The Assessment of ADL Among Frail Elderly in an Interview Survey: Self-report versus Performance-Based Tests and Determinants of Discrepancies

Gertrudis I.J.M. Kempen,1 Nardi Steverink,1 Johan Ormel,2 and Dorly J.H. Deeg3

1Northern Centre for Health Care Research and 2Department of Psychiatry and Department of Health Sciences, University of Groningen, The Netherlands.
3Department of Psychiatry and Department of Social Gerontology, Free University of Amsterdam, The Netherlands.

Although the interest in performance-based measurement of activities of daily living (ADLs) has increased during the last decade, relatively little is known about the relationship between self-report and performance-based measures of ADL. Even less is known about the factors responsible for the discrepancies between them. The purpose of this study was to identify determinants of discrepancies between self-report and performance-based measures of ADL.

Recently, several studies showed only moderate to weak associations between self-report and performance-based measures of functioning (Cress et al., 1995; Deeg, 1994; Jette & Branch, 1985; Kelly-Hayes, Jette, Wolf, D'Agostino, & Odell, 1992; Myers, Holliday, Harvey, & Hutchinson, 1993; Reuben, Valle, Hays, & Siu, 1995; Sager et al., 1992). From a conceptual or theoretical point of view, the discrepancies between both types of measurement can be explained by at least three factors. Firstly, compared to self-report measures, performance-based measures seem to be more sensitive to (a) fluctuations in time and (b) differences in equipment or aids used in the performance of daily activities. Subjective evaluation of functioning reflects findings over a period of at least a couple of days, and the judgments of functioning are generally based on the experiences with different aids or equipment. As was suggested by Myers et al. (1993) and Guralnik, Branch, Cummings, and Curb (1989), performance-based measures generally do not reflect adaptations made in a person's everyday living situation (e.g., proper footwear, residential features such as types of stairs and surfaces, compensatory mechanisms, e.g., handrails, and supportive relationships) that may affect the lack of concordance between the types of measurement.

Secondly, discrepancies can be due to the differences in concepts measured during the development from pathology to its final consequences (e.g., impairments vs functional limitations vs disability in Nagi’s scheme [Nagi, 1991]). Stronger associations can be expected among concepts which are situated closer together in this development. Most performance-based measures assess basic functional limitations and are not expected to directly assess ADLs. In contrast to most self-report measures, performance-based measures are, typically, much more tightly focused on a specific aspect of functioning and provide more precise information regarding that aspect of functioning (Evans, 1993). Finally, previous research suggests that cognitive functioning, affective functioning, personality characteristics, and sociodemographic variables may reduce the concordance between self-report and performance-based measures of ADL.

With respect to cognitive functioning, the lack of concordance may reflect a problem of memory or of judgment (Branch & Myers, 1987). Guralnik et al. (1989) stated that performance-based measures are less sensitive to poor cognitive functioning, educational level, and culture. Kelly-Hayes et al. (1992) and Sager et al. (1992) found significantly lower rates of agreement between performance-based and self-report measures for subjects with cognitive impairments. However, several studies showed positive associations of cognitive functioning with both self-reported functioning (Barberger-Gateau et al., 1992; Cress et al., 1995; Griffiths et al., 1987; Laukkanen, Kappinen, Era, & Heikkinen, 1993; Rozzini, Frisoni, Bianchetti, Zanetti, & Trabucchi, 1993) and performance-based functioning (Cress...
et al., 1995; Rozzini et al., 1993). The role of cognitive functioning in the discrepancies between the types of measurement is still unclear.

With respect to affective functioning, it was found that depressive symptoms independently predict lower levels of self-reported functioning among community-dwelling older subjects, although depressive symptoms were also negatively associated with performance-based functioning (Cress et al., 1995). In contrast, Rozzini et al. (1993) found no significant relationship between performance-based functioning and depressive symptomatology. Several studies showed negative associations between depressive symptomatology and self-reported functioning (Griffiths et al., 1987; Kempen & Suurmeijer, 1991; Laukkanen et al., 1993; Rozzini et al., 1993; Turner & Noh, 1988; VonKorff, Ormel, Katon, & Lin, 1992; Wells et al., 1989). Sager et al. (1992) reported that subjects with depressive symptomatology tended to have lower rates of agreement between self-report and performance-based measures, although these differences were not significant (Sager et al., 1992). These results suggest that depressive symptomatology is more strongly associated with self-reported functioning than with performance-based functioning. Subjects with depressive symptomatology tend to underestimate their level of functioning.

