

BACHELORARBEIT II

Titel der Bachelorarbeit

Inter- und Intratester-Reliabilität bei manuellen Zeitmessungen der Short Physical Performance Battery (SPPB)

Verfasserin

Maria Zamostny

angestrebter akademischer Grad

Bachelor of Science in Health Studies (BSc)

St. Pölten, 2019

Studiengang:

Studiengang Physiotherapie

Jahrgang

PT 16

Betreuer :

Dipl.-Sporting. Dr. Mario Heller

EHRENWÖRTLICHE ERKLÄRUNG

Ich erkläre, dass ich die vorliegende Bachelorarbeit selbstständig verfasst, andere als die angegebenen Quellen und Hilfsmittel nicht benutzt und mich auch sonst keiner unerlaubten Hilfe bedient habe.

Dieses Bachelorarbeitsthema habe ich bisher weder im In- noch im Ausland in irgendeiner Form als Prüfungsarbeit vorgelegt.

.....

Datum

.....

Unterschrift

I Zusammenfassung (Deutsch)

Inter- und Intratester-Reliabilität bei manuellen Zeitmessungen der Short Physical Performance Battery (SPPB)

Einleitung: Studien zufolge stürzen rund ein Drittel aller Menschen über 65 Jahre mindestens einmal pro Jahr. Wie groß mögliche Defizite sind und inwieweit sie den Menschen in seiner Selbständigkeit einschränken, misst die Short Physical Performance Battery (SPPB). Die SPPB ist eines der meistgenutzten Assessments zur Evaluierung der funktionellen Leistungsfähigkeit von Personen im fortgeschrittenen Alter. Generell müssen Assessments Voraussetzungen und Gütekriterien entsprechen. Die Hauptgütekriterien stellen die Validität, die Reliabilität und die Objektivität dar. Um bei einer Testbatterie eine hohe Validität zu erzielen, ist eine hohe Reliabilität erforderlich. Die Reliabilität gibt Auskunft darüber, wie gut sich die Ergebnisse eines Tests reproduzieren lassen. Die vorliegende Bachelorarbeit beschäftigt sich mit der Frage, ob unterschiedliche TesterInnen beziehungsweise Testwiederholungen durch einen TesterInnen Einfluss auf die Messergebnisse der SPPB ausüben und wie groß die Streuungen sind.

Methodik: Im Zuge dieser Studie wurden 21 ProbandInnen im Alter zwischen 60 und 85 Jahren bei der Durchführung der SPPB mit einer HDV-Kamera gefilmt. Aus den entstandenen Videos wurden mittels Zufallsgenerator für vier der fünf Subtests drei Videos ausgelost und fünf TesterInnen zur Bewertung durch Stoppuhren vorgespielt. Die erhobenen Daten wurden mit dem Statistikprogramm „SPSS Statistics“ analysiert und die Inter- und Intratester-Reliabilität über den ICC bestimmt.

Ergebnisse: Die Interrater-Korrelation ergab beim Chair Rising Test (CRT) sowie auch beim Tandemstand (TS) einen ICC von 0.999. Beim Gangtest (GT) lag dieser bei 0.972. Die Berechnungen der Intrarater-Korrelation ergab einen ICC von 0.998 für den CRT, 0.997 für den TS und 0.759 für den GT.

Schlussfolgerung: Wiederholungsmessungen gleicher TesterInnen führten nur beim Gangtest zu signifikanten Unterschieden in den Messergebnissen und daraus resultierend beim ICC. Bei Messungen verschiedener TesterInnen war dies bei keinem Subtest anhand des ICC zu sehen.

SCHLÜSSELWÖRTER: Geriatrie, ICC, Intertester-Reliabilität, Intratester-Reliabilität, Physiotherapie, Reliabilität, SPPB, Short Physical Performance Battery, manuelle Zeitmessung

I Abstract (English)

Inter- und intra-rater reliability of manual measurement of the short physical performance battery (SPPB)

Introduction: Based on recent studies every third person above 65 years of age falls at least once a year. The SPPB is one of the most widely used assessments to evaluate the functional performance of human of advanced age. In general assessments must meet prerequisites and quality criteria. The main quality criterias are validity, reliability and objectivity. To achieve a high validity in a test battery, a high reliability is required. Reliability tells us how well the results of a test can be reproduced. The present bachelor thesis deals with the question of whether different testers or test repetitions by one tester influence the measurement results of the SPPB and how large the scatters are.

Methods: In the course of this study 21 people between the ages of 60 and 85 were filmed with one HDV camera performing the SPPB. From the resulting videos three videos were randomly generated for four of the five subtests and then shown to five testers for evaluation by stopwatches. The collected data was analyzed with the statistics program "SPSS Statistics" and the inter- and intratester reliability were determined via the ICC.

Results: The Interrater-Correlation of the Chair Rising Test (CRT) and the Tandemstance (TS) was 0.999. The Correlation of the Walking Test was 0.972. The Results of the Interrater-Correlation were 0.998 for the CRT, 0.997 for the TS and 0.759 for the Walking Test.

Conclusio: Only for the Walking Test do Repeated Measures of the same Rater lead to significant Differences in the Results and the ICC. Measurements of different Raters show no significant Differences in the Results and the ICC for any Subtest of the SPPB.

KEYWORDS: Geriatrics, ICC, intertester reliability, intratester reliability, physiotherapy, reliability, SPPB, Short Physical Performance Battery, manual time measurement

II Inhaltsverzeichnis

1. EINLEITUNG	1
1.1 GÜTEKRITERIUM RELIABILITÄT	3
1.2 MANUELLE ZEITMESSUNGEN.....	7
1.3 SHORT PHYSICAL PERFORMANCE BATTERY (SPPB)	7
1.3.1 <i>Subtests</i>	8
1.3.2 <i>Anwendungsgebiet</i>	10
1.4 RELIABILITÄT DER SPPB	11
1.5 FRAGESTELLUNG UND HYPOTHESEN	12
2. METHODIK	13
2.1 STUDIENDESIGN / STUDIENABLAUF	13
2.2 INFORMATIONEN ZUR STUDIENTEILNEHMERINNENREKRUTIERUNG UND BESCHREIBUNG DER STUDIENTEILNEHMERINNENCHARAKTERISTIKA	14
2.4 ABLAUF DER MESSUNGEN	15
2.4.1 <i>Erstellung der Videos</i>	16
2.4.2 <i>Auswahl und Randomisierung</i>	18
2.4.3 <i>Bewertung durch TesterInnen</i>	18
2.5 MESSINSTRUMENTE	19
2.6 AUSWERTUNG DER DATEN.....	19
3. ERGEBNISSE	20
3.1 INTERRATER-RELIABILITÄTSANALYSE MIT SPSS.....	23
3.2 INTRARATER-RELIABILITÄTSANALYSE MIT SPSS.....	24
4. DISKUSSION	28
4.1 INTERPRETATION DER ERGEBNISSE.....	28
4.1.1 <i>ICC</i>	28
4.2 KLINISCHE RELEVANZ.....	31
4.3 LIMITATIONEN	32
5. SCHLUSSFOLGERUNGEN UND AUSBLICK	33
6. LITERATURVERZEICHNIS	35
A ANHANG – EINVERSTÄNDNISERKLÄRUNG	39

B	ANHANG – DATENBLATT	41
C	ANHANG – ANLEITUNG BALANCETESTUNGEN.....	42
D	ANHANG – ANLEITUNG CHAIR RISING TEST	44
E	ANHANG – ANLEITUNG GANGTESTUNG	45
F	ANHANG – SPPB PROTOCOL AND SCORE SHEET.....	46

III **Abbildungsverzeichnis**

ABBILDUNG 1 BEVÖLKERUNG NACH BREITEN ALTERSGRUPPEN 1950 BIS 2080, MITTLERE VARIANTE (STATISTIK AUSTRIA, 2017)	2
ABBILDUNG 2 BEISPIEL FÜR EINE TESTUNG MIT NIEDRIGEM ICC	6
ABBILDUNG 3 BEISPIEL FÜR EINE TESTUNG MIT HOHEM ICC	6
ABBILDUNG 4 ABLAUFPLAN DER ARBEITSSCHRITTE.....	14
ABBILDUNG 5 SETTING AUFSTEHTEST VON OBEN	16
ABBILDUNG 6 SETTING GLEICHGEWICHTSTESTS VON OBEN	17
ABBILDUNG 7 SETTING GANGTESTUNG VON OBEN (LINKS) UND SKIZZE EINES HOLZPFLOCKS VON DER SEITE (RECHTS)	18
ABBILDUNG 8 DARSTELLUNG DER ERGEBNISSE DES CRT DER ID 1 VON VIDEO 1	21
ABBILDUNG 9 INTRARATER-KORRELATIONSKOEFFIZIENTEN DES CRT VON ID 1 BIS ID 5.....	25
ABBILDUNG 10 INTRARATER-KORRELATIONSKOEFFIZIENTEN DES TS VON ID 1 BIS ID 5	26
ABBILDUNG 11 INTRARATER-KORRELATIONSKOEFFIZIENTEN DES GT VON ID 1 BIS ID 5.....	26
ABBILDUNG 12 ERRECHNETE INTRARATER-KORRELATIONSKOEFFIZIENTEN DER SUBTESTS (BEINHÄLTET ALLE IDS)	27

IV TABELLENVERZEICHNIS

TABELLE 1 AUSWERTUNG DER BALANCETESTUNG DER SPPB NACH GURALNIK ET AL. (1994) .	9
TABELLE 2 AUSWERTUNG DER GANGTESTUNG DER SPPB NACH GURALNIK ET AL. (1994)	10
TABELLE 3 AUSWERTUNG DES AUFSTEHTEST DER SPPB NACH GURALNIK ET AL. (1994).....	10
TABELLE 4 ARTEN UND AUSWAHL DES ICC NACH SHROUT & FLEISS (1979) BZW. MCGRAW & WONG (1996).....	20
TABELLE 5 URLISTE DER ROHDATEN DER TESTERINNEN [s]	22
TABELLE 6 MITTELWERTE UND STANDARDABWEICHUNG DER MESSERGEBNISSE [s].....	23
TABELLE 7 INTERRATER-KORRELATIONSKOEFFIZIENT DES CRT	24
TABELLE 8 INTERRATER-KORRELATIONSKOEFFIZIENT DES GT.....	24
TABELLE 9 INTERRATER-KORRELATIONSKOEFFIZIENT DES TS	24
TABELLE 10 INTRARATER-KORRELATIONSKOEFFIZIENTEN ALLER IDS	25

V Abkürzungsverzeichnis

∅	Durchschnitt
>	Größer-als
≥	Größer-gleich
<	Kleiner-als
≤	Kleiner-gleich
%	Prozent
σ_{zwischen}	Varianz zwischen den Beobachtungen
$\sigma_{\text{innerhalb}}$	Varianz innerhalb der Beobachtungen
6MWT	Six-minute walking test
ATL	Aktivitäten des täglichen Lebens
CRT	Chair Rising Test
GT	Gangtest
HDV	High Definition Video
ICC	Intraclass-Correlations-Coeffizient
k	Anzahl der BeobachterInnen
m	Meter
Mini-BEST	Mini-Balance Evaluation Systems Test
sek	Sekunde
SD	Standardabweichungen
SPPB	Short Physical Performance Battery
STS	Semitandemstand
TS	Tandemstand

Vorwort:

Die vorliegende Bachelorarbeit zum Thema „Inter- und Intratester-Reliabilität bei manuellen Zeitmessungen der Short Physical Performance Battery (SPPB)“ entstand im Zuge der Ausbildung zur Physiotherapeutin an der Fachhochschule St. Pölten. Ziel war es herauszufinden, ob es bei manuellen Zeitmessungen der Subtests der SPPB bei verschiedenen TesterInnen beziehungsweise bei Testwiederholungen Unterschiede gibt und wie groß die Streuungen sind.

Unter der Betreuung von Herrn Dipl.-Sporting. Dr. Mario Heller gelang es mir im Rahmen meiner Bachelorarbeit, eine klare Struktur zu schaffen und den gesamten Prozess über, die Übersicht zu behalten. Bei allfälligen Fragen stand er mir stets mit Tipps zur Seite. Somit hat er einen erheblichen Beitrag zum Gelingen dieser Bachelorarbeit geleistet, herzlichen Dank hierfür!

Ein besonderer Dank gilt ebenso meinen Eltern, die mir mein Studium ermöglichen und mich in all meinen Entscheidungen unterstützt haben. Sie zeigten stets großes Interesse und fanden bei Problemen und Rückschlägen aufbauende Worte. Weiters waren Sie mir bei der ProbandInnensuche äußerst behilflich.

Von Herzen möchte ich auch meinem Freund danken, der mich immer wieder auf meinem Weg bestärkte und mir mit vielen hilfreichen Tipps zur Seite stand. Auch beim Korrekturlesen war er eine große Hilfe!

Schließlich danke ich auch meinen Mitstudierenden, allen ProbandInnen und TesterInnen, durch deren Unterstützung die Durchführung der Studie möglich gewesen wäre.

St. Pölten am 29.01.2019

Maria Zamostny

1. Einleitung

Mit dem Älterwerden steigt das Risiko zu stürzen und den Alltag nicht mehr alleine bewältigen zu können (Rodrigues, Fraga, & Barros, 2014). Laut Woolf & Akesson (2003) stürzt rund ein Drittel aller Menschen über 65 Jahre mindestens einmal pro Jahr. Im häuslichen Umfeld zählen Stürze bei Menschen im fortgeschrittenen Alter zu den häufigsten und schwerwiegendsten Unfällen. Sie sind ein Hauptgrund für eine Einweisung ins Krankenhaus oder in ein Pflegeheim (Horne, Skelton, Speed, & Todd, 2013). Jedes zu Sturz kommen erhöht die Mortalität (das Verhältnis der Zahl der Todesfälle zur Zahl der statistisch berücksichtigten Personen) sowie die Morbidität (die Häufigkeit der Erkrankungen innerhalb einer Bevölkerungsgruppe) und stellt eine Belastung für die gestürzte Person, deren familiäres Umfeld sowie das Sozialsystem dar (Horne et al., 2013).

Im Buch „Gerontologie, Geriatrie und Gerontopsychiatrie: Ein Lehrbuch für Gesundheits- und Pflegeberufe“ (Steidl & Nigg, 2014, S.14) nennt der Wiener Soziologe Leopold Rosenmayr folgende, auch von der Weltgesundheitsorganisation verwendete, Einteilung für Alter:

- Ältere Menschen (61-74 Jahre)
- Alte Menschen (75-90 Jahre)
- Hochbetagte Menschen (91-100 Jahre)
- Langlebige Menschen (über 100 Jahre)

STATISTIK AUSTRIA 2017 zufolge waren 2016 18,5 % der österreichischen Bevölkerung (1,62 Millionen Menschen) 65 Jahre und älter. Im Jahr 2030 wird ihre Zahl um weitere 33 % auf 2,15 Millionen und bis zum Jahr 2050 um weitere 64 % auf 2,64 Millionen Menschen ansteigen. Demnach würden im Jahr 2080 2,89 Millionen Personen der Generation 65+ in Österreich leben, um 79 % mehr als heute (siehe Abbildung 1).

