Cognitive Testing in Large-Scale Surveys: Assessment by Telephone

Cognitive functioning is a key indicator of an individual's overall health and well-being, yet large-scale survey studies of aging typically do not examine cognition. Cognitive measures are seldom included in survey instruments, perhaps because it is assumed that reliable and valid assessments are too difficult and time consuming to administer in a survey format by lay interviewers. Assessment of cognitive functioning traditionally is carried out in person, usually in a laboratory or clinical setting by trained testers, using long, time-consuming batteries that include multiple measures of each cognitive domain of interest. Thus, many survey researchers have been reluctant to include cognitive assessment in their batteries even though there is increasing recognition of the importance of cognitive functioning for understanding overall functioning and health, especially in later life.

Psychologists, steeped in psychometric tradition, typically administer lengthy test batteries with multiple items and trials in order to achieve reliable and valid assessments of cognition. This approach to cognitive testing is not feasible for the cutting-edge research that emphasizes multidisciplinary perspectives, including a focus on brain and behavior or mind and body connections, and that must address multiple domains in addition to cognition. Thus, there is a growing demand for shorter cognitive batteries to include in these “big picture” studies. In current research paradigms with a focus on multiple aspects of functioning, it is often not feasible to spend more than 15 minutes on any given domain such as cognition. Thus, it is important to select brief but highly reliable tests that are sensitive to variations and individual differences within the full range of cognitive functioning. Because of rising costs and reluctance of respondents to talk with interviewers in person or in their homes, there has been increasing use of telephone rather than face-to-face data collection. Thus, it becomes critical to develop and validate cognitive batteries that are short enough to be included in national surveys but are also reliable, valid, sensitive to wide age variations, and appropriate for telephone administration. The MacArthur Foundation Research Network on Successful Midlife Development (Brim, Ryff, & Kessler, 2004) has made great strides in developing brief but psychometrically sound instruments for large survey administration in areas traditionally using very long batteries. This trend for brief but reliable and valid measures was begun in the first Midlife in the U.S. National Survey (MIDUS; Lachman & Firth, 2004) for measures of personality, sense of control, and well-being and health, and it was extended to cognition in the second wave, the MIDUS II.
Why Assess Cognition in Surveys?

Including cognitive measure in epidemiological and longitudinal surveys has multiple benefits. For survey researchers it is useful to describe cognitive functioning of participants and to explore links between cognition and other domains of interest, such as health and economic behavior. Although much of the previous survey work on cognitive aging has used basic measures for dementia screening, there is an increased effort to measure variations of cognitive functioning in the normal range. Even when the primary focus of a study is on other variables, it is important for researchers to verify that participants cognitive status is adequate to ensure valid responses to their questions. Moreover, there is evidence that age-related cognitive changes can impact several aspects of self-reports of behavior and opinion, including comprehension of questions and effects of question context and response formatting (Schwarz, 1999). It is also beneficial for researchers in the field of cognitive aging to have cognitive assessments with larger, more representative samples than are typically studied in the laboratory.

Understanding cognitive aging involves a focus on more than just age differences. Laboratory samples are often selected to have high education to match college students and to have few or no health problems; even when such samples are not deliberately recruited, the pool of older adults who volunteer to come into the laboratory are typically high functioning because of self-selection. Such studies typically involve the comparison of young college students and older adults, matched on educational level, using an extreme two-age-groups design. As a consequence, there is little within-group variability in health, socioeconomic status and other key factors. Survey research goes beyond the limits of the laboratory and enables data collection with more representative samples. Thus, inclusion of cognitive batteries in surveys allows for investigation of cognitive aging in relation to a broader range of dimensions, such as disease, education, mental health, and stress.

Effective cognitive functioning throughout adulthood is a key element not only in an individual's quality of life but also in the ability to remain an independently functioning member of society. Although some large-scale surveys have included telephone assessments of cognition, the focus has largely been on screening for dementia (see Table 30.1). Declines in cognitive function can impact older adults ability to perform instrumental activities of daily living such as managing finances, following medical instructions, and planning sequences of activities, with important implications for health care and both private and public resources (Herzog & Wallace, 1997). Effective functioning in the second half of the life span can be threatened not only by devastating cognitive declines and dementia such as Alzheimer's disease but also by mild cognitive impairment (Petersen
et al., 1999) and normative age-related loss.

