Original Article

Journal of Aging and Health
2020, Vol. 32(9) 1084–1097

DOI: 10.1177/0898264319884404
journals.sagepub.com/home/jah

Psychosocial and Health-Related Factors Associated With Discordance Between 13-Year Trajectories of Self-Reported Functional Limitations and Performance-Based Physical Functioning in Old Age

Almar A. L. Kok, PhD[1], Marieke J. Henstra, MD, MSc[2], Nathalie van der Velde, MD, PhD[2], Didi Rhebergen, MD, PhD[1], and Natasja M. van Schoor, PhD[1]

Abstract

Objective: The objective of this study was to examine correlates of discordance between 13-year trajectories of self-reported functional limitations and performance-based physical functioning in older adults. **Method:** We included 2,135 participants from the population-based Longitudinal Aging Study Amsterdam, the Netherlands, followed across 1995-2008. Self-reported functional limitations included six (instrumental) activities of daily living. Performance-based functioning was a composite of four tests. We used latent class growth analysis and multinomial logistic regression to examine discordance and its correlates. **Results:** Patterns of discordance and concordance (41% concordance, 46% "overestimation" of daily functioning, 13% "underestimation") appeared to be persistent over 13 years. Older age, male sex, cognitive impairment, absence of pain, and light alcohol use were associated with overestimation. Younger age, female sex, and lower self-rated health were associated with underestimation. **Discussion:** Factors associated with overestimation partly differ from those associated with underestimation. Factors that are highly stable over time are particularly good indicators of persistent discordance.

Keywords

disability, health behaviors, cognitive function, physical function, social factors

Background

Preservation of physical functioning is an important goal in old age, from a public health perspective as well as from older adults' own perspective (Bowling, 2006; Guralnik, Fried, & Salive, 1996). Physical performance tests and functional limitations are two dominant modes of assessing physical functioning. Self-reports of functional limitations measure the extent to which one experiences limitations in activities of daily living (ADLs) and/or instrumental ADLs (iADLs), whereas performance tests measure the actual performance of specific actions, such as walking and standing up from a chair (Wittink, Rogers, Sukiennik, & Carr, 2003). The correlation between these measures is usually moderately strong. This indicates that while there is overlap, there are also individuals with substantial discordance, meaning that the level of daily functioning differs from what would be expected given a certain level of performance on physical tests (Brach, VanSwearingen, Newman, & Kriska, 2002; Daltroy, Larson, Eaton, Phillips, & Liang, 1999; Kempen et al., 1996). When older adults experience more limitations than would be expected based on performance tests, the literature typically refers to this as "overestimation," whereas

[1]Amsterdam UMC—Vrije Universiteit Amsterdam, The Netherlands
[2]Amsterdam UMC—University of Amsterdam, The Netherlands

Corresponding Author:
Almar A. L. Kok, Amsterdam UMC—Vrije Universiteit Amsterdam, Epidemiology & Biostatistics, Medical Faculty, F-Wing, De Boelelaan 1089a, 1081 HV Amsterdam, The Netherlands.
Email: a.kok1@amsterdamumc.nl

the opposite is referred to as "underestimation," and we follow this terminology throughout the present study.

From a methodological point of view, the imperfect correlation between functional limitations and performance-based physical functioning is partly due to the fact that they capture different concepts. However, particularly if one is interested in global physical functioning, which includes ADLs and more complex iADLs, such as using own or public transportation, no performance-based tests are available that can incorporate the daily context in which such tasks are executed, and a certain extent of discrepancy is inevitable. Furthermore, specific events such as injuries or falls might affect one measure more strongly than another, leading to a temporary discordance between them (Feuering, Vered, Kushnir, Jette, & Melzer, 2014). Despite these methodological issues, discrepancies between the two measures can be informative for our theoretical understanding of heterogeneity in disability processes, and in clinical practice (Kempen et al., 1996). Longitudinal studies are needed to examine discordance beyond temporary fluctuations of functioning, across long-term aging processes.

A considerable amount of studies shows that discordance between functional limitations and performance-based functioning is systematically related to individual psychosocial and health-related characteristics (e.g., Cress et al., 1995; Feuering et al., 2014; Goverover et al., 2005; Shulman et al., 2006). Although the findings are inconsistent across studies, they indicate that methodological issues alone are not sufficient to fully explain discordance, and that it is a phenomenon with important theoretical and clinical implications. For example, when individuals report no functional limitations while demonstrating poor performance on tests (overestimation), they may be at increased risk of injuries or falls (Delbaere, Close, Brodaty, Perminder, & Lord, 2010; Sakurai et al., 2013). Vice versa, when individuals underestimate their capacities for daily functioning, this may indicate suboptimal utilization of one's physical potential (Kempen, Sullivan, van Sonderen, & Ormel, 1999).

However, although various models on disability emphasize gradual aging processes that may lead to the emergence of discordance between performance-based and self-reported daily functioning, previous studies have been cross-sectional. Therefore, instead of examining discordance observed at a single point in time, in this study we examined psychosocial and health-related correlates of discordance, based on 13-year longitudinal trajectories of self-reported functional limitations and performance-based functioning in a large, population-based cohort of older adults.

Conceptual Framework and Hypotheses

The National Health and Aging Trends Study (NHATS) disability framework (Freedman, 2009) posits that disability encompasses several domains. One domain is *capacity*—referring to basic bodily and mental functions required to perform daily activities, which can be measured through performance tests. A second domain is the ability to carry out essential self-care and domestic activities, which are often measured through self-reported functional limitations. So-called *accommodations* encompass compensatory strategies that link capacity to daily functioning, and these include assistive technology and adaptations in the direct environment (Freedman, 2009; Kasper et al., 2009). Researchers applying the NHATS framework have shown that physical performance tests are more sensitive to early declines in capacity that represent earlier stages of the disablement process (Kasper, Chan, & Freedman, 2016). This suggests that declines in performance-based functioning typically precede declines in self-reported daily functioning, leading to a gradual increase of "overestimation" with increasing age. Moreover, *accommodations* may influence this potential discordance, leading to the possibility that when examined over a longer period of time, the extent of overestimation may remain limited, increase or decrease, depending on the extent to which accommodations are successful in compensating for declines in physical capacity over time.

However, the NHATS model does not explicitly incorporate the possibility of *underestimation* of daily functioning. Nevertheless, empirical evidence clearly suggests that there are groups of older adults who report more functional limitations than would be expected given their level of performance-based functioning. This "underestimation" may indicate substantial underutilization of one's physical capacity. Empirical findings showing that underestimation of daily functioning is linked to factors such as depressive or anxiety symptoms and a low sense of control (Daltroy et al., 1999; Kempen et al., 1999; Shulman et al., 2006; Wittink et al., 2003) suggest that for some older adults, declines in emotional well-being with aging may partly explain the emergence of underestimation over time. However, as for overestimation, few studies have examined which factors are associated with patterns of discordance between performance-based and self-reported daily functioning over an extended period of time.

Nevertheless, based on the theoretical premise that declines in performance-based functioning tends to precede declines in self-reported daily functioning, and on cross-sectional findings that the opposite pattern may also apply to some individuals, we expect the following:

> **Hypothesis 1:** There are subgroups of older adults in whom their trajectory of self-reported functional limitations substantially differs from their trajectory of performance-based functioning, in level and/or shape.