In the area of personality characteristics, research outcomes supported the hypothesis that personality traits such as neuroticism are related to the perception of health rather than to objective health status (Smith & Williams, 1992), although the positive correlation between neuroticism and self-reported ADL dysfunctioning was not strong (Jorm et al., 1993). Tinetti, Mendes de Leon, Doucette, and Baker (1994) reported significant positive associations between efficacy (as assessed with the Falls Efficacy Scale) and ADL–IADL functioning while controlling for performance-based levels of functioning. The latter result suggests that subjects with lower levels of efficacy tend to underestimate their level of functioning.

With respect to sociodemographic variables, previous research showed that age was related to both self-reported and performance-based functioning: older subjects reported lower levels of functioning (Jette & Branch, 1985). The same researchers showed that age and gender were still significantly associated with self-reported functioning, while performance-based levels of functioning are controlled. Whether or not gender is related to self-reported and/or performance-based ADL functioning seems to be partly dependent on the items selected (Deeg, 1993). Parker, Thorslund, and Lundberg (1994) found significant associations between social class (as an indicator of socioeconomic status) and both self-reported and performance-based functioning; in this case former white-collar workers had higher levels of functioning than blue-collar workers.

The above findings suggest that levels of cognitive and affective functioning, personality characteristics, and sociodemographic variables may play an important role in the lack of concordance between performance-based and self-report measures of ADL.

The objective of this study was to investigate the relationship between performance-based and self-report measures of ADL among frail elderly persons and to identify factors responsible for discrepancies between them. The starting point is a set of performance-based ADL measures that can be administered during an in-home interview session. We expect (1) only moderate associations between self-reported and performance-based ADLs, and (2) larger discrepancies between self-reported and performance-based ADLs for subjects who show lower levels of cognitive and/or affective functioning (e.g., depressive symptomatology), lower levels of self-efficacy (e.g., physical competence, mastery), higher levels of neuroticism, and lower levels of education.

METHODS

Sample

The data are obtained from a subsample (n = 753) of the baseline participants (N = 5,279) in the Groningen Longitudinal Aging Study (GLAS). This is a population-based prospective follow-up study of the determinants of health-related quality of life of older people, with special emphasis on physical and social disability and well-being (Ormel et al., 1992). The primary objective of GLAS is to identify the psychosocial factors that influence the trajectory of the quality of life, either independently or in interaction with disease-related factors. The source population of GLAS is defined as being people aged ≥ 57 years who lived in the north of the Netherlands either independently or in adapted housing for elderly people in 1993. Subjects with severe cognitive impairments at baseline (MMSE < 17; Folstein, Folstein, & McHugh, 1975) were excluded (n = 78). The subsample in this paper comprises 753 frail elderly (14.3%) with the lowest scores on the 6-item SF-20 physical functioning scale (Stewart, Hays, & Ware, 1988). The selected subjects reported four (35.9%), five (45.7%), or six (18.5%) physical limitations on this scale. Additional data were collected in a face-to-face follow-up interview, approximately three weeks after the baseline. The subsample consisted of 544 females (mean age: 73.6; range: 57–91; SD: 7.6) and 209 males (mean age: 71.9; range: 57–93; SD: 8.8). Of all the subjects, 133 (17.7%) were younger than 65 years of age, and 311 subjects (41.3%) were 75 years of age or older.

Measures

As suggested by Guralnik et al. (1989), the application of performance-based tests has, compared to self-report measures, both advantages and disadvantages. Performance-based tests are time-consuming; they require the special training of examiners or interviewers, adequate space, and special equipment (e.g., stopwatches); also, there is a risk of injury. As a consequence, not all tests can easily be implemented during in-home interview sessions. On the other hand, it was suggested that performance-based tests have better reproducibility, have more face validity for the task being performed, have greater sensitivity to change, and are less influenced by poor cognitive functioning, culture, language, and education. However, empirical results do not lend support to the notions that performance-based measures are more acceptable to patients, clinically feasible, reproducible, or sensitive to change (Myers et al., 1993).

