



EUNAAPA – Work Package 4

**Expert Survey on Assessment Instruments  
for Physical Activity and Physical Functioning in Older People:**

**National Report Germany**

*Prof. Dr. Alfred Rütten*

*Peter Gelius*

*Dr. Ellen Freiburger*

*Dr. Karim Abu-Omar*

*Friedrich-Alexander-Universität Erlangen-Nürnberg  
Institut für Sportwissenschaft und Sport*



## 1 Introduction

The EUNAAPA project's work package No. 4 (Assessment of Physical activity and Physical Functioning in older People) is supposed to produce an inventory and a critical review of current instruments for physical activity and physical functioning assessment among older people (Project Deliverable D6). From February to March 2007, all Associated Partners of the project conducted an expert survey in their respective nations. The resulting data is to be submitted to the work package leader (Karolinska Institute, Stockholm) along with a national report that will be used to compile the final deliverable. According to the instructions given by the work package leader, the national reports are supposed to address the following issues:

- Which instruments are most common and why?
- Which instruments are not used at all and why?
- Are there any local instruments/instruments in the partner's native language?
- With respect to the knowledge of assessment instruments, are there any patterns if one compares the different disciplines and professionals?
- Associated Partner's comment on the results

## 2 Methods

### 2.1 Expert Selection

#### *Sampling Procedure*

The national survey for Germany was conducted between 12 February and 16 March 2007. Expert selection was based on the instructions and sampling matrix provided by the work package leader (see table 1). Selection partly took place by means of convenience sampling. Experts known to the authors were listed and then put into the fields that suited their background most. Simultaneously, institutions working in the relevant sectors were approached and asked to name experts willing and able to participate in the survey. This was especially necessary for the four matrix fields related to experts with a governmental background, which might in part be due to the German political and administrative structures dealing with physical activity, physical functioning, and senior citizens. There is no single German governmental institution responsible for physical activity/functioning in older people, and consequently there is no center of expertise in the governmental sector with experts generally known in the field.

#### *Experience with Sampling*

While the Federal Ministry of Health is partly responsible for institutionalized older adults, the Ministry of Family Affairs deals with senior citizen's concerns in general, while the Ministry of the Interior is responsible for sport. In addition, national ministries tend to refer requests to their counterparts on the regional level of the 16 Federal States (Länder), and vice versa. The situation of non-exclusive jurisdiction on the national level is mirrored on the federal state level, with the possible addition of the Länder's Ministries of Culture.

All in all, 17 experts and institutions were selected as potential respondents, covering 13 of the 16 fields of the expert selection matrix (see table 2). The results of the sampling process support the above-mentioned assumption on German governmental structures. On the one hand, there are more than enough respondents for the academics/professional education sector (5 respondents while only 4 would have been necessary). On the other hand, finding respondents with a governmental background proved to be difficult. Only 1 of the 4 fields relating to governmental institutions could be filled (No.1). The expert originally approached for this field felt unable to complete the questionnaire but recommended a colleague at the same institution who was able to step in and take part in the survey. On the regional level (field No. 13), the expert originally approached was unable to fill out the questionnaire and recommended another governmental section, which, however, did not feel competent enough to fill out the questionnaire either.

Naming potential respondents turned out to be easier for the community-dwelling setting (only one out of eight fields [No. 9] could not be filled) than for the institutionalized setting (two fields could not be filled [Nos. 14 and 15]; one field [No. 7] could only be filled by selecting a respondent that had already been selected for a field in the community-dwelling setting [No. 3]). Likewise, finding potential respondents for the national level was easier (no field missing; one field [No. 7] filled by a respondent that had already been selected for another field on the same level [No. 3]) than for the regional/local level (three fields missing [Nos. 9, 14, and 15]).

Experts from different German regions were approached: Berlin, Lower Saxony, and Bavaria with a special emphasis on the Greater Nuremberg region. Contacts could only be successfully established in 11 out of 17 cases, covering 9 fields of the matrix. While all 11 individuals that were eventually sent the questionnaire responded, one subject sent back the questionnaire too late to be included in the analysis. Thus, the final sample amounts to 10 subjects.

Table 1: Sampling matrix for Germany

	Community-dwelling older adults				Institutionalized older persons			
National level	Government	Health care/ social care	Commercial sector	Academics/ Professional Education	Government	Health care/ social care	Commercial sector	Academics/ Professional Education
	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>
	<b>Expert A</b>	X X	<b>Expert B</b>	<b>Expert C</b> <b>Expert D</b>	X	X*	<b>Expert B</b>	<b>Expert E</b>
Regional/ local level	Government	Health care/ social care	Commercial sector	Academics/ Professional Education	Government	Health care/ social care	Commercial sector	Academics/ Professional Education
	<b>9</b>	<b>10</b>	<b>11</b>	<b>12</b>	<b>13</b>	<b>14</b>	<b>15</b>	<b>16</b>
		<b>Expert F</b> <b>Expert G</b> X	X	<b>Expert H</b> <b>Expert I</b>	X			<b>Expert J</b>

*Actual respondents given in bold.*

*X denotes experts/institutions that were originally selected but that could not be successfully contacted*

*\* Expert responded but sent back questionnaire too late to be included in the analysis.*

## 2.2 Expert Self-Rating

In order to double-check the *ex ante* allocation of experts to certain fields of the matrix, respondents were asked to self-rate their organizational level, field of expertise, setting, and sector (see table 2). The rating shows that both settings (community-dwelling and institutionalized older adults), both organizational levels (national and regional/local), and both fields (physical activity and physical functioning) are well-represented in the sample. With respect to the sectors, however, it must be noted that none of the experts comes from the social care sector. This, however, may be partly due to the fact that this field only appears in the questionnaire while it shares a field with the health care sector in the sampling matrix. If the sampling matrix is taken as a standard, all sectors have been covered by the survey. While all experts gave non-ambiguous information on their sector, the self rating shows that there is a certain overlap concerning the expertise in the different settings, organizational levels, and especially fields: 4 out of 9 respondents claim to be experts for both settings, 3 out of nine rate themselves experts for both the national and regional/local level, and 5 out of 9 consider themselves to be knowledgeable in both physical activity and physical functioning.

Table 2: Expert self rating

	Setting		Organizational level		Field		Sector				
	Community-dwelling older adults	Institutionalized older adults	National level	Regional/local level	Physical activity	Physical functioning	Governmental sector	Health care	Commercial sector	Academic sector	Social care sector
Expert A			X		X	X	X				
Expert B	X	X	X	X	X	X			X		
Expert C	X		X		X	X				X	
Expert D	X	X	X	X		X		X			
Expert E		X	X		X	X				X	
Expert F	X			X	X	X		X			
Expert G								X			
Expert H	X	X	X		X					X	
Expert I	X			X	X					X	
Expert J	X	X	X	X	X	X		X			

## 2.3 Data Collection

As agreed upon at the special meeting of WP4 in January 2007, the initial contact with the survey subjects was made by phone in order to explain the objectives of the questionnaire and to ensure the experts' compliance and understanding of the survey process. The questionnaire supplied by the work package leader was then sent to the experts either as a hardcopy by mail or as a PDF document by email, to be printed out by the respondents. All subjects were asked to complete the questionnaire by 16 March and send it back to the national partner by mail.

