Published in final edited form as:
*J Aging Health*. 2005 April ; 17(2): 148–171. doi:10.1177/0898264304274184.

# Caregiver Appraisals of Functional Dependence in Individuals With Dementia and Associated Caregiver Upset: Psychometric Properties of a New Scale and Response Patterns by Caregiver and Care Recipient Characteristics

**LAURA N. GITLIN**
*Thomas Jefferson University*

**DAVID L. ROTH**
*University of Alabama at Birmingham*

**LOUIS D. BURGIO**
*University of Alabama at Birmingham*

**DAVID A. LOEWENSTEIN**
*University of Miami*

**LARAINE WINTER**
*Thomas Jefferson University*

**LINDA NICHOLS**
*University of Tennessee*

**SOLEDAD ARGÜELLES**
*University of Miami*

**MARY CORCORAN**
*George Washington University*

**ROBERT BURNS**
*University of Tennessee*

**JENNIFER MARTINDALE**
*University of Tennessee*

## Abstract

**Objective—**To evaluate psychometric properties and response patterns of the Caregiver Assessment of Function and Upset (CAFU), a 15-item multidimensional measure of dependence in dementia patients and caregiver reaction.

**Method—**640 families were administered the CAFU (53% White, 43% African American, and 4% mixed race and ethnicity). We created a random split of the sample and conducted exploratory factor analyses on Sample 1 and confirmatory factor analyses on Sample 2. Convergent and discriminant validity were evaluated using Spearman rank correlation coefficients.

**Results—**A two-factor structure for functional items was derived, and excellent factorial validity was obtained. Convergent and discriminant validity were obtained for function and upset measures. Differential response patterns for dependence and caregiver upset were found for caregiver race, relationship, and care recipient gender but not for caregiver gender.

**Discussion—**The CAFU is easily administered, reliable, and valid for evaluating appraisals of dependencies and upsetting care areas.

**Keywords**

home care; Alzheimer's disease and related disorders; caregiving

---

*Although individuals* with suspected cognitive impairments undergo neurological testing for a diagnostic determination, an important component of the evaluative process is the assessment of functional capabilities. For a diagnosis of a dementia syndrome, the *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.; American Psychiatric Association, 1994) requires at least one impairment in daily functioning. Numerous performance-based and patient self-report scales have been developed to assess physical functioning in individuals with dementia (for comprehensive reviews of these measures, see Carswell & Spiegal, 1999; Gitlin, in press; Loewenstein & Mogosky, 1999). However, administering performance-based measures is time consuming, and previous research indicates that persons with dementia overestimate their degree of functional capacity (Kiyak, Teri, & Borson, 1994; Rubenstein, Schairer, Wieland, & Kane, 1984). Thus, it is common for family members to be asked to provide clinical information about the person's functional performance as part of the diagnostic process (Carswell & Spiegal, 1999). Also, family reports are used to inform ongoing treatment and monitoring decisions as well as evaluate the effectiveness of various pharmacological interventions (Carswell & Spiegal, 1999). Nevertheless, despite the important role of family members in the diagnostic and disease-monitoring process, it is unclear how best to systematically obtain functional information from caregivers.

An important related point is that with the progression of the disease and deterioration of functional capacity, most family caregivers provide hands-on assistance with instrumental and basic activities of daily living (IADLs and ADLs). This is particularly the case at the moderate to severe stages of the disease, in which caregivers often report providing around-the-clock supervision or direct care (Mahoney et al., 2003). Although previous research has shown the negative health and psychological consequences for family caregivers providing protracted care, little is known about the association between specific areas of functional dependence and family caregiver reaction to providing physical assistance. Thus, the development of a simple-to-use valid instrument that measures both the patient's level of dependence and the caregiver's reaction to providing assistance with daily activities has utility for both research and clinical practice. A measure that captures the physical dependence of persons with dementia (referred to here as care recipient) and caregiver reactions to providing assistance would be helpful to guide an intervention approach that addresses both the physical functional needs of the care recipient and the emotional upset of the caregiver. A measure that focuses on the functioning of both caregiver and care recipient is particularly important in view of new evidence that caregiver appraisals and approaches to managing dementia affect life quality of the dementia patient (McClendon, Smyth, & Neundorfer, 2004).

This study examines the underlying structure of a new measure, the Caregiver Assessment of Functional Dependence and Caregiver Upset (CAFU), to assess both the caregiver appraisals of the level of physical dependence in 15 daily activities of persons with dementia and the caregiver's reaction to or upset with providing assistance in each area. Our approach is modeled after the Revised Memory and Behavior Problem Checklist (RMBPC) developed by Teri et al. (1992). The RMBPC assesses both caregiver report of the frequency of occurrence of 24 behaviors and the caregiver's reaction to such occurrences. This measure is specific to memory-related, disruptive, and depressive behaviors, such as repetitive vocalization or sad affect, and does not examine physical dependence in IADLs and its impact on caregivers.

To evaluate the psychometric properties of the CAFU, this study uses baseline data collected as part of the National Institutes of Health's Resources for Enhancing Alzheimer's Caregiver

Health (REACH) initiative. REACH is a 6-year initiative that was funded to characterize and test the feasibility of the most promising innovative interventions for family caregivers of persons with Alzheimer’s disease and related disorders (ADRD). Six sites developed and implemented unique interventions (Birmingham, AL; Boston; Memphis, TN; Miami, FL; Palo Alto, CA; and Philadelphia), and a seventh site (Pittsburgh, PA) served as the coordinating center to assure standardization and quality oversight of study protocols, data management, and analysis. Three of these sites (Birmingham, Memphis, and Philadelphia) enrolled both African American and White caregivers and administered the instrument that is evaluated in this study.

