Physical Predictors of Cognitive Performance in Healthy Older Adults: A Cross-Sectional Analysis

Christiaan G. Blankevoort1*, Erik J. A. Scherder1,2, Martijn B. Wieling3, Tibor Hortobágyi1, Wiebo H. Brouwer4, Reint H. Geuze5, Marieke J. G. van Heuvelen1

1 Center for Human Movement Sciences, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands, 2 Department of Neuropsychology, VU University, Amsterdam, The Netherlands, 3 Department of Quantitative Linguistics, University of Tübingen, Tübingen, Germany, 4 Department of Neurology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands, 5 Department of Clinical and Developmental Neuropsychology, University of Groningen, Groningen, The Netherlands

Abstract

There is ample evidence that physical and cognitive performance are related, but the results of studies investigating this relationship show great variability. Both physical performance and cognitive performance are constructs consisting of several subdomains, but it is presently unknown if the relationship between physical and cognitive performance depends on subdomain of either construct and whether gender and age moderate this relationship. The aim of this study is to identify the strongest physical predictors of cognitive performance, to determine the specificity of these predictors for various cognitive subdomains, and to examine gender and age as potential moderators of the relationship between physical and cognitive performance in a sample of community-dwelling older adults. In total, 98 men and 122 women (average age 74.0±5.6 years) were subjected to a series of performance-based physical fitness and neuropsychological tests. Muscle strength, balance, functional reach, and walking ability (combined score of walking speed and endurance) were considered to predict cognitive performance across several domains (i.e. memory, verbal attention, visual attention, set-shifting, visuo-motor attention, inhibition and intelligence). Results showed that muscle strength was a significant predictor of cognitive performance for men and women. Walking ability and balance were significant predictors of cognitive performance for men, whereas only walking ability was significant for women. We did not find a moderating effect of age, nor did we find support for a differential effect of the physical predictors across different cognitive subdomains. In summary, our results showed a significant relationship between cognitive and physical performance, with a moderating effect of gender.

Introduction

Demographic data suggest that the number of older adults will increase at an accelerating rate in the coming decades [1]. Age is a risk factor for different aspects of physical performance, such as muscle strength, endurance, and balance [2,3] and also for impairment in cognition, including episodic memory and executive function (e.g., inhibition, planning, and set-shifting [4]). Although epidemiological studies show a positive relationship between physical performance and cognition [5–9], a number of questions remain open [10]. We specifically address three issues that might affect the physical performance-cognition association in older adults: (1) the selection of cognitive domains, (2) gender, and (3) age.

As for test and domain selection, prior studies used a wide variety of methods and tests to quantify the association between physical and cognitive performance. Most prominent is the divergent selection of physical performance domains (e.g., mobility, balance, strength, or endurance) and cognitive performance domains (e.g., memory, global cognitive performance, fluency, attention, or executive functions) across studies to represent physical and cognitive performance [7–10]. While the use of many tests and domains is a logical consequence of the desire to assess multiple facets of physical and cognitive performance, this approach also increases the heterogeneity in predicting cognition from motor performance across studies. In addition to the issue of test and domain selection, there are also differences between studies in ethnicity, age, and the number of comorbidities. Due to variations between studies so far, multiple studies need to be taken into account to provide a coherent overview. Unfortunately, the differences between studies make it difficult to determine which physical performance tests are the strongest predictors of (individual measures of) cognitive performance. Indeed, the association between physical and cognitive performance varies widely between studies [10]. Furthermore, it is not well known if gender differences affect the association between measures of physical and cognitive performance. Imaging and neuroanatomical data provide a conceptual basis to expect a gender effect in the association between physical and cognitive performance in community-dwelling older adults [11,12]. The male brain is larger than the female brain, even after
controlling for height, but the decline in volume is also steeper for men than for women [12,13]. In addition, gender differences in cerebral blood flow after a cognitive task have been observed [14]. More specifically, gender differences are well documented in terms of maximal voluntary leg strength [13,16] and grip strength [16,17] in healthy older adults, with women also exhibiting greater reductions in motor coordination [16] than men. With respect to cognition, there is some support for a better overall cognitive performance in aging women versus men [18], especially in memory tasks [19]. Based on these findings it is conceivable that gender differences may influence the association between physical and cognitive performance.

Finally, it is unclear what the effect of age is on the relationship between physical and cognitive performance in healthy older adults. While there is a parallel increase in the variability of physical and cognitive performance with age, the rate of decline differs between the two domains: cognitive impairment accelerates after the age of 60 [4], but the decline in balance and muscle strength accelerates only markedly after the age of 75 [20,21]. Such a temporal dissociation can confound the associations between physical and cognitive performance in healthy older adults.