Additional questions on the questionnaire and survey procedures were answered by telephone. In an effort to speed up the implementation of the survey, a number of experts was interviewed by telephone a few days after receiving the questionnaire. The respective questionnaires were completed by the interviewer.

## 3 Results

### 3.1 Frequently used and common tests

Table 3 shows the number of tests rated as "used in my country", sorted by respondent and survey section. The number of tests known to be currently in use in Germany varies significantly between respondents, ranging from 1 out of a total of 84 tests in the survey (expert A) to 46 out of 84 (expert C). There are particularly pronounced differences in section B (Physical Activity): While 5 experts did not name a single test, 2 rated 10 and 13, respectively, as "currently used". A similar but less

pronounced pattern might be discerned for sections J and I. At the other end of the spectrum are sections F and G, where most experts do not name a single test as being currently used in Germany. However, this result might be attributable to the small number of tests dealt with in these sections. Section E (Balance) stands out in that 5 experts rate 7 or more of the 13 tests as currently in use. In sections D, E, and H, all but one expert know at least one test to be currently in use in Germany.

Table 3: Number of tests rated “used in my country” by respondent

Section	Expert									
	Expert A	Expert B	Expert C	Expert D	Expert E	Expert F	Expert G	Expert H	Expert I	Expert J
Section B: Physical Activity (17 tests)	1	0	13	5	10	0	0	0	0	4
Section C: Physical Functioning – Endurance (5 tests)	0	1	4	2	3	1	0	0	2	4
Section D: Physical Functioning – Mobility (9 tests)	0	5	6	4	5	3	2	2	4	6
Section E: Physical Functioning – Balance (13 tests)	0	7	7	8	7	2	5	1	5	7
Section F: Physical Functioning – Range of Motion (3 tests)	0	0	0	2	0	0	0	0	0	0
Section G: Physical Functioning – Dexterity (2 tests)	0	0	1	0	0	0	0	0	0	0
Section H: Physical Functioning – Muscle Strength (7 tests)	0	2	5	6	4	2	1	1	3	3
Section I: Physical Functioning – Overall Index Tests (15 tests)	0	2	6	4	5	1	0	1	4	7
Section J: Physical Functioning – Daily Living (13 tests)	0	6	4	4	6	0	0	3	7	6
<b>Overall (84 tests)</b>	<b>1</b>	<b>23</b>	<b>46</b>	<b>35</b>	<b>40</b>	<b>9</b>	<b>8</b>	<b>8</b>	<b>25</b>	<b>37</b>

Table 4 gives a combined overview over the tests that experts consider to be currently in use in Germany as well as their assessment of how common these tests are. While a test that is rated as “currently used” by a large number of respondents does not necessarily imply that this test is also “very common”, the figures below show that the tests that score highest in one category also appear in top positions in the other. Judging from this table, the *Timed Up and Go* test, the *Romberg Test*, the *Tandem Stance*, the *Chair Stand 5 Times*, and *Tinetti’s Performance-oriented Mobility Assessment* are the most common tests currently used in Germany.

Table 4: Tests rated as “currently used” and “very common”

Test name	Number of expert rating test as „currently used“	Number of experts rating test as „very common“
Timed Up and Go, Romberg Test	9	6
Tandem Stance, Chair Stand 5 Times	8	6
Stops Walking While Talking	8	2
Tinetti's Performance-oriented Mobility Assessment	7	6
Barthel Index	7	5
Get Up And Go, Grip Strength	7	4
Walking Speed 10m, Lawton IADL	7	3
6-Minute Walk	7	2
One Leg Stance	6	4
Katz ADL	6	2
Berg's Balance Scale	6	1
Functional Reach, Chair Stand 3 Times	5	3
Short Physical Performance Battery, ADL	5	2
Pedometer	4	1
IPAQ, 12-Minute Walk, 2-Minute Walk, FAQ	4	0
PAR, PASE	3	1
YPAS, FICSIT 3-Balance, FICSIT 4-Balance, Step Test, Climbing Boxes, Functional Fitness, AAPHER, Combined ADL-IADL	3	0
FIM	2	2
Modified Baecke Questionnaire, Step Test, 180 Degree Turn, Physical Performance Test	2	1
Minnesota Leisure-time Physical Activity Questionnaire, CHAMPS, Life Space, Accelerometer, Doubly-Labeled Water, TUG manual, Walking Speed 30m, Figure of Eight, Chair Stand Once, Physical Performance and Mobility Examination, Elderly Mobility Scale, Groningen Fitness Test	2	0
Hand In Neck, Hand In Back	1	1
Modified Dallosso Questionnaire, CHAMPS questionnaire for older adults, Zupthen Physical Activity, FAI, Energy Expenditure (METS), L Test, Dynamic Gait Index, TUSS, Box And Block Test, Chair Stand 10 Times, Chair Stand 30 Seconds, Physical Fitness Field Test, Clinical Outcomes, Nursing Home Disability Instrument, General Motor Function Assessment, Nottingham ADL, OARS-IADL	1	0
Walking Habits, Endurance Shuttle Walking Test, Functional Ambulation, Solec Test, Modified Figure Of Eight, Pour Out Of Pot, Nine Hole Peg Test, Timed Functional Movements, Functional Fitness in Daily Functioning, Modified Elderly Mobility Scale, Bristol ADL, Cleveland ADL, IAM, ADL Staircase	0	0

### 3.2 Tests not used<sup>1</sup>

9 tests were rated as “not used in Germany” by 3 respondents each, while 2 experts each rated 24 tests as not being used.

Table 5: Tests rated as “not used”

Number of experts indicating that the instrument was not used in Germany	Test name
3	Modified Baecke, Functional Ambulation, Solec Test, Modified Figure Of Eight, Pour Out Of Pot, Bristol ADL, Cleveland ADL, IAM, ADL Staircase
2	Minnesota Leisure Time PA Questionnaire, Modified Dallosso Questionnaire, PAR, Zupthen Physical Activity, Walking Habits, 12-Minute Walk, 2-Minute Walk, Endurance Shuttle Test, Step Test, TUG Manual, L Test, Walking Speed 30 m, TUSS, FICSIT 3-Balance, FICSIT 4-Balance, Figure Of Eight, Hand In Back, Nine Hole Peg Test, Chair Stand 30 Seconds, Timed Functional Movements, Functional Fitness, Modified Elderly Mobility Scale, ADL, Nottingham ADL

The results for this item, however, are not always congruent with the previous section: In some cases, an instrument is labeled “not used” by one expert but rated “currently used” by another. A proxy to finding out which tests are currently not used in Germany is a comparison of tables 4 and 5. It yields a list of tests that were mentioned as “not being used” while never being mentioned as “currently used”. These tests are thus most likely not to be currently in use in Germany:

Table 6: Tests most probably not in use in Germany

Tests rated as “currently used” by 0 respondents <i>and</i> rated “not used” by at least 2 respondents
Functional Ambulation, Solec Test, Modified Figure Of Eight, Pour Out Of Pot, Bristol ADL, Cleveland ADL, IAM, ADL Staircase, Walking Habits, Endurance Shuttle Test, Nine Hole Peg Test, Timed Functional Movements, Functional Fitness, Modified Elderly Mobility Scale

### 3.3 Main reasons why tests are not used

Assessing why tests are not used in German is difficult if one only looks at the “why not” questions of the questionnaire. In the entire survey (i.e. for all tests and by all respondents), only 25 answers were given. In 24 cases, experts indicated that a test was not used because it was “not known”. Only in one case, a test was rated as “too time-consuming”.

