

# Linking Self-Perceived Cognitive Functioning Questionnaires Using Item Response Theory: The Subjective Cognitive Decline Initiative

Laura A. Rabin<sup>1, 2, 3</sup>, Sietske A. M. Sikkes<sup>4, 5</sup>, Douglas Tommet<sup>6</sup>, Richard N. Jones<sup>6, 7</sup>, Paul K. Crane<sup>8</sup>, Milushka M. Elbulok-Charcape<sup>1, 2</sup>, Mark A. Dubbelman<sup>5</sup>, Rebecca Kosciak<sup>9, 10</sup>, Rebecca E. Amariglio<sup>11</sup>, Rachel F. Buckley<sup>11, 12, 13</sup>, Mercè Boada<sup>14</sup>, Gaël Chételat<sup>15</sup>, Bruno Dubois<sup>16, 17</sup>, Kathryn A. Ellis<sup>12, 13</sup>, Katherine A. Gifford<sup>18</sup>, Angela L. Jefferson<sup>18</sup>, Frank Jessen<sup>19, 20</sup>, Sterling Johnson<sup>21, 22</sup>, Mindy J. Katz<sup>3</sup>, Richard B. Lipton<sup>3, 23, 24</sup>, Tobias Luck<sup>25</sup>, Eleni Margioti<sup>26, 27</sup>, Paul Maruff<sup>28</sup>, Jose Luis Molinuevo<sup>29</sup>, Audrey Perrotin<sup>15</sup>, Ronald C. Petersen<sup>30, 31</sup>, Lorena Rami<sup>29</sup>, Barry Reisberg<sup>32, 33</sup>, Dorene M. Rentz<sup>11</sup>, Steffi G. Riedel-Heller<sup>34</sup>, Shannon L. Risacher<sup>35, 36</sup>, Octavio Rodriguez-Gomez<sup>14</sup>, Perminder S. Sachdev<sup>37</sup>, Andrew J. Saykin<sup>35, 36</sup>, Nikolaos Scarmeas<sup>38, 39</sup>, Colette Smart<sup>40, 41</sup>, Beth E. Snitz<sup>42</sup>, Reisa A. Sperling<sup>11</sup>, Vanessa Taler<sup>43, 44</sup>, Wiesje M. van der Flier<sup>4, 5, 45</sup>, Argonde C. van Harten<sup>5, 31, 45</sup>, Michael Wagner<sup>20, 46</sup>, Steffen Wolfgruber<sup>20, 47</sup>, the Alzheimer's Disease Neuroimaging Initiative, the Canadian Longitudinal Study on Aging, and the Health and Aging Brain Study: Health Disparities (HABS-HD) Study Team

<sup>1</sup> Department of Psychology, Brooklyn College, City University of New York

<sup>2</sup> Department of Psychology, The Graduate Center, City University of New York

<sup>3</sup> Saul R. Korey Department of Neurology, Albert Einstein College of Medicine

<sup>4</sup> Department of Epidemiology and Data Science, Vrije Universiteit Amsterdam

<sup>5</sup> Department of Neurology, Alzheimer Center Amsterdam, Vrije Universiteit Amsterdam

<sup>6</sup> Department of Psychiatry and Human Behavior, Warren Alpert Medical School, Brown University

<sup>7</sup> Department of Neurology, Warren Alpert Medical School, Brown University

<sup>8</sup> Division of General Internal Medicine, University of Washington

<sup>9</sup> Wisconsin Alzheimer's Institute, University of Wisconsin School of Medicine and Public Health

<sup>10</sup> Department of Medicine, University of Wisconsin School of Medicine and Public Health

<sup>11</sup> Department of Neurology, Brigham and Women's Hospital, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts, United States

<sup>12</sup> Department of Psychiatry, University of Melbourne

<sup>13</sup> Melbourne School of Psychological Sciences, University of Melbourne

<sup>14</sup> Alzheimer Research Center and Memory Clinic, Fundació Alzheimer Centre Educacional, Institut Català de Neurociències Aplicades, Universitat Internacional de Catalunya

<sup>15</sup> Normandie Université, UNICAEN, INSERM, U1237, Physiopathology and Imaging of Neurological Disorders, NeuroPresage Team

<sup>16</sup> Université Pierre et Marie Curie-Paris 6, Assistance Publique-Hôpitaux de Paris, Hôpital de la Salpêtrière, Paris, France

<sup>17</sup> Centre des Maladies Cognitives et Comportementales, Institut du Cerveau et de la Moelle épinière, UMR5975, Paris, France

<sup>18</sup> Department of Neurology, Vanderbilt Memory & Alzheimer's Center, Vanderbilt University Medical Center

<sup>19</sup> Department of Psychiatry and Psychotherapy, Medical Faculty, University of Cologne

<sup>20</sup> German Center for Neurodegenerative Diseases (Deutsches Zentrum für Neurodegenerative Erkrankungen), Bonn, Germany

<sup>21</sup> Geriatric Research Education and Clinical Center, William S. Middleton Memorial Veterans Hospital, Madison, Wisconsin, United States

<sup>22</sup> Department of Medicine, University of Wisconsin School of Medicine and Public Health

<sup>23</sup> Department of Psychiatry and Behavioral Medicine, Albert Einstein College of Medicine

<sup>24</sup> Department of Epidemiology & Population Health, Albert Einstein College of Medicine

<sup>25</sup> Faculty of Applied Social Sciences, University of Applied Sciences Erfurt

<sup>26</sup> Laboratory of Cognitive Neuroscience, School of Psychology, Aristotle University of Thessaloniki

<sup>27</sup> Aviv Clinics, Jumeirah Lake Towers, Dubai, United Arab Emirates

<sup>28</sup> Cogstate, Ltd., Melbourne, Victoria, Australia

<sup>29</sup> Alzheimer's Disease and Other Cognitive Disorders Unit, Institut D'Investigacions Biomèdiques August Pi i Sunyer, Hospital Clinic, Barcelona, Spain

<sup>30</sup> Department of Health Sciences Research, Mayo Clinic, Rochester, Minnesota, United States

<sup>31</sup> Department of Neurology, Mayo Clinic, Rochester, Minnesota, United States

<sup>32</sup> Department of Psychiatry, New York University Langone Medical Center

<sup>33</sup> Silberstein Aging and Dementia Research Center, New York University School of Medicine

<sup>34</sup> Institute of Social Medicine, Occupational Health and Public Health (Institut für Sozialmedizin, Arbeitsmedizin und Public Health), University of Leipzig

<sup>35</sup> Department of Radiology and Imaging Sciences, Indiana University School of Medicine

<sup>36</sup> Indiana Alzheimer's Disease Research Center, Indiana University School of Medicine

<sup>37</sup> Centre for Healthy Brain Ageing, School of Psychiatry, University of New South Wales

<sup>38</sup> 1st Department of Neurology, Aiginition Hospital, National and Kapodistrian University of Athens Medical School

<sup>39</sup> Department of Neurology, Columbia University

<sup>40</sup> Department of Psychology, University of Victoria

<sup>41</sup> Centre on Aging, University of Victoria

<sup>42</sup> Department of Neurology, University of Pittsburgh

<sup>43</sup> School of Psychology, University of Ottawa

<sup>44</sup> Bruyère Research Institute, Ottawa, Ontario, Canada

<sup>45</sup> Department of Neurology, Amsterdam Neuroscience–Neurodegeneration, Amsterdam Universitair Medische Centra

<sup>46</sup> Department of Neurodegenerative Diseases and Geriatric Psychiatry, University of Bonn

<sup>47</sup> Department of Psychiatry and Psychotherapy, University of Bonn

## on behalf of the ISTAART Subjective Cognitive Decline Professional Interest Area

**Objective:** Self-perceived cognitive functioning, considered highly relevant in the context of aging and dementia, is assessed in numerous ways—hindering the comparison of findings across studies and settings. Therefore, the present study aimed to link item-level self-report questionnaire data from international aging studies. **Method:** We harmonized secondary data from 24 studies and 40 different questionnaires with item response theory (IRT) techniques using a graded response model with a Bayesian estimator. We compared item information curves to identify items with high measurement precision at different levels of the self-perceived cognitive functioning latent trait. Data from 53,030 neuropsychologically intact older adults were included, from 13 English language and 11 non-English (or mixed) language studies. **Results:** We successfully linked all questionnaires and demonstrated that a single-factor structure was reasonable for the latent trait. Items that made the greatest contribution to measurement precision (i.e., “top items”) assessed general and specific memory problems and aspects of executive functioning, attention, language, calculation, and visuospatial skills. These top items originated from distinct questionnaires and varied in format, range, time frames, response options, and whether they captured ability and/or change. **Conclusions:** This was the first study to calibrate self-perceived cognitive functioning data of geographically diverse older adults. The resulting item scores are on the same metric, facilitating joint or pooled analyses across international studies. Results may lead to the development of new self-perceived cognitive functioning questionnaires guided by psychometric properties, content, and other important features of items in our item bank.

### Key Points

**Question:** Can item response theory be used to link questionnaires of self-perceived cognitive functioning with varied item properties? **Findings:** We harmonized and identified the psychometric properties of 601 self-report items from 40 different questionnaires of 24 aging studies in a secondary data analysis that used an IRT item banking approach. **Importance:** Resulting item scores were on the same metric, facilitating joint or pooled analyses across international aging studies; item data also may be useful to investigators proposing to measure self-perceived cognitive functioning in clinical and research settings. **Next Steps:** Future research will utilize the current results to develop new questionnaires and determine associations of self-perceived cognitive functioning with relevant outcomes in the context of aging and dementia research.

**Keywords:** subjective cognitive decline, measurement, item response theory, harmonization, self-perceived cognitive functioning

Richard N. Jones  <https://orcid.org/0000-0002-1049-218X>

Paul K. Crane  <https://orcid.org/0000-0003-4278-7465>

Milushka M. Elbulok-Charcape  <https://orcid.org/0000-0003-0674-8557>

Mark A. Dubbelman  <https://orcid.org/0000-0002-5708-4925>

Richard B. Lipton  <https://orcid.org/0000-0003-2652-2897>

Tobias Luck  <https://orcid.org/0000-0001-7551-4397>

Paul Maruff  <https://orcid.org/0000-0002-6947-9537>

Dorene M. Rentz  <https://orcid.org/0000-0002-6318-9213>

Perminder S. Sachdev  <https://orcid.org/0000-0002-9595-3220>

Andrew J. Saykin  <https://orcid.org/0000-0002-1376-8532>

Colette Smart  <https://orcid.org/0000-0002-2439-4425>

Wiesje M. van der Flier  <https://orcid.org/0000-0001-8766-6224>

Michael Wagner  <https://orcid.org/0000-0003-2589-6440>

The authors would like to thank the Alzheimer's Association and the International Society to Advance Alzheimer's Research and Treatment (ISTAART) Subjective Cognitive Decline Professional Interest Area (SCD-PIA). This research was made possible using the data collected by the Canadian Longitudinal Study on Aging. Funding for the Canadian Longitudinal Study on Aging is provided by the Government of Canada through the Canadian Institutes of Health Research (CIHR) under Grant LSA 94473 and the Canada Foundation for Innovation, as well as the following provinces, Newfoundland, Nova Scotia, Quebec, Ontario, Manitoba, Alberta, and British Columbia. This research has been conducted using the Canadian Longitudinal Study on Aging data set (Follow-Up 1 Comprehensive Data Set Version 3.2),

*continued*

Older adults commonly perceive changes in their own cognitive abilities, hereafter referred to as decline in *self-perceived cognitive functioning* (Jonker et al., 1996, 2000; Ponds et al., 1997; Röhr et al., 2020; van Harten et al., 2018). Although sometimes accompanied by impairment detected from assessment with objective neuropsychological tests, as in prodromal dementia conditions such as mild cognitive impairment (MCI), decline in self-perceived cognitive functioning can also be related to various medical conditions, mood or sleep disturbances, medications, personality traits, or physiological changes that occur with normal aging (Buckley et al., 2013; Comijs et al., 2002; Hill et al., 2016; Jessen et al., 2020; Rabin et al., 2017; Zlatar et al., 2014). For some older adults, a self-experienced persistent decline in cognitive functioning, in the absence of objective neuropsychological deficits (and unrelated to an acute event),

may be an early manifestation of underlying neurodegenerative changes associated with an increased risk of future objective cognitive decline (Jessen et al., 2014, 2020; Reisberg et al., 2008). This condition, referred to as subjective cognitive decline (SCD), may represent one of the first symptomatic manifestations of Alzheimer's disease and related disorders (Jessen et al., 2014, 2020; Molinuevo et al., 2017; Rabin et al., 2017; Reisberg & Gauthier, 2008; Slot et al., 2019).

The assessment of self-perceived cognitive functioning offers the advantage of directly gathering information about important and potentially clinically significant aspects of cognition and function from a first-person perspective (Lai et al., 2014; Lucas & Baird, 2006). Unfortunately, the field lacks a standardized approach for measuring self-perceived cognitive functioning in older adults,

under Application 2203008. The opinions expressed in this article are the authors' own and do not reflect the views of the Canadian Longitudinal Study on Aging.

Sterling C. Johnson serves on an advisory boards to Roche Diagnostics, Prothena, and Eisai in the past 2 years and has received research support from Cerveau Technologies. Jose Luis Molinuevo is a full-time employee at H. Lundbeck A/S. Ronald C. Petersen is a consultant for Roche, Inc., Merck, Inc., Biogen, Inc., Genentech, Inc., Eisai, Inc., Eli Lilly, and Nestle, Inc. Perminder S. Sachdev was a member of Advisory Committees for Biogen Australia and Roche Australia in 2020 and 2021. This study was supported by the National Institute on Aging (Grants P01 AG036694 and K24 AG035007 to Reisa A. Sperling; NIH Grants to K23 AG038479, P01 AG025204, R37 AG025516, and AG023651 to Beth E. Snitz; Grants K01 AG049050, R01 AG061788, R01 AG19771, and P30 AG10133; K01 AG049050, R01 AG061788, R01 AG19771, and P30 AG10133 to Shannon L. Risacher; NIA U01 AG006786, P50 AG016574, GHR Foundation, and Mayo Foundation to Ronald C. Petersen; NIA 2 P01 AG03949, the Leonard and Sylvia Marx Foundation, and the Czap Foundation to Richard B. Lipton; NIA 2 P01 AG03949, the Leonard and Sylvia Marx Foundation, and the Czap Foundation to Mindy J. Katz; Grant AG027161 to Sterling Johnson; Grants R01-AG034962, R01-AG056534, R01-NS100980, and K24-AG046373 to Angela L. Jefferson; Grants R01-AG062826, K23-AG045966 to Katherine A. Gifford; Grant AG061238 to Rachel F. Buckley; Grant R01AG058825 to Rebecca E. Amariglio; Grant AG027161 to Rebecca Kosciak; Grant AG051170 to Paul K. Crane; Grant AG051170 to Douglas Tommet). Laura A. Rabin received funding from the NIH/NIA and the National Science Foundation under Grants AG 066039; AG051170; NSF 2050755. Nikolaos Scarmeas received funding from Alzheimer's Association (IIRG-09133014), ESPA-EU Program Excellence Grant, ARISTEIA, co-funded by the European Social Fund and Greek National Resources (189 10,276/8/9/2011), Ministry for Health and Social Solidarity, Greece (DY2b/oik.51657/14.4.2009). Sietske A. M. Sikkes received funding from the National Institute of Health under public-private funding from Health-Holland, Topsector Life Sciences & Health (PPP-allowance; LSHM20084-SGF, project DEFEAT-AD), and the National Institutes of Health, as well as license fees from Green Valley, VtV Therapeutics, Alzheon, Vivoryon, and Roche, and honoraria from Boehringer and Toyama. All funding is paid to her institution. Perminder S. Sachdev received funding from NHMRC Australia Grants ID350833, ID568969, and APP1093083. Steffi G. Riedel-Heller received funding from the Interdisciplinary Centre for Clinical Research of the University of Leipzig (IZKF, Project C07). Barry Reisberg was supported by the U.S. DHHS Grants P30 AG08051, AG03051, AG12101, MH35976, and MH40410; and the Zachary and Elizabeth M. Fisher Education and Research Program at the New York University School of Medicine, and the Louis J. Kay and June E. Kay Foundation. Andrew J. Saykin received support from multiple National Institute of Health (NIH) Grants (P30 AG010133, P30 AG072976, R01

AG019771, R01 AG057739, R01 CA129769, U19 AG024904, R01 LM013463, R01 AG068193, T32 AG071444, U01 AG068057, and U01 AG072177). He has also received support from Avid Radiopharmaceuticals, a subsidiary of Eli Lilly (in kind contribution of Positron Emission Tomography tracer precursor); Bayer Oncology (Scientific Advisory Board); Eisai (Scientific Advisory Board); Siemens Medical Solutions USA, Inc. (Dementia Advisory Board); NIH National Heart, Lung, and Blood Institute (Multi-Ethnic Study of Atherosclerosis; MESA Observational Study Monitoring Board); and Springer-Nature Publishing (Editorial Office Support as Editor-in-Chief, Brain Imaging and Behavior). Nikolaos Scarmeas is chair of the Data Safety Monitoring Board for an NIH-funded study at The Albert Einstein College of Medicine. Nikolaos Scarmeas is also a local site Principal Investigator of a Phase III industry-sponsored study (Semaglutide-Novonordisc), with funding to the institution.

HABS-HD MPIs include Sid E. O'Bryant, Kristine Yaffe, Arthur Toga, Robert Rissman, and Leigh Johnson; and the HABS-HD investigators: Meredith Braskie, Kevin King, James R. Hall, Melissa Petersen, Raymond Palmer, Robert Barber, Yonggang Shi, Fan Zhang, Rajesh Nandy, Roderick McColl, David Mason, Bradley Christian, Nicole Philips, Stephanie Large, and Rocky Vig.

Laura A. Rabin played an equal role in conceptualization, funding acquisition, methodology, project administration, visualization, writing—original draft, and writing—review and editing. Sietske A. M. Sikkes played an equal role in conceptualization, data curation, formal analysis, visualization, writing—original draft, and writing—review and editing. Douglas Tommet played an equal role in conceptualization, data curation, formal analysis, methodology, writing—original draft, and writing—review and editing. Richard N. Jones played an equal role in conceptualization, data curation, formal analysis, methodology, supervision and writing—review and editing. Paul K. Crane played an equal role in writing—original draft, and writing—review and editing. Milushka M. Elbulok-Charcape played an equal role in visualization and writing—review and editing. Mark A. Dubbelman played an equal role in resources and writing—review and editing. Rebecca L. Kosciak played a supporting role in resources and writing—review and editing. Rebecca E. Amariglio played a supporting role in resources and writing—review and editing. Rachel F. Buckley played a supporting role in resources and writing—review and editing. Mercè Boada played a supporting role in resources and writing—review and editing. Gaël Chételat played a supporting role in resources and writing—review and editing. Bruno Dubois played a supporting role in resources and writing—review and editing. Kathryn A. Ellis played a supporting role in resources and writing—review and editing. Katherine A. Gifford played a supporting role in resources and writing—review and editing. Angela L. Jefferson played a supporting role in conceptualization, resources and writing—review and editing. Frank Jessen played a supporting role in resources and writing—review and editing. Sterling C. Johnson played a supporting role in resources and writing—review and editing. Mindy J. Katz played a supporting

*continued*

which greatly limits understanding of the construct of SCD and the ability to compare findings across studies and settings (Molinuevo et al., 2017; Röhr et al., 2020). Coincident with the conceptualization of the SCD condition as a manifestation of the late preclinical stage of Alzheimer's disease (AD; Jack et al., 2018), there has been a proliferation of new questionnaires or adoption of existing questionnaires within international aging studies (Abdulrab & Heun, 2008; Rabin et al., 2015; Reid & MacLulich, 2006). Some measures have been adapted for application in multiple languages (Raimo et al., 2016; Şahin et al., 2013; Slot et al., 2018) and/or used internationally with culturally and linguistically diverse populations (Brucki & Nitrini, 2009; Hao et al., 2017; Vlachos et al., 2019). A previous study documented the vast heterogeneity in questionnaires used to assess self-perceived cognitive functioning in 19 aging studies from Australia and countries in North America and Europe (Rabin et al., 2015). These aging studies were also diverse with respect to recruitment approaches, research settings, language in which assessments were conducted, and sample sizes. The study found that while items inquired about a variety of cognitive domains, those pertaining to memory were the most common (Rabin et al., 2015). Overall, results revealed little overlap among self-perceived cognitive functioning questionnaires and great variability in terms of key features of item format and content. Additionally, questionnaire and item selection decisions were mostly based on practical rather than psychometric considerations, with many measures lacking evidence of validity (Rabin et al., 2015).

The variability in questionnaires and item properties may account, in part, for discrepant findings with regard to whether decline in self-perceived cognitive functioning is associated with AD biomarkers and clinical progression (Amariglio et al., 2012, 2021; Amariglio, Mormino, et al., 2015; Buckley et al., 2016; Jessen et al., 2010; Kryscio et al., 2014; Mielke et al., 2012; Mitchell et al., 2014; Mosconi et al., 2008; Perrotin et al., 2012; Peter et al., 2014; Reisberg et al., 2010, 2019; Risacher et al., 2015; Saykin et al.,

2006; Scheef et al., 2012; Schultz et al., 2015; Slot et al., 2019; Snitz et al., 2018; van Harten et al., 2013; Visser et al., 2009) or is not associated with AD biomarkers and clinical progression (Chételat et al., 2010; Hollands et al., 2015; Zwan et al., 2016; see Colijn & Grossberg, 2015; Lin et al., 2019; Lista et al., 2015; Sun et al., 2015, for reviews). Initiatives to understand the condition of SCD have led to efforts to develop a unified conceptual framework (Jessen et al., 2014, 2020). Additionally, there have been calls to pool data across studies to examine the psychometric properties and quality of items used to classify SCD (Jessen et al., 2020; Molinuevo et al., 2017; Rabin et al., 2017; Röhr et al., 2020). Investigation into the reliability and validity of self-perceived cognitive functioning measures and improvements in measurement development are critical to promoting understanding and replicability of study conclusions (Flake & Fried, 2020; Fried & Flake, 2018).

The application of psychometric modeling techniques, such as item response theory (IRT), to self-perceived cognitive functioning measurement could enable harmonization of data across studies with diverse assessment approaches. Additionally, such an approach could facilitate the evaluation of the precision of items and subsequent modification of the instruments themselves. IRT comprises a collection of modeling techniques for the analysis of item-level data, which can be used to evaluate the psychometric properties of existing scales and items, shorten scales, and calibrate items from different scales (Edelen & Reeve, 2007). IRT modeling can help generate precise, valid, and relatively brief instruments that present minimal burden to respondents. One example of the use of IRT in health care research and clinical practice is the Patient-Reported Outcomes Measurement Information System, which consists of item banks and short forms of self-reported health measures that assess many aspects of physical, mental/cognitive, and social health in diverse adult and pediatric populations (Cella et al., 2010; Irwin et al., 2012; Northwestern University, 2022). The application of IRT to self-perceived cognitive functioning measurement has been

role in resources and writing-review and editing. Richard B. Lipton played a supporting role in resources and writing-review and editing. Tobias Luck played a supporting role in resources and writing-review and editing. Eleni Margioti played a supporting role in resources and writing-review and editing. Paul Maruff played a supporting role in resources and writing-review and editing. Jose Luis Molinuevo played a supporting role in resources and writing-review and editing. Audrey Perrotin played a supporting role in resources and writing-review and editing. Ronald C. Petersen played a supporting role in resources and writing-review and editing. Lorena Rami played a supporting role in resources and writing-review and editing. Barry Reisberg played a supporting role in resources and writing-review and editing. Dorene M. Rentz played a supporting role in resources and writing-review and editing. Steffi G. Riedel-Heller played a supporting role in resources and writing-review and editing. Shannon L. Risacher played a supporting role in resources and writing-review and editing. Octavio Rodriguez-Gomez played a supporting role in resources and writing-review and editing. Perminder S. Sachdev played a supporting role in conceptualization, resources and writing-review and editing. Andrew J. Saykin played a supporting role in resources and writing-review and editing. Nikolaos Scarmeas played a supporting role in resources and writing-review and editing. Colette Smart played a supporting role in resources and writing-review and editing. Beth E. Snitz played a supporting role in resources and writing-review and editing. Reisa A. Sperling played a supporting role in resources and writing-review and editing. Vanessa Taler played a supporting

role in resources and writing-review and editing. Wiesje M. van der Flier played a supporting role in resources and writing-review and editing. Argonde C. van Harten played a supporting role in resources and writing-review and editing. Michael Wagner played a supporting role in resources and writing-review and editing. Steffen Wolfgruber played a supporting role in resources and writing-review and editing.

In addition to the authors above, data were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (<http://adni.loni.usc.edu>). As such, numerous investigators within ADNI (in addition to those listed above) contributed to the design and implementation of ADNI and/or provided data but did not participate in the analysis or writing of this specific article. A complete listing of ADNI investigators can be found at: [http://adni.loni.usc.edu/wp-content/uploads/how\\_to\\_apply/ADNI\\_Acknowledgement\\_List.pdf](http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf)

In addition to the authors above, data were obtained from the Anti-Amyloid Treatment in Asymptomatic Alzheimer's study (A4; <https://a4study.org>). The A4 study included older individuals (ages 65–85) who have normal thinking and memory function but who may be at risk for memory loss due to Alzheimer's disease. The A4 study is coordinated by the University of Southern California Alzheimer's Therapeutic Research Institute, with study sites in multiple locations.