The goal of the present study was to re-examine the association between physical and cognitive performance in community-dwelling older men and women. In an effort to better understand the relationship between physical and cognitive performance, we examined this association using a wide array of important physical and cognitive domains which are known to be vulnerable for age-related decline. Concretely, for the physical domain we included measures of gait speed, endurance, grip strength, quadriceps strength, and balance. Each of these measures of physical performance are well-documented in terms of age-related decline [22–24] due to, for example, sarcopenia, and show a positive relationship with cognitive performance [5,7,9]. For the cognitive domain, we included tests assessing global cognitive performance, memory, processing speed and various aspects of executive function. A proper functioning of these cognitive domains is important for our functioning. Moreover, memory and executive function are well-documented in terms of age-related decline [25,26], and have a reported positive relationship with physical performance [5,9]. We addressed the following questions in this study: (1) Which physical performance measures are the strongest predictors of cognitive performance? (2) Do different physical performance tests predict different aspects of cognitive performance? (3) Do gender and age moderate the association between physical and cognitive performance?

Methods

Ethics statement

The local medical ethical committee of the university medical center of Groningen, the Netherlands, approved the study and all participants provided a signed informed consent prior to the assessments. The study was conducted in accordance with the Declaration of Helsinki (59th Amendment).

Subjects and Design

In this study, 220 older community-dwelling adults, with a mean age of 74 years (SD = 5.6; range 65–92) participated. The participants were drawn from the baseline measurement of the Groningen Intervention Study for Successful Aging [27], an intervention study with participants of 65 years and older, which in turn recruited its participants from a longitudinal cohort study [28]. A flow chart illustrating the participant selection procedure is presented in Figure 1. Inclusion criteria for the Groningen Intervention Study for Successful Aging and therefore our study were: (1) being older than 65, (2) having no cognitive decline, as indicated by a score of 24 or lower on the MMSE [6,29], (3) not exceeding the physical activity guidelines set by the American College of Sports and Medicine for healthy older adults, i.e. five times a week, 30 minutes of moderate intensity physical activity [27], and (4) having no medical condition preventing participation in a physical intervention study (e.g., severe heart problems). All 220 participants included in our study performed the neuropsychological and physical performance tests. The 33 participants who were excluded (see Figure 1) had withdrawn prior to the pretest or had only performed the neuropsychological or physical performance tests. Table 1 shows the characteristics of the participants. Women had a significantly lower level of education and income than men. Both men and women reported low numbers of chronic medical conditions, but women reported significant more chronic medical conditions than men. Women suffered significantly more from high blood pressure, rheumatoid arthritis, and neurologic diseases. Women also reported significantly more use of a walking aid.

Potential confounding variables

Various sociodemographic factors which may influence the risk for a decline in cognitive or physical performance (such as education and income) were measured [30]. The level of education was assessed on a seven-point scale suitable for the Dutch education system. The scores range from 1, less than primary school to 7, a university master’s degree. Income (after tax) was classified as below average (1), average (2), or above average (3) of the Dutch population according to Statistics Netherlands, an independent government-funded organization [31]. Other possible confounding factors included were the level of anxiety and depression, the amount of time people spent on physical activity in their spare time, and the number of comorbidities. These factors were measured using the scores on the Dutch version of the Hospital Anxiety and Depression Scale for Anxiety and Depression [32], the score on the Minnesota Leisure Time Physical Activity questionnaire (MLTPA) [33], and the number of comorbidities based on the international classification of diseases (these were summed).

Figure 1. Flowchart specifying the participant selection.
doi:10.1371/journal.pone.0070799.g001

Groningen Longitudinal Aging Study (GLAS)
N = 5279

220 participants successfully enrolled in the intervention program

All 220 participants that joined GISSA were used in the present study (N = 220)

1274 participants were invited for the Groningen Intervention Study of Successful Aging (GISSA)

253 participants agreed to participate in GISSA

33 participants were excluded
- due to disease (21)
- Not interested anymore (9)
- Not completed the pretest (1)

220 participants successfully enrolled in the intervention program

Potential confounding variables

Various sociodemographic factors which may influence the risk for a decline in cognitive or physical performance (such as education and income) were measured [30]. The level of education was assessed on a seven-point scale suitable for the Dutch education system. The scores range from 1, less than primary school to 7, a university master’s degree. Income (after tax) was classified as below average (1), average (2), or above average (3) of the Dutch population according to Statistics Netherlands, an independent government-funded organization [31]. Other possible confounding factors included were the level of anxiety and depression, the amount of time people spent on physical activity in their spare time, and the number of comorbidities. These factors were measured using the scores on the Dutch version of the Hospital Anxiety and Depression Scale for Anxiety and Depression [32], the score on the Minnesota Leisure Time Physical Activity questionnaire (MLTPA) [33], and the number of comorbidities based on the international classification of diseases (these were summed).