<sup>1</sup> In this overview, only test with two and more negative answers are listed as one respondent labelled most tests with “not used” while it must be assumed that he/she had intended to answer “don’t know”.

The picture might become slightly clearer if one takes into account the disadvantages of certain tests mentioned in the experts' comments (see 3.4). This might be problematic, however, as a negative comment on a test does not necessarily imply that the expert who made this comment rated the test as “not used”.

### 3.4 Test Ratings

Table 7 shows the respondents' ranking of the tests in the survey.<sup>2</sup> The tests with the best overall rankings are *Timed Up and Go*, the *Romberg Test*, and the *Get Up and Go* test.

Table 7: Tests rated “very good” or “fairly good”

Test name	Number of experts rating general opinion of test as „very good“	Number of experts rating general opinion of test as „fairly good“
Timed Up and Go	7	2
Romberg Test	5	1
Get Up and Go Test	4	2
Chair Stand 5 Times, Barthel Index	3	3
IADL, One Leg Stance	3	2
Tandem Stance	3	2
Grip Strength	3	0
Tinetti's Performance-oriented Mobility Assessment	2	4
Chair Stand 3 Times	2	1
Short Physical Performance Battery	1	4
Berg's Balance Scale	1	3
Combined ADL-IADL	1	0
Walking Speed 10 m	0	6
6-Minute Walking	0	5
Stops Walking While Talking, ADL Index, Katz ADL	0	3
Pedometer, PASE, YPAS, Accelerometer, OpinionC3, OpinionC5, OpinionD5, OpinionE1	0	2
Minnesota Leisure Time PA Questionnaire, Modified Baecke Questionnaire, PAR, CHAMPS, Zupthen Physical Activity, FAI, Energy Expenditure (METS), 12-Minute Walk, TUG Manual, TUSS, 180 Degree Turn, Hand in Neck, Hand in Back, Chair Stand 30 seconds, Physical Fitness Field Tests, Clinical Outcome Variables, Physical Performance Test, Functional Fitness, Physical Performance and Mobility Examination, Elderly Mobility Scale, General Motor Function Assessment, Functional Activities Questionnaire (FAQ), FIM	0	1

<sup>2</sup> The ranking follows the principle of an Olympic medal count: The table was first sorted by the number of times a test was rated “very good”, then by number of times rated “fairly good”.

Table 8 summarizes the tests rated either “very bad” or “rather bad” by the respondents. The *IPAQ* was the only test in the entire survey that was rated “very bad” by a respondent. 8 tests were rated “rather bad” by one respondent each. However, 6 of these had also been rated “fairly good” or “very good” by at least one other respondent.

Table 8: Tests rated “very bad” or “rather bad”

Test name	Number of experts rating general opinion of test as „very bad“	Number of experts rating general opinion of test as „rather bad“
IPAQ	1	0
Step Test, Chair Stand 10 Times, Romberg Test*, Lawton IADL*, 6-Minute Walk*, Stops Walking While Talking*, ADL Index*, Pedometer*	0	1

\* Test was also rated as being generally considered “fairly good” or “very good” by at least one respondent

Moreover, in their comments, some of the tests were criticized for various reasons:

- too time-consuming: *Accelerometer, Doubly-labeled Water, 6-Minute Walk Berg’s Balance Scale*
- too expensive: *Accelerometer, Doubly-labeled Water*
- measurement or validation problems: *Functional Reach, Accelerometer, Grip Strength, Minnesota Leisure-time PA Questionnaire*
- limited target group: *IPAQ* (age limit), *Berg’s Balance Scale* (only for frail people), *ADL Index* (rather bad if used by community-dwelling persons)
- general problems: The *IPAQ* was rated as being problematic because “participants are having trouble with this instrument”.

### 3.5 Translation:

Table 9 gives an overview of the tests and instruments that respondents rated as “translated into German”, while table 10 indicates which tests were most frequently rated as translated by means of scientific procedures. While these tables are not apt to clarify which tests have indeed been translated or if scientific procedures were truly used, they may give information on how well-known these translations are among German experts.

Table 9: Tests translated into German

Number of experts indicating that the instrument has been translated into German	Test name
9	Timed Up and Go
8	Romberg Test
7	Chair Stand 5 Times, Lawton IADL, Barthel Index
6	Get Up and Go Test, One Leg Stance, Tandem Stance, Tinetti's Performance-Oriented Mobility Assessment, Katz ADL
5	Walking Speed 10 m, Stops Walking While Talking, Berg's Balance Scale
4	Chair Stand 3 Times, ADL Index
3	6-Minute Walk, Functional Reach, Short Physical Performance Battery, Functional Activities Questionnaire (FAQ)
2	Modified Baecke Questionnaire, PAR, PASE, IPAQ, Walking Speed 30 m, Grip Strength, Combined ADL-IADL
1	Zupthen Physical Activity, Life Space, Modified Dalosso Questionnaire, CHAMPS, Step Test, TUG Manual, Figure of Eight, 180 Degree Turn, Climbing Boxes, Chair Stand 10 Times, Chair Stand 30 Seconds, Physical Performance Test, Functional Fitness, FIM

Table 10: Tests translated by means of scientific procedures

Number of experts indicating that scientific procedures were used for translation	Test name
4	Lawton IADL, Barthel Index, Get Up and Go Test, Tinetti's Performance-Oriented Mobility Assessment, Berg's Balance Scale
2	Timed Up and Go, Stops Walking While Talking, ADL Index
1	Chair Stand 5 Times, One Leg Stance, Tandem Stance, Katz ADL, Walking Speed 10 m, 6-Minute Walk, Functional Reach, Short Physical Performance Battery, Modified Baecke Questionnaire, Grip Strength, Combined ADL-IADL, Figure of Eight, FIM

### 3.6 Comparison of expert groups

The following section contains a comparison of the different expert groups, differentiated by sector, organizational level, and setting. Only the question “Is this instrument currently used in your country?” was taken into consideration for this part of the data analysis. For all other survey questions, the number of answers was too small to allow for a further subdivision of the sample.

#### 3.6.1 Comparison by Sector

Table 11 gives an overview over the number of tests rated “used in my country” by all experts combined, divided by sectors. The experts’ self-rating was used to allocate respondents to the different sectors.<sup>3</sup> Given in parentheses is the theoretical maximum value for the sector, which would be achieved if all experts in the sector would rate all 84 tests in the survey as “used” (e.g. health care sector (n=4):  $\max = 4 \cdot 84 = 336$ ).

Table 11: Number of tests rated „used in my country“ by all experts combined, by sector

Governmental Sector (n=1) (max=84)	Health Care Sector (n=4) (max=336)	Commercial Sector (n=1) (max=84)	Academic Sector (n=4) (max=336)
1 1.19%	89 26.48%	23 27.38%	120 35.71%

It is striking that the only respondent from the governmental sector only rated a single instrument as “currently used” – the lowest figure in the entire survey. However, due to the small n in the governmental and commercial sector (both n=1), a detailed comparison seems advisable only between the health care and the academic sector (both n=4).

At a theoretical maximum of 336 tests for both sectors, 120 tests were rated “used” by academic experts compared to 89 by the health care experts. As table 12 illustrates, both the mean ( $M=30.00$ ,  $SD=17.30$ ) and the median ( $Md=32.50$ ) are higher for the academic sector than for the health care sector ( $M=22.25$ ,  $SD=15.90$ ,  $Md=22.00$ ).

