAS 530](#)).

### 3.1.2 In Study P: Financial Literacy, Numeracy

The  $p<w>finlitscore$  variables contain the total number of correct answers each respondent gave to a set of 14 questions about financial literacy. These 14 questions are *l001- l005*, *d001 - d002*, and *p001-p007*, which are asked at the beginning of each wave of Study P. Higher scores in the  $p<w>finlitscore$  variables indicate more financial literacy.

The  $p<w>cog$  variables have each wave's numeracy scale score, which is an Item Response Theory (IRT)-based scaled score representing performance over eight questions (*lip001*, *lip002*, *lip003*, *lip008*, *lip009*, *lip012*, *lip015*, *lip017*) which asked respondents to solve problems designed to measure their proficiency working with probabilities and with other numerical information. The scores were derived using a two-parameter logistic IRT model. Higher scores in the  $p<w>cog$  variables indicate better performance at numeracy problems.

Note that as of wave 5 the financial literacy scale questions and the numeracy scale questions are administered in separate surveys (for wave 5 this is UAS 593 and UAS 594).

More on the methodology used for the  $p<w>cog$  scores can be found in the documents "UAS # Numeracy Scoring", currently available for download on each Study P wave's survey webpage ([UAS 1](#), [UAS 121](#), [UAS 237](#), [UAS 458](#), [UAS 593](#) (UAS593 *finlitscore* only), and [UAS 594](#) (UAS 594 *cog* score only)).

### 3.1.3 In Woodcock-Johnson Studies (Studies N, V, and A)

The variables in studies N, V, and A are measures taken from the Woodcock–Johnson Tests of Cognitive Abilities®. These tests were designed to measure a respondent’s quantitative reasoning (the number series module in Study N) and lexical knowledge (the picture vocabulary module in Study V and the verbal analogies module in Study A). Each module consists of 15-19 items which are scored dichotomously as correctly solved or incorrect.

Odd-numbered waves use Form A, with the total number of correct answers among the 15 items being stored in the variables *n<w>nsa\_score*, *v<w>pva\_score*<sup>6</sup>, and *a<w>vea\_score*. Even-numbered waves use a parallel Form B, and the total number of correct answers during even-numbered waves is stored in variables *n<w>nsb\_score*, *v<w>pvb\_score*, and *a<w>veb\_score*. The two forms are alternated in order to mitigate practice effects for repeated test-takers.

IRT-based scaled scores for the number series, picture vocabulary, and verbal analogy modules were derived using a two-parameter logistic IRT model and are found in the variables *n<w>cog*, *v<w>cog*, and *a<w>cog*.

For each Woodcock-Johnson based measure, higher scores indicate better performance. More on the methodology used for calculating these measures can be found in the documents "UAS # Scoring", currently available for download on the Wave 1 - Wave 5 survey webpages: for Study N ([UAS 42](#), [UAS 83](#), [UAS 292](#), [UAS 483](#), and [UAS 642](#)), Study V ([UAS 43](#), [UAS 84](#), [UAS 293](#), [UAS 484](#), and [UAS 643](#)), and Study A ([UAS 44](#), [UAS 85](#), [UAS 294](#), [UAS 485](#), and [UAS 644](#)).

### 3.1.4 In the Stop and Go Switch (Study G)

Wave 1 of Study G is administered in three surveys: a screener (UAS 323) and two implementations of the Stop and Go Switch Task, with game 1 in UAS 324 and game 2 in UAS 325. The CogCF includes variables from just the first implementation in UAS 324 (except for a handful of cases for which the variables are imputed from UAS 325 per the *g1\_imputed* variable). The reason is that UAS 325 was administered to examine device- and practice-effects in an experimental design, not for the purposes of providing cognitive scores for the full UAS panel. Wave 2 of Study G is administered in a single survey UAS 488. It conducts the same test as in UAS 324, but allows respondents to use the device of their own choosing.

The task is divided into four different types of trials: baseline, reverse, non-switch, and switch. For each trial type, a respondent's median reaction time (in seconds) is stored in four variables whose

---

<sup>7</sup> Not available for wave 3. For more on this see entry in Data Issues section below.

names end the with the text string "score\_all": *g1nb\_score\_all*, *g1rb\_score\_all*, *g1nsw\_score\_all*, and *g1sw\_score\_all*.

However, some respondents perform this task with low accuracy, calling into question whether they completed the task as directed. To account for this, a corresponding set of four variables (*g1nb\_score*, *g1rb\_score*, *g1nsw\_score*, and *g1sw\_score*) were derived by starting with the above "score\_all" variables and then imputing to missing the values for those respondents who appeared to have exceptionally low task fidelity, that is, those respondents who correctly completed less than 70% of all trials. Thus, these four additional variables contain the median reaction times in each trial type for those respondents who completed the task with at least 70% accuracy.

In the Study G measures, it is lower scores that indicate better (which here means faster) performance. Further description of these and other variables in the Stop and Go Switch Task are found in the survey codebooks (particularly pages 15 - 18), which are available for download on the survey webpage for UAS 324 (Wave 1) [here](#), UAS 488 (Wave 2) [here](#) and UAS 668 (Wave 3) [here](#).

### 3.1.5 In the Figure Identification Task (Study U)

Wave 1 of Study U is administered in three surveys: a screener (UAS 323) and two implementations of the Figure Identification Task, with game 1 in UAS 327 and game 2 in UAS 328. The CogCF includes variables from just the first implementation in UAS 327 (except for a handful of cases for which the variables are imputed from UAS 328 per the *u1\_imputed* variable). Wave 2 of Study U is administered in a single survey UAS 489. It conducts the same test as in UAS 327, but allows respondents to use the device of their own choosing.

#### *Wave 1*

The task consists of 60 trials where respondents match (with a mouse-click or touchscreen touch) which of five figures is identical to a target figure presented directly above the five-figure array. The 60 trials are split into two sets of 30 trials, with each set analyzed independently, producing two sets of results.