# Method

## PARTICIPANTS

The participants in this study were 640 family caregivers who were enrolled in one of three REACH sites (Birmingham, *n* = 140; Memphis, *n* = 245; and Philadelphia, *n* = 255). Each site recruited family caregivers of individuals with ADRD from multiple community sites, health and social agency settings, primary care clinics, and physician offices. A special effort was made at each site to enroll White and African American caregivers. The final sample included 340 White caregiver–care recipient dyads (53%), 277 African American caregiver–care recipient dyads (43%), and 23 caregiver–care recipient dyads of mixed racial and ethnic identities (4%). African Americans composed at least 40% of the sample at all three sites. Caregivers were eligible for study participation if they were older than 21 years of age and lived with and provided an average of 4 hours of supervision or direct care each day for a relative with ADRD for at least the past 6 months. Caregivers were excluded if they were involved in another caregiver intervention study or had an acute illness that would prevent them from participating for at least 6 months. Caregivers were also excluded if their care recipients had a terminal or severe illness or disability that would prohibit them from participating in the interventions.

Care recipients had to have a medical diagnosis of probable ADRD or attain a score of less than 24 (out of 30 points) on the Mini-Mental State Examination (MMSE; Folstein, Folstein, & McHugh, 1975). Additionally, they had to have at least one limitation in basic activities of daily living (ADLs; Katz, Ford, Moskowitz, Jackson, & Jaffe, 1963) or two dependencies in their IADLs (Lawton & Brody, 1969) as reported by the caregiver. These participation criteria were designed to ensure that caregivers were involved in daily tasks that were potentially burdensome. Other requirements were more logistical and included having a telephone, planning to remain in the geographic area for at least 6 months, and caregiver competency in English.

## PROCEDURE

Each site interviewed potential participants by telephone using screening questions shared across the sites. Eligible caregivers were subsequently administered the REACH core battery of measures at home (Birmingham and Philadelphia) or a physician’s office (Memphis) as well as the specific items described below related to the level of IADL and ADL dependence and caregiver reaction to providing assistance. These specific items were placed in the REACH core battery and were administered at each site in the same order by trained interviewers who had been certified by the REACH coordinating center. Each site obtained local Institutional Review Board approval for its specific site, and informed consent was obtained from each participating caregiver prior to administering the REACH interview. The REACH coordinating center (University of Pittsburgh) periodically visited each site to ensure adherence to study protocols and confirm the exclusive use of REACH-trained and REACH-certified interviewers.

It also conducted monthly monitoring of enrollment and data processing for quality control purposes.

## MEASURES

**CAFU—**To develop the CAFU, we used eight items from Lawton and Brody's (1969) instrumental ADL scale (telephone, shopping, meal preparation, housework, laundry, travel, medicine, and finances) and seven items from the Self-Care subscale (bathing, dressing upper body, dressing lower body, toileting, grooming, eating, and getting in and out of bed) of the Functional Independence Measure (FIM; Hamilton & Fuhrer, 1987). For each IADL and ADL item, we determined level of assistance required using a modification of the 7-point FIM response options (7 = *complete independence* and 1 = *total assistance*). Our modification to the FIM response options was as follows. For each IADL and ADL item, caregivers were asked the following: “During the past week, has (name of care recipient) needed any kind of help with (name of activity)?” As in the FIM, if the care recipient did not need any assistance, the caregiver was asked if the care recipient needed more than a reasonable amount of time, the care recipient needed an assistive device, or the caregiver had a concern for the person's safety. If no, then a score of 7 for complete independence was recorded; if yes, a score of 6 for modified assistance was recorded. If the care recipient needed assistance, the type and level of assistance provided was asked. Caregivers who reported that their care recipient needed supervision, set-up, or cueing but no physical help received a score of 5. If physical help was needed, we modified the FIM response set by asking the following question: “Which category best describes the amount of help your care recipient needs for (name of activity): 4 = *a little help or 25% assistance*, 3 = *moderate help or 50% assistance*, 2 = *a lot of help or 75% assistance*, or 1 = *complete help or more than 75% help with the task*.” The use of standardized questions to obtain the caregiver's appraisal of the amount of physical assistance eliminated deriving a rating from a caregiver's narrative and using the judgment of interviewers who were not clinicians. That is, the scores that were obtained reflected the caregiver's own appraisal of the amount of assistance that was being provided. This is in contrast to the FIM approach, in which expert raters determine the level-of-assistance score based on a respondent's description of how the task is performed.

For items with dependence scores of 5 or less (e.g., some form of assistance was provided) and for which caregivers indicated that they were responsible at least in part for providing assistance, caregivers were also asked the following: “How much does helping with this bother or upset you?” Responses were recorded on a 5-point Likert-type scale (0 = *not at all* and 4 = *extremely*).

For analysis purposes, the dependence ratings were reverse coded (7 recoded to 1, 6 to 2, 5 to 3, etc.) such that higher scores represented higher functional dependence. The upset ratings were summarized in two different ways. A total upset score was calculated by first coding as 0 (*no upset*) items for which no assistance was needed or provided and then adding upset ratings across all items in each category. This approach follows previous research that has examined emotional responses to caregiving in specific domains, such as behavioral occurrences (Gitlin, Corcoran, Winter, Boyce, & Hauck, 2001; Haley et al., 1996; Teri et al., 1992). The total upset score can be elevated by either having high-reaction ratings on a few items or by indicating the need for assistance in many IADL or ADL areas. A second score, mean upset, was also calculated to reflect how reactive the caregiver is, on average, to any one IADL or ADL area for which assistance is provided. Caregivers may have high average reaction ratings even though the number of reported areas of dependence may be small.

In summary, the CAFU items were selected from two standardized scales. We modified these scales in three ways. First, we modified the FIM response options. Second, we applied these modified options to both IADL and ADL items. Third, for each IADL and ADL item for which assistance was provided, we asked caregivers to rate their reaction or level of upset. Finally, scores of mean dependence, total upset, and mean upset per dependent item were calculated for analytic purposes.

**Caregiver variables—**In addition to basic demographic variables of caregiver age, race and ethnicity, and gender, we examined depression, caregiver upset associated with memory and behavioral problems, and variables related to vigilance. Caregiver depressive symptomatology was measured by the Center for Epidemiological Studies Depression Scale (CES-D; Radloff, 1977), a 20-item scale with high scores indicating increased depressive symptoms. Scores range from 0 to 60, with a score of 16 or greater indicating risk for clinical depression.