The detailed comparison of means and medians for the various sections of the survey shows that there is a considerable difference between the two groups in Section B (Physical Activity): The mean is considerably higher for the academic sector ( $M=6.00$ ,  $SD=7.12$ ) than for the health care sector ( $M=2.25$ ,  $SD=2.63$ ). The same applies to the Median (academic sector:  $Md=5.00$ , health care sector:  $Md=2.00$ ). Another section with a strong contrast is Section J (Activities of Daily Living). Again, the figures for the academic sector ( $M=5.00$ ,  $SD=1.83$ ,  $Md=5.00$ ) are higher than those for the health care sector ( $M=2.50$ ,  $SD=3.00$ ,  $Md=2.00$ ). In sections C, D, H, and I of the survey, the difference between the two sectors is much less pronounced, although the figures for the academic sector are slightly higher. In sections E, F, and G, there hardly any discernable difference between the academic and the health care sector. In the cases of sections F and G, this might be partly caused by the small number of tests covered in these parts of the survey.

<sup>3</sup> Only in two cases, the expert’s self-rating differed from the original allocation made in the sampling matrix. Both experts had originally been sampled as academic experts but rated themselves as health care experts. All respondents made a non-ambiguous self-allocation to a single sector.

Table 12: Number of tests rated “used in my country”: means and medians (health care and academic sector only).

Section		Health Care Sector (n=4)	Academic Sector (n=4)
Section B: Physical Activity (17 tests)	Mean	2,25	6,00
	(SD)	(2,63)	(7,12)
	Median	2,00	5,00
Section C: Physical Functioning – Endurance (5 tests)	Mean	1,75	2,25
	(SD)	(1,71)	(1,71)
	Median	1,50	2,50
Section D: Physical Functioning – Mobility (9 tests)	Mean	3,75	4,25
	(SD)	(1,71)	(1,71)
	Median	3,50	4,50
Section E: Physical Functioning – Balance (13 tests)	Mean	5,50	5,00
	(SD)	(2,65)	(2,83)
	Median	6,00	6,00
Section F: Physical Functioning – Range of Motion (3 tests)	Mean	0,50	0,00
	(SD)	(1,00)	(0,00)
	Median	0,00	0,00
Section G: Physical Functioning – Dexterity (2 tests)	Mean	0,00	0,25
	(SD)	(0,00)	(0,50)
	Median	0,00	0,00
Section H: Physical Functioning – Muscle Strength (7 tests)	Mean	3,00	3,25
	(SD)	(2,16)	(1,71)
	Median	2,50	3,50
Section I: Physical Functioning – Overall Index Tests (15 tests)	Mean	3,00	4,00
	(SD)	(3,16)	(2,16)
	Median	2,50	4,50
Section J: Physical Functioning – Activities of Daily Living (13 tests)	Mean	2,50	5,00
	(SD)	(3,00)	(1,83)
	Median	2,00	5,00
<b>Overall (84 tests)</b>	Mean	<b>22,25</b>	<b>30,00</b>
	(SD)	<b>(15,90)</b>	<b>(17,30)</b>
	Median	<b>22,00</b>	<b>32,50</b>

### 3.6.2 Comparison by organizational level

In their self-rating, many of the respondents allocated themselves both to the national and regional/local level, which is why the original allocation in the sampling matrix was used to divide respondents into two groups. Table 13 gives an overview over the number of tests rated “used in my country” by all experts combined, sorted by organizational level.

Table 13: Number of tests rated „used in my country“ by all experts combined, by organizational level

National Level (N=5) (max=420)	Regional/Local Level (N=5) (max=420)
146 34.76%	87 20.71%

The number of respondents is n=5 for both organizational levels, which leads to a theoretical maximum number of tests rated “used” of 420 in both cases. The experts on the national level rate a total of 146 tests as “currently used in Germany” compared to 87 tests for the regional/local experts. As illustrated in table 14, the overall means and medians, which are both higher for the national than for the regional/local expert group.

In the case of the national experts, the median (Md=35.00) is considerably higher than the mean (M=29.20, SD=18.03). A detailed analysis of the national experts shows that the figures for one respondent in this group were much lower than for the other four.

Likewise, the median (Md=9.00) is considerably lower than the mean (M=17.40, SD=13.13) for the regional/local experts. As the figures for the individual respondents in this group show, two regional/local experts rated far more tests as “currently” used than the other three.

Table 15 also shows that national experts rate more tests as “used” and consequently reach higher means than their regional counterparts in all survey sections. Once more, there is a considerable difference for section B (Physical Activity), with M=6.00 (SD=5.96) and Md=5.00 for the national level compared to M=0.80 (SD=1.79) and Md=0.80 for the regional/local level. Once more, the difference is least discernable in sections F and G, which might once more be partly due to the small number of tests in these two parts of the survey.

Table 14: Number of tests rated “used in my country”: means and medians, by organizational level

Section		National Level (n=5)	Regional/Local Level (n=5)
Section B: Physical Activity (17 tests)	Mean	6,00	0,80
	(SD)	(5,96)	(1,79)
	Median	5,00	0,80
Section C: Physical Functioning – Endurance (5 tests)	Mean	2,00	1,40
	(SD)	(1,58)	(1,67)
	Median	2,00	1,00
Section D: Physical Functioning – Mobility (9 tests)	Mean	4,00	3,40
	(SD)	(2,35)	(1,67)
	Median	5,00	3,00
Section E: Physical Functioning – Balance (13 tests)	Mean	5,80	4,00
	(SD)	(3,27)	(2,45)
	Median	7,00	5,00
Section F: Physical Functioning – Range of Motion (3 tests)	Mean	0,40	0,00
	(SD)	(0,89)	(0,00)
	Median	0,00	0,00
Section G: Physical Functioning – Dexterity (2 tests)	Mean	0,20	0,00
	(SD)	(0,45)	(0,00)
	Median	0,00	0,00
Section H: Physical Functioning – Muscle Strength (7 tests)	Mean	3,40	2,00
	(SD)	(2,41)	(1,00)
	Median	4,00	2,00
Section I: Physical Functioning – Overall Index Tests (15 tests)	Mean	3,40	2,60
	(SD)	(2,41)	(2,88)
	Median	4,00	1,00
Section J: Physical Functioning – Activities of Daily Living (13 tests)	Mean	4,00	3,20
	(SD)	(2,45)	(3,27)
	Median	4,00	3,00
<b>Overall (84 tests)</b>	Mean	<b>29,20</b>	<b>17,40</b>
	(SD)	<b>(18,03)</b>	<b>(13,13)</b>
	Median	<b>35,00</b>	<b>9,00</b>

### 3.6.3 Comparison by Setting

As in the case of organizational levels, many of the respondents allocated themselves both to the community-dwelling and institutionalized setting, which is why the original allocation in the sampling matrix was used to divide respondents into two groups. One respondent had to be excluded from the analysis as the expert had been sampled for both settings. This led to  $n=7$  for the community-dwelling and  $n=2$  for the institutionalized setting. As the latter figure is too small for a meaningful statistical analysis, the following results have to be treated with caution.

