

Aus dem Zentrum für Neurologie und Psychiatrie der Universität zu Köln
Abteilung für Medizinische Psychologie
Leiterin: Universitätsprofessorin Dr. rer. nat. E. Kalbe

**Computerisiertes Arbeitsgedächtnistraining bei
Patientinnen und Patienten mit Morbus Parkinson
ohne kognitive Beeinträchtigungen
Effekte auf das Arbeitsgedächtnis und Prädiktion
des Trainingsoutcomes**

Inaugural-Dissertation zur Erlangung der Doktorwürde
der Medizinischen Fakultät
der Universität zu Köln

vorgelegt von
Pia Schröder
aus Damme

promoviert am 09. Juni 2022

Dekan: Universitätsprofessor Dr. med. G. R. Fink

1. Gutachterin: Universitätsprofessorin Dr. rer. nat. E. Kalbe
2. Gutachter: Privatdozent Dr. med. H. Dafsari

Erklärung

Ich erkläre hiermit, dass ich die vorliegende Dissertationsschrift ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe; die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht.

Bei der Auswahl und Auswertung des Materials sowie bei der Herstellung des Manuskriptes habe ich Unterstützungsleistungen von folgenden Personen erhalten:

Dr. Anja Ophey

Weitere Personen waren an der Erstellung der vorliegenden Arbeit nicht beteiligt. Insbesondere habe ich nicht die Hilfe einer Promotionsberaterin/eines Promotionsberaters in Anspruch genommen. Dritte haben von mir weder unmittelbar noch mittelbar geldwerte Leistungen für Arbeiten erhalten, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertationsschrift stehen.

Die Dissertationsschrift wurde von mir bisher weder im Inland noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt.

Die dieser Arbeit zugrunde liegenden Daten wurden in der Abteilung für Medizinische Psychologie | Neuropsychologie & Gender Studies der Uniklinik Köln unter Leitung von Prof. Dr. rer. nat. Elke Kalbe in einem Kooperationsprojekt mit der Abteilung für Multimodale Bildgebung der Uniklinik Köln unter Leitung von Prof. Dr. med. Thilo van Eimeren erhoben. Part III der Unified Parkinson's Disease Rating Scale wurde anhand von Videoaufnahmen in der Klinik und Poliklinik für Neurologie durch Dr. Paul Reker ausgewertet.

Mein Eintritt in die Arbeitsgruppe der Abteilung für Medizinische Psychologie | Neuropsychologie & Gender Studies erfolgte im Oktober 2016 und damit zu Beginn der Durchführung der ersten Testungen. Mein Beitrag zur Durchführung der Studie lag in der Mithilfe bei der Probandenrekrutierung mittels Telefoninterviews sowie der anschließenden Terminvereinbarung für die Testungen, in der selbstständigen Durchführung eines Teils der klinischen und neuropsychologischen Testungen, in der Auswertung eines Teils der neuropsychologischen Tests und Fragebögen sowie im Einpflegen der Ergebnisse in unsere Datenbank.

Zur Entwicklung der Fragestellungen, die in Absprache mit der Studienleitung aufgestellt wurden, erfolgte eine eigenständige Literaturrecherche und eine darauf aufbauende Auswahl der Prädiktoren. Die Extraktion der für die Untersuchung der Fragestellungen relevanten Daten aus der Datenbank, die Auswertung der Daten mittels SPSS (Version 2017) sowie die schriftliche, tabellarische und graphische Darstellung der Ergebnisse wurden selbstständig von mir vorgenommen. Ebenfalls verlief der weitere Verschriftlichungsprozess inklusive Einführung in die Thematik, Einbettung in den aktuellen literarischen Kontext, Beschreibung der verwendeten Materialien und Methoden sowie sich einer dem Ergebnisteil anschließenden kritischen Betrachtung der Ergebnisse unter Bezugnahme zu vorangegangenen Studien und Analysen in Eigenleistung unter regelmäßiger Rücksprache mit der zuständigen Betreuerin.

Die in der Studie gesammelten Daten wurden von Kolleginnen und Kollegen des Teams unter verschiedenen Fragestellungen ausgewertet. Verweise zu den bisherigen Veröffentlichungen zur durchgeführten Studie mit Erstautorenschaften der beiden Studienleiterinnen Dr. Kathrin Giehl und Dr. Anja Ophey finden sich an entsprechenden Stellen.

Erklärung zur guten wissenschaftlichen Praxis:

Ich erkläre hiermit, dass ich die Ordnung zur Sicherung guter wissenschaftlicher Praxis und zum Umgang mit wissenschaftlichem Fehlverhalten (Amtliche Mitteilung der Universität zu Köln AM 132/2020) der Universität zu Köln gelesen habe und verpflichte mich hiermit, die dort genannten Vorgaben bei allen wissenschaftlichen Tätigkeiten zu beachten und umzusetzen.

Köln, den 29.04.2021

Unterschrift:

Danksagung

Ich danke der gesamten Abteilung für Medizinische Psychologie | Neuropsychologie & Gender Studies für die Möglichkeit, eine gute wissenschaftliche Praxis zu erlernen und meine medizinische Doktorarbeit zu verfassen. Im Besonderen danke ich der Leiterin Frau Prof. Dr. rer. nat. Elke Kalbe in ihrer Rolle als Doktormutter für die jahrelange Begleitung des Promotionsprozesses.

Weiterhin möchte ich meine große Dankbarkeit gegenüber der Arbeitsgruppe, in deren Rahmen die Durchführung und Auswertung der Studie stattgefunden hat, ausdrücken: Gegenüber der Mitinitiatorin Sarah Rehberg, M.Sc. für die Aufnahme in die Arbeitsgruppe und die Betreuung im ersten Jahr, Dr. Anja Opey für die Übernahme der Funktion als Betreuerin und insbesondere für den intensiven und unterstützenden Austausch während des Schreibprozesses, meinen Teamkolleginnen Dr. Ann-Kristin Folkerts, Hannah Liebermann-Jordanidis, M.Sc., Kathrin Krugmann, Clara Blome und Jana Nietzsche für die Unterstützung bei der Datenerhebung sowie den Kooperationspartnern Dr. Kathrin Giehl und Univ.-Prof. Dr. Thilo van Eimeren aus der Nuklearmedizin für die Mitinitiation und -arbeit an diesem gemeinsamen Projekt.

Natürlich geht mein aufrichtiger Dank zudem an die Probandinnen und Probanden, die sich zur Teilnahme an der Studie bereiterklärt und uns ihre Zeit und ihr Engagement geschenkt haben. Nicht zuletzt möchte ich NeuroNation (Synaptikon GmbH) für die Bereitstellung ihres Programms für das Arbeitsgedächtnistraining und für die Trainingslizenzen aller Studienteilnehmenden sowie der Schuhfried GmbH für die Bereitstellung des Wiener Testsystems danken.

Inhaltsverzeichnis

Abkürzungsverzeichnis	7
1. Zusammenfassung.....	8
2. Einleitung.....	9
2.1 Morbus Parkinson und Kognition	9
2.2 Arbeitsgedächtnis	9
2.3 Arbeitsgedächtnistraining	10
2.4 Neuronale Veränderungen durch Arbeitsgedächtnistraining	12
2.5 Prädiktion des Ansprechens auf Arbeitsgedächtnistraining	13
2.6 Zusammenfassung und Fragestellung.....	15
3. Material und Methoden	16
3.1 Studiendesign.....	16
3.2 Probanden.....	16
3.2.1 Rekrutierung	16
3.2.2 Schätzung der Stichprobengröße	17
3.2.3 Ein- und Ausschlusskriterien	17
3.3 Ablauf	18
3.3.1 Studienein- und ausschluss	18
3.3.2 Testungen	19
3.3.3 Klinische Verfahren	21
3.3.4 Neuropsychologische Verfahren	22
3.4 Training	24
3.5 Genehmigung, Registrierung und Patienteneinwilligung	27
3.6 Statistische Auswertung.....	27
4. Ergebnisse.....	28
4.1 Studienverlauf bezüglich der Probandenzahl	28
4.2 Stichprobenbeschreibung	29
4.2.1 Baseline-Charakteristika.....	29
4.2.2 Ausgangsniveau der Arbeitsgedächtnistests.....	30

4.3 Trainingseffekte.....	31
4.3.1 Mixed ANOVAs	31
4.3.2 t -Tests für abhängige Stichproben	34
4.4 Prädiktionsanalysen.....	38
5. Diskussion	40
5.1 Ergebnisse	40
5.1.1 Hypothese (a): Effektivität des computerisierten Arbeitsgedächtnistrainings	41
5.1.2 Hypothese (b): Prädiktionsanalysen.....	42
5.2 Stärken und Schwächen	44
5.3 Offene Fragen und Perspektiven	47
5.4 Fazit	48
6. Literaturverzeichnis	49
7. Anhang.....	56
7.1 Tabellenverzeichnis	56
7.2 Abbildungsverzeichnis	57
Zu 4.3.1 Mixed ANOVAs	58
Zu 4.3.2 t -Tests für abhängige Stichproben.....	59
Zu 4.4 Prädiktionsanalysen	61

Abkürzungsverzeichnis

ADHS	Aufmerksamkeitsdefizit-/Hyperaktivitätsstörung
CG	Kontrollgruppe (engl. control group)
CORSI	Corsi-Block-Tapping-Test
CWMT	computerisiertes Arbeitsgedächtnistraining (engl. computerised working memory training)
EG	Interventionsgruppe (engl. experimental group)
ES	Effektstärke
fMRT	funktionelle Magnetresonanztomographie
FU	Follow-Up
GDS	geriatrische Depressionsskala
H&Y	Hoehn & Yahr
IPS	idiopathisches Parkinson-Syndrom
LEDD	Levodopa equivalent daily dose
MCI	leichte kognitive Beeinträchtigung (engl. mild cognitive impairment)
MoCA	Montreal Cognitive Assessment
NBN	N-back nonverbal
NBV	N-back verbal
PD	Morbus Parkinson (engl. Parkinson's disease)
PET	Positronen-Emissions-Tomographie
PFC	präfrontaler Cortex
RCT	randomisierte kontrollierte Studie (engl. randomised controlled trial)
UPDRS	Unified Parkinson's Disease Rating Scale
WM	Arbeitsgedächtnis (engl. working memory)
WMS-R	Wechsler Memory Scale-Revised
WMT	Arbeitsgedächtnistraining (engl. working memory training)
WTS	Wiener Testsystem
WwW-Task	What-was-Where-Task

1. Zusammenfassung

Morbus Parkinson als zweithäufigste neurodegenerative Erkrankung ist sowohl durch motorische als auch vielfältige nicht-motorische Symptome wie kognitive Störungen charakterisiert. Eine der vulnerabelsten kognitiven Domänen stellt das Arbeitsgedächtnis dar, dessen Funktion in der kurzfristigen Speicherung und Manipulation von Informationen liegt. Nach Arbeitsgedächtnistrainings wurden mehrfach Leistungssteigerungen sowie erhöhte Dopaminspiegel in entsprechenden Hirnregionen beobachtet. Dies legt die Fragen nahe, ob ein Arbeitsgedächtnistraining für Patientinnen und Patienten mit Morbus Parkinson positive Effekte bringt und ob bestimmte Charakteristika vorhersagen können, ob eine Person besonders von dem Training profitiert.

Im Rahmen einer randomisierten, kontrollierten Studie wurde ein fünfwöchiges, adaptives, computerisiertes Arbeitsgedächtnistraining bei Probandinnen und Probanden mit Morbus Parkinson ohne kognitive Beeinträchtigungen durchgeführt. Die Probandinnen und Probanden wurden randomisiert in die Interventionsgruppe oder in die passive Kontrollgruppe eingeteilt. Sie wurden zur Baseline, nach der Interventions- bzw. Wartephase, sowie nach weiteren 14 Wochen klinisch und neuropsychologisch untersucht.

In dieser Arbeit wurden die Ergebnisse von sechs elaborierten Arbeitsgedächtnistests ausgewertet, um kurz- und mittelfristige direkte Trainingseffekte auf das verbale und non-verbale Arbeitsgedächtnis zu untersuchen. Anschließend wurde mit den Daten der Interventionsgruppe eine Prädiktionsanalyse vorgenommen, um Aussagen über den Einfluss der Variablen Alter, Baselinenniveau, Bildung, Krankheitsschwere im Sinne motorischer Beeinträchtigung sowie Menge der dopaminergen Medikation auf den Trainingserfolg treffen zu können. Die Trainingseffekte waren klein bis mittelgroß, nur teils signifikant und stärker im verbalen Arbeitsgedächtnis zu beobachten. Es ließ sich ein Trend hin zu mittelfristigen Trainingseffekten erkennen. Bei den Prädiktionsanalysen konnte lediglich das Baselinenniveau als zuverlässiger negativer Prädiktor identifiziert werden.

Wir konnten zeigen, dass computerisiertes Arbeitsgedächtnistraining eine sinnvolle Intervention bei Patientinnen und Patienten mit Morbus Parkinson darstellt, insbesondere bei denjenigen, die ein niedriges Ausgangsniveau im Arbeitsgedächtnis zeigen. Die eher kleinen Trainingseffekte sowie die limitierte Evidenz für Prädiktoren des Ansprechens auf das Training könnten unter anderem durch die Homogenität des Probandenkollektivs, der Auswahl zu wenig sensitiver Arbeitsgedächtnistests und einem zu kurzen Beobachtungszeitraum zur Detektion längerfristiger Trainingseffekte erklärt werden. Wir hoffen, den Weg für weiterführende Studien auf diesem Gebiet mit dem Einsatz sensitiver neuropsychologischer Tests und längeren Beobachtungszeiträumen sowie für eine Implementierung von Arbeitsgedächtnistrainings in den klinischen Alltag geebnet zu haben.

2. Einleitung

2.1 Morbus Parkinson und Kognition

Morbus Parkinson (engl. Parkinson's disease, PD) bzw. das idiopathische Parkinson-Syndrom (IPS) ist mit einer Prävalenz von 0,3% in der Gesamtbevölkerung und steigender Tendenz mit zunehmendem Lebensalter nach der Alzheimer-Erkrankung die zweithäufigste neurodegenerative Erkrankung in industrialisierten Ländern¹. Der Erkrankung liegt neben anderen Alterationen im Transmitterhaushalt des Gehirns vor allem eine Abnahme der dopaminergen Neuroonen in der Pars compacta der Substantia nigra und der damit einhergehende Dopaminmangel zugrunde².

Neben den Kardinalsymptomen Hypo-/Bradykinese, Rigor und Tremor, die die motorischen Symptome beschreiben³, treten auch nicht-motorische Symptome auf, die häufig schon in einem sehr frühen PD-Stadium beginnen⁴. Zu den nicht-motorischen Symptomen zählen Dysfunktionen des autonomen Nervensystems, Schlafstörungen, sensorische Störungen einschließlich Muskel- und Gelenkschmerzen, Veränderungen der Stimmung sowie kognitive Beeinträchtigungen¹. Mit einer gepoolten Prävalenz von 40% (95%-Konfidenzintervall = 36 – 45%) ist die leichte kognitive Beeinträchtigung (engl. mild cognitive impairment, MCI) ein häufiges nicht-motorisches Symptom bei Patienten* mit PD und stellt einen Risikofaktor für die Entwicklung einer Demenz dar⁵. So stellten Litvan et al.⁶ bei 20-60% ihrer Probanden eine dementielle Entwicklung innerhalb von 2-5 Jahren fest und Hely et al.⁷ konnten eine solche Entwicklung in 83% der Fälle innerhalb von 20 Jahren nach Diagnosestellung beobachten. Oftmals zeigen sich multiple Defizite in den verschiedenen kognitiven Domänen, beispielsweise im exekutiven, mnestischen, visuell-räumlichen sowie semantischen Bereich⁸, die teils auch schon früh im Verlauf der Erkrankung oder sogar zum Zeitpunkt der Diagnosestellung zu erkennen sind⁶. Besonders vulnerabel und damit häufig betroffen scheint das Arbeitsgedächtnis (engl. working memory, WM) zu sein^{9,10}.