Several types of performance-based tests can be distinguished (for reviews see Guralnik et al., 1989; Myers et al.,
1993). They may differ considerably in the degree to which they can be used during in-home interview sessions. In our study we used three simple, performance-based ADL tests in the follow-up. The tests were obtained from the Longitudinal Aging Study Amsterdam (Deeg, 1994; Deeg, Knipscheer, & Van Tilburg, 1993). These consist of (a) putting on and taking off a jacket, (b) walking 6 meters (including a 180° turn after 3 meters), and (c) standing up from a kitchen chair and sitting down 5 times, without using the arms. The number of seconds required for the performance of each test is recorded as the score. Although putting on a jacket and taking it off again were performed separately, the scores of both tests were added up. Higher scores reflect poorer performance. Subjects for whom no timed score was available (i.e., they tried but were unable to do the test, were told to stop for safety reasons, or they refused) were assigned a score equal to the worst actual score (putting on/taking off a jacket: n = 22 of whom n = 16 refused; walking 6 meters: n = 48 of whom 8 refused; standing up/sitting down from a kitchen chair: n = 180 of whom n = 99 refused). (The procedure conformed to Seeman et al., 1994.) Nonresponse analyses showed significantly poorer self-reported ADL functioning for nonresponders compared to responders for each of the three performance-based tests (Student's t-tests, p < .001).

Self-reported ADL has been assessed at the baseline with the 11-item ADL subscale of the Groningen Activity Restriction Scale (GARS; Kempen, Miedema, Ornelt, & Molenaar, in press; Kempen & Suurmeijer, 1990; Suurmeijer et al., 1994). The items and response categories are presented in Table 1. The theoretical range varies from 11 (no ADL limitations) to 44 (maximum of ADL limitations). The internal reliability estimate (Cronbach's alpha) was .84. The results of previous studies showed that the GARS meets the stochastic cumulative scalability criteria of the Mokken model (Kempen et al., in press; Kempen & Suurmeijer, 1990; Suurmeijer et al., 1994). Three GARS items, which will also be analyzed separately, correspond with the performance-based ADL items: “Can you, fully independently, dress yourself?” (item 1), “Can you, fully independently, get around in the house [if necessary with a cane]?” (item 8), and “Can you, fully independently, stand up from sitting in a chair?” (item 3).

Educational level has been assessed according to the International Standard Classification of Education (ISCED) (UNESCO, 1976). The six levels of education vary from “basic education, not finished” to “university level, finished.” The highest level of education reflects formal education as well as occupational courses followed during the course of life.

Two cognitive performance tests were administered in the follow-up. One test, which is an adapted version of the Auditory Verbal Learning Test (AVLT; Lezak, 1983), refers to the acquisition and retention of new information. The original test involves five presentations of 15 words with free recall after each presentation, followed by delayed recall of this list after 20 minutes. We have used a three-trial version with free recall after every presentation. The score in this study is the sum score of the correctly recalled words in the three trials. The second cognitive performance test is part of the Groningen Intelligence Test (GIT; Luteijn & Van der Ploeg, 1983); the test measures verbal comprehension (or verbal intelligence). It is a 20-item multiple choice test measuring synonyms of the 20 words; the test is not age-dependent. The theoretical score ranges from 0 (no synonyms correct) to 20 (all synonyms correct).

Affective functioning (e.g., depressive symptoms and feelings of anxiety) was assessed using the two subscales of the Hospital Anxiety and Depression Scale (HADS; Alyard, Gooding, McKenna, & Snaith, 1987). Both were assessed at the baseline. The theoretical range of these 7-item scales varies from 0 to 21; higher scores indicate more symptoms. The internal reliability estimates were .71 (depressive symptoms) and .84 (feelings of anxiety).

Four aspects of personality were measured: neuroticism and extraversion (both using subscales of the revised version of the Eysenck Personality Questionnaire (EPQ-R; Eysenck, Eysenck, & Barrett, 1985), perceived physical competence (Bosscher, Laurijssen, & De Boer, 1992; Ryckman, Robbins, Thornton, & Cantrell, 1982), and mastery or personal control (Pearlin & Schooler, 1978). Neuroticism, extraversion, and mastery were assessed at the baseline, perceived physical competence in the follow-up. Higher scores on the personality measures indicate higher levels of neuroticism, extraversion, perceived physical competence, and mastery, respectively. The internal reliability estimates were .83, .80, .87, and .73, respectively.