Furthermore, we argue that several individual characteristics and resources may affect the transition from declines in capacity to declines in daily functioning. For example, demographic characteristics (such as age and sex), psychosocial resources (such as anxiety), and health factors (such as

cognitive functioning) have been shown to affect physical functioning in general (Stuck et al., 1999) and the relationship between capacity and functional limitations in particular (Brach et al., 2002; Daltroy et al., 1999 Kempen et al., 1996). Moreover, psychosocial factors have been suggested as modifiers of the disablement process (Verbrugge & Jette, 1994). In addition to the accommodations mentioned in the NHATS model, such factors may be crucial to identify profiles of older adults who persistently overestimate or underestimate their daily functioning relative to their physical capacities. Therefore, we formulate the general expectation:

> **Hypothesis 2:** Demographic, psychosocial, and health-related characteristics (including health behaviors) are associated with discordance between trajectories of self-reported functional limitations and performance-based functioning.

Present Study

By adopting an exploratory approach, we aim to empirically highlight specific individual factors related to long-term patterns of discordance in old age, aiming to improve theoretical understanding of individual variation in the disablement process and deriving useful information for clinical practice.

First, in a large population-based sample, we compare 13-year trajectories of self-reported functional limitations with trajectories of performance-based functioning. This approach allows for discordance to decrease or increase over time and distinguishes subgroups with specific long-term patterns of discordance. Second, we provide a detailed profile of psychosocial and health-related characteristics in these subgroups. Third, in accordance with the NHATS instrument for measuring physical capacity (Kasper et al., 2016), we use a comprehensive measure of performance-based functioning based on four different tests. This approach acknowledges the fact that performance tests measure relatively specific, isolated functions, implying that information on multiple types of performance tests is needed to reflect the more complex daily activities on which measures of functional limitations are based (e.g., walking outside, using public transport) (Feuering et al., 2014). Finally, rather than focusing on specific high-risk groups, such as those with multiple sclerosis (Goverover et al., 2005), Parkinson's disease (Shulman et al., 2006), and severe functional disability (Kempen et al., 1999), our analysis is more generalizable to the aging population at large, and therefore relevant for public health in general.

Method

Study Sample

We used data from the Longitudinal Aging Study Amsterdam (LASA), a prospective cohort study carried out since 1992 in the Netherlands, aiming to study the course and consequences of physical, mental, and social functioning of older adults (Hoogendijk et al., 2016). At baseline, LASA included 3,107 respondents aged 55 to 84 years who were randomly sampled from the registries of 11 municipalities in three distinct geographic regions (response rate 62%). Data were gathered in face-to-face interviews at the participants' homes, medical interviews, and postal questionnaires. Participants were followed up every 3 years. The LASA study was approved by the Medical Ethics Review Board of the VU medical center, and all participants provided written informed consent.

As the second measurement wave was the first to provide all physical performance tests that we wanted to include, the baseline of the current study is the 1995/1996 wave (n = 2,545). Trajectories of functioning were based on this measurement and four follow-up measurements, in 1998/1999 (n = 2,076), 2002/2003 (n = 1,691), 2005/2006 (n = 1,257), and 2008/2009 (n = 985). Sample attrition in-between waves was on average 21%, which was mainly due to mortality (16%). The selection of psychosocial and health-related characteristics was guided by the previous literature on discordance and on data availability, and taken from the 1995/1996 wave. The analytic sample included respondents with data on at least three out of four performance tests (n = 2,135).

Measures

Self-reported functional limitations. For assessing self-reported functional limitations, we constructed a scale based on six items measuring whether one experienced limitations in carrying out the following ADLs and iADLs: dressing and undressing oneself, cutting one's own toenails, walking up and down a staircase of 15 steps without resting, sitting down and standing up from a chair, using own or public transportation, and walking outside for 5 min without resting. Answer categories ranged from 1 (no) to 4 (yes, without difficulty) and were summed to form a scale ranging from 6 to 30. Higher scores indicate better self-reported functioning.

Performance-based physical functioning. As we wanted the performance-based measure of physical functioning to reflect the capacities needed for daily functioning as closely as possible, we included four physical performance tests measuring lower and upper extremity function and balance: (a) chair stand: time to stand up from and sit down on a chair 5 times without using the arms; (b) walking: time to walk 3 meters and back; (c) cardigan: time to put a cardigan on and take it off; and (d) tandem stand: time the participant was able to stand still on a straight line with one foot directly before the other. Similar to previous research (Koster et al., 2006), we divided the scores on each of these measures into quartiles determined at baseline, and assigned scores to these quartiles: 0 (not able or not successfully completed), and 1 (lowest performing quartile) to 4 (highest performing quartile). As few participants could not hold the tandem stand for 10 s or more, this measure was

categorized differently: 0 (unable to 2 s), 2 (3-9 s), and 4 (10 s). We included participants with valid data on three out of four tests (n = 158), and then computed the mean score of all tests, because using a sum score would result in a different maximum score (i.e., 12 instead of 16 points) for those with only three valid tests. The inclusion of this group did not notably affect the mean scores and their variability.

Demographic factors. Age (continuous), sex (0 = male; 1 = female), and years of education (range between 5 and 18) were included.

Psychological factors. For depressive symptoms, we used the 20-item Center for Epidemiologic Studies Depression scale (CES-D; Radloff, 1977), measuring core symptoms of depression in the past week. Response categories ranged from 0 (<1 day) to 3 (5-7 days). The scale ranges from 0 to 60.

Anxiety symptoms were measured with a Dutch translation of the anxiety subscale of the Hospital Anxiety and Depression Scale (HADS; Zigmond & Snaith, 1983), including seven items on anxiety symptoms in the past 4 weeks, with a total scale score of 0 to 21.

Cognitive functioning was measured using the Mini-Mental State Examination (MMSE; Folstein, Folstein, & McHugh, 1975). Following conventions, we dichotomized this into MMSE <24 versus higher to indicate (mild) cognitive impairment (Tombaugh & McIntyre, 1992).

We included four psychological measures related to individuals' sense of control and self-esteem. Mastery reflects the extent to which one feels that one's life is under one's own control as opposed to being determined by external forces. It was measured by a five-item version of the Pearlin Mastery Scale (Pearlin & Schooler, 1978), ranging from 5 to 25. Global self-esteem indicates participants' overall evaluation of their own worth and was measured by a four-item version of the Rosenberg Self-Esteem Scale (Rosenberg, Schooler, Schoenbach, & Rosenberg, 1995), with a total scale score of 4 to 20. Self-efficacy reflects one's evaluation of the ability to attain goals. It was measured by a 12-item version of the General Self Efficacy Scale with a range of 12 to 60 (Sherer & Maddux, 1982). Finally, neuroticism was measured by 15 items from the Dutch Personality Questionnaire (Barelds & Luteijn, 2002) that indicate participants' tendency to experience distress ("emotional instability"). The total scale score was 0 to 30.

Social factors. Instrumental support reflected the extent to which the participant received help with tasks in and around the house from the nine most important social network members, excluding the partner (Van Tilburg, 1998). The scale ranged from 0 (no help from any network member) to 36 ("often" help from all network members). We dichotomized partner status as having a partner living inside the household (0) versus no partner or a partner living outside the household (1), assuming that partners living outside the household are less able than those inside the household to provide help with daily tasks.

Health-related factors. Self-rated health was asked with a single question, "how is your health in general," and recoded into 1 (poor) to 5 (excellent). The number of chronic diseases (observed range = 0-8) was based on the self-reported presence of cerebrovascular accident or stroke, diabetes mellitus, cardiac disease, peripheral arterial disease, asthma or chronic obstructive pulmonary disease, nonspecific lung disease, rheumatoid arthritis, osteoarthritis, cancer, or other diseases. Subjective pain was assessed with five items, for example, "I am in constant pain" (0: no; 1: yes), summed to a score of 0 to 5.