In the last two decades, a substantial body of research findings from the laboratory and from large longitudinal studies has documented age-related declines in cognitive abilities among adults over age 60 (e.g., Craik & Salthouse, 2000; Salthouse, 1996; Schaie, 1996). To date, much is less known about changes in cognitive abilities during midlife, even though a large proportion of the U.S. population is now entering the mature years with cognitive systems that will undergo age-related changes (Dixon, deFrias, & Maitland, 2001; Sternberg, Grigorenko, & Oh, 2001; Willis & Schaie, 1999). Baby Boomers, who represent some 77 million people born after World War II (between 1946 and 1964), now comprise one third of the American population and are entrenched in the middle years (Lemme, 1995). Midlife is a period characterized by myriad tasks (Lachman & James, 1997), including juggling career and family responsibilities (Lachman, 2004). These place heavy stress on the ability to divide attention between multiple concurrent activities, which often becomes more difficult with increased age (Tun & Wingfield, 1995). Nevertheless, there is a paucity of nationally representative data on cognitive functioning in mid-and later life, in part because it is difficult for this age group to come to the laboratory for testing given their busy schedules (Lachman, 2004).

<table>
<thead>
<tr>
<th>Cognitive Battery</th>
<th>Study</th>
<th>No. of Participants</th>
<th>Age of Participants</th>
<th>Cognitive Subtests</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blessed Telephone Information Memory Concentration (TIMC)</td>
<td>Kawas et al. (1995)</td>
<td>49</td>
<td>50–98 years</td>
<td>TIMC administered by phone and in person.</td>
</tr>
<tr>
<td>Mini-Mental State Examination (MMSE)</td>
<td>Jorm et al. (1993)</td>
<td>74+ years</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MMSE</td>
<td>Roccaforte et al. (1992)</td>
<td>100</td>
<td></td>
<td>Validity of telephone version of the MMSE; Brief Neuropsychological Screening Test; MMSE as part of the Adult Lifestyles and Function Interview (ALFI-MMSE)</td>
</tr>
<tr>
<td>MMSE</td>
<td>Monteiro et al. (1998)</td>
<td>34 (17 women, 17 men)</td>
<td>M = 76.8 (women) M = 77.6 (men)</td>
<td>• Global Deterioration Scale</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Functional Assessment Staging</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Behavioral Pathology in Alzheimer's Disease Rating Scale</td>
</tr>
<tr>
<td>Test</td>
<td>Reference</td>
<td>Age Range</td>
<td>Notes</td>
<td></td>
</tr>
<tr>
<td>-------------------------------------------------------------</td>
<td>---------------------------</td>
<td>-----------</td>
<td>----------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Minnesota Cognitive Acuity Screen</td>
<td>Knopman et al. (1999)</td>
<td>228</td>
<td>Orientation, Delayed Word Recall, Verbal fluency, Computation, Judgment</td>
<td></td>
</tr>
<tr>
<td>Modified MMSE (3MS)</td>
<td>Norton et al. (1999)</td>
<td>263</td>
<td>3MS and the Telephone Modified MMSE</td>
<td></td>
</tr>
<tr>
<td>Short Portable Mental Status Questionnaire (SPSMQ)</td>
<td>Roccaforte et al. (1992)</td>
<td>63–93</td>
<td>Tested reliability of telephone version of SPMSQ</td>
<td></td>
</tr>
<tr>
<td>Structured Telephone Interview for Dementia</td>
<td>Go et al. (1997)</td>
<td>60–88</td>
<td>The National Institute of Mental Health Genetics Initiative: Clinical Dementia Rating Scale</td>
<td></td>
</tr>
<tr>
<td>Telephone Interview of Cognitive Status (TICS)</td>
<td>Brandt et al. (1988)</td>
<td>133</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TICS</td>
<td>Desmond et al. (1994)</td>
<td>72</td>
<td>M = 72.1 years</td>
<td></td>
</tr>
<tr>
<td>TICS</td>
<td>Grodstein et al. (2000)</td>
<td>2,138</td>
<td>70–78 years</td>
<td></td>
</tr>
<tr>
<td>TICS</td>
<td>Jarvenpaa et al. (2002)</td>
<td>56</td>
<td>52–80 years</td>
<td></td>
</tr>
<tr>
<td>TICS Modified (TICS-M)</td>
<td>Buckwalter et al. (2002)</td>
<td>3,681</td>
<td>80+ years</td>
<td></td>
</tr>
<tr>
<td>TICS-M Computer Assisted Telephone Interview</td>
<td>Buckwalter et al. (2002)</td>
<td>3,681</td>
<td>80+</td>
<td></td>
</tr>
</tbody>
</table>

- Brief Cognitive Rating Scale
- MMSE
Most major epidemiological surveys, such as the Longitudinal Survey on Aging (M. E. Miller, Rejeski, Rebourssin, Ten Have, & Ettinger, 2000) do not measure cognitive function at all. One exception is the Asset and Health Dynamics Among the Oldest Old (AHEAD) study, a telephone survey of 6,500 adults age 70 and over that documented significant changes in mental status and memory function (Herzog & Wallace, 1997). Another is the multisite MacArthur Study of Successful Aging, which examined individuals between the ages of 70 and 79 (Albert et al., 1995). Although longitudinal studies have demonstrated significant changes in midlife for some mental abilities (the Baltimore Aging Study [Shock et al., 1984], Berlin Aging Study [Baltes & Mayer, 1999], Seattle Longitudinal Study [Schaie, 1996], Victoria Longitudinal Study [Hultsch, Hertzog, Dixon, & Small, 1998]), and for some specific groups, such as men (e.g., the Normative Aging Study; Aldwin, Spiro, Levenson, & Bosse, 1989) and postmenopausal women (Women's Health Initiative—Memory Study; Shumaker et al., 1998), no nationally representative, large-scale data have been available specifically for the cognitive functioning of middle-aged men and women until the MIDUS II survey. Results from this study are summarized below.