The four "score" variables (*u1figID\_score\_all*, *u1second\_figID\_score\_all*, *u1figID\_score*, *u1second\_figID\_score*) are measures of speed, penalized for incorrect responses. Each of the four variables count the number of figures a respondent correctly identified within a 90 second time window (after removing responses that are considered invalid<sup>7</sup>). Variables with prefix "u1figID"

---

<sup>8</sup> Responses to single trials that take longer than 30 seconds are considered invalid and removed from all analyses. They do not contribute to score totals, nor do they count against (that is, subtract from) the 90 second time limit.

represent scores for the first 30 trials, and variables with prefix “u1second” represent scores for the second 30 trials. The two "score" variables with annex “all” provide scores for all respondents; variables without the “all” annex set the scores to missing for those respondents who failed to produce correct and valid responses for at least 70% of the 30 trials.

There are also two variables (*u1totalcorrect\_first30items* and *u1totalcorrect\_second30items*) directly measuring accuracy. These contain the total number of correct and valid responses (out of 30), so not restricted by any time limit and not restricted by the number of valid responses.

For all of these variables, higher scores indicate better performance. Further description of these and other variables in Wave 1 of the Figure Identification Task is found in the survey codebook (particularly pages 15 - 16), which is available for download on the survey webpage for UAS 327 [here](#).

#### *Wave 2 and up*

The Wave 2 task consists of 30 trials where respondents match (with a mouse-click or touchscreen touch) which of five figures is identical to a target figure presented directly above the five-figure array. The 30 trials in Wave 2 are a subset of the 60 trials administered in Wave 1 and constitute form 1. The remaining 30 trials from Wave 1 comprise form 2 and are administered in Wave 3. Future waves alternate between form 1 and form 2.

The two "score" variables (*u1figID\_score\_all*, *u1figID\_score*) are measures of speed, penalized for incorrect responses. Each of the two variables count the number of figures a respondent correctly identified within a 90 second time window (after removing responses that are considered invalid<sup>9</sup>). The "score" variable with annex “all” provides the score for all respondents; the variable without the “all” annex set the scores to missing for those respondents who failed to produce correct and valid responses for at least 70% of the 30 trials.

There is also the variable *u1totalcorrect* directly measuring accuracy. This contains the total number of correct and valid responses (out of 30), so not restricted by any time limit and not restricted by the number of valid responses.

---

<sup>9</sup> Responses to single trials that take longer than 30 seconds are considered invalid and removed from all analyses. They do not contribute to score totals, nor do they count against (that is, subtract from) the 90 second time limit.

For all of these variables, higher scores indicate better performance. Further description of these and other variables in the Figure Identification Task is found in the survey codebook (particularly pages 15 - 16), which is available for download on the survey webpage for UAS 489 [here](#).

### 3.1.6 In the Mental Health core surveys: PROMIS® Perceived Cognitive Function (Study C)

The CogCF incorporates the Perceived Cognitive Function instrument from NIH Patient-Reported Outcomes Measurement Information System (PROMIS) administered in UAS604. The instrument assesses subjective cognitive abilities and concerns about cognitive difficulties in daily life. Participants complete 4 self-report items administered as Computerized Adaptive Test (CAT), which dynamically selects items from the larger PROMIS Cognitive Function item bank (32 items) based on each participant's pattern of responses. This adaptive approach optimizes measurement precision while minimizing respondent burden. Higher scores reflect better perceived cognitive function. Scores are provided as theta and T-score values. Thetas have an approximate mean of 0 and standard deviation of 1 in the general population. T-scores have an approximate mean of 50 and standard deviation of 10 in the general population. Variables representing the standard errors of the thetas and T-scores are provided, as well. Relevant variables are *c1theta*, *c1theta\_se*, *c1tscore* and *c1tscore\_se*.

---

## 3.2 Survey Participation Flags

Survey participation is represented in the CogCF with a set of participation flag variables that indicate whether respondents completed each of the available waves for each study (and thus completed the particular survey that is associated with the wave and that contributes variables to the CogCF--see Table 1 above for a mapping between survey wave and UAS number).

In four of the CogCF studies as well as in Wave 2 of Studies G and U, each wave consists of a single survey. Survey participation variables for these studies have the naming structure:

*in\_<s><w>*

where "<s>" indexes the particular study and "<w>" indexes the wave number (e.g., *in\_v2* for the 2<sup>nd</sup> wave of Study V).

All waves of the UAS HRS study, Wave 1 of the Stop and Go Switch study (Study G), and Wave 1 of the Figure Identification study (Study U) all consist of multiple surveys, but only contribute a single survey to the CogCF. For the UAS HRS, the CogCF only includes variables from the Section A - D survey. For Waves 1 of Studies G and U, the CogCF only includes variables from each study's Game 1 survey (UAS 324 and UAS 327, respectively). The participation flag variables for these study

waves represent this, listing both the study/wave and the relevant UAS survey number within that study. For the four waves of the UAS HRS, these survey participation flags are *in\_hrs1\_uas20*, *in\_hrs2\_uas95*, *in\_hrs3\_uas185*, *in\_hrs4\_uas396*, and *in\_hrs5\_uas530*. For Study G Wave 1, the survey participation flag is *in\_g1\_uas324*. For Study U Wave 1, the survey participation flag is *in\_u1\_uas327*.

---

### 3.3 Demographic and Default Variables

Each UAS survey dataset contains a set of default survey variables and demographic variables that together are termed the UAS standard variables (see the [UAS standard variables page for more](#)). Three of these standard variables (*uasid*, *uashhid*, and *end\_date*) are found in the CogCF, merged in with each survey that contributes data to the CogCF.

To avoid overshadowing the cognitive assessment variables (the principal variables of interest within the CogCF), the entire set of standard variables for each survey have been compiled into a companion dataset, the CogCF Demographic & Default Variables File.

