



Published in final edited form as:

*J Appl Gerontol.* 2005 ; 24(3): 211–230.

## The Revised Observed Tasks of Daily Living:

### A Performance-Based Assessment of Everyday Problem Solving in Older Adults

**Manfred Diehl, Michael Marsiske, and Ann L. Horgas**

*University of Florida*

**Adrienne Rosenberg,**

*Hebrew Rehabilitation Center for Aged*

**Jane S. Saczynski, and**

*National Institute on Aging*

**Sherry L. Willis**

*Pennsylvania State University*

### Abstract

The Revised Observed Tasks of Daily Living (OTDL-R), a performance-based test of everyday problem solving, was administered to a sample of community-dwelling older adults. The OTDL-R included nine tasks, representing medication use, telephone use, and financial management. The OTDL-R had a desirable range of difficulty and satisfactory internal consistency and showed a relatively invariant pattern of relations between measured tasks and the underlying latent dimensions they represent across White and non-White subsamples. The OTDL-R also correlated significantly with age, education, self-rated health, a paper-and-pencil measure of everyday problem solving, and measures of basic cognitive functioning. Thus, the OTDL-R is a reliable and valid objective measure of everyday problem solving that has great practical utility for assessing performance in diverse populations.

### Keywords

everyday problem solving; everyday competence; performance-based assessment; adulthood; aging

---

A number of investigations in the aging literature have considered the concepts of *everyday cognition* and *everyday problem solving* (Allaire & Marsiske, 1999; Berg & Klaczynski, 1996; Blanchard-Fields & Chen, 1996; Willis et al., 1998). Research on this topic has generally been motivated by the concern that traditional laboratory-based measures of everyday cognition, which are often context-free, may deny older adults the benefits of experience and familiarity that they bring to their daily lives (e.g., Dixon & Baltes, 1986; Sternberg, 1997). From a scientific perspective, a resulting question is whether typical studies of cognitive aging in the laboratory might overestimate the magnitude of age-related impairments, in part because they eliminate or reduce potentially supportive effects of the real-world context. From an applied perspective, a question is whether more ecologically valid measures of cognition might provide more accurate estimates of elders' functional competencies in their daily lives (see Baddeley, 1989; Bahrck, 1989; Landauer, 1989).

The need for objective measures of adults' everyday functioning is great. The largest proportion of gerontological research in this area has relied on self-report measures (e.g., Fillenbaum,

1988; Lawton & Brody, 1969) that, although widely accepted and predictive of outcomes like morbidity and mortality (Branch & Jette, 1982; Wolinsky, Coe, Miller, & Prendergast, 1984), are also subject to self-evaluation biases, including both overestimation and underestimation of functioning and lack of congruence with proxy or clinician ratings (Diehl, 1998; Kuriansky, Gurland, Fleiss, & Cowan, 1976; Rogers & Holm, 1990; Rubenstein, Schairer, Wieland, & Kane, 1984).

Consequently, a subset of the everyday cognition research with older adults has focused directly on the objective measurement of the cognitive components of everyday task performance (e.g., Allaire & Marsiske, 1999; Cornelius & Caspi, 1987; Denney & Pearce, 1989). Although this literature shows substantial heterogeneity with regard to the domains examined and measurement approaches used (Diehl, 1998; Marsiske & Willis, 1995), one commonality has been the use of paper-and-pencil and interview-based simulations of everyday problems. In most studies, participants are presented with a prototypical problem and asked to identify or generate solutions to the problem. In general, research on everyday cognition has not employed direct observation of older adults' performance of real-world tasks. This is a potential limitation of the research for several reasons. First, the paper-and-pencil or interview modes of administration introduce an artificiality and unfamiliarity that may influence elders' performance in a negative way. Second, by increasing the reading, writing, and speaking load of simulated everyday tasks, these administration modes may introduce education and literacy biases into everyday task performance. Third, paper-and-pencil and interview approaches cannot speak directly to the efficiency of behavioral execution in everyday problem situations, which actual behavioral observations come closer to simulating.

In the more recent psychological and gerontological literature, several groups of investigators (e.g., Diehl, Willis, & Schaie, 1995; Loewenstein et al., 1989; Mahurin, DeBettignies, & Pirozzolo, 1991; Owsley, Sloane, McGwin, & Ball, 2002) have implemented performance-based approaches to measuring activities of daily living (ADLs; Katz, Ford, Moskowitz, Jackson, & Jaffe, 1963) and instrumental activities of daily living (IADLs; Lawton & Brody, 1969). These studies have shown that performance-based measures are related to older adults' self-reported IADL performance and that they show substantial associations with measures of basic cognitive functioning (Diehl et al., 1995).

In the Diehl et al. (1995) study, the investigators used the Observed Tasks of Daily Living (OTDL), a measure they had developed to assess older adults' performance on selected IADLs. They presented participants with a set of 31 tasks drawn from the domains of food preparation (e.g., preparing cake mix), medication use (e.g., following medicine bottle instructions), and telephone use (e.g., using a telephone rate chart). Results showed that performance on the OTDL was characterized by substantial interindividual variation (fewer than 50% of participants were able to perform most tasks in a completely correct manner) and could be dimensionalized (via confirmatory factor analysis) into three reliable but highly related scales (i.e., food preparation, medication use, and telephone use). The OTDL was strongly related to a paper-and-pencil measure of everyday cognition, the Everyday Problems Test (EPT; Willis & Marsiske, 1993), and was also significantly related to a number of traditional measures of psychometric intelligence (ranging from  $r = .33$  with memory to  $r = .68$  with fluid intelligence; see Diehl et al., 1995).

The earlier research employing the OTDL had several limitations that were addressed by the current study. Scientifically, the original study used a small, racially homogeneous, and economically and educationally advantaged sample of older adults. Pragmatically, the original OTDL was very time consuming (31 tasks with an average completion time between 40 min to 60 min) and was, therefore, of limited usefulness in applied or research settings.

