The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences, 2025, **80**(S1), gbae149
https://doi.org/10.1093/geronb/gbae149
Advance access publication: 29 August 2024
Supplement Article: Studying Cognition in the National Social Life, Health, and Aging Project: Methodological and Analytic Considerations, Strategies and Findings From Round 4

Measuring Cognitive Function In-Person and Remotely in Round 4 of the National Social Life, Health, and Aging Project

Kelly Pudelek, MA,[1,*] L. Philip Schumm, MA,[2] Jennifer Hanis-Martin, PhD,[1] Melissa J. K. Howe, PhD,[1] and Terese Schwartzman, PhD[1]

[1]NORC at the University of Chicago, The Bridge, Chicago, Illinois, USA.
[2]University of Chicago, Center for Translational Data Science, Chicago, Illinois, USA.
*Address correspondence to: Kelly Pudelek, MA. E-mail: pudelek-kelly@norc.org

Decision Editor: Hui Liu, PhD (Social Sciences Section)

Abstract

Objectives: This paper describes the changes made to the collection of cognitive measures when the National Social Life, Health, and Aging Project (NSHAP) introduced remote modes of data collection.

Methods: In Round 4 (2021–2023), the longitudinal study transitioned from being conducted in-person to collecting data via multiple modes including in-person and remote modes: web, phone, and paper-and-pencil. The team began with the measures used in Rounds 2 and 3 of NSHAP—the survey-adapted Montreal Cognitive Assessment (MoCA-SA)—and evaluated which measures could be administered remotely, introducing new measures for each cognitive subdomain, as needed, to compensate for items that could not be administered remotely.

Results: Cognitive items used in Rounds 2 and 3 that could not be administered remotely were dropped from the respective modes, and items selected from the Rush Alzheimer's Disease Center's (RADC) global cognition battery were added as substitutes. For comparison, the RADC substitute items were added to the in-person mode making it longer in Round 4.

Discussion: The changes in cognitive measures resulted in different numbers of cognitive items across the 4 modes of survey administration in Round 4. Analysts should be aware of these changes when creating a single global cognition score for the entire NSHAP sample in Round 4, and aware that there may be mode effects that could affect cognition scores.

Keywords: Alzheimer's disease, Cognition measures, Longitudinal methods, Measurement

Cognition is a fundamental component of older adults' health and well-being and is an integral part of the survey measures collected by the National Social Life, Health, and Aging Project (NSHAP; Shega, et al., 2014). The proportion of the population that is aging into older adulthood and experiencing Alzheimer's disease (AD) and Alzheimer's disease-related dementias (ADRD) is growing. The Alzheimer's Association (2023) estimates that 6.7 million Americans 65 and older are living with AD and that this number may nearly double by 2060 given current trends. This anticipated growth in the number of individuals living with dementia has substantial implications for their loved ones and society at large. NSHAP data provide an opportunity to examine changes in cognitive functioning as people age along with changes in their physical and mental health and their social relationships.

Similar to other population-based studies of health and aging, which include assessments of cognitive function such as the Health and Retirement Study, NSHAP administered cognitive function measures in-person during its first three rounds of data collection in 2005–2006, 2010–2011, and 2015–2016 (Kotwal et al., 2015; Shega et al. 2014). However, increased travel costs associated with conducting face-to-face interviews across the United States together with new and improved methods for remote data collection led NSHAP to move to a hybrid strategy of remote and in-person (as necessary) data collection for Round 4.

Prior to the coronavirus disease 2019 (COVID-19) pandemic, NSHAP had planned to implement remote data collection in Round 4 among a subset of respondents in parallel with in-person data collection, to permit evaluating the feasibility, efficacy, and validity of collecting cognitive data remotely. These initial plans were then altered due to the Round 4 field period coinciding with the COVID-19 pandemic; all data collection was delayed a year, and in-home visits were not possible for the first several months of the field period. Thus remote administration became crucial for collecting Round 4 data.

In this paper, we describe our overall approach to adapting the NSHAP cognition module to permit both remote and in-person administration in a consistent way yielding data that are comparable with those collected in previous rounds. We discuss both the implementation of each specific measure

Received: April 1 2024; Editorial Decision Date: August 13 2024.

as well as general challenges encountered with remote administration using web, phone, and paper-and-pencil modes. Finally, we discuss the implications of the Round 4 data collection strategy for analyses of the resulting data, both cross-sectionally and together with the data from previous rounds.