#### 3.3.1 Default Survey Variables

Most default survey variables represent when and how the survey was administered. Examples are times/dates when respondents began and completed the survey, whether the survey was administered in English or Spanish, sample weights, and whether a tablet was ever provided to the respondent.

Of particular note are three identifier variables. Each respondent is uniquely identified with *uasid*. This variable, assigned to a respondent at recruitment into the UAS panel, stays fixed across surveys and is the key for linking the Demographic & Default Variables File with the CogCF, as well as with any other UAS dataset. Households are identified with *uashhid* (original household identifier), which also stays fixed, and *survhhid* (survey-specific identifier), which uniquely identifies the household a panel member belongs to at the time of a given survey, and so can change for respondents when they move between different households.

#### 3.3.2 Demographics from My Household

The demographic variables found in the Demographic & Default Variables File include respondents' age, gender, ethnicity, marital status, family structure, state of residence, education, marital status, employment characteristics, and household composition, among other attributes. All demographic variables were collected in the My Household survey, which is administered



quarterly to every respondent in the UAS panel, ensuring that these variables are regularly updated. The demographic variables are associated with each particular survey according to the following rule: variable values collected in the My Household survey completed closest in time before a given survey is begun by a respondent are associated with that survey and added to its data.

---

#### 4. PROBABILITY OF COGNITIVE IMPAIRMENT (PCI) SCORE

---

The Cognitive Comprehensive File contains a probability of cognitive impairment (PCI) score for use by investigators. This is a summary cognitive status score used for research and **should not** be interpreted as a clinical diagnosis. The description of the development of the PCI score is available in the publication:

"Gatz, M., Schneider, S., Meijer, E., Darling, J. E., Orriens, B., Liu, Y., & Kapteyn, A. (2022). Identifying cognitive impairment among older participants in a nationally representative internet panel. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences*, 78(2):201-209. <https://doi.org/10.1093/geronb/gbac172>"

PCI is a continuous score quantifying the probability of cognitive impairment. It is based on the established Langa-Weir classification used in the Health and Retirement Study, which distinguishes "normal cognition", "cognitively impaired/not demented" and "dementia".

Within the CogCF, the PCI score is derived using a combination of memory tests administered in HRS sections D and G: immediate and delayed recall (r4recall1\_score and r4recall2\_score), working memory and attention test (r4serialseven), and self-reported ability to manage daily activities that require cognitive health (r4iadlza). The memory tests are from UAS396; the self-reported ability scores from UAS397). These are summed (r4sumscore) and the PCI score is derived from the scoring crosswalk described in the publication.

Currently the PCI score is available for Study R wave 4 and wave 5. The relevant variables are:

- r4PCI: probability of cognitive impairment score based on r4sumscore and the scoring crosswalk described in the publication listed above.
- Where r4sumscore is the sum of r4recall1\_score, r4recall2\_score, r4serialseven and r4iadlza (here r4iadlza summarizes five instrumental activities of daily living (IADLs): shopping for groceries, preparing hot meals, using the phone, managing money, and taking medications).
- r5PCI: probability of cognitive impairment score based on r5sumscore and the scoring crosswalk described in the publication listed above.
- Where r5sumscore is the sum of r5recall1\_score, r5recall2\_score, r5serialseven and r5iadlza (here r5iadlza summarizes five instrumental activities of daily living (IADLs): shopping for groceries, preparing hot meals, using the phone, managing money, and taking medications).

The following notes and caveats apply:

- The PCI is a summary cognitive status score used for research and should not be interpreted as a clinical diagnosis.

- PCI scores take values within a bounded range from 0 to 1.0.
- Variables r4sumscore and r4PCI are only available for respondents who:
  - completed both UAS396 and UAS397 per variables in\_hrs4\_uas396 and in\_hrs4\_uas397 within 90 days
  - have non-missing values for r4recall1\_score, r4recall2\_score, r4serialseven and r4iadlza
- Variables r5sumscore and r5PCI are only available for respondents who:
  - completed both UAS530 and UAS531 per variables in\_hrs4\_uas530 and in\_hrs4\_uas531 within 90 days
  - have non-missing values for r5recall1\_score, r5recall2\_score, r5serialseven and r5iadlza

## 5. DATA ISSUES

Data quality issues during data collection and actions taken are listed in Table 3 below.

Table 3. Data Issues in the Cognitive Comprehensive File

Study	Wave	Survey	Variable(s) Affected	Issue	Action
N	1	UAS 42	- n1nsa_score	The question nsa41 was incorrectly administered up until July 21, 2016 at 7.59 AM.	For those respondents who completed UAS 42 before this date, nsa41 was not included in the calculation (a sum) for <i>n1nsa_score</i> , so their maximum sum possible is one less (=14) than it is for all other respondents (for whom it =15).
V	3	UAS 293	- v3pva_score	The question pva21 was not administered, so the overall score (which is a simple sum) represented in this variable would be biased in comparison to what it was in other waves.	The variable was never calculated and not included in the survey dataset.
G	1	UAS 324	- g1end_date, - g1nb_score, - g1nb_score_all, - g1rb_score, - g1rb_score_all, - g1nsw_score, - g1nsw_score_all, - g1sw_score, - g1sw_score_all	The stop and go switch task trials were not administered to five respondents. Consequently, these respondents' median reaction time scores were not recorded.	These nine variables were filled with values from UAS 325, which is a second implementation of the stop and go switch task. A variable, <i>g1_imputed</i> , was defined which flags the five respondents who have imputed data for Wave 1 of Study G.
U	1	UAS 327	- u1figID_score, - u1second_figID_score, - u1figID_score_all, - u1second_figID_score_all, - u1totalcorrect_first30items, - u1totalcorrect_second30items	The figure ID task trials were not administered to four respondents. Consequently, these respondents' results on the task were not recorded.	These six variables were filled with values from UAS 328, which is a second implementation of the figure ID task. A variable, <i>u1_imputed</i> , was defined which flags the four respondents who have imputed data for Wave 1 of Study U.