The present study preceded the Advanced Cognitive Training for Independent and Vital Elderly (ACTIVE; see Ball et al., 2002; Jobe et al., 2001) clinical trial and was designed to address these limitations. The ACTIVE trial was a randomized, double-blind trial designed to determine whether cognitive interventions (memory, reasoning, and speed of processing training), which have previously been found to be successful at improving functioning under laboratory or small-scale field conditions, can affect measures of daily functioning. The present study focused on three research questions. First, are the psychometric characteristics (reliability, factor structure) of the OTDL adequate when a short form (taking about 20 min to complete) is used? Second, do the OTDL maintain their psychometric adequacy in a more heterogeneous study population? Third, do the revised and shortened OTDL show criterion validity; that is, do the OTDL show substantial correlations with other measures of cognition, everyday problem solving, health, and functioning?

## Method

### Participants

Participants were drawn from the ACTIVE Pilot Study<sup>1</sup>, a multisite investigation of cognition, everyday functioning, health, and well-being, and consisted of 170 older adults (141 women, 29 men). Because of incomplete data, 44 participants were excluded from the analyses. The resulting sample of 126 individuals was ethnically heterogeneous and included 53 White participants and 73 non-White participants (65 participants were of African American descent). *T*-test comparisons revealed that the participants with complete data did not differ significantly in age or education from the participants with incomplete data. The sample was drawn from six geographic catchment areas in the United States (i.e., Alabama, Indiana, Maryland, Massachusetts, Michigan, and Pennsylvania). The mean age of participants was 74.6 years (*SD* = 6.00 years, range = 65 to 90 years). Mean educational level was 12 years of schooling (*SD* = 2.99).

We investigated whether there were significant differences between White and non-White participants in terms of age and education. Because the proportion of participants from each racial group varied by study site, site and a site-by-race interaction term were included as covariates in these analyses. White and non-White participants did not differ significantly in age,  $F(1, 122) = .24, p > .05$ , or education,  $F(1, 120) = 1.39, p > .05$ .

### Procedure

Participants in the ACTIVE Pilot Study were selected from a larger pool of potential participants identified by each field site. Specifically, participants in the pilot study underwent a two-step screening process. First, potential participants at each of the six field sites were telephoned and invited to participate in a brief telephone screening. Participants younger than 65 years of age; with substantial self-reported hearing, vision, or cognitive impairments; with diagnoses of terminal medical conditions; who were unable to communicate with interviewers; and with travel plans conflicting with the study schedule were excluded from the study.

---

<sup>1</sup>In the summer of 1997, the Advanced Cognitive Training for Independent and Vital Elderly (ACTIVE) Steering Committee conducted a feasibility and pilot study to evaluate the measures proposed for study inclusion as well as the proposed intervention programs. Because the current study focuses on the psychometric properties of the OTDL-R, data from the ACTIVE Pilot Study are reported. The ACTIVE trial has been funded by the National Institute on Aging (NIA) and the National Institute of Nursing Research (NINR) and is coordinated by a Coordinating Center based at the New England Research Institutes (NERI). Participating field sites include the University of Alabama–Birmingham (UAB), Hebrew Rehabilitation Center for the Aged (HRCA), Indiana University (IU), Johns Hopkins University (JHU), Pennsylvania State University (PSU), and University of Florida (UF)/Wayne State University (WSU). The ACTIVE Steering Committee consists of Karlene Ball (UAB), Kathy Mann-Koepke (NINR), Jeffrey Elias (NIA), Michael Marsiske (UF/WSU), John N. Morris (HRCA), George W. Rebok (JHU), Fred Unverzagt (IU), Sharon L. Tennstedt (NERI), and Sherry L. Willis (PSU) and includes a number of multidisciplinary coinvestigators at each of the field sites. See Jobe et al. (2001) and Ball et al. (2002) for a more complete description.

Interested, eligible participants were invited to an in-person screening visit. Participants with low cognitive status (i.e., scores of 23 or lower on the Mini Mental Status Examination; Folstein, Folstein, & McHugh, 1975) or severe objective vision performance (i.e., binocular distance visual acuity measured at 20/50 or lower) were excluded. All other participants were invited to enroll in the ACTIVE Pilot Study.

Eligible participants attended two testing sessions. In the first session, conducted with a single participant and an interviewer, selected measures of cognition, everyday functioning, and self-reported health and driving were administered. The OTDL-R was administered during this session. Administration of all measures lasted typically 90 min to 120 min. In the second session, small groups of participants (on average 4 to 6 elders) completed measures of psychometric intelligence and cognition as well as several self-report measures of mobility and functioning. On average, this session required 3 hrs to complete.

All interviewers/testers were trained and certified using a multistep protocol. First, all interviewers attended a 5-day workshop focusing on the study assessment protocol. This workshop was conducted by the Coordinating Center, the New England Research Institutes. As part of this workshop, interviewers performed many practice assessments, both with one another and with volunteer elders. Second, at the field sites, interviewers performed a prescribed number of practice assessments with mock participants, including several elders. These practice sessions were conducted under the observation of study coordinators and investigators. Interviewers received verbal and written feedback after each practice assessment. Third, after certification, interviewers received regular quality-control observations in the field sites with immediate feedback regarding correct and incorrect adherence to study protocol.

## Measures

### Revised Observed Tasks of Daily Living (OTDL-R)

The OTDL (Diehl et al., 1995) is a behavioral measure of everyday competence that requires adults to perform a set of discrete, observable actions in response to a question by an interviewer. The OTDL was revised for inclusion in this study. Specifically, one of the domains from the original OTDL (i.e., food preparation) was replaced with the domain of financial management. This change was made to allow the OTDL-R to better reflect the cognitive ADLs identified by Wolinsky and Johnson (1991), which have been shown to predict bed-disability days, hospital contact, and mortality.