Caregiver reaction to care recipient memory and behavior problems was measured by a modified version of the RMBPC (Teri et al., 1992). The REACH-modified RMBPC reports the presence or absence of 24 disturbing behaviors (7 memory, 8 depressive, and 9 disruptive behaviors) during the past week and yields the number of problem behaviors that occur (range of 0 to 24 behaviors; Roth et al., 2003). This is in contrast to the original scale that reports the actual frequency of occurrence. However, as in the original RMBPC, for each behavior that occurred, in the REACH-modified version, caregivers were also asked their level of upset using a 5-point Likert-type scale (0 = *not at all* and 4 = *extremely*). Thus, scores ranged from 0 to 96, and we refer to this as the RMBPC Reaction scale.

Caregiver vigilance refers to the amount of time caregivers perceive that they provide oversight or surveillance of care recipients in performing everyday tasks. Two vigilance items developed by REACH were used in this study: hours that the caregiver estimated being on duty and hours that the caregiver was involved in doing things for the care recipient (Mahoney et al., 2003). Scores for each item ranged from 0 to 24 hours.

**Care recipient variables—**Care recipient demographic variables included age, gender, and race and ethnicity as identified by the caregiver. Other variables examined were those that might influence the amount of care required, including number of medications, number of behaviors endorsed as occurring using the REACH-modified RMBPC (possible range = 0 to 24), and severity of dementia (MMSE).

## DATA ANALYSIS

Because our goal was to both explore and confirm the factor structure of this new instrument, we first created a random split of the total sample ($N = 640$) into two subsamples of 320 participants using the random sampling procedures of the Statistical Package for the Social Sciences. For the first subsample, we conducted standard exploratory factor analyses using principal components extraction on the dependence ratings. The scree test was used to determine the number of factors, and a varimax rotation was performed.

For the second subsample, we conducted confirmatory factor analyses using the structural equation modeling procedures of LISREL 8 (Jöreskog & Sörbom, 1993a). We sought to evaluate and compare the fit of a simple one-factor model, a two-factor model (ADL vs. IADL), and any other multifactor models that might be suggested from the exploratory factor analysis. Chi-square goodness-of-fit statistics and standard fit indices were used to evaluate model fit. Two different estimation methods were examined. In one set of analyses, maximum likelihood estimation was used to model the standard correlation matrix. In a second set of analyses, weighted least squares estimation was used to model the polychoric correlation matrix. The latter procedure is sometimes considered more appropriate when items with ordinal but not interval properties are the focus of the analysis (Jöreskog & Sörbom, 1993b). Cronbach’s alpha was used to assess internal consistency of the ADL and IADL factors.

We evaluated convergent and discriminant validity by examining the relationships between factor-derived scores and other criteria that were assessed simultaneously using Spearman rank correlation coefficients with the entire sample of 640 participants (Switzer, Wisniewski, Belle, Dew, & Schulz, 1999). Central to the validation effort is the basic idea that scores on a given construct should correlate with conceptually related constructs. To explore the convergence of the measures of dependence and upset, we selected other measures that theoretically should be empirically related to these constructs. We speculated that both IADL and ADL dependence would be empirically related to cognitive status as measured by the MMSE; the greater the dependence, the lower the cognitive status should be. Dependence should also be associated with the amount of time spent providing assistance or caregiving. Therefore, we examined its relationship with the two vigilance items (hours estimated being on duty and hours doing things for care recipient) and speculated that the greater the dependence, the more time will be spent in providing daily oversight. Also, we expected that IADL and ADL dependence would be related to years providing care because more years in caregiving would be an indicator of increasing progressive deterioration. To examine convergent validity for caregiver upset with IADL and ADL dependence, we selected depressive symptoms as assessed by the CES-D score, reaction to problem behaviors as measured by the REACH-modified RMBPC (Roth et al., 2003), and hours involved with care recipient as measured by the two vigilance items. We expected that greater upset with IADL and ADL assistance would be associated with higher depressive scores, upset with problem behaviors, and more time feeling on duty and doing things for care recipient.

To test discriminant validity of the dependence factors, we examined the relationship between IADL and ADL dependence and the occurrence of problem behaviors, number of care recipient medications, and caregiver age. We speculated that these factors would not be significantly associated with level of ADL or IADL dependence. For upset with IADL and ADL dependence, we examined relationships with caregiver age, years of caregiving, and number of care recipient medications.

Finally, we used analyses of variance to determine if functional dependence and caregiver upset were related to grouping variables, such as caregiver and care recipient race and ethnicity and gender or caregiver relationship to the care recipient.

# Results

## SAMPLE CHARACTERISTICS

Table 1 shows the basic caregiver and care recipient characteristics for the total sample and the two random subsamples. The random subsamples were not significantly different on any caregiver or care recipient variable.

There were roughly equivalent numbers of spouses (45%) and children (42%) in the total sample. The mean age of the caregivers was 61.9 (*SD* = 13.5). The majority (77%) of the caregivers were female, and most (80%) reported 12 or more years of education. Most (88%) reported an annual income of $20,000 or more. For care recipients, the mean age was 79.3 (*SD* = 7.9), and 39% were males. The mean MMSE scores for care recipients was 12.0 (*SD* = 7.1) of a possible 30 points. On average, caregivers had spent 4.1 years (*SD* = 3.8) providing care to their care recipients.

## EXPLORATORY FACTOR ANALYSIS

The principal component analysis from the first subsample suggested a two-factor structure, with IADL and ADL items loading on separate factors for both variables. The scree test strongly suggested a two-factor model, with the third largest eigenvalue only slightly greater than 1 (i.e., 1.01). The Varimax factor loadings for the two-and three-factor solutions are displayed in Table 2. All ADL items had their largest factor loadings on Factor 1 in both solutions. In the two-factor models, all IADL items had their largest loading on Factor 2, whereas in the three-factor solution, the IADL items split out into two different factors. Three IADL items (preparing meals, housework, and laundry) now showed their largest loadings on Factor 3. Because of the scree test, the more parsimonious two-factor solutions were endorsed by the results of the principal components analysis, but direct comparisons between the two- and three-factor models were still conducted using confirmatory factor analysis techniques on the independent data from Subsample 2.