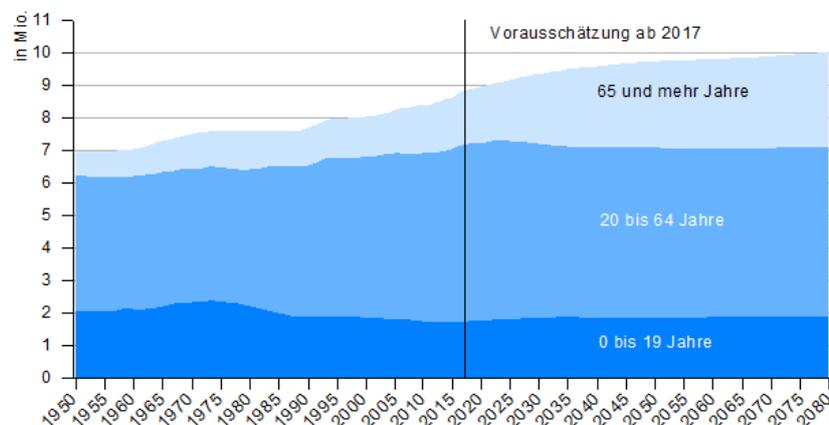


Abbildung 1 Bevölkerung nach breiten Altersgruppen 1950 bis 2080, mittlere Variante (Statistik Austria, 2017)

In Anbetracht der steigenden Lebenserwartung wird der Prävention von Stürzen eine hohe Wichtigkeit zugesprochen (Horne et al., 2013). Neben dem Sturzrisiko steigt mit dem Alter auch die Vulnerabilität gegenüber Begleiterkrankungen und Medikamentennebenwirkungen ebenso wie manifeste Syndrome mit Hypomobilität, Inkontinenz oder Gebrechlichkeit. Stürze können extrinsischer, synkopaler oder lokomotorischer Ursache sein. Ein Beispiel für einen extrinsischen Sturz wäre das Ausrutschen auf einer Bananenschale, ein von außen einwirkender Aspekt bringt den Menschen zu Fall. Synkopale Stürze sind auf definierbare Erkrankungen, wie zum Beispiel eine Herz-Rhythmusstörung oder einen epileptischer Anfall, zurückzuführen. Der Großteil an Stürzen im fortgeschrittenen Alter sind jedoch lokomotorische Stürze und stets multifaktorieller Genese. Beim älteren Menschen treffen hier reduzierte Funktion- und Leistungsfähigkeit diverser Organsysteme mit Faktoren, die endogen (im Körper inneren entstehend), exogen (von außen auf den Körper einwirkend), iatrogen (durch ärztliche Einwirkung) und situativ (die jeweilige Situation betreffend) sein können, zusammen und enden in einem Sturzereignis (Bernecker, 2010). Wie groß mögliche Defizite sind und inwieweit sie den Menschen in seiner Selbständigkeit einschränken, misst die Short Physical Performance Battery (SPPB). Die SPPB ist eines der meistgenutzten Assessments zur Evaluierung der funktionellen Leistungsfähigkeit der unteren Extremität von Personen im fortgeschrittenen Alter (Freire, Guerra, Alvarado, Guralnik, & Zunzunegui, 2012). Eine große Sorge von älteren Menschen ist es, die Aktivitäten des täglichen Lebens (ATL) nicht mehr eigenständig durchführen zu können (Roman de Mettelinge & Cambier, 2015). Für Menschen die bisher keine körperlichen Einschränkungen in ihren ATL aufwiesen, kann der Beginn einer Behinderung mittels SPPB vorausgesagt werden (Guralnik et al., 2000). Assessments, wie auch die SPPB müssen unter anderem reproduzierbar sein, um eine hohe Gültigkeit der Messergebnisse gewährleisten zu können (Oesch et al., 2011, S.27).

Die vorliegende Bachelorarbeit beschäftigt sich mit der Frage, ob unterschiedliche TesterInnen beziehungsweise Testwiederholungen durch eineN TesterIn Einfluss auf die Messergebnisse der SPPB ausüben und wie groß die Streuungen sind. Als theoretische Grundlage sind in den folgenden Kapiteln die Begriffsdefinitionen des Gütekriteriums Reliabilität, der manuellen Zeitmessung, der SPPB sowie der Reliabilität der SPPB dargestellt.

1.1 Gütekriterium Reliabilität

Generell müssen Assessments Voraussetzungen und Gütekriterien entsprechen. Die Hauptgütekriterien stellen die Validität (Gültigkeit), die Reliabilität (Zuverlässigkeit) und die Objektivität (Unabhängigkeit) dar (Flick, Kardoff, & Steinke, 2004, S.184).

Weitere Gütekriterien für ein gutes Assessment sind die Praktikabilität (Anwendbarkeit), die Responsivität (Empfindlichkeit), die Spezifität (Identifikation von Gesunden) und die Sensitivität (Identifikation von Kranken).

Um bei einer Testbatterie eine hohe Validität zu erzielen, ist eine hohe Reliabilität erforderlich (Mai, Greiff, & Weidemann-Wendt, 2017, S.106). Die Reliabilität gibt Auskunft darüber, wie gut sich die Ergebnisse eines Tests reproduzieren lassen. Synonyme dafür sind Begriffe wie Reproduzierbarkeit und Zuverlässigkeit. Messverfahren werden als reliabel angesehen, wenn Messwiederholungen (Test-Retest) unter denselben Bedingungen dasselbe Ergebnis liefern. Die Test-Retest Reliabilität wird mittels Korrelationskoeffizienten angegeben (Oesch et al., 2011, S.27). Korrelation und Übereinstimmung sind weit verbreitete Konzepte, die den Zusammenhang zwischen Variablen bewerten. Beide werden verwendet, um die Stärke der Assoziation zwischen Variablen anzuzeigen, sind jedoch konzeptionell verschieden und erfordern daher die Verwendung verschiedener Statistiken.

Zur Berechnung von Korrelationen stehen unter anderem folgende Verfahren zur Auswahl:

- Der parametrische Pearsons Korrelationskoeffizient
- Der nicht-parametrische Spearmans Rang-Korrelationskoeffizient
- Der parametrische Intraklassen-Korrelationskoeffizient
(Englisch: Intraclass-Correlations-Coeffizient; Abkürzung: ICC)

Der Korrelationskoeffizient gibt Aufschluss über die Stärke des Zusammenhangs zweier Testresultate und teilt mit, wie nahe Messpaare (Test-Retest) bei grafischen Darstellungen auf einer Linie liegen. Bei Pearsons Korrelationskoeffizient und Spearmans Rang-Korrelationskoeffizient kann dieser Werte von -1.00 (perfekte negative Korrelation) bis +1.00 (perfekte positive Korrelation) annehmen (Liu et al., 2016). Beim ICC hingegen kann der Korrelationswert nur positive Werte zwischen 0.00 und 1.00 annehmen. Bei einer optimalen Korrelation von 1.00 würden sich alle Messpaare auf der Geraden befinden (Oesch et al., 2011). Shrout & Fleiss (1979) geben keine absoluten Grenzwerte für den ICC in ihrer Studie an.

Ihrer Meinung nach sind die Anforderungen an den ICC von der jeweiligen Fragestellung abhängig. Das „Scientific Advisory Committee of Medical Outcomes Trust“ hat 2002 hingegen Richtlinien formuliert. Ein ICC von >0.7 wird als genügend erachtet. Bei den Verlaufsmessungen von einzelnen PatientInnen wird ein ICC von >0.9 empfohlen (Oesch et al., 2011).

Pearsons und Spearmans messen den Zusammenhang zwischen zwei Variablen, nicht aber deren Übereinstimmung. Pearsons setzt intervallskalierte Daten sowie einen linearen Zusammenhang dieser voraus. Er ist empfindlich gegenüber Ausreißern.

Spearmans setzt keinen linearen Zusammenhang der Daten voraus und ist robust gegenüber Ausreißern. Im Gegensatz zu Pearsons wird die Korrelation bei Spearmans nicht zwischen den Datenpunkten selbst, sondern zwischen ihren Rängen berechnet. Der ICC setzt wie Pearsons intervallskalierte Daten voraus, kann jedoch auch für mehr als zwei BeobachterInnen oder mehrere Beobachtungszeitpunkte herangezogen werden. Weiters misst er sowohl die Korrelation als auch die Übereinstimmung zwischen Messwerten (Liu et al., 2016). Aus obengenannten Gründen wird der ICC zur Datenauswertung vorliegender Arbeit herangezogen.

Dem ICC liegt ein varianzanalytischer Ansatz zu Grunde. Dieser lässt sich in die Varianz zwischen den Beobachtungen σ_{zwischen} (entsteht durch systematische Merkmalsvariation) und die Varianz innerhalb der Beobachtungen $\sigma_{\text{innerhalb}}$ (entsteht durch mangelhafte BeobachterInnenübereinstimmung) unterteilen. Gebildet wird der ICC aus dem Quotienten aus der Differenz von σ_{zwischen} und $\sigma_{\text{innerhalb}}$ der Beobachtungsfälle (systematische Varianz) und der Summe von σ_{zwischen} und der gewichteten $\sigma_{\text{innerhalb}}$ der Beobachtungsgefälle (Gesamtvarianz).

$$\text{ICC} = \frac{\text{systematische Varianz}}{\text{Gesamtvarianz}}$$

$$\text{ICC} = \frac{\sigma_{\text{zwischen}} - \sigma_{\text{innerhalb}}}{\sigma_{\text{zwischen}} + (k - 1) \times \sigma_{\text{innerhalb}}}$$

k = Anzahl der BeobachterInnen

σ_{zwischen} = Varianz zwischen den Beobachtungen

$\sigma_{\text{innerhalb}}$ = Varianz innerhalb der Beobachtungen

(Vincent & Weir, 2012, S.215)

Zur Veranschaulichung des ICC dienen Abbildung 1 und 2. Abbildung 1 zeigt eine Messung mit niedrigem ICC. Die TesterInnen kamen zu sehr unterschiedlichen Ergebnissen, die Messwerte zeigen starke Streuungen und liegen nicht nahe einer Geraden. Der ICC ist mit 0.48 gering. Abbildung 2 zeigt ein Assessment mit hohem ICC.

Die TesterInnen kamen zu sehr ähnlichen Ergebnissen, die Messwerte zeigen geringe Streuungen und liegen nahe einer Geraden. Der ICC ist mit 0.98 hoch.

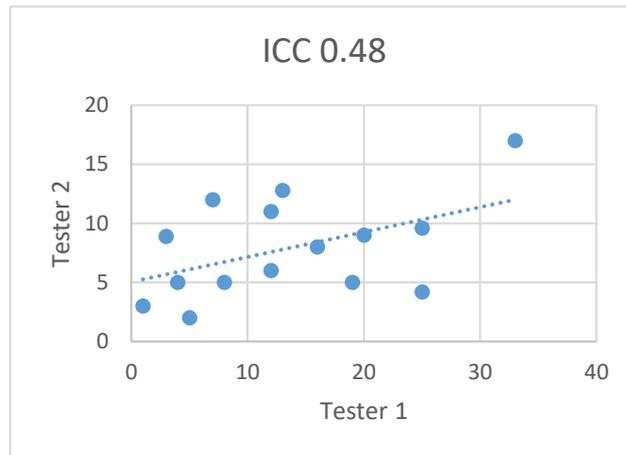


Abbildung 2 Beispiel für eine Testung mit niedrigem ICC

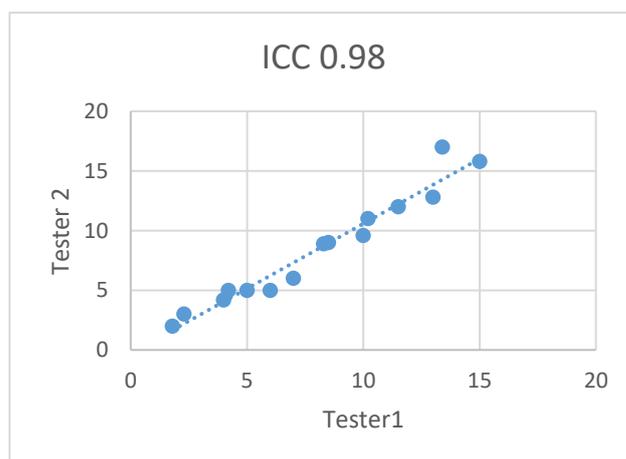


Abbildung 3 Beispiel für eine Testung mit hohem ICC

Bei der Reliabilität wird zwischen Inter- und Intratester-Reliabilität unterschieden. Als Intertester-Reliabilität wird das Ausmaß an Übereinstimmungen (=Konkordanzen) bei Testungen durch mindestens zwei UntersucherInnen (Rater) bezeichnet. Zeigt ein Test im Alltag eine hohe Intertester-Reliabilität, so kann man davon ausgehen, dass zwei TherapeutInnen denselben Test gleich bewerten.

Ist diese niedrig, so würden im Umkehrschluss zwei TherapeutInnen einen Test unterschiedlich bewerten und auch zu unterschiedlichen Ergebnissen kommen. Im therapeutischen Alltag wird der Reliabilität eines Assessments somit hohe Relevanz zugesprochen (Luomajoki & Pfeiffer, 2018, S.8). Als Intratester-Reliabilität wird die Übereinstimmung der Messergebnisse eines einzelnen Raters bezeichnet.

Zeigen die Messungen von nur einer Person in derselben Situation dargestellt dieselben Ergebnisse, so wird von einer hohen Intratester-Reliabilität ausgegangen. Es ist naheliegend anzunehmen, dass ein und derselbe Mensch einen Test unter gleichen Bedingungen gleich beurteilen würde.

Die Studie „Intrarater reliability of dual-energy X-Ray absorptiometry-based measures of vertebral height in postmenopausal women“, welche von Intrarater-Reliabilität bei Knochendichtemessung mittels Röntgenuntersuchung handelt, ergab jedoch Korrelationswerte zwischen 0.15 und 0.90. Es ist somit nachvollziehbar, dass ForscherInnen die Reliabilität der Datenerhebung ihrer Messergebnisse kritisch hinterfragen (McHugh, 2012).

1.2 Manuelle Zeitmessungen

In der Physiotherapie wird Zeitmessung als reliables, praktisches und aussagekräftiges Messinstrument und therapeutisches Hilfsmittel eingesetzt (Barth et al., S.86, 2005). Die Erhebung der Ganggeschwindigkeit bei geriatrischen Menschen wird als klinisch relevanter Parameter zur Erhebung der individuellen Gesundheit, des Wohlbefindens und des funktionellen Status angesehen. Elektronische Messmethoden, wie der Einsatz von Beschleunigungssensoren, sind potenziell reliabler, valider und höher in ihrer Präzision. Dennoch hat sich die manuelle Zeitmessung (Stoppuhr-Messung) als Goldstandard in physiotherapeutischen Assessments durchgesetzt. Stoppuhr-Messungen sind praktikabel, einfach in ihrer Handhabung, schnell durchführbar, reproduzierbar, kostengünstig und müssen nicht zwingend von Fachpersonal durchgeführt werden (Maggio et al., 2016). Fraglich ist jedoch, wie sich der Einfluss von Inter- und IntratesterInnen auf die Messergebnisse auswirkt.

1.3 Short Physical Performance Battery (SPPB)

Der körperlichen Funktionsfähigkeit wird im Alter eine bedeutende Rolle als Biomarker für die allgemeine Gesundheit eines Menschen zugesprochen. Zahlreiche standardisierte Tests ermöglichen durch festgelegte Kriterien, wie unter anderem Zeitmessung oder Wiederholungsanzahl, eine objektive Bewertung. Die frühzeitige Erkennung von Funktionsverlust ist wichtig, um Interventionen setzen und einen Leistungsabfall revidieren oder verhindern zu können (Olsen & Bergland, 2017).

Im Systematic Review von Freiburger et al. (2012) wurden zwölf unterschiedliche Assessments für die körperliche Aktivität bei älteren Menschen auf Validität, Reliabilität und Änderungssensibilität geprüft. Am besten schnitt die Short Physical Performance Battery (SPPB) gefolgt vom Physical Performance Test und der Continuous Scale Physical Functional Performance ab.

Die Short Physical Performance Battery (SPPB) ist eines der meistgenutzten Assessments zur Evaluierung der funktionellen Leistungsfähigkeit von Personen im fortgeschrittenen Alter (Freire et al., 2012). Sie dient als Funktionstest zur Messung von Ganggeschwindigkeit, Gleichgewicht im Stehen sowie Kraftausdauer der unteren Extremität (Stookey, Katzel, Steinbrenner, Shaughnessy, & Ivey, 2014). Guralnik et al. entwickelte die SPPB im Jahr 1994 für eine Kohortenstudie, um Testung im Zuge eines Hausbesuches und somit unter Platzmangel zu ermöglichen.

In seiner Studie wird beschrieben, dass diese Testbatterie ein effizienter, praktischer und sicherer Weg ist, um die Funktionalität der unteren Extremität zu prüfen (Maggio et al., 2016). Mittels SPPB lässt sich funktionelle Leistungsfähigkeit hierarchisch von hohen Funktionsniveaus hin zu Verschlechterungen der Funktion des Bewegungsapparates gliedern. Die daraus resultierenden Ergebnisse werden in engem Zusammenhang mit Funktionseinschränkungen sowie Einschränkungen der ATL gebracht. Höhere Ergebnisse lassen auf eine gute Funktion der unteren Extremität schließen. Niedrige Ergebnisse hingegen können Aufschluss über Mobilitätsverlust, Behinderung, Länge des Krankenhausaufenthaltes, Langzeitversorgung durch ein Pflegewohnheim und auch Tod eines/einer PatientIn liefern (Freire et al., 2012).