Remote Data Collection

Remote data collection has certain advantages over in-person when it comes to survey-based research. First, it can significantly reduce the costs of survey administration by eliminating the need for interviewers to travel to participants' homes to conduct face-to-face interviews. Data may instead be collected using a web-based instrument, a computer-assisted phone interview (sometimes abbreviated to CATI, but in this paper will be referred to simply as phone), or a paper-and-pencil interview (PAPI) returned by mail. Second, remote data collection by web and paper-and-pencil formats may improve response rates by allowing participants to complete the survey at a time of their choosing as opposed to having to schedule an in-person or phone interview with a field interviewer.

However, remote modes can also pose challenges to survey data collection and come with trade-offs. For instance, in web-based surveys, there is no opportunity for participants to ask interviewers to clarify questions, and this mode of survey administration may underrepresent individuals with lower levels of literacy (Hensen et al., 2021). In the case of phone-based surveys, similar to in-person interviews, there is potential for interviewer bias (Hensen et al., 2021). Both web and phone data collection require access to and effective use of technological devices, and in the case of web-based survey collection, Internet access is needed as well. All remote modes lack the potential benefits provided by an in-person interviewer whose personal touch and ability to build rapport can increase respondent cooperation and understanding.

Although most in-person survey questions can be adapted in a straightforward manner for web, phone, or paper-and-pencil administration, certain measures—including some of NSHAP's cognition measures—are more difficult to adapt. Yet with careful attention to question design and analytic strategy, it is possible to overcome these barriers to produce robust, valid, and reliable results. Previous studies provide examples highlighting the feasibility of remote administration of cognitive items. Collins et al. (2022) found that compliance, performance, and data quality of remote data collection were comparable to the in-person setting. Iiboshi et al. (2020) found that video teleconferencing was a viable alternative to face-to-face-administered Montreal Cognitive Assessment (MoCA) testing. Atkins et al. (2022) found equivalence for three of the four cognitive domains assessed between remote and site-based tests. Finally, D'Arcy et al. (2013) observed high completion rates for automated assessments of cognitive assessment tasks. These studies suggest that remote data collection of cognition measures is feasible and can produce results that are comparable to those conducted in-person.

To adapt NSHAP's in-person cognition measures for remote administration in Round 4, the team began with the measures used in Rounds 2 and 3 of NSHAP—the survey-adapted Montreal Cognitive Assessment (MoCA-SA; see Kotwal et al., 2015; Shega, et al., 2014). Because some of the items could not be administered by one or more of the web, phone, and paper-and-pencil remote modes, we introduced new measures of the same cognitive subdomains as the omitted items. The remaining items administered across all modes may be used to generate a harmonized measure of cognitive function for analysis, including longitudinal analyses using previous rounds. Further information that may be used to estimate and adjust for mode effects is available from a randomized experiment incorporated into the Round 4 data collection (see O'Muircheartaigh et al., 2021).

NSHAP Round 4: Data Collection Methods

NSHAP is a nationally representative longitudinal study of older community-dwelling adults in the United States that has conducted four rounds of data collection at approximately 5-year intervals since 2005 (NSHAP data sets are publicly available through the National Archive of Computerized Data on Aging, see Waite et al., 2017). In 2020, the year when Round 4 was scheduled to begin, a brief survey focusing on topics related to the COVID-19 pandemic was administered using the web only, and the fourth round of data collection was subsequently performed from 2021 to early 2023. Round 4 was therefore the first round to be administered using multiple modes. Of the 3,372 NSHAP R4 respondents, 1,903 (56%) completed the survey via a remote mode, web, phone, or PAPI, and 1,469 (44%) completed in-person (in-person questionnaire, or IPQ), in their home with an interviewer (see Table 1).

Initial mode assignments were made using available information (e.g., quality of address information, perceived level of cooperation, and incentive amount in previous rounds) in an attempt to maximize the overall response rate; specifically, those judged likely to complete via web were assigned initially to this mode because some respondents prefer this mode and it saves resources that can then be spent to complete more challenging cases. Those assigned initially to the web but who did not complete the survey were offered the opportunity to complete it using another remote mode (i.e., phone or PAPI) or an in-person interview. In addition, an experiment was performed whereby 400 randomly selected respondents initially assigned to the web proceeded directly to an in-person interview. This experiment permits unbiased evaluation and

Table 1. Number of Respondents by Mode of Completion

Mode of interview/questionnaire completion	Number completed[a]
Remote	**1,903**
Web	(1,245)
Phone	(469)
Paper-and-pencil interview (PAPI)	(144)
Hybrid[b]	(45)
In-person questionnaire (IPQ)[c]	**1,469**
Total completed	**3,372**

Notes:
[a]Numbers in bold represent totals. Total remote cases comprise Web, Phone, PAPI, and Hybrid, broken down below in parentheses. The total completed NSHAP cases are the sum of Remote + IPQ.
[b]Hybrid cases started in one of the remote modes but finished in another. These were mainly cases started on the web and finished on the phone with an interviewer.
[c]In four cases, respondents initially began in one of the remote modes, but later completed a full interview with an in-person interviewer, and counted as in-person.

estimation of mode effects on the responses to individual items.