Table 15: Number of tests rated „used in my country“ by all experts combined, by setting

Community-dwelling Older People (N=7) (max=588)	Institutionalized Older People (N=2) (max=168)
133 22.61%	77 45.83%

Table 15 gives an overview over the number of tests rated “used in my country” by all experts combined, sorted by setting. The figures seem to suggest that experts for the institutionalized setting rate far more tests as “currently used” (77 out of a maximum of 168, amounting to 45.83%) than experts for the community-dwelling setting (133 out of 588, 22.61%). As shown in table 16, the mean is  $M=38.50$  ( $SD=2.21$ ) for the institutionalized setting compared to  $M=19.00$  ( $SD=17.04$ ) for the community-dwelling setting. The low median ( $Md=9.00$ ) for the community-dwelling setting suggests that there are some outliers within this group. A detailed analysis shows that 3 out of 7 experts for community-dwelling older people rated far fewer tests as “currently used” than the rest of the group. Consequently, the medians are preferred to the means in the following comparison of the various sections of the survey.

The analysis shows that experts for the institutionalized setting rate more tests as “currently used” than experts for the community-dwelling setting in all sections except for sections F and G ( $Md=0.00$  for both groups). The difference is particularly stark in section B (institutionalized setting:  $Md=7.00$ , community-dwelling setting:  $Md=0.00$ ), section I (institutionalized setting:  $Md=6.00$ , community-dwelling setting:  $Md=1.00$ ), and section J (institutionalized setting:  $Md=6.00$ , community-dwelling setting:  $Md=3.00$ ). As mentioned above, however, due to the small  $n$  in the institutionalized setting it is questionable how much these results can tell us about the difference between the two groups.

Table 16: Number of tests rated “used in my country”: means and medians, by setting

Section		Community-dwelling Older People (n=7)	Institutionalized Older People (n=2)
Section B: Physical Activity (17 tests)	Mean	2,86	7,00
	(SD)	(5,24)	(4,24)
	Median	0,00	7,00
Section C: Physical Functioning – Endurance (5 tests)	Mean	1,29	3,50
	(SD)	(1,50)	(0,71)
	Median	1,00	3,50
Section D: Physical Functioning – Mobility (9 tests)	Mean	3,00	5,50
	(SD)	(1,91)	(0,71)
	Median	3,00	5,50
Section E: Physical Functioning – Balance (13 tests)	Mean	4,00	7,00
	(SD)	(3,06)	(0,00)
	Median	5,00	7,00
Section F: Physical Functioning – Range of Motion (3 tests)	Mean	0,29	0,00
	(SD)	(0,76)	(0,00)
	Median	0,00	0,00
Section G: Physical Functioning – Dexterity (2 tests)	Mean	0,14	0,00
	(SD)	(0,38)	(0,00)
	Median	0,00	0,00
Section H: Physical Functioning – Muscle Strength (7 tests)	Mean	2,57	3,50
	(SD)	(2,23)	(0,71)
	Median	2,00	3,50
Section I: Physical Functioning – Overall Index Tests (15 tests)	Mean	2,29	6,00
	(SD)	(2,36)	(1,41)
	Median	1,00	6,00
Section J: Physical Functioning – Activities of Daily Living (13 tests)	Mean	2,57	6,00
	(SD)	(2,70)	(0,00)
	Median	3,00	6,00
<b>Overall (84 tests)</b>	Mean	<b>19,00</b>	<b>38,50</b>
	(SD)	<b>(17,04)</b>	<b>(2,21)</b>
	Median	<b>9,00</b>	<b>38,50</b>

### 3.7 Comparison by Section<sup>4</sup>

#### 3.7.1 Section B: Physical Activity

A closer analysis of Section B indicates that a wide variety of tests is known and rated as “currently used” but that no particular test stands out. *Pedometer*, *Modified Baecke Questionnaire*, *PAR*, and *PASE* are rated “very common” by one respondent each. *Walking Habits* was the only test not rated “currently used” by any respondent.

The *IPAQ* is worth mentioning in that it is the only test in the entire survey that was rated “very bad” at all, albeit only by a single respondent. In a comment, one expert criticized that participants tended to have problems with the *IPAQ* and that there was an age limitation for the test. The *Doubly-labeled Water Test* was criticized for being too expensive and time-consuming. The same applied to the *Accelerometer*, which was also criticized because it was unclear against what this instrument had been validated.

In other comments, the *Minnesota Leisure Time PA Questionnaire* was rated as “useful for specific research” with the qualification that it required special skills. The *Zupthen Physical Activity* instrument was praised for being “one of the few instruments with analytic data” but also criticized for its restriction to males. The *Pedometer* was commended for working on a low budget and for not being time-consuming provided the right instruments were used.

#### 3.7.2 Section C: Physical Functioning – Endurance

All tests in Section C except for the *Endurance Shuttle Walking Test* were rated as “currently used” by the German experts. The *6-Minute-Walk* test reaches the top position in all rankings. It is the test that is most often rated as “currently used” (7 experts), it is most frequently rated as “very common” (2 experts), and it gets the best “grades” in the test rating (5 times “fairly good”).

In a comment, the *6-Minute-Walk* was considered too time-consuming in practical work, and one expert raised the “question of functional borderline for endurance” in connection with the *2-Minute-Walk*.

#### 3.7.3 Section D: Physical Functioning – Mobility

Except for the *Functional Ambulation* instrument, all tests in the section were rated as “currently used” at least once. There is a set of four tests that seem to be well-known, frequently used, and quite positively rated by the experts: *Timed Up and Go*, *Stops Walking While Talking*, *Get Up and Go*, and *Walking Speed 10m*. The *Timed Up and Go* reaches first place in both the “currently used” ranking (9 experts), the “very common” ranking (6 experts), and the test rating (7 experts “very good”, 2 experts “fairly good”).

In two comments, it was pointed out that *Stops Walking While Talking* is also referred to as “Lundin-Olson Test” and that its status as a real tests as well as its validation remain unclear.

#### 3.7.4 Section E: Physical Functioning – Balance

Except for the *Solec Test* and the *Modified Figure of Eight*, all tests were rated as “currently used” at least by one expert each. There seems to be a set of four tests that are commonly used and reach good grades in the ranking. The *Romberg Test* reaches top scores in all rankings: It is rated 9 times as “currently used”, 6 times as “very common”, and 5 times as “very good”/once as “fairly good”. The *Tandem Stance*, the *Once Leg Stance*, and *Berg’s Balance Scale* are comparatively common and positively rated as well.

One expert indicated that *Figure of Eight* had so far only been used by two centers. It was also criticized that *Berg’s Balance Scale* was time-consuming and could only used for frail persons. *Functional Reach* was assessed as “not easy to measure” by one expert.

<sup>4</sup> For detailed statistics, see Appendix A.

### 3.7.5 Section F: Physical Functioning – Range of Motion

Of the three tests in the section, two (*Hand in Neck*, *Hand in Back*) were rated as “currently used”, albeit only by one expert each. Each of these tests was also rated “very common” and “fairly good” once.

The comments on the Section F tests suggest a certain lack of formalization in this section: One comment stated that the *Hand in Neck* test had “probably not [been] formally translated”, while another indicated that *Hand in Back* has been used by general practitioners in a non-formalized form for a long time.

### 3.7.6 Section G: Physical Functioning – Dexterity

Of the two tests in Section G, only the *Box and Block Test* is mentioned as being “currently used” in Germany by one expert. There are no further data on the common usage of the tests, neither are there any assessments of the tests’ quality or any comments.