Correspondence concerning this article should be addressed to Laura A. Rabin, Department of Psychology, Brooklyn College, City University of New York, 2900 Bedford Avenue, Brooklyn, NY 11210, United States. Email: [lrabin@brooklyn.cuny.edu](mailto:lrabin@brooklyn.cuny.edu)



performed in the context of oncology, thereby supporting the suitability of measuring cognitive concerns and abilities with a unidimensional item bank (Lai et al., 2014). In the field of cognitive aging, IRT has been utilized in two published studies researching item-level self-perceived cognitive functioning data. Snitz et al. (2012) examined the utility of IRT for items probing self-perceived cognitive functioning in 3,495 older adult participants from four community-based studies in Southwestern Pennsylvania, United States. IRT scoring of questionnaire items was associated with objective cognitive test scores and provided additional information beyond a simple sum of items, making this approach ideal for capturing very subtle declines in self-perceived cognitive functioning. Gifford et al. (2015) sought to identify items that most reliably captured self-perceived cognitive functioning among 188 nondemented older adults recruited from a research registry in Boston, United States. After post hoc simulation using computerized adaptive testing, researchers identified nine of 21 items that represented the latent trait and differentiated normal controls from those with MCI. These items assessed self-perceived global and specific problems and changes related to memory functioning.

These previous studies supported the benefits of applying a modern psychometric approach to self-perceived cognitive functioning data. The present study expands upon this foundation by employing item banking and cocalibration to link questionnaires from international cognitive aging studies to provide a more generalizable model of self-perceived cognitive functioning. Using IRT methods, we sought to minimize measurement scale differences as

a potential impediment to understanding the coherence of results across studies. We further aimed to identify items with both strong and weak psychometric characteristics based on item information functions. Harmonization results may be used to compare data from studies that use items with varied content and measurement properties and to identify optimal items to include on new self-perceived cognitive functioning questionnaires.

## Method

Figure 1 provides an overview of the 10 primary data coding and analytic processing steps, which are discussed below in the same order.

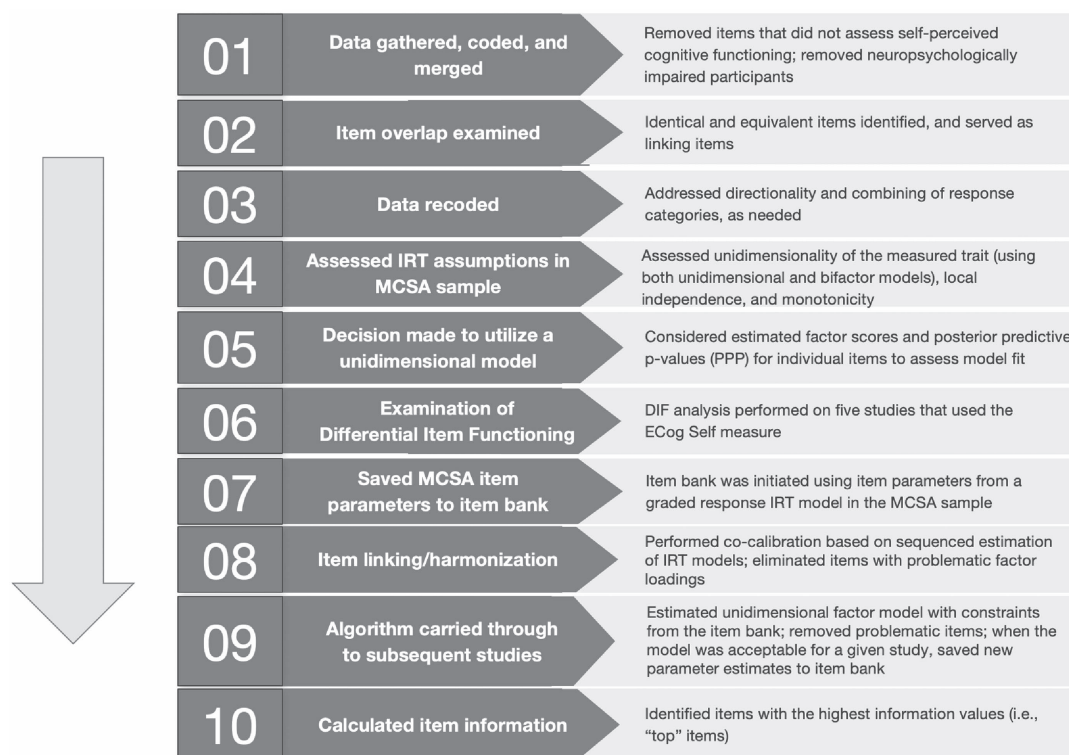
### (1) Data Gathered, Coded, and Merged

#### Data Sources and Questionnaires

Data were obtained from members of the Subjective Cognitive Decline Initiative (SCD-I), part of the SCD Professional Interest Area of the Alzheimer's Association International Society to Advance Alzheimer's Research and Treatment (ISTAART SCD-PIA). The SCD-I is a working group of AD researchers with a specific interest in self-perceived cognitive functioning.

Participating researchers sent electronic copies of their published and unpublished cognitive and functional self-report questionnaires. For studies conducted in languages other than English, researchers

**Figure 1**  
*Sequence of Data Coding and Analytic Procedures*



*Note.* IRT = item response theory; MCSA = Mayo Clinic Study of Aging; DIF = differential item functioning; ECog Self = Everyday Cognition-Subject/Self-Report.

sent translated versions of their questionnaires. For the Alzheimer's Disease Neuroimaging Initiative (ADNI), Anti-Amyloid Treatment in Asymptomatic Alzheimer's (A4), and Health and Aging Brain Study: Health Disparities (HABS-HD) studies, after obtaining permission, we retrieved the questionnaires directly from the study databases (<http://adni.loni.usc.edu>, <https://a4study.org>, and <https://apps.unthsc.edu/itr/researchers>, respectively).

### Categorization and Item Inclusion/Exclusion

The study authors reviewed all questionnaires to determine which merited inclusion by virtue of containing items related to self-perceived cognitive functioning. To make this determination, each item was assigned to one of eight primary cognitive domains: (a) *memory*, (b) *attention/working memory/processing speed*, (c) *language*, (d) *executive function*, (e) *basic calculation and arithmetic tasks*, (f) *orientation*, (g) *general cognitive ability*, and (h) *visuospatial skills*. Items that could not be assigned to a cognitive domain were excluded; such items assessed mood, health complaints, self-efficacy, perceived stress, personality, life satisfaction, basic activities of daily living, use of cognitive strategies, and other variables that are relevant to or associated with cognition but do not directly assess self-perceived cognitive functioning. Questionnaires that contained at least one self-perceived cognitive function item were retained.

For each self-perceived cognitive functioning item beyond cognitive domain, we coded other relevant features such as the specific type(s) of response options, scaling methods, response time frames, and whether item stems were multibarreled (i.e., containing two or more subquestions within a single item). We also determined whether item stems and response options together inquired about the ability to perform a target cognitive task versus change (improvement or decline).

For these various coding procedures, author LAR carried out the initial coding. Two additional authors (CMS and SAMS) subsequently identified items that they considered to be miscoded. Items in dispute were then reviewed and discussed until a consensus was reached, occasionally after discussing disputed items with authors DT, RNJ, and PKC.<sup>1</sup>

### Data Merged

Researchers from the different studies were asked to send item-level data—that is, de-identified files that included responses to self-report items. For the ADNI, A4, and HABS-HD studies, we retrieved these data directly from the study databases. We included only the baseline (or first) assessment for each participant to ensure that observations were independent and to avoid the potential that previous exposure to a questionnaire affected subsequent responses. As we were interested in self-perceived cognitive functioning in neuropsychologically intact individuals, data from participants with prevalent or concurrently classified diagnoses of MCI, AD, or other dementia were excluded. Appendix A displays the definition of *neuropsychologically intact* or *cognitively normal* for each study.

### (2) Item Overlap Examined: Determination of Identical, Equivalent, and Linking Items

After careful review of the database, a subset of authors (DT, LAR, SAMS, and RJ) noticed that some studies used multiple questionnaires (see Table 1), and there was overlap in questionnaires

used among studies (see Table 2). These authors next sought to identify items from different questionnaires that were exactly the same, or *identical*. To be deemed identical, items were required to contain precisely the same item stems, response options, temporal referents (recall intervals), and so forth. Close examination of all items revealed that no *identical* items appeared on different questionnaires (i.e., the same item did not appear on more than one questionnaire). These authors then considered whether items with overlapping content were similar enough to be treated as *equivalent* in the analyses despite minor differences in wording. Such items contained overlapping content—that is, addressed the same cognitive process, such as memory for recent events, but had slight differences in wording, and thus could not be deemed *identical*. *Equivalent items* were identified by consensus among authors DT, LAR, RNJ, and SAMS. In some cases, response options for equivalent items were combined so that they would match across items. Appendix B presents sample equivalent items and recoding of response options. Items deemed equivalent were assigned a single code and subsequently treated as identical in study analyses. Both *identical* and *equivalent* items served as *linking items* in the analyses (i.e., were entered into the analysis as the same item). Our approach to linking test scores is known as the nonequivalent anchor test approach (Dorans et al., 2007) and is commonly used when different tests are given to nonequivalent groups (but there are common items administered between the groups). As noted above, some of our links involved complete tests, others involved a small number of items, and sometimes the linking items were generated ad hoc using the process for determining item equivalence.

### (3) Data Recoded

Items were recoded, as needed, so that for every item a higher response option corresponded to better self-perceived cognitive functioning. As an example, the original response options for item 56 from the Metamemory in Adulthood Questionnaire (Dixon et al., 1988), “My memory for dates has greatly declined in the last 10 years” ranged from 1 = *strongly disagree*, 2 = *disagree*, 3 = *undecided*, 4 = *agree*, and 5 = *strongly agree*. These were recoded—essentially reversed—so that the higher option categories matched better self-perceived cognitive functioning 1 = *strongly agree*, 2 = *agree*, 3 = *undecided*,

<sup>1</sup> Coding of most items was carried out as part of a previous SCD-I study that served as a precursor to the current harmonization effort (Rabin et al., 2015). Two questionnaires included in the previous study were excluded from the current analysis: (a) Self-Evaluation of Cognition Questionnaire (Pre-AI SCQ, with 14 usable items) because the parent study did not enroll neuropsychologically intact participants and (b) Subjective Cognitive Decline Self-Identification Item (Victoria SCDS, with one usable item; Smart et al., 2014) because this item was used to classify participants and was not a self-report item per se. Also, for the current analysis, we added several new questionnaires: Amsterdam Dementia Cohort, Subjective Cognitive Concerns screener (van der Flier & Scheltens, 2018), Blessed Functional Activities Scale (Blessed et al., 1968), Canadian Longitudinal Study on Aging (CLSA) SCD/worry item (Raina et al., 2009, 2019), Cognitive Function Instrument; Amariglio, Donohue, et al., 2015), Hellenic Longitudinal Investigation of Aging and Diet Study (HELIAD) Subjective Cognitive Items (Margioti et al., 2020), Mayo Clinic Study of Aging (MCSA), adapted Blessed Memory scale (Mielke et al., 2012), Subjective Memory Complaints Questionnaire (SMCQ; Youn et al., 2009), and the WRAP Memory Item (Nicholas et al., 2017). Together these new questionnaires contributed 51 usable items. Authors LAR, MEC, and SAMS independently carried out cognitive domain coding for the new items, with complete agreement.

**Table 1**  
*Key Features of Participating Studies*

Study name (abbreviation)	Institutional affiliation(s)	Country (language)	Research environment	Number of participants: neuropsychologically intact/total	Number of questionnaires	Final usable number of items
1. Alzheimer's Disease Center Clinical Core and Center for Brain Health (ADC-NYU)	New York University School of Medicine	United States (English)	Volunteer	245/788	1	4
2. Alzheimer's Disease Neuroimaging Initiative (ADNI)	Over 100 international affiliates: <a href="http://adni.loni.usc.edu">http://adni.loni.usc.edu</a>	United States Canada Australia Germany United Kingdom (English, German) Netherlands (Dutch)	Volunteer and memory clinic	938/1,729	2	60
3. Amsterdam Dementia Cohort (Ams Dem Cohort)	Vrije Universiteit Amsterdam University Medical Center	United Kingdom (English, German) Netherlands (Dutch)	Memory clinic	569/569	3	27
4. Anti-Amyloid Treatment in Asymptomatic Alzheimer's Study (A4)	Multiple sites across the United States, Canada, and Japan: <a href="https://a4study.org/locations/">https://a4study.org/locations/</a>	United States Canada Japan (English, Spanish, Japanese) Australia (English)	Community-based	5,667/6,945	3	18
5. Australian Imaging Biomarkers and Lifestyle Flagship Study of Ageing (AIBL)	University of Melbourne	Spain (Spanish) Australia (English)	Volunteer and physician referred	595/851	2	6
6. Barcelona Group (Barcelona)	Institut D'Investigacions Biomèdiques August Pi i Sunyer, Hospital Clinic, Barcelona	Spain (Spanish)	Volunteer and memory clinic	186/186	1	26
7. Bonn Memory Clinic (Bonn)	University of Bonn and the German Center for Neurodegenerative Diseases, Bonn	Germany (German)	Memory clinic	121/121	1	32
8. Canadian Longitudinal Study on Aging (CLSA)	Multiple sites across Canada: <a href="https://www.clsa-elcv.ca">https://www.clsa-elcv.ca</a>	Canada (English, French)	Population-based	27,229/27,760	2	21
9. Dartmouth Memory and Aging Study/Indiana Memory and Aging Study (Dart-Indiana)	Dartmouth Medical School, Indiana University School of Medicine	United States (English)	Volunteer and physician referral	141/271	8	173
10. Einstein Aging Study (EAS)	Albert Einstein College of Medicine	United States (English)	Community-based	1,399/1,953 224/258 292/447	3 7 1	17 121 41
11. German Study on Ageing, Cognition, and Dementia in Primary Care Patients (AgeCoDe)	University of Bonn and five other German universities	Germany (German)	General practice registry-based	2,422/3,327	2	5
12. Health and Aging Brain Study, Health Disparities (HABS-HD)	University of North Texas Health Science Center	United States (English)	Community-based	1,842/2,370	2	18
13. Harvard Aging Brain Study (HABS)	Harvard Medical School, Brigham and Women's Hospital, and Massachusetts General Hospital	United States (English)	Volunteer	237/237	3	93

(table continues)

Table 1 (continued)

Study name (abbreviation)	Institutional affiliation(s)	Country (language)	Research environment	Number of participants: neuropsychologically intact/total	Number of questionnaires	Final usable number of items
14. Hellenic Longitudinal Investigation of Aging and Diet (HELLAD)	Aiginition Hospital, National and Kapodistrian University of Athens Medical School	Greece (Greek)	Population-based	1,550/1,943	3	15
15. Institut des Matériaux Poreux de Paris Caen Group	Inserm, University and Hospital of Caen	France (French)	Memory clinic	123/209	1	32
16. Indiana Alzheimer's Disease Center Cohort (IADC)	Indiana Alzheimer Disease Center; Indiana University School of Medicine	United States (English)	Memory clinic	226/350	2	60
17. Leipzig Longitudinal Study of the Aged (LEILA 75+)	University of Leipzig	Germany (German)	Population-based	732/1,692	1	1
18. Mayo Clinic Study of Aging (MCSA)	MCSA	United States (English)	Community-based	4,200/4,200	2	45
19. Memory Clinic-Fundació Alzheimer Centre Educacional (Fundació)	Fundació Alzheimer Centre Educacional Barcelona Alzheimer Treatment and Research Center	Spain (Spanish)	Community-based	132/201	1	28
20. Sydney Memory and Ageing Study (Sydney)	University of New South Wales	Australia (English)	Community-based	504/873	3	32
21. University of Pittsburgh Study/Monongahela-Youghiogheny Health Aging Team (MYHAT) Pittsburgh (Pitt) Substudy <sup>b</sup>	University of Pittsburgh	United States (English)	Clinic-community volunteer sample and population-based	1,916/1,916 101/174	1 3	17 81
22. Vanderbilt Memory and Alzheimer's Center (VMAC)	Vanderbilt University School of Medicine	United States (English)	Memory clinic, physician referral, and community-based	115/191	1	51
23. Victoria Subjective Cognitive Decline Study (Victoria SCDS)	University of Victoria	Canada (English)	Volunteer	42/42	4	53
24. Wisconsin Registry for Alzheimer's Prevention (WRAP)	University of Wisconsin-Madison	United States (English)	Community-based	1,282/1,538	2	35
Total (including substudies)				53,030/61,141	65	1,112

Note. ADL Abbrev = Activities of Daily Living Rating Scale Self-Version, Abbreviated; CAPM-C = Comprehensive Assessment of Prospective Memory-Section C; Einstein HSA = Albert Einstein Health Self-Assessment; MIA = Memory in Adulthood Questionnaire; MFQ = Memory Functioning Questionnaire; CFQ = Cognitive Function Index; SCCS = Subjective Cognitive Complaint Scale-also called the Subjective Memory Scale; BRIEF-A = Behavioral Rating Inventory of Executive Function-Adult version; CERAD-self = Consortium to Establish a Registry for Alzheimer's Disease-Self version; GDS Short = Geriatric Depression Scale-Short version.

<sup>a</sup> A subset of participants ( $n = 260$ ) received all seven questionnaires (ADL Abbrev, BRIEF-A, CAPM-C, CERAD-self, Einstein HSA, GDS Short, MIA); the remaining participants received three questionnaires (Einstein HSA, CERAD-self, GDS Short). <sup>b</sup> A subset of participants from the University of Pittsburgh (Pitt Substudy,  $n = 174$ ) received all three questionnaires (MFQ, CFQ, SCCS); the remaining participants (from the MYHAT study) received one questionnaire (SCCS).



**Table 2**  
*Key Features of Questionnaires Used by Participating Studies*

Questionnaire/item abbreviation	Questionnaire/item full name	Study abbreviation(s)	References	Total number of original items	Number (%) of items relevant for analysis <sup>a</sup>	Final number of items retained
1. ADL	Activities of Daily Living Rating Scale Self-Version	Dart-Indiana	Saykin (1992)	63	33 (52%)	33 (52%)
2. ADL Abbrev	Activities of Daily Living Rating Scale Self-Version, Abbreviated	EAS Substudy 1	Saykin (n.d.)	50	45 (90%)	43 (86%)
3. AgeCoDe MQ	AgeCoDe Study Memory Questions	AgeCoDe	Jessen et al. (2007)	2	2 (100%)	1 (50%)
4. AIBL Screen	Melbourne Australian Imaging Biomarkers and Lifestyle Flagship Study of Ageing, Memory Screen Item	AIBL	Ellis et al. (2009)	1	1 (100%)	1 (100%)
5. Ams Dem Cohort CC	Amsterdam Dementia Cohort, subjective cognitive concerns screener	Ams Dem Cohort	van der Flier and Scheltens (2018)	5	5 (100%)	5 (100%)
6. Blessed FAQ	Blessed Functional Activities Scale	HELLAD	Blessed et al. (1968)	8	6 (75%)	6 (75%)
7. BRIEF-A	Behavioral Rating Inventory of Executive Function-Adult version	Dart-Indiana; EAS Substudy 1	Roth et al. (2005)	75	36 (48%)	32 (43%)
8. CAPM-C	Comprehensive Assessment of Prospective Memory-Section C	EAS Substudy 1	Chau et al. (2007)	15	13 (87%)	10 (67%)
9. CCI	Cognitive Change Index	ADNI; Ams Dem Cohort; Dart-Indiana; IADC EAS Substudy 2	Rattanabannakit et al. (2016); Saykin et al. (2006) Nester et al. (2022)	20	20 (100%)	20 (100%)
10. CCI-40	Cognitive Change Index-extended 40-item version	EAS Substudy 2	Saykin et al. (2006)	41	41 (100%)	41 (100%)
11. CDS-Q	Cognitive Difficulties Scale	Insitut des Matériaux Poreux de Paris Caen Group	McNair and Kahn (1983)	39	33 (85%)	32 (82%)
12. CERAD-self	Consortium to Establish a Registry for Alzheimer's Disease-Self-Version	EAS/EAS Substudy 1	Morris et al. (1989)	36	22 (61%)	12 (33%)
13. CFI	Cognitive Function Instrument, Self-Report	A4	Amariglio, Donohue, et al. (2015)	15	12 (80%)	12 (80%)
14. CFQ	Cognitive Failures Questionnaire	Pitt Substudy	Broadbent et al. (1982)	25	17 (68%)	11 (44%)
15. CLSA SCD/worry	Canadian Longitudinal Study in Aging, SCD/worry item	CLSA	Raina et al. (2009, 2019)	1	1 (100%)	1 (100%)
16. ECog Self	Everyday Cognition-Subject/Self-Report	ADNI; Dart-Indiana; HABS; IADC; MCSA	Farias et al. (2008)	40	40 (100%)	40 (100%) <sup>b</sup>
17. Einstein HAS	Albert Einstein Health Self-Assessment	EAS/EAS Substudy 1	Derby et al. (2013)	20	6 (30%)	4 (20%)
18. GDS Long	Geriatric Depression Scale-Long version	Dart-Indiana; Victoria SCDS; HABS-HD	Yesavage et al. (1982-1983)	30	4 (13%)	4 (13%)
19. GDS Short	Geriatric Depression Scale-Short version	A4; EAS/EAS Substudy 1; HELIAD	Sheikh and Yesavage (1986)	15	1 (7%)	1 (7%)
20. HELIAD SCI	HELIAD Subjective Cognitive Items Short Form of the Informant	HELIAD	Margioli et al. (2020)	7	7 (100%)	7 (100%)
21. IQCODE Short	Questionnaire on Cognitive Decline in the Elderly, Self-Report version	Dart-Indiana; Sydney	Jorm (1994)	16	15 (94%)	15 (94%)
22. LEILA 75+ Questions	Leipzig Longitudinal Study of the Aged Questionnaire	LEILA 75+	Luck et al. (2010)	4	4 (100%)	1 (25%)
23. MAC-Q (a) <sup>c</sup>	Memory Complaint Questionnaire	A4; AIBL; ADC-NYU; Victoria SCDS	Riedel-Heller et al. (1999)	6	6 (100%)	5 (83%)

(table continues)

Table 2 (continued)

Questionnaire/item abbreviation	Questionnaire/item full name	Study abbreviation(s)	References	Total number of original items	Number (%) of items relevant for analysis <sup>a</sup>	Final number of items retained
24. MAC-Q (b)	Memory Complaint Questionnaire	Sydney	Crook et al. (1992)	6	6 (100%)	5 (183%)
25. MATS	Memory and Aging Telephone Screen	Dart-Indiana; Victoria SCDS	Rabin et al. (2007)	13	13 (100%)	12 (92%)
26. MCSA Blessed	Mayo Clinic Study of Aging, adapted Blessed Memory scale	MCSA	Blessed et al. (1968); van Harten et al. (2018)	6	5 (83%)	5 (83%)
27. MFE-30	Memory Failures Everyday	Fundació	Lozoya-Delgado et al. (2012)	30	29 (97%)	26 (87%)
28. MFQ	Memory Functioning Questionnaire	HABS; Pitt Substudy; WRAP	Gilewski et al. (1990)	63, 64, or 54 <sup>d</sup>	49 (78%); 50 (78%); or 39 (72%)	48 (78%); 49 (78%); or 35 (90%)
29. MIA	Metamemory in Adulthood	EAS Substudy 1; Victoria	Dixon et al. (1988)	108	35 (32%)	32 (30%)
30. MMQ	Questionnaire	SCDS				
31. SCCS	Multifactorial Memory Questionnaire	Bonn Memory Clinic; CLSA	Troyer and Rich (2002)	57; 20 <sup>e</sup>	32 (56%)	32 (56%)
	Subjective Cognitive Complaint Scale, also called the Subjective Memory Scale	MYHAT; Pitt Substudy	Ganguli et al. (2010)	30	27 (90%)	21 (70%) <sup>f</sup>
32. SCD-Q: Part 1 MyCog	Subjective Cognitive Decline Questionnaire: Part 1, My Cognition	Barcelona Group	Rami et al. (2014)	27	26 (96%)	26 (96%)
33. SMCQ	Subjective Memory Complaints Questionnaire	HABS-HD	Youn et al. (2009); O'Bryant et al. (2022)	14	14 (100%)	14 (100%)
34. Amsterdam SCF	Subjective Cognitive Functioning	Ams Dem Cohort	Aalten et al. (2014)	4	2 (50%)	1 (25%)
35. SMDS	Subjective Memory Decline Scale	AgeCoDe	Jorm et al. (2001)	4	4 (100%)	4 (100%)
36. Squire	Squire Memory Self-Rating Questionnaire	Dart-Indiana	Squire et al. (1979)	18	18 (100%)	18 (100%)
37. STDA	Adapted from Structured Telephone Dementia Assessment	HABS	Go et al. (1997)	7	7 (100%)	6 (86%)
38. Sydney-SCQ Wave 1	Sydney Subjective Complaint Questions, Wave 1	Sydney	Slavin et al. (2010)	18	16 (89%)	12 (66%)
39. VMAC CCQ	Vanderbilt Memory and Alzheimer's Center Cognitive Complaint Questionnaire	VMAC	Gifford et al. (2014)	57	56 (98%)	51 (90%)
40. WRAP Memory Item	Wisconsin Registry for Alzheimer's Prevention, Memory Problem Item	WRAP	Nicholas et al. (2017)	1	1 (100%)	1 (100%)
Total				963 <sup>g</sup>	690 <sup>h</sup>	635 <sup>i</sup>

Note. Dart-Indiana = Dartmouth Memory and Aging Study/Indiana Memory and Aging Study; EAS = Einstein Aging Study; HELLAD = Hellenic Longitudinal Investigation of Aging and Diet; ADNI = Alzheimer's Disease Neuroimaging Initiative; IADC = Indiana Alzheimer's Disease Center Cohort; A4 = Anti-Amyloid Treatment in Asymptomatic Alzheimer's study; SCD = subjective cognitive decline; HABS = Harvard Aging Brain Study; MCSA = Mayo Clinic Study of Aging; Victoria SCDS = Victoria Subjective Cognitive Decline Study; HABS-HD = Health and Aging Brain Study; Health Disparities; ADC-NYU = Alzheimer's Disease Center Clinical Core and Center for Brain Health; CLSA = Canadian Longitudinal Study on Aging; MYHAT = Monongahela-Youghiogheny Health Aging Team.