Physical activity expressed the total number of hours spent doing light or heavy household work, walking outside, biking, and performing up to two sports in the past 2 weeks, weighted according to Metabolic Equivalent of Task (MET) scores (Ainsworth et al., 2011). MET scores were divided by 10 for easier interpretation of effect sizes. Body mass index (BMI) was based on measured weight in kilograms divided by height in meters squared, and categorized as underweight or normal (BMI < 25), overweight (BMI = 25-30) and obese (BMI ≥ 30). Smoking status was categorized as never smoked or stopped ≥15 years ago, former smoker who stopped <15 years ago, and current smoker. Alcohol use was based on the frequency of drinking alcohol combined with the number of glasses a participant consumed on average each time (Garretsen & Knibbe, 1983), and categorized as no use, light, moderate, and (very) excessive.

Analytical Procedure

The first step in the analysis was to determine the extent to which participants' trajectories of functional limitations and performance-based physical functioning corresponded. Because of our longitudinal focus, we could not simply subtract performance scores from self-reported daily functioning scores to measure discordance. Therefore, we first classified individuals' separate trajectories of self-reported and performance-based functioning over time using latent class growth analysis (LCGA; Nagin, 1999) in Mplus v7 (Muthén & Muthén, 2012), and then constructed subgroups based on individuals' combination of both trajectories.

Classifying trajectories was an iterative process. We started with two subgroups and then repeatedly added an additional subgroup while comparing model fit statistics to the previous model with one subgroup less. Following recommendations from the literature, optimal fit was determined on the basis of lower sample size–adjusted Bayesian information criterion (SSA-BIC), likelihood ratio test (LRT) p value <.05, smallest latent class contains ≥5% of the sample, and interpretability (B. Muthén, 2003). LCGA is probabilistic, meaning that there may be statistical uncertainty in assigning participants to latent classes. Therefore, we evaluated classification accuracy, on the basis of the "entropy" value.

We estimated models with linear and quadratic slopes and then selected the model with the best fit. Furthermore, because for functional limitations at baseline, 49% of the participants had the maximum score of 30, we estimated a "censored" model that accounted for this ceiling effect. Missing follow-up data were handled by maximum likelihood estimation.

After deciding on the best fitting models, variables expressing the classes with the highest posterior probability for each participant were exported to SPSS v22 for further analysis. In SPSS, we made concordant and discordant groups based on individual combinations of latent class membership in self-reported functional limitations and performance-based functioning.

Finally, we obtained complete data on potential correlates of discordance using multiple imputation with 20 datasets. We examined unadjusted differences in these correlates between the groups based on discordance using *F*-tests and chi-square tests. Adjusted associations between the correlates and discordance were obtained using multinomial logistic regression models, with concordance as the reference group. These models were first adjusted for age and sex (Model 1). Subsequently, we used a forward selection procedure available in SPSS to examine a final model including the set of factors that explained the most variance in discordance patterns, setting statistical significance for the selection to $p < .10$ (Model 2). As the 20 imputed datasets showed some differences in selected variables, we included only those in the final model that were selected in at least five imputed datasets.

Results

Sample Selectivity

The analytic sample with sufficient data on functional limitations and performance-based physical functioning (n = 2,135) was younger, had more years of education, and had better physical and mental health than those excluded from the analyses (n = 410).

Descriptive Statistics

Mean age in the study sample was 72 years (SD = 8.5), 53% was female, and mean education was 9 years (SD = 3.4; Table 1). The mean functional limitations score was 27 (SD = 4.0) and the mean physical performance score was 2.6 (SD = 0.9). Descriptive statistics of psychosocial and health-related factors can be found in Table 1.

Latent Class Analyses

Model fit statistics (SSA-BIC and LRT p value) indicated that for both trajectories of functional limitations and performance-based functioning, a six-class model with a quadratic slope fitted best to the data, but had a very small class (<5% of the sample; Table 2). Therefore, we accepted the five class solutions. Entropy values were relatively low; 0.72 for functional limitations and 0.64 for performance-based functioning, indicating some uncertainty in classification. Because of this, we carried out two sensitivity analyses (see Supplemental Material). These showed that despite the low entropy values, the final grouping of participants into discordant and concordant groups was reasonably consistent with continuous baseline discordance scores.

We labeled the five trajectories of functional limitations as "high and stable" (25% of the sample), "high initial functioning with some decline" (37%), "some initial limitations with decline" (20%), "multiple initial limitations with decline" (11%) and "many initial limitations with some decline" (7%; Figure 1a). The trajectories of performance-based functioning were labeled "high and stable" (19%), "high initial functioning with decline" (27%), "intermediate initial functioning with decline" (25%), "intermediate initial functioning with steep decline" (10%), and "low and stable" (19%; Figure 1b). From this point forward we refer to these trajectories as "A" to "E" (arranged from high to low functioning).

Discordance Patterns

We identified five discordance patterns by cross-tabulating the classification of functional limitations and performance-based trajectories of functioning (Table 3). Participants who had the same classification in functional limitations and performance-based functioning were labeled "concordant" (n = 867; 41%). Those whose classification in self-reported daily functioning was one class higher than in performance-based functioning were labeled "slight overestimation" (n = 683; 32%), and if self-reported daily functioning was two or more classes higher than performance-based functioning, this was labeled "strong overestimation" (n = 314; 14%). The opposite was done when classification in self-reported daily functioning was lower than in performance-based functioning ("slight underestimation": n = 243; 11%; "strong underestimation": n = 38; 2%). The finding that the latent class analyses of both indicators of physical functioning yielded an equal number of subgroups with trajectories that did not cross each other suggests that discordance was relatively stable over time. Nevertheless, in the Supplemental Material, we provide detailed graphs which demonstrate that within the discordant groups based on combinations of the latent classes, overestimation and underestimation tend to increase and then decrease with aging, while the extent of discordance is small and stable in the "concordant" group.

Given the low prevalence of underestimation, which resulted in low statistical power for these groups, we decided to combine the slight and strong underestimation groups. For consistency and ease of interpretation, we also combined the slight and strong overestimation groups. Our main results are thus based on a three-group classification: overestimation, concordance, and underestimation. However, detailed results

Table 1. Baseline Characteristics of the Study Sample ($N = 2,135$).

	Observed				Imputed
Variable (observed range)	*n*	% missing	*M*/%	*SD*	*M*/%
Demographics					
Age (57-86)	2,135	0	72.1	8.5	n/a
Sex (% female)	2,135	0	52.6		n/a
Years of education (5-18)	2,133	<0.1	9.1	3.4	9.1
Physical functioning					
Functional limitations (6-30)	2,135	0	27.4	4.0	n/a
Physical performance (0-4)	2,135	0	2.6	0.9	n/a
Psychological factors					
Depressive symptoms (0-60)	2,126	0.4	7.8	7.7	7.8
Anxiety symptoms (0-21)	2,133	<0.1	2.7	3.3	2.7
Cognitive decline (% MMSE score ≤23)	2,133	<0.1	10.2		10.2
Mastery (5-25)	2,107	1.3	17.4	3.3	17.4
Self-esteem (5-20)	2,128	0.3	15.3	2.2	15.3
Self-efficacy (12-60)	2,124	0.5	42.0	5.3	41.9
Neuroticism (0-30)	1,874	12.2	5.8	5.7	5.9
Social factors					
Instrumental support (0-36)	2,116	0.9	14.8	6.7	14.8
Partner in the household (%)	2,135	0	60.5		n/a
Health-related factors					
Self-rated health (1-5)	2,134	<0.1	3.6	0.9	3.6
No. of chronic diseases (0-8)	2,134	<0.1	1.6	1.3	1.6
Pain (0-5)	1,731	18.9	0.6	1.3	0.8
Physical activity (MET / 10)	2,122	0.6	5.8	3.9	5.8
BMI (kg/m^2)	1,397	34.6			
<25	476		34.1		32.1
25-30	633		45.3		45.0
≥30	288		20.6		22.9
Smoking (%)	1,405	34.2			
Never/stopped >15 years	923		66.2		61.6
Former (stopped ≤15 years)	202		14.5		15.3
Current	270		19.4		21.4
Alcohol use	1,404	34.2			
Does not drink	344		24.5		22.9
Light	699		49.8		49.6
Moderate	277		19.7		20.3
(Very) Excessive	84		6.0		7.2

Note. MMSE = Mini-Mental State Examination; MET = Metabolic Equivalent of Task; BMI = body mass index.

from analyses with the original five groups are available in the Supplemental Material and are referred to here when informative.