---

## APPENDIX A. VARIABLE LISTS

---

### A.1 Cognitive Comprehensive File (CogCF) - Complete Variable List

Table A.1 below contains a list of all variables included in the CogCF, as well as each variable's name within the UAS Comprehensive File. CogCF variables not found in the Comprehensive File are marked with an "n/a."

Table A.1 Variables in the Cognitive Comprehensive File

Variable Name in CogCF	Variable Name in Comprehensive File	Variable Label
uasid	uasid	Individual identifier
uashhid	uashhid	Original household identifier
in_hrs1_uas20	inuas20	whether completed the 1st surv (UAS 20) of wave 1 of HRS study
in_hrs2_uas95	inuas95	whether completed the 1st surv (UAS 95) of wave 2 of HRS study
in_hrs3_uas185	inuas185	whether completed the 1st surv (UAS 185) of wave 3 of HRS study
in_hrs4_uas396	inuas396	whether completed the 1st surv (UAS 396) of wave 4 of HRS study
in_hrs4_uas397	inuas397	whether completed the 2nd surv (UAS 397) of wave 4 of HRS study
in_hrs5_uas530	inuas530	whether completed the 1st surv (UAS 530) of wave 5 of HRS study
in_p1	in_p12	whether completed wave 1 of study P (UAS 1)
in_p2	in_p13	whether completed wave 2 of study P (UAS 121)
in_p3	in_p14	whether completed wave 3 of study P (UAS 237)
in_p4	in_p15	whether completed wave 4 of study P (UAS 458)
in_p5_uas593	in_p16	whether completed wave 5 of study P part 1 (UAS 593)
in_p5_uas594	in_p16	whether completed wave 5 of study P part 2 (UAS 594)
in_n1	in_n12	whether completed wave 1 of study N (UAS 42)
in_n2	in_n13	whether completed wave 2 of study N (UAS 83)
in_n3	in_n14	whether completed wave 3 of study N (UAS 292)
in_n4	in_n15	whether completed wave 4 of study N (UAS 483)
in_n5	in_n16	whether completed wave 5 of study N (UAS 642)
in_v1	in_v12	whether completed wave 1 of study V (UAS 43)

in_v2	in_v13	whether completed wave 2 of study V (UAS 84)
in_v3	in_v14	whether completed wave 3 of study V (UAS 293)
in_v4	in_v15	whether completed wave 4 of study V (UAS 484)
in_v5	in_v16	whether completed wave 5 of study V (UAS 643)
in_a1	in_a12	whether completed wave 1 of study A (UAS 44)
in_a2	in_a13	whether completed wave 2 of study A (UAS 85)
in_a3	in_a14	whether completed wave 3 of study A (UAS 294)
in_a4	in_a15	whether completed wave 4 of study A (UAS 485)
in_a5	in_a16	whether completed wave 5 of study A (UAS 644)
in_g1_uas324	n/a	whether completed the game 1 surv (UAS 324) of wave 1 of study G
in_g2	n/a	whether completed wave 2 of study G (UAS 488)
in_g3	n/a	whether completed wave 3 of study G (UAS 668)
in_u1_uas327	n/a	whether completed the game 1 surv (UAS 327) of wave 1 of study U
in_u2	n/a	whether completed wave 2 of study U (UAS 489)
in_u3	n/a	whether completed wave 3 of study U (UAS 669)
r1end_date	r12_uas20iwend	w1 uas20:: survey end date
r1d101	n/a	w1 uas20::self-reported rating of current memory
r1d102	n/a	w1 uas20::self-reported comparison of current memory to previous UAS HRS wave
r1serialeven	r12serialeven	w1 uas20 Serial Sevens Test:: score, # of subtractions correct out of 5
r1d155_ok	n/a	w1 uas20 Gen Knowe Test:: whether ans correct, tool used to cut paper
r1d156_ok	n/a	w1 uas20 Gen Knowe Test:: whether ans correct, name of prickly desert plant
r1d157_ok	n/a	w1 uas20 Gen Knowe Test:: whether ans correct, who is president of US
r1d158_ok	n/a	w1 uas20 Gen Knowe Test:: whether ans correct, who is vp of US
r2end_date	r13_uas95iwend	w2 uas95:: survey end date
r2d101	n/a	w2 uas95::self-reported rating of current memory
r2d102	n/a	w2 uas95:: self-reported comparison of current memory to previous UAS HRS wave
r2serialeven	r13serialeven	w2 uas95 Serial Sevens Test:: score, # of subtractions correct out of 5
r2d155_ok	n/a	w2 uas95 Gen Knowe Test:: whether ans correct, tool used to cut paper
r2d156_ok	n/a	w2 uas95 Gen Knowe Test:: whether ans correct, name of prickly desert plant