<sup>a</sup> Items relevant for the current analysis relate to the self-perceived cognitive functioning (see Method and Results sections, for details about eliminated items). <sup>b</sup> In the HABS study, data were available for 39 of 40 ECog Self items. <sup>c</sup> The MAC-Q was separated into versions a and b because of different timeframes referenced (high school or college for A4, AIBL, NYU, and Victoria SCDS studies change versus 5 years ago for the Sydney study). <sup>d</sup> Also, the ADC-NYU version of the MAC-Q contains a different item stem for Item 4, which was not used in the present study (bringing the total number of MAC-Q items to one less for the NYU study). <sup>e</sup> The Pitt Substudy, HABS, and WRAP study versions of the MFQ contain 64, 63, and 54 items, respectively. <sup>f</sup> The MMQ has three subscales, of which the CLSA only includes one (i.e., Memory mistakes scale with 20 items). <sup>g</sup> In the University of Pittsburgh MYHAT study, data were available for 17 of the 21 SCCS items. <sup>h</sup> This total excludes 35 items from the CCI and GDS Short, as items from these questionnaires also appear on the lengthier versions of these questionnaires. <sup>i</sup> This total excludes 21 items from the CCI and GDS Short, as items from these questionnaires also appear on the lengthier versions of these questionnaires. <sup>j</sup> This total includes the 49-item version of the MFQ.

4 = *disagree*, and 5 = *strongly disagree*. Also, item response categories were combined if a category had fewer than five respondents. This was the case for items that had Likert-type response scales where few participants endorsed categories at the extreme end(s) of the scale. These sparse extreme response categories were combined with the adjacent, less extreme category/categories, until the new combined category had five or more respondents.

#### (4) Assessed IRT Assumptions in the Reference Sample

The reference sample defines the latent trait in IRT analyses. The Mayo Clinic Study of Aging (MCSA), a large community-based sample of neuropsychologically intact older adults, served as the reference study for this analysis. The MCSA utilizes a fairly extensive assessment of self-perceived cognitive functioning that includes the Mayo Clinic Study of Aging, adapted Blessed Memory scale (Blessed et al., 1968; van Harten et al., 2018), and the Everyday Cognition–Subject/Self-Report (ECog Self; Farias et al., 2008), the latter of which assesses self-perceived cognitive functioning in multiple domains that include memory, language, visuospatial/perceptual ability, and executive functioning/divided attention. Moreover, the ECog Self is used in multiple studies, which facilitates the linking of studies and items.

We used Mplus Version 8.5 (Muthén & Muthén, 1998–2017) for IRT analyses and the R statistical software package Version 4.0 (R Core Team, 2021) for data management and other study analyses. IRT is a latent variable technique used to describe the relationship between item responses and the latent trait presumed to underlie those responses, which in our case was the level of self-perceived cognitive functioning. The latent trait is assumed to be normally distributed with a mean of 0 and a *SD* of 1 (Lord, 1980). By fitting IRT models to item data, self-perceived cognitive functioning performance can be estimated conditional on the observed item responses and their respective item parameters; additionally, IRT methods can be used to generate scores that are comparable regardless of what items were utilized in a particular study.

Several key assumptions underlie the IRT framework, including unidimensionality of the measured trait, local independence, and monotonicity (McHorney & Cohen, 2000; Nguyen et al., 2014). We assessed the IRT assumptions in the MCSA sample by fitting a graded response IRT model (Samejima, 1997), which is typically used for ordered polytomous response data. All models were estimated using a Bayesian estimator (Haug, 2012; Winkler, 1993). A weakly informative prior distribution ( $N(0, 5)$ ) was assumed for the model parameters. The Bayesian estimator allows for better performance in the small sample sizes that several studies provided (Hox et al., 2012; Lee & Song, 2004).

#### Unidimensionality Assumption

Based on the previous literature (Gifford et al., 2015; Snitz et al., 2012), we hypothesized that a unidimensional (or single-factor) model might provide an adequate representation of the self-perceived cognitive functioning trait. In other words, the construct being measured, self-perceived cognitive functioning, is considered to have a single, dominant factor that accounts for item endorsement and covariation among the items (with small factors present beyond the major factor). This has been shown to be the case in previous work with cancer patients—where abilities and concerns about

cognition loaded on a single factor were conceptualized as a single dimension despite measuring different aspects of self-perceived cognitive functioning (Lai et al., 2014). We also considered the possibility that our data were multidimensional with strong factors beyond the major factor (i.e., differing cognitive domains might induce some meaningful secondary structure/residual covariance). For self-perceived cognitive functioning data, it is reasonable to expect that multidimensionality will arise from differences in content domain (e.g., some items assess the ability to remember recent conversations, others inquire about language problems, visuospatial ability, or language skills), temporal referents (e.g., “Compared to 10 years ago, do you notice change in ...” versus “Do you currently have difficulty with ...”), or other key features of the questionnaire and its items. In a given study, any of these differences in item properties could contribute to subdomain structures and cause an item to deviate from strict unidimensionality. Such deviations from unidimensionality complicate the linking of the latent self-perceived cognitive functioning trait across multiple studies that use similar but nonequivalent questionnaires/items such that using a unidimensional-based calibration in the IRT model results in bias.

Prior to data analysis, we categorized items into cognitive subdomains and considered a bifactor model, which allows for additional covariation among a set of items beyond the general factor based on these cognitive domains. Fitting a bifactor model proved to be an intractable problem in our analysis given the various possible subdomain structures (beyond cognitive domain) such as temporal referents or item content and structure. We therefore prioritized the unidimensional model, using two approaches to characterize the extent to which our inferences were likely to be biased due to unmeasured multidimensionality. Both approaches made use of data collected in the MCSA because, as noted, this was the study used to set the metric of our latent self-perceived cognitive functioning trait. Moreover, with one study, there was only one pattern of multidimensionality to be addressed for the item set.

**Item-Level Fit of Unidimensional and Bifactor Models: MCSA Study.** The first method used to characterize threats to interpreting parameter estimates due to ignoring multidimensionality involved contrasting the posterior predictive *p* values (PPP) for each of the items in the model under a unidimensional and a bifactor model. Our bifactor model specification was based on the cognitive domain assignments. To assess unidimensionality in the MCSA, we fit unidimensional and bifactor item response models and examined PPP for each item in the model. The PPP compares the  $\chi^2$  value from the model-generated data to the  $\chi^2$  from the observed data and is approximated by the proportion of times the model-generated data  $\chi^2$  is greater than the observed data  $\chi^2$  (Gelman, 2013; Meng, 1994). A low PPP implies that it is unlikely that the observed data derive from a data-generating model that is represented by the model used in the analysis. Conventionally, PPP less than .05 indicate poor fit. The distribution of PPP for all items is shown in Appendix Figure C1. This figure reveals that, overall, the fit was good. The mean PPP was higher in the unidimensional case relative to the bifactor case (.38 vs. .28, respectively). In both models for the MCSA, all items had PPP greater than .05. Therefore, we inferred that for the MCSA self-perceived cognitive functioning items, the pattern of correlations was consistent with a unidimensional model.

**Differences in Common Latent Trait Estimates Under Unidimensional and Multidimensional Models.** The second method used to probe the level of *misfit* introduced with the

assumption of unidimensionality involved examining the estimated factor scores derived from a unidimensional and a priori specified bifactor model. Results of these analyses are summarized in Appendix Figure C2 (left and right panels). These plots show the scatterplot of factor scores (left) and the difference in factor scores as a function of the average of the two-factor scores (right). In our work, we typically consider absolute differences in factor scores greater than 0.3 to be large and meaningful (Gibbons et al., 2017). Only 3% of the participants in the MCSA had a greater than 0.3 difference in factor scores. Therefore, the unidimensional model-derived factor score estimate would be deemed sufficiently precise for 97% of the sample. Further inspection of Appendix Figure C2 revealed that there was a region where the unidimensional factor scores seemed to produce a biased high estimate of the latent trait: values greater than +2 on the latent trait distribution, or the expected top 2.5% of scores. Thus, bias was apparent in the estimates from the unidimensional models, but it was restricted to the upper tail of self-perceived cognitive functioning (i.e., best performers/lowest cognitive concerns). As higher values on the latent trait indicated better self-perceived cognitive functioning, these 2.5% of respondents were of the least interest, clinically, if the goal was to identify individuals with higher levels of concern about their cognitive functioning.

**Item-Level Fit of Unidimensional and Bifactor Models: All Studies.** Further support for the assumption of unidimensionality came from examining the PPP for all items included in our analytic models, in all studies, from a unidimensional model and a bifactor model (Appendix Figure C3). The secondary structure was specified according to a priori assignment of items to specific domains based on manifest item content. For the PPP obtained, assuming unidimensionality, 96 of 1,112 had PPP less than .05, suggesting that for the vast majority of items, the data were consistent with the assumption of a data-generating model that was unidimensional. For the bifactor models, 76 of 1,112 had PPP less than .05, suggesting that for this fraction of items, even our bifactor specification was insufficient to account for the observed covariation. For these items, it is likely that additional covariation among items other than due to the manifest cognitive domain was responsible for the misfit, and we suspect that other item features (e.g., temporal referent, format and range of response options, whether items tapped ability or change) could be responsible or that our a priori domain assignments did not conform to observed patterns of covariation among the items.

### **Local Independence and Monotonicity Assumptions**

The next assumption, local independence, means that for a given latent trait (in our case, self-perceived cognitive functioning), item responses are unrelated (Edwards et al., 2018); the assumption is usually tested for pairs of items (Kim et al., 2011), and meeting this assumption can be an additional indication of adequate model fit. In the MCSA study, we checked for and eliminated items with factor loadings  $> 1$ , which implies that there was a strong residual correlation or logical dependency with another item. The monotonicity assumption means that as the trait level increases, the probability of a correct response also increases (Kang et al., 2018; Nguyen et al., 2014)—that is, the higher the factor score, the more likely a respondent is to endorse concerns about cognitive functioning. In the MCSA sample, we checked for and eliminated items that demonstrated a common factor loading qualitatively

different from theoretical expectations (loadings  $< 0$ ), which implies that the directionality of the item is incorrect and may indicate the measurement of a distinct construct.

### **(5) Decision to Utilize a Unidimensional Model**

In sum, across a set of models assuming unidimensionality and another set assuming multidimensionality (modeled with the bifactor approach), we compared item-level fit and values estimated for latent traits. We found that item-level fit was superior under the assumption of multidimensionality. However, we also found that in our reference sample, the item fits (PPP) were consistent with the assumption of unidimensionality, and, in our overall sample, only 9% of items demonstrated poor fit ( $PPP < .05$ ) under the assumption of unidimensionality, and 7% of items demonstrated poor fit in models intending to capture multidimensionality. While this result favors the bifactor approach, it also reveals that the bifactor approach alone is insufficient to account for the poor fit of all items in our large item set, and the gain in fit of a bifactor approach over a unidimensional approach was not large. We then examined differences in factor score estimates obtained in the reference sample under a unidimensional and bifactor model. For 97% of the sample, the differences were less than 0.3 standard errors, an acceptable level to assume equivalence. Bias was apparent in the estimates from the unidimensional models, but it was restricted to the upper tail of self-perceived cognitive functioning performance (i.e., best performers). Based on these results, we concluded that there was little evidence of important bias that would result from using a unidimensional model, and we proceeded with this approach.

### **(6) Examination of Differential Item Functioning**

We performed a brief analysis of uniform differential item functioning (DIF) with respect to study to determine whether participants with the same latent trait level across different studies differed in their response patterns on identical items. To be included in the DIF analysis, an item needed to be present across at least two studies. For most studies, the number of linking items was small and not conducive to formal DIF testing. This limited us to studies that used the same questionnaire. The studies that were most useful for a DIF analysis were the five that included the ECog Self: MCSA, ADNI, Dartmouth Memory and Aging Study/Indiana Memory and Aging Study (Dart-Indiana), HABS, and the Indiana Alzheimer's Disease Center Cohort (IADC).

### **(7) Saved MCSA Item Parameters to Item Bank**

An item bank is a collection of item parameters calibrated by using IRT models. It is created by saving the estimated item parameters from the IRT model with new items added to the bank from calibrating subsequent studies. Item banks can be used by instrument developers to refine existing measures by selecting the "best" set of questions for a particular purpose or for developing new instruments including using computerized adaptive testing (CAT) platforms and/or static, fixed-length short forms (Lai et al., 2003, 2014). The two sets of parameters were the item discrimination (or measurement slope) and item difficulties or thresholds (i.e., locations on an item-level latent response variable at which the most likely response moves from a lower to a higher



value). The item bank was initiated with parameter estimates, taken as the medians of the posterior distribution for each of the IRT item parameters, from the MCSA.

## (8) Item Linking/Harmonization in MCSA

We accomplished the statistical harmonization of questionnaires using a sequenced estimation of IRT models.

There are two broad approaches to item linking—common item anchoring and common group equating. In common group anchoring, an assumption is that the groups are equivalent on the trait of interest—and the different questionnaires can simply be linked. However, we could not make that assumption given inherent differences in our samples for key sociodemographic and other variables. Instead, we utilized the common items anchoring approach, where there are *identical* or *equivalent* items across different questionnaires, and we fit a series of IRT models using item banking. The linking transformation served to place the items on the same scale (Vale, 1986).

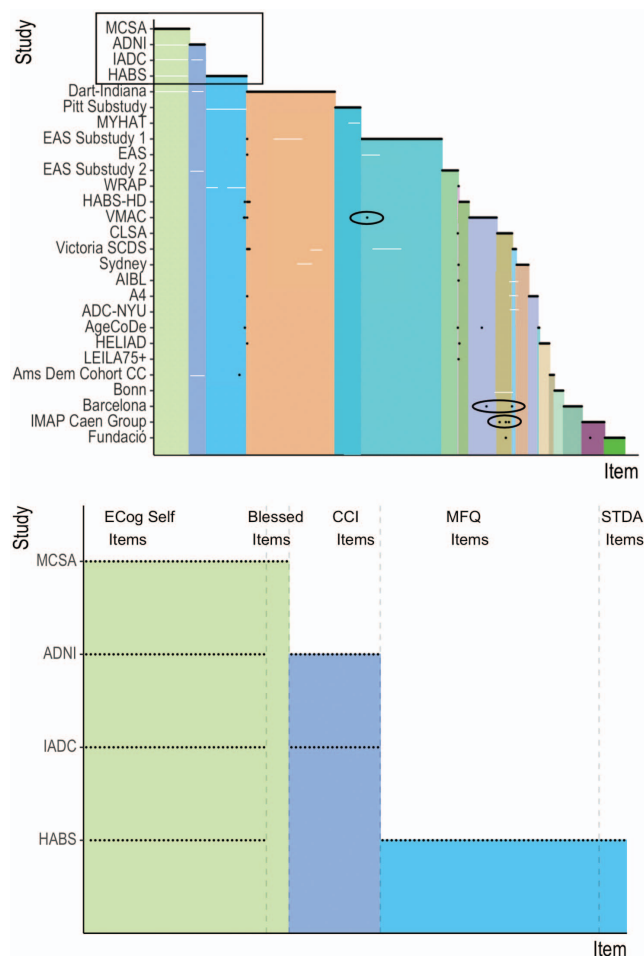
To perform statistical harmonization (or linking) of the questionnaires, we used cocalibration (i.e., putting the questionnaires on the same scale) based on IRT in a series of steps. We initiated the harmonization sequence with the MCSA and its two questionnaires, ECog Self and MCSA adapted Blessed Memory scale (MCSA Blessed; Blessed et al., 1968; van Harten et al., 2018). This approach anchored the self-perceived cognitive functioning latent trait to a community-based sample with no neuropsychological impairment. As the reference population, the mean trait level in MCSA was set to 0 (with *SD* set to 1), affording comparability with other studies as they were harmonized. As an example, if the mean trait level for a study entered after MCSA was 0.5, this would indicate that the average participant in that study was .5 *SD* higher on the self-perceived cognitive functioning trait than the average MCSA participant. Figure 2 shows the ordering of studies in the harmonization process, with MCSA and its questionnaires entered first. The order in which subsequent studies were harmonized after MCSA was determined by considering item overlap with previous studies and sample size. As described above, response categories for certain items were combined. Decisions related to combining response categories for a given item were made for the first study in the harmonization process to use the item. For example, for the ECog Self item “finding my way around a familiar neighborhood,” the response options of “consistently a little worse” and “consistently much worse” were combined. These decisions were then carried forward to subsequent studies that used the item (i.e., to the four other studies that used the ECog Self, see Table 2).

## (9) Algorithm Carried Through to Subsequent Studies

After MCSA, we analyzed other studies sequentially. If subsequent studies included items already calibrated in the item bank, we constrained those item parameters to values in the item bank and freely estimated parameters for items not already in the bank. New items and their parameters were thus added to the item bank. This approach (banking and linking) required that—apart from the reference study—each subsequent study had at least one item in common with a previously calibrated item in the bank. This implies that studies with no overlap of items could not be included in the linking process, and to improve the quality of the estimates, we

**Figure 2**

*Overlap Among Self-Perceived Cognitive Functioning Items Across Studies*



*Note.* Figure 2 displays the overlap between self-perceived cognitive functioning items in the studies (including substudies) and shows how the different SCD-I working group studies and items are linked together. In the top panel, the bars/dots represent items administered within a study. The bottom panel displays additional detail for the portion of the top panel enclosed in a gray rectangle. Overall, the figure reveals very little item overlap across studies and no single item represented across all studies. The figure displays 27 instead of 24 studies because the EAS Substudy 1, EAS Substudy 2, and Pitt Substudy were linked separately from EAS and University of Pittsburgh/MYHAT studies. Please refer to Table 1, for the definitions of the study abbreviations. Please refer to Table 2, for information about study items and questionnaires.

needed to balance the size of the study and number of items represented in the bank in choosing the order in which items were added. As was done in MCSA (discussed above), we performed data quality checks and preliminary model checks for items with measurement slopes that corresponded to standardized factor loadings that were problematic (i.e.,  $<0$  or  $>1$ ). This was an additional quality check on whether the direction and potential recoding for each item were done correctly. It also enabled us to identify items where we had made an error, necessitating reverse coding of items that we had erroneously failed to reverse code. In

this way, we were able to use the signs for the loadings as a further check on the directionality of each item as included in the final model. Beyond coding discrepancies, several other items were excluded via two iterations of a cycle of excluding items, reharmonizing remaining items, and examining factor loadings until all factor loadings were between 0 and 1 for each study.

### (10) Calculated Item Information and Identified Items With the Highest Information Values (“Top” Items)

Item information was determined by the derivative of the item response function—that is, the increasing probability of endorsing an item response in a higher category as a function of the underlying trait level. The amount of information at a given latent trait level is the inverse of its standard error of measurement (*SEM*) squared; therefore, the larger the amount of information provided by the item, the greater the precision of the measurement (Hays et al., 2000). The amount of information (and precision) varies across levels of the latent trait. The location on the latent trait with the maximum amount of information for a binary or dichotomous item is where the probability of positively or negatively endorsing a response category is equal.

We plotted item information curves (Hambleton & Cook, 1977), with item information plotted against the latent trait. Item information curves are useful because the item parameters and the latent trait are on the same scale, and items can be selected based on their measurement ability in regions of interest on the latent trait (Baker & Kim, 2017). For example, measuring self-perceived cognitive functioning on the “higher” end might be of interest when studying healthy older adults living in the community, whereas the “lower” end might be more interesting in individuals who present to a memory clinic with concerns about cognition. Some items or questionnaires may provide more information in one region of the latent trait and less in a different region. For the current analysis, we determined several regions of interest on the latent trait. The first region was from 0 to +1 *SD* above the mean of the latent trait, which may capture individuals with better self-perceived cognitive functioning—that is, no or few concerns about cognitive function. A second region was from the mean to –1 *SD* below the mean, which may capture individuals with a mild level of self-perceived cognitive functioning concerns. A third region was from –1 to –2 *SDs* below the mean of the latent trait, which may capture individuals with worse self-perceived cognitive functioning—that is, a moderate to high level

of cognitive concerns. From the regions of interest, we identified “top items” with the highest information values.

## Results

### Participant and Study Characteristics

Data were available for 61,141 participants. Because we were interested in individuals with normal performance on neuropsychological tests, we excluded one of the studies that only contributed data for cognitively impaired individuals (i.e., Pre-Al, Hôpital de la Salpêtrière, Assistance Publique–Hôpitaux de Paris). We also excluded subsets of cognitively impaired participants from 18 of the 24 remaining studies (see Table 1). After removing these individuals ( $n = 8,111$ ), the final sample contained 53,030 neuropsychologically intact participants with data on self-perceived cognitive functioning items. Table 1 presents key features of the 24 participating working group studies. Two of the participating sites (i.e., Einstein Aging Study and University of Pittsburgh) had substudies as shown in Table 1 and Figure 2, but for this study were considered a single site/study, unless otherwise specified. Contributing studies had sites in the United States, Canada, Australia, Germany, United Kingdom, the Netherlands, Spain, Greece, France, and Japan. In addition, various research environments were represented including memory clinics ( $n = 4$ ), community-based samples ( $n = 7$ ), volunteer samples ( $n = 3$ ), population-based samples ( $n = 3$ ), a general practice registry ( $n = 1$ ), and mixed sampling approaches ( $n = 6$ ).

### Questionnaire Characteristics and Linking Items

Forty questionnaires qualified for inclusion (Tables 1 and 2). As shown in Table 1, for the 24 primary studies (i.e., excluding the Einstein Aging Study and University of Pittsburgh substudies), the number of questionnaires used per study ranged from 1 to 8 (mean 2.3; median 2.0). Also, there was considerable variability in the number of usable items per study, ranging from 1 to 173 (mean 36.2; median 27.5; interquartile interval 46.5–17.0). There was minimal overlap of questionnaires across studies, with the majority ( $n = 30$ , 75%) used only within a single study. Appendix D shows the number of studies in which a given questionnaire was utilized. In terms of item content, as shown in Table 3, over half of the items related to memory

**Table 3**  
*Percentage of 601 Items Within Each of Eight Cognitive Domains*

Cognitive domain	Number of items (%)	Description
Memory	344 (57%)	Includes short-term/long-term/episodic/semantic/prospective memory, and learning new information.
Executive function	97 (16%)	Includes organizing, planning, switching, initiating, multitasking, reasoning, judgment, problem-solving, decision-making, handling emergencies, impulsivity and self-regulation, clarity of mind, learning or operating machinery, handling money, and self-awareness of problems.
Attention working memory processing speed	63 (10%)	Includes basic attention, sustained attention, focused attention, concentration, divided attention, and alertness.
Language	52 (9%)	Includes expressive and receptive language, word finding, reading, and spelling.
Visuospatial skills	23 (4%)	Includes visuoperception, route finding, and directional orientation.
Orientation	9 (2%)	Includes orientation to person, time, place, or situation.
Calculation	7 (1%)	Includes basic calculation and arithmetic tasks.
General cognitive ability	6 (1%)	Includes memory and other thinking abilities grouped together in a single item.

(57%), followed by executive function (16%), attention/working memory/processing speed (10%), and language (9%).

Table 2 presents key features of the questionnaires including the total number of original items from each questionnaire with a total across questionnaires of 963 items (this total includes the 64-item version of the MFQ). In the penultimate column, we present the number of items relevant for the current analysis because they relate to self-perceived cognitive functioning (total  $n = 690$ ). We excluded 273 items that related to other characteristics such as mood, personality, health complaints, basic activities of daily living, and use of memory/other cognitive strategies. Of the 690 items, 55 were subsequently discarded during the data analysis process, leaving 635 items. For example, we dropped items initially included but later determined to have critical content or response category limitations such as: “I am impulsive” and “Do you fail to listen to people’s names when you are meeting them?” These items could not be categorized within the eight cognitive domains and could capture something other than self-perceived cognitive functioning. Similarly, items such as “If your memory was 100% when you were 30 years old, what % would you say it is now?” and “How long ago did your memory problems start?” had free-response formats not easily handled by our modeling approach. As noted above, during statistical harmonization, we also dropped items with loadings  $<0$  or  $>1$ .