Unadjusted Analyses

Unadjusted analyses showed that, on average, the overestimation group had the highest age, the lowest proportion of women and persons with a partner in the household, the highest self-rated health, and the lowest levels of pain, physical activity, obesity and excessive drinking (Table 4). The underestimation group was youngest, had the highest level of depressive and anxiety symptoms and neuroticism, and relatively unfavorable characteristics in terms of health-related factors. Compared with the concordant group, mastery and self-efficacy were lower in both discordant groups.

Adjusted Analyses

We found that age had a quadratic rather than linear effect on overestimation, but not on underestimation. When divided into five quintiles, compared with the youngest quintile (ages 55-63), odds of overestimation increased in each subsequent quintile, ages 64-68: odds ratio (OR) = 1.6; ages 69-73: OR=2.5; ages 74-80: OR=3.8, and then declined in the fifth quintile (OR = 3.1). The linear effect of

Table 2. LCGA Model Fitting Process (n = 2,135).

				% in classes					
Classes	SSA-BIC	Entropy	LRT p value	1	2	3	4	5	6
Self-reported functional limitations (quadratic slope)									
2	30,468	0.77	<.001	67	33				
3	294,57	0.76	<.001	43	41	16			
4	29,013	0.73	<.001	38	29	20	13		
5[a]	28,894	0.72	0.03	37	25	21	11	7	
6	28,810	0.71	0.049	35	29	18	10	6	2[b]
Performance-based physical functioning (quadratic slope)									
2	15,625	0.76	<.001	53	47				
3	14,960	0.73	0.001	39	36	24			
4	14,714	0.68	<.001	29	28	22	20		
5[a]	14,647	0.64	0.001	28	25	19	19	10	
6	14,584	0.64	0.02	25	23	19	19	10	4[b]

Note. LCGA = latent class growth analysis; SSA-BIC = sample size–adjusted Bayesian information criterion; LRT = p value for the Vuong–Lo–Mendell–Rubin–adjusted likelihood ratio test.
[a]This is the selected model.
[b]Prevalence considered too low.

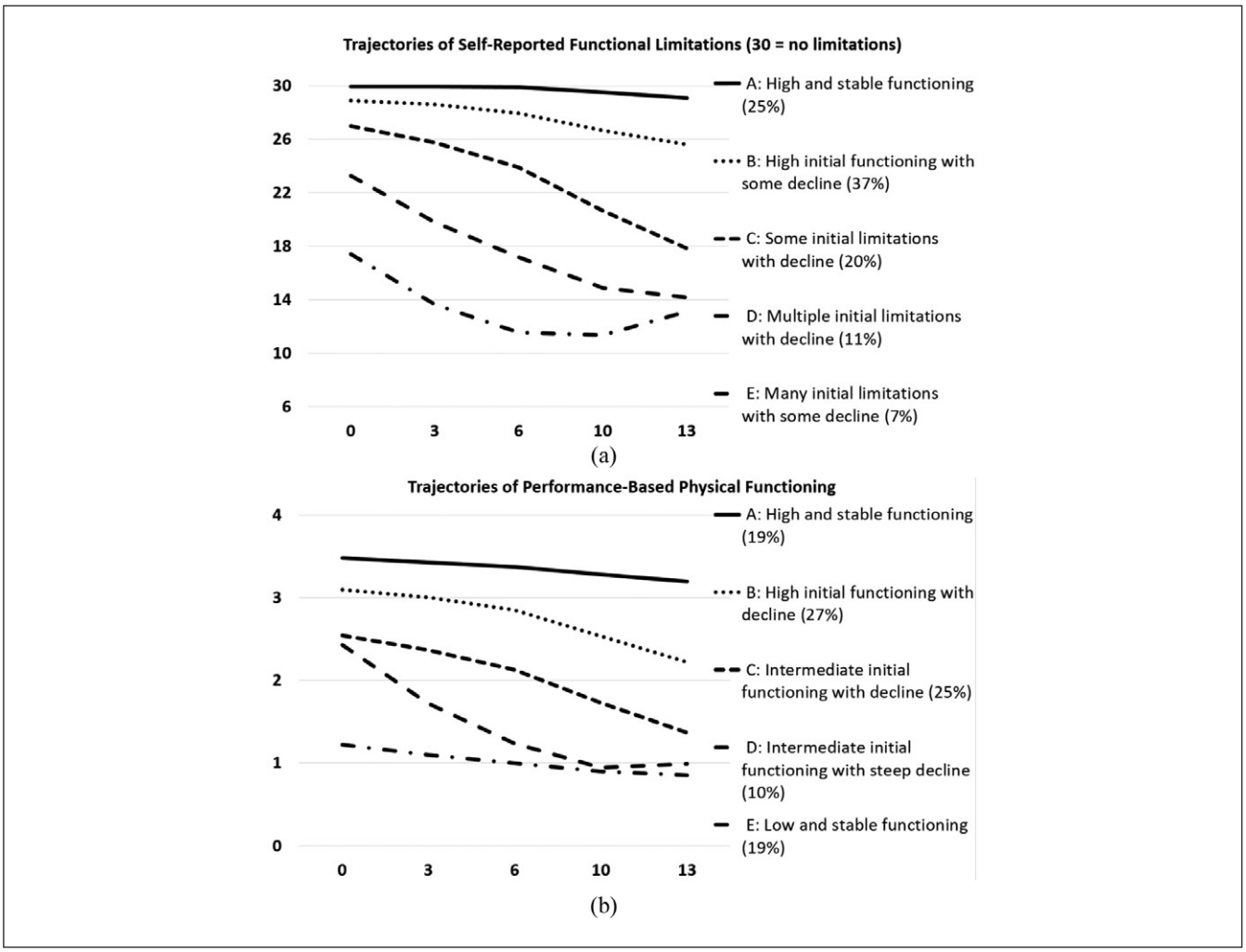

Figure 1. Typology of trajectories of self-reported functional limitations (a) and performance-based (b) physical functioning, based on latent class growth analysis.

Table 3. Cross-Tabulation of Trajectories of Functional Limitations and Performance-Based Physical Functioning and Prevalence of Concordance/Discordance Patterns.

		Performance-based (%)				
Trajectory[a]		A	B	C	D	E
Self-report (%)	A	12.4	11.1	4.0	0.6	0.7
	B	6.3	12.6	10.9	2.4	2.4
	C	0.5	3.0	7.4	4.2	4.1
	D	0.0	0.6	1.9	2.3	5.8
	E	0.0	0.1	0.4	0.2	5.9
Somer's *D*: .57, $p < .001$						

Concordance/discordance patterns				
Pattern	%	*n*	Three groups[b]	
Strongly overestimating	14.2	304	46.2%	$n = 987$
Slightly overestimating	32.0	683		
Concordant	40.6	867	40.6%	$n = 867$
Slightly underestimating	11.4	243	13.0%	$n = 281$
Strongly underestimating	1.6	38		

[a]Self-report trajectories: A = high and stable functioning; B = high initial functioning with some decline; C = some initial limitations with decline; D = multiple initial limitations with decline; E = many initial limitations with some decline; performance-based trajectories: A = high and stable functioning; B = high initial functioning with decline; C = intermediate initial functioning with decline; D = intermediate initial functioning with steep decline; E = low and stable functioning.
[b]Overestimation; concordance; underestimation.

age on underestimation showed lower odds of underestimation with older age (Table 5). For ease of interpretation, we present only linear effects of age, but as we found that the multinomial model with age-squared showed better overall fit than the model without (chi-square difference = 16, $df = 2$, $p < .001$), all models are additionally adjusted for age-squared.