2.2 Arbeitsgedächtnis

Bereits 1974 wurde das WM durch Baddeley definiert als ein System des Gehirns, in dem Informationen zur Durchführung kognitiver Aufgaben kurzfristig gespeichert und manipuliert würden¹¹. In seiner Arbeit teilte Baddeley das WM in Subsysteme ein: die „phonologische Schleife“ diene der Verarbeitung verbaler, der „visuell-räumliche Notizblock“ der Verarbeitung visuell-räumlicher Informationen. Ein drittes Subsystem stelle die „zentrale Exekutive“ dar, die

* Hinweis: Im Interesse einer besseren Lesbarkeit wird nicht ausdrücklich in geschlechtsspezifischen Personenbezeichnungen differenziert. Die gewählte männliche Form schließt alle Geschlechter gleichberechtigt ein.

die zur Ausführung kognitiver Prozesse notwendige Koordinierung leiste. In einer späteren Arbeit führte Baddeley außerdem ein viertes Subsystem ein, das der zentralen Exekutive als „episodischer Puffer“ diene und zusammenhängende Informationen episodisch einteile¹².

Andere Wissenschaftler vertreten die Hypothese, dass das WM ein Netzwerk multimodaler Neuronen sei¹³, in dem sich die Bereiche zur Durchführung unterschiedlich gearteter WM-Aufgaben überlappten¹⁴. Neuere Studien sehen außerdem einen engen Zusammenhang zwischen dem WM und der Aufmerksamkeit bzw. zwischen dem WM und der fluiden Intelligenz¹⁵. Insgesamt lässt sich festhalten, dass das WM im Gegensatz zum Kurz- und Langzeitgedächtnis nicht nur der Speicherung, sondern auch der Manipulation von Informationen dient, was es eine tragende Rolle in verschiedenen kognitiven Bereichen, etwa bei exekutiven Funktionen, einnehmen lässt¹⁵.

An der Funktionalität des WM sind verschiedene Hirnstrukturen beteiligt: McNab und Klingberg schrieben den Basalganglien, hier insbesondere der Aktivität des Globus pallidus, zusammen mit dem präfrontalen Cortex (PFC) eine filternde Funktion zu und sahen ihre Aufgabe in der Sondierung relevanter Informationen sowie deren Weiterleitung an höhere Hirnstrukturen¹⁶. Die Verbindung dieser dopaminergen Basalganglien und des PFC wird als „frontostriatale Schleifen“ beschrieben⁸. Höher geschaltete Zentren des WM fand man in Bereichen des Frontal- und Parietallappens¹⁷.

Weiterhin unterteilt wird das WM in der Literatur in zwei grundsätzliche Funktionen: das „Updating“ und die Speicherung von Informationen. „Updating“ benennt die Manipulation von Informationen und wird dem Striatum, einem Teil der Basalganglien, zugeordnet¹⁸. Die Hirnregion, der die Speicherfunktion des WM zugeschrieben wird, ist der Sulcus intraparietalis des Parietallappens. Eine weitere Rolle in der Speicherung von Informationen scheint der dorso-laterale PFC einzunehmen, die nicht nur in der Selektion und Weiterleitung von sensorischen Eindrücken im Sinne einer „Bottom-Up“-Kontrolle, sondern zusätzlich in einer „Top-Down“-Kontrolle von höher zu niedriger geschalteten Hirnstrukturen zu liegen scheint¹⁹.

Sowohl die „Updating“- als auch die Speicherfunktion des WM wurden von verschiedenen Autorengruppen untersucht und gemessen, wobei eine begrenzte Kapazität festgestellt wurde^{18,19}. Im Jahr 1956 legte Miller diese Kapazitätsgrenze der im WM zu verarbeitenden Informationen in seinem Aufsatz bei ungefähr sieben Einheiten fest²⁰. Neueren Erkenntnissen zufolge liegt diese Grenze bei ungefähr drei bis fünf Einheiten²¹.

2.3 Arbeitsgedächtnistraining

Das WM und seine Kapazitätsgrenzen stehen seit Langem im Fokus der Wissenschaft, insbesondere mit der Fragestellung, ob die WM-Kapazität statisch oder wie andere Hirnfunktionen modulierbar ist und ob ein Arbeitsgedächtnistraining (engl. working memory training, WMT) auch andere kognitive Funktionen verbessern kann. Es wurden verschiedene Arten von WMT

entwickelt, angewandt und auf ihre Effektivität hin untersucht. Ein häufig verwendeter Ansatz sind N-back-Aufgaben, bei denen ein Stimulus mit demjenigen verglichen werden muss, der N Positionen vorher angezeigt wurde. Ein weiteres Modul stellen Span-Aufgaben dar, wobei es sich eine bestimmte Spanne an z.B. Ziffern zu merken und anschließend wiederzugeben gilt. Des Weiteren wurden komplexere Aufgaben entwickelt, die eine Form von Ablenkung enthalten²².

In einigen Metaanalysen kam man zu dem Schluss, dass ein WMT lediglich kurzfristig positive Effekte zeige²² und dass sich diese Verbesserungen auf trainierte Aufgaben beschränkten²²⁻²⁴. Andere Metaanalysen legen jedoch nahe, dass die WM-Kapazität ebenso wie andere kognitive Funktionen durch gezieltes Training modulierbar ist. So konnten mehrere Befunde eine Steigerung der WM-Kapazität nach einem intensiven, mehrwöchigen, gezielten WMT bei unterschiedlichen Probandengruppen zeigen:

Karbach und Verhaeghen stellten in einer Metaanalyse über 49 Studien signifikante Trainingseffekte sowohl bei jüngeren als auch bei älteren Gesunden fest²⁵. Die Trainingseffekte beschränkten sich nicht nur auf die direkt trainierten WM-Bereiche, sondern wurden auch im Sinne von Nah-Transfereffekten für nicht-trainierte WM-Funktionen sowie in drei von vier verglichenen Fällen im Sinne von Fern-Transfereffekten für andere kognitive Bereiche wie Aufmerksamkeit, fluide Intelligenz, episodisches Gedächtnis und Schnelligkeit signifikant, wenngleich die Effektstärken (ES) kleiner waren als in den direkt trainierten Bereichen. Teixeira-Santos et al. konnten die Ergebnisse für Nah-Transfereffekte eines WMT bei älteren Gesunden reproduzieren²⁶.

Zu ähnlichen Ergebnissen kamen Weicker et al. in ihrer Übersichtsarbeit, in der 103 WMT-Studien vergleichend untersucht wurden²⁷. Zusätzlich zu gesunden Probanden verschiedenen Alters wurden auch solche Studien mit eingeschlossen, die Probanden mit WM-Defiziten untersuchten. Auch in dieser Probandengruppe konnte herausgearbeitet werden, dass ein WMT langfristige Effekte auf die WM-Leistung hat. Nah- und Ferntransfereffekte konnten ebenfalls bestätigt werden.

Eine der Probandengruppen mit WM-Defiziten, mit der bisher Studien zur Wirksamkeit eines WMT durchgeführt wurden, sind Personen mit einem stattgefundenen Schlaganfall. Hier konnten signifikante Effekte in trainierten und nicht-trainierten WM-Bereichen sowie in einer Verbesserung der Aufmerksamkeit festgestellt werden²⁸.

Eine weitere untersuchte Gruppe mit WM-Defiziten stellen Kinder mit der Diagnose Aufmerksamkeitsdefizit-/Hyperaktivitätsstörung (ADHS) dar. Eine schwedische Arbeitsgruppe um Klingberg fand nicht nur signifikante Trainingseffekte in trainierten und nicht-trainierten WM-Aufgaben, sondern außerdem einen signifikanten Effekt auf die motorische (Hyper-)Aktivität der Kinder²⁹. Wenige Jahre später konnte Klingberg seine Ergebnisse in einer größer angelegten Studie zur Wirkung eines WMT bei Kindern mit ADHS reproduzieren³⁰. Eine andere

Autorengruppe um Green et al. konnte die Verbesserungen in WM-Aufgaben nach WMT bei Kindern mit ADHS ebenfalls bestätigen und sah vor allem einen Effekt auf die Aufmerksamkeit³¹.

2.4 Neuronale Veränderungen durch Arbeitsgedächtnistraining

Einige der zum Thema WMT durchgeführten Studien maßen die Effektivität des Trainings nicht ausschließlich anhand des Outcomes der WM-Tests, sondern untersuchten etwaige Veränderungen auf neuronaler Ebene. In einer Studie von Olesen, Westerberg und Klingberg wurde anhand von funktionellen Magnetresonanztomographie (fMRT)-Untersuchungen nach einer abgeschlossenen WMT-Einheit eine Zunahme der Hirnaktivität derjenigen Regionen des Frontal- und Parietallappens, die dem WM zugeordnet werden, gemessen³². In einer weiteren Studie, in der der Einfluss eines WMT auf die Hirnaktivität mittels fMRT untersucht wurde, wurden die Ergebnisse ebenfalls dahingehend gedeutet, dass postinterventionell eine höhere Effizienz auf neuronaler Ebene des WM auszumachen war³³. In dem Review von Constantinidis und Klingberg, in dem verschiedene hirnbildgebende und neurophysiologische Untersuchungen an Primaten und Menschen nach WMT verglichen wurden, konnten diese Ergebnisse bestätigt werden³⁴. Ebenso hielten sie der Metaanalyse von Salmi et al. stand³⁵.

Weitere WMT-Effekte auf neuronaler Ebene wurden anhand von Veränderungen des Transmitterhaushalts im Gehirn hin untersucht. Insbesondere zu nennen sind hier die Untersuchungen des Neurotransmitters Dopamin, dem in der Funktionalität des WM eine tragende Rolle zugeschrieben wird⁸. Mithilfe von Positronen-Emissions-Tomographien (PET) konnte ein vermindertes Bindungspotenzial von kortikalen D1-Rezeptoren in den dem WM zugeordneten Bereichen des Frontal- und Parietallappens nach einem mehrwöchigen WMT nachgewiesen werden³⁶. Als eine zugrundeliegende Theorie wurde eine durch das WMT hervorgerufene Erhöhung des Dopaminspiegels aufgestellt. In einer anderen Studie wurde ebenfalls die PET-Methode genutzt, um biochemische Veränderungen des Dopaminhaushalts nach einem mehrwöchigen Training der „Updating“-Funktion des WM festzustellen. Hier zeigte sich ein erniedrigtes Bindungspotenzial von striatalen D2-Rezeptoren, was wiederum mit einer Erhöhung des verfügbaren Dopamins in Verbindung gebracht wurde³⁷. Auch die Veränderungen im dopaminergen System wurden im Review von Constantinidis und Klingberg aufgegriffen und fanden hier Bestätigung³⁴.

Interessant sind diese Beobachtungen für verschiedene Personengruppen. So stellten Kinder mit ADHS schon mehrfach das Probandenkollektiv für WMT-Studien²⁹⁻³¹, da die ADHS-Symptomatik und die damit einhergehenden WM-Defizite mit Alterationen des Dopaminhaushalts im Gehirn in Verbindung gebracht werden. Die Verbesserungen in den durchgeführten Tests sowie die Abmilderung der ADHS-Symptomatik lassen einen Zusammenhang mit der erhöhten Ausschüttung von Dopamin durch das WMT vermuten³¹.

Auch der PD-Symptomatik liegt ein Dopaminmangel zugrunde. Die WM-Defizite konnten in einer Studie an Probanden mit einer neu diagnostizierten PD-Erkrankung mittels fMRT mit einer verminderten Konnektivität der dem WM zugehörigen Hirnzentren in Verbindung gebracht werden. Die schlechtere Netzwerkverbindung wurde auf eine verminderte Ausschüttung des Transmitters Dopamin zurückgeführt³⁸. Diese Ergebnisse sind insofern interessant, als dass sich die Frage stellt, inwieweit ein WMT für diese Gruppe effektiv sein könnte.

Trotz der vielversprechenden Links zwischen WMT und PD wurde bislang nur eine randomisierte kontrollierte Studie (engl. randomised controlled trial, RCT) veröffentlicht, die die Wirksamkeit von WMT bei Probanden mit PD untersuchte³⁹. Die Befunde konnten eine Verbesserung in Testergebnissen bei trainierten und nicht-trainierten WM-Aufgaben, insbesondere in solchen, die das „Updating“ erfordern, zum Zeitpunkt der Nachtestung feststellen. Darüber hinaus fanden sie insgesamt eine Milderung der Depressionssymptomatik in der Trainingsgruppe. Jedoch zeigte das Studiendesign einige Schwächen: Die Probanden wurden lediglich mit kognitiven Screening-Tools charakterisiert und weder motorisch noch kognitiv elaboriert, die Studie verlief ausschließlich online und es gab keine Follow-Up (FU)-Untersuchung.

Patienten mit PD scheinen also eine vielversprechende Probandengruppe für die Durchführung eines WMT zu sein. Da viele Betroffene schon in einem sehr frühen Stadium, oftmals schon bei Diagnosestellung, WM-Defizite aufweisen³⁸, liegt eine Intervention noch vor einer diagnostizierbaren MCI nahe. Die bisher noch recht schwache Datenlage in diesem Feld lässt Raum für weitere Untersuchungen.

2.5 Prädiktion des Ansprechens auf Arbeitsgedächtnistraining

Es stellt sich die Frage, ob es in dieser trotzdem weiterhin heterogenen PD-Patientengruppe Personen gibt, die besonders stark von einem WMT profitieren könnten. In der Literatur werden verschiedene personenbezogene und klinische Prädiktoren für das Ansprechen auf WMT bzw. den Trainingserfolg diskutiert. So fanden Fernández-Ballesteros et al. einen positiven Einfluss von einem höheren Bildungsstand auf den Trainingserfolg bei älteren Gesunden sowie Probanden mit MCI⁴⁰. Dahlin, Nyberg et al. beobachteten einen größeren Erfolg eines WMT bei jüngeren Probanden, insbesondere in Bezug auf Transfereffekte⁴¹. Dahingegen konnte in einer anderen, vergleichenden Studie kein signifikanter Unterschied zwischen verschiedenen Altersgruppen festgestellt werden²⁵.

Neben Bildung und Alter der Probanden lässt sich das Baselinenniveau der verschiedenen Tests als personenbezogener Prädiktor aufnehmen. Die Richtung der Prädiktion ist in der Literatur umstritten, was das „Magnification and Compensation“-Modell veranschaulicht⁴². Im „Magnification“-Ansatz wird das Baselinenniveau als positiver Prädiktor für das Trainingsoutcome benannt. Das bedeutet, dass sich Probanden mit einem höheren Ausgangsniveau durch

ein kognitives Training stärker verbessern als solche mit einem niedrigeren Score zur Baseline, was mit überlegenen neurobiologischen Faktoren wie einer höheren neurologischen Plastizität erklärt wird. Dadurch soll sich der Unterschied im Abschneiden der Tests zwischen einer leistungsstärkeren und einer -schwächeren Gruppe nach dem Training noch vergrößern⁴². Als Beispiel einer diese Hypothese bejahenden Studie lässt sich die Arbeit von Nemmi et al. anführen⁴³. Weicker et al. sahen jedoch ein höheres Verbesserungspotenzial bei Probanden mit einem niedrigeren Ausgangsscore, was eher mit dem „Compensation“-Ansatz in Einklang zu bringen ist²⁷. Dieser geht davon aus, dass ein niedrigeres Baselinenniveau positiv auf das Trainingsoutcome wirkt, was mit einem größeren Spielraum für Verbesserungen bei Probanden mit niedrigerem Ausgangsniveau erklärt wird⁴².