r2d157_ok	n/a	w2 uas95 Gen Knowe Test:: whether ans correct, who is president of US
r2d158_ok	n/a	w2 uas95 Gen Knowe Test:: whether ans correct, who is vp of US
r3end_date	r14_uas185iwend	w3 uas185:: survey end date
r3d101	n/a	w3 uas185::self-reported rating of current memory
r3d102	n/a	w3 uas185:: self-reported comparison of current memory to previous UAS HRS wave
r3serialseven	r14serialseven	w3 uas185 Serial Sevens Test:: score, # of subtractions correct out of 5
r4end_date	r15_uas396iwend	w4 uas396:: survey end date
r4d101	n/a	w4 uas396::self-reported rating of current memory
r4d102	n/a	w4 uas396:: self-reported comparison of current memory to previous UAS HRS wave
r4serialseven	r15serialseven	w4 uas396 Serial Sevens Test:: score, # of subtractions correct out of 5
r4wordlistvalue	n/a	w4 uas396 Word Recall:: version of word list presented to R for memorizing
r4recall1_score	r15recall1_score	w4 uas396 Word Recall:: ques 1, score (# of words correctly recalled)
r4recall1_nowords	n/a	w4 uas396 Word Recall:: ques 1, whether dont knowed (dont remember any words)
r4recall1_refuse	n/a	w4 uas396 Word Recall:: ques 1, whether refused
r4recall2_score	r15recall2_score	w4 uas396 Word Recall:: ques 2, score (# of words correctly recalled)
r4recall2_nowords	n/a	w4 uas396 Word Recall:: ques 2, whether dont knowed (dont remember any words)
r4recall2_refuse	n/a	w4 uas396 Word Recall:: ques 2, whether refused
r4sumscore	n/a	w4 uas396/397: sum score of immediate & delayed recall, serial 7s and reverse IADL
r4box_click_score	n/a	w4 uas396: box clicking score
r4awpm	n/a	w4 uas396: typing speed score
r4lawpm	n/a	w4 uas396: log-transformed typing speed score
r4iadlza	n/a	w4 uas397: reverse IADL (0-5)
r4PCI	n/a	w4 uas396/397:: probability of cognitive impairment score
r5end_date	r16_uas530iwend	w5 uas530:: survey end date
r5d101	n/a	w5 uas530::self-reported rating of current memory
r5d102	n/a	w5 uas530:: self-reported comparison of current memory to previous UAS HRS wave
r5serialseven	r16serialseven	w5 uas530 Serial Sevens Test:: score, # of subtractions correct out of 5
r5wordlistvalue	n/a	w5 uas530 Word Recall:: version of word list presented to R for memorizing
r5recall1_score	r16recall1_score	w5 uas530 Word Recall:: ques 1, score (# of words correctly recalled)

r5recall1_nowords	n/a	w5 uas530 Word Recall:: ques 1, whether dont knowed (dont remember any words)
r5recall1_refuse	n/a	w5 uas530 Word Recall:: ques 1, whether refused
r5recall2_score	r16recall2_score	w5 uas530 Word Recall:: ques 2, score (# of words correctly recalled)
r5recall2_nowords	n/a	w5 uas530 Word Recall:: ques 2, whether dont knowed (dont remember any words)
r5recall2_refuse	n/a	w5 uas530 Word Recall:: ques 2, whether refused
r5sumscore	n/a	w5 uas530/531: sum score of immediate & delayed recall, serial 7s and reverse IADL
r5box_click_score	n/a	w5 uas530: box clicking score
r5awpm	n/a	w5 uas530: typing speed score
r5lawpm	n/a	w5 uas530: log-transformed typing speed score
r5iadlza	n/a	w5 uas531: reverse IADL (0-5)
r5PCI	n/a	w5 uas530/531:: probability of cognitive impairment score
p1end_date	r12_uas1iwend	w1 uas1:: survey end date
p1cog	p12cog	w1 uas1 Numeracy:: numeracy scale score (IRT-based, us pop mean=50)
p1finlitscore	p12finlitscore	w1 uas1 Financial Literacy:: score (n correct in ques l001-p007, max=14)
p2end_date	r13_uas121iwend	w2 uas121:: survey end date
p2cog	p13cog	w2 uas121 Numeracy:: numeracy scale score (IRT-based, us pop mean=50)
p2finlitscore	p13finlitscore	w2 uas121 Financial Literacy:: score (n correct in ques l001-p007, max=14)
p3end_date	r14_uas237iwend	w3 uas237:: survey end date
p3cog	p14cog	w3 uas237 Numeracy:: numeracy scale score (IRT-based, us pop mean=50)
p3finlitscore	p14finlitscore	w3 uas237 Financial Literacy:: score (n correct in ques l001-p007, max=14)
p4end_date	r15_uas458iwend	w4 uas458:: survey end date
p4cog	p15cog	w4 uas458 Numeracy:: numeracy scale score (IRT-based, us pop mean=50)
p4finlitscore	p15finlitscore	w4 uas458 Financial Literacy:: score (n correct in ques l001-p007, max=14)
p5end_date_uas593	r16_uas593iwend	w uas593:: survey end date
p5end_date_uas594	r16_uas594iwend	w uas594:: survey end date
p5cog	p16cog	w5 uas593 Numeracy:: numeracy scale score (IRT-based, us pop mean=50)
p5finlitscore	p16finlitscore	w5 uas594 Financial Literacy:: score (n correct in ques l001-p007, max=14)
n1end_date	r12_uas42iwend	w1 uas42:: survey end date
n1nsa_score	n12nsa_score	w1 uas42 Woodcock-Johnson:: Number Series, total raw score (max=15)