Age, age-squared, and sex adjusted effects of single factors on discordance showed that male sex, cognitive impairment, lower self-efficacy, less pain, normal weight, and light alcohol use were associated with higher odds of overestimation. Female sex, depressive symptoms, anxiety symptoms, lower self-efficacy, higher neuroticism, lower self-rated health, chronic diseases, and more pain were associated with underestimation.

Results in the models using five groups (see Supplemental Material) showed that most associations reflected a dose-response relationship with concordance, with odds ratios increasing or decreasing in the same direction as in the three-group models. There were some nuances, however; less pain and normal weight were specifically associated with strong overestimation, whereas light alcohol use was predominantly observed in the slight overestimation group. In addition, there was an association between higher self-rated health and strong overestimation. Age, sex, self-efficacy, and neuroticism were mainly associated with slight underestimation. In addition, higher self-esteem was particularly associated with lower odds of strong—rather than slight—underestimation.

The forward selection procedure indicated that cognitive impairment, self-efficacy, neuroticism, partner status, self-rated health, pain, BMI, and alcohol use should be included in the final model (Table 5; Model 2). In this model, older age, male sex, cognitive impairment, less pain, and light alcohol use were associated with overestimation. Marginally significant correlates of overestimation were lower self-efficacy and lower BMI. Younger age, female sex, and lower self-rated health were associated with underestimation. Higher neuroticism and having a partner in the household were marginally significant correlates.

Analyses with the five discordance categories (see Supplemental Material) were again largely in line with the three categories, and often showed gradients in the associations between covariates and different discordance groups. However, partner status did not emerge as an important predictor in the forward selection procedure, whereas smoking did, although this association was not statistically significant in the final model.

Discussion

In this population-based study, we examined which psychosocial and health-related factors were associated with discordance of patterns of 13-year change in self-reported functional limitations and measured physical performance. In line with our first hypothesis, we found that 46% showed a higher trajectory of self-reported functional limitations relative to performance-based functioning ("overestimation") and that 13%

Table 4. Unadjusted Characteristics of Groups Based on Discordance Patterns ($n = 2,135$).

	Overestimation ($n = 987$)	Concordance ($n = 867$)	Underestimation ($n = 281$)	
	M (*SE*) / %	*M* (*SE*) / %	*M* (*SE*) / %	*p*
Demographics				
Age	74.1 (0.3)	70.8 (0.3)	69.1 (0.5)	<.001
Sex (% female)	46.9	54.9	65.1	<.001
Education	9.0 (0.1)	9.2 (0.1)	9.1 (0.2)	.44
Physical functioning				
Functional limitations	28.0 (0.1)	26.9 (0.2)	27.0 (0.2)	<.001
Physical performance	2.3 (0.03)	2.8 (0.03)	3.2 (0.04)	<.001
Psychological factors				
Depressive symptoms	7.6 (0.2)	7.5 (0.3)	9.0 (0.5)	.01
Anxiety symptoms	2.5 (0.1)	2.6 (0.1)	3.5 (0.2)	<.001
MMSE score ≤23 (%)	14.0	7.4	5.7	<.001
Mastery	17.3 (0.1)	17.6 (0.1)	17.3 (0.2)	.08
Self-esteem	15.3 (0.1)	15.4 (0.1)	15.2 (0.1)	.24
Self-efficacy	41.6 (0.2)	42.5 (0.2)	41.7 (0.3)	.001
Neuroticism	5.9 (0.2)	5.5 (0.2)	7.0 (0.4)	.001
Social factors				
Instrumental support	14.7 (0.2)	14.9 (0.2)	14.9 (0.4)	.70
Partner in house (%)	57.5	61.7	66.9	.01
Health-related factors				
Self-rated health	3.7 (0.03)	3.6 (0.03)	3.4 (0.06)	.001
Chronic diseases	1.6 (0.04)	1.6 (0.04)	1.8 (0.08)	.06
Pain	0.7 (0.04)	0.8 (0.05)	1.0 (0.09)	<.001
Physical activity	5.5 (0.1)	5.9 (0.1)	6.2 (0.2)	.01
BMI				
<25	36.4	28.4	28.6	
25-30	44.4	45.9	44.3	
≥30	19.2	25.6	27.2	.01
Smoking				
No	62.7	62.6	65.1	
Former	15.2	16.2	15.3	
Current	22.0	21.2	19.7	.56
Alcohol				
Does not drink	22.0	24.7	22.3	
Light	52.7	46.2	46.7	
Moderate	18.6	21.6	21.6	
(Very) Excessive	6.8	7.5	9.4	.09

Note. MMSE = Mini-Mental State Examination; BMI = body mass index.

showed a lower trajectory of self-reported functional limitations relative to performance-based functioning ("underestimation"). Because the number, initial levels, and shapes of the trajectories identified with the latent class analysis were similar for self-reported functional limitations and performance-based functioning, the magnitude of discordance within the concordant, overestimation, and underestimation groups was relatively stable over time. Older age, male sex, cognitive impairment, experiencing less pain, and light alcohol use were independent factors associated with overestimation, while younger age, female sex, and lower self-rated health were associated with underestimation. In addition, we found evidence that low or normal BMI (as opposed to overweight or obesity) was associated with overestimation, that having a partner in the household was associated with underestimation, and that lower self-efficacy and higher neuroticism were associated with both overestimation and underestimation, although these associations were marginally significant in fully adjusted models.

Overestimation

Although the magnitude and development over time tended to depend on the general level of functioning (see Supplemental

Table 5. Adjusted Associations Between Risk Factors and Discordance ($n = 2,135$).