### Telephone Testing: Advantages and Limitations

In general, cognitive testing is done in person; face-to-face assessment is desirable because it gives the most flexibility in terms of testing equipment and stimuli. It also enables one to establish greater rapport with the participants and personalized treatment in terms of giving breaks and sensitivity to fatigue and understanding

<table>
<thead>
<tr>
<th>Test Method</th>
<th>Study/Author</th>
<th>Sample Size</th>
<th>Age Range</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>TELE self-report interview</td>
<td>Gatz et al.</td>
<td>65 pairs of twins</td>
<td>55+ years</td>
<td>TELE, including the Mental Status Questionnaire MMSE</td>
</tr>
<tr>
<td>Jarvenpaa et al. (2002)</td>
<td>56</td>
<td>52–80 years</td>
<td></td>
<td>Compared TELE with TICS and MMSE</td>
</tr>
<tr>
<td>Chumbler &amp; Zhang (1998)</td>
<td>48</td>
<td>65+ years</td>
<td></td>
<td>Validity of a modified telephone screening device (Gatz et al., 1995) against the MMSE</td>
</tr>
<tr>
<td>Telephone-Assessed Mental State Exam (TAMS)</td>
<td>Lanska et al. (1993)</td>
<td>30</td>
<td>59–88 years</td>
<td>Compared TAMS with MMSE and Alzheimer's Disease Assessment Scale</td>
</tr>
</tbody>
</table>
of instructions. With face-to-face testing it is possible to include a wider range of tests, including those with visual components, those that require written responses, or those with specialized stimuli or equipment that cannot be administered by phone. Nevertheless, the use of telephone assessment for cognitive testing has a good deal of promise.

Telephone testing offers advantages to both clinicians and researchers, including convenience; low expense; and the opportunity to test a greater number of individuals, including those who are unable or unwilling to be tested in person in a laboratory or clinical setting. Testing by telephone allows researchers to access a wider, more diverse range of respondents who vary in physical mobility and geographical distance from the investigator, as well as in health status, age, socioeconomic status, racial/ethnic background, and educational level. Some respondents may also feel more comfortable with the anonymity of phone testing, because they do not have to face an interviewer if they perform poorly.

There are some limitations associated with telephone testing. It is essential that the respondents can clearly hear the interviewer and test stimuli. Given that increased age is associated with declines in auditory acuity, this can present a challenge. Hearing problems can be exacerbated over the telephone because of variations in the quality of connection, the phone equipment, and technical difficulties. Thus, it is important to include a brief hearing test in telephone batteries to establish the effectiveness of hearing. Fortunately, hearing has not presented a significant problem in previous studies, and one telephone study reported that fewer than 4% of an older sample had hearing difficulties (Lipton et al., 2003).

The use of cell phones is not ideal for telephone testing given that they may introduce variable lags in the signal relating to delays in propagation, transmission, and processing, all of which could affect response latencies. This is especially problematic when timing is critical for a test, as is the case with the Stop and Go Switch Task, which we describe later. Some participants, especially those in the younger age ranges, may have only cell phone service. It may be possible to test these individuals on land line phones in their work setting. Some phones have the push buttons in the head set, which can make it difficult if push-button technology is used for the test responses. Portable phones also can have more interference than other types of phone equipment.

Background noise in the home or work setting also may interfere with the clarity of presentation for cognitive tests by telephone. It may be difficult to monitor background noise or other distractions, which can affect performance and are typically controlled in laboratory settings. However, we have been generally successful in arranging phone interviews at times that are convenient for respondents and that minimize distractions from
other people or activities.

Cheating is another potential caveat of telephone testing. Given that the respondent usually cannot be seen by the tester (unless a videophone or webcam is used), it is possible for the respondent to write words down during a memory test, or use paper and pencil to jot down notes. It is also possible for the respondent to get some type of assistance, either from another person or by looking something up in a dictionary or other source. In our experience, we believe the incidence of cheating is low based on the correlations between scores from in-person and telephone administrations in our pilot studies (Tun & Lachman, 2006b), and from the Health and Retirement Study (HRS) reports of concordance between both modes of assessment (Herzog & Rodgers, 1998). In the introduction phase of the survey, we emphasize to participants that no one is expected to be able to answer or complete all of the test questions. We have implemented two procedures to help guard against unwanted writing or cheating. We suggest to our respondents that they close their eyes during the testing and indicate that this can help with concentration and performance. Another strategy we have used is to ask the respondent to hold the phone receiver in one hand and a blank piece of paper in the other hand. We specify that looking at the blank page helps focus and avoid distraction from other sources. It also serves to prevent the respondent from writing. To investigate possible cheating we examine unusual patterns of performance, such as extreme scores, for example, 100% recall on a word list or scores greater than 2 standard deviations above the age group mean. Other unusual patterns include absence of forgetting between immediate and delayed testing of verbal memory or large improvements in performance with a retest. We also compare level of performance on tests that could be influenced by cheating and those that are less susceptible to cheating. If there is a wide disparity we take this as an indication of potential cheating. The decision as to whether and when to drop data from cases suspected of cheating will depend on the specific research circumstances.