1.3.1 Subtests

In der Literatur ist die Durchführung der Subtests der SPPB unterschiedlich beschrieben. Guralnik et al. (1994) endet den Aufstehetest im Stehen, im Fachartikel von Büsching (2015, S.42) wird er hingegen mit dem letzten Hinsetzen beendet. Ebenso zeigen sich Unterschiede bei der Ausübung des Semitandemstands. Guralnik et al. (1994) lässt den/die ProbandIn die Ferse des vorderen Fußes seitlich an den ersten Zeh stellen, Büsching (2015, S.42) hingegen gibt an, dass ein Fuß mit der Ferse an die Mitte der Innenseite des anderen Fußes gestellt wird. Dieser Arbeit wird die Testung der Short Physical Performance Battery nach Guralnik et al. (1994) zu Grunde gelegt, da es sich bei den AutorInnen um Entwicklerteam des Assessments handelt. Die SPPB besteht aus drei Balancetestungen, einer Gangtestung und einem Aufstehetest (Chair Rising Test). Die Balancetestungen sind in Semitandemstand, Tandemstand und Rombergstand untergliedert. Jede Position wird zuerst demonstriert, dann erhält die Testperson Hilfestellung an einer Hand, um die Position einzunehmen und versucht diese über zehn Sekunden zu halten. Begonnen wird mit dem Semitandemstand.

Bei diesem wird die Ferse des vorderen Fußes seitlich an den ersten Zeh des hinteren Fußes gesetzt. Kann diese Position nicht durchgeführt werden, wird der Rombergstand (zusammenstehende Füße) getestet. Ist der Semitandemstand möglich, wird darauffolgend der Tandemstand (Füße stehen hintereinander) untersucht.

Die Testungen enden, wenn die Zeit endet, die Testperson den Fuß von der Unterstützungsfläche wegbewegt oder nach Hilfe greift. Bei der Gangtestung wurde die Gehgeschwindigkeit ursprünglich auf einer Gehstrecke von 2,44 Metern (acht feet) gemessen. Auf der Homepage des „National Institute on Aging“ ist eine Beschreibung der Auswertungsmöglichkeiten der Short Physical Performance Battery (SPPB) von Guralnik für eine drei oder vier Meter lange Strecke zu finden. Bei dieser Testung werden zwei aufeinanderfolgende Messungen durchgeführt, wobei die schnellere Durchführung gewertet wird. Gestartet wird nach Kommando an der Startlinie, beendet wird die Messung, wenn ein Fuß zur Gänze die Ziellinie übertreten hat. Die Testperson soll dabei in einem für sie alltagsüblichen Gangtempo gehen, Hilfsmittel sind erlaubt.

Für den Aufstehetest wird die Testperson gebeten, die Arme vor der Brust zu verschränken und einmal von einem Sessel aufzustehen. Gelingt dies, so kommt es zur eigentlichen Testung, bei der die Testperson gebeten wird fünfmal hintereinander so schnell wie möglich aufzustehen und sich hinzusetzen. Die Zeitmessung fängt mit der ersten sitzenden Position an und endet mit der letzten Standposition. Bei Sicherheitsbedenken oder längerdauernder Durchführung als eine Minute wird der Test abgebrochen. Die Auswertung wird anhand einer Punktetabelle durchgeführt (Tabelle 1-3, S.8). Bei jedem Subtest kann der/die ProbandIn bis zu vier Punkte erreichen. Die höchste zu erreichende Gesamtwertung sind zwölf Punkte. Hilfsmittel beziehungsweise Testabbruch sind zu notieren. EinE PatientIn mit einem Endergebnis von null bis drei Punkten ist im Alltag beim Gehen von einigen hundert Metern, beim Treppensteigen und bei der Selbstversorgung stark beeinträchtigt. Vier bis sechs Punkte gelten als moderate und sieben bis neun Punkte als leichte Einschränkung. Menschen mit einer Punkteanzahl von zehn bis zwölf Punkten gelten als minimal bis gar nicht im Alltag beeinträchtigt.

Tabelle 1 Auswertung der Balancetestung der SPPB nach Guralnik et al. (1994)

Punkte	Rombergstand	Semitandemstand	Tandemstand
0	0-9 Sekunden oder Versuch nicht möglich oder nicht versucht	keine weitere Balancetestung	keine weitere Balancetestung
1	10 sek	0-9 sek oder Versuch nicht möglich oder nicht versucht	keine weitere Balancetestung
2	10 sek	10 sek	< 3 sek oder Versuch nicht möglich oder nicht versucht
3	10 sek	10 sek	3-9,99 sek
4	10 sek	10 sek	10 sek

Tabelle 2 Auswertung der Gangtestung der SPPB nach Guralnik et al. (1994)

Punkte	4 Meter	3 Meter
0	nicht möglich	nicht möglich
1	> 8,70 sek	>6,52 sek
2	6,21-8,70 sek	4,66-6,52 sek
3	4,82-6,20 sek	3,62-4,65 sek
4	< 4,82 sek	< 3,62 sek

Tabelle 3 Auswertung des Aufstehetest der SPPB nach Guralnik et al. (1994)

Punkte	
0	nicht möglich
1	≥ 16,70 sek
2	13,70-16,69 sek
3	11,20-13,69 sek
4	≤ 11,19 sek

1.3.2 Anwendungsgebiet

Latham et al. (2008) bestätigen in ihrer prospektiven Beobachtungsstudie, dass sich die Short Physical Performance Battery als Evaluierungsparameter eignet, um in der Therapie Veränderungen nach Interventionen abzubilden. Durch die Anwendung der SPPB konnte im Zuge einer randomisierten klinischen Studie bei PatientInnen mit Schenkelhalsfraktur sechs Monate nach Therapiebeginn die Funktionalität der unteren Extremität evaluiert werden (Latham et al., 2014).

Ebenso hilfreich zeigte sich die SPPB in der Abbildung von Veränderungen bei Menschen mit:

- Myokardinfarkt (Herzinfarkt) (Ostir et al., 2002)
- Myokardinsuffizienz (Herzschwäche) (Ostir et al., 2002)
- Insult (Schlaganfall) (Ostir et al., 2002)
- Sarkopenie (Muskelschwund) (Liu et al., 2014)
- Claudicatio intermittens (Schaufensterkrankheit) (Mockford et al., 2014)
- Chronic obstructive pulmonary disease (Chronisch obstruktive Lungenerkrankung) (Patel et al., 2014)

Puthoff (2008) berichtet, dass der Einsatz der SPPB ein nützliches Hilfsmittel für die Behandlung von PatientInnen mit zentraler Venendruck - Erkrankung darstellt. Jedoch wird in seiner Studie auch beschrieben, dass die Anwendung dieses Assessments für aktive Menschen im fortgeschrittenen Alter wenig Aussagekraft hat, da sich schnell ein Deckeneffekt einstellt. Die SPPB liefert zwar eine Prognose darüber, wie groß die Wahrscheinlichkeit ist, dass ein Mensch in den kommenden vier Jahren verstirbt, Umkehrschluss ist jedoch keiner möglich. Ein hohes Testergebnis gibt keinen Aufschluss darüber, ob die Testperson generell gesund ist. Büsching (2015) schreibt in seinem Fachartikel, dass die Anwendung eines 400-Meter-Gehtest bei geriatrischen Menschen in gutem körperlichen Zustand zur Evaluierung der Funktionalität der unteren Extremität besser geeignet ist.

1.4 Reliabilität der SPPB

Freire et al. (2012) beschreibt die SPPB als objektives Maß zur Beurteilung körperlicher Leistungsfähigkeit, welches wenig von Kultur, Bildung und Sprache beeinflusst wird.

Ebenso erklären Gómez, Curcio, Alvarado, Zunzunegui, & Guralnik (2013), dass die SPPB ein reliables und valides Testinstrument ist, um die Funktionalität von älteren Personen zu überprüfen. Test-Retest Reliabilität konnte in ihrer Pilotstudie einen Korrelationskoeffizienten von 0.87 erreichen. Ebenso guten ICC erlangte die SPPB in der Studie von Ostir et al. (2002). Dieser lag hier zwischen 0.88 bis 0.92. Auch zeigten die Subtests in ihren Einzelwertungen in dieser Studie gute Ergebnisse. In anderen Studien konnte der Gangtest 0.79 (Jette, Jette, Ng, Plotkin, & Bach, 1999) und der Chair Rising Test 0.80 erreichen. Beim Tandemstand hingegen lag dieser bei nur 0.22 (Curb et al., 2006).

Im Zuge dieser Arbeit gilt es herauszufinden, warum der Tandemstand vergleichsweise schlechte Reliabilität aufweist und wie sich diese bei mehreren TesterInnen und bei wiederholten Messungen der Testperson verändert.

1.5 Fragestellung und Hypothesen

Gibt es Unterschiede bei manuellen Zeitmessungen der Subtests der SPPB bei verschiedenen TesterInnen beziehungsweise bei Testwiederholungen und wie groß sind die Streuungen?

Daraus ergeben sich folgende Hypothesen:

H0¹: Es gibt keine signifikanten Unterschiede bei manuellen Zeitmessungen der Subtests der SPPB bei verschiedenen TesterInnen.

H1¹: Es gibt signifikante Unterschiede bei manuellen Zeitmessungen der Subtests der SPPB bei verschiedenen TesterInnen.

H0²: Es gibt keine signifikanten Unterschiede bei manuellen Zeitmessungen der Subtests der SPPB bei Wiederholungsmessungen gleicher TesterInnen.

H1²: Es gibt signifikante Unterschiede bei manuellen Zeitmessungen der Subtests der SPPB bei Wiederholungsmessungen gleicher TesterInnen.

2. Methodik

Im folgenden Kapitel wird die Methodik dieser Bachelorarbeit dargestellt. Diese beinhaltet das Studiendesign, die ProbandInnen- und TesterInnenrekrutierung, die Ein- und Ausschlusskriterien, den Ablauf der Messungen, die Messinstrumente sowie die Auswertung der Daten.

2.1 Studiendesign / Studienablauf

Bei vorliegender Bachelorarbeit handelt es sich um eine Querschnittstudie. Die StudienteilnehmerInnen bestanden aus ProbandInnen und TesterInnen, welche in Kapitel 2.2 näher erläutert werden.

Im Zuge dieser Studie wurden ProbandInnen bei der Durchführung der SPPB gefilmt und die daraus entstandenen Videos zur Bewertung mittels digitaler Handstoppuhr TesterInnen vorgespielt.

Die Zusammentragung und Verschriftlichung der theoretischen Grundlage dieser Studie wurde in den Monaten April, Mai und Juni 2018 vollzogen. Die Planung der Durchführung fand im Juli und die ProbandInnen- und TesterInnenrekrutierung im August 2018 statt. Ein Übungstag wurde am 20. Oktober 2018 im Laborraum des Bundes- und Innovationszentrum St. Pölten abgehalten. Etwaige Fehlerquellen wurden analysiert und beseitigt. Am 17. und 18. Oktober 2018 wurden im Zuge der Testungen der Bachelorarbeit „Einfluss von Schuhwerk auf die Ergebnisse der Short Physical Performance Battery im Vergleich zur Barfußstestung“ von Arkadiusz Wojciak 21 ProbandInnen bei der Ausführung der SPPB mit einer Canon Legria HV40 Kamera (HDV-Kamera) gefilmt. Am ersten Testtag wurden 15 Personen und am zweiten Testtag sechs Personen getestet.

Aus den daraus entstandenen Videos wurden Ende Oktober für vier der fünf Subtests je drei Videos mittels Research Randomizer¹ ausgewählt. Diese wurden in randomisierter Reihenfolge fünf PhysiotherapeutInnen an zwei Testtagen Mitte Dezember zur Bewertung vorgelegt. Die Bewertung fand unter der Verwendung einer handelsüblichen, digitalen Stoppuhr statt. Die Auswertung und Interpretation der daraus resultierenden Daten sowie das Verfassen der Bachelorarbeit II wurde im Dezember 2018 und Jänner 2019 durchgeführt. Grafisch dargestellt ist dieser Ablaufplan der Arbeitsschritte in Abbildung 4.

¹ <https://www.randomizer.org>

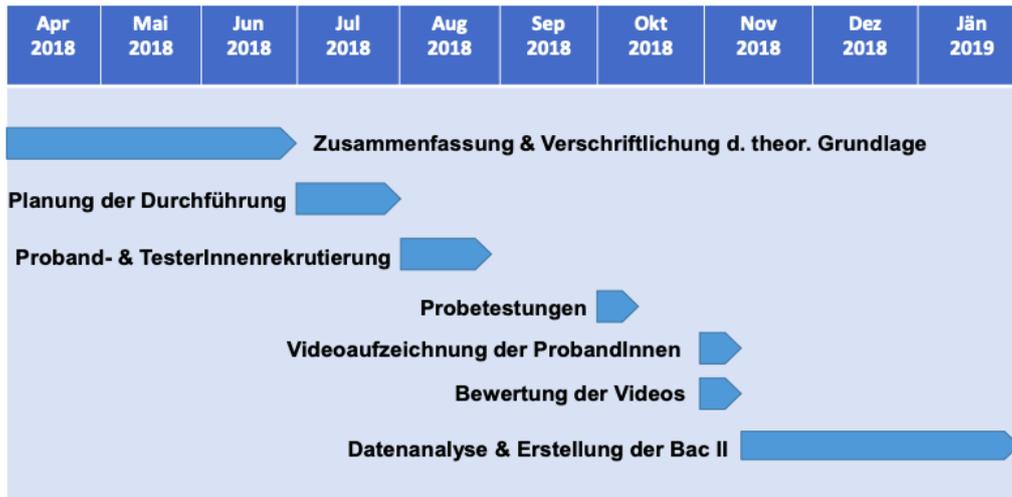


Abbildung 4 Ablaufplan der Arbeitsschritte

2.2 Informationen zur StudienteilnehmerInnenrekrutierung und Beschreibung der StudienteilnehmerInnencharakteristika

Die StudienteilnehmerInnen unterteilten sich in ProbandInnen und TesterInnen. Die ProbandInnenanzahl von 17 Personen resultierte aus der TeilnehmerInnenanzahl der Studie von Arkadiusz Wojciak, welche mittels Fallzahlschätzung mit dem Programm G*Power 3.1.9.2 berechnet wurde. Die tatsächliche ProbandInnenanzahl am Testtag lag bei 21 Personen. Für die Bewertung der Messergebnisse wurden fünf TesterInnen herangezogen.

Beide Gruppen wurden im Bekannten- und Freundeskreis über die Studie, deren Ablauf und Ein- und Ausschlusskriterien per Schreiben informiert und für eine potenzielle Teilnahme angefragt. Nach Einwilligung zur Studienteilnahme wurde den ProbandInnen eine schriftliche Einverständniserklärung sowie ein Datenblatt zur Unterzeichnung vorgelegt (Anhang A & B). In der schriftlichen Einverständniserklärung wurde über die Absichten und den Zweck dieser Studie aufgeklärt und mit dem Datenblatt Personendaten, wie unter anderem Gewicht, Körpergröße und Alter erhoben.

Die Ein- und Ausschlusskriterien für die ProbandInnen und TesterInnen lauteten folgendermaßen:

Das Einschlusskriterium für die ProbandInnen war das Alter:

- Die ProbandInnen mussten zwischen 60 und 85 Jahre alt sein

Zu den Ausschlusskriterien zählten:

- Verletzungen oder Operationen der unteren Extremität mit Auswirkungen auf deren motorische Funktion
- Erkrankungen oder Einnahme von Medikamenten, welche folgende Tätigkeiten einschränken:
 - zügiges Gehen über eine Strecke von vier Metern
 - aufrechtes Stehen ohne Unterstützung
 - mehrmaliges zügiges Aufstehen von einem Sessel hintereinander

Weiters mussten die ProbandInnen in der Lage sein, die Aktivitäten des täglichen Lebens ohne Unterstützung einer anderen Person auszuführen.

Die TesterInnen mussten einen Bachelorabschluss und mindestens zwei Jahre Berufspraxis aufweisen. Sie sollten zuvor keine Erfahrungen mit der SPPB gemacht haben.

2.4 Ablauf der Messungen

Mittels Research Randomizers wurde allen ProbandInnen eine ID-Nummer und eine der drei gleich großen Gruppen zugewiesen. Jede Gruppe unterlag einer individuellen Reihenfolge der Subtests der SBBP, um Lern- und Ermüdungseffekte zu vermeiden.

Die Reihenfolgen wurden folgendermaßen definiert:

Gruppe 1: Gleichgewichtstest → Gangtest → Aufstehtest

Gruppe 2: Gangtest → Aufstehtest → Gleichgewichtstest

Gruppe 3: Aufstehtest → Gleichgewichtstest → Gangtest

Den ProbandInnen wurde vor Beginn der Testungen eine standardisierte Anleitung (zu finden in den Anhängen C, D, E) vorgelesen. Diese beinhaltete die Aufgabenstellung der jeweiligen Testung, die Abbruchkriterien sowie Platz für individuelle Fragen.