	Model 1[a]		Model 2[b]	
	Overestimation	Underestimation	Overestimation	Underestimation
Ref = concordance	OR (CI)	OR (CI)	OR (CI)	OR (CI)
Demographic factors				
Age	1.59 [1.29, 1.97]*	0.97 [0.96, 0.99]*	1.66 [1.33, 2.07]*	0.98 [0.96, 1.00]*
Age-squared	0.997 [0.996, 0.999]*		0.997 [0.995, 0.998]*	
Sex (female)	0.71 [0.59, 0.86]*	1.55 [1.17, 2.05]*	0.71 [0.57, 0.89]*	1.68 [1.21, 2.34]*
Education	0.99 [0.96, 1.01]	1.00 [0.96, 1.04]		
Psychological factors				
Depressive symptoms	1.00 [0.99, 1.01]	1.02 [1.01, 1.04]*		
Anxiety symptoms	0.99 [0.96, 1.02]	1.07 [1.03, 1.11]*		
MMSE score ≤23	1.61 [1.16, 2.24]*	0.91 [0.51, 1.64]	1.63 [1.16, 2.28]*	0.89 [0.49, 1.61]
Mastery	0.99 [0.96, 1.02]	0.96 [0.93, 1.01]†		
Self-esteem	0.99 [0.94, 1.03]	0.96 [0.90, 1.02]		
Self-efficacy	0.98 [0.96, 0.99]*	0.97 [0.95, 1.00]*	0.98 [0.96, 1.00]†	0.99 [0.96, 1.02]
Neuroticism	1.01 [1.00, 1.03]	1.04 [1.02, 1.07]*	1.01 [0.99, 1.03]	1.03 [1.00, 1.06]†
Social factors				
Instrumental support	0.99 [0.98, 1.01]	1.00 [0.98, 1.02]		
Partner in house (%)	0.95 [0.77, 1.18]	1.28 [0.93, 1.75]	0.95 [0.76, 1.18]	1.34 [0.97, 1.85]†
Health-related factors				
Self-rated health	1.08 [0.97, 1.20]	0.78 [0.68, 0.90]*	1.04 [0.93, 1.18]	0.83 [0.71, 0.99]*
Chronic diseases	0.97 [0.90, 1.05]	1.16 [1.05, 1.29]*		
Pain	0.89 [0.82, 0.97]*	1.13 [1.03, 1.25]*	0.89 [0.81, 0.98]*	1.06 [0.94, 1.19]
Physical activity	1.01 [0.98, 1.04]	1.00 [0.96, 1.03]		
BMI				
25-30	0.77 [0.60, 1.00]†	0.94 [0.63, 1.39]	0.79 [0.61, 1.03]†	0.95 [0.64, 1.42]
≥30	0.66 [0.46, 0.95]*	0.95 [0.60, 1.49]	0.71 [0.49, 1.02]†	0.92 [0.57, 1.47]
Smoking				
Former	0.95 [0.67, 1.34]	0.98 [0.57, 1.70]		
Current	1.10 [0.81, 1.50]	0.92 [0.57, 1.48]		
Alcohol				
Light	1.36 [1.03, 1.80]*	1.13 [0.71, 1.79]	1.37 [1.02, 1.83]*	1.24 [0.78, 1.97]
Moderate	0.97 [0.67, 1.40]	1.25 [0.70, 2.21]	0.98 [0.68, 1.43]	1.41 [0.78, 2.53]
Excessive	1.15 [0.68, 1.96]	1.55 [0.71, 3.39]	1.16 [0.67, 2.00]	1.83 [0.83, 4.05]

Note. OR = odd ratio; CI = 95% confidence interval; MMSE = Mini-Mental State Examination; BMI = body mass index.
[a]Single risk factors adjusted for age, age-squared and sex.
[b]All risk factors from a forward selection procedure.
†$p < .10$. *$p < .05$.

Material), we found that about 46% of the older adults in our sample showed persistent overestimation and that older age was one of the strongest predictors of overestimation. This finding is in line with the NHATS conceptual framework (Freedman, 2009). It suggests that as people age, their experienced daily functioning tends to decline more slowly than does physical capacity, and only at very old ages this tendency may decline—an expectation that was supported by supplementary analyses of the dynamics in overestimation. Hence, given the conceptual distinction between capacity and functional limitations and their typical course in the disability process, finding a large "overestimation" group is not surprising; overestimation may at least partly reflect normal aging rather than a pathological process. This type of discordance may also reflect compensatory strategies or adaptation (Glass, 1998). For example, older adults might consider difficulties with daily tasks as a part of normal aging and therefore report no difficulties even if they use material or personal aids when executing specific tasks (Feuering et al., 2014).

However, in line with Hypothesis 2, several psychosocial and health-related factors were associated with overestimation, suggesting that there are factors beyond chronological age that are systematically related to overestimation. Particularly, our finding that the absence of pain and self-rated health was associated with (strong) overestimation suggests that as long as a loss of capacity does not translate into

pain or a conscious experience of bad health, older adults are less likely to experience limitations in daily functioning even when capacity declines. We also found the opposite pattern, that is, that more pain was associated with underestimation, but in our study this association was weak.

Previous findings on cognitive impairment are contradictory, demonstrating associations with underestimation (Shulman et al., 2006) as well as overestimation (Daltroy et al., 1999). At least for the general older population, our study strengthened the evidence for a link between general cognitive impairment and overestimation. Previous studies suggest that this association may be due to defects in executive functioning, which includes self-monitoring and working memory (Cahn-Weiner, Boyle, & Malloy, 2002; Okonkwo et al., 2009), which may affect the judgment of one's own physical capabilities. The specific subgroup with cognitive impairment might be particularly at risk of injuries or falls (Liu-Ambrose, Ahamed, Graf, Feldman, & Robinovitch, 2008; Sakurai et al., 2013).

Finally, while we would have expected the highest rates of overestimation in the excessive drinkers, only the light drinkers showed increased odds of overestimation compared with abstainers. In an explorative analysis, we found that nondrinkers were the oldest and least healthy, the moderate and excessive drinkers were the youngest and healthiest, and the light drinkers fell in-between. The intermediate level of health in the light drinkers may suggest that they already had some physical decline, but possibly not severe enough to translate into limitations in daily functioning, which might explain their overestimation of functioning. However, as surprisingly few studies on discordance have included alcohol use, we cannot externally verify this possibility.

Underestimation

Similar to overestimation, age and sex were important correlates of underestimation, but in the opposite direction. In addition, lower self-rated health was associated with underestimation. However, while only these three factors had predictive value independent from other included factors, many other factors were associated with underestimation in partially adjusted models; chronic disease, pain, depressive and anxiety symptoms, and lower self-efficacy. As such factors may contribute to more negative evaluations of one's own health (Daltroy et al., 1999; Shulman et al., 2006), it might be that mood states and factors reflecting self-regulation indicate specific mechanisms behind the associations of age, sex, and self-rated health with underestimation of daily functioning (Kempen et al., 1999; Wittink et al., 2003).

One surprising finding was that although partner status was not associated with discordance in partially adjusted models, it did emerge as a predictor of underestimation from the forward selection procedure in six out of 20 imputed datasets. Theoretically, it seems plausible that those with a partner in the household may be more inclined to receive help with daily functioning, even when they are physically able to carry out particular tasks. However, partner status was not identified as a correlate of underestimation in the sensitivity analysis with five discordance groups. Therefore, in accordance with previous studies (Wloch, Kuh, & Cooper, 2015), we consider our evidence for the role of partner status in underestimation to be weak.

Sex Differences

Our finding that women are more likely to underestimate their daily functioning relative to their capacity, while men tend to overestimate it, is in line with previous research (e.g., Merrill, Seeman, Kasl, & Berkman, 1997; Sainio et al., 2006). Explaining these differences in depth is beyond the scope of this article, but some hypotheses are worthwhile discussing. For example, it has been suggested that women are more sensitive to bodily symptoms and more inclined to acknowledge discomfort or limitations than men (Merrill et al., 1997). In addition, it is suggested that the lower average level of performance-based physical functioning in older women compared with men may by itself alter internal standards of functioning in women and influence the perception of functional limitations, such that men are more optimistic and women more pessimistic about their physical capabilities (Daltroy et al., 1999; Merrill et al., 1997). Discordance between self-reported functional limitations and performance-based functioning may provide a fruitful context to further explore sex differences in aging in future studies.

Contributions of Measuring Discordance Longitudinally

The majority of previous studies into discordance have been cross-sectional, hampering insight into the extent to which discordance is persistent across the aging process. Our results based on latent class analysis suggest that patterns of discordance and concordance are persistent over time. However, a visual comparison of the trajectories for each combination of performance-based and self-reported measure of functioning suggests that underlying this relative stability, overestimation and underestimation tend to increase and then decrease with aging.

Furthermore, cross-sectional studies might overestimate associations between discordance and factors that are likely to be dynamic, such as mood. Indeed, we found that stable or slowly changing characteristics such as age, sex, cognitive impairment, and self-rated health were the strongest correlates of long-term patterns of discordance. In addition, relatively stable psychological characteristics such as self-efficacy and neuroticism were more robust correlates than depressive and anxiety symptoms. Interestingly, within the NHATS framework, mood symptoms are conceptualized as "health conditions," located up front the disablement process (Kasper & Freedman, 2014, pp. S3-S4). However,

rather than acting as predisposing factors, such symptoms may be interacting with different stages of the process. It thus appears crucial that future studies on discordance are based on repeated measures of physical functioning as well as mood symptoms.