### 3.7.7 Section H: Physical Functioning – Muscle Strength

All tests in the section are mentioned as “currently used”. Three tests stand out from the rest, with a clear ranking: The *Chair Stand 5 Times* is most frequently rated as “currently used” (8 times), most often named as “very common” (6 times), and given the best ratings (3 times “very good” and 3 times “fairly good”). In second place is *Grip Strength* (7 times “currently used”, 4 times “very common”, 3 times “very good”); the *Chair Stand 3 Times* ranks third (5 times “currently used”, 3 times “very common”, 2 times “very good”, and once “fairly good”).

It was noted that *Grip Strength* was used by general practitioners in a non-formalized way and that it was “extremely dependent” on an “instrument screening test”.

### 3.7.8 Section I: Physical Functioning – Overall Index Tests

Of the section’s 15 tests, 3 (*Timed Functional Movements*, *Functional Fitness in Daily Functioning*, and the *Modified Elderly Mobility Scale*) were not mentioned as “currently used”. Most frequently mentioned were *Tinetti’s Performance-oriented Mobility Assessment* (7 times) and the *Short Physical Performance Battery* (5 times). The *Tinetti* test, however, was more often rated as “very common” (6 times) than the *SPPB* (2 times), and also received slightly better grades (2 times “very good”, 4 times “fairly good”) than *SPPB* (once “very good”, 4 times “fairly good”).

### 3.7.9 Section J: Physical Functioning – Activities of Daily Living

4 out of 13 tests (*IAM*, *Cleveland ADL*, *Bristol ADL*, *ADL Staircase*) were not mentioned as being “currently used” in Germany. All in all, the “best” and most frequently used test in this section seems to be the *Barthel Test* (7 times mentioned as “currently used”, 5 times as “very common”, 3 times as “very good”, and 3 times as “fairly good”), closely followed by the *Lawton IADL*. The “runners up” in this section are the *Combined ADL/IADL*, the *ADL*, and the *Katz ADL*.

It was commented that the *Barthel Index* was “used as a process determinant” and that the *ADL Index* was “rather bad if used by community-dwelling persons”.

## 3.8 Modifications

The following modifications of existing tests were indicated by the experts:

- Section B: *Minnesota Leisure Time PA Questionnaire*: a German modified version for youths (Fuchs et al. 1987)
- Section E: *Functional Reach*: a modification better for balance
- Section E: *Romberg Test*: used in many different modifications
- Section E: *FISCIT 3-Balance Scale*: the FISCIT dataset describes another protocol
- Section H: *Climbing Boxes*: a modified version with steps

### 3.9 Other Tests

The following additional tests and instruments were named by the experts:

- Section D: *Geh- und Zähltest* (Walk-and-Count Test, DMW, 2000), which might be a German modified version of the *Stops Walking While Talking/Ludin-Olson Test*
- Section F: several “quantitative measurements in physiotherapy” are being used
- Section I: there is a “quantitative anamnesis individual assessment related to Dr. Brügger”

### 3.10 Guidelines

Four major guidelines involving assessment instruments for physical activity and physical functioning in older people were mentioned by the respondents: The Guideline for Older Fall Patients by the DEGAM (2004), the Expert Standard for the Prevention of Falls in Care by the DNQP (2006), the Geriatric Basic Assessment by the AGAST (1997), and the guidelines issued by the AFGiB.

#### 3.10.1 DEGAM

In 2004, the Deutsche Gesellschaft für Allgemeinmedizin und Familienmedizin (DEGAM, German Society for General Medicine and Family Medicine) issued “Guideline No. 4” for older fall patients (Düsseldorf, 2004), which is supposed to help general practitioners assess and minimize the risk of falls among older patients. Among other instructions, the diagnostics section includes recommendations for the following tests: *Geh- und Zähltest* (Walk-and-Count Test, see 3.11), *Ludin-Olson-Test* (Stops Walking While Talking, see 3.9.3), *Up and Go Test*, *Modified Romberg Test*, and *Risk of Falls Score after Oliver*. In addition, the following recommendations are made for specialists and in-depth diagnosis: *Tinetti’s Performance-oriented Mobility Assessment*, *Self-selected Walking Speed*, *6-Minute Walk*, *Vitamin D3 in Serum*, as well as tests for diagnosis of osteoporosis. *Tinetti’s Performance-oriented Mobility Assessment* is described as the “golden standard” for the diagnosis of mobility impairments in the elderly. However, that the test is rated as time- and cost-intensive and only to be used by specialists.

#### 3.10.2 DNQP

The Deutsches Netzwerk für Qualitätsentwicklung in der Pflege (DNQP, German Network for Quality Development in Care) has developed a guideline entitled “Expertenstandard Sturzprophylaxe in der Pflege, Entwicklung - Konsentierung- Implementierung“ (Expert Standard for the Prevention of Falls in Care; Osnabrück, February 2006). It is designed to help caregivers of institutionalized older people assess and minimize the risk of falls and provides for the systematic assessment of risks for individual patients. Further details on this guideline are currently not available.

#### 3.10.3 AGAST

In 1997, the Arbeitsgruppe Geriatrisches Assessment (Workgroup for Geriatric Assessment, AGAST) published its “Geriatrisches Basisassessment” (Geriatric Basic Assessment, Munich 1997). It is a general guideline for geriatric assessment developed by a commission of experts from the Deutsche Gesellschaft für Geriatrie and the Deutsche Gesellschaft für Gerontologie). After a general screening on 21 main problem areas after Lachs, the second stage of the assessment contains the *Barthel-ADL*, *Timed Up and Go*, *Tandem Stance*, *Semi-Tandem Stance*, among others. In addition, AGAST recommends the *Hand Grip Test*.

#### 3.10.4 AFGiB

While the three above-mentioned guidelines were issued by or on behalf of national scientific and professional organizations, the Arbeitsgemeinschaft zur Förderung der Geriatrie in Bayern (AFGiB, Workgroup for the Advancement of Geriatrics in Bavaria) is a professional organization on the regional level. AFGiB analyzes a number of tests in its recommendations for a geriatric basic assessment. First, a basic screening after Lachs is recommended. For the second diagnostic stage

(see table 17), a distinction is made between instruments for the assessment of activities of daily living and mobility/risk of falls as well as between different degrees of frailty.

With respect to activities of daily living, the *Barthel Index* and *IADL* (as an additional instrument for less frail people) are recommended while *FIM* is considered adequate mainly for neurological/neuropsychological diseases but rated as considerably more time-consuming than the *Barthel Index*.

As to mobility/risk of falls, the *Timed Up and Go* is recommended while *Tinetti's Performance-oriented Mobility Assessment*, *Tandem Stance/Semi-Tandem Stance*, *Chair Rising Test*, and the *Esslingen Risk of Falls Assessment* are also mentioned. For frail people, a modified version of *Timed Up and Go* developed by the AFGiB is recommended.