n1num_cog	n12num_cog	w1 uas42 Woodcock-Johnson:: Number Series, IRT-based score (us pop mean=50)
n2end_date	r13_uas83iwend	w2 uas83:: survey end date
n2nsb_score	n13nsb_score	w2 uas83 Woodcock-Johnson:: Number Series, total raw score (max=19)
n2num_cog	n13num_cog	w2 uas83 Woodcock-Johnson:: Number Series, IRT-based score (us pop mean=50)
n3end_date	r14_uas292iwend	w3 uas292:: survey end date
n3nsa_score	n14nsa_score	w3 uas292 Woodcock-Johnson:: Number Series, total raw score (max=15)
n3num_cog	n14num_cog	w3 uas292 Woodcock-Johnson:: Number Series, IRT-based score (us pop mean=50)
n4end_date	r15_uas483iwend	w4 uas483:: survey end date
n4nsb_score	n15nsb_score	w4 uas483 Woodcock-Johnson:: Number Series, total raw score (max=19)
n4num_cog	n15num_cog	w4 uas483 Woodcock-Johnson:: Number Series, IRT-based score (us pop mean=50)
n5end_date	r16_uas642iwend	w5 uas642:: survey end date
n5nsa_score	n16nsa_score	w5 uas642 Woodcock-Johnson:: Number Series, total raw score (max=15)
n5num_cog	n16num_cog	w5 uas642 Woodcock-Johnson:: Number Series, IRT-based score (us pop mean=50)
n1_order		w1 :: indicates nth time R taken topic N when participating in uas42
n2_order		w2 :: indicates nth time R taken topic N when participating in uas83
n3_order		w3 :: indicates nth time R taken topic N when participating in uas292
n4_order		w4 :: indicates nth time R taken topic N when participating in uas483
n5_order		w5 :: indicates nth time R taken topic N when participating in uas642
v1end_date	r12_uas43iwend	w1 uas43:: survey end date
v1pva_score	v12pva_score	w1 uas43 Woodcock-Johnson:: Picture Vocab, total raw score (max=15)
v1pvoc_cog	v12pvoc_cog	w1 uas43 Woodcock-Johnson:: Picture Vocab, IRT-based score (us pop mean=50)
v2end_date	r13_uas84iwend	w2 uas84:: survey end date
v2pvb_score	v13pvb_score	w2 uas84 Woodcock-Johnson:: Picture Vocab, total raw score (max=19)
v2pvoc_cog	v13pvoc_cog	w2 uas84 Woodcock-Johnson:: Picture Vocab, IRT-based score (us pop mean=50)
v3end_date	r14_uas293iwend	w3 uas293:: survey end date
v3pva_score	v14pva_score	w3 uas293 Woodcock-Johnson:: Picture Vocab, total raw score (max=15)
v3pvoc_cog	v14pvoc_cog	w3 uas293 Woodcock-Johnson:: Picture Vocab, IRT-based score (us pop mean=50)
v4end_date	r15_uas484iwend	w4 uas484:: survey end date
v4pvb_score	v15pvb_score	w4 uas484 Woodcock-Johnson:: Picture Vocab, total raw score (max=19)

v4pvoc_cog	v15pvoc_cog	w4 uas484 Woodcock-Johnson:: Picture Vocab, IRT-based score (us pop mean=50)
v5end_date	r16_uas643iwend	w5 uas643:: survey end date
v5pva_score	v16pva_score	w5 uas643 Woodcock-Johnson:: Picture Vocab, total raw score (max=15)
v5pvoc_cog	v16pvoc_cog	w5 uas643 Woodcock-Johnson:: Picture Vocab, IRT-based score (us pop mean=50)
v1_order		w1 :: indicates nth time R taken topic V when participating in uas43
v2_order		w2 :: indicates nth time R taken topic V when participating in uas84
v3_order		w3 :: indicates nth time R taken topic V when participating in uas293
v4_order		w4 :: indicates nth time R taken topic V when participating in uas484
v5_order		w5 :: indicates nth time R taken topic V when participating in uas643
a1end_date	r12_uas44iwend	w1 uas44:: survey end date
a1vea_score	a12vea_score	w1 uas44 Woodcock-Johnson:: Verbal Analogies, total raw score (max=15)
a1vana_cog	a12vana_cog	w1 uas44 Woodcock-Johnson:: Verbal Analogies, IRT-based score (us pop mean=50)
a2end_date	r13_uas85iwend	w2 uas85:: survey end date
a2veb_score	a13veb_score	w2 uas85 Woodcock-Johnson:: Verbal Analogies, total raw score (max=19)
a2vana_cog	a13vana_cog	w2 uas85 Woodcock-Johnson:: Verbal Analogies, IRT-based score (us pop mean=50)
a3end_date	r14_uas294iwend	w3 uas294:: survey end date
a3vea_score	a14vea_score	w3 uas294 Woodcock-Johnson:: Verbal Analogies, total raw score (max=15)
a3vana_cog	a14vana_cog	w3 uas294 Woodcock-Johnson:: Verbal Analogies, IRT-based score (us pop mean=50)
a4end_date	r15_uas485iwend	w4 uas485:: survey end date
a4veb_score	a15veb_score	w4 uas485 Woodcock-Johnson:: Verbal Analogies, total raw score (max=19)
a4vana_cog	a15vana_cog	w4 uas485 Woodcock-Johnson:: Verbal Analogies, IRT-based score (us pop mean=50)
a5end_date	r16_uas644iwend	w5 uas644:: survey end date
a5vea_score	a16vea_score	w5 uas644 Woodcock-Johnson:: Verbal Analogies, total raw score (max=15)
a5vana_cog	a16vana_cog	w5 uas644 Woodcock-Johnson:: Verbal Analogies, IRT-based score (us pop mean=50)
a1_order		w1 :: indicates nth time R taken topic A when participating in uas44
a2_order		w2 :: indicates nth time R taken topic A when participating in uas85
a3_order		w3 :: indicates nth time R taken topic A when participating in uas294
a4_order		w4 :: indicates nth time R taken topic A when participating in uas485
a5_order		w5 :: indicates nth time R taken topic A when participating in uas644

