Langa-Weir Classification of CognitiveFunction (1995-2022)

Kenneth M. Langa, David R. Weir, Mohammed Kabeto, and Amanda Sonnega

December 2025
Survey Research Center
Institute for Social Research
University of Michigan

This document describes a researcher-contributed dataset that provides a total summary score for cognition using measures[1] from the core HRS interview as well as the three derived cognition categories: Normal, Cognitively Impaired but not Demented (CIND), and Demented. We refer to these as theLanga-Weir Classifications (Crimmins et al., 2011).[2] HRS also makes use of proxy respondents to reduce sample attrition.[3] Therefore, this dataset also includes scoring of cognition based on data from proxy respondent interviews.

Self-Respondents

Face-to-face / TEL Interview

HRS employs a range of measures to assess cognitive status that can be used to create a summary score. Depending on the study sample, researchers can use either a 35-point (range 0-35) or a 27-point (range 0-27) scale to determine cognitive functioning. The 35-point scale[4] includes a subset of cognitive measures questions that **were only asked of respondents age 65 and older,** whereasthe 27-point scale includes questions that were asked of all HRS respondents.

The 27-point scale includes: 1) an immediate and delayed 10-noun free recall test to measure memory (0 to 20 points); 2) a serial sevens subtraction test to measure working memory (0 to 5 points); and 3) a counting backwards test to measure speed of mental processing (0 to 2 points). The additional questions included in the 35-point scale are the three mental status questions that include: 1) date naming (0 to 4 points); 2) object naming (0 to 2 points); and 3) naming the president and the vice president of the United States (0 to 2 points). The Langa-Weir classification is based on the 27-point scale.

---

1 Much more detail on all of these measures is available in the user guide, *Documentation of Cognitive FunctioningMeasures in the Health and Retirement Study,* also available on the HRS website: https://hrs.isr.umich.edu/sites/default/files/biblio/dr-006.pdf

2 Crimmins, E.M., Kim, J.K., Langa, K.M., & Weir, D.R. (2011). Assessment of cognition using surveys and neuropsychological assessment: the health and retirement study and the aging, demographics, and memory study. TheJournals of Gerontology, Series B: Psychological Sciences and Social Sciences, 66B(S1), i162–i171, doi:10.1093/geronb/gbr048

3Weir DR, Faul JD, Langa KM. Proxy interviews and bias in cognition measures due to non-response in longitudinalstudies: a comparison of HRS and ELSA. Longitudinal and Life Course Studies. 2011;2(2):170-184. doi:10.14301/llcs.v2i2.116.

4 The 35-point scale is available in the RAND HRS dataset as RwCOGTOT.

Web Self-administered

Beginning the 2018 interview, HRS included web-based self-administered interviews for some participants. The variable "cogivemodeX" indicates mode of interview (1=Self-respondent interviewer administer FTF/TEL; 2=Self-respondent self-administer WEB, 3= Proxy-respondent interviewer administer FTF/TEL). Users should use this variable to identify web participants. Immediate and delayed word recall and serial subtraction were administered to web respondents. The 27-point scale for the web participants was created by summing the immediate word recall score, delayed word recall score, serial subtraction score, and a wholly imputed backward count score. As suggested by Domingue et al. (2023), the web-based summary score was adjusted by subtracting one from the total, and the Langa-Weir Classifications were done using this adjusted score.[5]

Self-respondent cognition classification

The Langa-Weir Classifications map onto the 27-point scale (variable name *cogtot27_imp*) thus: Normal (12 – 27); Cognitively Impaired but not Demented (CIND) (7 – 11); and Demented (0 – 6) (variable name *cogfunction*). Crimmins et al. (2011) described the methods used to make these classifications based on diagnostic information from the ADAMS. Another important feature of the present contributed dataset is that it makes use of imputed information for the HRS cognitive measures. Procedures for the imputation methodology are available on the HRS website.[6] It is important to note that if you choose to use raw (not imputed) data, you will be missing a large fraction of respondents with dementia.

Proxy Respondents

If you wish to include proxy respondents in your study sample, note that there is no direct assessment of cognition. Rather, we suggest using a combination of instrumental activities of daily living (IADLs), proxy assessment of the respondent's memory, and the interviewer's assessment of the respondent's cognition to create a total cognition score. IADLs include items for difficulty of preparing meals, shopping, making phone calls, taking medication, and managing money. Note that in this contributed dataset, we also imputed missing values for each IADL item, and proxy and interviewer's assessment.

