



HEALTH AND RETIREMENT STUDY
A Longitudinal Study of Health, Retirement, and Aging
Sponsored by the National Institute on Aging

*Protocols for Missing Data in the
Harmonized Cognitive Assessment
Protocol (HCAP)*

Documentation Report

Ryan J. McCammon
Steven G. Heeringa
Richard N. Jones
Lindsay C. Kobayashi
Kenneth M. Langa
Deborah A. Levine
Jennifer J. Manly
Lindsay R. Ryan
David R. Weir

Survey Research Center
Institute for Social Research
University of Michigan
Ann Arbor, Michigan

2025

Funding

The HCAP Network is supported by the National Institute on Aging through the Division of Social and Behavioral Research (NIA U24AG065182).

Suggested Citation

McCammon, R. J., Heeringa, S. G., Jones, R. N., Kobayashi, L. C., Langa, K. M, Levine, D. A., Manly, J. J., Ryan, L. R., & Weir, D. R. (2025). Protocols for Missing Data in the Harmonized Cognitive Assessment Protocol (HCAP). University of Michigan.

<https://hrs.isr.umich.edu/publications/biblio/12950>

Protocols for Missing Data in HCAP

The Harmonized Cognitive Assessment Protocol (HCAP) is part of an ongoing international research collaboration funded by the National Institute on Aging (NIA) (R01 AG051142) to measure and understand dementia risk within ongoing longitudinal studies of aging around the world. The HCAP collects a carefully selected set of established cognitive and neuropsychological assessments and informant reports to better characterize cognitive function in older adults.

In 2016, the Health and Retirement Study Harmonized Cognitive Assessment Protocol (HRS-HCAP) instrument was administered to a subsample of HRS study respondents aged 65 and older, yielding 3,347 respondent assessments and 3,183 informant reports, including 313 respondent-only cases and 149 informant-only cases, for a combined sample size of 3,496.

In contrast to the typical survey interview, item nonresponse in neuropsychological assessments is often informative of the respondent's task performance, particularly when the item responses are scored as either correct or incorrect. This is clearly the case when a "don't know" response is given to a factual question such as "What is the day of the week?" This logic can be extended to other tasks in neuropsychological exams and to other forms of nonresponse until the coding guidance becomes "a question asked is a question answered." The boundaries of this approach are necessarily quite nuanced, and the viability of the approach is contingent on the interview protocol.

HRS HCAP Imputation Philosophy

The purpose of this document is to report on a protocol for the handling of missing data in HRS-HCAP, applying deterministic assignment ("recoding") when the missingness is informative of the respondent's ability to answer correctly, while using probabilistic imputation when the interview protocol and response provided, if any, leave ambiguity as to the respondent's score on a particular task.

The literature on the coding of non-response to neurological test items in the survey setting is remarkably consistent. Based on a sample aged 60+ that was administered the MMSE, Fillenbaum, George and Blazer (1988) find that "in epidemiological surveys, scoring nonresponse as error is more likely to be correct." Similarly, A. Regula Herzog, one of the original architects of the HRS cognition assessment, analyzed the data from the 1993 AHEAD cohort (age 70+) and found that nonresponse relative to item difficulty varied with respondent's cognitive ability, with low cognition respondents being more likely to answer "don't know" to difficult questions (Knäuper, Belli, Hill, and Herzog, 1997). Using the same data, Herzog and Wallace (1997) found that respondents refusing cognition items had worse cognitive and physical health than did the lowest performing of non-refusing respondents, concluding that "imputing incorrect answers or relatively low scores for those who refused might represent a first approach to preserving the full data base, while not introducing major distortions into the data."

Following from these studies, the general presumption guiding this work is that, in the context of neuropsychological tests, once a stimulus is provided and the respondent has had an opportunity to answer, nonresponse is consistent with an incorrect response and no imputation is needed. More succinctly, for some tests, there is no meaningful distinction between a "don't know" and "refused" response. When a respondent had the opportunity to provide the correct answer but did not, a score of zero or incorrect can reasonably be assigned, while when the reluctance to complete a task (a statement of "don't know" or "refused" from the respondent or a complete breakoff of the interview) occurs *prior to the opportunity to answer*, then there is a need for imputation.

The level of engagement that constitutes the opportunity to respond varies by test. For instance, for word list and story recall tasks, having been presented with the list or story is sufficient. Any nonresponse after that presentation should be scored as incorrect. For task-based measures like verbal fluency, letter cancellation, trail making, and others, imputation is generally needed for all nonresponse prior to the respondent beginning the task, but once the respondent commences work on the task, it should be scored as complete without imputation. An exception to this rule occurs when there is repeated refusal of individual items within a test, suggesting that the respondent has ceased attempting to correctly answer the items. In such instances, the test is scored as incomplete, and the refused items are imputed.