Of the 2,794 respondents initially assigned to the web, 1,245 completed the web-based survey. Those who did not complete it were sent multiple reminders, and after 7 weeks were asked to complete the survey via an interviewer-administered phone interview. Four hundred sixty-nine additional respondents completed the phone survey. Those who did not complete either the web-based or phone-based survey within 12 weeks were then asked to complete the survey via a self-administered PAPI, and another 144 respondents completed the survey via this mode. There were also 45 cases that began in one mode and were completed in another; most of these began on the web and were finished on the phone (see Table 1). Remote data collection ended on February 15, 2022. Respondents who were initially assigned to remote but did not complete the survey were then visited by trained field interviewers to complete an in-person interview when the face-to-face interviews began in spring 2022.

The change to multimode administration required the NSHAP team to make decisions about how to administer each of the survey items remotely. Particularly challenging to adapt was the battery of cognition items. Those items designed to assess visuospatial ability (clock, trails) could not be implemented on the phone or web-based survey, and the verbal fluency item could not be implemented on the web or paper-and-pencil survey. To mitigate the loss of these items in the remote modes, several visuospatial ability/perceptual orientation items were chosen from the Rush Alzheimer's Disease Center's (RADC) global cognition battery as substitutes for the items lost and to fortify these subdomains.

Cognition Measures

In NSHAP Round 1, the Short Portable Mental Status Questionnaire was used to assess cognition, but due to the fact that it was not sensitive enough to detect mild-to-moderate cognitive decline, starting in Round 2, NSHAP switched to the use of a survey-adapted version of the MoCA, the MoCA-SA (Kotwal et al., 2015; previously referred to as the Chicago Cognitive Function Measure; see Shega et al., 2014). In both Rounds 2 and 3, the MoCA-SA was administered by trained field interviewers in respondents' homes. The respondents were asked to perform 18 different cognitive tasks including the modified trail-making test B, clock drawing test, naming a rhinoceros, forward and backward digit span, serial 7 subtractions, sentence repetition, phonemic fluency, abstraction, 5-word delayed recall, and orientation to assess executive function, language, working memory, episodic memory, temporal orientation, and visuospatial ability. These cognitive items, which had been successfully administered during in-person interviews, could not all be administered through web, phone, and paper-and-pencil. Items deemed too difficult to implement in a remote data collection context were excluded from the remote modes. Because all the cognitive items designed to assess the visuospatial ability cognitive domain were dropped from web-based and phone instruments, visual items including three progressive matrices and three-line orientation visual tasks were selected from RADC's battery of validated cognitive items to augment NSHAP's cognition measures. Items selected included three-line orientation items and three matrix items with various difficulty levels (based on Rush data) to provide a wide distribution of responses and higher sensitivity. The line and matrices were visual items with multiple choice answers so the respondent could see and complete them in web, paper, and in-person modes but not by phone. Refer to Table 2 for a full list of cognitive items used in Rounds 2 and 3, and in Round 4 across the different modes of administration. To help validate the inclusion of these items and to test for mode effects, all the newly added RADC cognitive items were also administered through the in-person interview (for discussion of mode effects, see Smith et al., this issue).

Cognitive Measures Across Modes

Each mode had different numbers and types of cognitive items based on which ones could be feasibly administered in each mode (Table 2). In Rounds 2 and 3, 12 cognitive items were administered in person. In Round 4, for the in-person survey, the cognitive items administered grew to 19, for the web-based survey, 12 cognitive items were administered, for the phone-based survey, 10 cognitive items were administered, and for the pencil-and-paper survey, 14 cognitive items were administered.