Administration and Scoring

There are two general administration modes for testing by telephone: (1) a live interviewer or (2) a computer. We have found that participants, especially older adults, prefer to talk with a person rather than a computer. The live interviewer can make sure that the person clearly hears the tests and can verify that the person understands the instructions before beginning the testing and stop the testing if something goes awry or if the respondent has further questions during the test process. Moreover, the live interviewer can detect problems
with the interview, including extraneous sounds, such as coughs or false starts, which can interfere and lead to errors or incorrect responses with computerized assessments using voice recognition or sound waves.

A live interviewer can use a computer-assisted telephone interview system or play a standardized recording to control the pace and sequencing of test questions. Entry of some simple responses can be done by the interviewer during the testing. Automated computerized administration, without a live interviewer, usually involves presentation by digitized voice recording, a method that usually cannot be responsive to individual participants needs. It is possible to have a person listening in on the telephone to make note of invalid trials and make necessary adjustments.

For both live and computer-assisted administrations, we recommend digital recording of the protocol so that it can be reviewed and scored later. Digitized files are ideal so that scoring can be done automatically using computer software such as for latencies; however, analog audio-taping is also useful for reviewing responses and checking accuracy. Processing and scoring can be done by person or computer. Computerized scoring can be applied manually to audio or text files generated by the interviewer, or automatically on line by means of voice recognition, as in the TELECOG battery described below. This works particularly well when response choices are finite and require only a few distinctive responses so accurate responses can be identified more effectively.

Telephone testing can be supplemented and enhanced with mailings. If testing requires stimuli that cannot be delivered orally, it is possible to send the visual materials to the respondents in advance and ask them not to open the envelope until the phone interview session. The interviewer can refer to the test materials and guide participants to look at the stimuli while answering orally administered test questions.

---

**Previous Surveys with Cognitive Assessments**

Most large studies collect cognitive data by face-to-face, in-person interview (e.g., Atherosclerosis Risk in Communities [Cerhan et al., 1998], the Cardiovascular Health Study [Haan, Shemanski, Jagust, Manolio, & Kuller, 1999], the Framingham Study [Elias, Elisa, D'Agostino, Silbershatz & Wolf, 1997], Medical Research Council's Cognitive Function and Ageing Study [Medical Research Council Cognitive Function and Ageing Study, 1998]). Attempts to assess cognitive functioning by telephone historically have focused on diagnosis of dementia and other cognitive pathologies. The Mini-Mental Status Examination (MMSE; Folstein, Folstein,
& McHugh, 1975) has been adapted for telephone administration in the Telephone Interview for Cognitive Status (TICS; Brandt, Spencer, & Folstein, 1988), which has been used successfully as a screening instrument for dementia. A summary of these and other telephone measures is provided in Table 30.1. The TICS, one of the sources for the HRS/AHEAD battery, is but one of a growing number of telephone assessments of cognitive status and dementia, including the Telephone Cognitive Assessment Battery, the modified TICS (TICS-M), the MMSE as part of the Adult Lifestyles and Function Interview, Blessed Telephone Information Memory Concentration, the Structured Telephone Interview for Dementia, the Telephone-Assessed Mental State Exam, and the Short Portable Mental Status Questionnaire (see Table 30.1). With relatively simple and brief measures, it is possible to obtain a reasonable estimate of dementia status, but such measures are not sensitive across a wider range of cognitive performance in normal healthy adults.

Nevertheless, as summarized in Table 30.2, a number of studies have demonstrated the efficacy of assessing the normal range of cognition by telephone (Herzog & Wallace, 1997; Kawas, Karagiozis, Resau, Corrada, & Brookmeyer, 1995; Nesselroade, Pederson, McClearn, Plomin, & Bergeman, 1988). The TELECOG (Tennstedt, Lachman, & Salthouse, 2004), a computerized telephone test that uses voice recognition to assess memory and attentional switching in adults, has shown similar performance in person and over the telephone and found few differences across testing mode. The large-scale HRS/AHEAD study (Herzog, Rodgers, & Kulka, 1983; Herzog & Wallace, 1997) has demonstrated the feasibility of a telephone survey of cognitive function in adults over the age of 70. It is important to note that the AHEAD study and the large-scale Nurses Health Study of 18,000 older women found no significant differences in performance between respondents tested by telephone and face-to-face assessments (Herzog & Rodgers, 1998). Another study that compared phone and face-to-face administration found correlations, adjusted for age and depression, ranging from .71 to .89 for the Age-Related Eye Disease cognitive battery (Petrill, Rempell, Oliver, & Plomin, 2002).