The OTDL was also shortened to include a total of nine tasks, three in each of the domains of medication use (following medicine label directions, understanding an aspirin leaflet, and completing a patient record form); telephone use (finding and dialing a number from the yellow pages, finding and dialing a number from a directory of social service resources from the phone book, and using a rate discount chart from a phone book); and financial management (making change with coins and bills, balancing a checkbook, and paying a utility bill with a check and mailing it). Each main task in the OTDL-R (e.g., finding and dialing a number from the yellow pages) has a number of subtasks or steps (e.g., first finding the number), yielding a total of 28 items on which scores were assigned. Analyses that follow are based on these items (i.e., subtask scores). For each task, participants received real-life materials (e.g., medicine bottles, a push-button phone, a lunch bill and coins) and a question that the interviewer presented on a 4 × 5 index card. Participants were then asked to perform the necessary steps to complete the task. The mean completion time for the OTDL-R was 26.5 min ( $SD = 8.6$  min; range = 10 min to 63 min). A full version of the OTDL, with administration and scoring instructions and training videos, is available at <http://www.phhp.ufl.edu/marsiskelab/otdl/otdl.html>.

Interviewers were trained in the administration of a standardized observation and data recording protocol. All interviewers received a detailed question-by-question manual for the OTDL-R and also viewed an instructional videotape about test administration. Prior to administering the OTDL-R to participants, interviewers had multiple practice opportunities and were certified according to a quality-control protocol monitored by the study's Coordinating Center. A random sample of 5% of all OTDL-Rs was also collected and examined by the study's investigators to ensure fidelity to the data recording and scoring protocol.

OTDL-R protocols were scored by independent, masked, and certified scorers. Maximum item scores could range from 1 to 4 (depending on the item), with higher scores reflecting better performance. For the purposes of the following analyses, participants' scores on each of the 28 items were recoded into a correct (1) or incorrect (0) format.

### **Additional Measures**

From the broader battery of ACTIVE pilot measures, the current study included measures of basic cognitive functioning (e.g., verbal ability, memory, inductive reasoning, and perceptual speed), a paper-and-pencil test of everyday problem solving (EPT; Willis & Marsiske, 1993), selected demographic variables (age, education, gender, and race), and a measure of general health. Table 1 provides a summary of all measures used in this study.

### **Results**

Results are reported in four sections. First, we will report findings with regard to the item and task difficulty of the OTDL-R. Second, we will report results on the reliability of the revised measure. Third, we will present findings with regard to the factor structure and the invariance of the factor structure of the revised measure across racial groups. Finally, we will present data showing the associations of OTDL-R performance with age, health, and measures of cognitive functioning.

#### **Item and Task Difficulty**

Analyses examining the difficulty of the behavioral steps involved in each OTDL showed that the items covered all difficulty levels. To be more specific, the difficulty level ranged from easy items (92% correct responses) to difficult items (24% correct responses). The mean item difficulty (the percentage of correct responses) for medication use items, telephone use items, and financial management items was 57%, 65%, and 72%, respectively.

Based on the correct-incorrect scoring format, individual task difficulties (i.e., the sum of all items within each task) ranged from easy (86.5% correct responses for financial management Task 1) to difficult (38.5% correct responses for telephone use Task 3). The mean task difficulty across all OTDL-R tasks was 64.7% correct responses. Table 2 shows the mean task difficulties and the mean task difficulty for each behavioral domain.

#### **Reliability**

Coefficients of internal consistency (Cronbach's  $\alpha$ ) were calculated to examine the reliabilities of the three scales and the total OTDL-R. Because items were scored dichotomously, the Kuder-Richardson 20 formula (Pedhazur & Schmelkin, 1991) was used to calculate  $\alpha$ .<sup>2</sup> The resulting  $\alpha$ s were .71 for medication use, .66 for telephone use, and .71 for financial management (see Table 2). The internal consistency for the total measure was .82.

---

<sup>2</sup>Please note that Kuder-Richardson 20 alphas tend to be lower than Cronbach's alpha because of the dichotomous item scoring. Thus, they represent a lower bound estimate of the scales' reliabilities.

## Factor Structure

LISREL 8 (Jöreskog & Sörbom, 1996) was used to perform confirmatory factor analyses examining the factor structure (i.e., the measurement model) of the OTDL-R. Two competing factor solutions were tested. First, because work with the original OTDL had provided support for a multidimensional structure of the measure (Diehl et al., 1995), we tested a three-factor model, specifying a separate factor for each domain of daily living. Second, because the scales of the OTDL have been shown to be moderately to highly correlated, we also tested a single-factor model as a potentially more parsimonious representation of the observed data.

The specified three-factor model provided a good fit to the observed data,  $\chi^2 (24, N = 126) = 16.50, p = .87$ , root mean square error of approximation (RMSEA) = .00, root mean square residual (RMR) = .01, goodness of fit index (GFI) = .97, adjusted goodness of fit index (AGFI) = .94, and comparative fit index (CFI) = .99. The estimated factor loadings of the individual tasks (i.e., standardized solution), the unique variances, and the factor intercorrelations are shown in Figure 1. As can be seen in Figure 1, all estimated factor loadings were significantly different from zero and ranged in size from .32 to .76. The three factors of the OTDL-R were highly correlated; correlations ranged from .78 to .91.

The high factor intercorrelations suggested that the scales of the OTDL-R shared a great deal of variance and that a single-factor solution may be a more parsimonious solution to account for the variability in the observed data. The single-factor model also provided a good fit to the observed data,  $\chi^2 (27, N = 126) = 20.96, p = .79$ , RMSEA = .00, RMR = .05, GFI = .96, AGFI = .94, CFI = 1.00. Moreover, the single-factor model did not result in a significant loss of goodness of fit as examined by the change in chi-square between the two competing models,  $\Delta\chi^2 (3) = 4.46, p > .05$ . The loadings and uniquenesses of the OTDL tasks on this single, general factor were almost identical (typically less than .05 different) to those reported in Figure 1. The only exception was the making change item, which had a loading of .49 on a specific financial management factor but had a loading of .29 on the general factor.