As described above, we also considered items with overlapping content that were similar enough to be treated as *equivalent* in the analyses despite minor differences in wording. Equivalent items were given the same item number and treated as a single item in the analysis; therefore, we lost one item for every replicate of an equivalent item. There were 27 sets of equivalent items: 22 items had one replicate, three items had two replicates, and two items had three replicates. In total, 34 ( $22 + 6 + 6$ ) replicates were lost from the item total because they were considered equivalent items. Thus, we began with 635 items but after the equivalence process that yielded 11 sets of equivalent items, our final item count was  $635 - 34 = 601$ . These *equivalent* items ( $n = 27$ ) in combination with the *identical* items ( $n = 220$ ), comprised the *linking items* ( $n = 247$ ), which were used to link studies in the IRT analyses. As noted above, identical items were items from questionnaires that were used in more than one study.

Every study had at least one item in common with another study (*equivalent* or *identical*)—that is, all studies had at least one linking item. Appendix E shows the number of linking items between pairs of studies. As shown in bold in Appendix E, the highest number of linking items between two studies was 60 for ADNI and the IADC, which both utilized the ECog Self and the Cognitive Change Index (CCI). For study pairs that had linking items among them (e.g., MCSA and IADC had 40, Memory Clinic—Fundació Alzheimer Centre Educacional and the Bonn Memory Clinic had one), the mean number of linking items between studies was 11.1 ( $SD = 12.6$ ). As shown in Appendix F, each study shared linking items with an average of 6.2 other studies,  $SD = 3.2$  (range 1–14).

Figure 2 shows the ordering of studies for the IRT analysis and the overlap among self-perceived cognitive functioning items across studies including the substudies. In the top panel, the horizontal black and white bars consist of multiple dots that represent items administered within a study. The bars that align vertically represent *identical* items that come from the same questionnaire. The bars that are black at the top of the vertical-colored regions indicate the study used to calibrate the items for the item bank. The white bars are identical to items in the black bars with which they vertically align

and represent items already calibrated by the top-most study above it. *Equivalent* items are single dots found within a colored region that indicate items from different instruments merged for study analysis. Items that are within a colored region (both black dots and white bars) provide the linkage for a study between previously calibrated items and items to be calibrated.

As shown in the top panel of Figure 2, there was very little item overlap across studies, and no single item was represented across all studies. In some cases, a small number of items provided the linkage between specific studies. For example, the Barcelona Group had two items that linked to other studies. This is represented by a circle around two dots: (a) the Barcelona Group study was linked to the Bonn Memory Clinic and Canadian Longitudinal Study on Aging (CLSA) by a single item (single dot in the tan region) and (b) the Barcelona Group was linked to the Vanderbilt Memory and Alzheimer’s Center (VMAC) study by a single item (single dot in the purple region). Another example is the three items from the Institut des Matériaux Poreux de Paris Caen Group, which were *equivalent* to those from the Bonn Memory Clinic (three dots circled in the tan region). A final example from the figure is the VMAC study, which had links to 14 other studies. Here, we highlight the single dot circled in the teal region, which is an item from the VMAC study that is *equivalent* to an item from the Einstein Aging Study.

The bottom panel displays additional detail for the portion of the top panel enclosed in a rectangle. In this panel, each dot reflects an item from a given questionnaire (dots are ordered so that all of a given questionnaire’s items are grouped together). MCSA was the first study to be harmonized and is at the top (MCSA has the ECog Self and MCSA Blessed). ADNI was then harmonized and has the ECog Self and CCI questionnaires. The ECog Self provides the link to MCSA. This is the first encounter with the CCI, and it will therefore be calibrated with the ADNI sample. The IADC has two questionnaires (ECog Self and CCI), both of which were previously calibrated; therefore, the IADC does not provide any new questionnaires to be calibrated. The HABS will be linked through the ECog Self and will calibrate the Memory Functioning Questionnaire (MFQ) and adapted from Structured Telephone Dementia Assessment (STDA).

### Differential Item Functioning Analysis

As discussed above, to lend credibility to the harmonization process, we conducted a DIF analysis with respect to study on the subset of items identified as linking items from the ECog Self in five studies ( $n = 5,742$ ). As shown in Appendix G, we found evidence of DIF using a statistical significance criterion for flagging items with DIF; 36 of 160 study-item pairs (23%) had evidence of DIF, which is not surprising given the large sample size. Accounting for DIF in the estimation of latent traits accounted for between 2% and 19% of the DIF-naïve group differences in the self-perceived cognitive functioning trait. Moreover, the absolute differences in the mean level of the ECog Self common latent trait were between 0.01 and 0.03  $SD$  units (Appendix G), which are considered very small (i.e., the latent trait estimates do not meaningfully change after accounting for DIF).

### Item Content Examination

Higher scores on the underlying latent *self-perceived cognitive functioning* trait represented better self-perceived cognitive functioning. In support of the idea that these items actually measured the

trait of interest, the vast majority of items (575 of 601) had standardized factor loadings greater than 0.3, and the median (and interquartile interval) was 0.61 (0.50–0.71). The ranges over which we present item information values for selected items reflected different levels of the latent trait, and these tables (see below) highlight items that might be most relevant for individuals within a given range.

Tables 4 and 5 present the 30 items with the highest information values and lowest *SEM* in the underlying latent trait ranges of  $-2$  to  $-1$  SDs and  $0$  to  $+1$  SD, reflecting psychometrically sound items likely to be endorsed by participants with high and low levels of concern about their cognitive functioning, respectively. Items in the middle range ( $-1$  to  $0$  SD) had almost complete content overlap with items in the  $-2$  to  $-1$  or  $0$  to  $+1$  SD ranges, and are presented in Appendix H. Of the top 15 items with good measurement precision in the higher concern range of  $-2$  to  $-1$ , memory accounted for 80% of items ( $n = 12$ ). By contrast, of the top 15 items in the higher concern range of  $0$  to  $+1$ , there was more variability in cognitive domains represented, with 33% ( $n = 5$ ) of items related to memory, and the remaining items related to executive function ( $n = 5$ ), language ( $n = 3$ ), calculation ( $n = 1$ ), and attention and concentration ( $n = 1$ ). Across the ranges of interest, most items (70%,  $n = 21$ ) assessed perceptions of change (“C”) in ability or function from some prior state or level, while 30% ( $n = 9$ ) assessed current ability (designated “A”). With respect to response option types, of the top 30 items 67% ( $n = 20$ ) used Likert-type and rank-ordered/categorical item scales; notably, all of the dichotomous items ( $n = 10$ ) came from the higher concern range of  $-2$  to  $-1$ .

Figure 3 presents the item-level information plots for all 601 items (i.e., those with both high and low information values). The items that appear in blue have the highest information in the two ranges, and these are the items presented in Tables 4 and 5). Also, items that are dichotomous have only one threshold and therefore a single peak. Polytomous/Likert-scaled items have multiple thresholds and therefore can have multiple peaks. For polytomous items with multimodal curves, information typically covers a broader range—that is, there is a higher level of information over a broader range of ability level. Notably, most of the peaks fell within the  $0$  to  $-2$  range, reflecting worse self-perceived cognitive functioning, consistent with the design of these items in terms of capturing self-perceived difficulties and/or declines in cognitive functioning.

Appendix I presents the 15 items with the highest *SEM* in the underlying latent trait range of  $-4$  to  $+4$  SDs, reflecting items with weak psychometric properties across the range of the latent trait. For the items included in this analysis, additional parameters and features (e.g., item discrimination, item difficulty thresholds, summary of item content) are available upon request from the corresponding author.

## Discussion

### Overview

The present study harmonized self-perceived cognitive functioning questionnaires across international aging studies, with over 53,000 neuropsychologically intact participants, using an IRT item banking approach. The measurement properties of 601 items from studies of various sizes and settings were characterized, and items with the strongest psychometric properties were identified. The challenges we encountered in linking items and harmonizing these data speak to the effort involved in reviewing item content,

determining equivalences among items, managing the data files, and performing statistical analyses across the 24 studies and 40 questionnaires. Results may lead to the development of new questionnaires guided by psychometric properties, content, and other important features of items in our item bank.

### Dimensionality

We were able to replicate and expand upon earlier work in linking self-perceived cognitive functioning items and placing them on a common scale using IRT (Gifford et al., 2015; Snitz et al., 2012). We selected a unidimensional measurement model as our final modeling strategy. Overall, results did not provide convincing evidence to go forward with a bifactor model over a unidimensional model based on the estimated factor scores (which did not significantly differ between the two models) and the PPP for the individual items (which suggested a good fit for both models). As the secondary structure did not markedly impact scores, loadings, or fit, we determined that the single-factor model was appropriate for modeling self-perceived cognitive functioning. However, we cannot rule out that a different approach with a multidimensional factor structure for the multiple domains of self-perceived cognitive functioning might have presented a more nuanced representation. Nonetheless, we found the specification and estimation of multidimensional models intractable given the sparse coverage of domains and limited overlap of items across studies. As we evaluated the approach in this study, we found that if the structure of the secondary or specific factors was different in different studies, the meaning (and expected population values) of secondary factor loadings were also different. The implication is that for each study, the common and specific factors must be included in the linking operation, and this is an important area for future methodological and applied research.

The finding that a unidimensional model fits the data lends support to the idea that self-perceived cognitive functioning can reasonably be conceptualized as a unitary construct for measurement purposes, regardless of the specific domains tapped by items on a given questionnaire. In fact, in the original criteria set forth by Jessen et al. (2014), the authors stated that “cognitive” refers to any cognitive domain and is not restricted to memory. The term cognitive, as opposed to memory, was specifically chosen for the SCD criteria because the first symptoms of AD may not be limited to memory decline and older adults often report memory decline when they actually experience decline in other cognitive domains such as language or executive functioning and vice versa (e.g., reporting poor memory when the difficulty is actually in language or attention). Additionally, work by van Harten et al. (2018) showed that all ECog Self subdomains and multidomain scores (generated from items across cognitive domains) were associated with MCI. In this study, the multidomain score was slightly better than the separate domain scores, lending support to the idea that a general self-perceived cognitive functioning score rather than a specific domain score might be most predictive of underlying neurodegenerative change (van Harten et al., 2018).

### Features of Psychometrically Most Robust Items

Examination of the most robust items—those with the lowest *SEM*/highest information values from the specified ranges of interest—revealed several notable features. Top items in the low self-perceived cognitive functioning (high concern) range were



**Table 4**  
*Fifteen Items With the Smallest SEM in the Range of -2 to -1*

Questionnaire abbreviation	Cognitive domain	Item stem	Response options <sup>a</sup>	Temporal referent	Ability versus change	SEM; information
1. VMAC CCQ	Memory	Do you think you have problems with your memory?	No; yes	Current	A	0.37; 7.20
2. VMAC CCQ	Memory	Do you have problems with your memory compared to the way it was 5 years ago?	No problems; some minor problems	Compared to 5 years ago	C	0.39; 6.53
3. VMAC CCQ	Memory	Do you feel that your everyday life is difficult now due to your memory decline?	No; yes	Current	A	0.42; 5.65
4. VMAC CCQ	Memory	Do you have difficulty remembering a conversation from a few days ago?	No; yes	Current	A	0.42; 5.56
5. ADL Abbrev	Memory	Remembering what I intended to do	Much better; slightly better; no change; slightly worse; slightly to moderately worse; moderately worse; much worse <sup>b</sup>	Ability compared to 5 years ago	C <sup>b</sup>	0.43; 5.29
6. CFI	Executive	Compared to 1 year ago, do you have more difficulty managing money (e.g., paying bills, calculating change, completing tax forms)?	No; maybe; yes	Compared to 1 year ago	C	0.44; 5.27
7. CFI	Memory	Compared to 1 year ago, do you feel that your memory has declined substantially?	No; maybe; yes	Compared to 1 year ago	C	0.44; 5.09
8. HELIAD SCI	Memory	Do you have difficulty remembering things that you just read or heard?	No problems; problems	Current	A	0.46; 4.69
9. CFI	Attention	Do you have more trouble following the news or the plots of books, movies, or TV shows?	No; maybe; yes	With reference to 1 year ago	C	0.47; 4.46
10. SMCQ	Memory	Do you have difficulty in remembering an appointment made a few days ago?	No; yes	Current	A	0.48; 4.41
11. VMAC CCQ	Memory	Do memory problems make it harder to complete tasks that used to be easy?	No; yes	Current	A	0.48; 4.39
12. VMAC CCQ	Memory	Do you consider your own memory to be worse than others that are your same age?	No; yes	Current	A	0.48; 4.26
13. VMAC CCQ	Memory	Do you have complaints about your memory in the last 2 years?	No; yes	In the last 2 years	A	0.49; 4.17
14. VMAC CCQ	Memory	Do you lose objects more often than you did previously?	No; yes	Current	C	0.49; 4.13
15. CFI	Visuospatial	Are you more likely to become disoriented, or get lost, for example, when traveling to another city?	No; yes	1 year ago	C	0.50; 3.94

*Note.* SEM = standard error of measurement; ADL Abbrev = Activities of Daily Living Rating Scale Self-Version, Abbreviated; CFI = Cognitive Function Index; HELIAD = Hellenic Longitudinal Investigation of Aging and Diet; HELIAD SCI = HELIAD Subjective Cognitive Items; SMCQ = Subjective Memory Complaints Questionnaire; VMAC CCQ = Vanderbilt Memory and Alzheimer's Center Cognitive Complaint Questionnaire.

<sup>a</sup>The italicized response options were combined for purposes of data harmonization due to the categories having fewer than five respondents. <sup>b</sup>The ADL Abbrev scale uses a time referent compared to 5 years ago to capture current ability level on certain tasks, skills, or problem areas in relation to previous ability. As such, even though labeled as "C," indicating that it assesses change, the ADL Abbrev also taps into ability.

**Table 5***Fifteen Items With the Smallest SEM in the Range of 0 to +1*

Questionnaire abbreviation	Cognitive domain	Item stem	Response options <sup>a</sup>	Temporal referent	Ability versus change	SEM; information
1. ADL Abbrev	Executive	Focusing on goals or carrying out a plan	<i>Much better;</i> <i>slightly better<sup>b</sup>;</i> no change; <i>slightly worse;</i> <i>slightly to moderately worse;</i> <i>moderately worse;</i> <i>much worse</i>	Ability compared to 5 years ago	C <sup>b</sup>	0.50; 4.08
2. CFI	Language	Do you have more trouble recalling names, finding the right word, or completing sentences?	No; maybe; yes	Reference to 1 year ago	C	0.52; 3.77
3. ADL Abbrev	Memory	Remembering appointments or meetings	<i>Much better;</i> <i>slightly better;</i> no change; <i>slightly worse;</i> <i>slightly to moderately worse;</i> <i>moderately worse;</i> <i>much worse</i>	Ability compared to 5 years ago	C	0.54; 3.48
4. MMQ	Memory	I worry about my memory ability	Strongly disagree; disagree; undecided; agree; strongly agree	Feelings over the past 2 weeks	A	0.55; 3.29
5. ADL Abbrev	Attention	Remaining on task and not getting distracted by external stimuli	<i>Much better;</i> <i>slightly better;</i> no change; <i>slightly worse;</i> <i>slightly to moderately worse;</i> <i>moderately worse;</i> <i>much worse</i>	Ability compared to 5 years ago	C	0.57; 3.03
6. ADL Abbrev	Executive	Shifting easily from one activity to the next	<i>Much better;</i> <i>slightly better;</i> no change; slightly worse; <i>slightly to moderately worse;</i> <i>moderately worse;</i> <i>much worse</i>	Ability compared to 5 years ago	C	0.60; 2.75
7. ADL Abbrev	Executive	Making decisions on everyday matters	<i>Much better;</i> <i>slightly better;</i> no change; <i>slightly worse;</i> <i>slightly to moderately worse;</i> <i>moderately worse;</i> <i>much worse</i>	Ability compared to 5 years ago	C	0.61; 2.65
8. MAC-Q (a)	Memory	Remembering specific facts from a newspaper or magazine article you have just finished reading	<i>Much better now</i> <i>somewhat better now;</i> about the same; somewhat poorer now; much poorer now	As compared to high school or college	C	0.62; 2.63
9. MAC-Q (a)	Memory	Remembering the name of a person just introduced to you	<i>Much better now;</i> <i>somewhat better now;</i> <i>about the same;</i> somewhat poorer now; much poorer now	As compared to high school or college	C	0.62; 2.59
10. CFI	Memory	Compared to 1 year ago, do you find that you are relying more on written reminders (e.g., shopping lists, calendars)?	No; maybe; yes	Reference to 1 year ago	C	0.62; 2.59

*(table continues)*

Table 5 (continued)

Questionnaire abbreviation	Cognitive domain	Item stem	Response options <sup>a</sup>	Temporal referent	Ability versus change	SEM; information
11. ADL Abbrev	Executive	Reasoning through a complicated problem	Much better; slightly better; no change; slightly worse; <i>slightly to moderately worse; moderately worse; much worse</i>	Ability compared to 5 years ago	C	0.63; 2.56
12. ADL Abbrev	Language	Expressing myself through speech	<i>Much better; slightly better;</i> no change; slightly worse; slightly to moderately worse; <i>moderately worse; much worse</i>	Ability compared to 5 years ago	C	0.63; 2.54
13. ADL Abbrev	Executive	Organizing and managing my medication schedule	<i>Much better; slightly better; no change;</i> <i>slightly worse; slightly to moderately worse; moderately worse; much worse</i>	Ability compared to 5 years ago	C	0.63; 2.54
14. ADL Abbrev	Language	Understanding instructions or directions	<i>Much better; slightly better;</i> no change; <i>slightly worse; slightly to moderately worse; moderately worse; much worse</i>	Ability compared to 5 years ago	C	0.63; 2.54
15. ADL Abbrev	Calculation	Handling everyday arithmetic problems (knowing how much food to buy, how long it's been between visits from family and friends)	<i>Much better; slightly better;</i> no change; <i>slightly worse; slightly to moderately worse; moderately worse; much worse</i>	Ability compared to 5 years ago	C	0.63; 2.49

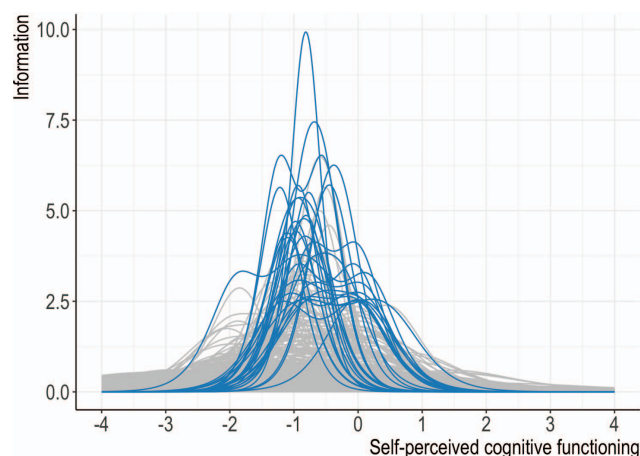
Note. SEM = standard error of measurement; ADL Abbrev = Activities of Daily Living Rating Scale Self-Version, Abbreviated; CFI = Cognitive Function Index; MAC-Q = Memory Complaint Questionnaire; MMQ = Multifactorial Memory Questionnaire.

<sup>a</sup>The italicized response options were combined for purposes of data harmonization due to the categories having fewer than five respondents. <sup>b</sup>The ADL Abbrev scale uses a time referent of compared to 5 years ago to capture current ability level on certain tasks, skills, or problem areas in relation to previous ability. As such, even though labeled as "C," indicating that it assesses change, the ADL Abbrev also taps into ability.

almost exclusively related to memory, both general (e.g., memory problems or decline relative to 1 year or 5 years ago) and specific (e.g., difficulty remembering conversations, remembering what you intended to, remembering appointments). By contrast, top items in the high self-perceived cognitive functioning (low concern) range assessed various aspects of cognition in addition to memory, most notably executive functioning (e.g., focusing on goals, decision-making, shifting between activities, managing a medication schedule) and language (e.g., word finding, self-expression, understanding instructions). Additionally, several items in the low self-perceived cognitive functioning (high concern) range tapped into the consequences of the memory problems (e.g., "Do you feel that your everyday life is difficult now due to your memory decline?"; "Do memory problems make it harder to complete tasks that used to be easy?"), which was not the case for the high self-perceived

cognitive functioning (low concern) range. Together, these results raise the possibility that different types of items might be most relevant for neuropsychologically intact older adults in different stages of SCD. If this is the case, then ultimately it may be possible to select items based on their ability to distinguish between (or capture concerns most associated with) specific levels of SCD. For example, items from the higher concern range, which focused on memory and included various comparisons and consequences or impact on everyday life, might be most suitable for identifying those at risk for future dementia. By contrast, items from the lower concern range, which related mainly to higher order executive cognitive abilities (e.g., carrying out plans, making decisions, shifting set, being distracted, managing medications) or language (e.g., word finding, self-expression), might be more suitable for older adults at lower risk of progression to AD. These various possibilities will need to be

**Figure 3**  
Item Information Plots



*Note.* Figure 3 presents the item information plots for the 601 items included in the analysis. Items presented in blue have the highest information in the ranges of interest ( $-2$  to  $-1$  SDs and  $0$  to  $+1$  SDs), and these are the items presented in Tables 4 and 5. All remaining items are illustrated with gray lines.

investigated through prospective data collection and longitudinal follow-up of neuropsychologically intact older adults.

Another finding was that items from the low self-perceived cognitive functioning (high concern) range tended to have fewer response options. Almost all items (14 of 15) were either dichotomous or had three response options (e.g., *no*, *maybe*, *yes*), with the reverse pattern occurring for the high self-perceived cognitive functioning (low concern) range, where almost all items (13 of 15) utilized 5- and 7-point Likert-type scales. Follow-up research should determine whether fewer response options accurately capture the experiences of neuropsychologically intact older adults with worse self-perceived cognitive functioning (such as those presenting to a memory clinic). By contrast, a wider range of options may be necessary to adequately assess concerns among those with better self-perceived cognitive functioning, as it may allow these individuals to endorse subtle levels of difficulty or change across time. Given that self-perceived cognitive functioning likely falls along a continuum, Likert-type scales with a sufficient number of response options (i.e., 5–7) might be most appropriate to capture variability and change over a broad range of function for this complex construct. There is also research suggesting that rating scales with a small number of response options may be less preferred by respondents (Jones, 1968; Preston & Colman, 2000), including older adults, who in one study perceived dichotomous scales as being too restrictive (Carp, 1989). Ultimately, the choice of response options will be based on multiple study-specific considerations, including but not limited to the purpose of the measurement and characteristics of the target sample (Lozano et al., 2008).

There were no identical items that appeared as top items in both the low and high concern ranges of interest. Despite this, items that captured the ability to remember appointments—from different questionnaires—appeared in both ranges, suggesting that this specific content warrants inclusion in questionnaires targeting older adults with

varying levels of self-perceived cognitive functioning. In addition, prospective memory items (“Remembering what I intended to do”; “Remembering appointments or meetings”) emerged as top items in both self-perceived cognitive functioning ranges. Although most aging studies objectively assess retrospective memory, prospective memory is integral to the maintenance of independent and successful living as it is the cognitive ability that maintains, updates, and monitors our to-do list (Brandimonte et al., 1996). Moreover, failures of prospective memory have important ramifications for older adults’ everyday functioning, emotional well-being, and safety (Kliegel et al., 2016; Mogle et al., 2019; Woods et al., 2012), suggesting a benefit for the inclusion of prospective memory items on questionnaires.

Also, while time referents for top items varied, comparisons to 5 years ago and the present (including current functioning and the past 2 weeks) were the most common and accounted for approximately three quarters of responses, followed by 1 year ago. Only two items referenced more distant time frames—that is, high school or college. Based on these results, it may be advisable to avoid items with referents that target the distant past, which could pose difficulties for older adults trying to recall specific instances of remembering or forgetting and lead to inaccurate or biased reports (Cavanaugh et al., 1998; Robinson & Clore, 2002a, 2002b). Another finding was that approximately one third of the top items captured *ability*, while the majority assessed *change*. Reporting ability requires appraisal of current status only; reporting change requires appraisal of current and past abilities and contrasting them. The predominance of items that focused on change is consistent with the idea that SCD assesses intraindividual change in cognitive functioning (Jessen et al., 2020).

An important issue is whether our results align with proposed SCD plus features—that is, characteristics of SCD presumed to be associated with increased risk of cognitive decline (Jessen et al., 2014, 2020). Consistent with the SCD plus framework, we found an overrepresentation of memory items among the top items, compared to other cognitive domains, particularly in the high concern range. Also, while time referents for top items varied, all but two targeted the last 5 years or a more recent time referent, consistent with the onset of SCD within the last 5 years. A top item related to worry about memory. This corresponds to the SCD plus feature of worry and with research showing that older adults who express concern or worry about their perceived cognitive difficulties have an increased risk of future objective cognitive decline or dementia (Jessen et al., 2010; St John & Montgomery, 2002; van Harten et al., 2018; Verfaillie et al., 2019). Although the present study cannot be used to validate the SCD plus concept, results may be used to inform measurement selection in research aimed at refining or expanding SCD plus criteria by suggesting possible items to measure self-perceived cognitive functioning (inclusion of some iteration of the item “I worry about my memory ability”).