g1end_date	n/a	w1 uas324 Stop Go:: survey end date
g1_imputed	n/a	w1 uas324 Stop Go:: whether g1 vars were imputed with values from UAS 325
g1nb_score	n/a	w1 uas324 Stop Go:: normal baseline score (median rt in sec), Rs w >= 70% corct
g1nb_score_all	n/a	w1 uas324 Stop Go:: normal baseline score (median rt in sec), all Rs
g1rb_score	n/a	w1 uas324 Stop Go:: reverse baseline score (median rt in sec), Rs w >= 70% corct
g1rb_score_all	n/a	w1 uas324 Stop Go:: reverse baseline score (median rt in sec), all Rs
g1nsw_score	n/a	w1 uas324 Stop Go:: non-switch score (median rt in sec), Rs with >= 70% corct
g1nsw_score_all	n/a	w1 uas324 Stop Go:: non-switch score (median rt in sec), all Rs
g1sw_score	n/a	w1 uas324 Stop Go:: switch score (median rt in sec), Rs with >= 70% correct
g1sw_score_all	n/a	w1 uas324 Stop Go:: switch score (median rt in sec), all Rs
g2end_date	n/a	w2 uas488 Stop Go:: survey end date
g2nb_score	n/a	w2 uas488 Stop Go:: normal baseline score (median rt in sec), Rs w >= 70% corct
g2nb_score_all	n/a	w2 uas488 Stop Go:: normal baseline score (median rt in sec), all Rs
g2rb_score	n/a	w2 uas488 Stop Go:: reverse baseline score (median rt in sec), Rs w >= 70% corct
g2rb_score_all	n/a	w2 uas488 Stop Go:: reverse baseline score (median rt in sec), all Rs
g2nsw_score	n/a	w2 uas488 Stop Go:: non-switch score (median rt in sec), Rs with >= 70% corct
g2nsw_score_all	n/a	w2 uas488 Stop Go:: non-switch score (median rt in sec), all Rs
g2sw_score	n/a	w2 uas488 Stop Go:: switch score (median rt in sec), Rs with >= 70% correct
g2sw_score_all	n/a	w2 uas488 Stop Go:: switch score (median rt in sec), all Rs
g3end_date	n/a	w3 uas668 Stop Go:: survey end date
g3nb_score	n/a	w3 uas668 Stop Go:: normal baseline score (median rt in sec), Rs w >= 70% corct
g3nb_score_all	n/a	w3 uas668 Stop Go:: normal baseline score (median rt in sec), all Rs
g3rb_score	n/a	w3 uas668 Stop Go:: reverse baseline score (median rt in sec), Rs w >= 70% corct
g3rb_score_all	n/a	w3 uas668 Stop Go:: reverse baseline score (median rt in sec), all Rs
g3nsw_score	n/a	w3 uas668 Stop Go:: non-switch score (median rt in sec), Rs with >= 70% corct
g3nsw_score_all	n/a	w3 uas668 Stop Go:: non-switch score (median rt in sec), all Rs
g3sw_score	n/a	w3 uas668 Stop Go:: switch score (median rt in sec), Rs with >= 70% correct
g3sw_score_all	n/a	w3 uas668 Stop Go:: switch score (median rt in sec), all Rs
u1end_date	n/a	w1 uas327 FigID:: survey end date

u1_imputed	n/a	w1 uas327 FigID:: whether u1 vars were imputed with values from UAS 328
u1figID_score	n/a	w1 uas327 FigID:: score (n of <30 s resp crct in 90 s), w 70% req, 1st 30 items
u1second_figID_score	n/a	w1 uas327 FigID:: score (n of <30 s resp crct in 90 s), w 70% req, 2nd 30 items
u1figID_score_all	n/a	w1 uas327 FigID:: score (n of <30 s resp crct in 90 s), no req, 1st 30 items
u1second_figID_score_all	n/a	w1 uas327 FigID:: score (n of <30 s resp crct in 90 s), no req, 2nd 30 items
u1totalcorrect_first30items	n/a	w1 uas327 FigID:: n of correct responses among 1st 30 items
u1totalcorrect_second30items	n/a	w1 uas327 FigID:: n of correct responses among 2nd 30 items
u2end_date	n/a	w2 uas489 FigID:: survey end date
u2figID_score	n/a	w2 uas489 FigID:: score (n of <30 s resp crct in 90 s), w 70% req, form1:30 items
u2figID_score_all	n/a	w2 uas489 FigID:: score (n of <30 s resp crct in 90 s), no req, form1:30 items
u2totalcorrect	n/a	w2 uas489 FigID:: n of correct responses among form 1 30 items
u3end_date	n/a	w3 uas669 FigID:: survey end date
u3figID_score	n/a	w3 uas669 FigID:: score (n of <30 s resp crct in 90 s), w 70% req, form1:30 items
u3figID_score_all	n/a	w3 uas669 FigID:: score (n of <30 s resp crct in 90 s), no req, form1:30 items
u3totalcorrect	n/a	w3 uas669 FigID:: n of correct responses among form 1 30 items
c1theta	n/a	w1 uas604 CogFun:: perceived cognitive function theta (mean=0)
c1theta_se	n/a	w1 uas604 CogFun:: standard deviation perceived cognitive function theta
c1tscore	n/a	w1 uas604 CogFun:: perceived cognitive function t-score (mean=50)
c1tscore_se	n/a	w1 uas604 CogFun:: standard deviation perceived cognitive function of T-score
date_cogcfCreated	n/a	date when the CogCF was created and when its input files were downloaded

## A.2 CogCF Demographic & Default Variables File - Summary of Variables Included

With three exceptions (*uasid*, *uashhid*, and *date\_ddvfCreated*--which remain fixed across all surveys), the variables listed in Table A.2 are included in the Demographic & Default Variables File once for each of the surveys. The variables for each survey are distinguished from one another within the file by the same study/wave prefix scheme used in the CogCF. For example, each survey's gender variable is included in the file as *r1\_gender*, *p1\_gender*, *n1\_gender*, *v1\_gender*, *a1\_gender*, *g1\_gender*, *u1\_gender*, *r2\_gender*, *p2\_gender*, etc.

In the below table, the string "<s>" in a variable name represents the one letter prefix for each of the six studies (r, g, u, p, n, v, a, c) in the CogCF; the string "<w>" in a variable name represents the set of all pertinent wave numbers (1-2 1-4, depending on how many waves the given study has); and the string "<#>" represents the set of UAS survey numbers in the CogCF (e.g., UAS #s 20, 1, 42, 43, 44, 324, 327, 185, 121, etc.). For wave 5 of topic P the variable names will be appended with *\_uas593* and *\_uas594* respectively in order to distinguish between the variables at the time of part 1 and part 2.