Der Aufstehtest und die Gleichgewichtstests wurden zusätzlich zu den verbalen Anleitungen demonstriert. Im Anschluss an die Instruktion, wurden die ProbandInnen gefragt, ob sie alle Anweisungen verstanden hatten und ob sie bereit seien mit der Testdurchführung zu beginnen. Gestartet wurde die Zeitmessung mit den Worten „3, 2, 1, los“.

Die Gleichgewichtstests endeten nach Ablauf der Zeit mit den Worten: „Danke, die Testung ist beendet. Sie können die Position nun auflösen!“, der Gangtest mit Übertreten eines Fußes der Ziellinie und der Aufstehtest mit dem fünften Niedersetzen des/der ProbandIn.

2.4.1 Erstellung der Videos

Die HDV-Kamera wurde zur Erstellung der Videos auf einem dafür vorgesehenen Stativ fixiert und bei den jeweiligen Tests, wie im Folgenden beschrieben, positioniert:

Aufstehtest: Wie in Abbildung 4 skizziert, wurde die Kamera (K) rechts neben den vorderen Beinen des Stuhls mit einem Abstand von 2,4 m positioniert. Die Abmessungen betragen bei der Sitzfläche des Stuhls 0,395 m in der Tiefe und 0,465 m in der Breite, beim Abstand vom Boden zur Sitzfläche 0,46 m, bei den vorderen Sesselbeinen zur vorderen Kante der Kraftmessplatte (KMP) 0,03 m und beim seitlichen Abstand der Sesselbeine zur seitlichen Kante der Kraftmessplatte jeweils 0,275 m.

Die Kraftmessplatte hatte eine Breite von 0,405 m und eine Länge von 0,61 m. Der Stuhl wurde, wie in Abbildung 5 ersichtlich, vor der KMP positioniert. Zur Standardisierung wurde die Position mit Klebestreifen am Boden markiert. Ebenso wurde die Position der Kamera am Boden markiert.

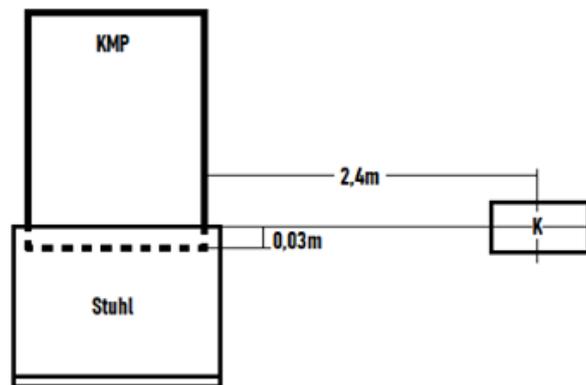


Abbildung 5 Setting Aufstehtest von Oben

Gleichgewichtstests: Bei den Gleichgewichtstests wurde die Kamera frontal vor den ProbandInnen mit einem Abstand von 2,35 m platziert. Zur Erleichterung der Einnahme der Position beziehungsweise zur Sicherheit wurde links neben der Kraftmessplatte ein Stuhl positioniert. Zur Veranschaulichung ist dies in Abbildung 6 skizziert.

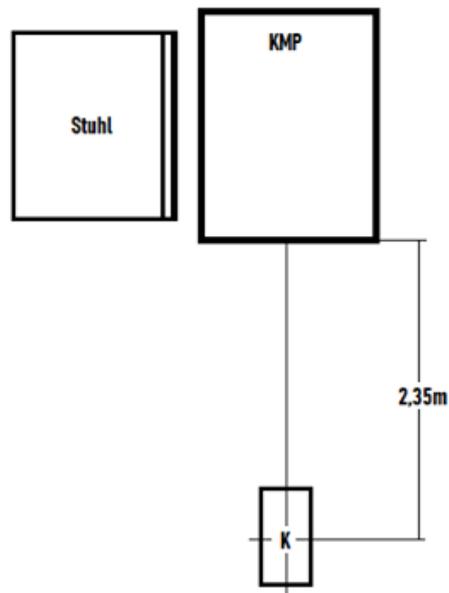


Abbildung 6 Setting Gleichgewichtstests von Oben

Gangtest: Beim Gangtest galt es eine Wegstrecke von vier Metern zu absolvieren. Die Lichtsensoren (in Abbildung 6 mit S für „Sensoren“ abgekürzt) an der Innenseite von vier 0,4 m großen Holzpflocken 0,1 m vom Boden entfernt befestigt. Diese waren paarweise je an der Start- und Ziellinie in einem Abstand von 0,7 m aufgestellt. Markiert wurde die Start- und Ziellinie jeweils durch einen vier Zentimeter breiten Klebestreifen. Der Vier Meter Abstand zwischen den Klebestreifen wurde von den Innenkanten gemessen. Die Kamera (K) war 0,9 m seitlich beziehungsweise 1,5m vor der Gangstrecke mit einem Winkel von 60° auf die gesamte Strecke positioniert.

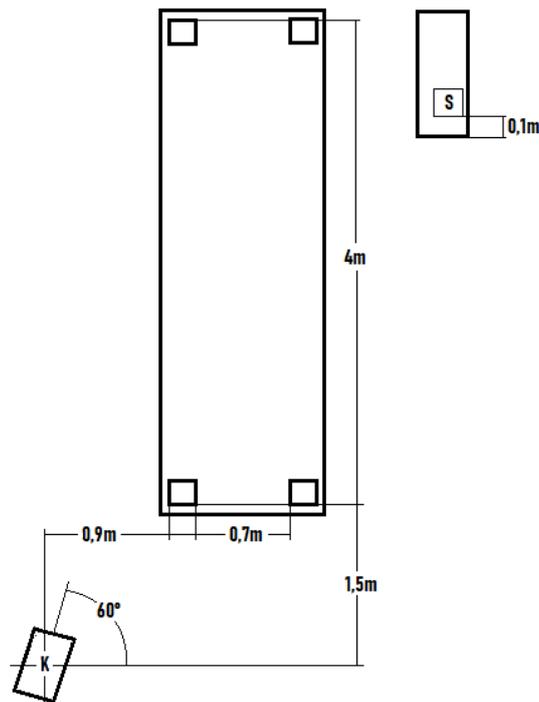


Abbildung 7 Setting Gangtestung von Oben (links) und Skizze eines Holzpflocks von der Seite (rechts)

2.4.2 Auswahl und Randomisierung

Von jedem der fünf Subtest der SPPB wurden je 63 Videos aufgezeichnet, welche mittels HDMI-Kabel automatisch auf den Computer des Labors und mittels USB-Stick auf ein MacBook Air transferiert wurden. Der Rhombengstand wurde aus dem weiteren Verlauf der Studie ausgeschlossen, da keineR der ProbandInnen diesen nicht erfolgreich beziehungsweise unsicher absolvierte. Bei der Durchführung des Tandemstands entstanden 15 Videos, bei denen die ProbandInnen Unsicherheiten zeigten beziehungsweise den Test nicht erfolgreich absolvierten. Beim Semitandemstand waren es drei Stück. Beim Gangtest beziehungsweise beim Aufstehetest wurde aus allen 63 entstanden Videos, wie auch aus ausgewählten Videos der Gleichgewichtstests, mittels Research Randomizer ausgelost. Diese bildeten eine Stichprobe der Grundgesamtheit aller ProbandInnen.

2.4.3 Bewertung durch TesterInnen

Um Lern- beziehungsweise Ermüdungseffekte zu vermeiden, wurden die zwölf, mittels Zufallsgenerator ausgewählten Videos, ebenfalls durch Zufallsgenerator in fünf verschiedene Rangreihenfolgen gebracht und fünf ID-Nummern zugeteilt. Auch die ID-Nummern wurden randomisiert den TesterInnen zugeteilt.

Nach einer Einführung anhand des "Short Physical Performance Battery Protocol and Score Sheets" wurde den fünf TesterInnen die zwölf Videos je zehnmal zur Bewertung mittels manueller, digitaler Handstoppung vorgelegt.

Die Messwerte wurden in einer Excel-Liste festgehalten und im Anschluss mit dem Statistikprogramm „SPSS Statistics“ analysiert.

2.5 Messinstrumente

Zur Aufzeichnung der Videos kam eine Canon Legria HV40 Kamera zum Einsatz. Die Videos wurden mittels HDMI-Kabel von der Kamera auf den Computer des Labors und mittels USB-Stick auf ein MacBook Air überspielt. Dieses wurde genutzt, um den TesterInnen die Videos zur Bewertung vorzuspielen. Für die Zeitmessungen durch die TesterInnen wurde eine handelsübliche Stoppuhr eingesetzt.

2.6 Auswertung der Daten

Die erhobenen Daten wurden mit dem Statistikprogramm „SPSS Statistics“ analysiert und die Inter- und Intratester-Reliabilität über den ICC bestimmt. JederR ProbandIn wurde von jedem Rater gemessen und die Rater wurden zufällig ausgewählt. So kamen für die Berechnungen der Interraterkorrelationen Shrout & Fleiss (1979) zufolge der ICC(3,k) beziehungsweise Mcgraw & Wong (1996) nach der ICC(C,k) zur Anwendung.

Bei der Interrater-Reliabilitätsanalyse wurde der ICC des Chair Rising Test, des Gangtest sowie des Tandemstandes berechnet. Der Rhombergstand und der Semitandemstand wurden aufgrund von mangelnden Daten aus dem weiteren Studienverlauf ausgeschlossen. Pro TesterIn und Test wurde aus je zehn Messwerten zu drei Videos der Mittelwert (Mean) bestimmt und in weiterer Folge mit den Means der restlichen Rater mittels SPSS verglichen. Bei den Berechnungen der Intraraterkorrelation wurden die Ratingrohwerte einzelner Rater verwendet. So kamen die ICC-Typen ICC(3,1) (Shrout & Fleiss, 1979) beziehungsweise ICC(C,1) (Mcgraw & Wong, 1996) zu tragen. Sowohl bei der Berechnung der Interrater- als auch der Intraraterkorrelation wurde mit dem SPSS-Modell „zweifach, gemischt“ gearbeitet. Zur Veranschaulichung ist dieses Auswahlverfahren in Tabelle 4 dargestellt.

Tabelle 4 Arten und Auswahl des ICC nach Shrout & Fleiss (1979) bzw. McGraw & Wong (1996)

Arten und Auswahl des ICC						
Frage 1	Wurde jeder Fall von jedem Rater eingeschätzt?					
	Nein		Ja			
Frage 2	Die Rater wurden zufällig ausgewählt.		Wurden alle Rater zufällig ausgewählt?			
			Ja		Nein	
Frage 3	Sind die Messwerte Rohwerte einzelner Rater oder Mittelwerte (k) verschiedenerer Rater?					
	Einzelwert	Mittelwert	Einzelwert	Mittelwert	Einzelwert	Mittelwert
ICC-Typ Shrout & Fleiss (1979)	ICC (1,1)	ICC (1,k)	ICC (2,1)	ICC (2,k)	ICC (3,1)	ICC (3,k)
ICC-Typ McGraw & Wong (1996)	ICC(1)	ICC(k)	ICC(A,1)	ICC(A,k)	ICC(C,1)	ICC(C,k)
SPSS-Modell	einfach, zufällig		zweifach, zufällig		zweifach, gemischt	
	einzelne Maße	∅ Maße	einzelne Maße	∅ Maße	einzelne Maße	∅ Maße

SPSS benötigt beim zweifach, gemischten Modell ebenso die Unterscheidung, ob die Berechnungen justiert oder unjustiert erfolgen sollen. Beim justierten Modell werden Mittelwertsunterschiede zwischen den Ratern (z.B. der Unterschied zwischen einem strengerem und einem milderem Rater) im Modell aus der Fehlervarianz herausgerechnet, während diese beim unjustierten Modell als Teil der Fehlervarianz erhalten bleiben (Wirtz & Caspar, 2002, S.189). Das unjustierte Modell entspricht der strengeren Prüfung und wird von SPSS als „Absolute Übereinstimmung“ bezeichnet. Dieses wurde für vorliegende Studie zur Prüfung gewählt.

3. Ergebnisse

Im folgenden Kapitel werden die Ergebnisse der statistischen Auswertung beschrieben und mittels Tabellen und Diagrammen grafisch dargestellt. Als Verfahren zur Berechnung der Korrelationen wurde der ICC gewählt. Der Korrelationswert des ICC kann nur positive Werte zwischen 0.00 und 1.00 annehmen. Bei einer optimalen Korrelation von 1.00 würden sich alle Messpaare auf der Geraden befinden. Weist der ICC einen Wert von >0.7 auf, so wird dieser als genügend erachtet. In der Literatur wird für Verlaufsmessungen von einzelnen PatientInnen ein ICC von $>0,9$ empfohlen.

Wie bereits in Kapitel 2.6 erwähnt, wurde bei der Interrater-Reliabilitätsanalyse der ICC der Messwerte des Chair Rising Tests, des Gangtests, sowie des Tandemstands berechnet. Jedes Video wurde von jedem Tester / jeder TesterIn (ID 1 - ID 5) zehnmal gestoppt. Es ergaben sich somit pro TesterIn und pro Video zehn Messwerte. Als Beispiel hierfür sind die Messwerte von ID1 von Video 1 des Chair Rising Tests in einem Punktestreudiagramm in Abbildung 8 dargestellt. In der Horizontalwertachse ist die Messwiederholung ersichtlich, in der Vertikalachse die gestoppte Zeit in Sekunden und Hundertstelsekunden.

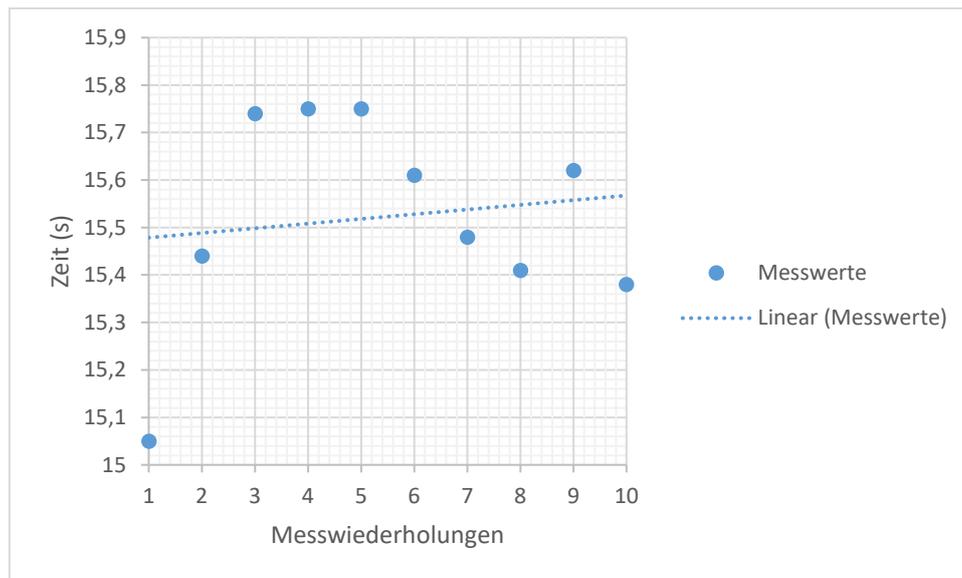


Abbildung 8 Darstellung der Ergebnisse des CRT der ID 1 von Video 1

Beim CRT sowie beim TS wurden aufgrund von manuellen Fehlern je zwei Datensätze aus dem weiteren Verlauf der Studie ausgeschlossen (ID 2, CRT 2, Durchgang 2; ID4, CRT2, Durchgang 3 und ID 3, CRT 3, Durchgang 1 + 2). Alle Messergebnisse zusammen sind in Tabelle 5 abgebildet. In der Horizontalwertachse sind die Testabkürzungen zu finden (Chair Rising Test - CRT, Tandemstand – TS und Gangtest – GT mit den Videos 1 – 3), in der Vertikalwertachse sind die ID Nummern beziehungsweise die Durchgangsnummer zu finden.