Statistical Analysis
Multiple regression analyses (SPSS/PC 5.0.1) were conducted to examine (a) the extent of variance explained in self-reported ADL by performance-based ADL only, and (b) whether explained variance is added by each of the four groups of determinants of discrepancies: sociodemographic, cognitive, affective, and personality measures. For all cognitive, affective, and personality measures, separate scores...
SELF-REPORTED VERSUS PERFORMANCE-BASED ADL

were used as predictors. To analyze whether self-reported ADL is affected by the selected determinants of discrepancy in a consistent manner (e.g., consistently higher or lower levels of self-reported compared to performance-based levels), partial correlation coefficients were computed between each of the determinants and self-reported ADL, while controlling for performance-based ADLs. To determine whether the selected determinants are associated with the magnitude of discrepancies (either positively or negatively, indicating inaccuracy) between both types of measurement, correlation coefficients were computed between each of the selected sociodemographic, affective, cognitive, and personality measures and the absolute residual scores (obtained after regression analysis with performance-based ADL as the predictor and self-reported ADL as the outcome). Regression analyses were conducted for the ADL sum scores (11 items) as the outcome measure with all three test scores as predictors, as well as for the three separate self-reported ADL items with each of the three corresponding test scores as the predictor. The latter were: (a) "putting on/taking off a jacket" as the predictor and the self-reported item "can you, fully independently, dress yourself?" as the outcome; (b) "walking 6 meters" as the predictor and the self-reported item "can you, fully independently, get around in the house [with a cane if necessary]?" as the outcome; and (c) "getting up 5 times from a kitchen chair" as the predictor and the self-reported item "can you, fully independently, stand up from sitting in a chair?" as the outcome. Because of their skewed distributions, all three self-reported items were transformed logarithmically.

RESULTS

Table 2 shows the variance explained in self-reported ADL by the performance-based ADL measures only, as well as the additional variance explained by each group of determinants. The three performance-based tests explain most variance in the self-report GARS ADL sum score. The contribution of personality and affective functioning, however, is also substantial. This suggests that objective capacities account for self-reported ADL but that other factors also play a significant role.

Table 3 presents the partial correlation coefficients between the self-report ADL measures and the individual sociodemographic, cognitive, affective and personality measures while controlling for performance-based ADL. Subjects with lower levels of perceived physical competence, and to some lesser extent mastery or personal control, generally report lower levels of ADL functioning while controlling for differences due to performance-based ADL ("underestimation"). The same holds for depressive symptoms, although for the ADL dressing item, the association is not significant.

Univariate analyses (not in the table) showed that depressive symptomatology was significantly associated with self-reported ADL (correlation with GARS sum score, \( r = .18 \)), but not with the performance-based ADL measures. In addition, the associations between self-reported ADL and perceived physical competence and mastery were much stronger (\( r = -.42 \) and \( r = -.24 \), respectively) than the associations between both personality measures and the three performance-based ADL measures (ranging from \(-.14 \) to \(-.35 \) for physical competence and from \(-.08 \) to \(-.10 \) for mastery, respectively).

Table 4 shows Pearson correlation coefficients between the absolute residuals of performance-based ADL on self-reported ADL with the determinants of discrepancies. These coefficients reflect the impact of the individual sociodemographic, cognitive, affective, and personality measures on the accuracy ("over-" or "underestimation") of self-reported ADL relative to performance-based ADL. The results show that physical competence in particular, and to a lesser extent, mastery, influence the accuracy of the self-reported ADL. This means that among subjects with lower levels of perceived physical competence and mastery, relatively high discrepancies (either positive or negative) are observed. The accuracy of the self-reported ADL "getting around in the house" item is also influenced by affective and cognitive functioning and by age.