Figure 1. Flowchart specifying the participant selection.
doi:10.1371/journal.pone.0070799.g001
### Assessment of cognitive functions

Neuropsychological tests assessed general intelligence and performance on various cognitive domains and the included tests have good reliability and validity [34–36]. For all tests, except those using time scores, higher scores indicate a better performance. Each test was administered by trained students from the Center for Human Movement Sciences or the Psychology department.

**Global cognitive performance** was assessed with the Cognitive screening test (CST) [35], an instrument to measure cognitive decline. Scores range from 0 to 20 [35].

**Verbal comprehension** was assessed with the digit symbol substitution test of the WAIS-III [37].

**Perceptual organization** was assessed with the matrix reasoning subtest of the WAIS-III [37].

**Memory** was assessed with the Dutch version of the Rey Verbal learning test [38]. This test was used for short-term and long-term memory function. A list of 15 words is presented five times. Short-term memory was assessed from direct recall (score range 0–75), long-term memory from the delayed recall after 15–20 minutes (score range 0–15), followed by a recognition test (score range 0–30).

Several executive functions were assessed: planning, inhibition, and set shifting.

**Planning** was assessed with the Zoo map, which is a subtest from the Behavioral Assessment of the Dysexecutive Syndrome (BADS) [36]. The Zoo map test required the participant to plan a course through a zoo while adhering to specific rules. Scores range from 0–16.

**Inhibition and processing speed** were assessed with the Stroop test [34,39]. First, participants had to read a list of one hundred words with the names of four different colors printed in black ink (i.e. Stroop ‘word’). Second, participants had to name the color of one hundred different squares (i.e. Stroop ‘color’) using the same four colors as on card 1. These two cards are thought to mainly measure processing speed. Finally, as a measure of inhibition, one hundred words of the same four colors were presented in different colors of ink and participants had to name the color the words were printed in (i.e. Stroop ‘word-color’). A difference score was also calculated and measures inhibition (Δ Stroop: time on Stroop ‘word-color’ minus time on Stroop ‘color’). The time to complete a card was noted in seconds and lower scores indicate a better performance.

**Visuomotor attention and set-shifting** were assessed with the Trail Making Test (TMT) [34,40]. In part A (measuring visuomotor attention) of the TMT, participants had to draw a line between encircled numbers. In part B, they had to alternate between circles with numbers and letters (1-A, 2-B) to assess set-shifting. A difference score was calculated to represent a measure of set-shifting (Δ TMT: time on TMT B minus time on TMT A). The time to complete the tasks was noted in seconds and lower scores indicate a better performance.

**Processing speed** was measured with the digit symbol substitution test (DSST) [37]. The DSST is a paper and pencil test of psychomotor performance. The test consists of a key grid of numbers (0–9) with corresponding symbols, followed by the test section. In the test section rows of numbers with empty spaces below them are provided and participants have to fill in as many corresponding symbols as possible in 120 seconds [37]. The score was equal to the number of correctly filled boxes after 120 seconds.

### Assessment of physical performance

The selected tests represent different aspects of physical and motor performance. The validity and reliability of these tests are acceptable and have been reported previously [41]. For all tests higher scores indicate a better performance.

**Grip strength** was assessed with a Jamar® hand dynamometer. Participants grasped the dynamometer with the preferred hand with the arm at the side of the body and the palm toward the thigh. Subjects were instructed to squeeze the dynamometer handle as hard as possible; the highest score (in kg) of three trials was recorded.

**Quadriceps strength** was assessed with a custom built dynamometer, the Quadriso-tester [42]. The favored leg is tested. Participants sat on a chair with knees in 90° flexion. The load cell was located in an ankle cuff that was placed above the ankle...
joint of the dominant leg and the participants were instructed to press as hard as possible for 3 seconds. The highest score of three trials was recorded.

**Balance** was assessed on an unstable platform that could tilt sideways [41]. For 30 seconds the subjects had to keep the platform in a position so that edge of the platform would not contact the floor. Ground contact was measured with pressure sensors. The time (s) in balance (i.e. the edges of the platform were not in contact with the floor) was recorded. The trial (out of three) with the longest time was selected as a measure of balance performance.