## Invariance of Factor Structure

When measures are used in qualitatively different groups of individuals (e.g., men vs. women, Whites vs. non-Whites), it is essential to assure that they measure the constructs of interest in the same way in each group. That is, it is important to examine that the underlying factor structure of a measure remains invariant across groups so that quantitative group comparisons can be considered meaningful (cf. Baltes & Nesselroade, 1973; Cunningham, 1991; Schaie, Willis, Jay, & Chipuer, 1989).

To examine the invariance of the factor structure of the OTDL-R across racial groups, we performed three different analyses (i.e., confirmatory factor analyses) separately for the three-factor model and the one-factor model. The analyses progressed from strong definitions of invariance to less stringent definitions of invariance. To be more specific, the analyses tested complete metric invariance (i.e., loadings and factor variances constrained to be equal across groups), loading invariance (i.e., only loadings constrained to be equal across groups), and configural invariance (i.e., same loading pattern across groups, but loadings not constrained to be equal across groups).

For the three-factor model, findings from these analyses supported configural invariance. The overall goodness of fit of the configural invariance model was good,  $\chi^2 (48, N = 126) = 49.55, p = .41$ , RMSEA = .00, CFI = .99. The goodness of fit in each racial group was good and slightly better in the non-White group (Whites: GFI = .90, RMR = .08; non-Whites: GFI = .94, RMR = .06). The completely standardized solutions within racial groups are shown in Table 3. Although the additional constraints associated with the invariance of the factor loadings and

with metric invariance did not result in significant losses of goodness of fit, some parameter estimates (i.e., factor intercorrelations) exceeded legitimate values, suggesting that the imposed constraints were untenable.

For the single-factor model, findings supported complete metric invariance across the two racial groups,  $\chi^2(63, N = 126) = 73.79, p = .17, RMSEA = .02, CFI = .93$ . Again, the goodness of fit was lower in the White subsample (GFI = .84; RMR = .02) and higher in the non-White sample (GFI = .93, RMR = .01).

### Correlates of OTDL Performance

The correlations of older adults' OTDL-R performance with age, educational level, general health, a paper-and-pencil measure of everyday problem solving, and basic cognitive abilities—and of these correlates with one another—were examined. A typical pattern of findings (i.e., significant negative correlations between most cognitive measures and age; significant positive correlations between most cognitive measures and education) and positive correlations among these cognitive measures and with self-reported health (i.e., the MOS-SF36 General Health scale) indicated that this sample was typical and representative of relationships generally reported in the aging research literature.

With respect to the OTDL-R, performance showed a significant negative correlation with age ( $r = -.24, n = 124$ ) and significant positive correlations with education ( $r = .44, n = 126$ ), the SF-36 general health scale ( $r = .25, n = 118$ ), and another objective measure of everyday problem solving (i.e., EPT;  $r = .64, n = 119$ ). In addition, OTDL-R performance showed significant positive correlations with basic cognitive abilities such as verbal ability, memory, inductive reasoning, and perceptual speed (range:  $r = .37$  to  $.59$ , range of  $n$ : 123–126).<sup>3</sup>

### Discussion

The results of this study showed that the revised version of the OTDL, the OTDL-R, is a reliable and valid measure of older adults' everyday competence. To be more specific, the tasks from the three IADL domains of medication use, telephone use, and financial management cover a desirable range of difficulty and show satisfactory internal consistency. One advantage of the OTDL-R is that it includes a greater range of task difficulties than some other existing measures of everyday competence (e.g., Loewenstein et al., 1989) and may, therefore, be more useful for assessments with community-residing older adults. A second strength of the OTDL-R is that it includes those IADLs for which poor self-reported performance has been shown to be associated with greater use of health care services (Wolinsky et al., 1983), increased risk of institutionalization (Branch & Jette, 1982), and increased risk of mortality (Fillenbaum, 1988; Manton, 1988). Thus, we believe that the OTDL-R is a valuable supplement to the existing self-report measures and provides gerontologists with a useful tool for research and applied purposes. The OTDL-R could be a useful alternative to paper-and-pencil or interview approaches to assess performance of older adults, especially in outpatient or in-home settings. Advantages of the tool include the reduction of biases related to education and literacy as well as allowing for actual behavioral observations. We note, in particular, that some elders (e.g., those with low literacy, severe arthritis, or patients who undergo rehabilitation after a stroke) might be compromised in their ability to complete paper-and-pencil tests, and the OTDL-R might serve as a practical alternative.

We note, however, that the clinical utility of the OTDL-R has not yet been demonstrated. We speculate that the measure could be used clinically in at least three contexts: (a) to validate

<sup>3</sup>A table with the zero-order correlations from this analysis can be obtained on request from the first author.

self-report, particularly to follow up those who are reporting functional or cognitive limitations; (b) to provide an independent assessment of performance in individuals for whom self-report is considered to have questionable validity (e.g., persons with depression, anosognosia, or who present in the early stages of progressive cognitive impairment); and (c) to serve as a face-valid adjunct to traditional and more abstract cognitive screening tools. We emphasize, however, that clinical research with the OTDL-R is a critical next step to validate this instrument further.

Related to the clinical uses of the OTDL, a concern for practitioners may be the amount of training and study that the OTDL-R requires. It is difficult to provide a concrete estimate of training time, but our impression is that the OTDL-R does not require much more preparation time than other widely used instruments (e.g., Mini Mental Status Examination, Older Adults' Resources and Services tool). To assist practitioners interested in exploring the tool, we have made a complete self-study protocol (including detailed instructions, kit assembly details, instructional videotapes, and a coding manual) available at the Web page that is maintained for the OTDL-R (<http://www.phhp.ufl.edu/marsiskelab/otdl/otdl.html>).