---

[5] Domingue BW, McCammon RJ, West BT, Langa KM, Weir DR, Faul J. The Mode Effect of Web-Based Surveying on the 2018 U.S. Health and Retirement Study Measure of Cognitive Functioning. J Gerontol B Psychol Sci Soc Sci. 2023 Aug 28;78(9):1466-1473. doi: 10.1093/geronb/gbad068. PMID: 37129872; PMCID: PMC10848225.

[6] Health and Retirement Study Imputation of Cognitive Functioning Measures: 1992 – 2014, available on the HRS website: http://hrsonline.isr.umich.edu/modules/meta/xyear/cogimp/desc/COGIMPdd.pdf

The total cognition score based on proxy information is an 11-point scale using proxy assessment of memory (excellent, very good, good, fair, poor), IADL limitations (ranging from no limitations to 5 limitations), and the interviewer's assessment of cognitive impairment (CI: no CI, may have CI, has CI). For proxy assessment of memory, "excellent" was assigned a value of 0, "very good" a value of 1, "good" a value of 2, "fair" a value of 3, and "poor" a value of 4. For the interviewer's assessment of CI, "no CI" was assigned a value of 0, "may have CI" a value of 1,and "has CI" a value of 2. In the 2000 interview, there is a fourth classification of interviewer assessment (please refer to variable G517), and it is classified as "has CI." Combining these three variables results in a scale ranging from 0 to 11, with higher scores indicating worse cognitive function.

The interviewer's assessment of cognitive impairment question was not asked for any wave prior to the 2000 interview; therefore, for waves 1995-1998, we created a 9-point scale by adding only the proxy assessment of memory and IADL limitations. The 11-point scale and the 9-point scale were combined into a single variable that ranges from 0-11 or 0-9, depending on the year cognition was assessed (variable name *prxyscore_imp*). Therefore, beginning the 2000 HRS interview, we used the 11-point scale and 9-point scale prior to the 2000 HRS interview to classify proxy-respondents into Normal, CIND, or Demented. This categorization is combined with the non-proxy categorization so that both proxy and non-proxy interview information is used to create the summary categorization (variable name: *cogfunction*).

Lastly, note that the dataset also contains all the variables used to create the summary scores with imputed values. For web cases, the mode adjustment results in a value for *cogtot27_imp* that is one less than the sum of the constituent measures.

The classifications and their relevant cut points are summarized in Table 1. The variable names and labels included in this dataset are in Table 2. The data are released in wide format, but the code to reshape into long format is included at the end of this document.

Note:

1) You can use the variable "interviewX = 1" for X = Study year to extract a cross-sectional data.

Papers that use the Langa-Weir Classifications

Davydow D, Levine D, Zivin K, Katon W, Langa KM. The association of depression, cognitive impairment without dementia, and dementia with risk of ischemic stroke: A cohort study. Psychosomatic Medicine. 2015; 77(2): 200-208.

Langa KM, Larson E, Crimmins E, Faul J, Levine D, Kabeto M, Weir D. A comparison of the prevalence of dementia in the United States in 2000 and 2012. JAMA Internal Medicine. 2017; 177(1): 51-58.

Wei M, Kabeto M, Langa KM, Mukamal K. Multimorbidity and physical and cognitive function: performance of a new multimorbidity-weighted index. J Gerontol A Biol Sci Med Sci. 2018; 73(2): 225-232.

Table 1. Cut-points on the self-respondent 27-point scale and proxy-respondent cognitive scales

	Self-Respondent	Proxy Respondent	
Cognitive Function	**27-point scale (1995 onward)**	**9-point Scale 1995 – 1998**	**11-point Scale (2000 onward)**
Normal	12 – 27	0 – 2	0 – 2
CIND	7 – 11	3 – 4	3 – 5
Demented	0 – 6	5 – 9	6 – 11