Implications

Discrepancies between physical capacity and self-reported functional limitations appear persistent over time, and easy to assess factors such as age, sex, cognitive impairment, pain, and self-rated health appear to be good starting points for estimating the accuracy of older adults' assessment of daily functioning. Admittedly, these characteristics are hard to change and can only serve as indications of who might be at risk. Therefore, psychological factors may still be important as mechanisms behind systematic discrepancies between experienced functional limitations and performance-based physical functioning, and they are more amenable to change. Although more evidence is needed on the consequences of over and underestimation of physical functioning for (clinical) outcomes, knowing of these discrepancies might ultimately stimulate interventions to reduce risk of future injuries or improve utilization of one's physical capabilities. Theoretically, we have highlighted several fixed ("predisposing") and dynamic factors that partly explain individual differences in the disablement process.

Strengths and Limitations

Strengths of this study are the 13-year longitudinal data from a large population-based sample of older adults and inclusion of a broad range of psychosocial and health-related factors. This enabled us to show to what extent discordance is persistent over time or changes across the aging process. Furthermore, whereas most previous studies have measured discordance as a continuous outcome, assuming that predictors affect both underestimation and overestimation, we showed that their predictors partly differ. Finally, we used an extensive set of physical performance tests measuring upper and lower extremity function, which aligned well with the diversity of bodily functions needed to execute more complex tasks in daily life.

One limitation of our study is that even with our relatively large sample, statistical power tended to become low, particularly in the underestimation group. This necessitated reduction of our sample to three groups. Nevertheless, sensitivity analyses using a five group classification showed that the main results were robust and largely a matter of degree, while also providing some relevant nuance to the main analysis. Furthermore, our categorization into discordance groups was based on group-level estimates from the latent class analysis, which might have resulted in observing more stability in discordance than if we had used individually calculated discordance scores. Second, there was substantial classification uncertainty in the latent class models, although sensitivity analyses using continuous discordance scores showed reasonably accurate distinction between subgroups. Finally, because we focused on baseline correlates of trajectories of functioning, an important area for a follow-up study is to provide insight into the interactions of dynamic correlates with discordance over time. These limitations should be attended to in future studies based on repeated measures of individually calculated discordance scores, linked to time-varying covariates.

Conclusion

Factors associated with overestimation of daily functioning relative to physical capacity partly differ from those associated with underestimation. Easy to assess characteristics such as chronological age, sex, cognitive impairment, pain, and self-rated health appear to be good indicators of persistent discordance across 13 years of time. In addition, lower self-regulation and increased affective symptoms may contribute to discordance. This set of stable and modifiable factors provides more understanding of how experiences of physical capability vary according to individual characteristics and across the aging process, and identifies groups that may be at risk of adverse consequences from overestimating their physical capabilities, or, vice versa, do not utilize these capabilities to their full potential.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The Longitudinal Aging Study Amsterdam is supported by a grant from the Netherlands Ministry of Health Welfare and Sports, Directorate of Long-Term Care.

ORCID iD

Almar A. L. Kok https://orcid.org/0000-0001-6398-2484

Data Accessibility Statement

More details on the research materials related to this article are available upon request, by email to the first author. Furthermore, access to data from the Longitudinal Aging Study Amsterdam can be requested by submitting a LASA analysis proposal form for evaluation. The LASA evaluation committee provides access to the data on the condition that the goals of the data request are in keeping with the overarching aims of LASA that its participants have provided consent for. The LASA analysis proposal template includes the option to request data for replication purposes. The template of the analysis proposal form can be obtained at www.

lasa-vu.nl, or by sending a request to the LASA secretariat, f.kursun@amsterdamumc.nl. Analysis proposals can be submitted to the LASA secretariat.

Supplemental Material

Results from additional analyses are available in the online supplemental material.

References

Ainsworth, B. E., Haskell, W. L., Herrmann, S. D., Meckes, N., Bassett, D. R., Tudor-Locke, C., . . . Leon, A. S. (2011). 2011 compendium of physical activities: A second update of codes and MET values. *Medicine and Science in Sports and Exercise*, *43*, 1575-1581. doi:10.1249/MSS.0b013e31821ece12

Barelds, D., & Luteijn, F. (2002). Measuring personality: A comparison of three personality questionnaires in the Netherlands. *Personality and Individual Differences*, *33*, 499-510.

Bowling, A. (2006). Lay perceptions of successful ageing: Findings from a national survey of middle aged and older adults in Britain. *European Journal of Ageing*, *3*, 123-136. doi:10.1007/s10433-006-0032-2

Brach, J. S., VanSwearingen, J. M., Newman, A. B., & Kriska, A. M. (2002). Identifying early decline of physical function in community-dwelling older women: Performance-based and self-report measures. *Physical Therapy*, *82*, 320-328. doi:10.1093/ptj/82.4.320

Cahn-Weiner, D. A., Boyle, P. A., & Malloy, P. F. (2002). Tests of executive function predict instrumental activities of daily living in community-dwelling older individuals. *Applied Neuropschology*, *9*, 187-191. doi:10.1207/S15324826AN0903

Cress, M. E., Schechtman, K. B., Mulrow, C. D., Fiatarone, M. A., Gerety, M. B., & Buchner, D. M. (1995). Relationship between physical performance and self-perceived physical function. *Journal of the American Geriatric Society*. doi:10.1111/j.1532-5415.1995.tb06372.x

Daltroy, L. H., Larson, M. G., Eaton, H. M., Phillips, C. B., & Liang, M. H. (1999). Discrepancies between self-reported and observed physical function in the elderly: The influence of response shift and other factors. *Social Science and Medicine*, *48*, 1549-1561. doi:10.1016/S0277-9536(99)00048-9

Delbaere, K., Close, J. C. T., Brodaty, H., Perminder, S., & Lord, S. R. (2010). Determinants of disparities between perceived and physiological risk of falling among elderly people : Cohort study. *BMJ*, *341*, c4165. doi:10.1136/bmj.c4165

Feuering, R., Vered, E., Kushnir, T., Jette, A. M., & Melzer, I. (2014). Differences between self-reported and observed physical functioning in independent older adults. *Disability and Rehabilitation*, *36*, 1395-1401. doi:10.3109/09638288.2013.828786

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state": A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12*, 189-198. doi:10.1016/0022-3956(75)90026-6

Freedman, V. A. (2009). Adopting the ICF language for studying late-life disability : A field of dreams ? *64A*, 1172-1174. doi:10.1093/gerona/glp095

Garretsen, H., & Knibbe, R. (1983). *Alkohol prevalentie onderzoek Rotterdam/Limburg; Landelijk eindrapport* [Alcohol prevalence study Rotterdam/Limburg; National end report]. Leidschendam, The Netherlands: Ministry of Welfare, Public Health and Culture.