Items identified as psychometrically strong should also be considered in relation to previous research and broader trends. Rabin et al. (2015) previously identified the 10 most frequently occurring item themes on questionnaires used by participating SCD-I working groups. Items that fell within eight of these previously identified themes (*memory change*, *memory for names of people*, *general memory problems*, *losing objects*, *word finding*, *remembering appointments*, *remembering recent conversations*, *memory for intentions/prospective memory*) appeared as top items in the present study. In addition, one of the most frequently occurring self-report items “Compared with one year ago, do you have trouble remembering



things?” was identified as a psychometrically poor item in the current analysis. However, a similar item capturing memory change over the same time interval (“Compared with one year ago, do you feel your memory has declined substantially?”) was identified as a psychometrically strong item. These results highlight the need to reexamine items and consider not only general content and psychometric issues but also nuances in wording and response options when modifying existing questionnaires or designing new ones.

Finally, approximately one third of our identified top items had stems that would typically be viewed as problematic by virtue of having multiple subquestions within a single item (Oppenheim, 1992). Some of these items would be considered multibarreled—that is, the item asks respondents to rate or respond to two or more different issues while allowing for only one response (Bradburn et al., 2004; Menold, 2020), such as “More trouble recalling names, finding the right word, or completing sentences.” Other top items had stems that inquired about a single aspect of cognition but were followed by a series of examples, such as “Compared to one year ago, do you have more difficulty managing money (e.g., paying bills, calculating change, completing tax forms)?” Although survey design researchers typically advise against the inclusion of multibarreled items, which can be confusing or difficult to interpret (DeWalt et al., 2007), they are still frequently used. After analysis of multibarreled items, Menold (2020) concluded that respondents likely attend to the aspect of the item that is most relevant to them, and that these items did not have inferior reliability compared to single-barreled questions (Menold, 2020). Additional research would be required to explore how older adults understand and respond to items that are single- versus multibarreled, and to guide decisions about the inclusion of multibarreled items in questionnaires targeting this population.

### Features of Psychometrically Poor Items

We identified items with poor psychometric properties (Appendix I), which had a few notable features. Some items inquired about issues that may not be highly relevant for most older adults (e.g., “Remembering trivia”). Time referents for some items were very different when compared to those found in psychometrically strong items; two specific time referents, appearing in six of 15 items, required ambiguous comparisons (“than you used to be” and “ever before”). Several items referenced very broad or vague areas that may have been difficult to interpret (e.g., “My ability to recall things that happened during my childhood is”; “Are you worried about these complaints”). One had a vague time referent: “Remembering things that happened a long time ago.” Another item was possibly confusing due to the wording: “When I forget to do something I had planned to do, it is usually not because I forgot what I had to do but because I forgot when I had to do it.” Two items came from a measure with a 9-point Likert-type scale, with response options ranging from –4 to +4, which might overwhelm some older adults. Finally, several items inquired about *complaints* related to the same aspects of cognition that also appear among the top items (language/word finding and attention/concentration), suggesting that older adults may better understand qualifiers such as difficulty, problems, or change (as opposed to complaints) when reporting on their cognitive functioning. Together, these results call attention to the importance of utilizing items with simple, straightforward language, and clear time frames, when assessing older adults (Jobe & Mingay, 1990).

### Strengths and Limitations

The present study has notable strengths. The large sample size was undeniably a strength because it diminished sampling variability effects on the population parameter estimates used in the reference study. We included numerous items covering multiple cognitive domains and participants with varying degrees of cognitive concerns. The application of an IRT model to self-perceived cognitive functioning measurement had several advantages. IRT models allow for a variety of data types to be included (Gibbons et al., 2007) and provide an acceptable trait estimate despite the use of nonidentical items (Hays et al., 2000). Moreover, IRT can keep item parameters constant even when utilizing different samples (Nguyen et al., 2014) and identify the problematic features of poorly performing items thereby helping researchers avoid these characteristics when developing future items and measures (Hays et al., 2000). IRT is also useful in implementing CAT, which has the possibility to reduce participant burden during assessments (Bandalos, 2018; Hays et al., 2000). However, items contained in the item bank for the present study cannot be used directly in a CAT without modification due to the vast heterogeneity in important item features (temporal referents, response options, etc.). In addition, although access to and facility with computers are widespread, this may not always be the case in clinical and research settings where diverse older adults are assessed—and fixed-length short forms, which can show strong precision along a trait continuum (Lai et al., 2011), may be useful alternatives.

Several limitations should be noted. Study analyses were specific to the group of items to which we had access, the samples in which they were tested, and the order in which we chose to proceed through our calibration steps. For example, the present study was anchored using the ECog Self measure in the MCSA. Using another measure may have yielded different results, and thus, we should consider the current results as preliminary. We chose to evaluate data only from the first occasion that each participant responded to the self-perceived cognitive functioning items. This choice ensured that all data would be independent. Some studies provided multiple waves of data, but others only provided a single wave of data. We could have incorporated all of these data with (much) more complicated models, but for this initial work, we focused on the simpler one-observation-per-participant data set.

We carried out a small DIF study for the five studies that used the ECog Self (i.e., in a subset of items that were identified as linking items) to assess how the different populations in the different studies would affect the estimation of the item parameters. Importantly, although our results demonstrated that cross-study DIF was present and could be detected at conventional levels of statistical significance, it was responsible for very small impact at the latent trait level. This lends support to the viability of harmonization across studies. Unfortunately, we were not able to carry out DIF testing across different language studies to ensure measurement invariance (or even linguistic equivalence) across different language versions of the questionnaires prior to merging the data. If the same questionnaires were to be used in different studies in different countries or linguistic groups, the approach to DIF analyses would be relatively straightforward. Also, the need for harmonization and cocalibration would be much simpler or unnecessary (in the absence of DIF). However, when different tests or measures are used across studies, to accomplish the harmonization goal, we must first identify linking items, which we assume are equivalent (and free from DIF) across country and language. For

many cross-study links, the number of linking items was small and not conducive to formal DIF testing.

Relatedly, the geographic and linguistic diversity of our sample, which can be viewed as a strength, also posed challenges in completing the cocalibration, and it placed limits on our ability to use psychometric methods to evaluate the assumptions used to complete the cocalibration activity. In practical terms, while items that were *identical* or similarly worded (deemed *equivalent*) may appear well-suited for analyses, subtleties in target languages and cultural differences can lead to participants interpreting and responding differently (Chan et al., 2015). These are important considerations to bear in mind as we use the current results to inform the creation and validation/norming of new self-perceived cognitive functioning questionnaires for use with culturally and linguistically diverse populations. Despite these challenges, we feel that the geographic and linguistic diversity provided inferential strength in the representativeness and generalizability of our results to the extent that our links were valid. Relatedly, although great effort was made to obtain and include questionnaires and participant data from culturally diverse countries, most of the contributing studies were North American and European; we recognize this as both a limitation and an important future direction. Additionally, diverse study settings were aggregated for the analysis. Given that rates of cognitive concerns are typically lower in population settings (Archer et al., 2015; Jonker et al., 1996), it is possible that conducting the analysis by separating memory clinic and population-based studies may have yielded different results, though IRT methods do not require any particular distribution of the underlying trait for calibration.

Finally, we limited our analyses to neuropsychologically intact participants because of the gradual and progressive lack of awareness of cognitive dysfunction that manifests in the middle to late stages of MCI through AD (Galeone et al., 2011; Vogel et al., 2004; Wolfgruber et al., 2014). We reasoned that the meaning of self-appraisals of cognition in individuals with objective cognitive impairment could be fundamentally different from self-appraisals in those who are neuropsychologically intact—and thus that separate models might be needed to capture the latent trait in these two populations. By focusing only on neuropsychologically intact individuals, we potentially restricted the range of self-perceived cognitive functioning that could be observed, as reflected by several items with very few responses in the category reflecting the greatest degree of self-perceived cognitive impairment. Nevertheless, there was considerable heterogeneity across neuropsychologically intact older adults in self-perceived cognitive functioning present in each study, including many with significant concerns about how their brains are functioning (while they perform well on objective cognitive tests).

### Additional Implications for Scale Construction and Future Directions

An important application of our harmonization effort, and one the field has repeatedly called for, is the utilization of standardized and valid questionnaires (Molinuevo et al., 2017; Röhr et al., 2020). Although our results can be used to inform questionnaire development, it is not possible to create a new questionnaire by simply combining all the items with the smallest *SEM* within particular ranges. Identification of top items does not imply that if taken together, these items would necessarily result in a reliable and valid questionnaire. Items with psychometric support in the present study necessitate independent validation studies to confirm adequate properties. This validation can

use the quality criteria as identified by the Consensus-based Standards for the selection of health Measurement Instruments (COSMIN) approach to quality of measurement (<https://www.cosmin.nl>), which was developed to standardize the evaluation of clinical outcome assessments. In addition, the top items had widely varying features (temporal referents, response option formats), which would make their combination into a single questionnaire problematic from both content and measurement perspectives. Also, these items may not comprehensively assess the self-perceived cognitive functioning construct, and feedback from content experts would be critical. Finally, the unidimensional model may have failed to adequately model non-memory items embedded within questionnaires with predominantly memory items, and future research might revisit and attempt to refine the latent factor using a bifactor model.

In light of these considerations, our immediate next step will be to use our item bank to derive and validate self-perceived cognitive functioning factor scores for existing questionnaires. When validated against relevant outcomes such as objective cognition, biomarkers, and clinical progression, these factor scores will be shared across international studies (Sikkes et al., 2021). After validation of the factor scores, researchers and clinicians from around the globe, who have questionnaire items in common with those from our bank will be able to derive factor scores for their own participants using our item bank. We also plan to use the current item bank for the construction of targeted questionnaires based on various item features and psychometric considerations. Prior to questionnaire development, we plan to carry out focus groups and cognitive interviewing with racial/ethnic and culturally diverse, community-dwelling older adults to enable informed decision-making about item content and key aspects of questionnaire structure/design (e.g., temporal referent, number, and wording of response options, whether and how to capture change, ability, or both). Also, the input from experts in the SCD field, as well as cocreation with participants experiencing low levels of self-perceived cognitive functioning, is necessary to construct a meaningful questionnaire. Finally, we will refine the new self-perceived cognitive functioning questionnaire(s) based on the results of pilot studies, with consideration of reliability and validity evidence.

### Conclusions

This study was the first to employ IRT methods to harmonize self-perceived cognitive functioning data across a large number of international studies. In carrying out the analyses, the authors encountered substantial measurement challenges that placed limits on the extent to which assumptions underlying the harmonization could be tested. An implication is that validity tests need to use external variables to further support the use of the resulting item bank. Another immediate future goal is to use the current results to derive screening measures that enable reliable identification of subtle cognitive changes in clinically normal individuals. Such tools can be utilized in multistudy initiatives and may facilitate large-scale screening efforts and targeting of at-risk populations for early intervention and tracking of treatment-related changes.

### References

- Aalten, P., Ramakers, I. H., Biessels, G. J., de Deyn, P. P., Koek, H. L., OldeRikkert, M. G., Oleksik, A. M., Richard, E., Smits, L. L., van Swieten, J. C., Teune, L. K., van der Lugt, A., Barkhof, F., Teunissen,

- C. E., Rozendaal, N., Verhey, F. R., & van der Flier, W. M. (2014). The Dutch Parnsnoer Institute—Neurodegenerative diseases; methods, design and baseline results. *BMC Neurology*, 14(1), Article 254. <https://doi.org/10.1186/s12883-014-0254-4>
- Abdulrab, K., & Heun, R. (2008). Subjective memory impairment. A review of its definitions indicates the need for a comprehensive set of standardised and validated criteria. *European Psychiatry*, 23(5), 321–330. <https://doi.org/10.1016/j.eurpsy.2008.02.004>
- Amariglio, R. E., Becker, J. A., Carnasin, J., Wadsworth, L. P., Lorus, N., Sullivan, C., Maye, J. E., Gidicsin, C., Pepin, L. C., Sperling, R. A., Johnson, K. A., & Rentz, D. M. (2012). Subjective cognitive complaints and amyloid burden in cognitively normal older individuals. *Neuropsychologia*, 50(12), 2880–2886. <https://doi.org/10.1016/j.neuropsychologia.2012.08.011>
- Amariglio, R. E., Donohue, M. C., Marshall, G. A., Rentz, D. M., Salmon, D. P., Ferris, S. H., Karantoulis, S., Aisen, P. S., Sperling, R. A., & the Alzheimer's Disease Cooperative Study. (2015). Tracking early decline in cognitive function in older individuals at risk for Alzheimer disease dementia: The Alzheimer's Disease Cooperative Study Cognitive Function Instrument. *JAMA Neurology*, 72(4), 446–454. <https://doi.org/10.1001/jama.neurol.2014.3375>
- Amariglio, R. E., Mormino, E. C., Pietras, A. C., Marshall, G. A., Vannini, P., Johnson, K. A., Sperling, R. A., & Rentz, D. M. (2015). Subjective cognitive concerns, amyloid- $\beta$ , and neurodegeneration in clinically normal elderly. *Neurology*, 85(1), 56–62. <https://doi.org/10.1212/WNL.0000000000001712>
- Amariglio, R. E., Sikkes, S. A. M., Marshall, G. A., Buckley, R. F., Gatchel, J. R., Johnson, K. A., Rentz, D. M., Donohue, M. C., Raman, R., Sun, C. K., Yaari, R., Holdridge, K. C., Sims, J. R., Grill, J. D., Aisen, P. S., & Sperling, R. A. (2021). Item-level investigation of participant and study partner report on the Cognitive Function Index from the A4 Study screening data. *The Journal of Prevention of Alzheimer's Disease*, 8(3), 257–262. <https://doi.org/10.14283/jpad.2021.8>
- Archer, H. A., Newson, M. A., & Coulthard, E. J. (2015). Subjective memory complaints: Symptoms and outcome in different research settings. *Journal of Alzheimer's Disease*, 48(Suppl. 1), S109–S114. <https://doi.org/10.3233/JAD-150108>
- Baker, F. B., & Kim, S.-H. (2017). *The basics of item response theory using R*. Springer. <https://doi.org/10.1007/978-3-319-54205-8>
- Bandalos, D. L. (2018). *Measurement theory and applications for the social sciences*. Guilford Press.
- Blessed, G., Tomlinson, B. E., & Roth, M. (1968). The association between quantitative measures of dementia and of senile change in the cerebral grey matter of elderly subjects. *The British Journal of Psychiatry*, 114(512), 797–811. <https://doi.org/10.1192/bjp.114.512.797>
- Bradburn, N. M., Sudman, S., & Wansink, B. (2004). *Asking questions: The definitive guide to questionnaire design: For market research, political polls, and social and health questionnaires*. Jossey-Bass.
- Brandimonte, M. A., Einstein, G. O., & McDaniel, M. A. (1996). *Prospective memory: Theory and applications*. Lawrence Erlbaum.
- Broadbent, D. E., Cooper, P. F., FitzGerald, P., & Parkes, K. R. (1982). The cognitive failures questionnaire (CFQ) and its correlates. *British Journal of Clinical Psychology*, 21(1), 1–16. <https://doi.org/10.1111/j.2044-8260.1982.tb01421.x>
- Brucki, S. M. D., & Nitrini, R. (2009). Subjective memory impairment in a rural population with low education in the Amazon rainforest: An exploratory study. *International Psychogeriatrics*, 21(1), 164–171. <https://doi.org/10.1017/S1041610208008065>
- Buckley, R., Saling, M. M., Ames, D., Rowe, C. C., Lautenschlager, N. T., Macaulay, S. L., Martins, R. N., Masters, C. L., O'Meara, T., Savage, G., Szoek, C., Villemagne, V. L., Ellis, K. A., & the Australian Imaging Biomarkers and Lifestyle Study of Aging (AIBL) Research Group. (2013). Factors affecting subjective memory complaints in the AIBL aging study: Biomarkers, memory, affect, and age. *International Psychogeriatrics*, 25(8), 1307–1315. <https://doi.org/10.1017/S1041610213000665>
- Buckley, R. F., Maruff, P., Ames, D., Bourgeat, P., Martins, R. N., Masters, C. L., Rainey-Smith, S., Lautenschlager, N., Rowe, C. C., Savage, G., Villemagne, V. L., Ellis, K. A., & the AIBL Study. (2016). Subjective memory decline predicts greater rates of clinical progression in preclinical Alzheimer's disease. *Alzheimer's & Dementia*, 12(7), 796–804. <https://doi.org/10.1016/j.jalz.2015.12.013>
- Carp, F. M. (1989). Maximizing data quality in community studies of older people. In M. P. Lawton & A. R. Herzog (Eds.), *Special research methods for gerontology* (pp. 93–122). Baywood Publishing.
- Cavanaugh, J. C., Feldman, J. M., & Hertzog, C. (1998). Memory beliefs as social cognition: A reconceptualization of what memory questionnaires assess. *Review of General Psychology*, 2(1), 48–65. <https://doi.org/10.1037/1089-2680.2.1.48>
- Cella, D., Riley, W., Stone, A., Rothrock, N., Reeve, B., Yount, S., Amtmann, D., Bode, R., Buysse, D., Choi, S., Cook, K., Devellis, R., DeWalt, D., Fries, J. F., Gershon, R., Hahn, E. A., Lai, J.-S., Pilkonis, P., Revicki, D., ... the PROMIS Cooperative Group. (2010). Initial adult health item banks and first wave testing of the Patient-Reported Outcomes Measurement Information System (PROMIS™) Network: 2005–2008. *Journal of Clinical Epidemiology*, 63(11), 1179–1194. <https://doi.org/10.1016/j.jclinepi.2010.04.011>
- Chan, K. S., Gross, A. L., Pezzin, L. E., Brandt, J., & Kasper, J. D. (2015). Harmonizing measures of cognitive performance across international surveys of aging using item response theory. *Journal of Aging and Health*, 27(8), 1392–1414. <https://doi.org/10.1177/0898264315583054>
- Chau, L. T., Lee, J. B., Fleming, J., Roche, N., & Shum, D. (2007). Reliability and normative data for the comprehensive assessment of prospective memory (CAPM). *Neuropsychological Rehabilitation*, 17(6), 707–722. <https://doi.org/10.1080/09602010600923926>
- Chételat, G., Villemagne, V. L., Bourgeat, P., Pike, K. E., Jones, G., Ames, D., Ellis, K. A., Szoek, C., Martins, R. N., O'Keefe, G. J., Salvado, O., Masters, C. L., Rowe, C. C., & the Australian Imaging Biomarkers and Lifestyle Research Group. (2010). Relationship between atrophy and  $\beta$ -amyloid deposition in Alzheimer disease. *Annals of Neurology*, 67(3), 317–324. <https://doi.org/10.1002/ana.21955>
- Colijn, M. A., & Grossberg, G. T. (2015). Amyloid and tau biomarkers in subjective cognitive impairment. *Journal of Alzheimer's Disease*, 47(1), 1–8. <https://doi.org/10.3233/JAD-150180>
- Comijs, H. C., Deeg, D. J. H., Dik, M. G., Twisk, J. W. R., & Jonker, C. (2002). Memory complaints: the association with psycho-affective and health problems and the role of personality characteristics. A 6-year follow-up study. *Journal of Affective Disorders*, 72(2), 157–165. [https://doi.org/10.1016/S0165-0327\(01\)00453-0](https://doi.org/10.1016/S0165-0327(01)00453-0)
- Crook, T. H., III, Feher, E. P., & Larrabee, G. J. (1992). Assessment of memory complaint in age-associated memory impairment: The MAC-Q. *International Psychogeriatrics*, 4(2), 165–176. <https://doi.org/10.1017/S1041610292000991>
- Derby, C. A., Burns, L. C., Wang, C., Katz, M. J., Zimmerman, M. E., L'italien, G., Guo, Z., Berman, R. M., & Lipton, R. B. (2013). Screening for predementia AD: Time-dependent operating characteristics of episodic memory tests. *Neurology*, 80(14), 1307–1314. <https://doi.org/10.1212/WNL.0b013e31828ab2c9>
- DeWalt, D. A., Rothrock, N., Yount, S., Stone, A. A., & the PROMIS Cooperative Group. (2007). Evaluation of item candidates: The PROMIS qualitative item review. *Medical Care*, 45(5, Suppl. 1), S12–S21. <https://doi.org/10.1097/01.mlr.0000254567.79743.e2>
- Dixon, R. A., Hulstsch, D. F., & Hertzog, C. (1988). The metamemory in adulthood (MIA) questionnaire. *Psychopharmacology Bulletin*, 24(4), 671–688.
- Dorans, N. J., Pommerich, M., & Holland, P. W. (Eds.). (2007). *Linking and aligning scores and scales*. Springer Science+Business Media. <https://doi.org/10.1007/978-0-387-49771-6>
- Edelen, M. O., & Reeve, B. B. (2007). Applying item response theory (IRT) modeling to questionnaire development, evaluation, and refinement.