**Functional reach** measures the maximal reach when standing. A subject reaches forward with the dominant arm, having the hand in a fist, the feet maintaining a fixed base of support, while sliding a measurement cube forward over a metal bar. The maximum distance (cm) was recorded and divided by the length of the participant. The trial (out of three) with the longest reach was selected as a measure of functional reach.

**Walking speed** was assessed over a 15 meter long level surface course. Participants were instructed to walk at a self-selected pace. The average duration in seconds of two trials was used as a measure of walking speed.

**Walking endurance** was assessed on an indoor walking track. Participants had to perform as many laps as possible on a 50-m-long rectangle track. Walking speed was increased by 1 km/h every 3 min, starting at a speed of 4 km/h and ending at a speed of 7 km/h. There were beeps between the four corners of the rectangle to guide the requested pace. Participants should reach the next corner on the following beep, if they failed to reach the corner in time the test was finished. The number of completed trajectories (i.e. a side of the rectangle) was recorded [43].

### Statistical analyses

SPSS 18.01 and R 2.10.1 were used to analyze the data. Means and standard deviations were calculated for neuropsychological and physical performance scores. A Student’s *t*-test and the chi-square test were used to determine differences between men and women. Grip strength, leg strength, gait speed, the trail making test, and the Stroop test were positively skewed and were, therefore, log-transformed.

The scores of all numeric variables were standardized by converting them to *z*-scores in order to facilitate comparison. The physical and cognitive test scores were inverted when a lower score indicated a higher performance (i.e. in this way a higher score always corresponds to a higher performance). Whenever physical measures correlated close to *r* = 0.7, we combined them in a single measure (i.e. by averaging their *z*-scores) to prevent multicollinearity.

Consequently, we combined walking speed and endurance (*r* = 0.67, *p* = .000, *N* = 220) into a variable Walking ability, and grip strength and quadriceps strength (*r* = 0.75, *p* = .000, *N* = 220) into a variable Strength. In sum, four physical factors were identified: Strength, Walking ability, Balance, and Functional reach. As the scores on grip strength, leg strength, walking speed and endurance were highly gender-dependent (with men having higher scores than women), we corrected those scores for men downwards before calculating the *z*-scores (i.e. new score men = (mean score women/mean score men) * original score men).

As the cognitive domain contained more tests than the physical domain, and there was substantial overlap between the tests, we conducted an exploratory factor analysis (maximum likelihood estimation with oblique rotation) to identify variables which could be grouped. In this way, the complexity of the dataset was reduced substantially as it yielded fewer measurement points per subject. A factor loading of .32 was set as the minimum to be printed in the output of the factor analysis [44]. Each neuropsychological test was uniquely assigned to the factor where it had the highest loading. Subsequently, for each of the resulting factors (described in the following section) the corresponding cognitive score was calculated as the average of the standardized scores of the neuropsychological tests linked to the factor. Since every participant had scores on multiple different cognitive factors, we used linear mixed-effects regression modeling (LME) with participant as a random-effect factor [45] to take the structural variation linked to each participant into account (i.e. participants who scored high on one cognitive factor are more likely to score high on another cognitive factor). In the analysis, the cognitive score was used as the dependent variable. By including the type of cognitive factor in our model, we were able to assess the precise effect of the physical performance measurements for each individual cognitive factor (i.e. we assessed the possible interaction between cognitive factor and each of the physical performance predictors).

The significance of fixed-effect predictors was evaluated by means of the *t*-test for the coefficients, in addition to model comparison likelihood ratio tests and AIC (Akaike Information Criterion; [46]). When the dataset is large enough (as in our case) the *t*-distribution resembles the normal distribution and variables are significant when their absolute *t*-value is at least 1.65 (when a directional hypothesis is used, i.e. applying a one-tailed test) or at least 2 (for a two-tailed test). As there is a large amount of evidence supporting a positive association between physical performance and cognitive performance we only used a one-tailed test for assessing the significance of the physical measures.

In addition, we conducted model comparison tests to assess if each predictor or interaction significantly improved the model by comparing the log-likelihood and AIC (Akaike Information Criterion) values of the more complex model (i.e. including the additional predictor) to a baseline model (the same as the more complex model, but excluding the additional predictor). A lower AIC (or higher log-likelihood) indicates a better model [46]. On the basis of the AIC values the evidence ratio can be calculated which expresses the relative probability that the model with the lowest AIC is more likely to provide a more precise model of the data. The evidence ratio is exponentially related to the AIC difference. For example if the AIC difference is 2 (generally used as the minimum required reduction), then the model is 2.7 times more likely to provide a precise model of the data, whereas an AIC difference of 6 indicates that the model with the lowest AIC is 34.6 times more likely to provide a precise model of the data [46].