Table 2. Variable names and labels for Langa-Weir classification dataset

Variable	Type	Label
HHID	Char	HRS HOUSEHOLD IDENTIFIER
PN	Char	HRS PERSON NUMBER IDENTIFIER
bwc20_impX	Num	BACKWARDS COUNT FROM 20
cogfunctionX	Num	Cognition Category: 1=Normal, 2=CIND, 3=Demented
cogivewmodeX*	Num	Cognition Interview mode (1=FTF/TEL, 2=WEB, 3=Proxy R)
cogtot27_impX	Num	TICS-m 27-point scale
dlrc_impX	Num	DELAYED WORD RECALL
fbwc20_impX	Num	IMPFLAG: BACKWARDS COUNT FROM 20
fdlrc_impX	Num	IMPFLAG: DELAYED WORD RECALL
fimparate_impX	Num	IMPFLAG: Interviewer assessment of proxy
fimrc_impX	Num	IMPFLAG: IMMEDIATE WORD RECALL
fmemoryp_impX	Num	IMPFLAG: Proxy Assessment of Memory
fnumiadl_impX	Num	IMPFLAG: NUMIADLs
fser7_impX	Num	IMPFLAG: SERIAL 7S
imparate_impX	Num	Interviewer's assessment of proxy
imrc_impX	Num	IMMEDIATE WORD RECALL
interviewX	Num	Interviewed this Wave
memoryp_impX	Num	Proxy Assessment of Memory
numiadl_impX	Num	NUMIADLs
proxyX	Num	PROXY TYPE: 1=Spouse, 2=Others, 5=Self-Resp
prxyscore_impX	Num	Total: Proxy Score (<2000 0-9; >=2000 0-11)
ser7_impX	Num	SERIAL 7S
Study_cohort	Num	Study cohort

X=1995, 1996, 1998, 2000, 2002, 2006, 2008, 2010, 2012, 2014, 2016, 2018, 2020, & 2022. For imparate_impX and fimparate_impX, HRS began collecting in 2000.

*This variable is available beginning with the 2018 interview. You may use this variable to identify those who completed the web-based interview.

SAS codes to reshape data from wide to long format

```
/**Macro to reshape data from wide to long format*/
libname cog "...\Cognitive function\data\";

*****************************************************************************;
*Note: 1) Interviewer assessment started in the HRS 2000  *;
*                  2) Web-based cognition measure started in 2018                                    *;
*****************************************************************************;
data work.cogfunc9522;
set cog.cogfinalimp_9522wide;
run;


*************************************************************;
***Wide to long*******************;
*************************************************************;
%macro cognl(year);
data work.cogd&year (keep= hhid pn studyyr Study_cohort firstiw numiadl_imp imparate_imp memoryp_imp fimrc_imp fdlrc_imp
fser7_imp fbwc20_imp imrc_imp dlrc_imp
ser7_imp bwc20_imp cogtot27_imp proxy cogfunction fnumiadl_imp fmemoryp_imp fimparate_imp prxyscore_imp cogivewmode
interview);
set cogd1.cogfinalimp_9522wide;
if interview&year=1;
studyyr=&year;

rename numiadl_imp&year=numiadl_imp imparate_imp&year=imparate_imp memoryp_imp&year=memoryp_imp
fimrc_imp&year=fimrc_imp
fdlrc_imp&year=fdlrc_imp fser7_imp&year=fser7_imp fbwc20_imp&year=fbwc20_imp imrc_imp&year=imrc_imp
dlrc_imp&year=dlrc_imp
ser7_imp&year=ser7_imp bwc20_imp&year=bwc20_imp cogtot27_imp&year=cogtot27_imp proxy&year=proxy
cogfunction&year=cogfunction fnumiadl_imp&year=fnumiadl_imp
fmemoryp_imp&year=fmemoryp_imp fimparate_imp&year=fimparate_imp prxyscore_imp&year=prxyscore_imp
cogivewmode&year=cogivewmode interview&year=interview;
run;
proc sort data=work.cogd&year;
 by hhid pn;
%mend;
```