## CONFIRMATORY FACTOR ANALYSIS

Confirmatory factor analyses were conducted on the data collected from Subsample 2 using LISREL 8. The fit statistics of one-factor, two-factor (IADL vs. ADL), and three-factor models were examined and compared. In multifactor models, items were constrained to have zero loadings on all factors, except the one factor that showed the largest loading from the principal components analysis in Subsample 1. Correlations between factors were allowed to be freely estimated.

The fit statistics obtained for these models are displayed in Table 3. The chi-square statistic tests the difference between the observed data matrix and the data that can be explained on the basis of the model. All chi-square tests were highly significant, indicating a lack of perfect fit for any model. Additional fit indices were also examined to compare these models. For the Goodness-of-Fit Index, the adjusted Goodness-of-Fit Index, and the Normed Fit Index, possible values ranged from 0 to 1, with higher values indicating better fit. For the root mean square error of approximation (RMSEA), lower values indicate better fit. Models with an RMSEA of less than .05 are generally considered to provide excellent fit to the observed data (Jöreskog & Sörbom, 1993a).

Nested comparisons of the chi-square statistics indicated highly significant improvements in fit for the two-factor model over the one-factor model ($p < .0001$). However, the three-factor model was not found to fit significantly better than the two-factor model. Consequently, the two-factor model with correlated IADL and ADL factors was adopted as the best fitting model for these data. Factor correlations between the IADL and ADL factors were moderate when standard maximum likelihood estimation was used ($r = .55$) and high when weighted least squares estimation was used ($r = .83$; $p < .0001$).

## DESCRIPTIVE STATISTICS AND CONVERGENT AND DISCRIMINANT VALIDITY

Table 4 contains the means, standard deviations, and Cronbach’s alpha measures of internal consistency for the dependence and upset ratings. As shown, caregivers reported that on average, care recipients needed a lot more help to complete IADLs than ADLs. Also, caregivers reported low average levels of upset for both IADL and ADL dependencies. That is, most caregivers reported little to no upset on these items. However, mean upset for each ADL item was slightly higher than the mean upset per IADL item. As shown in the appendix, the mean upset scores for each dependent item reveals that as expected, caregivers report the highest upset rating for toileting assistance, followed by bathing assistance. High internal consistency was observed for both the IADL and ADL dependence and upset scales. (Table 4)

Table 5 presents the Spearman rank correlations between dependence and upset ratings and other measures of caregiver and care recipient functioning. Correlations between total upset and mean upset scores were very high ($r$s > .96). Because both indices showed almost identical correlations with other measures, only the correlations with total upset are reported here. With regard to convergent validity for dependence factors, IADL and ADL dependence were both significantly associated with the number of hours the caregiver estimated being on duty and the number of hours estimated doing things for the care recipient. Similarly, IADL and ADL dependence were associated with the number of years of caregiving and with the care recipient's MMSE scores. As higher IADL and ADL scores indicate more dependence (or more assistance required), these correlations are consistent with theoretical predictions.

As with caregiver reactions, upset ratings for both IADLs and ADLs were associated with caregiver depressive symptoms and caregiver reaction to behaviors. These correlations are also consistent with theoretical predictions. Total upset with ADLs and not IADLs was associated with the number of hours doing things for the care recipient and number of hours estimated being on duty. The associations, although significant, are not as strong as the relationship between ADL upset and affective measures, suggesting that appraisal of upset is more closely linked with affective well-being rather than appraisals of perceived effort. We also found that IADL and ADL upset scores were both significantly associated with the number of behavioral occurrences; although, we did not initially consider the relationship for validity testing.

It should be noted that we did not consider the relationship between IADL and ADL upset scores and MMSE scores for validity testing. However, we found that lower MMSE scores were associated with greater caregiver reaction to ADLs, whereas the opposite was the case with IADL upset. We found a low but significant positive correlation between MMSE scores and IADL upset such that higher cognitive status was associated with greater upset in this area. It is unclear as to how to interpret this small but statistically significant relationship. It may be that persons with higher MMSE show initial signs of IADL performance difficulties and that this becomes particularly upsetting to caregivers as it represents concrete evidence of the beginning of physical and cognitive decline.

With regard to discriminant validity, as anticipated, dependence was not significantly associated with caregiver age, number of care recipient medications, or number of behavior problems (RMBPC). Also, as expected, upset ratings were not significantly related to years of caregiving or number of medications. There was a small but significant association between caregiver age and IADL upset such that older caregivers reported more upset with IADL dependencies. This was not the case, however, for the relationship with ADL upset. (Table 5)

## EFFECT OF CAREGIVER AND CARE RECIPIENT CHARACTERISTICS ON SUBSCALE SCORES

Finally, we examined whether subscale scores for dependence and upset differed as a function of gender, race and ethnicity, or the relationship between the caregiver and care recipient. Caregiver gender was not significantly related to either ADL or IADL dependence or upset scores. However, care recipient gender was significantly related to IADL upset and ADL dependence scores. Female care recipients were rated by caregivers as causing less total IADL upset than male care recipients, $F(1, 638) = 11.50$, $p = .0007$, *M*s = 3.41 and 4.72, respectively. Also, female care recipients were rated as more dependent on assistance for ADLs than male care recipients, $F(1, 638) = 4.45$, $p = .035$, *M*s = 3.82 and 3.50, respectively. With regard to race and ethnicity, of the 640 caregiver–care recipient dyads, 617 were classified as African American ($n = 277$) or White ($n = 340$). The remaining 23 dyads consisted of other ethnic classifications or interracial dyads, and these cases were excluded when examining differences between African American and White dyads. Using this simple dichotomy, no statistically significant differences were observed on IADL and ADL dependence scores. However, African American caregivers reported less total upset than White caregivers on both IADLs, $F(1, 615) = 8.80$, $p = .003$, *M*s = 3.23 and 4.38, respectively, and ADLs, $F(1,615) = 8.10$, $p = .005$, *M*s = 2.49 and 3.50, respectively.