DISCUSSION

The results show that the association between self-report and performance-based measures of ADL is not strong. This

Table 2. Multiple Regression Analyses: Amount of Variance Explained in Self-Reported ADL Sum Score and (logarithmically transformed) ADL Items by Performance-Based Measures of ADL and Variance Added by Sociodemographic, Cognitive, Affective, and Personality Measures Each to Performance-Based Measures

<table>
<thead>
<tr>
<th></th>
<th>Self-report: ADL Sum Score</th>
<th>Dressing</th>
<th>Getting Around in the House</th>
<th>Standing Up From Sitting in Chair</th>
</tr>
</thead>
<tbody>
<tr>
<td>R²:</td>
<td>38.8%*</td>
<td>12.5%*</td>
<td>15.3%*</td>
<td>13.5%*</td>
</tr>
<tr>
<td>R²-change after performance-based measures included in the regression equation:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sociodemographic</td>
<td>0.7%</td>
<td>0.3%</td>
<td>0.1%</td>
<td>1.2%</td>
</tr>
<tr>
<td>Cognitive functioning</td>
<td>0.6%</td>
<td>0.5%</td>
<td>0.6%</td>
<td>0.3%</td>
</tr>
<tr>
<td>Affective functioning</td>
<td>1.6%*</td>
<td>0.6%</td>
<td>1.7%*</td>
<td>1.6%*</td>
</tr>
<tr>
<td>Personality</td>
<td>0.5%*</td>
<td>8.2%*</td>
<td>5.6%*</td>
<td>2.7%*</td>
</tr>
</tbody>
</table>

*Percentages reflect amount of variance explained by performance-based ADL measures in self-report ADL measures. For ADL sum scores all three performance-based measures have been included in the regression equation as independent variables. For self-report ADL items only, the equivalent performance-based ADL tests have been included as independent variables.

*p < .01.
Table 3. Partial Correlation Coefficients of Self-Reported ADL Sum Score and (logarithmically transformed) ADL Items with Sociodemographic, Cognitive, Affective, and Personality Measures While Controlling for Performance-Based Measures of ADL

<table>
<thead>
<tr>
<th></th>
<th>Self-report: *</th>
<th>ADL Sum Score</th>
<th>Dressing</th>
<th>Getting Around in the House</th>
<th>Standing Up From Sitting in Chair</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sociodemographic:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>.09</td>
<td>.04</td>
<td>.03</td>
<td></td>
<td>-.03</td>
</tr>
<tr>
<td>Gender *</td>
<td>.03</td>
<td>-.04</td>
<td>-.01</td>
<td></td>
<td>.11*</td>
</tr>
<tr>
<td>Educational level †</td>
<td>-.07</td>
<td>.00</td>
<td>-.01</td>
<td></td>
<td>-.03</td>
</tr>
<tr>
<td>Cognitive functioning:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AVLT total score of 3 trials ‡</td>
<td>-.10*</td>
<td>-.06</td>
<td>-.05</td>
<td></td>
<td>-.06</td>
</tr>
<tr>
<td>GIT verbal comprehension ‡</td>
<td>-.03</td>
<td>.01</td>
<td>.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Affective functioning:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depressive symptoms §</td>
<td>.16*</td>
<td>.08</td>
<td>.13*</td>
<td></td>
<td>.13*</td>
</tr>
<tr>
<td>Feelings of anxiety §</td>
<td>.06</td>
<td>.04</td>
<td>.12*</td>
<td></td>
<td>.09</td>
</tr>
<tr>
<td>Personality:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neuroticism †</td>
<td>-.01</td>
<td>-.03</td>
<td>.03</td>
<td></td>
<td>.04</td>
</tr>
<tr>
<td>Extraversion †</td>
<td>-.07</td>
<td>-.01</td>
<td>-.03</td>
<td></td>
<td>-.04</td>
</tr>
<tr>
<td>Physical competence †</td>
<td>-.28*</td>
<td>-.27*</td>
<td>-.23*</td>
<td></td>
<td>-.17*</td>
</tr>
<tr>
<td>Mastery †</td>
<td>-.20*</td>
<td>-.16*</td>
<td>-.16*</td>
<td></td>
<td>-.08</td>
</tr>
</tbody>
</table>

*Higher self-report scores indicate poorer functioning. Coefficients are correlation coefficients between self-report ADL and the 11 selected predictors of discrepancies while controlling for all three performance-based ADL measures in case of self-report ADL sum scores and while controlling for each equivalent performance-based ADL measure in the case of self-report ADL items.