Potenzielle klinische Prädiktoren sind unter anderem die PD-bezogene dopaminerge Medikation. Die „dopamine overdose“-Hypothese lässt annehmen, dass die vermutete, durch das WMT hervorgerufene erhöhte Menge an verfügbarem Dopamin zusätzlich zu der „Levodopa equivalent daily dose“ (LEDD), das heißt zu der täglich eingenommenen Dosis an medikamentösem L-Dopa, einen Überschuss an Dopamin im frontostriatalen Netzwerk verursachen könnte. Die daraus folgende Überstimulation der dem WM zugeschriebenen Regionen könnte zu Einbußen der Funktionalität des WM führen. Dies würde folglich einen negativen Einfluss von der LEDD auf den Trainingserfolg bedeuten⁴⁴.

Ein weiterer klinischer Prädiktor ist die Krankheitsschwere. Weicker et al. arbeiteten in ihrer Metaanalyse heraus, dass Probanden mit Hirnverletzungen mehr von einem WMT profitierten als gesunde²⁷. In einer anderen Studie wiesen Probanden mit MCI jedoch eine langsamere Lernkurve auf als die kognitiv unbeeinträchtigten Teilnehmer⁴⁰.

Insgesamt lässt sich festhalten, dass Prädiktoren bei WMT bereits gängiger Gegenstand der Forschung sind, sich jedoch noch keine richtungsweisenden Ergebnisse herauskristalisieren konnten. Gleiches stellten Ophay, Roheger et al. in ihrem Review fest, die 16 Studien zur Prädiktionsanalysen bei WMT mit älteren gesunden Probanden vergleichend untersuchten⁴⁵. Die wenigen Studienergebnisse zur Prädiktion des Outcomes eines kognitiven Trainings bei Probanden mit PD sind ebenfalls uneinheitlich. Edwards et al. fanden einen positiven Einfluss von der LEDD auf den Trainingserfolg, der jedoch unterhalb des Signifikanzniveaus blieb⁴⁶. Außerdem zeigten signifikant gewordene Regressionen, dass länger betroffene bzw. jünger erkrankte Probanden den größten Nutzen aus dem kognitiven Training zogen. Die motorische Beeinträchtigung gemessen anhand der Hoehn & Yahr-Skala (H&Y; Part V der Unified Parkinson's Disease Rating Scale⁴⁷, UPDRS) ließ sich nicht als signifikanter Prädiktor ausmachen.

Fellman et al. sahen im Gegensatz dazu keinen Einfluss von der LEDD auf das Outcome des durchgeführten WMT³⁹. Außerdem gegensätzlich zeigten sich die Ergebnisse des Prädiktors

Länge der Erkrankung, da das WMT einen größeren Erfolg bei kürzer Erkrankten ergab. Zudem arbeiteten sie einen höheren kognitiven Status zur Baseline als positiven Prädiktor für den Trainingserfolg heraus. Zimmermann et al. wiederum sahen eine geringere Verbesserung durch kognitives Training bei höherem WM-Status zur Baseline⁴⁸.

2.6 Zusammenfassung und Fragestellung

Zusammenfassend lässt sich sagen, dass ein intensives, mehrwöchiges, adaptives WMT eine sinnvolle Maßnahme erscheint, um die Kapazität des WM zu steigern. Verschiedene Studien in Nicht-PD-Populationen legen nahe, dass ein WMT nicht nur die trainierten WM-Leistungen verbessern, sondern sich durch Nah- und Fern-Transfereffekte auch auf nicht-trainierte WM-Bereiche sowie anderen kognitive Funktionen positiv auswirken könnte. Durch FU-Untersuchungen konnten mittel- bzw. langfristige Effekte eines solchen Trainings ebenfalls teilweise erfasst werden^{25,27}.

Defizite in der Funktionalität des WM scheinen durch eine dopaminerge Unterversorgung in den entsprechenden Regionen des Gehirns verursacht zu werden. Durch ein WMT konnten nachweislich höhere Spiegel an verfügbarem Dopamin in ebendiesen Hirnregionen erzielt werden^{36,37}.

Da PD neuronal durch dopaminerge Defizite und Veränderungen in dem WM assoziierten kortikalen Netzwerken charakterisiert ist, liegt die Vermutung nahe, dass diese Gruppe von einem intensiven WMT profitieren würde. Eine möglichst frühe Intervention erscheint wegen der schon im Prodromalstadium auftretenden kognitiven Symptomatik und frühen neuronalen Veränderungen¹⁰ sinnvoll, insbesondere auch im Hinblick auf eventuelle präventive Möglichkeiten. Um die kurz- und mittelfristige Wirksamkeit eines computerisierten WMT (engl. computerised working memory training, CWMT) bei Patienten mit PD ohne kognitive Einschränkungen zu untersuchen, wurde für diese Analysen eine RCT inklusive FU-Untersuchung durchgeführt⁴⁹. Die Probanden wurden neuropsychologisch nach elaborierten Kriterien ausgewählt.

In dieser Doktorarbeit wird ein Teilbereich der groß angelegten RCT fokussiert. Es werden folgende zwei Hypothesen verfolgt: Wir vermuten, dass (a) Patienten mit PD ohne kognitive Einschränkungen nach Level-II-Kriterien⁵⁰ kurz- und mittelfristig von einem mehrwöchigen, intensiven, adaptiven CWMT im Vergleich zu einer passiven Kontrollgruppe sowohl im verbalen als auch non-verbalen WM profitieren und untersuchen (b) die Variablen Alter, Baselinenniveau der WM-Tests, Bildung, Krankheitsschwere im Sinne der motorischen Beeinträchtigung sowie LEDD als Prädiktoren des Trainingserfolgs.

3. Material und Methoden

3.1 Studiendesign

Die Studie entsprach einer einfach-verblindeten RCT mit den drei Erhebungszeitpunkten Vor-, Nach- und FU-Testung. Nach Durchführung der Vortestung wurden die Probanden randomisiert der Interventionsgruppe (engl. experimental group, EG) oder Kontrollgruppe (engl. control group, CG) zugeteilt. Dazu wurde eine Randomisierungsliste verwendet, die zuvor von einer ansonsten nicht an der Studie beteiligten Person über das Online-Tool Research-Randomizer (<https://www.randomizer.org>) erstellt und während der Studie verwaltet wurde. Die Studienleiter waren zu jedem Zeitpunkt der von ihnen durchgeführten Testungen in Bezug auf die Gruppenzugehörigkeit der Probanden verblindet. Die Probanden selbst erfuhren nach der Vortestung ihre Zugehörigkeit zur EG oder CG und waren somit ab diesem Zeitpunkt nicht mehr verblindet.

Neben den oben genannten drei Erhebungszeitpunkten, die dem Modul „Neuropsychologie“ angehörten, beinhaltete die Studie außerdem das optionale Modul „Bildgebung“. In diesem wurde während einer jeweils separat stattfindenden Vor- und Nachttestung die Hirnaktivität mittels fMRT erfasst, um mögliche neuronale Veränderungen zu untersuchen. Ergebnisse dieser Untersuchungen werden in dieser Arbeit nicht berichtet, finden sich jedoch an anderer Stelle⁵¹. Das Studiendesign wird in unten stehender Tabelle 1 veranschaulicht.

Tabelle 1. Studiendesign

Modul	Erhebungszeitpunkt			
	Vortestung	Training (EG) oder Warteliste (CG)	Nachttestung	FU-Testung
Neuropsychologie	X		X	X
Bildgebung	(X)		(X)	
Woche	1	2-6	7	21

Notiz: FU, Follow-Up.

3.2 Probanden

3.2.1 Rekrutierung

Die Studie fand im Zeitraum von September 2016 (erster Probandeneinschluss) bis Juli 2018 (letzte FU) statt. Die Kontaktaufnahme erfolgte über Telefonlisten des Kölner Parkinson Netzwerkes, das Auslegen von Flyern und den Aushang von Plakaten in Praxen niedergelassener Neurologen in der Umgebung sowie über Besuche von Parkinsongruppen durch die Studienleitung. Außerdem wurde Informationsmaterial durch die Deutsche Parkinson Vereinigung e.V. an Patienten mit PD übermittelt.

3.2.2 Schätzung der Stichprobengröße

Um die erforderliche Stichprobengröße zu berechnen, wurde eine a-priori Poweranalyse mit G*Power⁵² durchgeführt. Da noch keine zuverlässigen ES für die Effektivität eines WMT bei Patienten mit PD ohne MCI vorlagen, wurde sich zur Schätzung der Stichprobengröße an einer Metaanalyse zu einem allgemeinen kognitiven Training bei Patienten mit PD orientiert⁵³. In dieser Metaanalyse von vier Studien, in denen ein kognitives Training mit unterschiedlich großen Probandenkollektiven durchgeführt worden war, zeigte sich bezogen auf die WM-Leistung eine kombinierte ES $g = .74$ (95% Konfidenzintervall 0.32 – 1.17, $p < .001$), die nach den Richtlinien von Cohen⁵⁴ als moderat und statistisch signifikant eingeordnet werden kann. Nach Umrechnung der ES g ⁵⁵ in das ES-Maß f ergab sich eine ebenfalls moderate ES $f = .37$ ⁵⁴. Mit dieser ES, einer Power von 80%, einer Fehlerwahrscheinlichkeit $\alpha = .05$ und einer Dropout-Schätzung von 20% ergab sich über G*Power⁵² ein geschätzter erforderlicher Stichprobenumfang von insgesamt $N = 72$ Probanden und somit je $n = 36$ Probanden in der EG und CG.

3.2.3 Ein- und Ausschlusskriterien

Zur Auswahl des Probandenkollektivs wurden vorab Ein- und Ausschlusskriterien festgelegt. Diese Kriterien sind in Tabelle 2 zur Übersicht zusammenfassend aufgeführt:

Tabelle 2. Zur Studienteilnahme erforderliche Ein- und Ausschlusskriterien

Einschlusskriterien	Ausschlusskriterien
Diagnose: IPS	Kognitive Beeinträchtigung: PD-MCI oder PD-Demenz
Alter: 45-85 Jahre	Andere diagnostizierte neurologische Erkrankungen
Unbeeinträchtigte/ausreichend korrigierte Seh- und Hörfähigkeit	Schwere depressive Symptomatik (GDS ≥ 11)
Ausreichende Mobilität	Andere akute psychiatrische Erkrankungen
Unterschreiben der Einwilligungserklärung	Körperliche Beeinträchtigungen, die die Bedienung eines Computers verhindern
	Einsatz einer Tiefen Hirnstimulation
	Lebensbedrohliche Erkrankung

Notiz. GDS, geriatrische Depressionsskala; IPS, idiopathisches Parkinson-Syndrom; MCI, leichte kognitive Beeinträchtigung (engl. mild cognitive impairment); PD, Morbus Parkinson (engl. Parkinson's disease).

Zu den Einschlusskriterien zählten die Diagnose eines IPS⁵⁶, ein Alter zwischen 45 und 85 Jahren, eine unbeeinträchtigte bzw. ausreichend korrigierte Seh- und Hörfähigkeit, eine ausreichende Mobilität, um zu den Testungen in die Uniklinik gelangen zu können, sowie das Unterschreiben der Einwilligungserklärung.

Die Ausschlusskriterien umfassten kognitive Beeinträchtigungen nach den Level-II-Diagnosekriterien für PD-MCI⁵⁰ oder PD-Demenz⁵⁷, die Diagnose einer anderen neurologischen, akuten psychiatrischen oder lebensbedrohlichen Erkrankung sowie eine schwere depressive Symptomatik (geriatrische Depressionsskala⁵⁸, GDS \geq 11). Außerdem schloss eine körperliche Beeinträchtigung, die die Bedienung eines Computers verhinderte, und der Einsatz einer Tiefen Hirnstimulation die Studienteilnahme aus.

3.3 Ablauf

3.3.1 Studienein- und ausschluss

Die oben aufgeführten Kriterien, wie demographische Daten, PD-Diagnose, Komorbiditäten, Mobilität und die Ausstattung mit einem Computer, wurden soweit möglich bereits in einem Telefoninterview erfragt. Erfüllte der Studieninteressent alle Einschluss- und verneinte alle Ausschlusskriterien, wurde die Person zu einer Eingangsvsiste an der Uniklinik Köln eingeladen, um die Studieneignung zu prüfen. Hier wurde das schriftliche Einverständnis der Studieninteressenten eingeholt sowie ihr klinischer und neuropsychologischer Status evaluiert.

Die klinische Beurteilung erfolgte anhand einer Anamnese zum allgemeinen gesundheitlichen Zustand sowie zur PD-Diagnose im Speziellen. Außerdem wurde eine Untersuchung der Motorik vorgenommen.

Die Kognition der Interessenten wurde anhand einer neuropsychologischen Testbatterie evaluiert. Hierzu wurden die PD-spezifischen Kriterien der Level-II-Diagnostik zur Detektion einer MCI der MDS Task Force⁵⁰ angewandt. Es wurde die Leistung in den Bereichen Gedächtnis, exekutive Funktion, Aufmerksamkeit, visuell-räumliche Funktion sowie Sprache überprüft.

Stellte sich nach dieser Eingangsvsiste eine Studieneignung heraus, d.h. wick die Leistung in höchstens einem Test maximal 1.5 Standardabweichungen von der Norm ab, wurde die Person als Proband in die Studie aufgenommen und im Anschluss der restlichen Vortestung randomisiert der EG oder CG zugeteilt.

Die Probanden der EG erhielten bei Bedarf innerhalb von sieben Tagen einmalig Besuch von der Studienleitung, um den Zugang zum CWMT am eigenen oder an einem Leihgerät einrichten zu lassen oder sie richteten den Zugang eigenständig ein. In dem sich maximal sieben Tage nach der ersten Testung anschließenden fünfwoöchigen Interventionszeitraum führten die Probanden der EG zwischen der Vor- und Nachttestung ein fünfwoöchiges CWMT durch, wohingegen die CG keiner Intervention unterzogen wurde. Anschließend folgte die Nachttestung, die der Erfassung der unmittelbaren Trainingseffekte diente, in einem Zeitraum von maximal sieben Tagen nach Beendigung der Intervention. Nach weiteren 14 Wochen fand abschließend eine FU-Testung statt, die die mittelfristigen Trainingseffekte erfassen sollte.

Aus ethischen Gründen wurde der CG nach Beendigung der Datenerhebungsphase ein Zugang zum CWMT gegeben. Außerdem erhielten beide Gruppen als Dankeschön eine sechsmonatige Lizenz für das volle kognitive Trainingsprogramm der Internetplattform NeuroNation (Synaptikon GmbH, Berlin, Germany, <https://www.neuronation.com/>).

3.3.2 Testungen

Die Testungen fanden an drei Erhebungszeitpunkten statt. An den jeweiligen Erhebungszeitpunkten wurde eine Fülle von Testungen durchgeführt, die eine Beurteilung des klinischen und kognitiven Status der Probanden ermöglichen sollte. Außerdem wurden den Probanden nach Abschluss des jeweiligen Testtages Fragebögen ausgehändigt, die sie zu Hause ausfüllen und innerhalb von vier Tagen zurückschicken sollten.