The Karolinska Institute Twins Study in Sweden (Nesselroade et al., 1988) also assessed cognitive functioning in older adults by telephone using reduced versions of standardized tests. They found that shortening the standard versions of established cognitive tests to half the original length and administering them by telephone only minimally compromised the overall reliability and validity of the instrument.

<table>
<thead>
<tr>
<th>Cognitive Battery</th>
<th>Study</th>
<th>No. of Participants</th>
<th>Age of Participants</th>
<th>Cognitive Subtests</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study/Method</td>
<td>Participants</td>
<td>Age Range</td>
<td>Tests and Procedures</td>
<td></td>
</tr>
<tr>
<td>----------------------</td>
<td>--------------</td>
<td>-----------</td>
<td>-------------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>AREDS Telephone Battery</td>
<td>Rankin et al., 2005</td>
<td>1,738</td>
<td>55–80 years</td>
<td>WMS-R Logical Memory I and II, TICS-M, Letter fluency FAS, Animal Category Fluency, digits backward.</td>
</tr>
<tr>
<td>BTACT</td>
<td>Tun &amp; Lachman (2006b)</td>
<td>84 in pilot, 4,000 in MIDUS II</td>
<td>23–85 years</td>
<td>Free Recall Immediate and Delayed (15 words), Backward Digit Span, Category Fluency, Number Series, Speed, Task Switching</td>
</tr>
<tr>
<td>HRS/AHEAD Study</td>
<td>Herzog &amp; Wallace (1997)</td>
<td>6,500+</td>
<td>70+ years</td>
<td>Immediate free recall test, delayed free recall test, Serial 7s test, counting backwards, naming the day of the week and date, naming objects, naming President and Vice President of the United States, Modified Similarities test from WAIS-R, Self-rating of memory</td>
</tr>
<tr>
<td>Nurses Health Study</td>
<td>Grodstein et al. (2000)</td>
<td>2,138 women</td>
<td>70–78 years</td>
<td>East Boston Memory Test: Immediate and delayed recall, TICS, 10-word list for immediate recall, Verbal Fluency</td>
</tr>
<tr>
<td>TELECOG</td>
<td>Tennstedt et al. (2004)</td>
<td>120</td>
<td>18–87 years</td>
<td>Free Recall Immediate and Delayed (10 words), Working Memory—N-Back, Task Switching</td>
</tr>
<tr>
<td>Telephone-Assessed Cognitive Ability</td>
<td>Nesselroade et al. (1988)</td>
<td>194 pairs of twins</td>
<td>27.5–82 years</td>
<td>Analogies, figure logic (fluid intelligence), Forward and backward digit span (short-term memory, Information, synonyms (crystallized intelligence)</td>
</tr>
<tr>
<td>Telephone-Assessed Measure of Cognitive Ability</td>
<td>Petrill et al. 2002</td>
<td>52</td>
<td>6–8 years</td>
<td>Two verbal ability tests, 3 nonverbal ability tests, phonological awareness measure; correlated with Stanford-Binet</td>
</tr>
<tr>
<td>Telephone Cognitive Battery</td>
<td>Kent &amp; Plomin (1987)</td>
<td>212</td>
<td>9–15 years</td>
<td>Verbal, spatial, perceptual speed, and memory abilities</td>
</tr>
</tbody>
</table>

**NOTE:** AREDS = Age-Related Eye Disease; WMS-R = Wechsler Memory Scale—Revised; RICS-M = Telephone Interview of Cognitive Status Modified; BTACT = Brief Test of Adult Cognition by Telephone; MIDUS II = Midlife in the U.S. National Survey; HRS/AHEAD = Health and Retirement Study/Asset and Health Dynamics Among the Oldest Old; WAIS-R = Wechsler Adult Intelligence Scale—Revised; TICS = Telephone Interview of Cognitive Status.
Hrs/Ahead Study

The HRS/AHEAD study designers (Herzog & Wallace, 1997) recognized the central role of cognitive functioning in relation to functional impairment, disability, and health care utilization among the elderly. They also considered the possible economic consequences of limitations in cognitive abilities, especially involving work and decision making and planning for retirement. Finally, it was recognized that cognitive difficulties needed to be identified, because they could compromise the data quality for the entire survey (Schwarz, 1999).

The HRS/AHEAD cognitive battery (Ofstedal, McAuley, & Herzog, 2002) was designed to be administered by telephone, and it includes measures of memory, mental status, and verbal ability. For respondents who were unable to respond, a proxy informant was asked to rate the respondent's memory, judgment, organization of time, and complete Jorm's 16-item IQCODE (Jorm, Scott, Cullen, & MacKinnon, 1991), which is used to assess dementia.