Respondent-only & Informant-only cases

Assignment and imputation of HRS-HCAP scores is limited to cases in which the respondent or informant completed at least the portion of the interview required for inclusion in the dataset as a partial interview. Scores for the respondent interview were not imputed when the respondent assessment was not conducted (the 149 informant-only cases in the 2016 HRS-HCAP). Values in the informant interview were not imputed for cases in which the informant assessment was not conducted (the 313 respondent-only cases in the 2016 HRS-HCAP).

Partial interviews

On rare occasions, interviews may end before the instrument is completely administered. This typically results from the informant or respondent refusing or otherwise being unable to continue with the interview. There are a small number of cases (< 1% of respondent interviews and even fewer informant interviews) that breakoff prior to completion of the interview. In order to be included in the respondent dataset, the interview had to include a completed MMSE. In order to be included in the informant dataset, the interview had to include a completed Jorm IQCODE. Missing data for partial interviews is imputed in the 2016 HRS-HCAP final release datasets.

Multiple item tests with repeated refusals

Word list recognition, logical memory recognition, number series, and Raven's progressive matrices are tests that involve multiple discrete items of the same format. If the respondent began a test with multiple discrete items, but at some point refused each new item and ceased to give any effort to provide a correct response, then the refused items should be imputed. Specific guidance on this circumstance is provided later in this document.

Imputation of "Skipped by design", "Did Not Understand Instruction", "Not Attempted, Interviewer decision", and "R is unable to complete test"

If the completion code for a test indicates that the respondent did not understand the instruction, or that the test was skipped by design or by interviewer decision, then the respondent did not attempt to complete the task, and the test score should be imputed.

2016 HRS-HCAP recoding of missing values summary

Mini-Mental State Exam (MMSE)

Interviewers were given the following coding instructions, per the administration guidelines from Psychological Assessment Resources, the distributor of the MMSE:

Code 97= Not assessed is used when:

- the respondent has not even attempted to work through an answer, or
- the answer is Don't Know or Refused

Accordingly, in the 2016 HRS-HCAP, most item-missing data for the MMSE was coded "97. Not Assessed." Consistent with the interviewer instructions, item-level codes of 97 (NA), 98 (DK) and 99 (RF) should be treated as "incorrect." The MMSE total score is the total number of correct responses to the MMSE items.

SELECTED NAMING ITEMS FROM THE HRS

This series of three items has an introduction read by the interviewer - "Now I'm going to ask you for the names of some people and things." If the respondent declines the test at this point with a DK or RF response, or refuses all three individual items, then the total score should be imputed. For all others, an item asked of the respondent is scored as "correct" if the correct response is given, and as "incorrect" if otherwise.

CERAD WORD LIST IMMEDIATE (3 trials)

Each of the 3 trials has an introduction prior to the presentation of the words (Trial 1: "I am going to show you ten printed words. Read each word out loud as I show it to you. Later I will ask you to recall all ten words. Is this clear?"). For responses of DK or RF at the time of the introduction, and prior to the presentation of the words for a trial, imputation should be used. If the word list is presented, then the score for the trial is the number of words correctly recalled.

ANIMAL FLUENCY

As a task-based exercise, respondents who begin the task get the score they achieve, while those who do not start the task receive an imputed score. If there is interference or interruption during the timed portion of the task that affects administration or performance, then the total score should be imputed. If the respondent declines the task at the introduction (DK/RF) or provides no words (correct or incorrect) or does not understand the task, then the Animal Fluency total score should be imputed.

LETTER CANCELLATION

As a task-based exercise, respondents who begin the task get the score they achieve, while those who do not (e.g. due to refusing before having the opportunity to begin the task) receive an imputed score. If there is interference or interruption during the timed portion of the task that affects administration or performance, then the total score should be imputed. Letter Cancellation is an unusual task-based exercise in that the respondent is asked to write on a worksheet rather than responding verbally. Given the task and its complexity, imputation should be used for all for all cases who do not start the task (refused, did not understand instructions, partial IW, form lost, not attempted by decision of IWER).

BACKWARDS COUNT

As a task-based exercise, respondents who begin the task get the score they achieve, while those who do not receive an imputed score. If there is interference or interruption during the timed portion of the task that affects administration or performance, then the total score should be imputed. Any refusal, don't know, or Did Not Understand Instruction response prior to beginning the task should be imputed, along with partial interviews, and Not attempted, IWER decision cases.