† Higher scores indicate higher educational levels.

‡ Higher scores indicate better functioning.

§ Higher scores indicate poorer functioning.

* p < .01.

Table 4. Correlation Coefficients of Absolute Residuals of Self-Reported ADL Sum Score and (logarithmically transformed) ADL Items and Performance-Based Measures of ADL with Sociodemographic, Cognitive, Affective, and Personality Measures

<table>
<thead>
<tr>
<th></th>
<th>Self-report: *</th>
<th>ADL Sum Score</th>
<th>Dressing</th>
<th>Getting Around in the House</th>
<th>Standing Up From Sitting in Chair</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sociodemographic:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>-.04</td>
<td>.08</td>
<td>.10*</td>
<td></td>
<td>.00</td>
</tr>
<tr>
<td>Gender *</td>
<td>-.05</td>
<td>-.06</td>
<td>-.01</td>
<td></td>
<td>.06</td>
</tr>
<tr>
<td>Educational level †</td>
<td>-.05</td>
<td>.00</td>
<td>-.07</td>
<td></td>
<td>-.04</td>
</tr>
<tr>
<td>Cognitive functioning:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AVLT total score of 3 trials ‡</td>
<td>-.04</td>
<td>-.10*</td>
<td>-.10*</td>
<td></td>
<td>-.06</td>
</tr>
<tr>
<td>GIT verbal comprehension ‡</td>
<td>-.06</td>
<td>-.02</td>
<td>-.10*</td>
<td></td>
<td>.00</td>
</tr>
<tr>
<td>Affective functioning:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depressive symptoms §</td>
<td>.12*</td>
<td>.08</td>
<td>.16*</td>
<td></td>
<td>.07</td>
</tr>
<tr>
<td>Feelings of anxiety §</td>
<td>.05</td>
<td>.01</td>
<td>.10*</td>
<td></td>
<td>.05</td>
</tr>
<tr>
<td>Personality:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neuroticism †</td>
<td>.01</td>
<td>-.06</td>
<td>.01</td>
<td></td>
<td>.02</td>
</tr>
<tr>
<td>Extraversion †</td>
<td>-.04</td>
<td>-.01</td>
<td>-.04</td>
<td></td>
<td>-.02</td>
</tr>
<tr>
<td>Physical competence †</td>
<td>-.18*</td>
<td>-.26*</td>
<td>-.31*</td>
<td></td>
<td>-.14*</td>
</tr>
<tr>
<td>Mastery †</td>
<td>-.18*</td>
<td>-.16*</td>
<td>-.18*</td>
<td></td>
<td>-.09</td>
</tr>
</tbody>
</table>

*Higher scores indicate poorer functioning. Coefficients are Pearson correlation coefficients between the 11 selected predictors of discrepancies and the absolute residuals obtained after four separate multiple regression analyses with performance-based and self-report ADL measures: one analysis with all three performance-based ADL measures as independent variables and the self-report ADL sum score as dependent variable, and three analyses each with one performance-based ADL measure as independent variable and one equivalent self-report ADL item as dependent variable.

† = male, 2 = female.

‡ Higher scores indicate higher educational levels.

* Higher scores indicate better functioning.

† Higher scores indicate poorer functioning.

§ Higher scores indicate higher levels of extraversion, physical competence, or mastery/personal control.

* p < .01.
is consistent with recent reports (Cress et al., 1995; Reuben et al., 1995). We found that depressive symptoms, in addition to perceived physical competence and mastery, enlarge discrepancies between self-reported and performance-based ADL. Overestimation, i.e., higher self-reported levels of ADL compared to performance-based levels, occurs particularly among subjects with high perceptions of physical competence and mastery, and, to a lesser extent, with low levels of depressive symptomatology. This indicates that subjects with lower levels of perceived physical competence and mastery (or personal control), and higher levels of depressive symptomatology, tend to report lower levels of ADL compared to their performance-based ADL level. Furthermore, our results show that the perception of physical competence in particular and, to a lesser extent, mastery, affect the accuracy of the self-perceived level of ADL functioning. This means that among subjects with lower levels of perceived physical competence and mastery, relatively high discrepancies (either positive or negative) are observed. In contrast to what was previously suggested (Guralnik et al., 1989), we found no effects of education on the discrepancy between performance-based and self-report measures.