Table 17: Tests mentioned in the AFGiB guideline, 2<sup>nd</sup> stage

Field	Degree of Frailty	
	Light	Severe
Activities of Daily Living	Barthel*, IADL* (optional), FIM	Barthel*, FIM
Mobility/Risk of Falls	Timed Up and Go*, Tinetti's Performance-oriented Mobility Assessment, Tandem Stance, Chair Rising Test	Timed Up and Go (modified by AFGiB)*, Esslingen Risk of Falls Assessment

\*instrument recommended by AFGiB

## 4 Discussion/Conclusions

### 4.1 Sampling and Expert Selection

The main insight to be derived from the sampling process is that it is extremely difficult to find experts on physical activity/physical functioning in the elderly from German national or regional governmental institutions. This is most probably due to the federalist structure of the German state and to the distribution of responsibilities in the fields of physical activity, physical functioning, and senior citizens across various ministries, as discussed in detail in section 2.1.

While the theoretical distinction between sectors made in the sampling matrix turned out to be useful, differentiating between organizational levels, settings, and fields was problematic. Many respondents considered themselves to be experts for both the national and regional levels, both settings, and both physical activity and physical functioning.

The group of respondents was rather heterogeneous with respect to their knowledge on the current use of tests and assessment instruments in Germany. While 4 respondents knew of less than 10 tests to be currently in use, 3 stated that almost half or even more of the 84 tests polled in the survey were currently used in Germany.

### 4.2 General Results

While a rather clear ranking of tests that are currently and commonly used in Germany can be drawn up, a proxy was needed to discern the tests that are most probably not currently used in this country. Likewise, compiling a “top ten” of tests that are generally considered to be good is easier than finding out the tests that are widely regarded as deficient. In many cases, a negative opinion given by one expert is counterbalanced by a positive one from another respondent. The criticisms of experts expressed in their comments on particular tests might be helpful in this context. Reasons for criticizing tests include costs, the time needed, issues of validation, and the lacking ease of use. The

IPAQ stands out somewhat in that it is heavily criticized by one respondent and that there are no positive statements by other experts to counterbalance this opinion.

With respect to the translation of assessment instruments, the situation is diffuse. Experts often seem to have different information as to whether a translation of a certain test exists and if it is scientific or not. One might conclude for the future that experts ought to be provided with better information on existing German translations of assessment instruments.

### 4.3 Comparison of expert groups

In general, results of the comparison of different sectors/levels should be cautiously interpreted due to the small sample sizes. The comparison of expert groups along sectors and organizational levels shows that experts from the academic sector know more tests to be currently used than experts from health care; likewise, experts from the national level rate more tests as “currently used” than their regional counterparts. One possible conclusion is that experts with a national and/or academic background tend to be better informed about a large variety of assessment instruments than experts with a health care and/or regional background. However, this finding is hardly surprising given the large number of tests polled in this survey and the fact that the members of the regional/health care group are more likely to be practitioners, while those belonging to the national/academic group tend to be researchers and administrators. In Germany, health practitioners receive relatively limited theoretical training, as such these results might reflect aspects of the German educational system.

In spite of the experts’ self-rating, physical activity seems to be a field about which only part of the respondents is truly knowledgeable. The differences between national vs. regional experts and academic vs. health care sector are nowhere as pronounced as in section B of the survey. A possible reason might be that local and health care experts deal less often with this field in their daily work than national and academic experts. Venturing a bold conclusion, one might argue that physical activity is an issue that, while being dealt with on the national and academic stage, has not reached the regional level and the work of practitioners in health care.

Differences between the various expert groups are less visible in sections D, E, and H of the survey, which might suggest that regional and health care experts deal more frequently with issues of mobility, balance, and muscle strength and consequently know a larger number of tests relevant to their own work.

### 4.3. Analysis by Survey Section

A wide variety of tests is mentioned without a clear pattern or favorites in section B (Physical Activity). On the one hand, this might support the view that at least part of the experts questioned is not well-acquainted with this field. On the other hand, it might quite as well suggest that there is currently no really good test available for the field of physical activity.

The more “equal” results for sections F (Range of Motion) and G (Dexterity) could imply that there is little knowledge of these fields in Germany. However, the result is likely to be distorted by the small number of tests polled in these two sections.

Sections C (Endurance), D (Mobility), E (Balance), and H (Muscle Strength) all yield a rather clearly-discernable set of frequently-used tests and also a clear favorite: The *6-Minute Walk* (section C), the *Timed Up and Go* (section D), the *Romberg Test* (section E), and the *Chair Stand 5 Times* (Section H). The picture is less clear in sections I (Overall Index Tests) and J (Activities of Daily Living), but there are still two most-preferred tests each: *Tinetti’s Performance-oriented Mobility Assessment* and the *Short Physical Performance Battery* for section I and the *Barthel Index* and *Lawton IADL* for section J.

### 4.4 Guidelines

There are at least 4 guidelines concerning assessment instruments for physical activity and physical functioning in older people in Germany. All were issued by professional/academic organizations, three of which are operating on the national level. There seem to be no guidelines issued by a governmental institution, which once more underscores the absence of governmental institutions from the German discourse in the field. The guidelines have a strong focus on physical functioning,

either as a subset of a general geriatric assessment or as a means to assess and minimize the risk of falls. Physical activity tests are not explicitly mentioned, and neither is the general issue of physical activity.

To conclude, the following results of the German survey seem to be of special note:

- Sampling of German experts in governmental institutions has been difficult and cumbersome. This might be due to the fact that, on the national level, responsibilities for the area of physical activity/physical functioning are not clear-cut, and that the German federal system allows only for state level regulation in this area.
- There seem to be deficits in the field of physical activity. There are no clear favorite tests, experts on the local level and from the health care sector seem to have a limited knowledge of the field, and physical activity is so far not part of any German guideline.
- There seem to be deficits with respect to the translation of tests into German, or at least with providing experts with palpable information about existing translations.
- Most importantly, the survey shows that there is a number of tests (especially in sections C, D, E, and H but also in sections I and J) that are both commonly used in Germany and considered to be working well by German experts. These findings should be taken into account when formulating a European recommendation for assessment instruments for physical activity and physical functioning in older people.

## Appendix A: Statistics by Section

### Section B: Physical Activity

Table B.1: Tests most frequently used

Number of experts rating test as „currently used”	Test name
4	IPAQ, Pedometer
3	PAR, PASE, YPAS
2	Minnesota Leisure Time PA Questionnaire, Modified Baecke Questionnaire, CHAMPS, Life Space, Accelerometer, Doubly Labeled Water
1	Modified Dallosso Questionnaire, CHAMPS Questionnaire for older adults, Zupthen Physical Activity, Frenchay’s Activity Index, Energy Expenditure (METS)
0	Walking Habits

Table B.2: Tests most frequently rated as “very common”

Number of experts rating test as „very common“	Test name
1	Pedometer
1	Modified Baecke Questionnaire
1	PAR
1	PASE

Table B.3: Test Ratings

Test name	Number of experts rating general opinion of test as „very good“	Number of experts rating general opinion of test as „fairly good“
Pedometer, Accelerometer, PASE, YPAS	0	2
Minnesota Leisure Time PA Questionnaire, Zupthen Physical Activity, Frenchay’s Activity Index, Energy Expenditure (METS), Modified Baecke Questionnaire, PAR, CHAMPS	0	1

Test name	Number of experts rating general opinion of test as „very bad“	Number of experts rating general opinion of test as „rather bad“
IPAQ	1	0
Pedometer*	0	1

\* Test was also rated as being generally considered “fairly good” or “very good” by at least one respondent

### Section C: Physical Functioning – Endurance

Table C.1: Tests most frequently used

Number of experts rating test as „currently used”	Test name
7	6-Minute Walk
4	12-Minute Walk, 2-Minute Walk
2	Step Test
0	Endurance Shuttle Walking Test