Für die Beurteilung des motorischen Status wurden die Skalen III und V der UPDRS⁴⁷ verwendet. Die neuropsychologische Testbatterie bestand aus dem Montreal Cognitive Assessment⁵⁹ (MoCA), vier Subtests des Wiener Testsystems^{60–62} (WTS), dem Subtest „Zahlenspanne“ der Wechsler Memory Scale-Revised⁶³ (WMS-R), der nach Vorbild von Pertzov und Kollegen^{64–66} für die Studie eigens entworfenen What-was-Where-Task (WwW-Task), den Subtests 4 und 7 des Leistungsprüfungssystems⁶⁷ bzw. der Version 50+ für über 50-Jährige⁶⁸, der kompletten Testbatterie des Consortium to Establish a Registry for Alzheimer Diagnosis-Plus⁶⁹, dem Stroop-Interferenz-Test⁷⁰, dem Brief Test of Attention⁷¹ sowie dem Subtest „auditives Sprachverständnis“ der Aphasie-Check-Liste⁷². Für die Fragebögen wurden die GDS⁵⁸, die Skala zur Selbstwirksamkeit⁷³, der Freezing of Gait-Questionnaire⁷⁴, der Pittsburgh Schlafqualitätsindex⁷⁵, der Subjective Cognitive Impairment-Questionnaire (modifiziert nach Jessen et al.⁷⁶), die Skala der Every Day Cognition-12⁷⁷ sowie für die EG ein für die Studie konzipiertes Trainingstagebuch herangezogen.

Zusammenfassend sind alle motorischen und neuropsychologischen Testungen sowie Fragebögen, die im Rahmen der Studie zum Einsatz kamen, tabellarisch aufgeführt (Tabellen 3 und 4). Der Tabelle 3 kann außerdem entnommen werden, welche Tests für den Ausschluss kognitiver Beeinträchtigungen nach Level-II-PD-MCI-Diagnosekriterien⁵⁰ verwendet wurden.

Tabelle 3. Übersicht aller klinischen und neuropsychologischen Verfahren

Kategorie	Subkategorie	Verfahren	Subtest
Klinik			
Motorik		UPDRS	Skala III: Motorik Skala V: H&Y-Skala
Neuropsychologie			
Kognitiver Gesamtstatus		MoCA*	
Arbeitsgedächtnis	Verbal	WTS WMS-R	NBV Zahlenspanne -vorwärts -rückwärts
	Non-verbal/ Visuell-räumlich	WTS WwW-Task	NBN CORSI Blockspanne -vorwärts -rückwärts
Exekutive Funktionen	Logisches Schlussfolgern	LPS (50+)	Subtest 4
	Mentale Flexibilität	CERAD+	TMT-B TMT-B/A+
	Wortflüssigkeit	CERAD+	Semantische Wortflüssigkeit+ Phonetische Wortflüssigkeit
Gedächtnis	Verbal	CERAD+	Wortliste Lernen Wortliste verzögerte Abfrage+ Wortliste wiedererkennen
	Non-verbal	CERAD+	Figuren verzögerte Abfrage+
Aufmerksamkeit	Mentale Verarbeitungsgeschwindigkeit	CERAD+ Stroop	TMT-A Farbwort lesen Farbe benennen
	Inhibition	Stroop	Interferenz+
	Geteilte Aufmerksamkeit	BTA	BTA Total+
Visuell-räumliche Funktionen	Visuo-Konstruktion	CERAD+	Figuren abzeichnen+
	Räumliche Rotation	LPS (50+)	Subtest 7* +
Sprache	Benennen	CERAD+	BNT+
	Verstehen	ACL	Auditives Sprachverständnis* +

Notiz. ACL, Aphasie-Check-Liste; BNT, Boston Naming Test; BTA, Brief Test of Attention; CERAD+, Consortium to Establish a Registry for Alzheimer Diagnosis-Plus; CORSI, Corsi-Block-Tapping-Test; H&Y, Hoehn & Yahr; LPS, Leistungsprüfungssystem, Version 50+ für Probanden ≥ 50 Jahre; MoCA, Montreal Cognitive Assessment; NBN, N-back nonverbal; NBV, N-back verbal; TMT-A, Trial Making Test Version A; TMT-B, Trial Making Test Version B; TMT- B/A, Quotient des Wertes der B-Version und A-Version; UPDRS, Unified Parkinson's Disease Rating Scale; WMS-R, Wechsler Memory Scale-Revised; WTS, Wiener Testsystem; WwW-Task, What was Where-Task.

* Durchführung nur einmalig in der Vortestung.

+ Relevant zum Ausschluss leichter kognitiver Beeinträchtigungen bei Morbus Parkinson⁵⁰ (PD-MCI).

Tabelle 4. Übersicht aller Fragebögen

Kategorie	Subkategorie	Verfahren
Depression		GDS
Selbstwirksamkeit		SWE*
Freezing of Gait		FOG-Q
Schlafqualität		PSQI
Subjektiv kognitive Beschwerden		SCI-Q
Geistige Leistungen im Alltag		ECog-12 (Auszufüllen durch Angehörige)
Trainingstagebuch	Motivation	6-Punkt Likert-Skala (Auszufüllen nur durch EG während des WMT)
	Zufriedenheit	

Notiz. ECog-12, Every Day Cognition; EG, Interventionsgruppe (engl. experimental group); FOG-Q, Freezing of Gait-Questionnaire; GDS, geriatrische Depressionsskala; PSQI, Pittsburgh Schlafqualitätsindex; SCI-Q, Subjective Cognitive Impairment-Questionnaire; SWE, Selbstwirksamkeitserwartungsfragebogen; WMT, working memory training.

* Durchführung nur einmalig in der Vortestung.

Es folgt eine nähere Erläuterung derjenigen Verfahren, die in dieser Arbeit in der statistischen Auswertung berücksichtigt werden.

3.3.3 Klinische Verfahren

UPDRS III

Die motorische Untersuchung erfolgte anhand des Part III der UPDRS⁴⁷. In dem Bogen wird zunächst die Medikation des Probanden erfragt. Hier wird speziell auf eine mögliche Levodopa-Einnahme eingegangen. Anschließend folgt die Untersuchung zur klinischen Beurteilung. Sie wurde während der Testungen mittels Videoaufnahme protokolliert, um sie durch einen Neurologen, der zur Auswertung bezüglich der Zugehörigkeit der Probanden zur EG oder CG sowie zum Erhebungszeitpunkt Vor-, Nach- oder FU-Testung verblindet war, analysieren zu lassen. Die Testleitung gab den Probanden Anweisungen zur Durchführung verschiedener Bewegungen. Durch die Beobachtung dieser Bewegungen lassen sich Aussagen zu Sprache, Gesichtsausdruck, Bewegungen der Extremitäten, Aufstehen, Gangbild, posturaler Instabilität, Körperhaltung sowie Halte- und Ruhetremor der Hände treffen. Außerdem beinhaltet die Untersuchung eine Rigortestung des Nackens und der Extremitäten, die die Testleitung durchführte und selbst bewertete.

Die sich ergebene Summe aus den Werten der einzelnen Bereiche gibt die Schwere der motorischen Beeinträchtigung an (Score 0 – 132). Anhand dieser wurde die Krankheitsschwere bemessen.

UPDRS V

Die Probanden wurden außerdem anhand des Part V der UPDRS, der H&Y-Skala⁴⁷, einer Stufe zugeordnet. Die Einteilung reicht hierbei von 0 bis 5 mit aufsteigender motorischer Symptomatik und der Unterscheidung zwischen symmetrischem und asymmetrischem bzw. beid- und einseitigem Befall.

LEDD

Zur Berechnung der LEDD wurde den Probanden nach der Vortestung ein Bogen ausgeteilt, in dem sie ihre aktuelle Medikation mit Levodopa angeben sollten. Die LEDD wurde anschließend anhand der von Tomlinson et al. eingeführten Formel berechnet⁷⁸.

Sowohl bei der Nach- als auch bei der FU-Testung erfragte die Testleitung etwaige Änderungen in der Medikation. In einem solchen Fall wurde die LEDD dementsprechend für den jeweiligen Erhebungszeitpunkt erneut ermittelt.

3.3.4 Neuropsychologische Verfahren

Für die statistische Auswertung dieser Arbeit wurden die Ergebnisse der durchgeführten etablierten WM-Tests verwendet. Die Ergebnisse der WwW-Task wurden nicht berücksichtigt, da es sich hierbei um ein exploratives Verfahren handelt. Eine separate Betrachtung dieses Verfahrens nahmen Giehl, Ophey, Reker et al. vor⁷⁹.

Wechsler Memory Scale-Revised – Subtest „Zahlenspanne“

Der Subtest „Zahlenspanne“ der WMS-R⁶³ beinhaltet zwei Teile. Im ersten Teil, „Zahlenspanne vorwärts“, werden dem Probanden Zahlenfolgen vorgelesen, beginnend mit drei aufeinander folgenden Zahlen. Jeder Durchgang besteht aus zwei Versuchen mit verschiedenen Zahlenfolgen, jeweils mit der gleichen Anzahl an Ziffern. Wird mindestens eine der beiden Folgen korrekt wiederholt, wird mit dem Vorlesen der Zahlenfolgen des nächsten Durchgangs begonnen, jeweils gesteigert durch eine zusätzliche Ziffer. Der erste Teil wird beendet, sobald beide Zahlenfolgen eines Durchgangs nicht korrekt wiederholt werden. Es werden höchstens sechs Durchgänge bis zu einer Zahlenspanne von acht aufeinander folgenden Ziffern durchgeführt. Im zweiten Teil sollen die vorgelesenen Zahlenfolgen in umgekehrter Reihenfolge wiedergegeben werden. Der erste Durchgang beginnt mit einer Spanne von zwei Zahlen, aufsteigend bis zu sieben aufeinander folgenden Ziffern im sechsten Durchgang. Auch hier wird der Test beendet, sobald zwei Folgen eines Durchgangs inkorrekt wiedergegeben werden.

Die WMS-R erfasst die Speicherkapazität des verbalen Kurzzeitgedächtnisses und WM. Für beide Teile wurden jeweils separat die korrekt wiedergegebenen Zahlenspannen summiert und der dazugehörige alterskorrigierte Prozentrang der entsprechenden Normtabelle entnommen.

Wiener Testsystem

Das WTS beinhaltet eine Fülle von verschiedenen, am Computer durchführbaren Tests. Folgende vier Subtests fanden in dieser Studie Verwendung:

Subtest „N-Back Verbal“ (NBV)

Im Subtest NBV⁶⁰ wird ein 2-Back-Paradigma angewandt. Aufeinander folgend werden einzelne Konsonanten für kurze Zeit auf dem Bildschirm eingeblendet. Die Aufgabe besteht darin, eine Taste zu betätigen, wenn der aktuell eingeblendete Buchstabe mit demjenigen übereinstimmt, der zwei Stellen zuvor eingeblendet worden ist. Dieses Verfahren testet das verbale WM, insbesondere die „Updating“-Komponente.

Subtest „N-Back Nonverbal“ (NBN)

Der Subtest NBN⁶¹ beruht auf demselben Prinzip wie der NBV. Anstatt der Konsonanten werden abstrakte Figuren präsentiert, welche es ebenfalls im 2-Back-Paradigma mit einem Tastendruck zu markieren gilt. Erfasst wird mit diesem Subtest die Leistungsfähigkeit des visuell-räumlichen WM, auch hier erneut die „Updating“-Funktion im Besonderen.

Die Rohwerte der Tests wurden jeweils nach publizierten Normen nach Alter, Geschlecht und Bildung korrigiert. Von beiden Subtests ging jeweils der Prozentrang der Anzahl der Richtigen in die statistische Auswertung ein.

Subtest „Corsi-Block-Tapping-Test“ (CORSI)

Der Subtest CORSI⁶² stellte die letzten beiden verwendeten Einheiten des WTS dar. Es werden dem Probanden neun unregelmäßig auf dem Monitor verteilte Würfel präsentiert, die von einem durch das Programm gesteuerten Cursor in einer bestimmten Reihenfolge angeklickt werden. Nach einem auditiven Signal soll mithilfe der Computermaus die Reihenfolge wiedergegeben werden; in der ersten Einheit vorwärts, in der zweiten rückwärts. Begonnen wird mit einer Sequenz von drei angeklickten Würfeln. Die Anzahl bleibt über drei Durchgänge hinweg identisch und wird im folgenden Durchgang um einen Würfel gesteigert, sofern mindestens einer der drei Durchgänge korrekt vorwärts bzw. rückwärts reproduziert werden konnte. Maximal lässt sich die Sequenz auf acht angeklickte Würfel steigern. Dieses Verfahren evaluiert die Speicherkapazität des visuell-räumlichen Kurzzeitgedächtnisses und WM. Dem Programm wurden jeweils der Prozentrang für die Variable „unmittelbare Blockspanne“, d.h. die Anzahl der Würfel der längsten Sequenz, bei der die Probanden zwei von drei Durchgängen korrekt wiedergeben konnten, für die Auswertung entnommen.

3.4 Training

Das CWMT wurde in Kooperation mit NeuroNation (Synaptikon GmbH, Berlin, Deutschland, <https://www.neuronation.com/>) entwickelt und auf der Website dieses Anbieters durchgeführt. Das CWMT erfolgte an fünf Tagen pro Woche über einen Zeitraum von fünf Wochen. Eine Einheit umfasste ungefähr 40 bis 45 Minuten, wobei 30 Minuten die reine Trainingszeit und die restlichen zehn bis 15 Minuten das Lesen der Anleitungen zu den einzelnen Übungen sowie das Ausfüllen des Trainingstagebuchs ausmachten. Das Training galt als erfolgreich absolviert, sofern insgesamt mindestens 75% der Trainingseinheiten durchgeführt wurden.

Die Übungen waren adaptiv programmiert und passten sich der Leistung der Probanden an. Nach mehreren sukzessiv korrekt gelösten Aufgaben einer Schwierigkeitsstufe wurde die Schwierigkeit nach Erreichen einer Grenze erhöht. Wurden Aufgaben inkorrekt beantwortet, wurde die Schwierigkeit nach Unterschreiten dieser Grenze zurückgesetzt. Das CWMT am folgenden Tag passte sich der Leistung des Vortags an und begann mit der zuletzt festgelegten Schwierigkeitsstufe, sodass das CWMT im Verlauf des Interventionszeitraums bei erfolgreich gelösten Aufgaben zunehmend anspruchsvoller wurde. Es fand ein wöchentliches Telefonmonitoring durch die Studienleitung statt, um eventuelle Fragen der Probanden klären und somit eine sichere Durchführung des CWMT gewährleisten zu können.

Insgesamt beinhaltete das zusammengestellte CWMT neun verschiedene Übungen. Zu Beginn jedes Trainingstages wurde die Übung „Pfadfinder“ vierminütig als „warming-up“ durchgeführt. Es folgten pro Einheit jeweils vier weitere aus den übrigen acht Übungen à sechseinhalb Minuten, sodass insgesamt fünf Übungen pro Tag eine Einheit bildeten. Die einzelnen Übungen werden im Folgenden näher beschrieben:

Pfadfinder

Auf dem Monitor erscheinen Punkte, die durch gerade Linien nacheinander miteinander verbunden werden. Die Reihenfolge gilt es sich zu merken und anschließend durch Anklicken der Punkte nachzuahmen. Begonnen wird mit drei zu verbindenden Punkten. Nach drei aufeinander folgenden korrekt gelösten Aufgaben erhöht sich die Schwierigkeit, indem ein Punkt hinzugefügt wird. Bei drei Fehlern wird die Übung um eine Stufe, d.h. um einen zu verbindenden Punkt, zurückgesetzt.