Scoring note: The scored number of correct and incorrect counts is based on the number of occurrences of a correct or incorrect count, not the numeric size of any gaps. For example, 100 – 99 – 98 – 96 – 95 and 100 – 99 – 98 – 87 – 86 are both scored as 3 correct counts and 1 incorrect count.

1066 NAMING AND REASONING

For each of these four items, treat DK/RF as incorrect. Impute an item score only if the item is not asked.

CERAD WORD LIST DELAYED

If the interview reaches the point of the delayed word recall task (not a partial interview with a prior breakoff point) and the respondent was presented with the word list in all three of the immediate recall trials, then the delayed word recall task can be scored without need for imputation. Under these circumstances, all DK/RF should be coded 0 (zero – none remembered). If the respondent did not receive the word list in all three immediate trails, then the delayed word recall score should be imputed.

BRAVE MAN IMMEDIATE / LOGICAL MEMORY IMMEDIATE

As a recall exercise, imputation should be used only for nonresponse indicated prior to the presentation of the story. Any nonresponse after the respondent has heard the story should be scored as incorrect.

Scoring notes:

Respondents not attempting Brave Man Immediate (DK/RF prior to the presentation of the story) were skipped out of the Logical Memory Immediate test and coded “7. Skipped by design” on the Logical Memory completion indicator (R1LMB_IMMTEST). For these cases, the Logical Memory score should be imputed.

The HRS-HCAP test completion indicators for Logical memory and Brave Man (immediate and delayed) include a code of “9. Imputed”. This imputation is distinct from the imputation of missing data indicated by the standard imputation flags (R1FBM_DELScore, for example). For the story recall tests, HRS-HCAP revised the field scoring of the items by listening to recordings of the interviews. For those cases coded as “9. Imputed”, an intelligible recording was not available, and the revised story recall score was imputed based on the field scoring of the number of story details correctly recalled.

CERAD WORD LIST RECOGNITION

For this recognition task, we impute a total score for DK/RF responses that occur during the test introduction, before the administration of the actual test questions. At the item level, treat DK/RF as incorrect responses, but if a respondent refuses all twenty items, then the word recognition summary scores should be imputed. In the case of repeated refusal of items after beginning the test, see Section III of the HRS-HCAP imputation summary later in this document.

CONSTRUCTIONAL PRAXIS IMMEDIATE

Respondents are asked to copy 4 shapes, with 2, 3, 2, and 4 scoring points assigned to each, respectively. For each shape, there is a code “97. Respondent Cannot Draw.” The number of respondents with this code varies between shapes, and while it may include respondents who are physically incapable of drawing, the prevalence of this coding suggests that it also includes many others. If the respondent refuses all four tasks or is coded as “Cannot Draw” for all four tasks, then impute. Otherwise, treat DK/RF/cannot draw codes as incorrect responses.

SYMBOL DIGIT MODALITIES TEST

As a task-based exercise, respondents who begin the task get the score they achieve, while those who do not attempt the test receive an imputed score. If there is interference or interruption during the timed

portion of the task that affects administration or performance, then the total score should be imputed. Impute the total number correct and total number attempted for all respondents not attempting at least one item.

Note: The number incorrect is calculated as the difference between the total number attempted and the total number correct.

CONSTRUCTIONAL PRAXIS DELAYED

Respondents are asked to draw the four shapes from memory. A number of respondents indicated that they could not remember any of the shapes and should be given a score of zero. All cases with a non-missing score for Constructional Praxis Immediate should be assigned a score for Constructional Praxis Delayed. Individuals who did not attempt Constructional Praxis Immediate should be given an imputed score for the delayed task.

BRAVE MAN DELAYED / LOGICAL MEMORY DELAYED

If the respondent cannot remember the story, then score as zero. Apart from partial interviews, all cases that heard the story as part of the immediate story recall task should be assigned a score for the delayed story recall task. If the immediate task was not administered, the delayed task score is imputed.

LOGICAL MEMORY RECOGNITION

Impute for DK/RF responses that occur during the task introduction, prior to the presentation of the first story point in the recognition test. At the item level, treat DK/RF as incorrect responses, however, if a respondent refuses all fifteen items, then the recognition summary score should be imputed. In the case of repeated refusal of items after beginning the test, see Section III of the HRS-HCAP imputation summary later in this document.