To examine differences resulting from the relationship between the caregiver and the care recipient, five different groups were defined: wives ($n = 197$), husbands ($n = 91$), daughters ($n = 223$), sons ($n = 45$), and other ($n = 84$). One-way ANOVAs indicated that significant differences were evident among these five groups on IADL upset scores only, $F(4, 635) = 7.20$, $p < .0001$. Wives reported the highest level of upset ($M = 5.28$), and daughters ($M = 2.99$) and sons ($M = 3.04$) reported the lowest levels of upset. The level of upset for husbands ($M = 4.30$) was between that of wives and children.

## Discussion

This study reports the psychometric analyses of a new instrument designed to measure both caregiver appraisal of the amount of IADL and ADL assistance required by individuals with ADRD and also the caregiver's reaction to providing this assistance. Also, differential response patterns for both dependence and upset levels were examined. Although there are numerous existing instruments designed to measure functional dependence in persons with ADRD, this is the first measure involving a caregiver report that has been systematically evaluated with a geographically and racially diverse sample of caregivers. Also, although traditional measures of caregiver burden may include items related to providing hands-on care (Vitaliano, Russo, Young, Becker, & Maiuro, 1991; Zarit, Reever, & Bach-Peterson, 1980), they do not systematically assess caregiver reaction to providing assistance with IADL and ADL areas. Thus, the CAFU provides a reliable and valid approach to evaluating 15 traditional areas of physical function as well as which areas of functional dependence are upsetting to caregivers. The advantage of this approach is twofold. First, it provides the researcher and clinician a quick way of evaluating specific areas of daily functioning that require caregiver assistance and also present a source of distress. In this respect, it is more comprehensive than other traditional functional status measures. Also, given that we have standardized the response set and probes, it is shorter than the FIM and does not require clinical training for its use. Thus, it can easily be used by research interviewers who may not have clinical training. Second, the CAFU facilitates developing targeted interventions to enhance education and skills training for caregivers in specified areas in which assistance is provided and ways of managing the associated upset. As such, the instrument goes beyond other existing functional status measures that rely on self-reports or direct observation but that do not account for the caregiver's role in providing assistance and the associated emotional consequences.

To our knowledge, this study, conducted with 640 regionally and racially diverse caregiver–care recipient dyads, is one of the largest validation studies for a measure of functional dependence ever conducted with ADRD patients and their caregivers. The obtained findings demonstrated good internal consistency for both measures of functional capacity and caregiver upset. Excellent factorial validity was obtained as well as convergent and discriminant validity for both the functional capacity and upset measure. More specifically, factor analyses confirmed that ADLs loaded on a different factor than IADLs for dependence scores. ADLs are based on rudimentary over-learned abilities, whereas IADLs are more influenced by a range of factors, such as motivation, social expectations, the environment and prospective memory (e.g., remembering an intended action), which is significantly affected in Alzheimer's Disease (Carswell & Spiegal, 1999). The dissociation between ADLs and IADLs is also consistent with data obtained by Loewenstein et al. (in press) on a more mildly impaired ADRD sample than assessed in the current study.

The dependence measures also demonstrated excellent convergent validity with results showing higher IADL and ADL dependence associated with higher MMSE scores and caregiver reports of more time devoted to surveillance or oversight of daily performance. Also, upset scores were associated with higher depression and behavioral reaction ratings as predicted. Discriminant validity was established by the finding that dependence and upset were unrelated to other care recipient characteristics, such as number of behavior problems and number of medications.

Of interest is the relationship between level of assistance caregivers provide and their level of upset. We found that caregiver upset with IADL dependence appeared to be relatively independent of perceived level of IADL and ADL dependence of the care recipient. However, upset with ADL capacity was significantly associated with both IADL and ADL dependence level of the care recipient. One possible explanation for this finding is that the current sample consisted of moderately to severely impaired ADRD patients (Mean MMSE score = 12.0). In this population, difficulties in ADL performance might have been more emergent and caused more acute distress. Another explanation may be that caregiver distress is independent of IADL impairment level when ADL upset is present. The data, however, do not support a difference in the mean degree of upset associated with IADL and ADL assistance.

We also found differential response patterns for dependence and upset by caregiver–care recipient characteristics. There did not appear to be significant differences in reported functional capacities of African American and White care recipients. That is, African American and White caregivers report similar levels of assistance provided in IADLs and ADLs. However, African American caregivers report less upset with deficits in each of these domains than White caregivers do. This is consistent with previous literature showing that African American caregivers evidence less subjective burden with disruptive behaviors of care recipients and less depressive symptoms than White caregivers (Haley et al., in press; Haley et al., 1996). A potential limitation of this study is that we were unable to examine the factor structure using a confirmatory factor analysis approach for each racial group separately because of an insufficient number of participants for these multivariate comparisons in Subsample 2. Although we do not anticipate differences in the covariance structure for African American and White caregivers, nor do we have any hypotheses to support such differences, future research with a larger normative sample may want to examine this further.

With regard to gender, we found that care recipient gender was related to caregiver ratings, whereas caregiver gender did not appear to affect ratings of either assistance or burden. Deficits in IADLs of female care recipients were rated less burdensome than the same deficits in males. This is probably because of premorbid gender differences in the performance of IADLs, such as check writing and driving, found for this age cohort. Nevertheless, female care recipients were reported by caregivers to be more dependent on assistance with ADLs than male care recipients. However, when relationship of caregiver and care recipient was examined, differences were found for IADL upset ratings. Specifically, wife caregivers reported the greatest amount of burden associated with deficits in IADL, and sons and daughters reported the least amount of burden. Husband caregiver ratings of IADL upset ranked between the wives and children’s ratings.

The reliance of self-report on functional dependence may represent a potential limitation in the use of the CAFU. Previous research has identified potential biases when caregivers report the physical functioning of their family members and a tendency to overestimate functional ability (Loewenstein et al., in press; Magaziner, Simonsick, Kashner, & Hebel, 1988; Ostbye, Tyas, McDowell, & Koval, 1997; Weinberger et al., 1992; Zimmerman & Magaziner, 1994). More recently, Argüelles, Loewenstein, Eisdorfer, and Argüelles (in press) show that ADRD caregivers’ overestimation of functional capacities is associated with very low depressive scores on the CES-D scale. This suggests that nondepressed caregivers may minimize difficulties, whereas depressed caregivers may underestimate functional capacities. Another potential limitation of this study is that we did not evaluate test-retest reliability. Future research is in order to further evaluate the psychometric properties of the CAFU.