Confirmatory factor analyses showed that both a three-factor model and a single-factor model of everyday competence provided a good fit to the observed data. The three-factor model provided additional support for the multidimensionality of older adults' everyday problem solving and was consistent with findings reported by Diehl et al. (1995) on the original version of the OTDL. In contrast to the earlier findings, however, the intercorrelations among the latent factors in this study were considerably larger, suggesting that a more parsimonious representation of the factor structure may be psychometrically more appropriate.

Because the higher factor intercorrelations can be due to a number of different factors (e.g., smaller number of tasks compared to the original OTDL, different task domains, and different study sample), we consider it premature to abandon the multidimensional structure of the OTDL-R. Instead, we would like to advocate the continued use of scale-specific scores of everyday competence for practical purposes (e.g., applied and clinical research) as well as the parsimonious description of older adults' performance on everyday problem-solving tasks in the form of a summary score. The latter solution takes a psychometrically more rigorous approach and acknowledges that the factors in the current study shared a great deal of variance.

Although findings from these analyses are, in general, reassuring and attest to the psychometric soundness of the measure, they should still be considered preliminary. Future factor analytic work with larger independent samples will be required to show whether a similar pattern of factor loadings can be obtained and whether the high factor intercorrelations found in the present study generalize to other samples.

Examination of the factor structure across the two major racial groups (i.e., Whites vs. non-Whites) provided support for configural invariance for the three-factor solution and support for complete metric invariance for the single-factor solution. That is, for both factor solutions the assumption of identical factor loadings across racial subsamples was considered tenable. Thus, the results of this study suggest relative invariance of the factor structure of the OTDL-R across racial/ethnic groups with only small differences between groups. This is consistent with a larger body of research on behavioral and health factorial similarity across European American and African American older adults. Similar invariance findings (i.e., configural similarity between groups, with small group differences that can often be localized in specific factors) have been reported, for example, with regard to mental status (Rapp, Espeland, Hogan, Jones, & Dugan, 2003), depression (Blazer, Landerman, Hays, Simonsick, & Saunders, 1998; but see Callahan & Wolinsky, 1994, for a substantial exception), health status (Gibson, 1991; Stump, Clark, Johnson, & Wolinsky, 1997), menopausal symptoms (Avis et al., 2001), and a cancer-specific quality-of-life questionnaire (Kart & Ford, 2002).



Like the previous findings, these analyses should be interpreted with caution because of the relatively small sample. Specifically, with a larger sample, the study would have had greater power to detect significant differences in factor analytic parameters between racial subgroups. Although this evidence does not constitute a sufficient basis to claim that the tasks of the OTDL-R are culturally invariant, it is certainly encouraging to note that the pattern of item interrelationships was similar across racial groups. It also needs to be noted that these findings do not permit the conclusion that the factor structure of the OTDL-R extends to other racial groups such as Hispanics. At this point in time, the utility of the instrument with non-English speakers and non-U.S. cultures has not been examined. However, we recognize that this possibility should be explored in future research.

Correlational analyses revealed a significant negative association between OTDL-R performance and age. This means that older individuals tended to perform poorer than younger individuals. Of course, as is always true for such cross-sectional findings, true maturational differences are confounded with generational or cohort differences in relevant life and acculturation experiences in this association with age.

In addition, OTDL-R performance showed significant positive correlations with education and health, suggesting that a higher level of education and better health are associated with better everyday problem solving. Finally, performance on the OTDL-R was positively related with older adults' performance on tests that assessed basic cognitive functioning such as verbal ability, memory, inductive reasoning, and perceptual speed. These findings are consistent with previous work with the OTDL (Diehl et al., 1995) and consistent with work conducted with other measures of everyday problem solving (Allaire & Marsiske, 2002; Marsiske & Willis, 1995; Willis & Schaie, 1986). In general, a growing body of empirical studies supports the notion that cognition is an important component of older adults' everyday functioning and provides support for a hierarchical model of everyday cognition (Willis & Schaie, 1993). The core premise of this model is that cognition in everyday contexts is likely to be related to basic or laboratory-based measures of psychometric intelligence and cognition (Allaire & Marsiske, 2002; Willis & Schaie, 1993). In this framework, basic or academic intelligence is seen as a resource set that is drawn on in the management of everyday challenges. Moreover, because everyday problems are likely to be complex and multidimensional, the model also hypothesizes that broad measures of everyday problem solving (like the OTDL-R) will be related to multiple intellectual abilities. In general, we find the current results to be strongly supportive of these assertions.

In conclusion, we believe that the OTDL-R has great practical usefulness. It is a highly standardized, behavioral performance measure that can be used with persons of low literacy or with persons with other impairments that might compromise their ability to take paper-and-pencil tests. As a behavioral performance measure, it has high face validity (both for practitioners and elders themselves) and is less susceptible to a common criticism directed at many paper-and-pencil measures of everyday problem solving (i.e., that it shares substantial method variance with traditional psychometric intelligence tests).

An interesting next question for the OTDL-R, given its strong relatedness to psychometric intellectual abilities, is whether performance on the measure will be sensitive to cognitive interventions. The ACTIVE clinical trial, a study of the effects of cognitive interventions with older adults, will permit an examination of this question. The ACTIVE trial investigates the longitudinal effects of several cognitive interventions on study participants' health and well-being, including their everyday functioning. In this context, an important question will be whether the OTDL-R can uniquely predict long-term functional outcomes for older adults.

### Acknowledgements

AUTHORS' NOTE: This research was supported by grants U01AG14260, U01AG14263, U01AG14276, U01AG14282, U01AG14289, U01NR04507, and U01NR04508 from the National Institute on Aging and the National Institute of Nursing Research. We thank our many colleagues in the ACTIVE clinical trial for their assistance.