- Quality of Life Research*, 16(Suppl. 1), 5–18. <https://doi.org/10.1007/s11136-007-9198-0>
- Edwards, M. C., Houts, C. R., & Cai, L. (2018). A diagnostic procedure to detect departures from local independence in item response theory models. *Psychological Methods*, 23(1), 138–149. <https://doi.org/10.1037/met0000121>
- Ellis, K. A., Bush, A. I., Darby, D., De Fazio, D., Foster, J., Hudson, P., Lautenschlager, N. T., Lenzo, N., Martins, R. N., Maruff, P., Masters, C., Milner, A., Pike, K., Rowe, C., Savage, G., Szoek, C., Taddei, K., Villemagne, V., Woodward, M., ... the AIBL Research Group. (2009). The Australian Imaging, Biomarkers and Lifestyle (AIBL) study of aging: Methodology and baseline characteristics of 1112 individuals recruited for a longitudinal study of Alzheimer's disease. *International Psychogeriatrics*, 21(4), 672–687. <https://doi.org/10.1017/S1041610209009405>
- Farias, S. T., Mungas, D., Reed, B. R., Cahn-Weiner, D., Jagust, W., Baynes, K., & Decarli, C. (2008). The measurement of everyday cognition (ECog): Scale development and psychometric properties. *Neuropsychology*, 22(4), 531–544. <https://doi.org/10.1037/0894-4105.22.4.531>
- Flake, J. K., & Fried, E. I. (2020). Measurement schmeasurement: Questionable measurement practices and how to avoid them. *Advances in Methods and Practices in Psychological Science*, 3(4), 456–465. <https://doi.org/10.1177/2515245920952393>
- Fried, E. I., & Flake, J. K. (2018). Measurement matters. *Observer*. <https://www.psychologicalscience.org/observer/measurement-matters>
- Galeone, F., Pappalardo, S., Chieffi, S., Iavarone, A., & Carlomagno, S. (2011). Anosognosia for memory deficit in amnesic mild cognitive impairment and Alzheimer's disease. *International Journal of Geriatric Psychiatry*, 26(7), 695–701. <https://doi.org/10.1002/gps.2583>
- Ganguli, M., Chang, C. C. H., Snitz, B. E., Saxton, J. A., Vanderbilt, J., & Lee, C. W. (2010). Prevalence of mild cognitive impairment by multiple classifications: The Monongahela-Youghiogheny Healthy Aging Team (MYHAT) project. *The American Journal of Geriatric Psychiatry*, 18(8), 674–683. <https://doi.org/10.1097/JGP.0b013e3181cdee4f>
- Gelman, A. (2013). Two simple examples for understanding posterior p-values whose distributions are far from uniform. *Electronic Journal of Statistics*, 7, 2595–2602. <https://doi.org/10.1214/13-EJS854>
- Gibbons, L. E., Fredericksen, R., Batey, D. S., Dant, L., Edwards, T. C., Mayer, K. H., Mathews, W. C., Morales, L. S., Mugavero, M. J., Yang, F. M., Paez, E., Kitahata, M. M., Patrick, D. L., Crane, H. M., Crane, P. K., & the Centers for AIDS Research Network of Integrated Clinical Systems (CNICS). (2017). Validity assessment of the PROMIS fatigue domain among people living with HIV. *AIDS Research and Therapy*, 14(1), Article 21. <https://doi.org/10.1186/s12981-017-0146-y>
- Gibbons, R. D., Immekus, J. C., Bock, R. D., & Gibbons, R. D. (2007). The added value of multidimensional IRT models. *Multidimensional and hierarchical modeling monograph* (Vol. 1).
- Gifford, K. A., Liu, D., Lu, Z., Tripodis, Y., Cantwell, N. G., Palmisano, J., Kowall, N., & Jefferson, A. L. (2014). The source of cognitive complaints predicts diagnostic conversion differentially among nondemented older adults. *Alzheimer's & Dementia*, 10(3), 319–327. <https://doi.org/10.1016/j.jalz.2013.02.007>
- Gifford, K. A., Liu, D., Romano, R., III, Jones, R. N., & Jefferson, A. L. (2015). Development of a subjective cognitive decline questionnaire using item response theory: A pilot study. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, 1(4), 429–439. <https://doi.org/10.1016/j.dadm.2015.09.004>
- Gilewski, M. J., Zelinski, E. M., & Schaie, K. W. (1990). The Memory Functioning Questionnaire for assessment of memory complaints in adulthood and old age. *Psychology and Aging*, 5(4), 482–490. <https://doi.org/10.1037/0882-7974.5.4.482>
- Go, R. C. P., Duke, L. W., Harrell, L. E., Cody, H., Bassett, S. S., Folstein, M. F., Albert, M. S., Foster, J. L., Sharrow, N. A., & Blacker, D. (1997). Development and validation of a structured telephone interview for dementia assessment (STIDA): The NIMH Genetics Initiative. *Journal of Geriatric Psychiatry and Neurology*, 10(4), 161–167. <https://doi.org/10.1177/089198879701000407>
- Hambleton, R. K., & Cook, L. L. (1977). Latent trait models and their use in the analysis of educational test data. *Journal of Educational Measurement*, 14(2), 75–94. <https://doi.org/10.1111/j.1745-3984.1977.tb00030.x>
- Hao, L., Wang, X., Zhang, L., Xing, Y., Guo, Q., Hu, X., Mu, B., Chen, Y., Chen, G., Cao, J., Zhi, X., Liu, J., Li, X., Yang, L., Li, J., Du, W., Sun, Y., Wang, T., Liu, Z., ... Han, Y. (2017). Prevalence, risk factors, and complaints screening tool exploration of subjective cognitive decline in a large cohort of the Chinese population. *Journal of Alzheimer's Disease*, 60(2), 371–388. <https://doi.org/10.3233/JAD-170347>
- Haug, A. (2012). *Bayesian estimation and tracking: A practical guide*. Wiley. <https://doi.org/10.1002/9781118287798>
- Hays, R. D., Morales, L. S., & Reise, S. P. (2000). Item response theory and health outcomes measurement in the 21st century. *Medical Care*, 38(9, Suppl.), II28–II42. <https://doi.org/10.1097/00005650-200009002-00007>
- Hill, N. L., Mogle, J., Wion, R., Munoz, E., DePasquale, N., Yevchak, A. M., & Parisi, J. M. (2016). Subjective cognitive impairment and affective symptoms: A systematic review. *The Gerontologist*, 56(6), e109–e127. <https://doi.org/10.1093/geront/gnw091>
- Hollands, S., Lim, Y. Y., Buckley, R., Pietrzak, R. H., Snyder, P. J., Ames, D., Ellis, K. A., Harrington, K., Lautenschlager, N., Martins, R. N., Masters, C. L., Villemagne, V. L., Rowe, C. C., Maruff, P., & the AIBL Research Group. (2015). Amyloid- $\beta$  related memory decline is not associated with subjective or informant rated cognitive impairment in healthy adults. *Journal of Alzheimer's Disease*, 43(2), 677–686. <https://doi.org/10.3233/JAD-140678>
- Hox, J. J. C. M., van de Schoot, R., & Matthijsse, S. (2012). How few countries will do? Comparative survey analysis from a Bayesian perspective. *Survey Research Methods*, 6(2), 87–93. <https://doi.org/10.18148/srm/2012.v6i2.5033>
- Irwin, D. E., Gross, H. E., Stucky, B. D., Thissen, D., DeWitt, E. M., Lai, J. S., Amtmann, D., Khostou, L., Varni, J. W., & DeWalt, D. A. (2012). Development of six PROMIS pediatrics proxy-report item banks. *Health and Quality of Life Outcomes*, 10(1), Article 22. <https://doi.org/10.1186/1477-7525-10-22>
- Jack, C. R., Jr., Bennett, D. A., Blennow, K., Carrillo, M. C., Dunn, B., Haeblerlein, S. B., Holtzman, D. M., Jagust, W., Jessen, F., Karlawish, J., Liu, E., Molinuevo, J. L., Montine, T., Phelps, C., Rankin, K. P., Rowe, C. C., Scheltens, P., Siemers, E., Snyder, H. M., ... Silverberg, N. (2018). NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease. *Alzheimer's & Dementia*, 14(4), 535–562. <https://doi.org/10.1016/j.jalz.2018.02.018>
- Jessen, F., Amariglio, R. E., Buckley, R. F., van der Flier, W. M., Han, Y., Molinuevo, J. L., Rabin, L., Rentz, D. M., Rodriguez-Gomez, O., Saykin, A. J., Sikkes, S. A. M., Smart, C. M., Wolfsgruber, S., & Wagner, M. (2020). The characterisation of subjective cognitive decline. *Lancet Neurology*, 19(3), 271–278. [https://doi.org/10.1016/S1474-4422\(19\)30368-0](https://doi.org/10.1016/S1474-4422(19)30368-0)
- Jessen, F., Amariglio, R. E., van Bostel, M., Breteler, M., Ceccaldi, M., Ch  telat, G., Dubois, B., Dufouil, C., Ellis, K. A., van der Flier, W. M., Glodzik, L., van Harten, A. C., de Leon, M. J., McHugh, P., Mielke, M. M., Molinuevo, J. L., Mosconi, L., Osorio, R. S., Perrotin, A., ... the Subjective Cognitive Decline Initiative (SCD-I) Working Group. (2014). A conceptual framework for research on subjective cognitive decline in preclinical Alzheimer's disease. *Alzheimer's & Dementia*, 10(6), 844–852. <https://doi.org/10.1016/j.jalz.2014.01.001>
- Jessen, F., Wiese, B., Bachmann, C., Eifflander-Gorfer, S., Haller, F., K  lsch, H., Luck, T., M  sch, E., van den Bussche, H., Wagner, M., Wollny, A., Zimmermann, T., Pentzek, M., Riedel-Heller, S. G., Romberg, H. P., Weyerer, S., Kaduszkiewicz, H., Maier, W., Bickel, H., ... the German Study on Aging, Cognition and Dementia in Primary Care Patients Study Group. (2010). Prediction of dementia by subjective memory impairment: Effects of severity and temporal association with cognitive impairment.



- Archives of General Psychiatry*, 67(4), 414–422. <https://doi.org/10.1001/archgenpsychiatry.2010.30>
- Jessen, F., Wiese, B., Cvetanovska, G., Fuchs, A., Kaduszkiewicz, H., Kölsch, H., Luck, T., Mösch, E., Pentzek, M., Riedel-Heller, S. G., Werle, J., Weyerer, S., Zimmermann, T., Maier, W., & Bickel, H. (2007). Patterns of subjective memory impairment in the elderly: Association with memory performance. *Psychological Medicine*, 37(12), 1753–1762. <https://doi.org/10.1017/S0033291707001122>
- Jobe, J. B., & Mingay, D. J. (1990). Cognitive laboratory approach to designing questionnaires for surveys of the elderly. *Public Health Reports*, 105(5), 518–524.
- Johnson, S. C., Kosciak, R. L., Jonaitis, E. M., Clark, L. R., Mueller, K. D., Berman, S. E., Bendlin, B. B., Engelman, C. D., Okonkwo, O. C., Hogan, K. J., Asthana, S., Carlsson, C. M., Hermann, B. P., & Sager, M. A. (2018). The Wisconsin Registry for Alzheimer's Prevention: A review of findings and current directions. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, 10(1), 130–142. <https://doi.org/10.1016/j.dadm.2017.11.007>
- Jones, R. R. (1968). *Differences in response consistency and subjects' preferences for three personality inventory response formats* [Conference session]. Proceedings of the 76th Annual Convention of the American Psychological Association, San Francisco, CA, United States.
- Jonker, C., Geerlings, M. I., & Schmand, B. (2000). Are memory complaints predictive for dementia? A review of clinical and population-based studies. *International Journal of Geriatric Psychiatry*, 15(11), 983–991. [https://doi.org/10.1002/1099-1166\(200011\)15:11<983::AID-GPS238>3.0.CO;2-5](https://doi.org/10.1002/1099-1166(200011)15:11<983::AID-GPS238>3.0.CO;2-5)
- Jonker, C., Launer, L. J., Hooijer, C., & Lindeboom, J. (1996). Memory complaints and memory impairment in older individuals. *Journal of the American Geriatrics Society*, 44(1), 44–49. <https://doi.org/10.1111/j.1532-5415.1996.tb05636.x>
- Jorm, A. F. (1994). A short form of the Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE): Development and cross-validation. *Psychological Medicine*, 24(1), 145–153. <https://doi.org/10.1017/S003329170002691X>
- Jorm, A. F., Christensen, H., Korten, A. E., Jacomb, P. A., & Henderson, A. S. (2001). Memory complaints as a precursor of memory impairment in older people: A longitudinal analysis over 7–8 years. *Psychological Medicine*, 31(3), 441–449. <https://doi.org/10.1017/S0033291701003245>
- Kang, H. A., Su, Y. H., & Chang, H. H. (2018). A note on monotonicity of item response functions for ordered polytomous item response theory models. *British Journal of Mathematical & Statistical Psychology*, 71(3), 523–535. <https://doi.org/10.1111/bmsp.12131>
- Kim, D., De Ayala, R., Ferdous, A. A., & Nering, M. L. (2011). The comparative performance of conditional independence indices. *Applied Psychological Measurement*, 35(6), 447–471. <https://doi.org/10.1177/0146621611407909>
- Kliegel, M., Ballhausen, N., Hering, A., Ihle, A., Schnitzspahn, K. M., & Zuber, S. (2016). Prospective memory in older adults: Where we are now and what is next. *Gerontology*, 62(4), 459–466. <https://doi.org/10.1159/000443698>
- Kryscio, R. J., Abner, E. L., Cooper, G. E., Fardo, D. W., Jicha, G. A., Nelson, P. T., Smith, C. D., Van Eldik, L. J., Wan, L., & Schmitt, F. A. (2014). Self-reported memory complaints: Implications from a longitudinal cohort with autopsies. *Neurology*, 83(15), 1359–1365. <https://doi.org/10.1212/WNL.0000000000000856>
- Lai, J. S., Cella, D., Chang, C. H., Bode, R. K., & Heinemann, A. W. (2003). Item banking to improve, shorten and computerize self-reported fatigue: An illustration of steps to create a core item bank from the FACIT-Fatigue Scale. *Quality of Life Research*, 12(5), 485–501. <https://doi.org/10.1023/A:1025014509626>
- Lai, J. S., Cella, D., Choi, S., Junghaenel, D. U., Christodoulou, C., Gershon, R., & Stone, A. (2011). How item banks and their application can influence measurement practice in rehabilitation medicine: A PROMIS fatigue item bank example. *Archives of Physical Medicine and Rehabilitation*, 92(10, Suppl.), S20–S27. <https://doi.org/10.1016/j.apmr.2010.08.033>
- Lai, J. S., Wagner, L. I., Jacobsen, P. B., & Cella, D. (2014). Self-reported cognitive concerns and abilities: Two sides of one coin? *Psycho-Oncology*, 23(10), 1133–1141. <https://doi.org/10.1002/pon.3522>
- Lee, S.-Y., & Song, X.-Y. (2004). Evaluation of the Bayesian and maximum likelihood approaches in analyzing structural equation models with small sample sizes. *Multivariate Behavioral Research*, 39(4), 653–686. [https://doi.org/10.1207/s15327906mbr3904\\_4](https://doi.org/10.1207/s15327906mbr3904_4)
- Lin, Y., Shan, P. Y., Jiang, W. J., Sheng, C., & Ma, L. (2019). Subjective cognitive decline: Preclinical manifestation of Alzheimer's disease. *Neurological Sciences*, 40(1), 41–49. <https://doi.org/10.1007/s10072-018-3620-y>
- Lista, S., Molinuevo, J. L., Cavedo, E., Rami, L., Amouyel, P., Teipel, S. J., Garaci, F., Toschi, N., Habert, M. O., Blennow, K., Zetterberg, H., O'Bryant, S. E., Johnson, L., Galluzzi, S., Bokde, A. L., Broich, K., Herholz, K., Bakardjian, H., Dubois, B., ... Hampel, H. (2015). Evolving evidence for the value of neuroimaging methods and biological markers in subjects categorized with subjective cognitive decline. *Journal of Alzheimer's Disease*, 48(Suppl. 1), S171–S191. <https://doi.org/10.3233/JAD-150202>
- Lord, F. (1980). *Applications of item response theory to practical testing problems*. Lawrence Erlbaum.
- Lozano, L. M., García-Cueto, E., & Muñoz, J. (2008). Effect of the number of response categories on the reliability and validity of rating scales. *Methodology*, 4(2), 73–79. <https://doi.org/10.1027/1614-2241.4.2.73>
- Lozoya-Delgado, P., Ruiz-Sánchez de León, J. M., & Pedrero-Pérez, E. J. (2012). Validación de un cuestionario de quejas cognitivas para adultos jóvenes: Relación entre las quejas subjetivas de memoria, la sintomatología prefrontal y el estrés percibido [Validation of cognitive complaints questionnaire for young adults: Relationship between subjective memory complaints, prefrontal symptomatology and perceived stress]. *Revista de Neurología*, 54(3), 137–150. <https://doi.org/10.33588/rn.5403.2011283>
- Lucas, R. E., & Baird, B. M. (2006). Global self-assessment. In M. Eid & E. Diener (Eds.), *Handbook of multimethod measurement in psychology* (pp. 29–42). American Psychological Association. <https://doi.org/10.1037/11383-003>
- Luck, T., Luppa, M., Briel, S., Matschinger, H., König, H. H., Bleich, S., Villringer, A., Angermeyer, M. C., & Riedel-Heller, S. G. (2010). Mild cognitive impairment: Incidence and risk factors: Results of the Leipzig longitudinal study of the aged. *Journal of the American Geriatrics Society*, 58(10), 1903–1910. <https://doi.org/10.1111/j.1532-5415.2010.03066.x>
- Margioti, E., Kosmidis, M. H., Yannakoulia, M., Dardiotis, E., Hadjigeorgiou, G., Sakka, P., Ntanasi, E., Vlachos, G. S., & Scarmeas, N. (2020). Exploring the association between subjective cognitive decline and frailty: The Hellenic Longitudinal Investigation of Aging and Diet Study (HELIAD). *Ageing & Mental Health*, 24(1), 137–147. <https://doi.org/10.1080/13607863.2018.1525604>
- McHorney, C. A., & Cohen, A. S. (2000). Equating health status measures with item response theory: Illustrations with functional status items. *Medical Care*, 38(9, Suppl.), II43–II59. <https://doi.org/10.1097/00005650-200009002-00008>
- McNair, D. M., & Kahn, R. J. (1983). Self-assessment of cognitive deficits. In T. Crook, S. Ferris, & R. Bartus (Eds.), *Assessment in geriatric psychopharmacology* (pp. 119–136). Mark Powley.
- Meng, X.-L. (1994). Posterior predictive p-values. *Annals of Statistics*, 22(3), 1142–1160. <https://doi.org/10.1214/aos/1176325622>
- Menold, N. (2020). Double barreled questions: An analysis of the similarity of elements and effects on measurement quality. *Journal of Official Statistics*, 36(4), 855–886. <https://doi.org/10.2478/jos-2020-0041>
- Mielke, M. M., Wiste, H. J., Weigand, S. D., Knopman, D. S., Lowe, V. J., Roberts, R. O., Geda, Y. E., Swenson-Dravis, D. M., Boeve, B. F., Senjem, M. L., Vemuri, P., Petersen, R. C., & Jack, C. R., Jr. (2012). Indicators of amyloid burden in a population-based study of cognitively normal elderly. *Neurology*, 79(15), 1570–1577. <https://doi.org/10.1212/WNL.0b013e31826e2696>

- Mitchell, A. J., Beaumont, H., Ferguson, D., Yadegarfar, M., & Stubbs, B. (2014). Risk of dementia and mild cognitive impairment in older people with subjective memory complaints: Meta-analysis. *Acta Psychiatrica Scandinavica*, 130(6), 439–451. <https://doi.org/10.1111/acps.12336>
- Mogle, J., Muñoz, E., Hill, N. L., Smyth, J. M., & Sliwinski, M. J. (2019). Daily memory lapses in adults: Characterization and influence on affect. *The Journals of Gerontology: Series B*, 74(1), 59–68. <https://doi.org/10.1093/geronb/gbx012>
- Molinuevo, J. L., Rabin, L. A., Amariglio, R., Buckley, R., Dubois, B., Ellis, K. A., Ewers, M., Hampel, H., Klöppel, S., Rami, L., Reisberg, B., Saykin, A. J., Sikkes, S., Smart, C. M., Snitz, B. E., Sperling, R., van der Flier, W. M., Wagner, M., Jessen, F., ... the Subjective Cognitive Decline Initiative (SCD-I) Working Group. (2017). Implementation of subjective cognitive decline criteria in research studies. *Alzheimer's & Dementia*, 13(3), 296–311. <https://doi.org/10.1016/j.jalz.2016.09.012>
- Morris, J. C., Heyman, A., Mohs, R. C., Hughes, J. P., van Belle, G., Fillenbaum, G., Mellits, E. D., & Clark, C. (1989). The consortium to establish a registry for Alzheimer's disease (CERAD): I. Clinical and neuropsychological assessment of Alzheimer's disease. *Neurology*, 39(9), 1159–1165. <https://doi.org/10.1212/WNL.39.9.1159>
- Mosconi, L., De Santi, S., Brys, M., Tsui, W. H., Pirraglia, E., Glodzik-Sobanska, L., Rich, K. E., Switalski, R., Mehta, P. D., Pratico, D., Zinkowski, R., Blennow, K., & de Leon, M. J. (2008). Hypometabolism and altered cerebrospinal fluid markers in normal apolipoprotein E E4 carriers with subjective memory complaints. *Biological Psychiatry*, 63(6), 609–618. <https://doi.org/10.1016/j.biopsych.2007.05.030>
- Muthén, L. K., & Muthén, B. O. (1998–2017). *Mplus user's guide* (8th ed.).
- Nester, C. O., Mogle, J. A., Katz, M. J., Saykin, A. J., Lipton, R. B., & Rabin, L. A. (2022, February). *The Cognitive Change Index (CCI-40): Broadening the assessment of subjective cognitive concerns in demographically diverse community dwelling older adults* [Conference presentation]. International Neuropsychological Society 50th Annual Meeting, New Orleans, LA, United States.
- Nguyen, T. H., Han, H. R., Kim, M. T., & Chan, K. S. (2014). An introduction to item response theory for patient-reported outcome measurement. *Patient*, 7(1), 23–35. <https://doi.org/10.1007/s40271-013-0041-0>
- Nicholas, C. R., Dowling, N. M., Racine, A. M., Clark, L. R., Berman, S. E., Kosciak, R. L., Asthana, S., Hermann, B., Sager, M. A., & Johnson, S. C. (2017). Longitudinal assessment of self- and informant-subjective cognitive complaints in a sample of healthy late-middle aged adults enriched with a family history of Alzheimer's disease. *Journal of the International Neuropsychological Society*, 23(8), 617–626. <https://doi.org/10.1017/S1355617717000509>
- Northwestern University. (2022). *Health measures: Transforming how health is measured*. <https://www.healthmeasures.net/>
- O'Bryant, S. E., Zhang, F., Petersen, M., Hall, J. R., Johnson, L. A., Yaffe, K., Mason, D., Braskie, M., Barber, R. A., Rissman, R. A., Mapstone, M., Mielke, M. M., Toga, A. W., & the HABLE Study Team. (2022). A blood screening tool for detecting mild cognitive impairment and Alzheimer's disease among community-dwelling Mexican Americans and non-Hispanic Whites: A method for increasing representation of diverse populations in clinical research. *Alzheimer's & Dementia*, 18(1), 77–87. <https://doi.org/10.1002/alz.12382>
- Oppenheim, A. N. (1992). *Questionnaire design, interviewing, and attitude measurement*. St. Martin's Press.
- Perrotin, A., Mormino, E. C., Madison, C. M., Hayenga, A. O., & Jagust, W. J. (2012). Subjective cognition and amyloid deposition imaging: A Pittsburgh Compound B positron emission tomography study in normal elderly individuals. *Archives of Neurology*, 69(2), 223–229. <https://doi.org/10.1001/archneurol.2011.666>
- Peter, J., Scheef, L., Abdulkadir, A., Boecker, H., Heneka, M., Wagner, M., Koppara, A., Klöppel, S., Jessen, F., & the Alzheimer's Disease Neuroimaging Initiative. (2014). Gray matter atrophy pattern in elderly with subjective memory impairment. *Alzheimer's & Dementia*, 10(1), 99–108. <https://doi.org/10.1016/j.jalz.2013.05.1764>
- Ponds, R. W. M. H., Commissaris, K. J. A. M., & Jolles, J. (1997). Prevalence and covariates of subjective forgetfulness in a normal population in The Netherlands. *International Journal of Aging & Human Development*, 45(3), 207–221. <https://doi.org/10.2190/MVQ1-WB58-875H-Y4X0>
- Preston, C. C., & Colman, A. M. (2000). Optimal number of response categories in rating scales: Reliability, validity, discriminating power, and respondent preferences. *Acta Psychologica*, 104(1), 1–15. [https://doi.org/10.1016/S0001-6918\(99\)00050-5](https://doi.org/10.1016/S0001-6918(99)00050-5)
- Rabin, L. A., Saykin, A. J., Wishart, H. A., Nutter-Upham, K. E., Flashman, L. A., Pare, N., & Santulli, R. B. (2007). The Memory and Aging Telephone Screen: Development and preliminary validation. *Alzheimer's & Dementia*, 3(2), 109–121. <https://doi.org/10.1016/j.jalz.2007.02.002>
- Rabin, L. A., Smart, C. M., & Amariglio, R. E. (2017). Subjective cognitive decline in preclinical Alzheimer's disease. *Annual Review of Clinical Psychology*, 13(1), 369–396. <https://doi.org/10.1146/annurev-clinpsy-032816-045136>
- Rabin, L. A., Smart, C. M., Crane, P. K., Amariglio, R. E., Berman, L. M., Boada, M., Buckley, R. F., Chételat, G., Dubois, B., Ellis, K. A., Gifford, K. A., Jefferson, A. L., Jessen, F., Katz, M. J., Lipton, R. B., Luck, T., Maruff, P., Mielke, M. M., Molinuevo, J. L., ... Sikkes, S. A. (2015). Subjective cognitive decline in older adults: An overview of self-report measures used across 19 international research studies. *Journal of Alzheimer's Disease*, 48(Suppl. 1), S63–S86. <https://doi.org/10.3233/JAD-150154>
- Raimo, S., Trojano, L., Siciliano, M., Cuoco, S., D'Iorio, A., Santangelo, F., Abbamonte, L., Grossi, D., & Santangelo, G. (2016). Psychometric properties of the Italian version of the multifactorial memory questionnaire for adults and the elderly. *Neurological Sciences*, 37(5), 681–691. <https://doi.org/10.1007/s10072-016-2562-5>
- Raina, P., Wolfson, C., Kirkland, S., Griffith, L. E., Balion, C., Cossette, B., Dionne, I., Hofer, S., Hogan, D., van den Heuvel, E. R., Liu-Ambrose, T., Menec, V., Mugford, G., Patterson, C., Payette, H., Richards, B., Shannon, H., Sheets, D., Taler, V., ... Young, L. (2019). Cohort profile: The Canadian Longitudinal Study on Aging (CLSA). *International Journal of Epidemiology*, 48(6), 1752–1753j. <https://doi.org/10.1093/ije/dyz173>
- Raina, P. S., Wolfson, C., Kirkland, S. A., Griffith, L. E., Oremus, M., Patterson, C., Tuokko, H., Penning, M., Balion, C. M., Hogan, D., Wister, A., Payette, H., Shannon, H., & Brazil, K. (2009). The Canadian Longitudinal Study on Aging (CLSA). *Canadian Journal on Aging*, 28(3), 221–229. <https://doi.org/10.1017/S0714980809990055>
- Rami, L., Mollica, M. A., García-Sánchez, C., Saldaña, J., Sanchez, B., Sala, I., Valls-Pedret, C., Castellví, M., Olives, J., & Molinuevo, J. L. (2014). The subjective cognitive decline questionnaire (SCD-Q): A validation study. *Journal of Alzheimer's Disease*, 41(2), 453–466. <https://doi.org/10.3233/JAD-132027>
- Rattanabannakit, C., Risacher, S. L., Gao, S., Lane, K. A., Brown, S. A., McDonald, B. C., Unverzagt, F. W., Apostolova, L. G., Saykin, A. J., & Farlow, M. R. (2016). The cognitive change index as a measure of self and informant perception of cognitive decline: Relation to neuropsychological tests. *Journal of Alzheimer's Disease*, 51(4), 1145–1155. <https://doi.org/10.3233/JAD-150729>
- R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Reid, L. M., & MacLulich, A. M. J. (2006). Subjective memory complaints and cognitive impairment in older people. *Dementia and Geriatric Cognitive Disorders*, 22(5–6), 471–485. <https://doi.org/10.1159/000096295>
- Reisberg, B., & Gauthier, S. (2008). Current evidence for subjective cognitive impairment (SCI) as the pre-mild cognitive impairment (MCI) stage of subsequently manifest Alzheimer's disease. *International Psychogeriatrics*, 20(1), 1–16. <https://doi.org/10.1017/S1041610207006412>
- Reisberg, B., Pritch, L., Mosconi, L., John, E. R., Glodzik-Sobanska, L., Boksay, I., Monteiro, I., Torossian, C., Vedvyas, A., Ashraf, N., Jamil, I. A., &