Rückfinder

Die Übung folgt im Wesentlichen dem Aufbau der oben beschriebenen. Jedoch sollen die Punkte in umgekehrter Reihenfolge angeklickt werden.

Platzmerker

In dieser Übung werden verdeckte Karten präsentiert, wobei eine dieser Karten für kurze Zeit offen gezeigt wird, um das darauf abgebildete Symbol sichtbar zu machen. Nach Verdecken des Symbols wechseln die Karten mehrmals ihren Platz, wobei immer zwei Karten miteinander getauscht werden. Es folgt die Frage, an welcher Stelle sich das eingangs gezeigte Symbol nun befindet. Begonnen wird mit einer aufgedeckten von insgesamt vier Karten. Nach drei sukzessiven fehlerlosen Durchgängen erhöht sich die Anzahl sowohl der verdeckten als auch der anfangs offen gezeigten Karten. Es wird jedoch immer lediglich nach einer der gezeigten Karten gefragt. Nach drei inkorrekt gelösten Aufgaben wird die Schwierigkeit um eine Stufe zurückgesetzt.

Serienmerker

Hier werden im Wechsel Buchstaben bzw. Zahlen und Gleichungen präsentiert. Der gezeigte Buchstabe bzw. die gezeigte Ziffer gilt es sich einzuprägen, die mathematischen Gleichungen zu lösen. Im Anschluss jeder Gleichung erscheint auf dem Bildschirm eine Zahl. Nun wird gefragt, ob diese Zahl kleiner, gleich oder größer dem Ergebnis der Gleichung ist. Nach jedem Durchgang sollen daraufhin die gezeigten Buchstaben bzw. Ziffern angegeben werden. Werden drei Durchgänge hintereinander richtig bearbeitet, erhöht sich die Anzahl der sich zu merkenden Buchstaben bzw. Zahlen wie die der zu lösenden Gleichungen um eins. Zudem werden die mathematischen Gleichungen komplexer. Andersherum verringert sich die Anzahl der Items und vereinfachen sich die Gleichungen nach drei inkorrekt bearbeiteten Durchgängen.

Merkfluss

In dieser Übung werden auf dem Bildschirm aufgereihete Symbole präsentiert. Begonnen wird mit zwei Items. Nach einigen Sekunden wird das linke Symbol ausgeblendet, sodass nur noch das rechte sichtbar ist. Es gilt nun durch Betätigung einer Taste zu entscheiden, ob das sichtbare gleich dem unsichtbaren Symbol ist. Nach wenigen Sekunden wird das linke ausgeblendete Symbol entfernt, das rechte an seine Stelle gerückt und unkenntlich gemacht. Folgend wird rechts ein neues Symbol eingeblendet, woraufhin eine neuerliche Entscheidung bezüglich der Gleichheit der Symbole gefällt werden soll. Nach 20 korrekten Reaktionen, die im Drücken der Taste bei Gleichheit bzw. bei Unterlassen der Betätigung der Taste bei Ungleichheit bestehen, wird die Schwierigkeit erhöht, indem ein weiteres Item eingeblendet wird. Somit liegt zwischen den beiden Symbolen, die gleich oder ungleich sein können, ein weiteres. Nach fünf begangenen Fehlern wird die Schwierigkeit zurückgesetzt.

Parita

Diese Übung folgt im Aufbau der oben beschriebenen. Zusätzlich zu den präsentierten Symbolen wird eingangs eine Zahl eingeblendet. Bei jedem Wechsel der Symbole wird nun durch das Programm laut eine Zahl zwischen eins und neun genannt. Aufgabe ist es, nicht nur bei gleichen Symbolen die Taste zu betätigen, sondern auch zu reagieren, wenn die genannte Zahl gleich der anfangs gezeigten ist. Die Schwierigkeit wird nach 25 richtigen Durchgängen erhöht und nach vier falschen erniedrigt.

Multimerker

Hier werden drei Quadrate präsentiert, die Kästen darstellen sollen. Nacheinander werden auf dem Bildschirm Bälle von links und rechts kommend eingeblendet, die in diesen Kästen landen, damit nicht mehr sichtbar sind und diese umgekehrt auch wieder verlassen können. Im Anschluss jeden Durchgangs soll die Anzahl der Bälle in den einzelnen Kästen angegeben werden. Sie kann zwischen null und neun Bällen pro Kasten liegen. Ein Durchgang wird nur dann als korrekt bewertet, wenn die Anzahl aller drei Kästen richtig angegeben wird. Nach zwei aufeinander folgenden richtig beantworteten Durchgängen erhöht sich die Schwierigkeit, indem die Bälle in höherer Geschwindigkeit erscheinen und zudem zwischen den verschiedenen Kästen wechseln. Nach zwei begangenen Fehlern wird die Schwierigkeit um eine Stufe zurückgesetzt.

Drehmory

Auf dem Bildschirm wird ein Gitter präsentiert. In einer Parzelle wird ein Symbol ein- und kurz darauf wieder ausgeblendet. Anschließend dreht sich das Gitter mehrmals im oder gegen den Uhrzeigersinn um jeweils 90° . Aufgabe ist es, nach Ausführung dieser Drehungen die Lokalisation des anfänglich eingeblendeten Symbols auszumachen. Der Einstieg erfolgt mit einem Item in einem 2x2-Gitter. Nach zwei korrekt angegebenen Lokalisationen erhöht sich die Schwierigkeit, indem anfangs mehr Symbole eingeblendet werden und anschließend auch nach der Lokalisation mehrerer Symbole gefragt wird. Zudem vergrößern sich die Ausmaße des Gitters. Umgekehrt verringert sich die Schwierigkeit nach zwei begangenen Fehlern.

Blitzmerker

Auch in dieser Übung wird ein Gitter genutzt, um Symbole einblenden zu lassen. Hier werden in schneller Folge anfangs drei Symbole präsentiert. Anschließend wird nach der Lokalisation eines dieser Symbole gefragt. Nach drei korrekten Durchgängen wird die Schwierigkeit hinaufgesetzt, indem mehrere Symbole hintereinander und in schnellerer Abfolge gezeigt und die

Ausmaße des Gitters vergrößert werden. Ab der zweiten Schwierigkeitsstufe werden außerdem Items zum Teil zeitgleich präsentiert. Nach einer inkorrekt angegebenen Lokalisation wird die Übung um eine Stufe erleichtert.

3.5 Genehmigung, Registrierung und Patienteneinwilligung

Die Studie wurde durchgeführt unter Einhaltung der Deklaration von Helsinki des Weltärztebundes. Das Studienprotokoll wurde von der ethischen Kommission der medizinischen Fakultät der Universität zu Köln genehmigt (16-043) und im Deutschen Register Klinischer Studien registriert (drks.de, DRKS00009379). Alle Studienteilnehmer wurden vorab über den Ablauf und den Zweck der Studie informiert und gaben vor Durchführung der ersten Testung eine schriftliche Einwilligungserklärung ab.

3.6 Statistische Auswertung

Für die statistische Auswertung dieser Arbeit wurde die IBM SPSS Statistics Software, Version 25 (2017) für Windows genutzt. Zur Überprüfung der Daten auf Normalverteilung wurde der Shapiro-Wilk-Test herangezogen, zur Überprüfung auf Varianzhomogenität der Levene-Test. Die Gruppenvergleiche zwischen der EG und CG zu den Baseline-Charakteristika wurden für normalverteilte Daten mit dem t -Test für unabhängige Stichproben vorgenommen, für nicht-normalverteilte Daten mit dem Mann-Whitney-U-Test.

Zur Auswertung der Trainingseffekte wurden die Ergebnisse der sechs etablierten WM-Tests herangezogen. Dabei gingen jeweils die nach Alter, Bildung und Geschlecht korrigierten Prozentränge der erreichten Punktzahlen in den beiden Subtests der WMS-R, der Anzahl der Richtigen in den Subtests NBV und NBN des WTS sowie der unmittelbaren Blockspanne in den beiden Subtests des CORSI des WTS in die Analysen ein. Mithilfe dieser Daten wurden mixed 2x2-ANOVAs berechnet. Als Zwischensubjektfaktor diente die Gruppenzugehörigkeit zur EG bzw. zur CG, als Innersubjektfaktor die Zeitpunkte der Testungen. Die mixed ANOVAs wurden jeweils einzeln vorgenommen für den Vergleich der Testungszeitpunkte Vor- und Nachtestung ($t1 \times t2$) bzw. Vor- und FU-Testung ($t1 \times t3$). Betrachtet wurden die Haupteffekte *Zeit* und *Gruppenzugehörigkeit* sowie der Interaktionseffekt *Zeit x Gruppenzugehörigkeit*. Zur eingehenderen Untersuchung der einzelnen WM-Tests wurden anschließend t -Tests für abhängige Stichproben herangezogen, jeweils durchgeführt nach Zugehörigkeit zur EG bzw. CG und nach Vergleich der Testungszeitpunkte $t1 \times t2$ bzw. $t1 \times t3$. Die ES d wurde für die t -Tests berechnet nach Borenstein⁵⁴.

Die Prädiktionsanalysen wurden explorativ mit den Daten der EG anhand von multiplen linearen Regressionen durchgeführt. Dazu wurden für die sechs WM-Tests jeweils die Differenzen zwischen den Scores der Vor- und Nachtestung ($t2 - t1$) bzw. der Vor- und FU-Testung ($t3 - t1$) bestimmt. Als abhängige Variable galt die jeweils bestimmte Differenz der Testscores, als

unabhängige Variablen wurden die personenbezogenen und klinischen Prädiktoren Alter, Baselineniveau des jeweiligen WM-Tests, Bildung anhand der Anzahl der Bildungsjahre, Krankheitsschwere im Sinne der motorischen Beeinträchtigung gemessen anhand des Scores im Part III der UPDRS sowie die LEDD ausgewählt.

Als Maß der Signifikanz wurde für alle Tests ein p -Wert von $\alpha = .05$ festgesetzt, wobei $\alpha \leq .10$ als statistischer Trend gewertet wurde. Die ES d und η_p^2 für die mixed ANOVAs wurden interpretiert nach den Referenzwerten von Cohen, abgewandelt nach Ellis⁸⁰ mit $d > .2$ als kleiner, $d > .5$ als mittlerer und $d > .8$ als großer Effekt sowie mit $\eta_p^2 > .01$ als kleiner, $\eta_p^2 > .06$ als mittlerer und $\eta_p^2 > .14$ als großer Effekt. Als Maß für die Varianzaufklärung der multiplen linearen Regressionen galten nach Cohen, abgewandelt nach Ellis⁸⁰, $\text{korr. } R^2 > .02$ als geringe, $\text{korr. } R^2 > .13$ als mittlere und $\text{korr. } R^2 > .26$ als starke Aufklärung der Varianz.

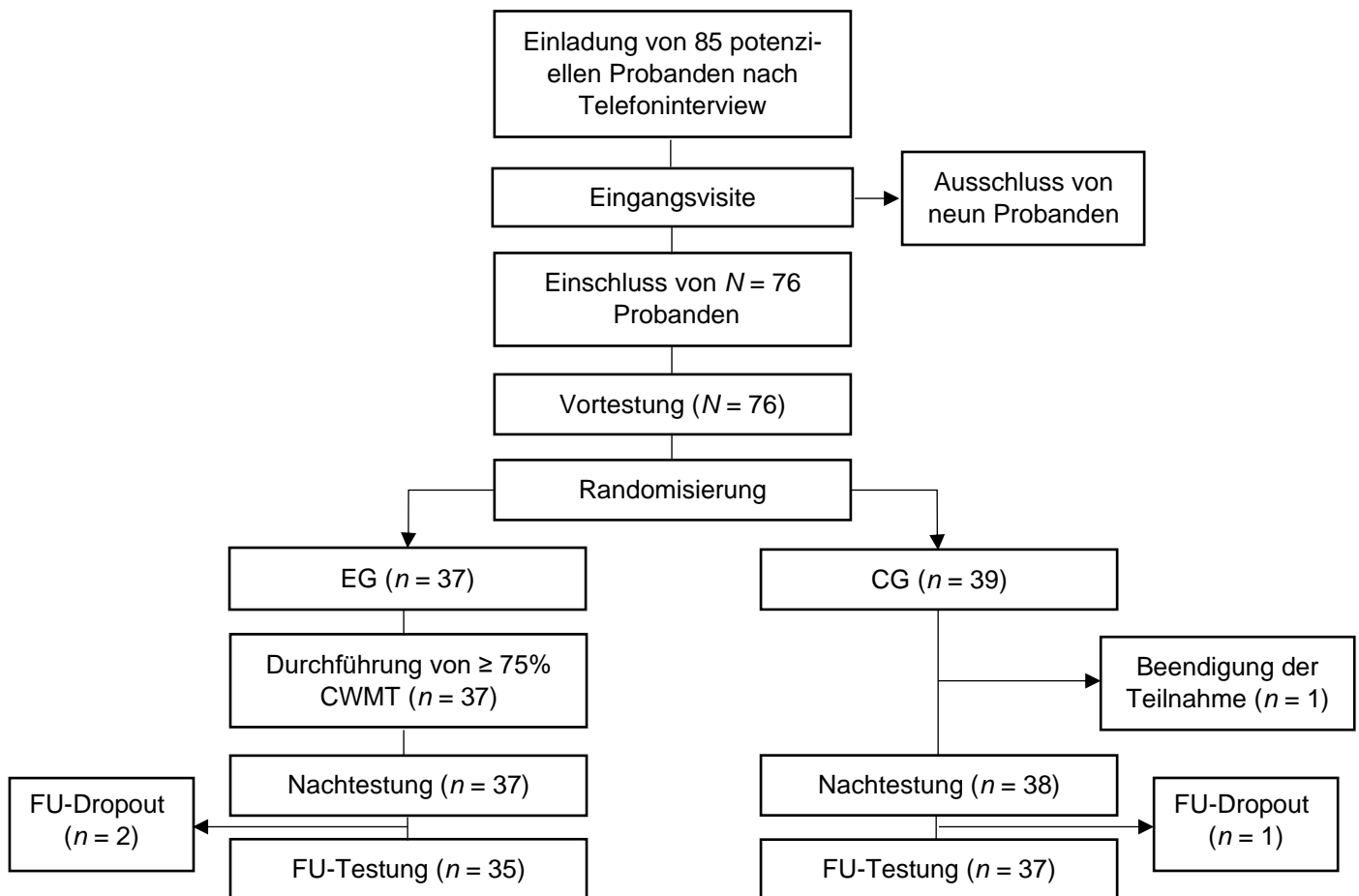
4. Ergebnisse

4.1 Studienverlauf bezüglich der Probandenzahl

Nach der ersten Kontaktaufnahme mittels Telefoninterview wurden 85 potenzielle Probanden zur Eingangsvsiste an der Uniklinik Köln eingeladen. Davon erfüllten neun Probanden nicht die Einschlusskriterien dieser Studie, sodass $N = 76$ Probanden in die Studie eingeschlossen wurden. Die eingeschlossenen Probanden wurden randomisiert auf die EG und CG verteilt. Die EG umfasste $n = 37$, die CG $n = 39$ Probanden. Nach der Vortestung und Randomisierung beendete ein Proband der CG die Teilnahme mit der Begründung eines zu hohen zeitlichen Aufwands, sodass $N = 75$ Probanden mit der Interventions- bzw. Wartephase begonnen; $n = 37$ in der EG und $n = 38$ in der CG. Die Ergebnisse der Vortestung des ausgeschiedenen Probanden wurden für die Auswertung nicht weiter berücksichtigt.