NUMBER SERIES

For those completing the test, item responses of DK/RF are treated as incorrect. In the case of repeated refusal of items after beginning the test, see Section III of the HRS-HCAP imputation summary later in this document.

RAVEN'S PROGRESSIVE MATRICES

At the item level, treat DK/RF as incorrect responses, however, if a respondent refuses all seventeen items, then the Raven's summary score should be imputed. In the case of repeated refusal of items after beginning the test, see Section III of the HRS-HCAP imputation summary later in this document.

TRAIL MAKING A&B

For the trail making tests, missingness is handled via imputation. Impute for all cases with completion statuses of RF/ DK/Did Not Understand Instruction/Partial IW/ Not Attempted, Iwer Decision/ Skipped by Design/ Form Lost/Not Recorded. Guidance on the imputation can be found in Section IV of the HRS-HCAP imputation summary later in this document.

2016 HRS-HCAP imputation summary

When direct assignment of values for missing data is inappropriate or the information necessary for doing so is lacking, then imputation is used to fill-in missing values. In the HRS-HCAP respondent assessment, this typically occurs when a test is not attempted, and the imputation is carried out at the test-level. In contrast, missingness in the informant interview is generally found for only a few, if any, of the items in a given battery, and item-level imputation is more appropriate. The imputation sequence described below is to first impute as needed to replace missing values in the variables from the 2016 HRS core interview that are used solely as covariates in the imputation model, and then move on to imputing the HRS-HCAP respondent test scores, and finally the HRS-HCAP informant items.

The imputation strategy for the 2016 HRS HCAP utilizes fully conditional specification (FCS) methods to singly impute all modeled variables using SAS PROC MI. The method included using predictive mean matching (PPM) for continuous variables, and logit and generalized logits models for binary and categorical variables, respectively.

I. Impute any missing 2016 HRS core items used in the 2016 HCAP imputation.

The HCAP respondent imputation model uses several variables from the 2016 HRS core as covariates. 2016 HRS core cognition is based on a scaled version of the Langa-Weir score (LW2SELF) that includes the eight orientation items and a four-level classification based on this score.

$LW2SELF = 15 + (3 * \text{orientation (8 pts)}) + (3 * \# \text{ of IADL impairments (5 pts)}) + \text{immediate recall (10 pts)} + \text{delayed recall (10 pts)} + \text{serial 7s (5 pts)} + \text{backward count (2 pts)}$

LW2DX is a four-level classification variable based on LW2SELF with categories 0-36, 37-47, 48-52, and 53+.

For HRS core proxy cases, LW2DX is calculated based on the proxy's assessment of the respondent's memory, the interviewer's assessment of the respondent's cognitive impairment, IADL limitations, and the Jorm score.

The same cognition variables calculated from 2014 HRS core data were also included in the 2016 HRS core imputation model.

The imputation model also includes other variables from the HRS Tracker file and the 2016 HRS core interview: age, race, gender, IW language, coupleness, income quintile, wealth quintile, schooling, parent's schooling, subjective health, health conditions (hypertension, heart disease, stroke, diabetes, dementia), rate vision, rate hearing, assisted living/nursing home residence, Nagi score, IADLs, and ADLs.

The model includes linear, squared, and cubic age terms. Respondent schooling included years of schooling as well as categorical degree statuses (high school, college degree, advanced degree).

II. Impute missing 2016 HRS-HCAP respondent scores.

Typically, when an HCAP respondent cognitive test has missing data, all item variables for the given test are missing. In this case, no data from partial completion of the test can inform the test score imputation. For this reason and for ease of estimation, the next step in the 2016 HCAP imputation sequence involves the imputation of summary scores for the cognitive tests.

In addition to the variables from the 2016 HRS core (Step I), the model includes the following measures from the 2016 HCAP informant interview: informant reports of stroke, Parkinson's, Alzheimer's, and memory difficulties, along with the CSID relscore, Blessed, and Jorm. For purposes of the respondent scores imputations, these informant summary measures were simply imputed in the same model for respondent-only cases and other instances of missing informant data.