Tabelle 5 Urliste der Rohdaten der TesterInnen [s]

		CRT1	CRT2	CRT3	TS1	TS2	TS3	GT1	GT2	GT3
ID1	1	15.05	25.75	20.63	2.44	10.14	2.31	2.14	2.37	3
	2	15.44	21	20.51	2.69	10	2.71	2	2.02	2.98
	3	15.74	24.95	20.23	2.26	10.02	2.08	2.75	2.34	2.89
	4	15.75	25.62	20.49	1.99	9.86	1.75	2.83	2.23	3.04
	5	15.75	25.14	20.75	1.98	9.28	1.96	2.81	2.3	3
	6	15.61	25.2	19.8	2.01	9.33	2.18	3	2.2	2.66
	7	15.48	25.5	20.51	1.9	9.39	1.56	2.83	2.32	2.89
	8	15.41	25.42	20	1.8	9.56	1.68	2.64	2.31	2.75
	9	15.62	25.6	20.3	1.77	8.94	1.7	2.64	2.67	2.93
	10	15.38	25.6	20.24	1.94	9.38	1.83	2.74	2.18	2.26
ID2	1	15.1	25.02	19.99	2.74	10.19	3.14	2.89	2.16	2.73
	2	15.1	24.93	19.77	2.22	10.37	2.94	2.88	2.17	2.74
	3	14.98	25.99	19.92	2	10.17	2.8	2.88	2.19	2.68
	4	14.93	24.98	19.94	2.43	10.3	2.92	2.86	2.13	2.79
	5	15.15	24.94	20.24	2.48	10.39	2.92	2.9	2.14	2.7
	6	15.1	24.84	20.13	2.38	10.23	2.94	2.9	2.2	2.68
	7	15.14	24.95	20.2	2.43	10.28	2.83	2.89	2.2	2.64
	8	15	24.88	20.26	2.42	10.19	2.83	2.97	2.19	2.66
	9	15.14	24.94	20.24	2.63	10.19	2.96	2.85	2.28	2.64
	10	15.14	24.92	20.11	2.47	10.35	2.95	2.87	2.25	2.69
ID3	1	15.01	24.87	20.12	1.94	9.74	3.66	2.39	2.27	2.62
	2	15	24.99	19.99	2.24	9.88	4.64	3.02	2.34	2.86
	3	15.3	24.97	20.02	2.03	10.05	1.99	3.06	2.36	2.9
	4	14.83	24.83	20.56	2.03	9.98	2.08	3.01	1.36	2.95
	5	15.17	25.14	20.21	1.94	9.93	1.75	2.94	2.41	2.69
	6	15.07	24.96	19.7	1.95	9.87	1.39	3.01	1.76	2.7
	7	15.21	24.67	20.12	2.1	9.9	1.32	3.03	2.41	2.69
	8	15.24	25.29	20.39	2.09	9.98	1.63	2.93	2.07	2.8
	9	15.28	24.78	19.9	2.03	9.88	1.81	2.99	2.31	2.83
	10	15.4	24.82	20.06	2.07	10.11	1.69	3.04	2.33	2.83
ID4	1	15.4	25.17	20.51	2	10.04	2.62	3	2.51	2.88
	2	15.4	25.11	20.63	2.71	9.88	2.04	2.89	2.31	2.97
	3	15.32	19.56	20.57	2.47	9.84	2.2	2.99	2.3	2.95
	4	14.92	25.14	20.23	2.57	9.86	1.88	2.99	2.18	3.12
	5	15.55	24.89	20.34	2.17	9.05	1.98	3.11	2.31	2.9
	6	15.19	25.1	20.38	2.2	9.6	1.87	3.1	2.23	3.18
	7	15.33	25.09	20.39	2.3	9.76	1.9	3.1	2.31	3.14
	8	15.6	24.94	20.28	2.12	9.89	1.98	3.02	2.2	3.24
	9	15.69	24.91	20.03	1.9	9.88	2.4	3.07	2.26	3
	10	15.38	25.1	20.21	2.25	9.84	1.8	3.2	2.3	3.09
ID5	1	14.9	24.29	19.77	1.83	9.84	2.13	3.24	2.57	2.93
	2	15.14	24.37	19.55	1.98	9.74	1.96	3.13	2.58	2.88
	3	15.08	24.3	19.58	2	9.75	1.87	3.08	2.29	2.7
	4	15.12	24.24	19.71	1.95	9.7	1.88	3.13	2.26	2.82
	5	14.9	24.36	19.48	1.85	9.57	1.88	2.91	2.24	2.74
	6	14.84	24.47	19.38	1.99	9.68	1.89	2.97	2.25	2.84
	7	14.98	24.53	19.75	1.94	9.77	1.45	2.9	2.28	2.77
	8	14.95	24.34	19.57	1.9	9.78	1.85	3.08	2.21	2.74
	9	14.96	24.3	19.44	1.93	9.76	1.78	3.94	2.3	2.78
	10	14.95	24.47	19.58	1.91	9.76	1.89	2.8	2.26	2.89

3.1 Interrater-Reliabilitätsanalyse mit SPSS

Für die Interrater-Reliabilitätsanalyse wurde pro TesterIn (ID) und pro Video aus den zehn erhobenen Messwerten der Mittelwert gebildet. Weiters wurden die Standardabweichungen (SD) der Mittelwerte berechnet, um die Streubreite dieser aufzeigen zu können. Die Standardabweichungen zeigen sich alle niedrig, was auf eine gute Repräsentativität der Mittelwerte schließen lässt (Tabelle 6).

Tabelle 6 Mittelwerte und Standardabweichung der Messergebnisse [s]

	ID 1		ID 2		ID 3		ID 4		ID 5	
	CRT	SD								
Video 1	15.52	+/- 0.21	15.08	+/- 0.07	15.15	+/- 0.16	15.38	+/- 0.21	14.98	+/- 0.09
Video 2	24.98	+/- 0.25	25.42	+/- 0.32	24.93	+/- 0.17	25.05	+/- 0.10	24.37	+/- 0.09
Video 3	20.35	+/- 0.28	20.08	+/- 0.16	20.11	+/- 0.23	20.36	+/- 0.17	19.58	+/- 0.12
	GT	SD								
Video 1	2.64	+/- 0.30	2.89	+/- 0.03	2.94	+/- 0.19	3.05	+/- 0.08	3.12	+/- 0.30
Video 2	2.29	+/- 0.16	2.19	+/- 0.04	2.16	+/- 0.33	2.29	+/- 0.09	2.32	+/- 0.13
Video 3	2.84	+/- 0.22	2.70	+/- 0.04	2.79	+/- 0.10	3.05	+/- 0.12	2.81	+/- 0.07
	TS	SD								
Video 1	2.08	+/- 0.28	2.42	+/- 0.19	2.04	+/- 0.09	2.27	+/- 0.24	1.93	+/- 0.05
Video 2	9.59	+/- 0.37	10.27	+/- 0.08	9.93	+/- 0.10	9.76	+/- 0.26	9.74	+/- 0.07
Video 3	1.98	+/- 0.33	2.92	+/- 0.09	1.71	+/- 0.25	2.07	+/- 0.25	1.86	+/- 0.16

In weiterer Folge wurden die Mittelwerte der einzelnen TesterInnen mittels SPSS verglichen. Wie bereits in Kapitel 2.6 beschrieben, wurde die Auswahl der ICC-Typen Shrout & Fleiss (1979) beziehungsweise Mcgraw & Wong (1996) durchgeführt. Aufgrund der Tatsache, dass sich die Autorin dieser Arbeit dafür entschieden hat, die Mittelwerte der einzelnen TesterInnen zu vergleichen, wurden der ICC-Typ ICC(3,k) beziehungsweise ICC(C, k) angewendet. Somit kam das SPSS-Modell für durchschnittliche Werte zur Anwendung. In den Tabellen 7, 8 und 9 sind die Ergebnisse der Interrater-Reliabilitätsanalyse dargestellt. Die Korrelation innerhalb der Klasse ergab beim Chair Rising Test (CRT) sowie auch beim Tandemstand (TS) 0.999. Beim Gangtest (GT) lag diese bei 0.972.

Tabelle 7 Interrater-Korrelationskoeffizient des CRT

	Korrelation innerhalb der Klasse ^b	95%-Konfidenzintervall		F-Test mit wahrem Wert 0			Sig.
		Untergrenze	Obergrenze	Wert	df1	df2	
Einzelne Maße	.996 ^a	0,974	1,000	2826,264	2	8	0,000
Durchschnittliche Maße	.999 ^c	0,995	1,000	2826,264	2	8	0,000

Modell mit gemischten Zwei-Weg-Effekten, bei dem die Personeneffekte zufällig und die Maßeffekte fest sind.

- a. Der Schätzer ist derselbe, unabhängig davon, ob ein Wechselwirkungseffekt vorliegt oder nicht.
- b. Korrelationskoeffizienten des Typs A innerhalb der Klasse unter Verwendung einer Definition der absoluten Übereinstimmung.
- c. Die Schätzung wird unter der Annahme berechnet, dass kein Wechselwirkungseffekt vorliegt, da anderenfalls keine Schätzung durchgeführt werden könnte.

Tabelle 8 Interrater-Korrelationskoeffizient des GT

	Korrelation innerhalb der Klasse ^b	95%-Konfidenzintervall		F-Test mit wahrem Wert 0			Sig.
		Untergrenze	Obergrenze	Wert	df1	df2	
Einzelne Maße	.875 ^a	0,548	0,996	46,284	2	8	0,000
Durchschnittliche Maße	.972 ^c	0,858	0,999	46,284	2	8	0,000

Modell mit gemischten Zwei-Weg-Effekten, bei dem die Personeneffekte zufällig und die Maßeffekte fest sind.

- a. Der Schätzer ist derselbe, unabhängig davon, ob ein Wechselwirkungseffekt vorliegt oder nicht.
- b. Korrelationskoeffizienten des Typs A innerhalb der Klasse unter Verwendung einer Definition der absoluten Übereinstimmung.
- c. Die Schätzung wird unter der Annahme berechnet, dass kein Wechselwirkungseffekt vorliegt, da anderenfalls keine Schätzung durchgeführt werden könnte.

Tabelle 9 Interrater-Korrelationskoeffizient des TS

	Korrelation innerhalb der Klasse ^b	95%-Konfidenzintervall		F-Test mit wahrem Wert 0			Sig.
		Untergrenze	Obergrenze	Wert	df1	df2	
Einzelne Maße	.994 ^a	0,964	1,000	2423,165	2	8	0,000
Durchschnittliche Maße	.999 ^c	0,993	1,000	2423,165	2	8	0,000

Modell mit gemischten Zwei-Weg-Effekten, bei dem die Personeneffekte zufällig und die Maßeffekte fest sind.

- a. Der Schätzer ist derselbe, unabhängig davon, ob ein Wechselwirkungseffekt vorliegt oder nicht.
- b. Korrelationskoeffizienten des Typs A innerhalb der Klasse unter Verwendung einer Definition der absoluten Übereinstimmung.
- c. Die Schätzung wird unter der Annahme berechnet, dass kein Wechselwirkungseffekt vorliegt, da anderenfalls keine Schätzung durchgeführt werden könnte.

3.2 Intrarater-Reliabilitätsanalyse mit SPSS

Auch bei der Intrarater-Reliabilitätsanalyse wurde der ICC des Chair Rising Tests (CRT), des Gangtests (GT) sowie des Tandemstands (TS) berechnet (siehe Kapitel 2.6). Wie im vorigen Unterkapitel bereits beschrieben, wurde jedes Video von jedem Tester / jeder TesterIn (ID 1 - ID 5) zehnmal gestoppt. Es ergaben sich somit pro TesterIn und pro Video zehn Messwerte. Aus diesen Messwerten wurde pro Rater pro Video der ICC berechnet. Der ICC des CRT lag bei ID 1 bei 0.994, bei ID3 bei 0.998 und bei ID 2, ID 4 und ID 5 bei 1.000. Der ICC des TS lag bei ID 1 bei 0.994, bei ID 4 bei 0.996, bei ID 2 und 5 bei 0.999 und bei ID3 bei 1.000.

Der ICC des GT lag bei ID 1 bei 0.531, bei ID 2 bei 0.986, bei ID 3 bei 0.745, bei ID 4 bei 0.948 und bei ID 5 bei 0.790 (Tabelle 10).

Tabelle 10 Intrarater-Korrelationskoeffizienten aller IDs

	ID 1	ID 2	ID 3	ID 4	ID 5
ICC CRT	0.994	1.000	0.998	1.000	1.000
ICC TS	0.994	0.999	1.000	0.996	0.999
ICC GT	0.531	0.986	0.745	0.948	0.790

Zur Veranschaulichung der Intrarater-Korrelationskoeffizienten der einzelnen Tests dienen die Abbildungen neun bis elf. In der Horizontalwertachse ist die ID Nummer ersichtlich, in der Vertikalachse der Korrelationskoeffizient.