Our results suggest that discrepancies between performance-based and self-reported ADL are not strongly associated with cognitive functioning. However, 78 subjects with severe cognitive impairments (MMSE < 17) were excluded at the baseline. Although the MMSE scores in our subsample ranged from 17 to 30 (mean: 27.2; SE: 2.2) and the number of subjects excluded at baseline was rather low, the association of the size of the discrepancies with cognitive functioning is probably reduced by this baseline exclusion criterion.

Subjects for whom no timed score was available were assigned a performance-based test score equal to the worst actual score. In contrast to performance-based ADL, no missing scores were observed for self-reported ADL. Nonresponse analyses showed significantly poorer self-reported ADL functioning for nonresponders compared to responders, for each of the three performance-based tests. This supports the hypothesis that, in particular, elderly people with high “frailty” levels were not able or refused to participate in the performance-based tests. In contrast, Tinetti and Ginter (1988) and Myers et al. (1993) found no difference in frailty between responders and nonresponders, but nonresponse rates in their studies were considerably higher than in our study. However, the level of frailty was probably higher in our sample.

It should be noted that only three very basic ADL items were involved in our study. The main reason for this was that we intended to focus on self-report and performance-based measures which reflect identical aspects of functioning (e.g., disability). However, it is of interest whether the selected determinants of discrepancies also affect discrepancies between self-report and performance-based measures of more complex activities (i.e., instrumental activities of daily living; IADLs). Additional research is necessary to study this issue.

The question arises as to which type of measurement is most appropriate in which circumstance. Recently, it was shown that performance-based measures should not be viewed as being “superior” (i.e., more acceptable to patients, clinically feasible, reproducible, sensitive to change) compared to self-report measures (Myers et al., 1993). Because performance-based tests are less easy to carry out in most data-collection circumstances (e.g., special equipment and special training is needed), self-report measures will continue to be used. The results showed that self-reported levels of ADL were independently influenced by depressive symptomatology, mastery, and perceived physical competence. In contrast to the results of Cress et al. (1995), but in line with the results of Rozzini et al. (1993), we found no significant associations between depressive symptomatology and performance-based measures of ADL. Self-reported ADL and depressive symptoms were significantly associated. In addition, we found some evidence that cognitive functioning plays a role in the concordance between self-report and performance-based measures of ADL, but the effects were relatively small. The effect of age, gender, and educational level on the discrepancies is even smaller or absent. Stronger associations of perceived physical competence and mastery with self-reported ADL were found than with performance-based ADL. These results suggest that self-report measures of ADL are more strongly “confounded” by personality and affective factors than performance-based measures. Although self-report ADL measures are easier to administer and less sensitive to nonresponse than performance-based ADL measures, the effects of perceived physical competence, mastery, and depressive symptomatology should be considered in any application of self-report ADL measures among frail elderly.

ACKNOWLEDGMENTS

This research is part of the Groningen Longitudinal Aging Study (GLAS), conducted by the Northern Centre for Health Care Research (NCG) and various departments of the University of Groningen (RUG). The primary departments involved are Health Sciences (Prof. dr W.J.A. van den Heuvel), Family Medicine (Prof. dr B. Meyboom-de Jong), Psychiatry (Prof. dr J. Ormel), Sociology (ICS) (Prof. dr S.M. Lindenberg) and Human Movement Sciences (Prof. dr P. Rispens). Director of GLAS is prof. dr J. Ormel (Health Sciences and Psychiatry). GLAS and its substudies are financially supported by the Dutch government (through NESTOR), the University of Groningen, the School of Medicine, the Dutch Cancer Foundation (NKB/KWF), and the Netherlands Organization for Scientific Research (NWO).

Address correspondence to Dr. G.J.M. Kempen, Northern Centre for Health Care Research, School of Medicine, University of Groningen, A. Deusinglaan 1, 9713 AV Groningen, The Netherlands. E-mail: G.J.M.Kempen@med.rug.nl

REFERENCES


Received June 28, 1995
Accepted March 5, 1996