### Missing values

Inspection of our dataset revealed only a few missing values. Less than 1% percent of the data was missing with respect to the physical performance test data, whereas only 0.1% of the data was missing for all neuropsychological tests. The limited amount of missing data is ‘trivial’ [47] and we used regression substitution to replace these missing values. This method is preferred over replacing missing values by their mean or deleting the cases with missing values [48]. Whereas multiple imputation [49] is the preferred method to deal with missing values, the amount of missing data was very limited in this study and we therefore opted for the simpler method of regression substitution.
Results

Table 2 presents the descriptive data for the neuropsychological and physical performance measures. Women performed significantly better than men with respect to Stroop ‘color’, Stroop ‘word-color’, and verbal memory. Men performed significantly better with respect to the WAIS subtest information, and also across the complete physical domain: men were stronger, faster, and had greater endurance than women. Only functional reach, after correcting for body height, was not significantly different between men and women.

The factor analysis of all 15 cognitive tests revealed seven factors for the cognitive performance tests. The seven factors were: verbal attention (Stroop ‘word’), visual attention (Stroop ‘color’), visuo-motor attention (TMT-A), set-shifting (TMT-B and Δ TMT), inhibition (Stroop ‘word-color’ and Δ Stroop), memory (direct recall, delayed recall, and recognition), and intelligence (WAIS subtest information, WAIS subtest matrices, and CST). As zoo time and the digit symbol substitution test did not reach the loading threshold of .32 for any of the factors, we excluded both from the analysis. Table 3 shows the loadings.

Prediction of cognitive performance by physical performance

Table 4 shows the best mixed-effects regression model (explained variance: 38.2%). This model shows that age ($\beta = -15$, $t = -3.98$) and being male ($\beta = -23$, $t = -3.44$) have a negative impact on all cognitive factors, whereas education ($\beta = 21$, $t = 6.33$) has a positive effect on these factors. The other potentially confounding variables (i.e. income, comorbidity, depression, anxiety, walking aid and the score on the MLTPA) did not reach significance by themselves or in interaction with any other variables and were therefore not included in the model.

Physical predictors of cognitive performance significantly improving the fit of the model were walking ability ($\beta = .15$, $t = 3.57$; Table 4 shows the effect moderated by gender), balance ($\beta = .11$, $t = 3.00$; Table 4 shows the effect moderated by gender) and strength ($\beta = .07$, $t = 1.84$). Functional reach did not reach significance and was excluded from the final model.

The results show that there was no variation in the effect of the physical performance measures on the different factors of cognitive performance. The model did not improve by allowing for a varying effect of the physical performance predictors on each individual cognitive factor (i.e. no interaction reduced the AIC compared to the simpler model with at least 2).

Finally, the model shows a specific interaction between balance and gender and walking ability and gender. For balance, there appears to be no significant effect for women ($\beta = .06$, $t = 1.32$), but a clear significant effect for men ($\beta = .16$, $t = 3.17$). For walking ability, only a small significant effect could be observed for women ($\beta = .09$, $t = 1.91$), but a more pronounced effect for men ($\beta = .24$, $t = 4.20$). No other significant interactions with age or gender were found. Note that the inclusion of the variable indicating that the participant is male (1, or not: 0) improves the model and does not lead to a clear significant effect for men in any of the cognitive factors.

Table 2. Means and standard deviations (SD) for the neuropsychological and physical performance tests.