### Biography

Manfred Diehl's program of research focuses on two main areas of adult development and aging: (a) development of self and personality across the adult life span and (b) everyday problem solving, everyday competence, and functional health in late life. His recent publications include "Agency and Communion Attributes in Adults' Spontaneous Self-Representations" (coauthored with S. Owens & L. Youngblade), published in the *International Journal of Behavioral Development*. He also contributed a chapter (coauthored with S. L. Willis) on "Everyday Competence and Everyday Problem Solving in Aging Adults: Role of the Physical and Social Context" to a recent volume of the *Annual Review of Gerontology and Geriatrics*.

Michael Marsiske's research focuses on three major areas: (a) evaluation of training and practice interventions to enhance cognitive performance in later adulthood, (b) assessment of everyday problem solving in older adults, and (c) the interrelation between cognitive changes with age and sensorimotor functioning. He is the author (with Jennifer A. Margrett) of the chapter on everyday problem solving and decision making in the forthcoming *Handbook of the Psychology of Aging* and has published several recent papers on intraindividual variability in older adults and training and practice in later life.

Ann L. Horgas's program of research focuses on chronic pain in late life and its impact on physical health, mental health, and everyday functioning. She has done extensive research in nursing homes and assisted living facilities, focusing primarily on strategies for improving care, managing pain, and promoting physical functioning and well-being.

Adrienne Rosenberg's current research focuses on (a) the development of a prospective payment system for inpatient psychiatric services, (b) testing cognitive interventions in later adulthood, and (c) implementation of a program to optimize the physical, mental, and social well-being of older adults living in continuing care retirement communities.

Jane S. Saczynski's current research interests are in three major areas: (a) physical health, brain aging, and the plasticity or improvement of cognitive performance in adulthood; (b) risk factors associated with cognitive impairment and protective mechanisms for cognitive resiliency; and (c) cognitive strategy use and intellectual performance.

Sherry L. Willis's research focuses on adult cognitive development and cognitive training, midlife development, everyday problem solving in late life, and antecedents of cognitive impairment. Recent publications include "Technology and Learning in Current and Future Elder Cohorts" in R. Pew & S. Van Hemel's *Technology for Adaptive Aging*. She is also coeditor of a volume on midlife development (with Mike Martin) to be published by Sage Publications.

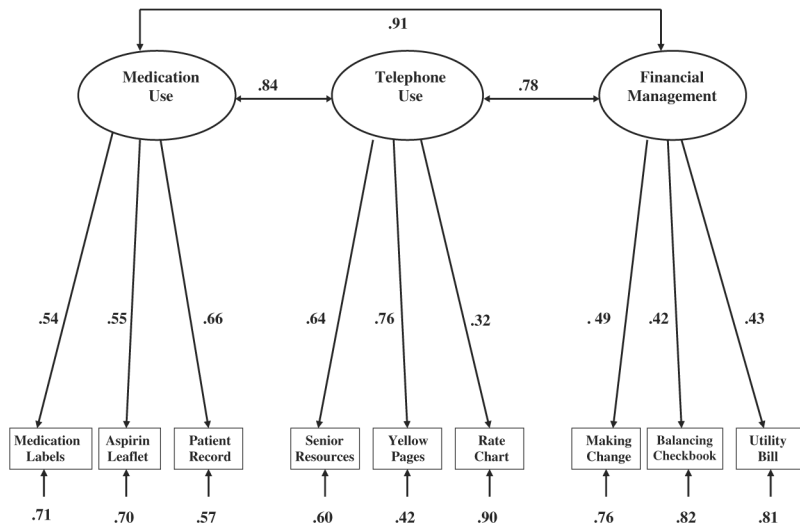
### References

- Allaire JC, Marsiske M. Everyday cognition: Age and intellectual ability correlates. *Psychology and Aging* 1999;14:627–644. [PubMed: 10632150]
- Allaire JC, Marsiske M. Well- and ill-defined measures of everyday cognition: Relationship to older adults' intellectual ability and functional status. *Psychology and Aging* 2002;17:101–115. [PubMed: 11931279]