Alle Probanden der EG absolvierten $\geq 75\%$ des CWMT. Somit gab es an dieser Stelle keine Dropouts und jeder der $N = 75$ Probanden durchlief die sich anschließende Nachtstung. Nachfolgend waren drei FU-Dropouts zu verzeichnen, begründet einmal durch eine sich zugezogene Fraktur, einmal durch einen längeren Klinikaufenthalt und einmal durch einen Interessenverlust an der weiteren Studienteilnahme. Daraus ergaben sich $n = 35$ Probanden in der EG und $n = 37$ Probanden in der CG für die FU-Testung. Der Studienverlauf bezüglich der Probandenzahl ist in Abbildung 1 in einem Flowchart veranschaulicht.

Abbildung 1. Studienverlauf bezüglich der Probandenzahl



Notiz. CG, Kontrollgruppe (engl. control group); EG, Interventionsgruppe (engl. experimental group); FU, Follow-Up; CWMT, computerisiertes Arbeitsgedächtnistraining (engl. computerised working memory training).

4.2 Stichprobenbeschreibung

4.2.1 Baseline-Charakteristika

Die Prüfung der Baseline-Charakteristika auf Normalverteilung erfolgte getrennt nach EG und CG. Es ergaben sich normalverteilte Daten für die Variablen Alter und Testergebnisse des UPDRS Part III (Shapiro-Wilk-Test, $p_s \geq .05$). Die restlichen zu den Baseline-Charakteristika herangezogenen Variablen Geschlecht, Anzahl der absolvierten Schul- und Ausbildungsjahre, Scores im MoCA, im UPDRS Part V (H&Y-Skala) sowie in der GDS, Anzahl der Jahre seit der PD-Diagnose sowie LEDD waren nicht-normalverteilt (Shapiro-Wilk-Test, $p_s < .05$).

Der Gruppenvergleich zwischen EG und CG bezogen auf die genannten Baseline-Charakteristika zeigte in keiner der untersuchten Variablen signifikante Unterschiede (t -Test für unabhängige Stichproben bzw. Mann-Whitney-U-Test, $p_s > .05$). Ein zusammenfassender Bericht der Ergebnisse findet sich in Tabelle 5.

Tabelle 5. Gruppenvergleich der Interventions- und Kontrollgruppe – Baseline-Charakteristika

	EG		CG		p
	Md	Range	Md	Range	
Geschlecht	m: 19 (51.40%); w: 18 (48.60%)		m: 21 (55.30%); w: 17 (44.70%)		.736 ⁺
Alter (Jahre)	64.09 [*]	8.56 [*]	63.88 [*]	8.39 [*]	.916 [°]
Bildung (Jahre)	15	11 – 22	15.50	10 – 23	.221 ⁺
MoCA (0 – 30)	27	24 – 30	28	25 – 30	.119 ⁺
UPDRS Part III (0 – 132)	28.70 [*]	7.97 [*]	29.45 [*]	8.73 [*]	.701 [°]
UPDRS Part V (0 – 5)	2: 35 (94.60%); 3: 2 (5.40%)		2: 36 (94.70%); 3: 2 (5.30%)		.978 ⁺
PD-Diagnose (Jahre)	5.20	0.47 – 22.05	4.62	0.35 – 27.04	.791 ⁺
LEDD (mg)	600	0 – 1785	590	100 – 2120	.309 ⁺
GDS (0 – 15)	1	0 – 7	1.50	0 – 9	.862 ⁺

Notiz. CG, Kontrollgruppe (engl. control group); EG, Interventionsgruppe (engl. experimental group); GDS, Geriatric Depression Scale; LEDD, Levodopa equivalent daily dose; m, männlich; Md, Median; MoCA, Montreal Cognitive Assessment; PD, Morbus Parkinson (engl. Parkinson's disease); UPDRS, Unified Parkinson's Disease Rating Scale; w, weiblich.

* Für die normalverteilten Daten der Variablen Alter und UPDRS Part III sind anstelle von Median und Range Mittelwert und Standardabweichung angegeben.

+ Gruppenvergleich der Baseline-Charakteristika mit dem Mann-Whitney-U-Test. Angegeben ist die zweiseitige Signifikanz.

° Gruppenvergleich der Baseline-Charakteristika mit dem t-Test für unabhängige Stichproben. Angegeben ist die zweiseitige Signifikanz.

4.2.2 Ausgangsniveau der Arbeitsgedächtnistests

Die Ergebnisse der einzelnen WM-Tests waren zum Zeitpunkt der Vortestung (t_1) für keine der beiden Gruppen normalverteilt (Shapiro-Wilk-Test, $p_s < .05$). Der Gruppenvergleich der EG und CG ergab insgesamt ein höheres Ausgangsniveau der CG gegenüber der EG in den Scores der WM-Tests. Anhand des Mann-Whitney-U-Tests zeigte sich ein marginal signifikanter Unterschied in den Testergebnissen des Subtests WMS-R vorwärts, $p = .05$. Für alle anderen Tests wurde der Unterschied zwischen den Gruppen nicht signifikant ($p_s > .05$). Die Ergebnisse werden zusammenfassend in Tabelle 6 berichtet.

Tabelle 6. Gruppenvergleich der Interventions- und Kontrollgruppe im Ausgangsniveau der Arbeitsgedächtnistests

Domäne	EG		CG		p
	Md	Range	Md	Range	
Verbales Arbeitsgedächtnis					
WMS-R vw.	53	13 – 98	76	13 – 98	.050
WMS-R rw.	53	12 – 92	67	2 – 97	.159
NBV (WTS)	27	5 – 84	58	6 – 84	.104
Non-verbales Arbeitsgedächtnis					
NBN (WTS)	43	12 – 100	53	5 – 100	.312
CORSI vw. (WTS)	44	4 – 96	48	2 – 100	.555
CORSI rw. (WTS)	34	1 – 87	34	1 – 87	.414

Notiz. CG, Kontrollgruppe (engl. control group); CORSI, Corsi-Block-Tapping-Test; EG, Interventionsgruppe (engl. experimental group); Md, Median; NBN, N-back nonverbal; NBV, N-back verbal; rw., rückwärts; vw., vorwärts; WM, Arbeitsgedächtnis (engl. working memory); WMS-R, Wechsler Memory Scale-Revised; WTS, Wiener Testsystem.

Angegeben sind Median und Range der Testergebnisse zur Vortestung. Als Ergebnisse wurden die Prozentränge der erreichten Punktzahlen in den beiden Subtests der WMS-R, der Anzahl der Richtigen in den Subtests NBV und NBN des WTS sowie der unmittelbaren Blockspanne in den Subtests des CORSI des WTS herangezogen.

Der Gruppenvergleich erfolgte mit dem Mann-Whitney-U-Test. Angegeben ist die zweiseitige Signifikanz.

4.3 Trainingseffekte

4.3.1 Mixed ANOVAs

Wie oben berichtet, waren die Ergebnisse der sechs der Domäne WM zugehörigen Tests zur Vortestung nicht-normalverteilt (Shapiro-Wilk-Test, $p_s < .05$). Varianzhomogenität war in allen Tests für beide Gruppen gegeben (Levene-Test, $p_s > .05$). Da keine weitere Annahme verletzt wurde, wurden trotz Verletzung der Normalverteilungsannahme zur Auswertung der Trainingseffekte mixed 2x2-ANOVAs ohne zusätzliche Transformierung der Daten gerechnet⁸¹.

Die Scores der sechs WM-Tests galten als abhängige Variablen. Als Zwischensubjektfaktor wurde die Zugehörigkeit zur EG bzw. CG berücksichtigt (*Gruppenzugehörigkeit*), als Innersubjektfaktor der Testungszeitpunkt (*Zeit*). Verglichen wurden sowohl die Testungszeitpunkte Vor- und Nachtstung ($t1 \times t2$) als auch Vor- und FU-Testung ($t1 \times t3$).

Die deskriptive Statistik ist in den Tabellen 7 und 8 getrennt nach EG und CG tabellarisch zusammengefasst. Für die Ergebnisse der Vortestung wurden die Werte des Vergleichs $t1 \times t2$ verwendet. Die Tabellen 16 und 17 mit den Werten des Vergleichs $t1 \times t3$ für die Vortestung finden sich im Anhang.

Tabelle 7. Mixed 2x2-ANOVAs: Deskriptive Statistik der Interventionsgruppe

Domäne	t1		t2		t3	
	M	SD	M	SD	M	SD
Verbales Arbeitsgedächtnis						
WMS-R vw.	55.62	26.15	57.16	27.87	64.54	27.30
WMS-R rw.	48.92	27.38	52.62	28.02	54.11	30.79
NBV (WTS)	38.51	26.74	50.22	23.62	57.20	24.85
Non-verbales Arbeitsgedächtnis						
NBN (WTS)	46.65	25.99	54.11	22.44	54.11	28.63
CORSI vw. (WTS)	42.05	26.66	47.16	24.18	44.37	23.84
CORSI rw. (WTS)	40.92	25.08	47.41	25.86	49.97	24.43

Notiz. CORSI, Corsi-Block-Tapping-Test; M, Mittelwert; NBN, N-back nonverbal; NBV, N-back verbal; rw., rückwärts; SD, Standardabweichung; vw., vorwärts; WMS-R, Wechsler Memory Scale-Revised; WTS, Wiener Testsystem.

Die Zeitpunkte *t1*, *t2* und *t3* entsprechen der Vor-, Nach- und Follow-Up-Testung. Angegeben sind die Mittelwerte und deren Standardabweichung der Prozenträge der erreichten Punktzahlen in den beiden Subtests der WMS-R, der Anzahl der Richtigen in den Subtests NBV und NBN des WTS sowie der unmittelbaren Blockspanne in den Subtests des CORSI des WTS. Die Werte von *t1* entsprechen den Werten beim Vergleich *t1* x *t2*.

Tabelle 8. Mixed 2x2-ANOVAs: Deskriptive Statistik der Kontrollgruppe

Domäne	t1		t2		t3	
	M	SD	M	SD	M	SD
Verbales Arbeitsgedächtnis						
WMS-R vw.	67.58	24.79	70.82	25.15	67.46	27.14
WMS-R rw.	57.08	29.92	64.26	25.50	57.73	28.15
NBV (WTS)	49.95	28.10	56.27	26.78	56.89	27.19
Non-verbales Arbeitsgedächtnis						
NBN (WTS)	53.49	29.12	61.35	23.31	60.16	26.10
CORSI vw. (WTS)	46.22	28.00	49.73	26.80	51.05	28.55
CORSI rw. (WTS)	45.51	24.64	48.22	25.85	46.57	27.61

Notiz. CORSI, Corsi-Block-Tapping-Test; M, Mittelwert; NBN, N-back nonverbal; NBV, N-back verbal; rw., rückwärts; SD, Standardabweichung; vw., vorwärts; WMS-R, Wechsler Memory Scale-Revised; WTS, Wiener Testsystem.

Die Zeitpunkte *t1*, *t2* und *t3* entsprechen der Vor-, Nach- und Follow-Up-Testung. Angegeben sind die Mittelwerte und deren Standardabweichung der Prozenträge der erreichten Punktzahlen in den beiden Subtests der WMS-R, der Anzahl der Richtigen in den Subtests NBV und NBN des WTS sowie der unmittelbaren Blockspanne in den Subtests des CORSI des WTS. Die Werte von *t1* entsprechen den Werten beim Vergleich *t1* x *t2*.

Der Haupteffekt des Innersubjektfaktors *Zeit* wird für den Subtest NBV des WTS sowohl für den Vergleich *t1* x *t2* als auch für *t1* x *t3* statistisch signifikant, $F(1,72) = 7.598$, $p = .007$, $\eta_p^2 = .013$ bzw. $F(1, 70) = 15.393$, $p < .001$, $\eta_p^2 = .180$. Beide Gruppen verbesserten sich in der

Nach- und FU-Testung gegenüber der Vortestung. Für den Subtest NBN des WTS zeigte sich durch die Verbesserung beider Gruppen in der Nach- und FU-Testung für den Vergleich $t1 \times t2$ ein signifikantes Ergebnis, $F(1, 72) = 5.223$, $p = .025$, $\eta_p^2 = .068$, für den Vergleich $t1 \times t3$ ein marginal signifikantes Ergebnis, $F(1,70) = 3.756$, $p = .057$, $\eta_p^2 = .051$. Für alle anderen WM-Tests wurde der Haupteffekt des Innersubjektfaktors *Zeit* nicht signifikant ($p_s > .05$).

Für den zweiten Haupteffekt des Zwischensubjektfaktors *Gruppenzugehörigkeit* zeigte sich lediglich im Subtest WMS-R vorwärts zum Zeitpunkt $t1 \times t2$ eine statistische Signifikanz, $F(1, 73) = 5.403$, $p = .023$, $\eta_p^2 = .069$, wobei die CG gegenüber der EG höhere Scores erzielte. Für alle anderen WM-Tests zeigten sich nicht signifikante Ergebnisse ($p_s > .05$). Die Haupteffekte des Innersubjektfaktors *Zeit* und des Zwischensubjektfaktors *Gruppenzugehörigkeit* werden in den Tabellen 9 sowie 10 berichtet.

Der Interaktionseffekt wurde in keiner der mixed 2x2-ANOVAs signifikant. Für zwei der sechs Subtests wurde der Interaktionseffekt jedoch für $t1 \times t3$ marginal signifikant, während sich zeitgleich die Effektgröße η_p^2 von $t1 \times t2$ zu $t1 \times t3$ erhöhte. Für den Subtest WMS-R vorwärts ergab sich für $t1 \times t2$ $F(1, 73) = 0.126$, $p = .724$, $\eta_p^2 = .002$ und für $t1 \times t3$ $F(1, 70) = 3.016$, $p = .087$, $\eta_p^2 = .041$. Für den Subtest NBV des WTS ergab sich für $t1 \times t2$ $F(1, 72) = 0.676$, $p = .414$, $\eta_p^2 = .009$ sowie für $t1 \times t3$ $F(1,70) = 2.909$, $p = .093$, $\eta_p^2 = .040$. Die Ergebnisse für den Interaktionseffekt zwischen den Faktoren *Zeit* x *Gruppenzugehörigkeit* sind in Tabelle 11 dargestellt.

Tabelle 9. Trainingseffekte: Haupteffekt *Zeit*

Domäne	<i>t1 x t2</i>				<i>t1 x t3</i>			
	<i>df</i>	<i>F</i>	<i>p</i>	η_p^2	<i>df</i>	<i>F</i>	<i>p</i>	η_p^2
Verbales Arbeitsgedächtnis								
WMS-R vw.	1, 73	0.999	.321	.013	1, 70	1.646	.204	.023
WMS-R rw.	1, 73	2.935	.091	.039	1, 70	0.515	.475	.007
NBV (WTS)	1, 72	7.598	.007	.095	1, 70	15.393	.000	.180
Non-verbales Arbeitsgedächtnis								
NBN (WTS)	1, 72	5.223	.025	.068	1, 70	3.756	.057	.051
CORSI vw. (WTS)	1, 72	2.361	.129	.032	1, 70	1.013	.318	.014
CORSI rw. (WTS)	1, 72	2.722	.103	.036	1, 70	2.923	.092	.040

Notiz. CORSI, Corsi-Block-Tapping-Test; NBN, N-back nonverbal; NBV, N-back verbal; rw., rückwärts; vw., vorwärts; WMS-R, Wechsler Memory Scale-Revised; WTS, Wiener Testsystem. Die Zeitpunkte $t1$, $t2$ und $t3$ entsprechen der Vor-, Nach- und Follow-Up-Testung.