<table>
<thead>
<tr>
<th>Neuropsychological tests</th>
<th>Men</th>
<th>Women</th>
<th>p-value assessing difference between men and women</th>
<th>All participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>CST</td>
<td>19.17 (1.52)</td>
<td>19.05 (1.15)</td>
<td>.51</td>
<td>19.11 (1.33)</td>
</tr>
<tr>
<td>WAIS information</td>
<td>16.17 (5.32)</td>
<td>13.23 (4.89)</td>
<td>&lt;.001</td>
<td>14.54 (5.28)</td>
</tr>
<tr>
<td>WAIS matrix reasoning</td>
<td>12.23 (5.47)</td>
<td>12.28 (5.36)</td>
<td>.95</td>
<td>12.26 (5.40)</td>
</tr>
<tr>
<td>15 WT direct recall</td>
<td>31.47 (9.51)</td>
<td>38.63 (9.60)</td>
<td>.001</td>
<td>35.44 (10.18)</td>
</tr>
<tr>
<td>15 WT delayed recall</td>
<td>6.49 (2.70)</td>
<td>8.07 (2.55)</td>
<td>&lt;.001</td>
<td>7.37 (2.73)</td>
</tr>
<tr>
<td>15 WT recognition</td>
<td>27.15 (3.47)</td>
<td>28.53 (2.03)</td>
<td>&lt;.001</td>
<td>27.92 (2.84)</td>
</tr>
<tr>
<td>Zoo map 1</td>
<td>.06 (4.75)</td>
<td>.07 (4.83)</td>
<td>.12</td>
<td>.51 (4.81)</td>
</tr>
<tr>
<td>TMT B (s)</td>
<td>126.76 (59.13)</td>
<td>127.36 (62.67)</td>
<td>.94</td>
<td>127.09 (60.98)</td>
</tr>
<tr>
<td>Δ TMT (s)</td>
<td>73.52 (44.93)</td>
<td>74.77 (53.89)</td>
<td>.85</td>
<td>74.21 (49.99)</td>
</tr>
<tr>
<td>Stroop ‘word-color’ (s)</td>
<td>137.39 (49.72)</td>
<td>122.98 (39.31)</td>
<td>.02</td>
<td>129.40 (44.72)</td>
</tr>
<tr>
<td>Δ Stroop (s)</td>
<td>66.74 (39.37)</td>
<td>58.21 (34.00)</td>
<td>.09</td>
<td>62.01 (36.65)</td>
</tr>
<tr>
<td>TMT A (s)</td>
<td>53.23 (21.61)</td>
<td>52.59 (18.77)</td>
<td>.81</td>
<td>52.88 (20.04)</td>
</tr>
<tr>
<td>DSST (score)</td>
<td>45.43 (14.09)</td>
<td>45.70 (11.58)</td>
<td>.87</td>
<td>45.58 (12.73)</td>
</tr>
<tr>
<td>Stroop ‘word’ (s)</td>
<td>53.31 (11.10)</td>
<td>52.25 (9.90)</td>
<td>.46</td>
<td>52.72 (10.44)</td>
</tr>
<tr>
<td>Stroop ‘color’ (s)</td>
<td>71.37 (16.97)</td>
<td>64.77 (13.21)</td>
<td>.001</td>
<td>67.71 (15.32)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Physical performance tests</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Grip strength (kg)</td>
<td>39.56 (9.21)</td>
<td>23.59 (5.48)</td>
<td>&lt;.001</td>
<td>30.70 (10.84)</td>
</tr>
<tr>
<td>Quadriceps strength (kg)</td>
<td>41.17 (14.90)</td>
<td>21.19 (7.32)</td>
<td>&lt;.001</td>
<td>30.09 (15.07)</td>
</tr>
<tr>
<td>Balance (*3)</td>
<td>68.98 (9.73)</td>
<td>67.84 (9.76)</td>
<td>.39</td>
<td>68.35 (9.74)</td>
</tr>
<tr>
<td>Functional reach (cm/length in cm)</td>
<td>.21(0.04)</td>
<td>.20 (0.04)</td>
<td>.73</td>
<td>.20 (0.04)</td>
</tr>
<tr>
<td>Walking speed (s)</td>
<td>20.92 (4.57)</td>
<td>23.37 (7.79)</td>
<td>&lt;.001</td>
<td>22.28 (6.65)</td>
</tr>
<tr>
<td>Endurance</td>
<td>34.73 (16.79)</td>
<td>24.02 (13.06)</td>
<td>&lt;.001</td>
<td>28.79 (15.74)</td>
</tr>
</tbody>
</table>

*a lower score indicates better performance; CST, cognitive screening test; WAIS, Wechsler Adult Intelligence Scale; 15 WT, Fifteen word test; TMT, trail making test; ΔTMT, TMT B–TMT A; Δ Stroop, Stroop ‘word-color’–Stroop ‘color’; DSST, digit symbol substitution test.

doi:10.1371/journal.pone.0070799.t002
alter the moderating effect of gender on balance and walking ability.

To illustrate the contribution of each predictor (or interaction) to the fit of the model, Table 5 shows the increase in goodness of fit when adding each predictor to the simpler model without the predictor. Given that each variable improves the fit significantly, as can be seen by the log-likelihood ratio test and the decrease in AIC (and associated high evidence ratios), the inclusion of each of the variables reported in this section is warranted.