- Avis NE, Stellato R, Crawford S, Bromberger J, Ganz P, Cain V, et al. Is there a menopausal syndrome? Menopausal status and symptoms across racial/ethnic groups. *Social Science & Medicine* 2001;52:345–356. [PubMed: 11330770]
- Baddeley, A. Finding the bloody horse. In: Poon, LW.; Rubin, DC.; Wilson, BA., editors. *Everyday cognition in adulthood and late life*. Cambridge University Press; New York: 1989. p. 104-115.
- Bahrick, HP. The laboratory and ecology: Supplementary sources of data for memory research. In: Poon, LW.; Rubin, DC.; Wilson, BA., editors. *Everyday cognition in adulthood and late life*. Cambridge University Press; New York: 1989. p. 73-83.
- Ball K, Berch DB, Helmers KF, Jobe JB, Leveck MD, Marsiske M, et al. Effects of cognitive training interventions with older adults: A randomized controlled trial. *Journal of the American Medical Association* 2002;288:2271–2281. [PubMed: 12425704]
- Baltes, PB.; Nesselroade, JR. The developmental analysis of individual differences on multiple measures. In: Nesselroade, JR.; Reese, HW., editors. *Life-span developmental psychology: Methodological issues*. Academic Press; New York: 1973. p. 219-252.
- Berg, CA.; Klaczynski, PA. Practical intelligence and problem solving. In: Blanchard-Fields, F.; Hess, TM., editors. *Perspectives on cognitive change in adulthood and aging*. McGraw-Hill; New York: 1996. p. 323-357.
- Blanchard-Fields F, Chen Y. Adaptive cognition and aging. *American Behavioral Scientist* 1996;39:231–248.
- Blazer DG, Landerman LR, Hays JC, Simonsick EM, Saunders WB. Symptoms of depression among community-dwelling elderly African-American and White older adults. *Psychological Medicine* 1998;28:1311–1320. [PubMed: 9854272]
- Branch LG, Jette AM. A prospective study of long-term care institutionalization among the aged. *American Journal of Public Health* 1982;72:1373–1379. [PubMed: 6814269]
- Brandt J. The Hopkins Verbal Learning Test: Development of a new memory test with six equivalent forms. *Clinical Neuropsychologist* 1991;5:125–142.
- Callahan CM, Wolinsky FD. The effect of gender and race on the measurement properties of the CES-D in older adults. *Medical Care* 1994;32:341–356. [PubMed: 8139299]
- Cornelius SW, Caspi A. Everyday problem solving in adulthood and old age. *Psychology and Aging* 1987;2:144–153. [PubMed: 3268204]
- Cunningham, WR. Issues in factorial invariance. In: Collins, LM.; Horn, JL., editors. *Best methods for the analysis of change: Recent advances, unanswered questions, future directions*. American Psychological Press; Washington, DC: 1991. p. 106-113.
- Denney NW, Pearce KA. A developmental study of practical problem solving in adults. *Psychology and Aging* 1989;4:438–442. [PubMed: 2619950]
- Diehl M. Everyday competence in later life: Current status and future directions. *The Gerontologist* 1998;38:422–433. [PubMed: 9726129]
- Diehl M, Willis SL, Schaie KW. Everyday problem solving in older adults: Observational assessment and cognitive correlates. *Psychology and Aging* 1995;10:478–491. [PubMed: 8527068]
- Dixon, RA.; Baltes, PB. Toward life-span research on the functions and pragmatics of intelligence. In: Sternberg, RJ.; Wagner, RK., editors. *Practical intelligence: Nature and origins of competence in the everyday world*. Cambridge University Press; New York: 1986. p. 203-235.
- Ekstrom, RB.; French, JW.; Harman, HH.; Dermen, D. *Kit of factor-referenced cognitive tests*. Educational Testing Service; Princeton, NJ: 1976.
- Fillenbaum, GG. *Multidimensional functional assessment of older adults: The Duke Older Americans Resources and Services Procedures*. Lawrence Erlbaum; Hillsdale, NJ: 1988.
- Folstein MF, Folstein SE, McHugh PR. “Mini mental state”: A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research* 1975;12:189–198. [PubMed: 1202204]
- Gibson RC. Race and the self-reported health of elderly persons. *Journals of Gerontology* 1991;46:S235–S242. [PubMed: 1890294]
- Jobe JB, Smith DM, Ball K, Tennstedt SL, Marsiske M, Rebok GR, et al. ACTIVE: A cognitive intervention trial to promote independence in older adults. *Controlled Clinical Trial* 2001;22:453–479.

- Jöreskog, K.; Sörbom, D. LISREL 8: User's reference guide. Scientific Software International; Chicago: 1996.
- Kart CS, Ford ME. Exploring the factorial structure of the EORTC QLQ-C30: Racial differences in measuring health-related quality of life in a sample of urban, older adults. *Journal of Aging and Health* 2002;14:399–421. [PubMed: 12146514]
- Katz S, Ford AB, Moskowitz RW, Jackson BS, Jaffe MW. Studies of illness in the aged. The index of ADL: A standardized measure of biological and psychosocial function. *Journal of the American Medical Association* 1963;185:914–919. [PubMed: 14044222]
- Kuriansky J, Gurland B, Fleiss JL, Cowan D. The assessment of self-care capacity in geriatric patients by objective and subjective methods. *Journal of Clinical Psychology* 1976;32:95–102. [PubMed: 1249244]
- Landauer, TK. Some bad and some good reasons for studying memory and cognition in the wild. In: Poon, LW.; Rubin, DC.; Wilson, BA., editors. *Everyday cognition in adulthood and late life*. Cambridge University Press; New York: 1989. p. 116-125.
- Lawton MP, Brody EM. Assessment of older people: Self-maintaining and instrumental activities of daily living. *The Gerontologist* 1969;9:179–185. [PubMed: 5349366]
- Loewenstein DA, Amigo E, Duara R, Guterman A, Hurwitz D, Berkowitz N, et al. A new scale for the assessment of functional status in Alzheimer's disease and related disorders. *Journal of Gerontology: Psychological Sciences* 1989;44:P114–P121.
- Mahurin RK, DeBettignies BH, Pirozzolo FJ. Structured assessment of independent living skills: Preliminary report of a performance measure of functional abilities in dementia. *Journal of Gerontology: Psychological Sciences* 1991;46:P58–P66.
- Manton KG. A longitudinal study of functional change and mortality in the United States. *Journal of Gerontology: Social Sciences* 1988;43:S153–S161.
- Marsiske M, Willis SL. Dimensionality of everyday problem solving in older adults. *Psychology and Aging* 1995;10:269–283. [PubMed: 7662186]
- Owsley C, Sloane M, McGwin G, Ball K. Timed instrumental activities of daily living tasks: Relationship to cognitive function and everyday performance assessments in older adults. *Gerontology* 2002;48:254–265. [PubMed: 12053117]
- Pedhazur, EJ.; Schmelkin, LP. *Measurement, design, and analysis: An integrated approach*. Lawrence Erlbaum; Hillsdale, NJ: 1991.
- Rapp SR, Espeland MA, Hogan P, Jones BN, Dugan E. Baseline experience with Modified Mini Mental State Exam: The Women's Health Initiative Memory Study (WHIMS). *Aging & Mental Health* 2003;7:217–223. [PubMed: 12775404]
- Rey A. L'examen psychologique dans les cas d'encephalopathie traumatique. *Archives de Psychologie* 1941;28:21.
- Rogers, JC.; Holm, MB. Measuring functional status outcomes: Objective and subjective methods; Paper presented at the Annual American Congress of Rehabilitation Medicine; Phoenix, AZ. 1990, October;
- Rubenstein LZ, Schairer C, Wieland GD, Kane R. Systematic biases in functional status assessment of elderly adults: Effects of different data sources. *Journal of Gerontology* 1984;39:686–691. [PubMed: 6436360]
- Schaie, KW. *Schaie-Thurstone Adult Mental Abilities Test*. Consulting Psychologists Press; Palo Alto, CA: 1985.
- Schaie KW, Willis SL, Jay G, Chipuer H. Structural invariance of cognitive abilities across the life span: A cross-sectional study. *Developmental Psychology* 1989;25:652–662.
- Sternberg RJ. The concept of intelligence and its role in lifelong learning and success. *American Psychologist* 1997;52:1030–1037.
- Stump TE, Clark DO, Johnson RJ, Wolinsky FD. The structure of health status among Hispanic, African American, and White older adults. *Journals of Gerontology: Psychological and Social Sciences* 1997;52B:49–60.
- Ware JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36): Conceptual framework and item selection. *Medical Care* 1992;30:473–483. [PubMed: 1593914]
- Wechsler, D. *WAIS-R: Wechsler Adult Intelligence Scale-Revised*. The Psychological Corporation; New York: 1981.