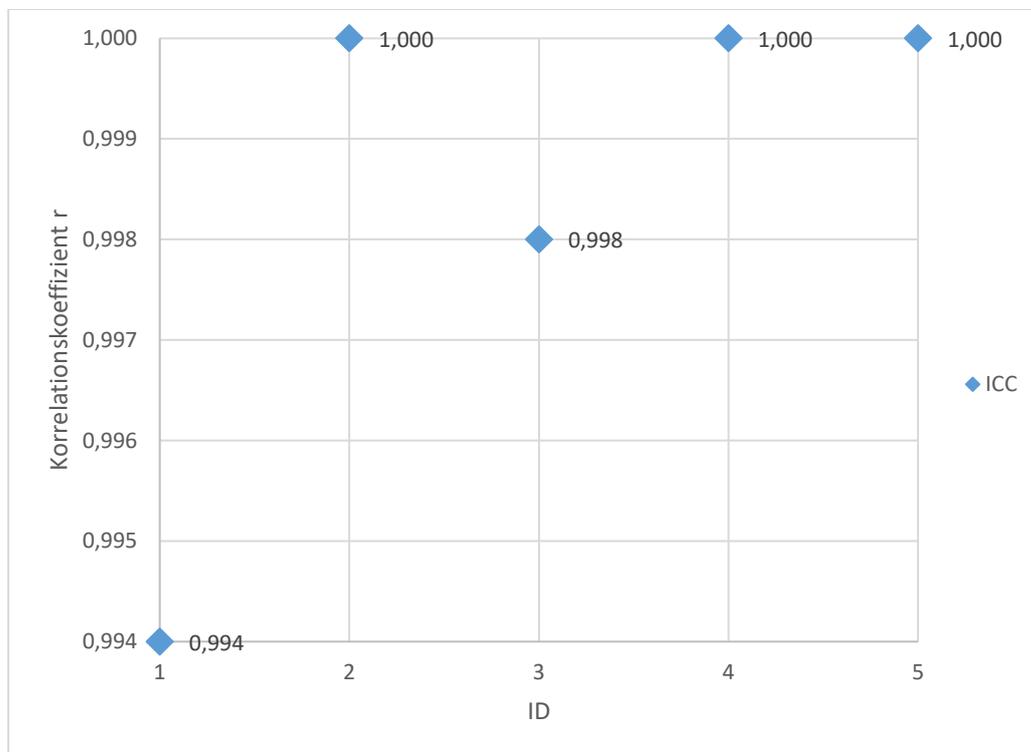


Abbildung 9 Intrarater-Korrelationskoeffizienten des CRT von ID 1 bis ID 5

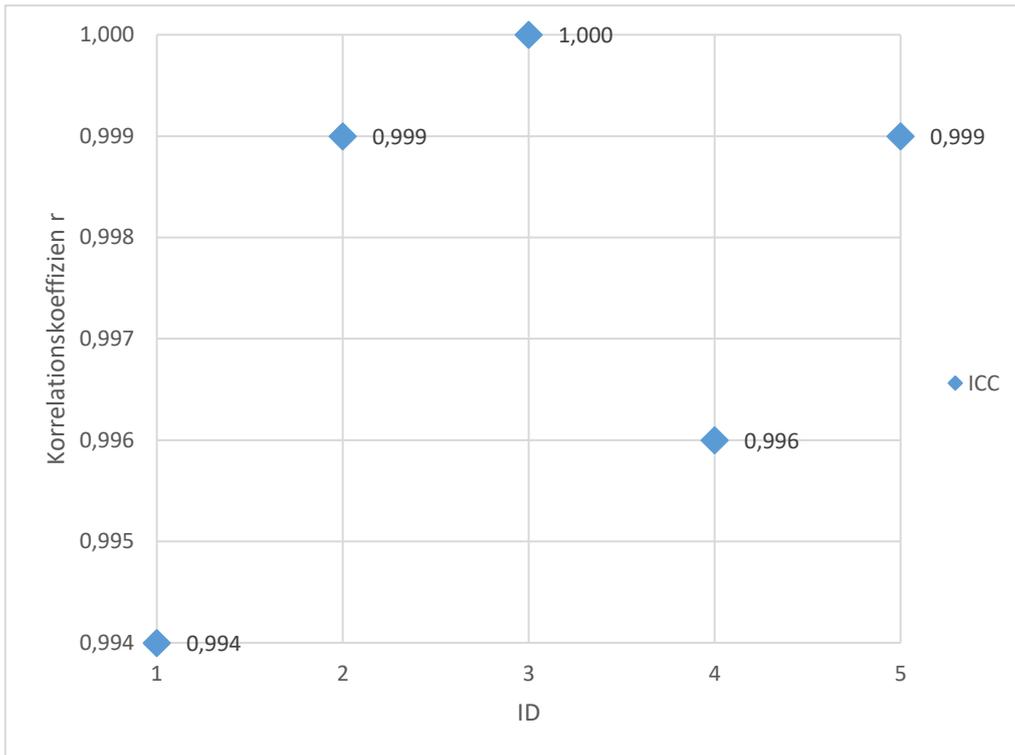


Abbildung 10 Intrarater-Korrelationskoeffizienten des TS von ID 1 bis ID 5

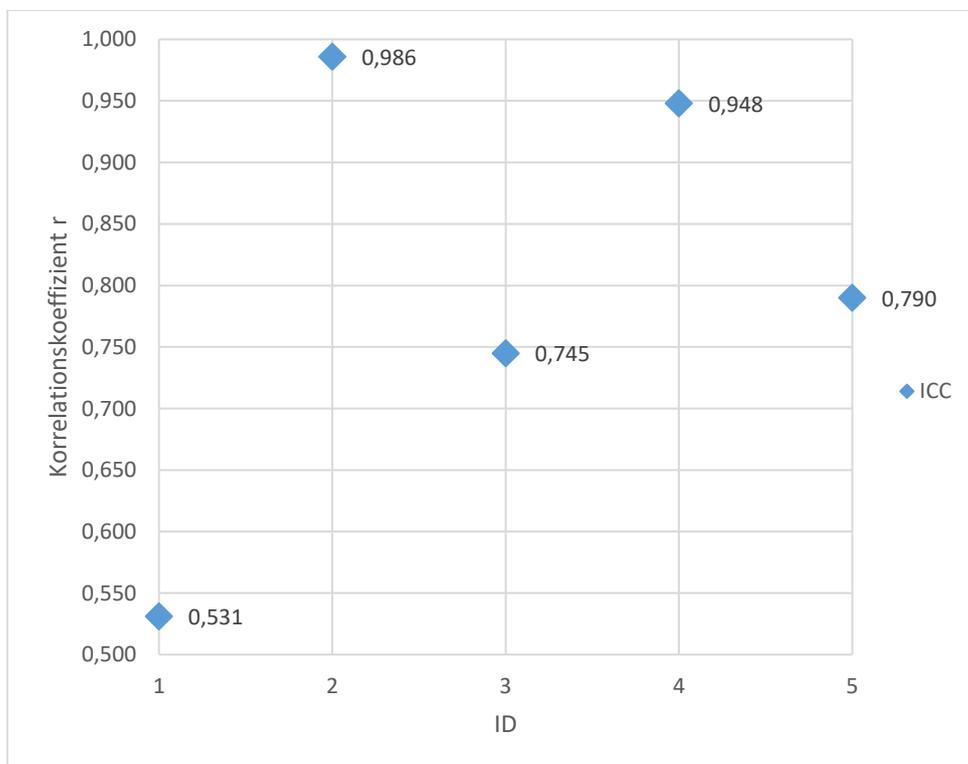


Abbildung 11 Intrarater-Korrelationskoeffizienten des GT von ID 1 bis ID 5

Um einen gemeinsamen Intrarater-Korrelationskoeffizienten der Rater mit der Literatur vergleichen zu können, hat sich die Autorin dieser Studie dazu entschieden, diesen ebenso zu berechnen. Die Berechnungen ergaben einen ICC von 0.998 für den CRT, 0.997 für den TS und 0.759 für den GT (Abbildung 12).



Abbildung 12 Errechnete Intrarater-Korrelationskoeffizienten der Subtests (beinhaltet alle IDs)

4. Diskussion

In der Einleitung dieser Bachelorarbeit wurden das Risiko und die Auswirkungen von Stürzen im fortgeschrittenen Alter beschrieben. Eine Möglichkeit das Sturzrisiko zu messen, stellt die SPPB dar. In den Unterkapiteln wurden das Gütekriterium Reliabilität, manuelle Zeitmessungen und die SPPB mit ihren Subtests, ihrem Anwendungsgebiet sowie ihrer Reliabilität erläutert.

Dies führte zur der Forschungsfrage dieser Bachelorstudie, welche folgend erneut dargestellt wird. Die Forschungsfrage war es herauszufinden, ob es Unterschiede bei manuellen Zeitmessungen der Subtests der SPPB bei verschiedenen TesterInnen beziehungsweise bei Testwiederholungen gibt und wie groß die Streuungen sind.

Aufgrund der Fragestellung und deren Hauptziele lassen sich folgende Hypothesen ableiten: Die H_0^1 -Hypothese lautet, dass es keine signifikanten Unterschiede bei manuellen Zeitmessungen der Subtests der SPPB bei verschiedenen TesterInnen gibt.

Die H_1^1 -Hypothese lautet, dass es signifikante Unterschiede bei manuellen Zeitmessungen der Subtests der SPPB bei verschiedenen TesterInnen gibt. H_0^2 Hypothese lautet, dass es keine signifikanten Unterschiede bei manuellen Zeitmessungen der Subtests der SPPB bei Wiederholungsmessungen gleicher TesterInnen gibt. Die H_1^2 -Hypothese lautet, dass es signifikante Unterschiede bei manuellen Zeitmessungen der Subtests der SPPB bei Wiederholungsmessungen gleicher TesterInnen gibt.

4.1 Interpretation der Ergebnisse

Zur Prüfung der oben genannten Forschungshypothesen wurde, wie im Kapitel Ergebnisse bereits ausführlich berichtet, der ICC für Inter- und Intrarater mittels SPSS berechnet.

4.1.1 ICC

Sowohl der Chair Rising Test (CRT) als auch der Tandemstand (TS) zeigten sich in vorliegender Studie mit nahezu perfekten **Interrater-Korrelationen**. Mit einem ICC von je 0.999 wiesen diese jeweils eine sehr hohe *Interrater-Übereinstimmung* auf. Der Gangtest (GT) erreichte einen ICC von 0.972. Ein ICC unter 0.7 wird laut „Scientific Advisory Committee of Medical Outcomes Trust“ als nicht genügend angesehen wird und für Verlaufsmessungen von einzelnen PatientInnen wird ein ICC >0.9 empfohlen (Ostir et al., 2002).

Bei den Studien von Duncan, Leddy & Earhart (2011), Lummel et al. (2016) und Telenius, Engedal & Bergland (2015) zeigte sich, dass handgestoppte Funktionstests hohe *Interrater-Korrelationen* aufweisen.

Duncan, Leddy & Earhart (2011) untersuchten in ihrer Studie „Five times sit-to-stand tests performance in Parkinson’s disease“ die *Interrater-Reliabilität* von Funktionstests bei 88 Menschen mit idiopathischer Parkinson-Erkrankung. Primäres Ziel war es, den CRT auf seine Zuverlässigkeit zu prüfen.

Die Messungen der *Interrater-Reliabilität* wurden von zwei PhysiotherapeutInnen gleichzeitig ausgeführt. Die PatientInnen waren während der Testungen in ihrer Medikationsphase. Zur Testbatterie gehörten neben dem CRT, der Mini-Balance Evaluation Systems Test (Mini-BEST, eine aus 14 Testungen bestehende Balance Batterie) sowie der Six-minute walking test (6MWT). Alle drei Tests konnten hohe ICC-Werte (CRT: 0.99, Mini-BEST: 0.92, 6MWT: 0.96) erreichen. In ihrem Conclusio beschrieben Duncan, Leddy & Earhart (2011) den CRT als schnelle, einfach durchgeführte und reliable Maßnahme zur Evaluation von Sturzrisiko bei Menschen mit Parkinson. Ihre Studie und die vorliegende Studie gleichen sich in der ähnlich geringen ProbandInnenanzahl, nicht jedoch beim Testsetting der gleichzeitigen Messungen. Ebenso mit der *Interrater-Reliabilität* einer Testbatterie beschäftigten sich Telenius, Engedal & Bergland (2015) in ihrer Studie „Inter-rater reliability of the Berg Balance Scale, 30s chair stand test, and 6m walking test and construct validity of the Berg Balance Scale in nursing home residents with mild-to-moderate dementia“.

In ihrer Studie wurden 33 PflegeheimpatientInnen mit leichter bis mäßiger Demenz einmalig von zwei Ratern getestet. Einer der Rater gab die Anweisungen, beide Rater bewerteten die Leistung des Patienten / der Patientin. Es wurden drei der SPPB sehr ähnliche Tests überprüft: der Berg-Balance-Test (Testbatterie mit 14 Balance Testungen), ein 30 Sekunden Aufstehtest sowie ein sechs Meter GT. Der ICC für den Berg-Balance-Test lag bei 0.99, für den 30 Sekunden Aufstehtest bei 1.00 und für den sechs Meter GT bei 0.97. Die Studie von Lummel et al. (2016) konnte für den „Timed Up and Go“ (Aufstehen, drei Meter gehen, umdrehen, zurückgehen, wieder niedersetzen) einen *Interrater-Korrelationswert* von 0.99 bei Menschen mit diversen medizinischen Diagnosen erzielen. Schlussfolgernd ist zu sagen, dass es bei vorliegender Studie sowohl beim CRT, beim TS als auch beim GT keine signifikanten Unterschiede bei manuellen Zeitmessungen der Subtests der SPPB bei verschiedenen TesterInnen gibt.

Bei der ***Intrarater-Reliabilitätsanalyse*** zeigten der CRT sowie der TS jeweils hohe Übereinstimmung. Mit Werten von 0.994, 0.998 und dreimal 1.00 beim CRT sowie 0.994, 0.996, zweimal 0.999 sowie 1.00 beim TS konnten beide Tests äußerst hohe Korrelationen erzielen. Der GT hingegen schnitt deutlich schlechter als die anderen Tests ab.

Die Tatsache, dass der ICC des GT in vorliegender Studie vergleichsweise so schlecht abschneidet, könnte daran liegen, dass die TesterInnen die ProbandInnen nur über Videos und nicht real sehen konnten. Der Winkel der Kamera könnte hierbei eine wesentliche Rolle gespielt haben.

Bei ID 1 erreichte der ICC nur 0.531, bei ID 3 0.745, bei ID 5 0.790, bei ID 4 0.948 und bei ID 2 0.986. Wie in den Abbildungen 9, 10 und 11 in Kapitel 3.2 ersichtlich, präsentierte sich ID 1 mit deutlich schlechteren *Intraklassen-Korrelationen*, als die anderen Rater. Beim CRT beziehungsweise beim TS ist der Unterschied zu den restlichen TesterInnen minimal, beim GT jedoch wird die genannte Beobachtung deutlich ersichtlich.

Die Erkenntnis der schlechten Korrelation beim GT machten auch Jette, Jette, Ng, Plotkin, & Bach (1999) in ihrer Studie „Are Performance-Based Measures Sufficiently Reliable for Use in Multicenter Trials?“. Ihre Ambition war es, eine aus sieben Tests bestehende Performance Batterie auf ihre Reliabilität zu überprüfen. Eingesetzt wird diese zur Erhebung von funktionellen Einschränkungen von älteren, gebrechlichen Menschen. Demnach wurden 105 ProbandInnen im Alter zwischen 65 - 94 Jahren (Altersdurchschnitt: 78 Jahre) zweimal im Abstand von zwei Wochen getestet. Test und Retest wurden bei allen ProbandInnen von einem Rater mit handelsüblichen Stoppuhren durchgeführt. Neben Tests wie dem „Timed Up and Go“, beinhaltete die Performance Batterie unter anderem den CRT sowie den Eight-Foot-walk Test.

Wie in Kapitel 1.3.1 beschrieben, ist der Eight-Foot-walk Test die ursprüngliche, 2,44 Meter lange Version der heutigen, mit drei beziehungsweise vier Meter üblichen Gangtestung der SPPB. Der Rater in der Studie von Jette, Jette, Ng, Plotkin, & Bach (1999) konnte bei dem GT einen ICC von 0.79 erreichen. Dieser Wert ist ähnlich gering wie die *Intraklassen-Korrelationskoeffizienten* des GT in vorliegender Studie. ID 5 erzielte exakt denselben ICC Wert. Die ICCs von ID 1 und ID 3 präsentierten sich noch geringer, ID 2 und 4 zeigten akzeptable ICC Werte. Während vorliegende Studie bei der *Intrarater-Reliabilitätsanalyse* für den CRT gute Ergebnisse erzielen konnte, erlangte die Studie von Jette, Jette, Ng, Plotkin, & Bach (1999) lediglich einen ICC von 0.67. Dadurch, dass ihre Testungen mit zwei Wochen Abstand und auch nur von einem Rater durchgeführt wurde, sind die beiden Studien nur bedingt miteinander vergleichbar. Einen etwas besseren, jedoch immer noch ungenügenden ICC zur Wiederbefundung von PatientInnen konnten Curb et al. (2006) in ihrer Studie „Performance-Based Measures of Physical Function for High-Function Population“ erreichen. Ihr Ziel war es leistungsbasierte Funktionsmessungen für den Einsatz in klinischen Forschungseinrichtungen zu verbessern und zu erweitern.

Die ProbandInnen waren in zwei Gruppen im Alter zwischen 35 und 55 beziehungsweise zwischen 56 und 71 aufgeteilt und wiesen keine signifikanten Funktionseinschränkungen auf. Die beiden Altersgruppen wurden verwendet, um die Altersspanne und die funktionale Vielfalt der Stichprobe zu erhöhen. Die Standarddurchführung von fünf Wiederholungen beim CRT wurde auf zehn Wiederholungen angehoben, um den Test noch schwerer zu gestalten und eventuell Streuungen besser aufzeigen zu können.

Der CRT konnte einen ICC von 0.80 erreichen und wurde somit als reliabel befunden. Vor allem die Balancetestungen, vorallem der TS, schnitten schlecht ab. Mit einem ICC von nur 0,22 wurde dieser als nicht reliabel eingestuft. Diese Erkenntnis wurde in vorliegender Studie nicht gemacht. Mit ICCs von 0.994, 0.998 und dreimal 1.000 wurde dieser als sehr reliabel bewertet. Jedoch sind auch diese Studien nur bedingt zu vergleichen, da es sich bei der Studie von Curb et al. (2006) um eine Test-Retest Studie handelt. Ostir et al. (2002) beschrieben die Subtests der Short Physical Performance Batterie in ihrer Studie als sehr reliabel.

Sie wendeten das für die SPPB übliche Punktesystem an. Für jeden der fünf Tests der SPPB konnten maximal vier Punkte erreicht werden, zwölf Punkte war die maximale Punkteanzahl. Der *Intraklassen-Korrelationskoeffizient* bei der Test-Retestung mit dem Abstand von einer Woche konnte ICCs von 0.88-0.92 erreichen. Der ICC zum Zeitpunkt des Studienendes (nach sechs Monaten) lag bei 0.77 (mit einer Spannweite zwischen 0.72 und 0.79). Der ICC aller TesterInnen bei vorliegender Studie bei CRT, GT und TS war 0.999 (mit einer Spannweite von 0.759 und 0.998).

Auch diese beiden Tests sind nicht unmittelbar vergleichbar, da es sich auch hier um ein Test-Retest Testsetting gehandelt hat. Schlussfolgernd ist zu sagen, dass es bei vorliegender Studie sowohl beim CRT als auch beim TS keine signifikanten Unterschiede bei manuellen Zeitmessungen der Subtests der SPPB bei Wiederholungsmessungen gleicher TesterInnen gibt. Beim GT jedoch ergab sich die Erkenntnis, dass es bei manuellen Zeitmessungen der Subtests der SPPB bei Wiederholungsmessungen gleicher TesterInnen signifikante Unterschiede gibt.

4.2 Klinische Relevanz

Für den klinischen Alltag kann nun gesagt werden, dass Testsituationen unterschiedlich bewerten werden können. Diese Problematik ist vor allem im Zuge von TherapeutInnenwechsel relevant. Ebenso kann gesagt, dass klare Beurteilungskriterien wichtig sind, um die Reliabilität von Tests zu gewähren. Diese sollten keinen Spielraum zulassen und klar verständlich definiert sein. Trotz klarer Definitionen in diese Studie, zeigten sich insbesondere beim GT signifikante Unterschiede in der Bewertung der Tests.

Dies ist womöglich auf das Testsetting zurückzuführen, welches aufgrund der Kameraposition nicht optimal war. Trotz einer Standardabweichung von maximal 0.37 wurde deutlich, dass grenznahe Messergebnisse eine Problematik darstellen können. In vorliegender Studie zeigte sich genannte Problematik beim TS Video 2. Den Messergebnissen einiger TesterInnen zu Folge, hätte der/die ProbandIn den Test nicht bestanden und somit eine andere Punktebewertung erhalten. Dies gilt es bei der Interpretation solcher Testbatterien zu beachten. Manuelle Zeitmessungen werden immer fehleranfälliger sein, als elektronische. Es gilt abzuwägen, was im Alltag praktikabler ist.