### Table 3. Results of the factor analysis (using oblique rotation).

<table>
<thead>
<tr>
<th></th>
<th>Factor 1</th>
<th>Factor 2</th>
<th>Factor 3</th>
<th>Factor 4</th>
<th>Factor 5</th>
<th>Factor 6</th>
<th>Factor 7</th>
</tr>
</thead>
<tbody>
<tr>
<td>TMT-B</td>
<td>0.74</td>
<td></td>
<td></td>
<td></td>
<td>0.36</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TMT Delta</td>
<td>1.13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15WT Dir recall</td>
<td></td>
<td>0.79</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15WT Del recall</td>
<td></td>
<td>0.90</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15 WT Rec</td>
<td></td>
<td>0.58</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stroop ‘word-color’</td>
<td></td>
<td>0.71</td>
<td></td>
<td></td>
<td>0.51</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stroop Delta</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.09</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WAIS information</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.05</td>
<td></td>
</tr>
<tr>
<td>WAIS matrices</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.37</td>
<td></td>
</tr>
<tr>
<td>CST</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.39</td>
<td></td>
</tr>
<tr>
<td>TMT A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.11</td>
</tr>
<tr>
<td>Stroop ‘word’</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.89</td>
</tr>
<tr>
<td>Stroop ‘color’</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.87</td>
</tr>
<tr>
<td>DSST</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SS loadings</td>
<td>1.88</td>
<td>1.78</td>
<td>1.75</td>
<td>1.53</td>
<td>1.49</td>
<td>1.25</td>
<td>0.95</td>
</tr>
<tr>
<td>Proportion Var</td>
<td>0.13</td>
<td>0.12</td>
<td>0.12</td>
<td>0.10</td>
<td>0.10</td>
<td>0.08</td>
<td>0.06</td>
</tr>
<tr>
<td>Cumulative Var</td>
<td>0.13</td>
<td>0.24</td>
<td>0.36</td>
<td>0.46</td>
<td>0.56</td>
<td>0.65</td>
<td>0.71</td>
</tr>
</tbody>
</table>

Test of the hypothesis that 7 factors are sufficient. The chi square statistic is 17.71 on 21 degrees of freedom. The p-value is 0.6.

TMT, Trail Making Test; TMT Delta, TMT B – TMT A; 15WT, fifteen word test; Dir, direct; Del, delayed; Rec, recognition; Stroop Delta, Stroop ‘word-color’ – Stroop ‘color’; WAIS, Wechsler Adult Intelligence Scale; CST, Cognitive Screenings Test; DSST, Digit Symbol Substitution, Test; BADS, Behavioural Assessment of Dysexecutive Syndrome; Var, Variance.

doi:10.1371/journal.pone.0070799.t003

### Table 4. Linear mixed effects regression model predicting cognitive performance.

<table>
<thead>
<tr>
<th>Fixed effects</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>0.09946</td>
<td>0.04431</td>
<td>2.245</td>
<td>&lt;.05</td>
</tr>
<tr>
<td>Age</td>
<td>−0.15174</td>
<td>0.03813</td>
<td>−3.980</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Education</td>
<td>0.21125</td>
<td>0.03336</td>
<td>6.333</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Male</td>
<td>−0.23001</td>
<td>0.06679</td>
<td>−3.444</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Strength*</td>
<td>0.07180</td>
<td>0.03897</td>
<td>1.842</td>
<td>&lt;.05</td>
</tr>
<tr>
<td>Balance*Female</td>
<td>0.06152</td>
<td>0.04661</td>
<td>1.320</td>
<td>.09</td>
</tr>
<tr>
<td>Balance*Male</td>
<td>0.16405</td>
<td>0.05172</td>
<td>3.172</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Walking ability* Female</td>
<td>0.09238</td>
<td>0.04840</td>
<td>1.909</td>
<td>&lt;.05</td>
</tr>
<tr>
<td>Walking ability* Male</td>
<td>0.24888</td>
<td>0.05933</td>
<td>4.195</td>
<td>&lt;.01</td>
</tr>
</tbody>
</table>

*Scores of men were normalized to be comparable with women’s scores: see text for details.

doi:10.1371/journal.pone.0070799.t004

### Discussion

The goal of the present study was threefold. First, to determine which domains of physical performance are the strongest predictors of cognitive performance. Second, to identify whether these physical predictors vary for different aspects of cognitive performance, and third, to determine whether age and gender moderate the relation between physical and cognitive performance.