Nevertheless, we contend that the CAFU is a measure of an important aspect of home life, the functional context of caregiving, and caregiver emotional reaction. As such, we maintain that obtaining a caregiver's appraisal of the care recipient's level of dependence and the emotional reaction to providing assistance provides knowledge of real-world daily care provided at home (i.e., as it is perceived and acted on by family caregivers). This stance is supported by recent evidence that caregiver self-reported FIM ratings correspond to FIM ratings that are derived by direct observation of self-care performance by objective raters (Cotter, Burgio, Stevens, Roth, & Gitlin, 2002). That research shows that caregivers were able to realistically report the level of assistance they provided. Perhaps both the caregiver report as well as the objective measures of functional status would be important to obtain to derive the most accurate representation of the caregivers participation in IADL and ADL care and reactions to providing assistance.

Given the factor structure and validity testing, we recommend that users of the CAFU derive four ratings, IADL and ADL dependence and IADL and ADL reaction scores, to physical dependence. Also, within clinical situations, it may be helpful to examine item-by-item reaction scores to evaluate the caregiver's emotional capacity to provide assistance and develop targeted interventions to address the particular areas of physical function that provide the most difficulty to the family.

## Acknowledgements

AUTHORS'NOTE: This research was supported through the Resources for Enhancing Alzheimer's Caregiver Health project, which is supported by the National Institute on Aging and the National Institute of Nursing Research (Grants Burgio U01-NR13269, Burns U01-AG13313, Eisdorfer U01-AG13297, Gallagher-Thompson U01-AG13289, Gitlin U01-AG13265, Mahoney U01-AG13255, and Schulz U01-AG13305).

## References

American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 4. Washington, DC: Author; 1994.

Argüelles S, Loewenstein DA, Eisdorfer C, Argüelles T. Caregivers' judgements of the functional abilities of the Alzheimer's disease patient. The impact of caregivers' depression and perceived burden. Journal of Geriatric Psychiatry and Neurology 14(2)in press

Carswell, A.; Spiegal, R. Functional assessment. In: Wilcock, G.; Bucks, R.; Rockwood, K., editors. Diagnosis and management of dementia: A manual for memory disorders teams. Oxford, UK: Oxford University Press; 1999.

Cotter EM, Burgio LD, Stevens AB, Roth DL, Gitlin LN. Correspondence of the Functional Independence Measure (FIM) self-care subscale with real-tine observations of dementia patients' ADL performance in the home. Clinical Rehabilitation 2002;16:36–45. [PubMed: 11837524]

Folstein M, Folstein S, McHugh P. Mini-mental state: A practical method for grading cognitive state of care recipients for the clinician. Journal of Psychiatry Research 1975;12:189–198.

Gitlin, LN. Physical function in the elderly: A comprehensive guide to its measurement. Austin, TX: ProPress; in press

Gitlin LN, Corcoran M, Winter L, Boyce A, Hauck W. A randomized, controlled trial of a home environmental intervention: Effect on efficacy and upset in caregivers and on daily function of persons with dementia. The Gerontologist 2001;41:4–14. [PubMed: 11220813]

Haley WE, Gitlin LN, Wisniewski SR, Mahoney DF, Coon DW, Winter L, et al. Well-being, appraisal, and coping in African American and Caucasian dementia caregivers: Findings from the REACH study. Aging & Mental Health. in press

Haley WE, Roth DL, Coleton MI, Ford GR, West CAC, Collins RP, et al. Appraisal, coping, and social support as mediators of well-being in black and white family caregivers of patients with Alzheimer's Disease. Journal of Consulting and Clinical Psychology 1996;64:121–129. [PubMed: 8907091]

Hamilton, BB.; Fuhrer, MJ., editors. Rehabilitation outcomes: Analysis and measurement. Baltimore: Brooks; 1987.

Jöreskog, KG.; Sörbom, DG. LISREL 8 user's reference guide. Chicago: Scientific Software International; 1993a.

Jöreskog, KG.; Sörbom, DG. PRELIS 2 user's reference guide. Chicago: Scientific Software International; 1993b.

Katz S, Ford AB, Moskowitz RW, Jackson BA, 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]

Kiyak HA, Teri L, Borson S. Physical and functional health assessment in normal aging and in Alzheimer's disease: Self reports versus family reports. The Gerontologist 1994;34:324–330. [PubMed: 8076873]

Lawton MP, Brody E. Assessment of older people: Self-maintaining and instrumental activities of daily living. The Gerontologist 1969;9:179–186. [PubMed: 5349366]

Loewenstein, DA.; Mogosky, B. Functional assessment of the older adult patient. In: Lichtenberg, P., editor. Handbook of assessment in clinical gerontology. New York: John Wiley; 1999. p. 529-554.

Loewenstein DA, Ownby R, Schram L, Acevedo A, Rubert M, Argüelles T. An evaluation of the NINCDS-ADRDA neuropsychological criteria for the assessment of Alzheimer's disease: A confirmatory factor analysis of single versus multi-factor models. Journal of Clinical & Experimental Neuropsychology. in press

Magaziner J, Simonsick EM, Kashner TM, Hebel JR. Patient-proxy comparability on measures of patient health and functional status. Journal of Clinical Epidemiology 1988;41:1065–1074. [PubMed: 3204417]

Mahoney DF, Jones RN, Coon D, Mendelsohn AB, Gitlin LN, Ory M. The Caregiver Vigilance Scale: Application and validation in the Resources for Enhancing Alzheimer's Caregiver (REACH) project. American Journal of Alzheimer's Disease and Other Dementias 2003;18:39–48. [PubMed: 12613133]

McClendon MJ, Smyth KA, Neundorfer MM. Survival of persons with Alzheimer's disease: Caregiver coping matters. The Gerontologist 2004;44:508–519. [PubMed: 15331808]

Ostbye T, Tyas S, McDowell I, Koval J. Reported activities of daily living: Agreement between elderly subjects with and without dementia and their caregivers. Age and Aging 1997;26:99–106.