The different numbers of cognitive items across modes mean that adjustments need to be made to enable comparability of cognition scores across modes and to generate a single global cognition score for the entire sample which will be forthcoming in a future article. Additionally, analysts need to be aware of mode effects that could bias the responses. For instance, for the web and pencil-and-paper survey, both these modes were self-administered by respondents. Self-administration may make it more challenging

Table 2. Cognition Items Across Rounds 2 and 3, and Four Modes of Administration in Round 4

Cognitive items	Rounds 2 and 3	Round 4 remote modes			Round 4 IPQ
	IPQ	Web	Phone	PAPI	IPQ
Date	X		X	X	X
Month	X		X	X	X
Rhino	X			X	X
Clock	X			X	X
Trail	X			X	X
Delayed Recall (5 words)	X	X	X		X
Forwards Digits	X	X	X		X
Backward Digits	X	X	X		X
Subtract 7s	X	X	X	X	X
Repeat sentence	X		X		X
Verbal Fluency	X		X		X
Watch/ruler similarity	X	X	X	X	X
Train/bike similarity		X	X	X	X
Progressive Matrices: A8		X		X	X
Progressive Matrices: A11		X		X	X
Progressive Matrices: A12		X		X	X
Line orientation: line1a/b		X		X	X
Line orientation line8a/b		X		X	X
Line orientation line13a/b		X		X	X
Total # of cognition items	12	12	10	14	19

Notes: IPQ = in-person questionnaire; PAPI = paper-and-pencil interview.

for respondents to complete the item because respondents cannot ask for clarification; on the other hand, without an interviewer observing the respondent completing the task, respondents could have received help from others in the home or from the Internet to complete the cognitive items without reporting they had done so.

In addition, items were changed for the web. For example, memory words were flashed on the screen with short intervals and the respondent was asked to type in the words they saw. Although only a perfect repetition was to be scored as correct (no mistakes allowed), we thought that was too much to ask respondents to type in, so in NSHAP R4, they were asked to type in the words they saw in any order. This also was the case for the digits forward and digits backward. They flashed on screen and respondents were asked to type them in. When respondents' responses were reviewed and scored, words that were coded as having typos but were phonetically correct were counted.

For paper-and-pencil mode, instead of typing, we had respondents write their responses down. They were able to do the visual items on paper-and-pencil, but not any items that required memory or digits forward or backward because it would have to have been written on paper instead of saying or flashing it. In the case of the phone, items were mostly the same as they were in-person, except for the visual items.

An issue for web and paper-and-pencil was the timed items. We did not want to do time responses for items on the web or paper-and-pencil because we thought it was too much to ask respondents to be under time pressure for typing or writing words, as opposed to saying them aloud as in the case of in-person administration.

On top of mode effects, there may be a selection bias for each mode. Respondents were not randomly assigned to each mode but were assigned based on an evaluation of each respondent's level of cooperation in previous surveys. More cooperative respondents were assigned to the web-based survey. Additionally, respondents who completed the web survey may have more proficiency with technology and have access to the Internet—factors that could yield a web-based sample with higher socioeconomic status. Because socioeconomic status is associated with MoCA-SA scores (Freitas et al. 2012), higher cognition scores in the web survey could indicate true cognitive ability or it could also indicate higher socioeconomic status. Analysts using the Round 4 cognition data must be vigilant about mode effects (Smith et al., this issue) and how socioeconomic status may affect the cognition results in each mode.

Discussion

NSHAP has always included cognition measures. In Rounds 1–3 of data collection, NSHAP exclusively employed face-to-face interviews to collect data. However, the combination of the availability of cost-efficient remote data collection modes, the exigency of the COVID-19 pandemic, and the need to reach a larger number of younger respondents led the NSHAP team to pivot to multimode data collection including remote components in Round 4. The remote data collection was successfully implemented, with 1,916 respondents completing the NSHAP questionnaire remotely, including web, phone, and paper-and-pencil modes, and 1,469 respondents completing the interview in-person.

Because some of the cognition items in the MoCA-SA, especially the items assessing visuospatial ability, could not be adapted to remote data collection methods, NSHAP incorporated visual cognitive items (progressive matrices and line orientation) from the RADC's cognition measures. These changes resulted in different numbers of cognitive items across the four modes of survey administration in Round 4 (Table 2). Analysts should be aware of these changes when creating a single global cognition score for the entire NSHAP sample in Round 4, and aware that there may be mode effects that could affect cognition scores.

Further guidance on these issues and proposed strategies for harmonization will be forthcoming in a future manuscript. This includes several proposals for analysts to harmonize the R4 cognition scores across the four different modes to create a single global cognition score for the entire Round 4 sample. Additionally, we plan to compare the difficulty of each cognitive item across the modes to determine how mode effects shape performance on each cognitive item to inform researchers using the Round 4 data set how to best manage mode effects. Because we would like to support researchers to use the NSHAP data as soon as it becomes publicly available, in this supplemental issue paper, we describe the changes to the MoCA-SA in Round 4 and urge researchers to take into account these changes when analyzing the cognition data in Round 4 whereas the NSHAP team prepares an additional paper to suggest ways in which the global score can be harmonized across modes.