The strongest physical predictors of cognitive performance

In our study muscle strength and a gender-moderated effect of balance and walking ability were significant predictors of cognitive performance. The predictive value of walking ability, balance and muscle strength for different cognitive tests, such as the TMT (set-shifting), MMSE (global cognition), and Stroop (inhibition) has been observed previously [5,50–54]. Moreover, in the recently published ‘central benefit model’ of Liu-Ambrose and colleagues [55] the importance of the association between walking ability (gait speed), balance and executive functions is postulated as well. For example, falls are not only related to a decline in gait, balance, and muscle strength, but also to a decline in executive functions [55]. Gait is not a simple motor task for older adults. With aging, gait increasingly demands cognitive control [56]. Gait speed, an important component of gait, is associated with executive functions (Stroop test) and also with other cognitive performance such as global cognitive functioning (MMSE) [52,53,57]. It has been argued that a higher gait speed increases the cerebral blood flow especially in the prefrontal cortex, a brain region that plays a crucial role in executive functions [58,59].
Moreover, to be able to maintain balance it is important to have the capability to activate muscles properly, to respond to balance threats, and to possess sufficient levels of muscle strength [60]. Such a role of strength and balance in physical performance might explain why both are predictive of performances across different cognitive components. More specifically, the balance task used in this study also appeals to the executive functions, such as inhibition (i.e. not being distracted by noise) and cognitive flexibility (i.e. being able to compensate for errors).

As strength, walking ability, and balance can be trained in older adults [61–63], future studies that focus on the causal relationship between these physical domains and cognitive performance are necessary. Besides the physical performance measures, several other variables were significant predictors of cognitive performance. Not surprisingly, older participants and participants with a lower education level showed reduced cognitive performance compared to younger and higher educated participants. In addition, men showed lower cognitive performance than women, which is in line with previous findings [18,19].

**Physical performance and different aspects of cognitive performance**

In our study we did not find a differential effect of the physical performance measures balance, strength, and walking ability on the different domains of cognitive performance. This finding might suggest that the link between physical and cognitive performance is relatively similar across cognitive domains for healthy older adults. However, as we did not assess all cognitive domains (e.g., non-verbal memory and planning), further studies are needed to assess if these results also extend to other cognitive domains.

**The moderating effect of gender and age on the association between physical and cognitive performance**

In line with other studies that detected gender differences in cognitive decline and physical decline [11,14,16–19], we identified a moderating effect of gender on the relationship between cognitive performance and the physical measures of walking ability and balance. For men, both physical measures strongly predicted cognitive performance, but for women only walking ability was a significant predictor of cognitive performance (albeit more reduced than for men).

It is possible that differences in brain morphology between men and women [11–14] contribute to these gender effects, or that these sex differences are caused by different metabolic and hormonal responses between men and women [64]. We therefore recommend that future studies specifically test for a possible gender effect.

There is strong evidence for a temporal dissociation between the decline in cognition, muscle strength, and balance [4,20,21]. We expected that age would influence the association between physical and cognitive performance, but we did not find such a moderating effect. The lack of additional interaction effects was not in line with our expectations and previously reported age effects [65]. Perhaps participation bias might have attenuated the expected age interaction effects, as very fit participants were excluded by design and many older adults with cognitive or physical difficulties normally refrain from participating in such studies [43].

**Frailty**

Our study consisted of relatively healthy elderly. Given that the number of elderly increases rapidly in the Netherlands [66], the number of frail elderly will probably increase even more in the following decades [67]. Especially frail elderly are at risk for adverse events such as falls, hospital admission and cognitive decline [68]. Although our present findings (i.e. the link between physical and cognitive performance measures) cannot be generalized to frail elderly (see below), they do fit the discussion about the concept of frailty. Our findings support the idea of Rockwood and colleagues [69] that frailty is an accumulation of deficits, and should not only consist of physical parameters [70] but also of other parameters such as cognitive performance.

**Limitations**

The present study also has several limitations. First, this is a cross-sectional study and therefore conclusions about causality cannot be drawn. Second, the participants in this study (aged over 65) formed a rather homogeneous group. The healthiest elderly were not included and their non-responding peers, we suspect, would have been more frail (i.e. having lower physical and cognitive performance) than those who were included. These biases obviously restrict the generalizability of our results to other subgroups. Although the physical and neuropsychological test scores were similar to those reported in other cross-sectional studies in healthy older adults [9,53,57,71–74], and similar prediction accuracies were found compared with previous studies [9,75], further studies need to assess if our results (presented in Table 4) are valid for other subgroups, such as participants suffering from cognitive decline, frail participants, or participants under the age of 65.
Conclusions

To the best of our knowledge, this is the first comprehensive assessment of the relationship between physical and cognitive performance in healthy older adults. We identified walking ability, balance and strength to be significant predictors of cognitive performance. Our finding that walking ability and balance are stronger predictors of cognitive performance for men than women, suggests that the effect of strength and balance training in older men might have a larger impact on cognitive performance than for women.

References