4.3 Limitationen

In diesem Unterkapitel sind einige der Limitationen erwähnt, welche im Zuge dieser Studie aufgekommen sind.

1. Die Short Physical Performance Battery (SPPB) dient zur Evaluierung der funktionellen Leistungsfähigkeit von Personen im fortgeschrittenen Alter. Das Alter der ProbandInnen lag zwischen 60 und 82 Jahren, das Durchschnittsalter der ProbandInnen betrug 69 Jahre \pm 6 Jahre.
2. Somit zählten die ProbandInnen zwar zur Gruppe der älteren Menschen (61-74 Jahre), nicht jedoch zur Gruppe der alten Menschen (75-90 Jahre), der eigentlichen Zielgruppe der SPPB. Die ProbandInnen zeigten sich großteils körperlich fit, somit entstanden nur wenige Videos mit Auffälligkeiten.
3. Die Anzahl der ProbandInnen war mit 21 Personen relativ klein und ist deshalb nur eingeschränkt repräsentativ für Menschen in diesem Alter.
4. Die Anzahl der TesterInnen war mit 5 Personen ebenso relativ klein und somit sind ihre Ergebnisse nur bedingt aussagekräftig.
5. Der Ausbildungsweg der TesterInnen gleichte sich im Abschluss eines FH-Studiums der Physiotherapie. Die Testergebnisse lassen sich womöglich nicht mit jenen von PhysiotherapeutInnen ohne FH – Abschluss vergleichen.
6. Die Tatsache, dass die TesterInnen ihre Messungen anhand von Videos durchführten ist nicht alltagsnahe. Besonders ist hier das Testsetting des Gangtests aufgrund des Blickwinkels durch die Kamera womöglich nicht optimal.

5. Schlussfolgerungen und Ausblick

Bei den Untersuchungen vorliegender Studie zeigte sich nur bei Wiederholungsmessungen gleicher TesterInnen des Gangtests, signifikante Unterschiede in den Messergebnissen und daraus resultierend im ICC.

Bei Messungen verschiedener TesterInnen war dies beim ICC nicht zu sehen. Da bei der Interrater-Reliabilität Mittelwerte zur Berechnung herangezogen wurden, war kein signifikanter Unterschied zu den anderen Tests ersichtlich. Zurückzuführen könnte dies jedoch auch auf das ungünstige, nicht alltagsnahe Testsetting sein. Beim Tandemstand und auch beim Chair Rising Test kam es weder bei Messungen verschiedener TesterInnen noch bei Wiederholungsmessungen gleicher TesterInnen zu signifikanten Unterschieden. Der Rhombergstand und der Semitandemstand wurden aufgrund von mangelnden Daten aus dem weiteren Studienverlauf ausgeschlossen, somit kann auch keine Aussage über ihre Reliabilität getroffen werden.

Schlussfolgernd ist zu sagen, dass jedes Assessment nur so reliabel ist, wie die Bewertung durch seinen Tester / seine Testerin.

Resultierend aus dieser Studie hat sich die Autorin überlegt, welche zukünftigen Studien sinnvoll wären und was es bei diesen zu beachten gilt:

- Das Ziel einer dieser Studien könnte es sein, die Reliabilität des Timed-Up-And-Go Tests zu überprüfen und die Ergebnisse mit jenen des Gangtests der SPPB zu vergleichen. Beim Timed-Up-And-Go-Test handelt es sich um ein gängig genutztes Assessment zur Beurteilung der Mobilität und des Sturzrisikos einer von Menschen im fortgeschrittenen Alter. Hierbei wird der Patient aufgefordert, von einem Stuhl aufzustehen, drei Meter zu gehen, umzudrehen und sich wieder niederzusetzen. Das Testsetting ist für geriatrische Personen somit alltagsnäher. Interessant wäre festzustellen, ob der ICC des Timed-Up-And-Go-Tests ein besseres Ergebnis, als der Gangtest der SPPB erzielt.
- Eine weitere Studie könnte das Thema „Sequenz- und Lerneffekte bei Messwiederholungen“ untersuchen. Generell kann gesagt werden, dass Messwiederholungen als Versuchsdesign als äußerst effiziente Art der Forschung gelten. Der Vorteil dieser Art der Testung ist es, dass eine kleine Gruppe an Versuchspersonen ausreicht, über die Zeit Veränderungen gut messbar sind und die Fehlervarianz, durch die Varianz zwischen den Personen, reduziert wird.

Interessant wäre festzustellen, ob es bei Messwiederholungen zu Sequenz- und Lerneffekten der TestetInnen kommt. Für die Berechnung der Daten würde sich der P-Wert eignen.

6. Literaturverzeichnis

- Barth, C. A., Freivogel, S., Hüter-Becker, A., Hirsch, H. V., Pott, C., & Hirsch, M. A. (2005). *Bewegungsentwicklung und Bewegungskontrolle* (Vol. 2). Stuttgart, Deutschland: Georg Thieme Verlag KG.
- Bernecker, P. (2010). Sturzrisiko und Medikation. *Journal Für Mineralstoffwechsel & Muskuloskelettale Erkrankungen*, 17(2), 52–54.
- Büsching, G. (2015). Short Physical Performance Battery Test - Ein Muss in der Geriatrie. *Physiopraxis*, 13(1), 42–43.
- Curb, J. D., Ceria-Ulep, C. D., Rodriguez, B. L., Grove, J., Guralnik, J., Willcox, B. J., ... Chen, R. (2006). Performance-based measures of physical function for high-function populations. *Journal of the American Geriatrics Society*, 54(5), 737–742. <https://doi.org/10.1111/j.1532-5415.2006.00700.x>
- Duncan, R. P., Leddy, A. L., & Earhart, G. M. (2011). Five times sit-to-stand test performance in Parkinson's disease. *Archives of Physical Medicine and Rehabilitation*, 92(9), 1431–1436. <https://doi.org/10.1016/j.apmr.2011.04.008>
- Flick, U., Kardoff, E. von, & Steinke, I. (2004). *A Companion to Qualitative Research*. SAGE.
- Freiberger, E., de Vreede, P., Schoene, D., Rydwik, E., Mueller, V., Frändin, K., & Hopman-Rock, M. (2012). Performance-based physical function in older community-dwelling persons: a systematic review of instruments. *Age and Ageing*, 41(6), 712–721. <https://doi.org/10.1093/ageing/afs099>
- Freire, A. N., Guerra, R. O., Alvarado, B., Guralnik, J. M., & Zunzunegui, M. V. (2012). Validity and reliability of the short physical performance battery in two diverse older adult populations in Quebec and Brazil. *Journal of Aging and Health*, 24(5), 863–878. <https://doi.org/10.1177/0898264312438551>
- Gómez, J. F., Curcio, C.-L., Alvarado, B., Zunzunegui, M. V., & Guralnik, J. (2013). Validity and reliability of the Short Physical Performance Battery (SPPB). *Colombia Médica: CM*, 44(3), 165–171.
- Guralnik, J. M., Ferrucci, L., Pieper, C. F., Leveille, S. G., Markides, K. S., Ostir, G. V., ... Wallace, R. B. (2000). Lower extremity function and subsequent disability: consistency across studies, predictive models, and value of gait speed alone compared with the short physical performance battery. *The Journals of Gerontology. Series A, Biological Sciences and Medical Sciences*, 55(4), M221-231.

- Guralnik, J. M., Simonsick, E. M., Ferrucci, L., Glynn, R. J., Berkman, L. F., Blazer, D. G., ... Wallace, R. B. (1994). A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. *Journal of Gerontology*, *49*(2), M85-94.
- Horne, M., Skelton, D. A., Speed, S., & Todd, C. (2013). Falls Prevention and the Value of Exercise: Salient Beliefs Among South Asian and White British Older Adults. *Clinical Nursing Research*, *23*(1), 94–110. <https://doi.org/10.1177/1054773813488938>
- Jette, A. M., Jette, D. U., Ng, J., Plotkin, D. J., & Bach, M. A. (1999). Are performance-based measures sufficiently reliable for use in multicenter trials? Musculoskeletal Impairment (MSI) Study Group. *The Journals of Gerontology. Series A, Biological Sciences and Medical Sciences*, *54*(1), M3-6.
- Latham, N. K., Harris, B. A., Bean, J. F., Heeren, T., Goodyear, C., Zawacki, S., ... Jette, A. M. (2014). Effect of a home-based exercise program on functional recovery following rehabilitation after hip fracture: a randomized clinical trial. *JAMA*, *311*(7), 700–708. <https://doi.org/10.1001/jama.2014.469>
- Latham, N. K., Mehta, V., Nguyen, A. M., Jette, A. M., Olarsch, S., Papanicolaou, D., & Chandler, J. (2008). Performance-based or self-report measures of physical function: which should be used in clinical trials of hip fracture patients? *Archives of Physical Medicine and Rehabilitation*, *89*(11), 2146–2155. <https://doi.org/10.1016/j.apmr.2008.04.016>
- Liu, C. K., Leng, X., Hsu, F.-C., Kritchevsky, S. B., Ding, J., Earnest, C. P., ... Fielding, R. A. (2014). The impact of sarcopenia on a physical activity intervention: the Lifestyle Interventions and Independence for Elders Pilot Study (LIFE-P). *The Journal of Nutrition, Health & Aging*, *18*(1), 59–64. <https://doi.org/10.1007/s12603-013-0369-0>
- Liu, J., Tang, W., Chen, G., Lu, Y., Feng, C., & Tu, X. M. (2016). Correlation and agreement: overview and clarification of competing concepts and measures. *Shanghai Archives of Psychiatry*, *28*(2), 115–120. <https://doi.org/10.11919/j.issn.1002-0829.216045>
- Lummel, R. C. van, Walgaard, S., Hobert, M. A., Maetzler, W., Dieën, J. H. van, Galindo-Garre, F., & Terwee, C. B. (2016). Intra-Rater, Inter-Rater and Test-Retest Reliability of an Instrumented Timed Up and Go (iTUG) Test in Patients with Parkinson's Disease. *PLOS ONE*, *11*(3), e0151881. <https://doi.org/10.1371/journal.pone.0151881>
- Luomajoki, H., & Pfeiffer, F. (2018). *Fallbuch Physiotherapie Muskuloskelettales System* (Vol. 1). München, Deutschland: Elsevier Health Sciences.
- Maggio, M., Ceda, G. P., Ticinesi, A., De Vita, F., Gelmini, G., Costantino, C., ... Lauretani, F.

- (2016). Instrumental and Non-Instrumental Evaluation of 4-Meter Walking Speed in Older Individuals. *PLoS One*, 11(4), e0153583. <https://doi.org/10.1371/journal.pone.0153583>
- Mcgraw, K., & Wong, S. P. (1996). Forming Inferences About Some Intraclass Correlation Coefficients. *Psychological Methods*, 1, 30–46. <https://doi.org/10.1037/1082-989X.1.1.30>
- McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia Medica*, 22(3), 276–282.
- Mockford, K. A., Gohil, R. A., Mazari, F., Khan, J. A., Vanicek, N., Coughlin, P. A., & Chetter, I. C. (2014). Effect of supervised exercise on physical function and balance in patients with intermittent claudication. *The British Journal of Surgery*, 101(4), 356–362. <https://doi.org/10.1002/bjs.9402>
- Oesch, P., Hilfiker, R., Keller, S., Kool, J., Luomajoki, H., Schädler, S., ... Leu, C. W. (2011). *Assessments in der Rehabilitation: Band 2: Bewegungsapparat* (2nd ed.). Bern: Hogrefe, vorm. Verlag Hans Huber.
- Olsen, C. F., & Bergland, A. (2017). Reliability of the Norwegian version of the short physical performance battery in older people with and without dementia. *BMC Geriatrics*, 17(1), 124. <https://doi.org/10.1186/s12877-017-0514-4>
- Ostir, G. V., Volpato, S., Fried, L. P., Chaves, P., Guralnik, J. M., & Women's Health and Aging Study. (2002). Reliability and sensitivity to change assessed for a summary measure of lower body function: results from the Women's Health and Aging Study. *Journal of Clinical Epidemiology*, 55(9), 916–921.
- Patel, M. S., Mohan, D., Andersson, Y. M., Baz, M., Samantha Kon, S. C., Canavan, J. L., ... Polkey, M. I. (2014). Phenotypic characteristics associated with reduced short physical performance battery score in COPD. *Chest*, 145(5), 1016–1024. <https://doi.org/10.1378/chest.13-1398>
- Puthoff, M. L. (2008). Outcome Measures in Cardiopulmonary Physical Therapy: Short Physical Performance Battery. *Cardiopulmonary Physical Therapy Journal*, 19(1), 17–22.
- Rodrigues, I. G., Fraga, G. P., & Barros, M. B. de A. (2014). Falls among the elderly: risk factors in a population-based study. *Revista Brasileira De Epidemiologia = Brazilian Journal of Epidemiology*, 17(3), 705–718.
- Roman de Mettelinge, T., & Cambier, D. (2015). Understanding the relationship between walking aids and falls in older adults: a prospective cohort study. *Journal of Geriatric Physical Therapy* (2001), 38(3), 127–132. <https://doi.org/10.1519/JPT.0000000000000031>

Short Physical Performance Battery (SPPB). (n.d.). Retrieved May 31, 2018, from <http://www.nia.nih.gov/research/labs/leps/short-physical-performance-battery-sppb>

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: uses in assessing rater reliability. *Psychological Bulletin*, *86*(2), 420–428.

Statistik Austria. (2017). Bevölkerungsprognosen. Retrieved May 26, 2018, from http://www.statistik.at/web_de/statistiken/menschen_und_gesellschaft/bevoelkerung/demografische_prognosen/bevoelkerungsprognosen/index.html#index1

Steidl, S., & Nigg, B. (2014). *Gerontologie, Geriatrie und Gerontopsychiatrie: Ein Lehrbuch für Gesundheits- und Pflegeberufe*. Wien: Facultas Verlags- und Buchhandels AG.

Stokey, A. D., Katzel, L. I., Steinbrenner, G., Shaughnessy, M., & Ivey, F. M. (2014). The short physical performance battery as a predictor of functional capacity after stroke. *Journal of Stroke and Cerebrovascular Diseases: The Official Journal of National Stroke Association*, *23*(1), 130–135. <https://doi.org/10.1016/j.jstrokecerebrovasdis.2012.11.003>

Telenius, E. W., Engedal, K., & Bergland, A. (2015). Inter-rater reliability of the Berg Balance Scale, 30 s chair stand test and 6 m walking test, and construct validity of the Berg Balance Scale in nursing home residents with mild-to-moderate dementia. *BMJ Open*, *5*(9), e008321. <https://doi.org/10.1136/bmjopen-2015-008321>

Vincent, W. J., & Weir, J. P. (2012). Statistics in Kinesiology. *Human Kinetics Verlage*, *4*, 392.

Wirtz, M., & Caspar, F. (2002). *Beurteilerübereinstimmung und Beurteilerreliabilität: Methoden zur Bestimmung und Verbesserung der Zuverlässigkeit von Einschätzungen mittels Kategoriensystemen und Ratingskalen*. Göttingen: Hogrefe, Verlag für Psychologie.

Wolf, A., & Akesson, K. (2003). *Wolf AD, Akesson K Preventing fractures in elderly people*. *BMJ* *327*(7406): 89-95 (Vol. 327). <https://doi.org/10.1136/bmj.327.7406.89>

A Anhang – Einverständniserklärung

Vereinbarung Physiotherapie

Auszufüllen von dem/der Teilnehmer/in

Name:

Anschrift:

Auszufüllen vom Studiengang Physiotherapie

Gegenstand der Vereinbarung:

Bachelorarbeiten:

„Zahts di? – Jo mir taugts voi“ Motivationspsychologische Aspekte bei heimbasiertem Training mit digitalen Bewegungsprogrammen für die ältere Generation – Lisa Maria Winkler

Einfluss von Schuhwerk auf die Ergebnisse der Short Physical Performance Battery (SPPB) im Vergleich zur Barfußstestung - Arkadiusz Wojciak

Inter- und Intratester-Reliabilität bei manuellen Zeitmessungen der Short Physical Performance Battery (SPPB) – Maria Zamostny

Auszufüllen vom Studiengang Physiotherapie

Folgende Maßnahmen werden von den Studierenden durchgeführt:

- Anleitung von Kraft-, Ausdauer- und Gleichgewichtstests
- Datenerhebung mittels Bodenreaktionskraftplatten und Lichtschranken
- Veranschaulichung digitaler Bewegungsprogramme anhand eines Demonstrationsfilms
- Praktische Einführung in das Fit Daheim Programm
- Datenerhebung zum Thema Motivationspsychologie mittels Fragebogen

Ärztliche Abklärung und Freigabe notwendig?