Many of the items in the HRS/AHEAD cognitive battery were adapted from the TICS (Brandt et al., 1988; Breitner, Welsh, Robinette, & Gau, 1995), or from the Iowa Established Populations for Epidemiologic Study of the Elderly (Purser et al., 2005). Some HRS/AHEAD items were later modified (e.g., the use of four alternate word lists for the recall tasks). The battery is heavily focused on knowledge and orientation items, which are most useful for identifying persons with some degree of cognitive impairment. The immediate and delayed free recall tests, the serial 7s, and the counting backwards test are indexes of episodic and working memory. These are important dimensions to include because there is strong evidence to suggest that these are among the first cognitive functions to decline during healthy aging (Bckman, Small, & Wahlin, 2000).

Telecog

TELECOG (Tennstedt, Lachman, & Salthouse, 2004) is a brief, computerized telephone-administered instrument that validly and reliably measures normal cognitive functioning using voice recognition technology. TELECOG focuses on two cognitive domains, memory and cognitive processing speed, which are highly age sensitive and represent basic capacity and processing skills that underlie higher-level cognitive performance and have been related to daily functional activities.

TELECOG includes tests of working memory and episodic memory. Working memory is assessed with the N-
Back task (Gevins et al., 1990), which requires the respondent to solve a set of one-digit arithmetic problems and to recall the last addend in each problem. Four different versions of the task are used: 0-Back, 1-Back, 2-Back, and 3-Back. These different trial types allow for increasing the working memory demands without altering the stimulus condition.

TELECOG measures episodic memory with the auditory list learning task, including tests of both immediate and delayed recall. In this task, the respondent hears a list of 10 unrelated words and then is asked to say as many words as he or she can remember (immediate recall). The order in which the words are said is not of concern. Immediately after a respondent is finished with the immediate recall, the 10-word list is presented again but in a different serial order. A total of three trials of the list are presented to the respondent in the immediate word recall task, each in a different order, and the delayed recall is assessed after 20 minutes.

TELECOG measures cognitive processing speed with a switching task (Salthouse, Fristoe, McGuthry, & Hambrick, 1998). The switching task is a test of odd/even or more/less determinations. A respondent hears a string of numbers presented one at a time and is asked to make one of two decisions about each number: (1) “whether the number is odd or even” or (2) whether the number is more or less than the number five.” At random times during the digit presentations, the respondent is told to switch from one type of decision to the other. There were only a few significant within-subject differences in performance between face-to-face computerized testing and telephone assessment, and age patterns were similar across testing modes. One exception is that the older adults were a little faster on the switching tasks by phone than in person, perhaps because the motoric response of key pressing on a computer takes more time than a verbal response.

**Brief Test of Adult Cognition by Telephone**

A relatively new telephone measure, the Brief Test of Adult Cognition by Telephone (BTACT), was used in the MIDUS II national survey (Tun & Lachman, 2006b). The range of cognitive domains tested includes key abilities critical to adult functioning based on cognitive aging theory, such as reasoning, executive function, attention, and speed of processing. The BTACT is appropriate for testing a wide range of the population, including well-functioning younger and middle-aged adults as well as older adults. This allows for sensitivity to individual differences in cognition that may be associated with a large array of biological, social, health, and psychological factors. The BTACT requires less than 20 minutes to administer and includes an optional attention test. We developed an alternative form of the BTACT for studies that involve repeated measures or
longitudinal designs, to address retest effects.

The BTACT (see Table 30.3) was designed to tap areas of cognitive function that are sensitive to the effects of aging, including episodic verbal memory (Craik & Anderson, 1999), working memory span and executive function (Baddeley, 1986, 1996), reasoning (L. S. Miller & Lachman, 2000; Schaie, 1996), and speed of processing (Meyerson, Hale, Wagstaff, Poon, & Smith, 1990; Salthouse, 1996; Verhaeghen & Salthouse, 1997). An optional supplementary test records response latencies, which afford a measure of speed of processing, attention, and switch costs. The use of latency data adds an extra dimension to the cognitive measures, providing greater sensitivity to subtle individual differences in speed of processing that may not be revealed by accuracy measures alone (e.g., Salthouse, 1996). Latencies were calculated by measuring the distance in milliseconds in the speech signal between stimulus onset and response onset using sound editing software. Latencies calculated from recordings of phone interviews have been shown to be similar to those taken from in-person interviews. One possible caveat is that wireless phones may introduce variable delays, and for this reason we caution against testing over cell phones.