Tabelle 10. Trainingseffekte: Haupteffekt Gruppenzugehörigkeit

Domäne	t1 x t2				t1 x t3			
	df	F	p	η_p^2	df	F	p	η_p^2
Verbales Arbeitsgedächtnis								
WMS-R vw.	1, 73	5.403	.023	.069	1, 70	1.768	.188	.025
WMS-R rw.	1, 73	3.163	.079	.042	1, 70	1.095	.299	.015
NBV (WTS)	1, 72	2.845	.096	.038	1, 70	0.835	.364	.012
Non-verbales Arbeitsgedächtnis								
NBN (WTS)	1, 72	2.110	.151	.028	1, 70	1.151	.287	.016
CORSI vw. (WTS)	1, 72	0.378	.540	.005	1, 70	0.838	.363	.012
CORSI rw. (WTS)	1, 72	0.270	.605	.004	1, 70	0.003	.954	.000

Notiz. CORSI, Corsi-Block-Tapping-Test; NBN, N-back nonverbal; NBV, N-back verbal; rw., rückwärts; vw., vorwärts; WMS-R, Wechsler Memory Scale-Revised; WTS, Wiener Testsystem.

Die Zeitpunkte *t1*, *t2* und *t3* entsprechen der Vor-, Nach- und Follow-Up-Testung.

Tabelle 11. Trainingseffekte: Interaktionseffekt Zeit x Gruppenzugehörigkeit

Domäne	t1 x t2				t1 x t3			
	df	F	p	η_p^2	df	F	p	η_p^2
Verbales Arbeitsgedächtnis								
WMS-R vw.	1, 73	0.126	.724	.002	1, 70	3.016	.087	.041
WMS-R rw.	1, 73	0.300	.585	.004	1, 70	0.585	.447	.008
NBV (WTS)	1, 72	0.676	.414	.009	1, 70	2.909	.093	.040
Non-verbales Arbeitsgedächtnis								
NBN (WTS)	1, 72	0.004	.952	.000	1, 70	0.001	.978	.000
CORSI vw. (WTS)	1, 72	0.081	.777	.001	1, 70	0.263	.610	.004
CORSI rw. (WTS)	1, 72	0.462	.499	.006	1, 70	1.625	.207	.023

Notiz. CORSI, Corsi-Block-Tapping-Test; NBN, N-back nonverbal; NBV, N-back verbal; rw., rückwärts; vw., vorwärts; WMS-R, Wechsler Memory Scale-Revised; WTS, Wiener Testsystem.

Die Zeitpunkte *t1*, *t2* und *t3* entsprechen der Vor-, Nach- und Follow-Up-Testung.

4.3.2 t-Tests für abhängige Stichproben

Nach der Durchführung der mixed 2x2-ANOVAs wurden mit den Ergebnissen der sechs WM-Tests *t*-Tests für abhängige Stichproben durchgeführt. Es erfolgte eine getrennte Betrachtung der EG und CG sowie der Testungszeitpunkte Vor- und Nachtestung (*t1* x *t2*) bzw. Vor- und FU-Testung (*t1* x *t3*).

Eine signifikante Verbesserung in den Testergebnissen zeigte sich für die EG im Subtest NBV sowohl in der Nach- als auch in der FU-Testung, $t(36) = -2.499$, $p = .017$, $d = -.463$ bzw. $t(34) = -3.961$, $p < .001$, $d = -.676$. Eine weitere signifikante Verbesserung ließ sich in dieser Gruppe im Subtest CORSI rückwärts in der FU-Testung finden, $t(34) = -2.263$, $p = .030$, $d = -.295$. Zudem ergab sich für die EG in der FU-Testung eine annähernd signifikante Steigerung der

Testergebnisse im Subtest WMS-R vorwärts, $t(34) = -1.830$, $p = .076$, $d = -.294$. Deskriptiv zeigt sich auch in allen anderen WM-Tests für die EG in der Nach- und FU-Testung eine Verbesserung im Vergleich zur Vortestung, wenngleich die Steigerung nicht signifikant wurde ($p_s > .05$).

In der CG zeigten sich keine signifikanten Veränderungen zwischen den Zeitpunkten ($p_s > .05$). Deskriptiv zeigten sich in den Tests der WMS-R in der FU-Testung in beiden Versionen marginale Verschlechterungen, in den übrigen Tests erzielten die Probanden der CG nicht signifikante Verbesserungen. Die Ergebnisse der t -Tests werden in den Tabellen 12 und 13 berichtet.

Tabelle 12. Trainingseffekte: t -Tests für abhängige Stichproben: Interventionsgruppe

Domäne	$t1 \times t2$				$t1 \times t3$			
	<i>df</i>	<i>t</i>	<i>p</i>	<i>d</i>	<i>df</i>	<i>T</i>	<i>p</i>	<i>d</i>
Verbales Arbeitsgedächtnis								
WMS-R vw.	36	-0.370	.713	.057	34	-1.830	.076	.294
WMS-R rw.	36	-0.784	.438	.134	34	-0.952	.348	.173
NBV (WTS)	36	-2.499	.017	.463	34	-3.961	.000	.676
Non-verbales Arbeitsgedächtnis								
NBN (WTS)	36	-1.614	.115	.281	34	-1.418	.165	.235
CORSI vw. (WTS)	36	-1.283	.208	.200	34	-0.304	.763	.061
CORSI rw. (WTS)	36	-1.648	.108	.254	34	-2.263	.030	.295

Notiz. CORSI, Corsi-Block-Tapping-Test; NBN, N-back nonverbal; NBV, N-back verbal; rw., rückwärts; vw., vorwärts; WMS-R, Wechsler Memory Scale-Revised; WTS, Wiener Testsystem. Die Zeitpunkte $t1$, $t2$ und $t3$ entsprechen der Vor-, Nach- und Follow-Up-Testung.

Tabelle 13. Trainingseffekte: t -Tests für abhängige Stichproben: Kontrollgruppe

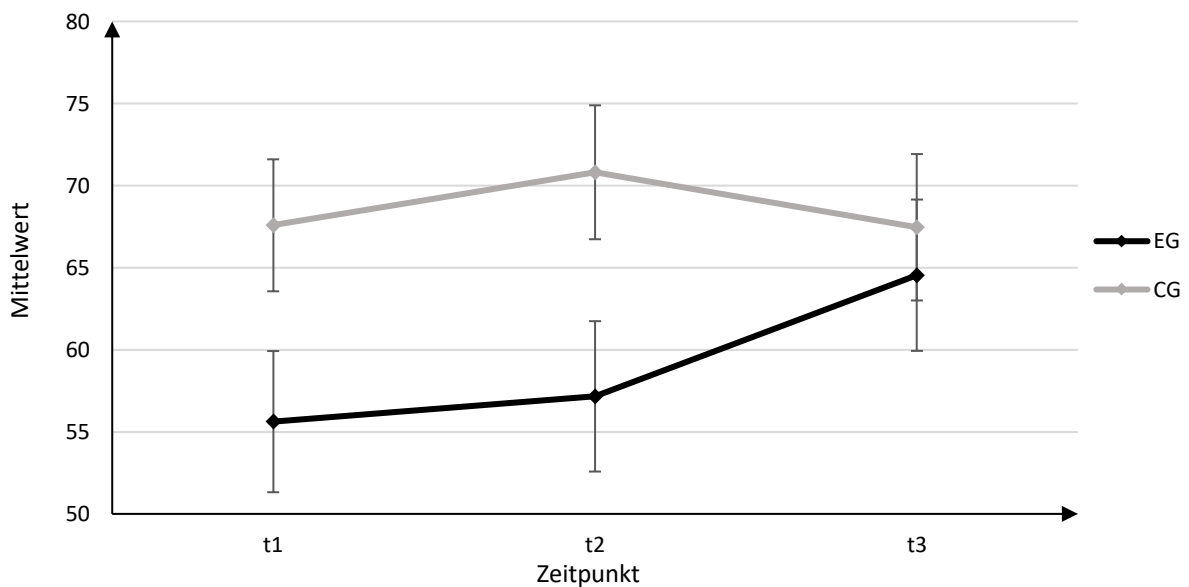
Domäne	$t1 \times t2$				$t1 \times t3$			
	<i>df</i>	<i>t</i>	<i>p</i>	<i>d</i>	<i>df</i>	<i>T</i>	<i>p</i>	<i>d</i>
Verbales Arbeitsgedächtnis								
WMS-R vw.	37	-1.337	.189	.129	36	0.391	.698	.046
WMS-R rw.	37	-1.684	.101	.256	36	0.037	.971	.006
NBV (WTS)	36	-1.385	.175	.230	36	-1.578	.123	.251
Non-verbales Arbeitsgedächtnis								
NBN (WTS)	36	-1.619	.114	.296	36	-1.338	.189	.241
CORSI vw. (WTS)	36	-0.888	.380	.128	36	-1.273	.211	.171
CORSI rw. (WTS)	36	-0.686	.497	.107	36	-0.292	.772	.040

Notiz. CORSI, Corsi-Block-Tapping-Test; NBN, N-back nonverbal; NBV, N-back verbal; rw., rückwärts; vw., vorwärts; WMS-R, Wechsler Memory Scale-Revised; WTS, Wiener Testsystem. Die Zeitpunkte $t1$, $t2$ und $t3$ entsprechen der Vor-, Nach- und Follow-Up-Testung.

Zusammenfassend lassen sich kaum signifikante Unterschiede zwischen den Testergebnissen der EG und CG berichten. Dennoch zeigen sich deskriptiv Verbesserungen in den Testergebnissen besonders in der EG, was eine Effektivität des durchgeführten CWMT vermuten lässt.

Zur Veranschaulichung dieser Trainingseffekte wurden die Ergebnisse der Subtests WMS-R vorwärts, NBV sowie CORSI rückwärts graphisch dargestellt (Abbildungen 2, 3 und 4). Die Graphen der Subtests WMS-R rückwärts, NBN sowie CORSI vorwärts finden sich im Anhang (Abbildungen 5, 6 und 7).

Abbildung 2. Vergleich der Mittelwerte von Interventions- und Kontrollgruppe im Wechsler Memory Scale-Revised vorwärts (Prozentränge)

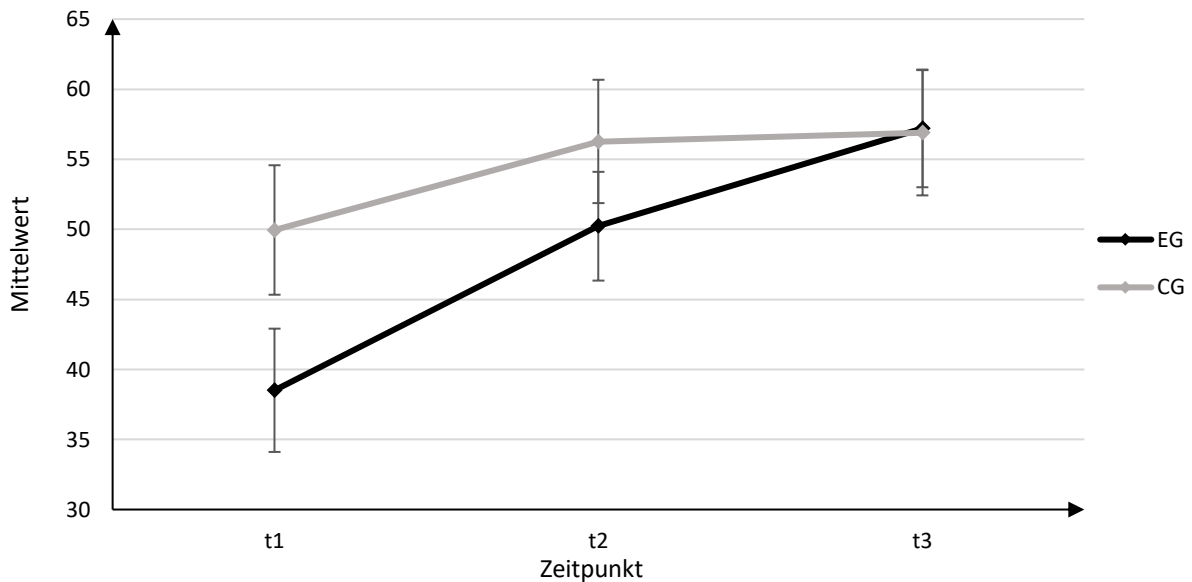


Notiz. CG, Kontrollgruppe (engl. control group); EG, Interventionsgruppe (engl. experimental group); PR, Prozentrang; WMS-R, Wechsler Memory Scale-Revised.

Die Zeitpunkte $t1$, $t2$ und $t3$ entsprechen der Vor-, Nach- und Follow-Up-Testung.

Verwendet wurden die errechneten Mittelwerte und Standardfehler aus den t -Tests für abhängige Stichproben. Für die Mittelwerte und Standardfehler des Zeitpunkts $t1$ wurden die Werte des Vergleichs $t1 \times t2$ verwendet.

Abbildung 3. Vergleich der Mittelwerte von Interventions- und Kontrollgruppe im N-back verbal (Prozentränge der Richtigen)

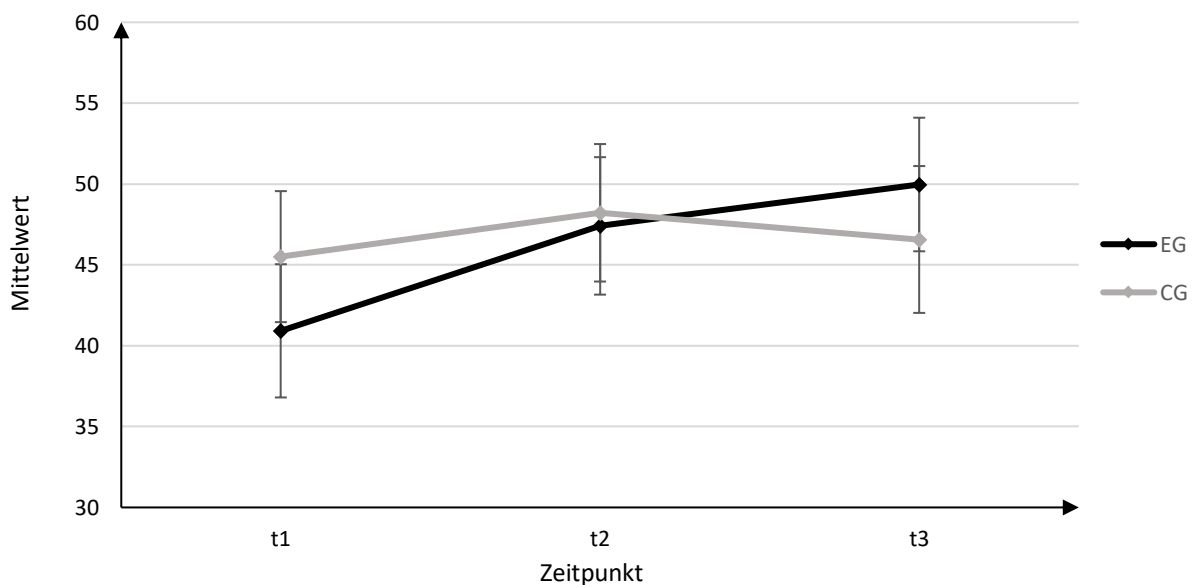


Notiz. CG, Kontrollgruppe (engl. control group); EG, Interventionsgruppe (engl. experimental group); NBV, N-back verbal; PR, Prozentrang.