Using the variable names from the 2016 HRS HCAP public release, the model includes all R1xxx_SCORE and R1xxx_DELSCORE variables with the following exceptions:

1. CESD and smells: The imputation model does not include R1CESD11_SCORE or R1SMELL_SCORE.
2. Number series: R1NS_SCORE. Given the block adaptive design of this test, the model includes the number series test as two variables – the number of correct responses (0-3) in each of the problem sets, along with a binary indicator of the NS list assignment. These values are later recoded to the NS scale scores.
3. Word list recognition: The model imputes word list recognition using R1WREC_ORG and R1WREC_FOIL, which sum to R1WREC_TOTSCORE.
4. Letter cancellation: The model includes R1LC_SCORE, along with R1LC_INCORR and R1LC_MISS: Due to difficulties with model estimation, these variables were imputed as categorical variables and then randomly assigned a discrete value within the range of each category.
5. Brave man: The model includes both gist and exact scoring summary variables: R1BM_IMMSCORE, R1BM_IMMEX, R1BM_DELSCORE, and R1BM_DELEX.

III. Special handling of multiple item tests with repeated refusals.

The codeframe for the R1xxx_TEST variable includes a code “12. Test scored as missing based on repeated refusal of individual test items.” This code value, used only for word list recognition, logical memory recognition, number series, and Raven's, indicates that the respondent began a test with multiple discrete items, but at some point, refused each new item and ceased to give any effort to providing a correct response. In these instances, the imputed summary score from Step II is discarded, and the individual test items are imputed as needed. The specific conditions that trigger this coding all involve the respondent beginning the task (providing correct/incorrect responses, including DK, and then refusing to attempt any further items in the test as follows:

1. Word list recognition. Refusal of the last three or more items (items 18-20).
2. Logical memory recognition. Refusal of the last three or more items (items 13-15).
3. Number series. Refusal of the last three items (items 4-6).
4. Raven's. Refusal of items B6, B8, and B10.

In these circumstances, the individual items were imputed, starting with the first item in the final series of refusals. The imputations were conducted separately for each test, using the dataset created in Step II, with the specific test summary score replaced by the individual test items. For example, R1LMB_RECOSCORE was removed from the model and replaced by R1LMB_RECO1 - R1LMB_RECO15.

For the number series items, first impute items 1-3 and then impute items 4-6, including the total score for items 1-3 as a predictor in the model.

IV. Special handling of Trails A & B.

Since the respondent may not have completed Trails A and B in the allotted five-minute task time, the imputed values from Step II were discarded and instead imputed as follows:

1. Create a binary test completion indicator (TA_COMP) for Trails A, recoding R1TMA_TEST as “Complete” (1) = 1, “Unable to complete in 5 minutes” (13) = 0, with all other R1TMA_TEST values set missing on the completion indicator.
2. Impute TA_COMP, excluding all other Trails variables from the model.
3. Impute R1TMA_SCORE among Trails A completers (based on the imputed value of TA_COMP).
4. Create TB_COMP in the same way as TA_COMP, using R1TMB_TEST while also coding “7. Skipped By Design” = 0.
5. Impute TB_COMP, including the imputed R1TMA_SCORE as a predictor in the model.
6. Impute R1TMB_SCORE among Trails B completers (based on the imputed value of TB_COMP).

V. Impute missing 2016 HRS-HCAP informant scores.

In contrast to the HCAP respondent imputations, primarily carried out at the test-level, the nature of the items and the patterns of item completion in the informant interview make imputation at the item-level the preferred approach for the informant data. For example, although most informants answered all or nearly all the Jorm items, almost 7% did not report on the respondent’s handling of money for shopping. Given the mostly complete item-level data, it makes sense to impute items as needed, making use of the other item-level responses provided.

The 2016 HRS HCAP informant data imputations proceeded in two basic steps. First, the bulk of the informant items were imputed using a model including the items asked of all informants (including the “stem” items for questions that involve branching). Other predictors in the model included the 2016 HRS core items from Step I and the final versions of the imputed respondent tests scores from Step IV. For purposes of the informant item imputations, the respondent summary scores for informant-only cases were imputed as necessary using the same model.

Based on the results of these imputations, the conditional, or “branch” items were then imputed, using similar logic to the handling of the Trails scores in Step IV. This includes items that ask about the circumstances of an activity, given that the activity occurs (INF1ACT13 and INF1ACT14, for example), as well as items that attribute cause to an impairment (like INF1BL1_1 and INF1BL1_1A).

References

Fillenbaum GG, George LK, Blazer DG. Scoring Nonresponse on the Mini-Mental State Examination. *Psychological Medicine*. 1988;18(4):1021-1025.

Herzog AR, Wallace RB. Measures of Cognitive Functioning in the Ahead Study. *The Journals of Gerontology: Series B*. 1997;52B(Special Issue):37-48.

Knäuper B, Belli RF, Hill DH, Herzog AR. Question Difficulty and Respondents' Cognitive Ability: The Effect on Data Quality. *Journal of Official Statistics*. 1997;13(2):181.