Table C.2: Tests most frequently rated as “very common”

Number of experts rating test as „very common“	Test name
2	6-Minute Walk
1	Step Test

Table C.3: Test Ratings

Test name	Number of experts rating general opinion of test as „very good“	Number of experts rating general opinion of test as „fairly good“
6-Minute Walk	0	5
2-Minute Walk, Step Test	0	2
12-Minute Walk	0	1

Test name	Number of experts rating general opinion of test as „very bad“	Number of experts rating general opinion of test as „rather bad“
6-Minute Walk*	0	1

\* Test was also rated as being generally considered “fairly good” or “very good” by at least one respondent

**Section D: Physical Functioning – Mobility**

*Table D.1: Tests most frequently used*

Number of experts rating test as „currently used”	Test name
9	Timed Up and Go (TUG)
8	Stops Walking While Talking
7	Get Up And Go Test, Walking Speed 10m
2	TUG Manual, Walking Speed 30m
1	L Test, Dynamic Gait Index
0	Functional Ambulation

*Table D.2: Tests most frequently rated as “very common”*

Number of experts rating test as „very common“	Test name
6	Timed Up and Go
4	Get up and Go
3	Walking Speed 10m
2	Stops Walking While Talking

*Table D.3: Test Ratings*

Test name	Number of experts rating general opinion of test as „very good“	Number of experts rating general opinion of test as „fairly good“
Timed Up and Go	7	2
Get Up and Go Test	4	2
Walking Speed 10 m	0	6
Stops Walking While Talking	0	3
Walking Speed 30 m	0	2
TUG Manual	0	1

Test name	Number of experts rating general opinion of test as „very bad“	Number of experts rating general opinion of test as „rather bad“
Stops Walking While Talking*	0	1

\* Test was also rated as being generally considered “fairly good” or “very good” by at least one respondent

## Section E: Physical Functioning – Balance

Table E.1: Tests most frequently used

Number of experts rating test as „currently used”	Test name
9	Romberg Test
8	Tandem Stance
6	One Leg Stance, Berg’s Balance Scale
5	Functional Reach
3	FICSIT 3-Balance Scale, FICSIT 4-Balance Scale, Step Test
2	Figure Of Eight, 180 Degree Turn
1	TUSS
0	Solec Test, Modified Figure Of Eight

Table E.2: Tests most frequently rated as “very common”

Number of experts rating test as „very common“	Test name
6	Romberg Test, Tandem Stance
4	One Leg Stance
3	Functional Reach
1	180 Degree Turn, Berg’s Balance Scale

Table E.3: Test Ratings

Test name	Number of experts rating general opinion of test as „very good“	Number of experts rating general opinion of test as „fairly good“
Romberg Test	5	1
One Leg Stance, Tandem Stance	3	2
Berg’s Balance Scale	1	3
Functional Reach	0	2
180 Degree Turn, TUSS	0	1

Test name	Number of experts rating general opinion of test as „very bad“	Number of experts rating general opinion of test as „rather bad“
Romberg Test*, Step Test	0	1

\* Test was also rated as being generally considered “fairly good” or “very good” by at least one respondent

## Section F: Physical Functioning – Range of Motion

Table F.1: Tests most frequently used

Number of experts rating test as „currently used”	Test name
1	Hand In Neck, Hand In Back
0	Pour Out Of Pot

Table F.2: Tests most frequently rated as “very common”

Number of experts rating test as „very common“	Test name
1	Hand in Neck
1	Hand in Back

Table F.3: Test Ratings

Test name	Number of experts rating general opinion of test as „very good“	Number of experts rating general opinion of test as „fairly good“
Hand in Neck, Hand in Back	0	1

## Section G: Physical Functioning – Dexterity

Table G.1: Tests most frequently used

Number of experts rating test as „currently used”	Test name
1	Box And Block Test
0	Nine Hole Peg Test

## Section H: Physical Functioning – Muscle Strength

Table H.1: Tests most frequently used

Number of experts rating test as „currently used”	Test name
8	Chair Stand 5 Times
7	Grip Strength
5	Chair Stand 3 Times
3	Climbing Boxes
2	Chair Stand Once
1	Chair Stand 10 Times, Chair Stand 30 Seconds

Table H.2: Tests most frequently rated as “very common”

Number of experts rating test as „very common”	Test name
6	Chair Stand 5 Times
4	Grip Strength
3	Chair Stand 3 Times

Table H.3: Test Ratings

Test name	Number of experts rating general opinion of test as „very good”	Number of experts rating general opinion of test as „fairly good”
Chair Stand 5 Times	3	3
Grip Strength	3	0
Chair Stand 3 Times	2	1
Chair Stand 30 seconds	0	1

Test name	Number of experts rating general opinion of test as „very bad”	Number of experts rating general opinion of test as „rather bad”
Chair Stand 10 Times	0	1

## Section I: Physical Functioning – Overall Index Tests

Table I.1: Tests most frequently used

Number of experts rating test as „currently used”	Test name
7	Tinetti’s Performance-oriented Mobility Assessment
5	Short Physical Performance Battery
3	Functional Fitness, AAPHERD
2	Physical Performance Test, Physical Performance and Mobility Examination, Elderly Mobility Scale, Groningen Fitness Test
1	Physical Fitness Field Test, Clinical Outcome Variables, Nursing Home Disability Instrument, General Motor Function Assessment
0	Timed Functional Movements, Functional Fitness in Daily Functioning, Modified Elderly Mobility Scale

Table I.2: Tests most frequently rated as “very common”

Number of experts rating test as „very common”	Test name
6	Tinetti’s Performance-oriented Mobility Assessment
2	Short Physical Performance Battery
1	Physical Performance Test

Table I.3: Test Ratings

Test name	Number of experts rating general opinion of test as „very good”	Number of experts rating general opinion of test as „fairly good”
Tinetti’s Performance-oriented Mobility	2	4
Short Physical Performance Battery	1	4
Physical Fitness Field Tests, Physical Performance and Mobility Examination, Elderly Mobility Scale, General Motor Function Assessment, Clinical Outcome Variables, Physical Performance Test, Functional Fitness	0	1

## Section J: Physical Functioning – Activities of Daily Living

Table J.1: Tests most frequently used

Number of experts rating test as „currently used”	Test name
7	Lawton IADL, Barthel Index
6	Katz ADL
5	ADL
4	Functional Activities Questionnaire FAQ
3	Combined IADL
2	FIM
1	OARS-IADL, Nottingham ADL
0	IAM, Cleveland ADL, Bristol ADL, ADL Staircase

Table J.2: Tests most frequently rated as “very common”

Number of experts rating test as „very common“	Test name
5	Barthel Index
3	Lawton IADL
2	ADL
2	FIM
2	Katz ADL

Table J.3: Test Ratings

Test name	Number of experts rating general opinion of test as „very good“	Number of experts rating general opinion of test as „fairly good“
Barthel Index	3	3
Lawton IADL	3	2
ADL-IADL	1	0
ADL Index, Katz ADL	0	3
Functional Activities Questionnaire, FIM	0	1

Test name	Number of experts rating general opinion of test as „very bad“	Number of experts rating general opinion of test as „rather bad“
Lawton IADL*	0	1
ADL Index*	0	1

\* Test was also rated as being generally considered “fairly good” or “very good” by at least one respondent