%*cognl*(1995); %*cognl*(1996); %*cognl*(1998); %*cognl*(2000); %*cognl*(2002); %*cognl*(2004); %*cognl*(2006);
%*cognl*(2008); %*cognl*(2010); %*cognl*(2012); %*cognl*(2014); %*cognl*(2016); %*cognl*(2018); %*cognl*(2020); %*cognl*(2022);

```
data cogd1.cognfinalimp_9522long;
set work.cogd1995 work.cogd1996 work.cogd1998 work.cogd2000 work.cogd2002 work.cogd2004 work.cogd2006 work.cogd2008
   work.cogd2010 work.cogd2012 work.cogd2014 work.cogd2016 work.cogd2018 work.cogd2020 work.cogd2022;

label   numiadl_imp          ="NUMIADLs"
                memoryp_imp         ="Proxy Assessment of Memory"
                imparate_imp  ="Interviewer assessment of proxy"
                fimrc_imp           ="IMPFLAG: IMMEDIATE WORD RECALL"
                fdlrc_imp           ="IMPFLAG: DELAYED WORD RECALL"
                fser7_imp           ="IMPFLAG: SERIAL 7S"
                fbwc20_imp          ="IMPFLAG: BACKWARDS COUNT FROM 20"
                imrc_imp            ="IMMEDIATE WORD RECALL"
                dlrc_imp            ="DELAYED WORD RECALL"
                ser7_imp            ="SERIAL 7S"
                bwc20_imp           ="BACKWARDS COUNT FROM 20"
                cogtot27_imp  ="TICS-m 27-point scale"
                proxy               ="PROXY TYPE: 1=Spouse, 2=Others, 5=Self-Resp"
                cogfunction         ="Cognition Category: 1=Normal, 2=CIND, 3=Demented"
                fnumiadl_imp ="IMPFLAG: NUMIADLs"
                fmemoryp_imp        ="IMPFLAG: Proxy Assessment of Memory"
                fimparate_imp="IMPFLAG: Interviewer assessment of proxy"
                prxyscore_imp="Total: Proxy Score (<2000 0-9; >=2000 0-11)"
                cogivewmode         ="Cognition Interview mode  (1=FTF/TEL, 2=WEB, 3=Porxy R)"
                interview           ="Interviewed this Wave"
                studyyr                     ="Study year";

run;
```

Stata codes to reshape data from wide to long format

```
use ". . .\cogfinalimp_9522wide.dta"

reshape long bwc20_imp cogfunction cogivewmode cogtot27_imp dlrc_imp fbwc20_imp fdlrc_imp fimparate_imp fimrc_imp
fmemoryp_imp fnumiadl_imp fser7_imp imparate_imp imrc_imp interview ///
        memoryp_imp numiadl_imp proxy prxyscore_imp ser7_imp, i(hhid pn) j(studyyr)

*Reshape create the same number of observation for each wave. Thus, needs to keep only those interviewed in that wave****
keep if interview==1  /*keeping only those interviewed**/

label var  bwc20_imp       "BACKWARDS COUNT FROM 20"
label var  cogfunction     "Cognition Category: 1=Normal, 2=CIND, 3=Demented"
label var  cogtot27_imp    "TICS-m 27-point scale"
label var  dlrc_imp        "DELAYED WORD RECALL"
label var  fbwc20_imp      "IMPFLAG: BACKWARDS COUNT FROM 20"
label var  fdlrc_imp       "IMPFLAG: DELAYED WORD RECALL"
label var  fimparate_imp   "IMPFLAG: Interviewer assessment of proxy"
label var  fimrc_imp       "IMPFLAG: IMMEDIATE WORD RECALL"
label var  fmemoryp_imp    "IMPFLAG: Proxy Assessment of Memory"
label var  fnumiadl_imp    "IMPFLAG: NUMIADLs"
label var  fser7_imp       "IMPFLAG: SERIAL 7S"
label var  hhid            "HOUSEHOLD IDENTIFICATION NUMBER"
label var  imparate_imp    "Interviewer assessment of proxy"
label var  imrc_imp        "IMMEDIATE WORD RECALL"
label var  interview       "Interviewed this Wave"
label var  memoryp_imp     "Proxy Assessment of Memory"
label var  numiadl_imp     "NUMIADLs"
label var  pn              "RESPONDENT PERSON IDENTIFICATION NUMBER"
label var  proxy           "PROXY TYPE: 1=Spouse, 2=Others, 5=Self-Resp"
label var  prxyscore_imp   "Total: Proxy Score (<2000 0-9; >=2000 0-11)"
label var  ser7_imp        "SERIAL 7S"
label var  cogivewmode            "2018-:Cog Interview mode (1=Self-R (FTF/TEL), 2=Self-R WEB, 3=Porxy-R (FTF/TEL)"
label var  studyyr         "Study year"
label var  study_cohort                    "Study Cohort"
```