Nein

Auszufüllen von dem/der Teilnehmer/in

Ich bin damit einverstanden, dass in Zusammenhang mit dieser Teilnahme gemachte Fotos oder Videos von meiner Person

nicht veröffentlicht²

nur anonymisiert und unter Unkenntlichmachung meines Gesichts veröffentlicht

vollständig veröffentlicht

werden dürfen. Diese Zustimmung kann jederzeit widerrufen werden.

Der/die oben angeführte Teilnehmer/in stimmt nachfolgenden, für die Teilnahme erforderlichen Bedingungen zu:

Diese Maßnahmen werden ausschließlich von Studierenden unter Aufsicht bzw. mit Rücksprache von externen BetreuerInnen oder hauptberuflich Lehrenden durchgeführt und ersetzen keine ärztliche Therapie oder Medikamente. Während der Teilnahme ist selbständig auf eine mögliche Überbelastung zu achten. Bei jeglichen Anzeichen ist sofort die Maßnahme abubrechen und der/die betreuende Studierende zu informieren.

Die Teilnahme ist freiwillig, kostenlos und erfolgt gegebenenfalls erst nach ärztlicher Abklärung und Freigabe. In diesem Fall ist dieser Vereinbarung eine Zustimmung der/des behandelnden Ärztin/Arztes beizulegen.

Die Tests werden ausschließlich von Studierenden abgewickelt, die sich noch in Ausbildung befinden. Durch die Betreuung durch hauptberuflich Lehrende ist eine professionelle Abwicklung zwar weitgehend gesichert, für unvorhersehbare Fehler aufgrund des Kenntnisstandes der Studierenden können jedoch keine verbindlichen gesundheitlichen Aussagen gemacht werden und kann die Fachhochschule St. Pölten keine Haftung übernehmen.

Alle vom Teilnehmer/von der Teilnehmerin bekanntgegebenen Informationen und Daten werden seitens der Studierenden und der FH St. Pölten vertraulich behandelt und nicht an unberechtigte Dritte weitergegeben, sofern dafür keine Zustimmung vorliegt.

Die Ergebnisse werden ausschließlich anonymisiert veröffentlicht.

Datum, Unterschrift

² Veröffentlichen: auf Homepage, Kongressen, in Lehre und Forschung innerhalb der FH

B Anhang – Datenblatt

DEMOGRAPHISCHES DATENBLATT:

Datum:

ID:

(Auszufüllen von der Studienleitung)

Vor- und Zuname:

Alter:

Jahre

Geschlecht:

männlich

weiblich

Wohnort:

Körpergröße:

cm

Gewicht:

kg

Schuhgröße:

Dominantes Bein:

links

rechts

(Auszufüllen von der Studienleitung)

Kreuzen Sie bitte das für Sie Zutreffende an!

Wie viel Zeit verbringen Sie am Tag durchschnittlich in Bewegung?

0 – 10 Minuten

11 – 30 Minuten

31 - 60 Minuten

61 - 120 Minuten

mehr als 120 Minuten

Wie viele Schritte legen Sie am Tag durchschnittlich zu Fuß zurück?

0 – 50 Meter

51 - 200 Meter

201 – 500 Meter

501 - 2000 Meter

mehr als 2000 Meter

Wie sicher fühlen Sie sich in Bewegung?

sehr sicher

größtenteils sicher

teilweise sicher

wenig sicher

unsicher

C Anhang – Anleitung Balancetestungen

Die Demonstrationen durch den Testleiter / die Testleiterin findet zeitgleich mit der Testerklärung statt.

Rhombergstand

Der erste Balancetest ist der Rhombergstand. Stellen Sie sich hierfür mit geschlossenen Füßen auf die Markierung. Die Innenkanten der Füße sollen sich dabei berühren und sich auf derselben Höhe befinden. Beim Einnehmen der Position dürfen Sie sich am Sessel anhalten.

Der Test startet mit meinem Kommando „3, 2, 1, los“. Mit dem „Los“ bitte ich Sie den Sessel loszulassen und die Position so lange wie möglich zu halten, bis ich Sie auffordere diese aufzulösen.

Der Test wird frühzeitig beendet, wenn Sie davor den Fuß aus der Position bewegen oder nach dem Sessel greifen.

Haben Sie noch Fragen? ...

Nehmen Sie nun die vorgezeigte Position ein. Sind sie bereit? ...

3, 2, 1, los

...

Danke, die Testung ist beendet. Sie können die Position nun auflösen!

Semitandemstand

Der zweite Balancetest ist der Semitandemstand. Hierfür würde ich Sie bitten die Ferse des vorderen Fußes seitlich an den ersten Zeh des hinteren Fußes zu stellen. Sie dürfen frei wählen, welcher Fuß dabei vorne ist und welcher hinten. Beim Einnehmen der Position dürfen Sie sich am Sessel anhalten.

Der Test startet mit meinem Kommando „3, 2, 1, los“. Mit dem „Los“ bitte ich Sie den Sessel loszulassen und die Position so lange als möglich zu halten, bis ich Sie auffordere diese aufzulösen.

Der Test wird frühzeitig beendet, wenn Sie davor den Fuß aus der Position bewegen oder nach dem Sessel greifen.

Haben Sie noch Fragen? ...

Nehmen Sie nun die vorgezeigte Position ein. Sind sie bereit? ...

„3, 2, 1, los“

...

Danke, die Testung ist beendet. Sie können die Position nun auflösen!

Tandemstand

Der dritte Balancetest ist der Tandemstand. Hierfür würde ich Sie die Ferse von einem Fuß direkt vor die Zehen des Anderen zu stellen. Sie dürfen frei wählen, welcher Fuß dabei vorne ist und welcher hinten. Beim Einnehmen der Position dürfen Sie sich am Sessel anhalten.

Der Test startet mit meinem Kommando „3, 2, 1, los“. Mit dem „Los“ bitte ich Sie den Sessel loszulassen und die Position so lange wie möglich zu halten, bis ich Sie auffordere die Position aufzulösen.

Der Test wird frühzeitig beendet, wenn Sie davor den Fuß aus der Position bewegen oder nach dem Sessel greifen.

Haben Sie noch Fragen? ...

Nehmen Sie nun die vorgezeigte Position ein. Sind sie bereit? ...

3, 2, 1, los

...

Danke, die Testung ist beendet. Sie können die Position nun auflösen!

D Anhang – Anleitung Chair Rising Test

Chair Rising Test

Bei diesem Test wird gemessen, wie lange Sie brauchen, um 5x von einem Sessel aufzustehen und sich wieder hinzusetzen – alles so schnell wie möglich. Gemessen wird die Zeit von der ersten sitzenden Position bis zur letzten sitzenden Position.

Sobald Sicherheitsbedenken auftreten oder die Durchführung das Zeitlimit von einer Minute überschritten hat, wird der Test abgebrochen.

Zuvor würde ich Sie bitten mit Ihren vor der Brust verschränkten Armen vom Sessel einmal aufzustehen und sich wieder niederzusetzen.

...

Vielen Dank – wir können nun mit dem Test starten.

Der Test startet mit meinem Kommando „3, 2, 1, los“.

Haben Sie noch Fragen? ...

Sind sie bereit? ...

3, 2, 1, los

...

Der Testleiter / die Testleiterin zählt die Wiederholungen mit.

...

Danke, die Testung ist beendet.

E Anhang – Anleitung Gangtestung

Gangtestung

Bei der Gangtestung wird die Zeit gemessen, welche Sie benötigen, um eine Strecke von 4 Metern zurückzulegen. Gehen Sie in einer Ihnen gewohnte Geschwindigkeit, so als ob Sie auf der Straße in ein Geschäft gehen würden.

Die Messung endet, wenn Sie mit einem Fuß die Ziellinie komplett überquert haben. Gehen Sie nach Überquerung der Ziellinie bitte noch bis zur zweiten Markierung weiter. Die Durchführung erfolgt zwei Mal, die schnellere Durchführung wird gewertet.

Der Test startet mit meinem Kommando „3, 2, 1, los“.

Haben Sie noch Fragen? ...

Sind sie bereit? ...

„3, 2, 1, los“

...

Vielen Dank – Versuch 1 ist nun beendet. Bitte kehren Sie neben der Gangstrecke zur Startlinie zurück und stellen Sie sich wieder vor die Markierung hin.

Sind sie bereit? ...

„3, 2, 1, los“

...

Vielen Dank – Versuch 2 ist nun ebenso beendet.

F Anhang – SPPB Protocol and Score Sheet

Short Physical Performance Battery Protocol and Score Sheet

Participant Name: _____ Date: _____

All of the tests should be performed in the same order as they are presented in this protocol. Instructions to the participants are shown in bold italic and should be given exactly as they are written in this script.

1. BALANCE TESTS

The participant must be able to stand unassisted without the use of a cane or walker. You may help the participant to get up.

Now let's begin the evaluation. I would now like you to try to move your body in different movements. I will first describe and show each movement to you. Then I'd like you to try to do it, tell me and we'll move on to the next one. Let me emphasize that I do not want you to try to do any exercise that you feel might be unsafe.

do it if y

Do you have any questions before we begin?

A. Side-by-Side Stand

1. ***Now I will show you the first movement.***
2. (Demonstrate) ***I want you to try to stand with your feet together, side-by-side, for about 10 seconds.***
3. ***You may use your arms, bend your knees, or move your body to maintain your balance, but try not to move your feet. Try to hold this position until I tell you to stop.***
4. Stand next to the participant to help him/her into the side-by-side position.
5. Supply just enough support to the participant's arm to prevent loss of balance.
6. When the participant has his/her feet together, ask ***"Are you ready?"***
7. Then let go and begin timing as you say, ***"Ready, begin."***
8. Stop the stopwatch and say ***"Stop"*** after 10 seconds or when the participant steps out of position or grabs your arm.
9. If participant is unable to hold the position for 10 seconds, record result and go to the gait speed test.

B. Semi-Tandem Stand

1. *Now I will show you the second movement.*
2. (Demonstrate) *Now I want you to try to stand with the side of the heel of one foot touching the big toe of the other foot for about 10 seconds. You may put either foot in front, whichever is more comfortable for you.*
3. *You may use your arms, bend your knees, or move your body to maintain your balance, but try not to move your feet. Try to hold this position until I tell you to stop.*
4. Stand next to the participant to help him/her into the semi-tandem position
5. Supply just enough support to the participant's arm to prevent loss of balance.
6. When the participant has his/her feet together, ask **"Are you ready?"**
7. Then let go and begin timing as you say **"Ready, begin."**
8. Stop the stopwatch and say **"Stop"** after 10 seconds or when the participant steps out of position or grabs your arm.
9. If participant is unable to hold the position for 10 seconds, record result and go to the gait speed test.

C. Tandem Stand

1. *Now I will show you the third movement.*
2. (Demonstrate) *Now I want you to try to stand with the heel of one foot in front of and touching the toes of the other foot for about 10 seconds. You may put either foot in front, whichever is more comfortable for you.*
3. *You may use your arms, bend your knees, or move your body to maintain your balance, but try not to move your feet. Try to hold this position until I tell you to stop.*
4. Stand next to the participant to help him/her into the tandem position.
5. Supply just enough support to the participant's arm to prevent loss of balance.
6. When the participant has his/her feet together, ask **"Are you ready?"**
7. Then let go and begin timing as you say, **"Ready, begin."**
8. Stop the stopwatch and say **"Stop"** after 10 seconds or when the participant steps out of position or grabs your arm.

2. GAIT SPEED TEST

Now I am going to observe how you normally walk. If you use a cane or other walking aid and you feel you need it to walk a short distance, then you may use it.

A. First Gait Speed Test

1. ***This is our walking course. I want you to walk to the other end of the course at your usual speed, just as if you were walking down the street to go to the store.***
2. Demonstrate the walk for the participant.
3. ***Walk all the way past the other end of the tape before you stop. I will walk with you. Do you feel this would be safe?***
4. Have the participant stand with both feet touching the starting line.
5. ***When I want you to start, I will say: "Ready, begin."*** When the participant acknowledges this instruction say: ***'Ready, begin.'***
6. Press the start/stop button to start the stopwatch as the participant begins walking.
7. Walk behind and to the side of the participant.
8. Stop timing when one of the participant's feet is completely across the end line.

B. Second Gait Speed Test

1. ***Now I want you to repeat the walk. Remember to walk at your usual pace, and go all the way past the other end of the course.***
2. Have the participant stand with both feet touching the starting line.
3. ***When I want you to start, I will say: "Ready, begin."*** When the participant acknowledges this instruction say: ***'Ready, begin.'***
4. Press the start/stop button to start the stopwatch as the participant begins walking.
5. Walk behind and to the side of the participant.
6. Stop timing when one of the participant's feet is completely across the end line.

Short Physical Performance Battery Protocol and Score Sheet

GAIT SPEED TEST SCORING:

Length of walk test course: Four meters Three meters

A. Time for First Gait Speed Test (sec)

1. Time for 3 or 4 meters _____.____ sec
2. If participant did not attempt test or failed, circle why:
 - Tried but unable 1
 - Participant could not walk unassisted 2
 - Not attempted, you felt unsafe 3
 - Not attempted, participant felt unsafe 4
 - Participant unable to understand instructions 5
 - Other (Specify) _____ 6
 - Participant refused 7
3. Aids for first walk..... None Cane Other

Comments: _____

B. Time for Second Gait Speed Test (sec)

1. Time for 3 or 4 meters _____.____ sec
2. If participant did not attempt test or failed, circle why:
 - Tried but unable 1
 - Participant could not walk unassisted 2
 - Not attempted, you felt unsafe 3
 - Not attempted, participant felt unsafe 4
 - Participant unable to understand instructions 5
 - Other (Specify) 6
 - Participant refused 7
3. Aids for second walk..... None Cane Other

What is the time for the faster of the two walks?
Record the shorter of the two times _____.____ sec
[If only 1 walk done, record that time] _____.____ sec
If the participant was unable to do the walk: 0 points

Short Physical Performance Battery Protocol and Score Sheet

For 4-Meter Walk:

- If time is more than 8.70 sec: 1 point
- If time is 6.21 to 8.70 sec: 2 points
- If time is 4.82 to 6.20 sec: 3 points
- If time is less than 4.82 sec: 4 points

For 3-Meter Walk:

- If time is more than 6.52 sec: 1 point
 - If time is 4.66 to 6.52 sec: 2 points
 - If time is 3.62 to 4.65 sec: 3 points
 - If time is less than 3.62 sec: 4 points
-

3. CHAIR STAND TEST

Single Chair Stand

1. **Let's do the last movement test. Do you think it would be safe for you to try to stand up from a chair without using your arms?**
2. **The next test measures the strength in your legs.**
3. (Demonstrate and explain the procedure.) **First, fold your arms across your chest and sit so that your feet are on the floor; then stand up keeping your arms folded across your chest.**
4. Please stand up keeping your arms folded across your chest. (Record result).
5. If participant cannot rise without using arms, say "Okay, try to stand up using your arms." This is the end of their test. Record result and go to the scoring page.

Repeated Chair Stands

1. **Do you think it would be safe for you to try to stand up from a chair five times without using your arms?**
2. (Demonstrate and explain the procedure): **Please stand up straight as QUICKLY as you can five times, without stopping in between. After standing up each time, sit down and then stand up again. Keep your arms folded across your chest. I'll be timing you with a stopwatch.**
3. When the participant is properly seated, say: **"Ready? Stand"** and begin timing.
4. Count out loud as the participant arises each time, up to five times.
5. Stop if participant becomes tired or short of breath during repeated chair stands.
6. Stop the stopwatch when he/she has straightened up completely for the fifth time.
7. Also stop:
 - If participant uses his/her arms
 - After 1 minute, if participant has not completed rises
 - At your discretion, if concerned for participant's safety
8. If the participant stops and appears to be fatigued before completing the five stands, confirm this by asking **"Can you continue?"**
9. If participant says "Yes," continue timing. If participant says "No," stop and reset the stopwatch.

Short Physical Performance Battery Protocol and Score Sheet

Participant ID: _____ Date: _____ Tester Initials: _____

Scoring for Complete Short Physical Performance Battery

Test Scores

Total Balance Test score _____ points

Gait Speed Test score _____ points

Chair Stand Test score _____ points

Total Score _____ **points** (sum of points above)