Finally, found in the five household member-level variable groups only

(<s><w>\_hhmemberage\_<n>, <s><w>\_hhmembergen\_<n>, <s><w>\_hhmemberin\_<n>, <s><w>\_hhmemberrel\_<n>, <s><w>\_hhmemberuasid\_<n>),

the string "<n>" represents the 10-20 variables (it varies by survey) a given survey has for each group. For example, respondents had at most 11 household members in Wave 1 of Study P. Thus, "<n>" in this survey ranges from 1-11, and this survey's hhmemberage variables in the Standard Variables File are *p1\_hhmemberage\_1* - *p1\_hhmemberage\_11*.

**Table A.2 Variables in the CogCF Demographic & Default Variables File**

Variable Name in Demographic & Default Variables File	Variable Label
uasid	Individual identifier
uashhid	Original household identifier
<s><w>_survhhid	w<w> uas<#>:: survey-specific household composition identifier
<s><w>_uasmembers	w<w> uas<#>:: number of household UAS members besides the respondent
<s><w>_samplotype	w<w> uas<#>:: sample type
<s><w>_batch	w<w> uas<#>:: recruiting batch
<s><w>_primary_respondent	w<w> uas<#>:: primary respondent or not
<s><w>_hardware	w<w> uas<#>:: hardware provided
<s><w>_language	w<w> uas<#>:: survey language
<s><w>_start_date	w<w> uas<#>:: survey start date
<s><w>_start_year	w<w> uas<#>:: survey start year
<s><w>_start_month	w<w> uas<#>:: survey start month
<s><w>_start_day	w<w> uas<#>:: survey start day

<s><w>_start_hour	w<w> uas<#>:: survey start hour
<s><w>_start_min	w<w> uas<#>:: survey start minute
<s><w>_start_sec	w<w> uas<#>:: survey start second
<s><w>_end_date	w<w> uas<#>:: survey end date
<s><w>_end_year	w<w> uas<#>:: survey end year
<s><w>_end_month	w<w> uas<#>:: survey end month
<s><w>_end_day	w<w> uas<#>:: survey end day
<s><w>_end_hour	w<w> uas<#>:: survey end hour
<s><w>_end_min	w<w> uas<#>:: survey end minute
<s><w>_end_sec	w<w> uas<#>:: survey end second
<s><w>_gender	w<w> uas<#>-MyHH:: gender
<s><w>_dateofbirth_year	w<w> uas<#>-MyHH:: year of birth
<s><w>_age	w<w> uas<#>-MyHH:: age
<s><w>_agerange	w<w> uas<#>-MyHH:: age range for missing age
<s><w>_citizenus	w<w> uas<#>-MyHH:: US citizen
<s><w>_bornus	w<w> uas<#>-MyHH:: born in the US
<s><w>_stateborn	w<w> uas<#>-MyHH:: state born - FIPS coding
<s><w>_countryborn	w<w> uas<#>-MyHH:: country born
<s><w>_countryborn_other	w<w> uas<#>-MyHH:: unlisted country born
<s><w>_immigrant_status	w<w> uas<#>-MyHH:: immigrant status
<s><w>_statereside	w<w> uas<#>-MyHH:: state residence - FIPS coding
<s><w>_maritalstatus	w<w> uas<#>-MyHH:: marital status
<s><w>_livewithpartner	w<w> uas<#>-MyHH:: living with partner
<s><w>_education	w<w> uas<#>-MyHH:: highest level of education
<s><w>_hisplatino	w<w> uas<#>-MyHH:: Spanish/Hispanic/Latino
<s><w>_hisplatino_group	w<w> uas<#>-MyHH:: Spanish/Hispanic/Latino group
<s><w>_white	w<w> uas<#>-MyHH:: White
<s><w>_black	w<w> uas<#>-MyHH:: Black
<s><w>_nativeamer	w<w> uas<#>-MyHH:: American Indian or Alaska Native

<s><w>_asian	w<w> uas<#>-MyHH:: Asian
<s><w>_pacific	w<w> uas<#>-MyHH:: Hawaiian/Pacific Islander
<s><w>_race	w<w> uas<#>-MyHH:: race
<s><w>_working	w<w> uas<#>-MyHH:: currently working
<s><w>_sick_leave	w<w> uas<#>-MyHH:: on sick or other leave
<s><w>_unemp_layoff	w<w> uas<#>-MyHH:: unemployed - on layoff
<s><w>_unemp_look	w<w> uas<#>-MyHH:: unemployed - looking
<s><w>_retired	w<w> uas<#>-MyHH:: retired
<s><w>_disabled	w<w> uas<#>-MyHH:: disabled
<s><w>_lf_other	w<w> uas<#>-MyHH:: other labor force status
<s><w>_laborstatus	w<w> uas<#>-MyHH:: labor force status
<s><w>_employmenttype	w<w> uas<#>-MyHH:: employment type
<s><w>_workfullpart	w<w> uas<#>-MyHH:: work full-time or part-time
<s><w>_hourswork	w<w> uas<#>-MyHH:: hours of work per week
<s><w>_hhincome	w<w> uas<#>-MyHH:: household income
<s><w>_anyhhmember	w<w> uas<#>-MyHH:: whether any other HH member
<s><w>_hhmembernumber	w<w> uas<#>-MyHH:: number of household members besides R
<s><w>_hhmemberage_<n>	w<w> uas<#>-MyHH:: HH member <n> age
<s><w>_hhmembergen_<n>	w<w> uas<#>-MyHH:: HH member <n> gender
<s><w>_hhmemberin_<n>	w<w> uas<#>-MyHH:: whether HH member <n> is in the HH
<s><w>_hhmemberrel_<n>	w<w> uas<#>-MyHH:: HH member <n> relationship
<s><w>_hhmemberuasid_<n>	w<w> uas<#>-MyHH:: HH member <n> UAS id
date_ddvfCreated	date when Demog & Default Vars File was created and its input files downloaded