- de Leon, M. J. (2008). The pre-mild cognitive impairment, subjective cognitive impairment stage of Alzheimer's disease. *Alzheimer's & Dementia*, 4(Suppl. 1), S98–S108. <https://doi.org/10.1016/j.jalz.2007.11.017>
- Reisberg, B., Shulman, M. B., Torossian, C., Leng, L., & Zhu, W. (2010). Outcome over seven years of healthy adults with and without subjective cognitive impairment. *Alzheimer's & Dementia*, 6(1), 11–24. <https://doi.org/10.1016/j.jalz.2009.10.002>
- Reisberg, B., Torossian, C., Shulman, M. B., Monteiro, I., Boksay, I., Golomb, J., Guillo Benarous, F., Ulysse, A., Oo, T., Vedvyas, A., Rao, J. A., Marsh, K., Kluger, A., Sangha, J., Hassan, M., Alshalabi, M., Arain, F., Shaikh, N., Buj, M., ... Shao, Y. (2019). Two year outcomes, cognitive and behavioral markers of decline in healthy, cognitively normal older persons with global deterioration scale stage 2 (subjective cognitive decline with impairment). *Journal of Alzheimer's Disease*, 67(2), 685–705. <https://doi.org/10.3233/JAD-180341>
- Riedel-Heller, S. G., Matschinger, H., Schork, A., & Angermeyer, M. C. (1999). Do memory complaints indicate the presence of cognitive impairment? Results of a field study. *European Archives of Psychiatry and Clinical Neuroscience*, 249(4), 197–204. <https://doi.org/10.1007/s004060050087>
- Risacher, S. L., Kim, S., Nho, K., Foroud, T., Shen, L., Petersen, R. C., Jack, C. R., Jr., Beckett, L. A., Aisen, P. S., Koeppe, R. A., Jagust, W. J., Shaw, L. M., Trojanowski, J. Q., Weiner, M. W., Saykin, A. J., & the Alzheimer's Disease Neuroimaging Initiative (ADNI). (2015). APOE effect on Alzheimer's disease biomarkers in older adults with significant memory concern. *Alzheimer's & Dementia*, 11(12), 1417–1429. <https://doi.org/10.1016/j.jalz.2015.03.003>
- Robinson, M. D., & Clore, G. L. (2002a). Episodic and semantic knowledge in emotional self-report: Evidence for two judgment processes. *Journal of Personality and Social Psychology*, 83(1), 198–215. <https://doi.org/10.1037/0022-3514.83.1.198>
- Robinson, M. D., & Clore, G. L. (2002b). Belief and feeling: Evidence for an accessibility model of emotional self-report. *Psychological Bulletin*, 128(6), 934–960. <https://doi.org/10.1037/0033-2909.128.6.934>
- Röhr, S., Pabst, A., Riedel-Heller, S. G., Jessen, F., Turana, Y., Handajani, Y. S., Brayne, C., Matthews, F. E., Stephan, B. C. M., Lipton, R. B., Katz, M. J., Wang, C., Guerchet, M., Preux, P. M., Mbelesso, P., Ritchie, K., Ancelin, M. L., Carrière, I., Guaita, A., ... the for Cohort Studies of Memory in an International Consortium (COSMIC). (2020). Estimating prevalence of subjective cognitive decline in and across international cohort studies of aging: A COSMIC study. *Alzheimer's Research & Therapy*, 12(1), Article 167. <https://doi.org/10.1186/s13195-020-00734-y>
- Roth, R. M., Gioia, G. A., & Isquith, P. K. (2005). *BRIEF-A: Behavior Rating Inventory of Executive Function—Adult Version*. Psychological Assessment Resources.
- Şahin, S., Yüksel, N., Utku, Ç., Bodur, N. E., Karapınar, Ö. K., Birer, N. Ç., & Kaya, D. (2013). Bellek İşlevselliği Anketi'nin Türkçe Geçerlik ve Güvenilirlik Çalışması [Validity and Reliability of Turkish version of the Memory Functioning Questionnaire]. *Klinik Psikiyatri*, 16, 135–147.
- Samejima, F. (1997). Graded response model. In W. van der Linden & R. Hambleton (Eds.), *Handbook of modern item response theory* (pp. 85–100). Springer. [https://doi.org/10.1007/978-1-4757-2691-6\\_5](https://doi.org/10.1007/978-1-4757-2691-6_5)
- Saykin, A. J. (n.d.). *Activities of Daily Living Rating Scale Self-Version, Abbreviated*.
- Saykin, A. J. (1992). *Neurobehavioral function and activities of daily living rating scale (NBFADL-63 item version)*. Dartmouth Medical School.
- Saykin, A. J., Wishart, H. A., Rabin, L. A., Santulli, R. B., Flashman, L. A., West, J. D., McHugh, T. L., & Mamourian, A. C. (2006). Older adults with cognitive complaints show brain atrophy similar to that of amnesic MCI. *Neurology*, 67(5), 834–842. <https://doi.org/10.1212/01.wnl.0000234032.77541.a2>
- Scheef, L., Spottke, A., Daerr, M., Joe, A., Striepen, N., Kölsch, H., Popp, J., Daamen, M., Gorris, D., Heneka, M. T., Boecker, H., Biersack, H. J., Maier, W., Schild, H. H., Wagner, M., & Jessen, F. (2012). Glucose metabolism, gray matter structure, and memory decline in subjective memory impairment. *Neurology*, 79(13), 1332–1339. <https://doi.org/10.1212/WNL.0b013e31826c1a8d>
- Schultz, S. A., Oh, J. M., Kosciak, R. L., Dowling, N. M., Gallagher, C. L., Carlsson, C. M., Bendlin, B. B., LaRue, A., Hermann, B. P., Rowley, H. A., Asthana, S., Sager, M. A., Johnson, S. C., & Okonkwo, O. C. (2015). Subjective memory complaints, cortical thinning, and cognitive dysfunction in middle-aged adults at risk for AD. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, 1(1), 33–40. <https://doi.org/10.1016/j.da.2014.11.010>
- Sheikh, J. I., & Yesavage, J. A. (1986). Geriatric Depression Scale (GDS): Recent evidence and development of a shorter version. *Clinical Gerontologist: The Journal of Aging and Mental Health*, 5(1–2), 165–173. [https://doi.org/10.1300/J018v05n01\\_09](https://doi.org/10.1300/J018v05n01_09)
- Sikkies, S. A. M., Dubbelman, M. A., Tommet, D., Jones, R. N., Crane, P. K., & Rabin, L. (2021, July). *The relationship between subjective cognitive decline and demographic characteristics is moderated by study setting: Findings from the SCD-1 item analysis working group* [Conference presentation]. 2021 Alzheimer's Association International Conference, Virtual Conference. <https://doi.org/10.1002/alz.056317>
- Slavin, M. J., Brodaty, H., Kochan, N. A., Crawford, J. D., Trollor, J. N., Draper, B., & Sachdev, P. S. (2010). Prevalence and predictors of “subjective cognitive complaints” in the Sydney Memory and Ageing Study. *The American Journal of Geriatric Psychiatry*, 18(8), 701–710. <https://doi.org/10.1097/JGP.0b013e3181df49fb>
- Slot, R. E. R., Sikkies, S. A. M., Berkhof, J., Brodaty, H., Buckley, R., Cavado, E., Dardiotis, E., Guillo-Benarous, F., Hampel, H., Kochan, N. A., Lista, S., Luck, T., Maruff, P., Molinuevo, J. L., Kornhuber, J., Reisberg, B., Riedel-Heller, S. G., Risacher, S. L., Roehr, S., ... van der Flier, W. M. (2019). Subjective cognitive decline and rates of incident Alzheimer's disease and non-Alzheimer's disease dementia. *Alzheimer's & Dementia*, 15(3), 465–476. <https://doi.org/10.1016/j.jalz.2018.10.003>
- Slot, R. E. R., Verfaillie, S. C. J., Overbeek, J. M., Timmers, T., Wesselman, L. M. P., Teunissen, C. E., Dols, A., Bouwman, F. H., Prins, N. D., Barkhof, F., Lammertsma, A. A., Van Berckel, B. N. M., Scheltens, P., Sikkies, S. A. M., & Van der Flier, W. M. (2018). Subjective Cognitive Impairment Cohort (SCIENCe): Study design and first results. *Alzheimer's Research & Therapy*, 10(1), Article 76. <https://doi.org/10.1186/s13195-018-0390-y>
- Smart, C. M., Segalowitz, S. J., Mulligan, B. P., & MacDonald, S. W. (2014). Attention capacity and self-report of subjective cognitive decline: A P3 ERP study. *Biological Psychology*, 103, 144–151. <https://doi.org/10.1016/j.bio.2014.08.016>
- Snitz, B. E., Wang, T., Cloonan, Y. K., Jacobsen, E., Chang, C. H., Hughes, T. F., Kamboh, M. I., & Ganguli, M. (2018). Risk of progression from subjective cognitive decline to mild cognitive impairment: The role of study setting. *Alzheimer's & Dementia*, 14(6), 734–742. <https://doi.org/10.1016/j.jalz.2017.12.003>
- Snitz, B. E., Yu, L., Crane, P. K., Chang, C. C. H., Hughes, T. F., & Ganguli, M. (2012). Subjective cognitive complaints of older adults at the population level: An item response theory analysis. *Alzheimer Disease and Associated Disorders*, 26(4), 344–351. <https://doi.org/10.1097/WAD.0b013e3182420bdf>
- Squire, L. R., Wetzel, C. D., & Slater, P. C. (1979). Memory complaint after electroconvulsive therapy: Assessment with a new self-rating instrument. *Biological Psychiatry*, 14(5), 791–801.
- St John, P., & Montgomery, P. (2002). Are cognitively intact seniors with subjective memory loss more likely to develop dementia? *International Journal of Geriatric Psychiatry*, 17(9), 814–820. <https://doi.org/10.1002/gps.559>
- Sun, Y., Yang, F. C., Lin, C. P., & Han, Y. (2015). Biochemical and neuroimaging studies in subjective cognitive decline: Progress and perspectives. *CNS Neuroscience & Therapeutics*, 21(10), 768–775. <https://doi.org/10.1111/cns.12395>
- Troyer, A. K., & Rich, J. B. (2002). Psychometric properties of a new metamemory questionnaire for older adults. *The Journals of Gerontology: Series B*, 57(1), P19–P27. <https://doi.org/10.1093/geronb/57.1.P19>



- Vale, C. D. (1986). Linking item parameters onto a common scale. *Applied Psychological Measurement*, 10(4), 333–344. <https://doi.org/10.1177/014662168601000402>
- van der Flier, W. M., & Scheltens, P. (2018). Amsterdam Dementia Cohort: Performing research to optimize care. *Journal of Alzheimer's Disease*, 62(3), 1091–1111. <https://doi.org/10.3233/JAD-170850>
- van Harten, A. C., Mielke, M. M., Swenson-Dravis, D. M., Hagen, C. E., Edwards, K. K., Roberts, R. O., Geda, Y. E., Knopman, D. S., & Petersen, R. C. (2018). Subjective cognitive decline and risk of MCI: The Mayo Clinic Study of Aging. *Neurology*, 91(4), e300–e312. <https://doi.org/10.1212/WNL.0000000000005863>
- van Harten, A. C., Visser, P. J., Pijnenburg, Y. A., Teunissen, C. E., Blankenstein, M. A., Scheltens, P., & van der Flier, W. M. (2013). Cerebrospinal fluid A $\beta$ 42 is the best predictor of clinical progression in patients with subjective complaints. *Alzheimer's & Dementia*, 9(5), 481–487. <https://doi.org/10.1016/j.jalz.2012.08.004>
- Verfaillie, S. C. J., Timmers, T., Slot, R. E. R., van der Weijden, C. W. J., Wesselman, L. M. P., Prins, N. D., Sikkes, S. A. M., Yaqub, M., Dols, A., Lammertsma, A. A., Scheltens, P., Ossenkoppele, R., van Berckel, B. N. M., & van der Flier, W. M. (2019). Amyloid- $\beta$  load is related to worries, but not to severity of cognitive complaints in individuals with subjective cognitive decline: The SCIENCe project. *Frontiers in Aging Neuroscience*, 11, Article 7. <https://doi.org/10.3389/fnagi.2019.00007>
- Visser, P. J., Verhey, F., Knol, D. L., Scheltens, P., Wahlund, L. O., Freund-Levi, Y., Tsolaki, M., Minthon, L., Wallin, A. K., Hampel, H., Bürger, K., Pirttilä, T., Soininen, H., Rikkert, M. O., Verbeek, M. M., Spira, L., & Blennow, K. (2009). Prevalence and prognostic value of CSF markers of Alzheimer's disease pathology in patients with subjective cognitive impairment or mild cognitive impairment in the DESCRIPA study: A prospective cohort study. *Lancet Neurology*, 8(7), 619–627. [https://doi.org/10.1016/S1474-4422\(09\)70139-5](https://doi.org/10.1016/S1474-4422(09)70139-5)
- Vlachos, G. S., Cosentino, S., Kosmidis, M. H., Anastasiou, C. A., Yannakouli, M., Dardiotis, E., Hadjigeorgiou, G., Sakka, P., Ntanas, E., & Scarmeas, N. (2019). Prevalence and determinants of subjective cognitive decline in a representative Greek elderly population. *International Journal of Geriatric Psychiatry*, 34(6), 846–854. <https://doi.org/10.1002/gps.5073>
- Vogel, A., Stokholm, J., Gade, A., Andersen, B. B., Hejl, A. M., & Waldemar, G. (2004). Awareness of deficits in mild cognitive impairment and Alzheimer's disease: Do MCI patients have impaired insight? *Dementia and Geriatric Cognitive Disorders*, 17(3), 181–187. <https://doi.org/10.1159/000076354>
- Winkler, R. L. (1993). Bayesian statistics: An overview. In G. Keren & C. Lewis (Eds.), *A handbook for data analysis in the behavioral sciences: Statistical issues* (pp. 201–232). Lawrence Erlbaum.
- Wolfsgruber, S., Wagner, M., Schmidtke, K., Frölich, L., Kurz, A., Schulz, S., Hampel, H., Heuser, I., Peters, O., Reischies, F. M., Jahn, H., Luckhaus, C., Hüll, M., Gertz, H. J., Schröder, J., Pantel, J., Rienhoff, O., Rütger, E., Henn, F., ... Jessen, F. (2014). Memory concerns, memory performance and risk of dementia in patients with mild cognitive impairment. *PLOS ONE*, 9(7), Article e100812. <https://doi.org/10.1371/journal.pone.0100812>
- Woods, S. P., Weinborn, M., Velnoweth, A., Rooney, A., & Bucks, R. S. (2012). Memory for intentions is uniquely associated with instrumental activities of daily living in healthy older adults. *Journal of the International Neuropsychological Society*, 18(1), 134–138. <https://doi.org/10.1017/S1355617711001263>
- Yesavage, J. A., Brink, T. L., Rose, T. L., Lum, O., Huang, V., Adey, M., & Leirer, V. O. (1982–1983). Development and validation of a geriatric depression screening scale: A preliminary report. *Journal of Psychiatric Research*, 17(1), 37–49. [https://doi.org/10.1016/0022-3956\(82\)90033-4](https://doi.org/10.1016/0022-3956(82)90033-4)
- Youn, J. C., Kim, K. W., Lee, D. Y., Jhoo, J. H., Lee, S. B., Park, J. H., Choi, E. A., Choe, J. Y., Jeong, J. W., Choo, I. H., & Woo, J. I. (2009). Development of the Subjective Memory Complaints Questionnaire. *Dementia and Geriatric Cognitive Disorders*, 27(4), 310–317. <https://doi.org/10.1159/000205512>
- Zlatar, Z. Z., Moore, R. C., Palmer, B. W., Thompson, W. K., & Jeste, D. V. (2014). Cognitive complaints correlate with depression rather than concurrent objective cognitive impairment in the successful aging evaluation baseline sample. *Journal of Geriatric Psychiatry and Neurology*, 27(3), 181–187. <https://doi.org/10.1177/0891988714524628>
- Zwan, M. D., Villemagne, V. L., Doré, V., Buckley, R., Bourgeat, P., Veljanoski, R., Salvado, O., Williams, R., Margison, L., Rembach, A., Macaulay, S. L., Martins, R., Ames, D., van der Flier, W. M., Ellis, K. A., Scheltens, P., Masters, C. L., & Rowe, C. C. (2016). Subjective memory complaints in APOE  $\epsilon$ 4 carriers are associated with high amyloid- $\beta$  burden. *Journal of Alzheimer's Disease*, 49(4), 1115–1122. <https://doi.org/10.3233/JAD-150446>

## Appendix A

### Definitions of “Neuropsychologically Intact” by Participating Studies

Study name	Criteria for defining “neuropsychologically intact”
Alzheimer's Disease Center Clinical Core and Center for Brain Health (ADC-NYU)	No objective evidence of memory deficit in clinical interview. Performed normally on the objective portion of the Brief Cognitive Rating Scale, which includes tests of concentration and calculation, memory, orientation, and functional abilities, carried out during the clinical interview.
Alzheimer's Disease Neuroimaging Initiative (ADNI)	Normal subjects: Clinical Dementia Rating (CDR) score = 0; Mini-Mental State Exam (MMSE) range 24–30, delayed recall one paragraph from the Logical Memory II subscale of the Wechsler Memory Scale–Revised (maximum score of 25) with cutoffs as follows: $\geq 9$ for 16 years of education, $\geq 5$ for 8–15 years of education, $\geq 3$ for 0–7 years of education.
Amsterdam Dementia Cohort (Ams Dem Cohort)	No objective impairment as judged by a multidisciplinary team in the university memory clinic, based (amongst others) on the neuropsychological evaluation.
Anti-Amyloid Treatment in Asymptomatic Alzheimer's Study (A4)	Cognitively unimpaired defined by scores $>25$ on the MMSE and CDR = 0.
Australian Imaging Biomarkers and Lifestyle Flagship Study of Ageing (AIBL)	MMSE scores 27–30; normal performance on Logical Memory; no evidence of significant difficulty on standardized neuropsychological tests of memory, language, attention, executive functioning, and psychomotor speed; CDR = 0.
Barcelona Group (Barcelona)	Healthy control participants were volunteers presenting no cognitive decline and normal scores on two screening tests: MMSE and Memory Alteration Test.

(Appendices continue)



## Appendix A (continued)

Study name	Criteria for defining “neuropsychologically intact”
Bonn Memory Clinic (Bonn)	Also performed in the normal range for age and education on standardized tests of global cognition, verbal and visual episodic memory, visuospatial functioning, language, verbal fluency, working memory, and executive functions. No cognitive deficit: no score greater than 1.5 <i>SD</i> below age- and education-adjusted norms on at least one subtest of the Consortium to Establish a Registry for Alzheimer’s Disease Neuropsychological battery.
Canadian Longitudinal Study on Aging	Participants were defined as cognitively unimpaired if trained interviewers determined at baseline assessment that they were able to understand the purpose of the study and could provide reliable data. Participants who self-reported at Follow-Up 1 that they had received a formal diagnosis of memory problems or dementia by a physician were excluded.
Dartmouth Memory and Aging Study/Indiana Memory and Aging Study (Dart-Indiana)	Objective cognition: scores within 1.5 <i>SDs</i> of the mean established for age- and education-matched controls on standardized neuropsychological testing of memory, attention, executive function, language, spatial ability, general intellectual functioning, psychomotor speed, and standard dementia screens.
Einstein Aging Study (EAS) and substudies	No score greater than 1.5 <i>SD</i> below the age-adjusted mean on neuropsychological tests of memory, attention, executive function, visuospatial ability, or language. Tests included: Free and Cued Selective Reminding Test, Logical Memory I from the Wechsler Memory Scale–Revised, Trail Making Test, Digit Span, Letter Fluency, Block Design, Digit Symbol, Category Fluency task (animals, vegetables, and fruits), and the Boston Naming Test.
German Study on Ageing, Cognition, and Dementia in Primary Care Patients (AgeCoDe)	Performed no greater than 1.0 <i>SD</i> below the normative domain scores on the neuropsychological test battery of the Structured Interview for Diagnosis of Dementia of Alzheimer type, Multiinfarct dementia, and dementia of other etiology according to Diagnostic Statistical Manual-III-Revised, Diagnostic Statistical Manual-IV, and <i>International Classification of Diseases-10</i> (SIDAM).
Harvard Aging Brain Study (HABS)	Cognitively unimpaired: global CDR = 0; MMSE $\geq 27$ with educational adjustment (for low education, scores $\geq 25$ ); and performance in the normal range within education-adjusted norms on the Logical Memory II Delayed Recall Index from the Wechsler Memory Scale.
Health and Aging Brain Study, Health Disparities (HABS-HD)	CDR sum of boxes score = 0 and cognitive test scores broadly within normal limits as defined by performance of no more than 1.5 <i>SDs</i> below the mean of the normative range on any neuropsychological test.
Hellenic Longitudinal Investigation of Aging and Diet (HELIAD)	No objective impairment on any cognitive domain ( $>1.5$ <i>SD</i> below age and education expected levels) as part of an extensive neuropsychological battery covering global cognition, verbal and visual memory, visuospatial functioning, language, attention speed, and executive functions.
Institut des Matériaux Poreux de Paris Caen Group	Performed in the normal range for age and education on standardized tests of global cognition, verbal and visual episodic memory, visuospatial functioning, language, verbal fluency, working memory, and executive functions.
Indiana Alzheimer’s Disease Center Cohort (IADC)	Without a measurable cognitive deficit.
Leipzig Longitudinal Study of the Aged (LEILA 75+)	Performed no greater than 1.0 <i>SD</i> below the normative domain scores on the neuropsychological test battery of the Structured Interview for Diagnosis of Dementia of Alzheimer type, multiinfarct dementia, and dementia of other etiology according to <i>DSM-III-R</i> , <i>DSM-IV</i> , and <i>ICD-10</i> (SIDAM).
Mayo Clinic Study of Aging (MCSA)	Cognitively unimpaired participants were cognitively and functionally normal as determined by a consensus panel based on clinical and neuropsychological assessment.
Memory Clinic—Fundació Alzheimer Centre Educacional (Fundació)	Performed in the normal range for age and education on the standardized Fundació Alzheimer Centre Educacional Neuropsychological Battery.
Sydney Memory and Ageing Study (Sydney)	Classified as cognitively normal if performance on all test measures was above the 6.68 percentile ( $-1.5$ <i>SDs</i> ) or equivalent score compared to normative published values, controlling for age, sex, and education.
University of Pittsburgh Study/Monongahela-Youghiogheny Health Aging Team (MYHAT); Pittsburgh (Pitt) Substudy	Cognitively normal: composite scores in all domains within 1.0 <i>SD</i> of the mean for the individual’s age–sex–education group. Domains included attention/processing speed, executive function, language, memory, and visuospatial function—with a composite score for each domain.
Vanderbilt Memory and Alzheimer’s Center (VMAC)	National Alzheimer’s Coordinating Center diagnostic criteria: cognitively unimpaired on neuropsychological tests covering 3 cognitive domains, including memory (Logical Memory I and II, Neuropsychological Assessment Battery List Learning), executive function (Trail Making Test parts A, B, Digits Backward), and language (Animal Fluency Test, Boston Naming Test).
Victoria Subjective Cognitive Decline Study (Victoria SCDS)	Cognitively intact defined as scored $>136$ on the Dementia Rating Scale-2 Total Score; no score less than $-1.5$ <i>SD</i> on the delayed recall trials of the California Verbal Learning Test-2; scored $>25/30$ on the MMSE-2.
Wisconsin Registry for Alzheimer’s Prevention (WRAP)	Cognitively unimpaired determined in one of two ways as described in <a href="#">Johnson et al. (2018)</a> : (a) perform within 2 <i>SD</i> on all cognitive tests relative to robust norms accounting for age, sex, Wide Range Achievement Test 3 reading and (b) individuals were flagged for committee review but determined to be unimpaired after considering all available information.

*Note.* Neuropsychologically intact includes individuals with varying levels of self-perceived cognitive functioning difficulties.