Radloff LS. The CES-D Scale: A self-report depression scale for research in the general population. Applied Psychological Measures 1977;1:385–401.

Roth DL, Burgio LD, Gitlin LN, Gallagher-Thompson D, Coon DW, Belle SH, et al. Psychometric analysis of the Revised Memory and Behavior Problems Checklist: Factor structure of occurrence and reaction ratings. Psychology and Aging 2003;18(4):906–915. [PubMed: 14692875]

Rubenstein LZ, Schairer C, Wieland GD, Kane R. Systematic biases in the functional status of elderly adults: Effects of different data sources. Journal of Gerontology 1984;39:686–691. [PubMed: 6436360]

Switzer GE, Wisniewski SR, Belle SH, Dew MA, Schulz R. Selecting, developing, and evaluating research instruments. Social Psychiatry Epidemiology 1999;34:399–409.

Teri L, Truax P, Logsdon R, Uomoto J, Zarit S, Vitaliano PP. Assessment of behavioral problems in dementia: The Revised Memory and Behavior Problems Checklist. Psychology and Aging 1992;7:622–631. [PubMed: 1466831]

Vitaliano PP, Russo J, Young HM, Becker J, Maiuro RD. The screen for caregiver burden. The Gerontologist 1991;31:76–83. [PubMed: 2007478]

Weinberger M, Samsa G, Schmader K, Greenberg S, Carr D, Wildman DS. Comparing proxy and patients'perceptions of patients' functional status: Results from a geriatric outpatient clinic. Journal of the American Geriatrics Society 1992;40:585–588. [PubMed: 1587975]

Zarit SH, Reever KE, Bach-Peterson J. Relatives of the impaired elderly: Correlates of feelings of burden. The Gerontologist 1980;20:649–655. [PubMed: 7203086]

Zimmerman SI, Magaziner J. Methodological issues in measuring the functional status of cognitively impaired nursing home residents: The use of proxies and performance-based measures. Alzheimer Disease and Associated Disorders 1994;8(1):S281–S290. [PubMed: 8068270]

## Table 1

Descriptive Data for Total Sample and Random Subsamples

| | Total (N = 640) | | Subsample 1 (n = 320) | | Subsample 2 (n = 320) | | |
|---|---|---|---|---|---|---|---|
| ***Variable*** | **M** | **SD** | **M** | **SD** | **M** | **SD** | **p**$^{a}$ |
| CG age | 61.9 | 13.5 | 62.0 | 13.8 | 61.9 | 13.2 | .92 |
| CR age | 79.3 | 7.9 | 78.9 | 7.9 | 79.8 | 7.9 | .15 |
| CR MMSE | 12.0 | 7.1 | 11.7 | 7.3 | 12.3 | 6.9 | .27 |
| | n | % | n | % | n | % | p$^{b}$ |
| CG-CR dyad race | | | | | | | .74 |
| White | 340 | 53 | 172 | 54 | 168 | 53 | |
| African American | 277 | 43 | 135 | 42 | 142 | 44 | |
| Interracial and other | 23 | 4 | 13 | 4 | 10 | 3 | |
| CG sex | | | | | | | .26 |
| Male | 148 | 23 | 80 | 25 | 68 | 21 | |
| Female | 492 | 77 | 240 | 75 | 252 | 79 | |
| CR sex | | | | | | | .87 |
| Male | 250 | 39 | 126 | 39 | 124 | 39 | |
| Female | 390 | 61 | 194 | 61 | 196 | 61 | |
| CG relationship | | | | | | | .14 |
| Wives | 197 | 31 | 96 | 30 | 101 | 32 | |
| Husbands | 91 | 14 | 55 | 17 | 36 | 11 | |
| Daughters | 223 | 35 | 104 | 33 | 119 | 37 | |
| Sons | 45 | 7 | 19 | 6 | 26 | 8 | |
| Others | 84 | 13 | 46 | 14 | 38 | 12 | |
| Site | | | | | | | .92 |
| Birmingham, AL | 140 | 22 | 69 | 22 | 71 | 22 | |
| Memphis, TN | 245 | 38 | 125 | 39 | 120 | 38 | |
| Philadelphia | 255 | 40 | 126 | 39 | 129 | 40 | |

[a] Independent sample *t* test.

[b] Pearson chi-square test

*Note.* CG = caregiver; CR = care recipient; MMSE = Mini-Mental State Examination.

## Table 2

Principal Components Analysis of ADL and IADL Dependence Scores With Varimax Factor Loadings Greater Than or equal to .30 (Subsample 1, *n* = 320)

| | Two-Factor Model | | Three-Factor Model | | |
|---|---|---|---|---|---|
| **Variable** | **Factor 1** | **Factor 2** | **Factor 1** | **Factor 2** | **Factor 3** |
| Phone | .376 | .540 | .380 | .507 | — |
| Shopping | — | .774 | — | .734 | .307 |
| Food | — | .664 | — | .410 | .579 |
| Housework | .358 | .513 | — | — | .760 |
| Laundry | — | .622 | — | — | .829 |
| Travel | — | .607 | — | .675 | — |
| Medicine | — | .576 | — | .484 | — |
| Finances | — | .587 | — | .730 | — |
| Bed | .783 | — | .784 | — | — |
| Eating | .695 | — | .699 | — | — |
| Bathing | .723 | — | .718 | — | — |
| Dress, above the waist | .847 | — | .839 | — | — |
| Dress, below the waist | .843 | — | .830 | — | — |
| Toilet | .843 | — | .842 | — | — |
| Grooming | .747 | — | .740 | — | — |
| Percentage of variance | 31.95 | 22.76 | 31.14 | 16.30 | 14.00 |