Glass, T. A. (1998). Conjugating the "tenses" of function: Discordance among hypothetical, experimental, and enacted function in older adults. *The Gerontologist*, *38*, 101-112. doi:10.1093/geront/38.1.101

Goverover, Y., Kalmar, J., Gaudino-Goering, E., Shawaryn, M., Moore, N. B., Halper, J., & DeLuca, J. (2005). The relation between subjective and objective measures of everyday life activities in persons with multiple sclerosis. *Archives of Physical Medicine and Rehabilitation*, *86*, 2303-2308. doi:10.1016/j.apmr.2005.05.016

Guralnik, J., Fried, L., & Salive, M. (1996). Disability as a public health outcome in the aging population. *Annual Review of Public Health*, *17*, 25-46. doi:10.1146/annurev.pu.17.050196.000325

Hoogendijk, E. O., Deeg, D. J. H., Poppelaars, J., van der Horst, M., Broese van Groenou, M. I., Comijs, H. C., . . . Huisman, M. (2016). The Longitudinal Aging Study Amsterdam: Cohort update 2016 and major findings. *European Journal of Epidemiology*, *31*, 927-945. doi:10.1007/s10654-016-0192-0

Kasper, J. D., Chan, K. S., & Freedman, V. A. (2016). Measuring physical capacity: An assessment of a composite measure using self-report and items. *Journal of Aging and Health*, *29*, 289-309. doi:10.1177/0898264316635566

Kasper, J. D., & Freedman, V. A. (2014). Findings from the 1st round of the National Health and Aging Trends Study (NHATS): Introduction to a special issue. *The Journals of Gerontology: Series B, Psychological Sciences and Social Sciences*, *69*, S1-S7. doi:10.1093/geronb/gbu125

Kasper, J. D., Edwards, B., Freedman, V. A., Seplaki, C. L., Weiss, C., Carlson, M., Bruce, T., Guralnik, J. M., Plassman, B. L., Wallace, R., Niefeld, M., & Varma, V. (2009). Advances in survey assessment of disability in older adults: Measuring physical and cognitive capacity in the National Health and Aging Trends Study. *Proceedings of the 10th Conference on Health Survey Research Methods* (pp. 11-18). Baltimore: Johns Hopkins University School of Public Health.

Kempen, G. I. J. M., Sullivan, M., van Sonderen, E., & Ormel, J. (1999). Performance-based and self-reported physical functioning in low-functioning older persons: Congruence of change and the impact of depressive symptoms. *The Journals of Gerontology: Series B, Psychological Sciences and Social Sciences*, *54*, 380-386. doi:10.1093/geronb/54B.6.P380

Kempen, G. I. J. M., Van Heuvelen, M., Van Den Brink, R., Kooijman, A., Klein, M., Houx, P., & Ormel, J. (1996). Factors affecting contrasting results between self-reported and performance-based levels of physical limitations. *Age and Ageing*, *25*, 458-464. doi:10.1093/ageing/25.6.458

Koster, A., Bosma, H., Broese van Groenou, M. I., Kempen, G. I. J. M., Penninx, B. W., van Eijk, J. T., & Deeg, D. J. (2006). Explanations of socioeconomic differences in changes in physical function in older adults: Results from the Longitudinal Aging Study Amsterdam. *BMC Public Health*, *6*, Article 244. doi:10.1186/1471-2458-6-244

Liu-Ambrose, T., Ahamed, Y., Graf, P., Feldman, F., & Robinovitch, S. N. (2008). Older fallers with poor working memory overestimate their postural limits. *Archives of Physical*

Medicine and Rehabilitation, *89*, 1335-1340. doi:10.1016/j.apmr.2007.11.052

Merrill, S. S., Seeman, T. E., Kasl, S. V., & Berkman, L. F. (1997). Gender differences in the comparison of self-reported disability and performance measures. *The Journals of Gerontology: Series A, Biological Sciences and Medical Sciences*, *52*, 19-26. doi:10.1093/gerona/52A.1.M19

Muthén, B. O. (2003). Statistical and substantive checking in growth mixture modeling: Comment on Bauer and Curran. *Psychological Methods*, *8*, 369-377. doi:10.1037/1082-989X.8.3.369

Muthén, L. K., & Muthén, B. O. (2012). M*plus* 7.0 user's guide. Retrieved from http://www.statmodel.com/download/usersguide/Mplus%20user%20guide%20Ver_7_r3_web.pdf

Nagin, D. S. (1999). Analyzing developmental trajectories: A semiparametric, group-based approach. *Psychological Methods*, *4*, 139-157. doi:10.1037/1082-989X.4.2.139

Okonkwo, O., Griffith, H., Vance, D., Marson, D., Ball, K., & Wadley, V. (2009). Awareness of functional difficulties in mild cognitive impairment: A multi-domain assessment approach. *Journal of the American Geriatrics Society*, *57*, 978-984. doi:10.1111/j.1532-5415.2009.02261.x

Pearlin, L. I., & Schooler, C. (1978). The structure of coping. *Journal of Health and Social Behavior*, *19*, 2-21. doi:10.2307/2136319

Radloff, L. S. (1977). The CES-D Scale: A Self-Report Depression Scale for research in the general population. *Applied Psychological Measurement*, *1*, 385-401. doi:10.1177/014662167700100306

Rosenberg, M., Schooler, C., Schoenbach, C., & Rosenberg, F. (1995). Global self-esteem and specific self-esteem: Different concepts, different outcomes. *American Sociological Review*, *60*, 141-156.

Sainio, P., Koskinen, S., Heliövaara, M., Martelin, T., Härkänen, T., Aromaa, A., & Miilunpalo, S. (2006). Self-reported and test-based mobility limitations in a representative sample of Finns aged 30+. *Scandinavian Journal of Public Health*, *34*, 378-386. doi:10.1080/14034940500489859

Sakurai, R., Fujiwara, Y., Ishihara, M., Higuchi, T., Uchida, H., & Imanaka, K. (2013). Age-related self-overestimation of step-over ability in healthy older adults and its relationship to fall risk. *BMC Geriatrics*, *13*, 15-17. doi:10.1186/1471-2318-13-44

Sherer, M., & Maddux, J. (1982). The self-efficacy scale-construction and validation. *Psychological Reports*, *51*, 663-671. doi:10.2466/pr0.1982.51.2.663

Shulman, L. M., Pretzer-Aboff, I., Anderson, K. E., Stevenson, R., Vaughan, C. G., Gruber-Baldini, A. L., . . . Weiner, W. J. (2006). Subjective report versus objective measurement of activities of daily living in Parkinson's disease. *Movement Disorders*, *21*, 794-799. doi:10.1002/mds.20803

Stuck, A. E., Walthert, J. M., Nikolaus, T., Büla, C. J., Hohmann, C., & Beck, J. C. (1999). Risk factors for functional status decline in community-living elderly people: A systematic literature review. *Social Science & Medicine*, *48*, 445-469. doi:10.1016/S0277-9536(98)00370-0

Tombaugh, T. N., & McIntyre, N. J. (1992). The mini-mental state examination: A comprehensive review. *Journal of the American Geriatrics Society*, *40*, 922-935. doi:10.1111/j.1532-5415.1992.tb01992.x

Van Tilburg, T. (1998). Losing and gaining in old age: Changes in personal network size and social support in a Four-Year Longitudinal Study. *The Journal of Gerontology: Series B, Psychological Sciences and Social Sciences*, *53*, 313-323. doi:10.1093/geronb/53B.6.S313

Verbrugge, L. M., & Jette, A. M. (1994). The disablement process. *Social Science & Medicine*, *38*, 1-14. doi:10.1016/0277-9536(94)90294-1

Wittink, H., Rogers, W., Sukiennik, A., & Carr, D. B. (2003). Physical functioning: Self-report and performance measures are related but distinct. *Spine*, *28*, 2407-2413. doi:10.1097/01.BRS.0000085304.01483.17

Wloch, E., Kuh, D., & Cooper, R. (2015). Which factors across life are associated with discordance between self-reported and performance measures of physical capability? *Journal of Epidemiology & Community Health*, *69*(Suppl. 1), A21. doi:10.1136/jech-2015-206256.29

Zigmond, A. S., & Snaith, R. P. (1983). The Hospital Anxiety and Depression Scale. *Acta Psychiatrica Scandinavica*, *67*, 361-370. doi:10.1111/j.1600-0447.1983.tb09716.x