<table>
<thead>
<tr>
<th>Task</th>
<th>Theoretical Construct(s)</th>
<th>Test Used</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word list recall (immediate and delayed)</td>
<td>Episodic verbal memory</td>
<td>Free recall of a list of 15 words drawn from the Rey Auditory-Verbal Learning Test (Lezak, 1995; Rey, 1964)</td>
</tr>
<tr>
<td>Backward digit span</td>
<td>Working memory span</td>
<td>Highest span achieved in repeating strings of digits backwards (Wechsler, 1997)</td>
</tr>
<tr>
<td>Category fluency</td>
<td>Verbal fluency: Executive function, semantic memory retrieval</td>
<td>Number of animal names produced in 1 minute (after Borkowski et al., 1967; see also Tombaugh et al., 1999)</td>
</tr>
<tr>
<td>Number series</td>
<td>Inductive reasoning</td>
<td>Complete the pattern in a series of 5 numbers with a final number (e.g., 2, 4, 6, 8, 10… 12). Five problems include 3 types of patterns (after Schaie, 1996; Salthouse &amp; Prill, 1987)</td>
</tr>
<tr>
<td>Backward counting</td>
<td>Processing speed</td>
<td>Maximum number of items produced counting backwards from 100 in 30 seconds (after AHEAD study; Herzog &amp; Wallace, 1997)</td>
</tr>
</tbody>
</table>
The BTACT battery includes *Episodic Verbal Memory*, which includes immediate recall and delayed recall of a 15-word list; *Working Memory Span*, reflecting a system that stores and manipulates information, measured with backward digit span; *Verbal Fluency*, assessed by category fluency, an index of executive function that is linked with frontal lobe function; *Inductive Reasoning*, a measure of fluid intelligence assessed with number series completion; and *Speed of Processing*, measured with a backward counting task requiring rapid generation of a nonautomatic sequence. In addition, an optional Stop and Go Switch Task test yields measures of reaction time and *task-switching* costs, as well as inhibitory control.

Based on factor analysis of the BTACT tests, we found a one-factor solution. On this basis a composite measure can be computed. This is useful for research in which it is suitable to use one general measure, especially if there are not specific hypotheses about differential relationships with the individual subtests or dimensions.

Many of the BTACT tests have been used previously in neuropsychological and laboratory applications. To confirm that our telephone measures yield results similar to the more standard in-person tests, we carried out a validation study on the BTACT both in-person and by telephone, and found no significant effect of mode of testing for any of the subtests (see Table 30.4). We also demonstrated the expected significant correlations between BTACT measures and standardized in-person assessments on other tests of similar cognitive domains administered in person (Tun & Lachman, 2006a, 2006b).

<table>
<thead>
<tr>
<th>Table 30.4 Correlations Between Telephone and Face-to-Face Versions of the Brief Test of Adult Cognition by Telephone (BTACT) Subtests</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>BTACT Subtest</strong></td>
</tr>
<tr>
<td>----------------------------------</td>
</tr>
<tr>
<td>Word List Immediate</td>
</tr>
<tr>
<td>Word List Delayed</td>
</tr>
</tbody>
</table>

NOTE: AHEAD = Asset and Health Dynamics Among the Oldest Old.
The BTACT and the Stop and Go Switch Task were administered to the MIDUS II nationally representative sample of 4,014 adults ages 32 to 84, with a mean age of 55.26 (SD = 12.29). The sample was 54% women, and education ranged from 6 to 20 years (Mean education = 14.23, SD = 2.6), with less education in the older age cohorts. Older age cohorts rated their health below the younger groups on a self-report scale, with an overall mean of 3.63 (SD = 0.96; 1 = poor to 5 = excellent). For analysis purposes, the sample was divided into five age groups: G1 (32–44), G2 (45–54), G3 (55–64), G4 (65–74), and G5 (75–85). Results from this sample with the BTACT tests are presented in Figure 30.1. All tests showed significant age differences, with older adults performing more poorly than middle-aged and younger adults. The results for the composite scale are presented in Figure 30.2. Significant differences were found among all age groups.

Switching ability has been assessed primarily in laboratory settings, with typically small numbers of research participants and selectivity bias of the participants who are willing to come into a laboratory. However, innovative methods of testing can provide new insights into this paradigm (Reimers & Maylor, 2005). With the Stop and Go Switch Task, we demonstrated how using a novel method—telephone technology—can expand the
range of participants tested and shed new light on individual differences in executive processes.

Figure 30.1 Brief Test of Adult Cognition by Telephone Accuracy Scores (z Scores) by Age Group
As shown in Figure 30.3, for all age groups latencies increase with the complexity of the task. In the Stop and Go Switch Task, participants hear the words “red” and “green” and make simple speeded responses of either “stop” or “go,” depending on the response rule. The response rule can indicate either a congruent response or an incongruent response that requires inhibitory control of the prepotent response. Latencies are smallest for single-task trials that involve one response rule. On the mixed-task trials that require alternating between two different response rules, nonswitch trials are faster than switch trials. Older adults were slower than young and middle-aged participants on all task conditions, but the magnitude of age differences was greater when switching was involved.