## Table 3

LISREL Fit Statistics (Subsample 2, n = 320)

| Model | Method | $\chi^2$ | df | p | GFI | AGFI | NFI | RMSEA |
|---|---|---|---|---|---|---|---|---|
| One factor | Corr and ML | 607.7 | 89 | < .001 | .70 | .60 | .75 | .14 |
| Two factor | Corr and ML | 195.3 | 88 | < .001 | .92 | .90 | .92 | .06 |
| Three factor | Corr and ML | 193.3 | 86 | < .001 | .92 | .89 | .92 | .06 |
| One factor | Poly and WLS | 188.5 | 89 | < .001 | .98 | .97 | .97 | .06 |
| Two factor | Poly and WLS | 141.6 | 88 | < .001 | .98 | .98 | .98 | .04 |
| Three factor | Poly and WLS | 137.7 | 86 | < .001 | .98 | .98 | .98 | .04 |

*Note*. Corr and ML = Pearson’s correlation matrix and maximum likelihood estimation; poly and WLS = polychoric correlation matrix and weighted least squares estimation; GFI = Goodness-of-Fit Index; AGFI = Adjusted Goodness-of-Fit Index; NFI = Normed Fit Index; RMSEA = root mean square error of approximation.

## Table 4

Subscale Descriptive Data for IADL and ADL Dependence and Upset With Total Sample ($N = 640$)

| ***Measure*** | **M** | **SD** | **Mdn** | ***Minimum*** | ***Maximum*** | ***Cronbach's α*** |
|---|---|---|---|---|---|---|
| IADL | | | | | | |
| Dependence | 6.14 | 1.05 | 6.50 | 7.00 | 1.37 | .81 |
| Total upset | 3.92 | 4.83 | 2.00 | 0.00 | 32.00 | .80 |
| Mean upset per dependent item | 0.58 | 0.69 | 0.33 | 0.00 | 4.00 | .84 |
| ADL | | | | | | |
| Dependence | 3.70 | 1.85 | 3.29 | 7.00 | 1.00 | .91 |
| Total upset | 3.09 | 4.41 | 1.00 | 0.00 | 28.00 | .83 |
| Mean upset per dependent item | 0.70 | 0.89 | 0.40 | 0.00 | 4.00 | .90 |

*Note*. Higher scores indicate greater dependence and greater upset.

## Table 5

### Spearman Correlations Between CAFU Subscales and Other Measures

| | IADL | | ADL | |
|---|---|---|---|---|
| **Variable** | **Dependence** | **Total Upset** | **Dependence** | **Total Upset** |
| CAFU | | | | |
| IADL dependence | 1.00 | | | |
| IADL upset | −0.04 | 1.00 | | |
| ADL dependence | 0.65** | −0.04 | 1.00 | |
| ADL upset | 0.31** | 0.43** | 0.46** | 1.00 |
| Caregiver | | | | |
| CG age | 0.01 | 0.14** | −0.08 | 0.00 |
| Years caregiving | 0.16* | 0.04 | 0.18** | 0.05 |
| CES-D | 0.07 | 0.32** | 0.17** | 0.29** |
| Hours feel on duty | 0.24** | 0.01 | 0.21** | 0.09* |
| Hours doing things for CR | 0.24** | 0.06 | 0.30** | 0.21** |
| RMBPC reaction | −0.05 | 0.47** | −0.06 | 0.34** |
| CR | | | | |
| CR age | 0.10* | −0.09* | 0.05 | 0.03 |
| Modified RMBPC occurrence | −0.00 | 0.26** | −0.01 | 0.26** |
| MMSE | −0.45** | 0.13* | −0.48** | −0.19** |
| Number of medications | 0.02 | −0.01 | 0.04 | 0.00 |

*Note.* CAFU = Caregiver Assessment of Function and Upset; IADL = instrumental activity of daily living; ADL = activity of daily living; CG = caregiver; CR = care recipient; CES-D = Center for Epidemiological Studies Depression; RMBPC = Revised Memory and Behavior Problem Checklist; MMSE = Mini-Mental State Examination.

* $p < .05$.

** $p < .001$.

## Appendix

Caregiver Assessment of Functional Dependence and Upset: items, Means, and Standard Deviations

| | **Mean Dependence** | **SD** | **n With Dependence Greater Than or Equal to 3** | **%** | **Mean Upset per Dependent Item Greater Than or Equal to 3** | **SD** |
|---|---|---|---|---|---|---|
| IADL | | | | | | |
| Phone | 5.23 | 2.17 | 417 | 65 | 0.58 | 0.93 |
| Shopping | 6.37 | 1.39 | 564 | 88 | 0.62 | 0.98 |
| Food | 6.31 | 1.50 | 582 | 91 | 0.57 | 0.97 |
| Housework | 6.03 | 1.78 | 551 | 86 | 0.75 | 1.08 |
| Laundry | 6.45 | 1.40 | 558 | 87 | 0.44 | 0.89 |
| Travel | 6.19 | 1.46 | 528 | 83 | 0.64 | 1.06 |
| Medicine | 5.83 | 1.82 | 603 | 94 | 0.50 | 0.94 |
| Finances | 6.73 | 0.94 | 572 | 89 | 0.51 | 0.98 |
| ADL | | | | | | |
| Bed transfer | 3.05 | 2.28 | 252 | 39 | 0.89 | 1.13 |
| Eating | 2.32 | 1.81 | 198 | 31 | 0.87 | 1.10 |
| Bathing | 4.79 | 2.28 | 391 | 61 | 0.95 | 1.18 |
| Dress, above the waist | 4.03 | 2.33 | 395 | 62 | 0.69 | 1.03 |
| Dress, below the waist | 4.00 | 2.36 | 382 | 60 | 0.88 | 1.11 |
| Toilet | 3.51 | 2.52 | 311 | 49 | 1.19 | 1.18 |
| Grooming | 4.16 | 2.32 | 394 | 62 | 0.58 | 0.98 |

*Note*. For each item, caregivers were asked the following: “During the past week, has (care recipient name) needed any kind of help with (name of activity)?” Dependence was rated on a scale of 1 to 7, with high scores indicating greater dependence. For items with dependence ratings of 3 or greater, caregivers were asked the following: “How upset are you with providing help on a 0 to 4 scale?” High scores indicated greater upset.