Funding

Research reported in this publication was supported by the National Institute on Aging of the National Institutes of Health under Awards R01AG043538, and R01AG048511. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or NORC.

Conflict of Interest

None.

References

2023 Alzheimer's disease facts and figures. (2023). *Alzheimer's & Dementia*, 19(4), 1598–1695. https://doi.org/10.1002/alz.13016

Atkins, A. S., Kraus, M. S., Welch, M., Yuan, Z., Stevens, H., Welsh-Bohmer, K. A., & Keefe, R. S. E. (2022). Remote self-administration of digital cognitive tests using the brief assessment of cognition: Feasibility, reliability, and sensitivity to subjective cognitive decline. *Frontiers in Psychiatry*, 13, 910896. https://doi.org/10.3389/fpsyt.2022.910896

Collins, C. L., Pina, A., Carrillo, A., Ghil, E., Smith-Peirce, R. N., Gomez, M., Okolo, P., Chen, Y., Pahor, A., Jaeggi, S. M., & Seitz, A. R. (2022). Video-based remote administration of cognitive assessments and interventions: A comparison with in-lab administration. *Journal of Cognitive Enhancement*, 6(3), 316–326. https://doi.org/10.1007/s41465-022-00240-z

D'Arcy, S., Rapcan, V., Gali, A., Burke, N., O'Connell, G. C., Robertson, I. H., & Reilly, R. B. (2013). A study into the automation of cognitive assessment tasks for delivery via the telephone: Lessons for developing remote monitoring applications for the elderly. *Technology and Health Care*, 21(4), 387–396. https://doi.org/10.3233/THC-130740

Freitas, S., Simões, M. R., Alves, L., & Santana, I. (2012). Montreal cognitive assessment: Influence of sociodemographic and health variables. *Archives of Clinical Neuropsychology*, 27(2), 165–175. https://doi.org/10.1093/arclin/acr116

Hensen, B., Mackworth-Young, C. R. S., Simwinga, M., Abdelmagid, N., Banda, J., Mavodza, C., Doyle, A. M., Bonell, C., & Weiss, H. A. (2021). Remote data collection for public health research in a COVID-19 era: Ethical implications, challenges and opportunities. *Health Policy and Planning*, 36(3), 360–368. https://doi.org/10.1093/heapol/czaa158

Iiboshi, K., Yoshida, K., Yamaoka, Y., Eguchi, Y., Sato, D., Kishimoto, M., Funaki, K., Mimura, M., & Kishimoto, T. (2020). A validation study of the remotely administered Montreal cognitive assessment tool in the elderly Japanese population. *Telemedicine and e-Health*, 26(7), 920–928. https://doi.org/10.1089/tmj.2019.0134

Kotwal, A. A., Schumm, P., Kern, D. W., McClintock, M. K., Waite, L. J., Shega, J. W., Huisingh-Scheetz, M. J., & Dale, W. (2015). Evaluation of a brief survey instrument for assessing subtle differences in cognitive function among older adults. *Alzheimer Disease & Associated Disorders*, 29(4), 317–324. https://doi.org/10.1097/WAD.0000000000000068

O'Muircheartaigh, C., English, N., Pedlow, S., & Schumm, L. P. (2021). Sample design and estimation in the national social life, health, and aging project: Round 3 (2015–2016). *Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*, 76(Suppl 3), S207–S214. https://doi.org/10.1093/geronb/gbab182

Shega, J. W., Sunkara, P. D., Kotwal, A., Kern, D. W., Henning, S. L., McClintock, M. K., Schumm, L. P., Waite, L. J., & Dale, W. (2014). Measuring cognition: The Chicago Cognitive Function Measure in the National, Social Life, Health, and Aging Project, Wave 2. *Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*, 69(Suppl 2), S166–S176. https://doi.org/10.1093/geronb/gbu106

Waite, L.J., Cagney, K., Dale, W., Hawkley, L., Huang, E., Lauderdale, D., Laumann, E.O., McClintock, M., O'Muircheartaigh, C.A., & Schumm, L.P. (2017). *National Social Life, Health and Aging Project (NSHAP): Round 3. ICPSR36873*. Inter-university Consortium for Political and Social Research [distributor]. https://doi.org/10.3886/ICPSR36873