(Appendices continue)

## Appendix B

### Sample Equivalent Items

Set	Measure	Study, language	Item	Response options	Time frame
1	Multifactorial Memory Questionnaire (MMQ)	Bonn Memory Clinic, German	A: Forget a telephone number you use frequently	0 = always; 1 = often; 2 = sometimes; 3 = <i>rarely</i> ; 4 = <i>never</i>	Previous 2 weeks
	Cognitive Difficulties Scale (CDS-Q)	Institut des Matériaux Poreux de Paris Caen Group, French	B: Trouble recalling frequently used phone numbers	0 = <i>never</i> ; 1 = <i>rarely</i> ; 2 = sometimes; 3 = <i>often</i> ; 4 = <i>very often</i>	Lately
2	Structured Telephone Dementia Assessment (STDA)	Harvard Brain Aging Study, English	C: Do you have trouble remembering things from 1 s to the next?	0 = no; 1 = yes	Current
	VMAC Cognitive Complaint Questionnaire (VMAC CCQ)	Vanderbilt Memory and Alzheimer's Center, English	D: Do you have trouble remembering things from one moment to the next?	0 = no; 1 = yes	Current

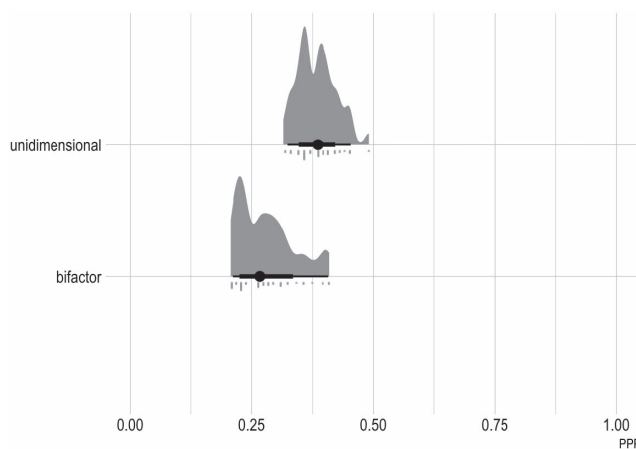
*Note.* The italicized response options (*rarely*, *never*) for Items A and B in Set 1 were combined for purposes of data harmonization due to the categories having fewer than five respondents (see Data Recoded). In addition, italicized response options (*often*, *very often*) for Item B in Set 1 were combined to establish equivalent response categories (see Item Overlap). VMAC = Vanderbilt Memory and Alzheimer's Center.

## Appendix C

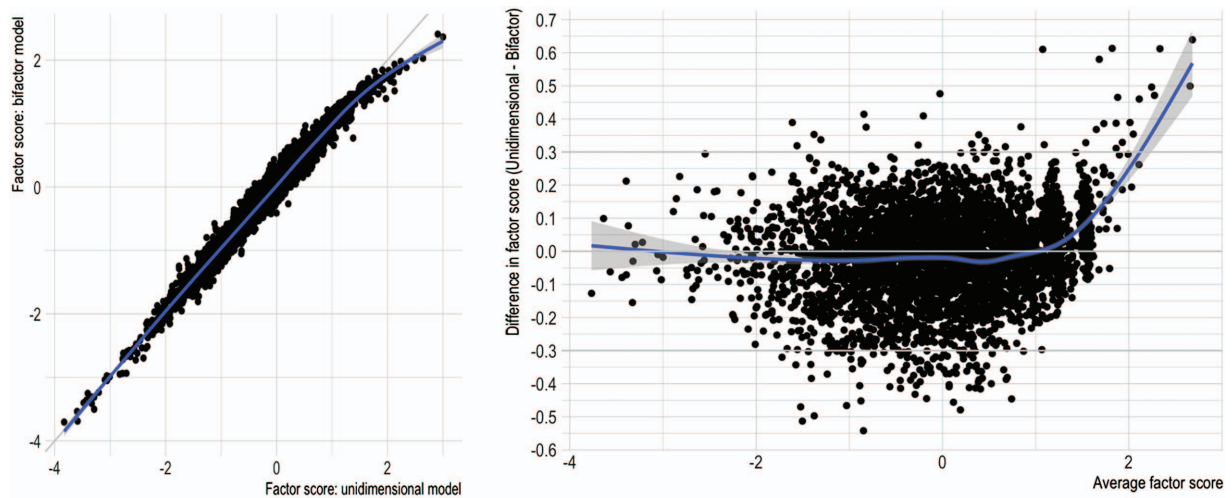
### Distribution of Posterior Predictive $P$ Values for all Items and Estimated Factor Score Estimates

**Figure C1**

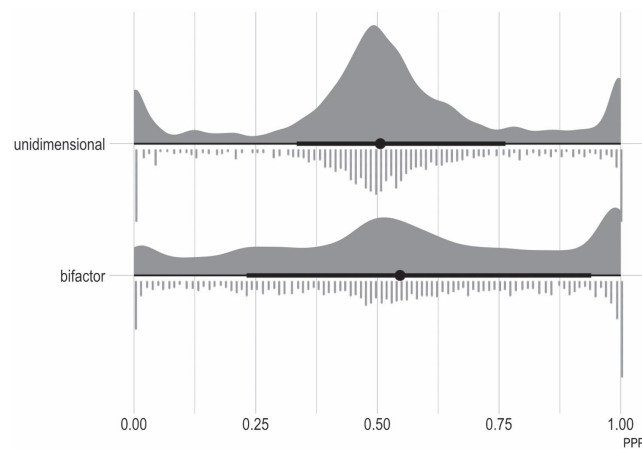
*Raincloud Plot of Posterior Predictive  $P$  Values (PPP) for Items the Mayo Clinic Study of Aging (MCSA) From a Unidimensional and Bifactor Model*



(Appendices continue)

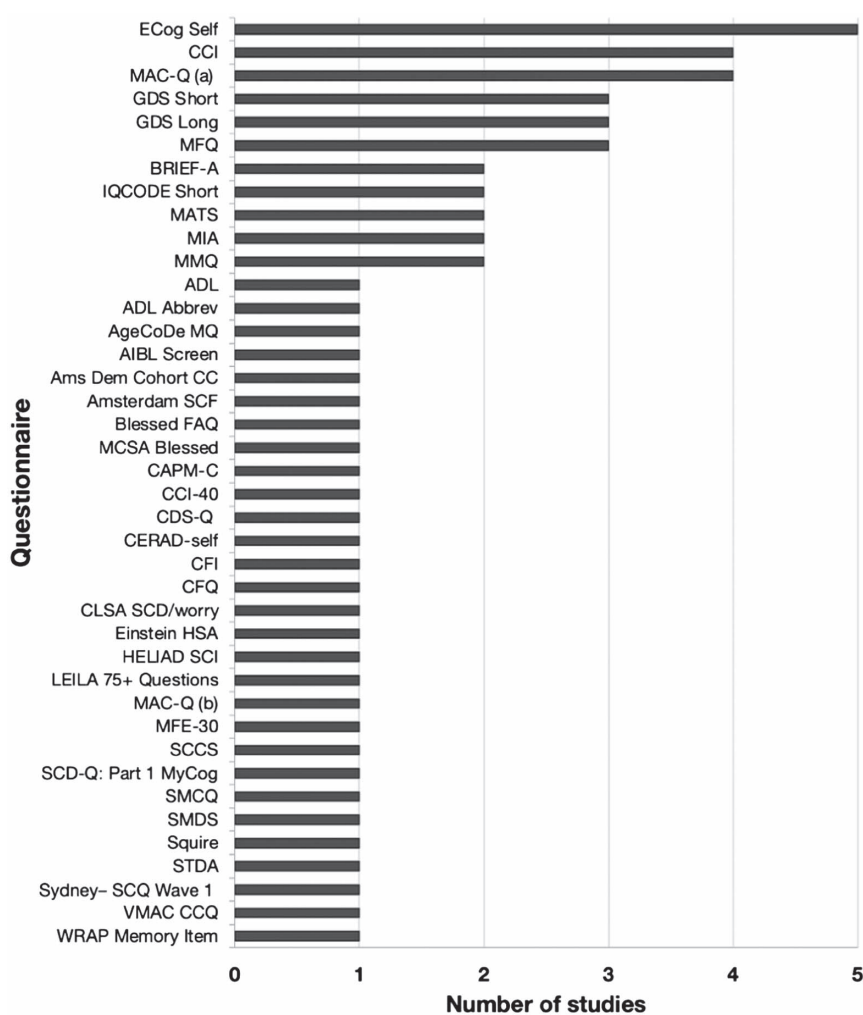
**Figure C2***Examining Estimated Factor Scores From a Unidimensional and a Priori Specified Bifactor Model*

*Note.* Left: Scatterplot of factor score estimates under bifactor (y-axis) and unidimensional model (x-axis). Right: Scatterplot of differences in factor score estimates as a function of the average of the two-factor scores.

**Figure C3***Raincloud Plot of Posterior Predictive P Values (PPP) for All Items From All Studies From a Unidimensional and Bifactor Model**(Appendices continue)*

### Appendix D

#### Number of Studies in Which Each of the 40 Study Questionnaires Was Utilized



*Note.* The total number of studies = 24 (substudies are not included); please refer to Table 2, for the definition of questionnaire abbreviations.

*(Appendices continue)*



Appendix E  
Number of Linking Items Between Studies

Study	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
1. ADC-NYU	4	—	—	4	4	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
2. ADNI	—	60	20	—	—	—	—	—	59	—	—	20	—	—	39	—	—	60	—	40	—	—	—	—	—	—	4
3. Ans Dem Cohort	—	20	27	—	—	—	—	—	19	—	—	20	—	—	1	—	—	20	—	—	—	—	—	1	—	—	1
4. A4	4	—	—	18	5	—	—	—	1	1	1	—	—	1	—	—	—	—	—	—	—	—	—	—	1	6	—
5. AIBL	4	—	—	5	6	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	—	—	1	5	1
6. Barcelona	—	—	—	—	—	26	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
7. Bonn	—	—	—	—	—	1	32	20	—	—	—	—	—	—	—	—	3	—	—	—	1	—	—	—	—	—	—
8. CLSA	—	—	—	—	—	—	20	21	—	—	—	1	1	—	—	—	2	—	—	—	1	—	—	—	—	—	—
9. Dart-Indiana	—	59	19	1	—	—	—	—	173	1	33	19	4	4	39	1	—	59	—	40	—	15	—	—	—	1	16
10. EAS	—	—	—	1	—	—	—	—	1	17	4	—	—	1	—	—	—	—	—	—	—	—	—	—	2	1	—
11. EAS Substudy 1	—	—	—	1	—	—	—	—	33	4	121	—	—	1	—	—	—	—	—	—	—	—	—	—	1	33	—
12. EAS Substudy 2	—	20	20	—	—	—	—	1	19	—	—	42	1	—	—	—	—	20	—	—	—	—	—	—	—	—	—
13. AgeCoDe	—	—	—	—	—	—	—	1	—	—	—	1	5	1	1	—	—	—	—	—	—	—	—	—	2	—	—
14. HABS-HD	—	—	—	1	—	—	—	—	4	1	1	—	1	18	1	1	—	—	—	—	—	—	—	—	11	4	—
15. HABS	—	39	1	—	—	—	—	—	39	—	—	—	1	1	93	—	—	39	—	39	—	—	—	48	2	—	34
16. HELIAD	—	—	—	1	—	—	—	—	—	1	1	—	—	—	—	15	—	—	—	—	—	—	—	—	2	1	—
17. IMAP Caen Group	—	—	—	—	—	—	3	3	—	—	—	—	—	—	—	—	32	—	—	—	2	—	—	—	—	—	—
18. IADC	—	60	20	—	—	—	—	59	—	—	—	20	—	—	39	—	—	60	—	40	—	—	—	—	—	—	—
19. LEILA75+	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	7	—	—	1	—	—	1	—	1
20. MCSA	—	40	—	—	—	—	—	40	—	—	—	—	—	—	39	—	—	40	—	45	—	—	—	—	—	—	—
21. Fundació	—	—	—	—	—	—	1	1	—	—	—	—	—	—	—	—	2	—	—	—	28	—	—	—	—	—	—
22. Sydney	—	—	—	—	1	—	—	15	—	—	—	—	—	—	—	—	—	—	1	—	—	32	—	—	1	—	1
23. MYHAT	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	17	17	—	—	—
24. Pitt Substudy	—	—	1	—	—	—	—	—	—	—	—	—	—	—	48	—	—	—	—	—	—	—	17	81	—	—	34
25. VMAC	—	—	—	1	1	1	—	—	1	2	1	—	2	11	2	2	—	—	1	—	—	1	—	—	51	1	1
26. Victoria SCDS	4	—	—	6	5	—	—	—	16	1	33	—	—	4	—	1	—	—	—	—	—	—	—	—	—	1	53
27. WRAP	—	—	1	—	1	—	—	—	—	—	—	—	—	—	34	—	—	—	1	—	—	1	—	34	1	—	35
Average Number of Linking Items <sup>a</sup>	4.0	39.7	11.7	2.3	2.6	1.0	6.3	5.2	21.9	1.6	10.6	13.5	1.2	2.8	24.3	1.1	2.3	40.0	1.0	40.0	1.3	3.8	17.0	25.0	2.0	7.9	10.4

*Note.* Because substudies are included, the total number of studies = 27; italicized numbers on the diagonal show the number of items administered by a given study. ADC-NYU = Alzheimer's Disease Center Clinical Core and Center for Brain Health; ADNI = Alzheimer's Disease Neuroimaging Initiative; Ans Dem Cohort = Amsterdam Dementia Cohort; A4 = Anti-Amyloid Treatment in Asymptomatic Alzheimer's study; AIBL = Australian Imaging Biomarkers and Lifestyle Flagship Study of Ageing; CLSA = Canadian Longitudinal Study on Aging; Dart-Indiana = Dartmouth Memory and Aging Study/Indiana Memory and Aging Study; EAS = Einstein Aging Study; AgeCode = German Study on Ageing, Cognition, and Dementia in Primary Care Patients; HABS-HD = Health and Aging Brain Study, Health Disparities; HABS = Harvard Aging Brain Study; HELIAD = Hellenic Longitudinal Investigation of Aging and Diet; IMAP = Institut des Matériaux Poreux de Paris; IADC = Indiana Alzheimer's Disease Center Cohort; LEILA 75+ = Leipzig Longitudinal Study of the Aged; MCSA = Mayo Clinic Study of Aging; MYHAT = Monongahela-Youghiogheny Health Aging Team; Pitt = Pittsburgh; VMAC = Vanderbilt Memory and Alzheimer's Center; Victoria SCDS = Victoria Subjective Cognitive Decline Study; WRAP = Wisconsin Registry for Alzheimer's Prevention.

<sup>a</sup> Does not include items on the diagonal. The mean number of linking items between studies was 11.1.

(Appendices continue)

## Appendix F

### Number of Studies With Which Each Study Shares Linking Items

Study abbreviation	Number of studies
ADC-NYU	3
ADNI	6
Ams Dem Cohort	7
A4	9
AIBL	7
Barcelona	2
Bonn	4
CLSA	5
Dart-Indiana	14
EAS	7
EAS Substudy 1	7
EAS Substudy 2	5
AgeCoDe	5
HABS-HD	9
HABS	10
HELIAD	7
IMAP Caen Group	3
IADC	6
LEILA 75+	4
MCSA	4
Fundació	3
Sydney	5
MYHAT	1
Pitt Substudy	4
VMAC	14
Victoria SCDS	9
WRAP	7
Average	6.2

*Note.* Because substudies are included, the total number of studies = 27. ADC-NYU = Alzheimer's Disease Center Clinical Core and Center for Brain Health; ADNI = Alzheimer's Disease Neuroimaging Initiative; Ams Dem Cohort = Amsterdam Dementia Cohort; A4 = Anti-Amyloid Treatment in Asymptomatic Alzheimer's study; AIBL = Australian Imaging Biomarkers and Lifestyle Flagship Study of Ageing; CLSA = Canadian Longitudinal Study on Aging; Dart-Indiana = Dartmouth Memory and Aging Study/Indiana Memory and Aging Study; EAS = Einstein Aging Study; AgeCode = German Study on Ageing, Cognition, and Dementia in Primary Care Patients; HABS-HD = Health and Aging Brain Study, Health Disparities; HABS = Harvard Aging Brain Study; HELIAD = Hellenic Longitudinal Investigation of Aging and Diet; IMAP = Institut des Matériaux Poreux de Paris; IADC = Indiana Alzheimer's Disease Center Cohort; LEILA 75+ = Leipzig Longitudinal Study of the Aged; MCSA = Mayo Clinic Study of Aging; MYHAT = Monongahela-Youghiogheny Health Aging Team; Pitt = Pittsburgh; VMAC = Vanderbilt Memory and Alzheimer's Center; Victoria SCDS = Victoria Subjective Cognitive Decline Study; WRAP = Wisconsin Registry for Alzheimer's Prevention.

(Appendices continue)

## Appendix G

**Differences in Mean Level of ECog Self Common Latent Variable Means (and Standard Errors)  
Relative to the MCSA Under Assumptions of Measurement Invariance (MI) and Partial  
Measurement Invariance ( $N = 5,742$ )**

Study	MI	Partial MI	% difference	Absolute difference
ADNI	-.580 (.040)	-.569 (.042)	2	.011
Dart-Indiana	-.329 (.191)	-.317 (.193)	4	.012
HABS	-.103 (.049)	-.121 (.064)	17	.018
IADC	-.171 (.078)	-.139 (.079)	19	.032

*Note.* Entries under MI report the estimated mean difference (and standard error) for the named study relative to the MCSA from a model that assumes measurement invariance (i.e., no DIF). Entries under partial MI reported estimated mean differences from a model that allows for some items to have DIF across study. ECog Self = Everyday Cognition-Subject/Self-Report; MCSA = Mayo Clinic Study of Aging; ADNI = Alzheimer's Disease Neuroimaging Initiative; Dart-Indiana = Dartmouth Memory and Aging Study/Indiana Memory and Aging Study; HABS = Harvard Aging Brain Study; IADC = Indiana Alzheimer's Disease Center Cohort.

## Appendix H

**Fifteen Items With the Smallest SEM in the Range of -1 to 0**

Questionnaire abbreviation	Cognitive domain	Item stem	Response options <sup>a</sup>	Temporal referent	Ability versus change	SEM; information
1. VMAC CCQ	Memory	Do you think you have problems with your memory? <sup>b</sup>	No; yes	Current	A	0.32; 9.93
2. CFI	Memory	Compared to 1 year ago, do you feel that your memory has declined substantially? <sup>b</sup>	No; maybe; yes	1 year ago	C	0.37; 7.45
3. VMAC CCQ	Memory	Do you have problems with your memory compared to the way it was 5 years ago? <sup>b</sup>	No problems; some minor problems; major problems	5 years ago	C	0.39; 6.53
4. CFI	Memory	Have you been misplacing things more often? <sup>d</sup>	No; maybe; yes	1 year ago	C	0.39; 6.49
5. CFI	Language	Do you have more trouble recalling names, finding the right word, or completing sentences? <sup>c</sup>	No; maybe; yes	1 year ago	C	0.40; 6.26
6. CFI	Memory	Do you find that you are relying more on written reminders (e.g., shopping lists, calendars)? <sup>c</sup>	No; maybe; yes	1 year ago	C	0.42; 5.71
7. VMAC CCQ	Memory	Do you have difficulty remembering a conversation from a few days ago? <sup>b</sup>	No; yes	Current	A	0.42; 5.69
8. VMAC CCQ	Memory	Overall, do you feel you can remember things as well as you used to? <sup>d</sup>	No; yes	Unclear	C	0.42; 5.65
9. VMAC CCQ	Memory	Do you have complaints about your memory in the last 2 years? <sup>b</sup>	No; yes	Past 2 years	A	0.43; 5.50
10. ADL Abbrev	Memory	Remembering what I intended to do <sup>b</sup>	<i>Much better;</i> <i>slightly better;</i> <i>no change;</i> <i>slightly worse;</i> <i>slightly to moderately worse;</i> <i>worse;</i> <i>moderately worse;</i> <i>much worse</i> <sup>a</sup>	5 years ago	C	0.43; 5.37
11. CFI	Executive	Compared to 1 year ago, do you have more difficulty managing money (e.g., paying bills, calculating change, completing tax forms)? <sup>b</sup>	No; maybe; yes	1 year ago	C	0.43; 5.35

(Appendices continue)

**Appendix H** (continued)

Questionnaire abbreviation	Cognitive domain	Item stem	Response options <sup>a</sup>	Temporal referent	Ability versus change	SEM; information
12. VMAC CCQ	Memory	Do you think that your memory is worse than 5 years ago? <sup>d</sup>	No; yes	5 years ago	C	0.45; 5.00
13. VMAC CCQ	Memory	Do you lose objects more often than you did previously? <sup>b</sup>	No; yes	Unclear	C	0.45; 4.87
14. CFI	Attention	Do you have more trouble following the news, or the plots of books, movies, or TV shows? <sup>b</sup>	No; maybe; yes	1 year ago	C	0.46; 4.78
15. HELIAD SCI	Memory	Do you have difficulty remembering things that you just read or heard? <sup>b</sup>	No problems; problems	Current	A	0.46; 4.70

*Note.* SEM = standard error of measurement; CFI = Cognitive Function Index; ADL Abbrev = Activities of Daily Living Rating Scale Self-Version, Abbreviated; HELIAD SCI = HELIAD Subjective Cognitive Items; VMAC CCQ = Vanderbilt Memory and Alzheimer's Center Cognitive Complaint Questionnaire.

<sup>a</sup> The italicized response options were combined for purposes of data harmonization due to the categories having fewer than five respondents. <sup>b</sup> Item also found among the top 15 items within the -2 to -1 range ( $n = 10$  in total). <sup>c</sup> Item also found among the top 15 items within the 0-1 range ( $n = 2$  in total). <sup>d</sup> Item not found among the top 15 items for either the -2 to -1 range or 0-1 range ( $n = 3$  in total).

**Appendix I****Fifteen Items With the Highest SEM in the Range of -4 to +4**

Questionnaire abbreviation	Cognitive domain	Item stem	Response options <sup>a</sup>	Temporal referent	Ability versus change	SEM; information
1. IQCODE-Short	Executive	Handling money for shopping	<i>Much improved;</i> <i>a bit improved;</i> <sup>b</sup>  <i>not much change;</i> <i>a bit worse;</i> <i>much worse</i>	Compared with 10 years ago	C	73.40; 0.00
2. CAPM-C	Memory	When I forget to do something I had planned to do, it is usually not because I forgot what I had to do but because I forgot when I had to do it	<i>Strongly disagree;</i> <i>disagree;</i>  <i>agree;</i> <i>strongly agree;</i>	Current	A	13.79; 0.01
3. Einstein HSA	Memory	Compared with 1 year ago, do you have trouble remembering things?	Less often; about the same; more often	Compared with 1 year ago	C	10.22; 0.01
4. MIA	Memory	I am poor at remembering trivia	<i>Disagree;</i> <i>strongly disagree;</i>  undecided; <i>strongly agree;</i> <i>agree</i>	Current	A	8.14; 0.02
5. Squire	Memory	My ability to recall things that happened during my childhood is	<i>Better than ever before;</i> <i>almost better than ever before;</i> <i>slightly more of a change;</i> <i>a little bit of a change</i>  no change; <i>almost no change;</i> <i>slightly worse than ever before;</i> <i>a little worse than ever before;</i> <i>worse than ever before</i>	"Ever before"     "Ever before" "Ever before"	C   C C	7.84; 0.02   7.84; 0.02 7.84; 0.02

(Appendices continue)



## Appendix I (continued)

Questionnaire abbreviation	Cognitive domain	Item stem	Response options <sup>a</sup>	Temporal referent	Ability versus change	SEM; information
6. Ams Dem Cohort CC	General	Are you worried about these complaints?	No; yes	Current	A	7.33; 0.02
7. SCCS	Orientation	Remembering what day/date/month it is?	Better; worse; the same	"Than you used to be"	C	6.81; 0.02
8. MATS	Memory	Any other problems with your memory?	No; yes	Current	A	6.04; 0.03
9. Squire	Memory	My ability to recall things that happened a long time ago is	<i>Better than ever before; almost better than ever before; slightly more of a change; a little bit of a change;</i>  no change;  <i>almost no change; slightly worse than ever before; a little worse than ever before; worse than ever before</i>	"Ever before"	C	5.76; 0.03
10. SCCS	Memory	Remembering where you've put things that you use often? (keys, watch, glasses, etc.)	Better; the same; worse	"Than you used to be"	C	5.71; 0.03
11. Ams Dem Cohort CC	Language	Do you have complaints about language or difficulty finding the right words?	No; yes	Current	A	5.54; 0.03
12. CFQ	Language	Do you find yourself suddenly wondering whether you've used a word correctly?	Never; very rarely; <i>occasionally; quite often; very often</i>	In the past 6 months	A	5.43; 0.03
13. Ams Dem Cohort CC	Attention	Do you have complaints about your attention and concentration?	No; yes	Current	A	5.41; 0.03
14. SCCS	Memory	Remembering things (events, people, etc.) from a long time ago?	Better; the same; worse	"Than you used to be"	C	5.33; 0.04
15. SCCS	Memory	Finding the right word to use to describe something you know well? (names of familiar objects, etc., not names of people)	Better; the same; worse	"Than you used to be"	C	5.04; 0.04

*Note.* Ams Dem Cohort CC = Amsterdam Dementia Cohort, subjective cognitive concerns screener; CAPM-C = Comprehensive Assessment of Prospective Memory Section C; CFQ = Cognitive Function Index; Einstein HSA = Albert Einstein Health Self-Assessment; IQCODE Short = Short Form of the Informant Questionnaire on Cognitive Decline in the Elderly, Self-Report Version; MATS = Memory and Aging Telephone Screen; MIA = Metamemory in Adulthood Questionnaire; SCCS = Subjective Cognitive Complaint Scale, also called the Subjective Memory Scale; Squire = Squire Memory Self-Rating Questionnaire.

<sup>a</sup> The italicized response options were combined for purposes of data harmonization due to the categories having fewer than five respondents.

<sup>b</sup> Highlighted response options were combined for purposes of data harmonization due to the categories having fewer than five respondents.

Received October 24, 2021

Revision received November 8, 2022

Accepted November 11, 2022 ■