The cost of switching is illustrated in Figure 30.4, which shows age differences in local switch cost across five age groups. Local switch costs (the increase in mixed task latencies on switch trials compared with nonswitch trials) showed robust effects of age, beginning in middle age. We also found age differences in general or global switch costs (the increase in mixed-task latencies compared with single-task latencies). It is important to note that these age-related increases in both local and general switch costs persisted even after control-
ling for baseline slowing, suggesting that switching was more difficult with age. In addition, older age groups showed an increase in the effect of congruency (congruent responses relative to incongruent responses), which we take as a reflection of inhibitory control processes (Tun & Lachman, 2006a).

**Figure 30.3 Stop and Go Switch Task Reaction Times by Trial Type and Age Group (With Standard Errors)**

We also found different effects of other demographic variables, which can shed light on the central executive process involved in switching and inhibitory control. Specifically, gender and level of education have different associations with task switching and with inhibitory control. Female gender was associated both with larger switching costs and poorer inhibitory control on incongruent trials; however, education showed a different pattern of effects, such that higher levels of education were associated with smaller switching costs but not with consistently better inhibitory function (Tun & Lachman, 2006a).

**Conclusion**

Given the state of the art of research on cognition and aging, there is a need to incorporate cognitive assessments into large-scale surveys on topic areas such as health and well-being, economics, and stress. Testing
by telephone greatly increases the range of possible participants, including people who cannot or will not come into the laboratory or clinical settings. Surveys typically cannot accommodate lengthy cognitive batteries even though the studies may benefit from understanding individual differences in cognition and longitudinal changes in cognition in response to time, experimental manipulations, or interventions. Cognition may also be useful as a covariate in studies of health and functional behavior. Another important focus has been to use cognitive data to assess the validity of survey responses, presuming that individuals with cognitive deficits might not provide valid responses to other items in the questionnaire (Knauper, Belli, Hill, & Herzog, 1997).

**Figure 30.4 Stop and Go Switch Task Local Switch Cost by Age Group (With Standard Errors)**

NOTE: Switch costs = median of mixed task switch trials – median of mixed tasks non-switch trials.

Many existing telephone batteries in surveys focus on detecting signs of dementia, but there is a need to include cognitive batteries that are more sensitive to the subtle changes associated with normal aging. Given that in adult samples a relatively small proportion of survey participants are expected to have or develop dementia, it is important to focus on those within normal levels of cognitive aging, or those who maintain adequate cognitive function well into their later years. In long-term longitudinal studies, especially with older samples, it may be possible to identify precursors of dementia or diagnostic signs of pre-clinical conditions.
In future studies, we will focus on standardizing and further validating the BTACT battery, with the goal of establishing norms and diagnostic criteria. Preliminary data look very promising, and we hope to further validate these measures with the in-person data from a subsample of the MIDUS from Boston who were tested with standard in-person cognitive measures. In the MIDUS II survey the cognitive assessment was limited to 15 minutes, but in pilot testing we found that we could also successfully administer paired associate learning and alternating fluency tasks over the phone. Another likely candidate for telephone assessment would be text recall tasks. The use of brief phone batteries may be beneficial for the cognitive aging research field, especially if this assessment approach is associated with less fatigue and anxiety than more traditional long batteries given in person.

Individual differences in cognitive ability have important implications for understanding differences in behaviors and outcomes in multiple domains (e.g., health, economics, social domains). If one is interested in prevention, it becomes important to think about identifying risk factors for cognitive decline, rather than focusing only on the consequences of decline. If individual differences in cognitive functioning for persons in the normal range can be measured well, it would enable tracking over time and identification of predisposing factors associated with decline. This would provide much-needed information about the emergence of cognitive impairments and preclinical status in later life.

Little attention has been given to cognitive functioning as an important resource for later life functioning. For example, cognitive abilities can serve as a moderator of social class differences in health and retirement outcomes. Relationships between cognitive functioning and other important factors, such as health, economic well-being, stress, and depression, need further exploration. Cognitive functioning and changes are interesting outcomes in their own right. There is still a great deal to be learned about predictors of change over time in cognitive functioning, or the transition from normal cognition to cognitive impairment to dementia. All of these and many other potentially important research questions may be more realistically and effectively addressed by including cognitive assessments over the telephone in large-scale surveys.

Margie E. Lachman Patricia A. Tun

AUTHORS NOTE: This research was supported by a grant from the National Institute on Aging (PO1-AG020166) to conduct a longitudinal follow-up of the MIDUS (Midlife in the U.S.) investigation. We appreciate Chandra Murphy’s valuable assistance with data analysis and manuscript preparation.
References


Kent, J., Plomin, R. Testing of specific cognitive abilities by telephone and mail. *Intelligence* 11(1987) 391–400


- telephones
- inhibitory control
- word lists
- dementia
- batteries
- cognitive assessment
- episodic memory

https://doi.org/10.4135/9781412976589