- Willis SL, Allen-Burge R, Dolan MM, Bertrand RM, Yesavage J, Taylor JL. Everyday problem solving among individuals with Alzheimer's disease. *The Gerontologist* 1998;38:569–577. [PubMed: 9803645]
- Willis, SL.; Marsiske, M. *Manual for the Everyday Problems Test*. Pennsylvania State University; University Park: 1993.
- Willis, SL.; Schaie, KW. Practical intelligence in later adulthood. In: Sternberg, RJ.; Wagner, RK., editors. *Practical intelligence: Nature and origins of competence in the everyday world*. Cambridge University Press; New York: 1986. p. 236-268.
- Willis, SL.; Schaie, KW. Everyday cognition: Taxonomic and methodological considerations. In: Puckett, JM.; Reese, HW., editors. *Mechanisms of everyday cognition*. Lawrence Erlbaum; Hillsdale, NJ: 1993. p. 33-53.
- Wilson, B.; Cockburn, J.; Baddeley, AD. *The Rivermead Behavioral Memory Test: Manual*. Thames Valley Test Company; Suffolk, United Kingdom: 1985.
- Wolinsky FD, Coe RM, Miller DK, Prendergast JM. Measurement of the global and functional dimensions of health status in the elderly. *Journal of Gerontology* 1984;39:88–92. [PubMed: 6690593]
- Wolinsky FD, Coe RM, Miller DK, Prendergast JM, Creel MJ, Chavez MN. Health services utilization among the noninstitutionalized elderly. *Journal of Health and Social Behavior* 1983;24:325–337. [PubMed: 6668412]
- Wolinski FD, Johnson RJ. The use of health services by older adults. *Journal of Gerontology: Social Sciences* 1991;46:S345–S357.



**Figure 1.** Standardized Parameter Estimates for the Three-Factor Model of the Revised Observed Tasks of Daily Living

**Table 1**  
 Relevant Measures From the ACTIVE Test Battery

<i>Domain</i>	<i>Instrument</i>
General health	Medical Outcome Study: General Health (Ware & Sherbourne, 1992)
Everyday problem solving	Everyday Problems Test (Willis & Marsiske, 1993)
Verbal ability	ETS Vocabulary Test (Ekstrom, French, Harman, & Dermen, 1976)
Memory	Auditory Verbal Learning Test (Rey, 1941) Hopkins Verbal Learning Test (Brandt, 1991) Rivermead Behavioral Memory Test (Wilson, Cockburn, & Baddeley, 1985)
Inductive reasoning	Word Series (Schaie, 1985) Letter Sets (Ekstrom et al. 1976)
Perceptual speed	Wechsler Adult Intelligence Scales: Digit Symbols Test (Wechsler, 1981) Identical pictures (Ekstrom et al., 1976)

NOTE: ACTIVE = Advanced Cognitive Training for Independent and Vital Elderly.

**Table 2**  
Task Difficulties and Reliabilities by Behavioral Domain

<i>Domain</i>	<i>% Correct Responses</i>	<i>Reliability (Kuder-Richardson's Corrected <math>\alpha</math>, KR-20)</i>
Medication use		
Medication labels	44.6	
Aspirin leaflet	72.8	
Patient record form	63.2	
Total scale	57.4	.71
Telephone use		
Senior resources directory	78.2	
Yellow pages	77.0	
Rate discount chart	38.5	
Total scale	64.6	.66
Financial management		
Making change	86.5	
Balancing a checkbook	40.9	
Paying a utility bill	78.2	
Total scale	72.1	.71

NOTE: KR-20 = Kuder-Richardson 20 correction (see Pedhazur & Schmelkin, 1991, pp. 97–100).



**Table 3**  
 Testing Configural Invariance Across Racial Groups: Completely Standardized Solution Within Groups

Variable	Medication Use		Domain of Daily Living Telephone Use		Financial Management	
	White	Non-White	White	Non-White	White	Non-White
Medication use						
Medication labels	.53	.52				
Aspirin leaflet	.59	.59				
Patient record form	.49	.73				
Telephone use						
Senior resources directory			.95	.56		
Yellow pages			.59	.75		
Rate discount chart			.06	.40		
Financial management						
Making change					.11	.44
Balancing a checkbook					.41	.37
Paying a utility bill					.30	.56
<i>Factor Intercorrelations</i>						
Medication use	—	—	—	—	—	—
Telephone use	.55	.95	.41	—	—	—
Financial management	.93	.79		.62	—	—