Die Zeitpunkte t_1 , t_2 und t_3 entsprechen der Vor-, Nach- und Follow-Up-Testung.

Verwendet wurden die errechneten Mittelwerte und Standardfehler aus den t -Tests für abhängige Stichproben. Für die Mittelwerte und Standardfehler des Zeitpunkts t_1 wurden die Werte des Vergleichs t_1 x t_2 verwendet.

Abbildung 4. Vergleich der Mittelwerte von Interventions- und Kontrollgruppe im Corsi-Block-Tapping-Test rückwärts (Prozentränge der unmittelbaren Blockspanne)



Notiz. CG, Kontrollgruppe (engl. control group); CORSI, Corsi-Block-Tapping-Test; EG, Interventionsgruppe (engl. experimental group); PR, Prozentrang.

Die Zeitpunkte t_1 , t_2 und t_3 entsprechen der Vor-, Nach- und Follow-Up-Testung.

Verwendet wurden die errechneten Mittelwerte und Standardfehler aus den t -Tests für abhängige Stichproben. Für die Mittelwerte und Standardfehler des Zeitpunkts t_1 wurden die Werte des Vergleichs t_1 x t_2 verwendet.

4.4 Prädiktionsanalysen

Um die Effektivität des Trainings für verschiedene Personengruppen zu untersuchen, wurden anhand der Testscores der EG explorative Prädiktionsanalysen durchgeführt. Als Prädiktoren des Trainingsoutcomes wurden die unabhängigen klinischen und personenbezogenen Variablen Alter des Probanden, Baselineniveau des jeweiligen WM-Tests, Anzahl der Bildungsjahre, Score im UPDRS Part III als Maß der motorischen Beeinträchtigung im Sinne der Krankheitsschwere sowie die LEDD untersucht. Die Differenzen der Testscores zwischen Vor- und Nach- ($t_2 - t_1$) bzw. Vor- und FU-Testung ($t_3 - t_1$) wurden bestimmt und als abhängige Variablen für die multiplen linearen Regressionen für die kurz- bzw. mittelfristigen Effekte des CWMT verwendet. Eine detaillierte Ergebnisdarstellung findet sich in den Tabellen 14 und 15.

Tabelle 14. Prädiktionsanalyse: $t_2 - t_1$

Domäne	R^2 (korr. R^2)	F-Test	standardisierter Regressionskoeffizient β				
			Alter	Bildung	LEDD	UPDRS Part III	Baseline- niveau
Verbales Arbeitsgedächtnis							
WMS-R vw.		$F(5, 31)$					
	.262 (.143)	= 2.201, $p = .080$	0.201	-0.193	0.068	-0.146	-0.347*
WMS-R rw.		$F(5, 31)$					
	.325 (.216)	= 2.980, $p = .026$	-0.009	0.025	0.271+	0.068	-0.428**
NBV (WTS)		$F(5, 31)$					
	.463 (.376)	= 5.332, $p = .001$	-0.004	0.175	0.138	0.016	-0.653***
Non-verbales Arbeitsgedächtnis							
NBN (WTS)		$F(5, 31)$					
	.541 (.467)	= 7.307, $p < .001$	-0.134	-0.103	-0.209	0.170	-0.655***
CORSI vw. (WTS)		$F(5, 31)$					
	.431 (.340)	= 4.702, $p = .003$	-0.323*	-0.156	0.066	-0.067	-0.605***
CORSI rw. (WTS)		$F(5, 31)$					
	.406 (.311)	= 4.245, $p = .005$	0.120	-0.040	-0.227	-0.408**	-0.418**

Notiz. CORSI, Corsi-Block-Tapping-Test; LEDD, Levodopa equivalent daily dose; NBN, N-back nonverbal; NBV, N-back verbal; rw., rückwärts; UPDRS, Unified Parkinson's Disease Rating Scale; vw., vorwärts; WMS-R, Wechsler Memory Scale-Revised; WTS, Wiener Testsystem.

Die Zeitpunkte t_1 und t_2 entsprechen der Vor- bzw. Nachttestung.

Für die Prädiktoren angegeben sind die standardisierten Regressionskoeffizienten β mit dem Signifikanzniveau α , eingeteilt wie folgt: *** $\leq .001$; ** $\leq .010$; * $\leq .050$; + $\leq .100$.

Zehn der zwölf berechneten Modelle (sechs pro Zeitvergleich) wurden signifikant ($p \leq .037$). Die Regression des Subtests WMS-R vorwärts mit der Differenz $t2 - t1$ ergab eine annähernde Signifikanz mit $p = .080$. Das Modell des Subtests CORSI rückwärts zum Zeitvergleich $t3 - t1$ wurde nicht signifikant ($p = .241$). Dieses Modell wurde daher in der weiteren Ergebnisdarstellung nicht berücksichtigt. Die elf berücksichtigten Regressionen ergaben mittlere bis starke Varianzaufklärungen ($.143 \geq \text{korr. } R^2 \leq .467$).

Tabelle 15. Prädiktionsanalyse: $t3 - t1$

Domäne	R^2 (korr. R^2)	F-Test	standardisierter Regressionskoeffizient β				
			Alter	Bildung	LEDD	UPDRS Part III	Baseline- niveau
Verbales Arbeitsgedächtnis							
WMS-R vw.		$F(5, 29)$					
	.323 (.206)	= 2.762, $p = .037$	-0.195	0.132	-0.174	-0.186	-0.456**
WMS-R rw.		$F(5, 29)$					
	.339 (.225)	= 2.980, $p = .027$	-0.137	0.109	0.181	0.287*	-0.380*
NBV (WTS)		$F(5, 29)$					
	.361 (.251)	= 3.277, $p = .018$	-0.015	0.140	-0.074	0.138	-0.596***
Non-verbales Arbeitsgedächtnis							
NBN (WTS)		$F(5, 29)$					
	.367 (.258)	= 3.362, $p = .016$	-0.002	0.207	0.039	0.396*	-0.382*
CORSI vw. (WTS)		$F(5, 29)$					
	.523 (.441)	= 6.372, $p < .001$	-0.196	0.055	-0.057	-0.166	-0.665***
CORSI rw. (WTS)		$F(5, 29)$					
	.199 (.060)	= 1.437, $p = .241$					

Notiz. CORSI, Corsi-Block-Tapping-Test; LEDD, Levodopa equivalent daily dose; NBN, N-back nonverbal; NBV, N-back verbal; rw., rückwärts; UPDRS, Unified Parkinson's Disease Rating Scale; vw., vorwärts; WMS-R, Wechsler Memory Scale-Revised; WTS, Wiener Testsystem.

Die Zeitpunkte $t1$ und $t3$ entsprechen der Vor- bzw. Follow-Up-Testung.

Für die Prädiktoren angegeben sind die standardisierten Regressionskoeffizienten β mit dem Signifikanzniveau α , eingeteilt wie folgt: *** $\leq .001$; ** $\leq .010$; * $\leq .050$; + $\leq .100$.

Ein geringes Baselineniveau des jeweiligen WM-Tests war in allen Prädiktionsmodellen ein signifikanter positiver Prädiktor für das Trainingsoutcome. Der Score zur Baseline im UPDRS

Part III zeigte als Prädiktor in drei Regressionen ebenfalls marginal signifikante und signifikante Werte: Für die kurzfristigen Trainingseffekte im Subtest CORSI rückwärts zeigte sich eine geringere motorische Beeinträchtigung als positiver Prädiktor. Im Gegensatz dazu zeigte sich für die mittelfristigen Trainingseffekte im Subtest NBN eine stärkere motorische Beeinträchtigung als signifikanter positiver Prädiktor für das Trainingsoutcome. Für den Subtest WMS-R rückwärts lag hier ein vergleichbares, wenn auch nur marginal signifikantes Ergebnis vor.

Das Alter des Probanden war im Modell des Subtests CORSI vorwärts ($t2 - t1$) ein signifikanter negativer Prädiktor und deutet somit auf bessere kurzfristige Trainingseffekte bei jüngeren Probanden hin. Ebenfalls für die kurzfristigen Trainingseffekte zeigte sich in der Regression des Subtests WMS-R rückwärts die LEDD als positiver Prädiktor mit annähernder Signifikanz. Für den Prädiktor Anzahl der Bildungsjahre ergaben die Regressionen keine Signifikanzen. Die Ergebnisdarstellung mit den nicht standardisierten Regressionskoeffizienten B findet sich im Anhang in den Tabellen 18 und 19.

5. Diskussion

Ziel der Studie war es, die Effektivität eines mehrwöchigen, intensiven, adaptiven CWMT bei Probanden mit PD ohne kognitive Beeinträchtigungen (Ausschluss von PD-MCI oder PD-Demenz nach Level-II-Kriterien⁵⁰) auf die WM-Leistung zu untersuchen. Dafür wurden $N = 75$ Probanden in die Studie eingeschlossen, wobei $n = 37$ Probanden das CWMT absolvierten (EG) und $n = 38$ Probanden als passive CG fungierten. Zur Auswertung des Trainingsoutcomes wurden die Ergebnisse sechs etablierter WM-Tests zur Vor-, Nach- und FU-Testung betrachtet. Zusätzlich wurden Prädiktionsanalysen durchgeführt, um den Einfluss der Variablen Alter, Baselinenniveau, Bildung, Krankheitsschwere im Sinne der motorischen Beeinträchtigung sowie LEDD auf das Ansprechen des CWMT zu untersuchen.

In dieser Arbeit wurden die Trainingseffekte des CWMT auf Einzeltestebene untersucht. Eine Auswertung der Effekte hinsichtlich der verschiedenen kognitiven Domänen im Allgemeinen und der Unterscheidung zwischen verbalem sowie non-verbalem WM im Besonderen findet sich in der Publikation von Ophey, Giehl et al.⁴⁹. Die Prädiktionsanalysen erfolgten explorativ unter alleiniger Betrachtung der EG und ebenfalls auf Einzeltestebene. Weitere Prädiktionsanalysen, mit denen der Einfluss verschiedener Variablen auf Nah- (WM) und Ferntransfereffekte (exekutive Funktionen) untersucht sowie ein Vergleich zur CG vorgenommen wurde, sind in einer weiteren Veröffentlichung zu unserer Studie nachzulesen⁸².

5.1 Ergebnisse

Im Folgenden werden die Ergebnisse getrennt nach den beiden Hypothesen (a) und (b) zusammengefasst und diskutiert.

5.1.1 Hypothese (a): Effektivität des computerisierten Arbeitsgedächtnistrainings

Es konnten lediglich in zwei der sechs WM-Tests, in den Subtests WMS-R vorwärts und NBV, mittelfristige Trainingseffekte festgestellt werden, wobei diese nur ein marginales Signifikanzniveau erreichten. Opehy, Giehl et al.⁴⁹ sahen signifikante Verbesserungen durch das CWMT einzig im verbalen WM. Da beide Subtests WMS-R vorwärts sowie NBV das verbale WM prüfen, gehen diese Ergebnisse miteinander einher. Kurz- und mittelfristige Retest-Effekte ergaben sich marginal signifikant bis signifikant ebenfalls für zwei WM-Tests, den NBN und den NBV des WTS. Demnach verbesserten sich sowohl die EG als auch die CG über den gesamten Zeitraum der Studie hinweg in diesen beiden Subtests. Die Gruppenzugehörigkeit der Probanden zur CG oder EG zeigte lediglich einen kurzfristigen signifikanten Effekt auf die Ergebnisse des Subtests WMS-R vorwärts, wobei die CG gegenüber der EG bessere Ergebnisse erzielte. Für alle anderen WM-Tests und Zeitpunkte zeigte sich kein signifikanter Unterschied zwischen beiden Gruppen. Deskriptiv ergaben sich in der EG in allen sechs WM-Tests Verbesserungen in der Nach- und FU-Testung im Vergleich zur Vortestung. Im Rahmen abhängiger *t*-Tests zeigten sich signifikante Ergebnisse jedoch nur für den Subtest NBV zu beiden Vergleichszeitpunkten und den Subtest CORSI rückwärts zur FU-Testung, die Ergebnisse des Subtests WMS-R vorwärts zeigten zur FU-Testung immerhin eine annähernde Signifikanz. Für die CG ergaben sich keine signifikanten Verbesserungen und teilweise sogar marginale Verschlechterungen in den FU-Testungen.

Insgesamt ergab diese Studie also keine eindeutig signifikanten Ergebnisse mit mittleren bis großen ES, wie vorangegangene Forschungsarbeiten²⁵⁻²⁷ es vorab vermuten ließen. Jedoch ließen sich die Ergebnisse von Weicker et al., die herausarbeiteten, dass ein WMT langfristige Effekte zeige²⁷, durch den Trend zu besseren Ergebnissen der EG in der FU- gegenüber der Nachtestung zumindest tendenziell bestätigen. Eine Erklärung für die mangelnde Signifikanz der in dieser Studie gefundenen Ergebnisse im Hinblick auf eine Effektivitätssteigerung des WM wäre, dass die gewählte intensive Form des CWMT mit hohem Trainingspensum sowie anspruchsvollen Aufgaben zu Überforderung und dadurch zu motivationaler und funktionaler Erschöpfung geführt haben könnte. Dies könnte womöglich das Abschneiden in den WM-Tests der sich der Interventionsphase anschließenden Nachtestung negativ beeinflusst haben⁸³. Hinweise auf die hohe Beanspruchung durch das CWMT und die damit einhergehende Erschöpfung ergeben sich aus Angaben der Probanden nach Abschluss der Studie bzw. im Trainingstagebuch. Zudem bestand im Gegensatz zu vielen anderen WMT-Studien das Training nicht hauptsächlich aus N-back- oder WM-Span-Aufgaben^{18,28,30,32}, sondern es kamen komplexere WM-Aufgaben zum Einsatz. Diese Aufgaben mit stark exekutiver Komponente könnten eventuell dazu geführt haben, dass kurzfristige Trainingseffekte durch Erschöpfung redu-

ziert wurden bzw. sich Effekte erst verzögert als mittelfristige positive Effekte entwickeln konnten⁴⁹. Für diese These spräche ebenfalls, dass die Probanden der EG sich nach einer Erholungsphase von drei Monaten zur FU-Testung insgesamt noch einmal verbesserten.

Wenngleich das Gros der Ergebnisse unterhalb des Signifikanzniveaus blieb, ließ sich deskriptiv jedoch eine stärkere Verbesserung der EG gegenüber der CG über den Studienzeitraum hinweg erkennen, was mit einer Effektivität des durchgeführten CWMT in Verbindung gebracht werden kann. Somit lassen sich die Erkenntnisse anderer wissenschaftlicher Arbeiten bestätigen, dass ein WMT bei Probanden mit (beginnenden bzw. zu erwartenden) WM-Defiziten effektiv sein kann, um die WM-Funktionen zu verbessern^{27,28,30}. Explizit scheint das auch für Probanden mit PD zu gelten, was die Studie von Fellman et al. schon vermuten ließ³⁹. Demnach lässt sich die eingangs aufgestellte Hypothese (a), dass Probanden mit PD ohne kognitive Einschränkungen nach Level-II-Kriterien⁵⁰ von einem mehrwöchigen, intensiven, adaptiven CWMT profitieren, bestätigen, wobei die Verbesserungen in den verbalen WM-Tests tendenziell stärker waren als in den non-verbalen.

5.1.2 Hypothese (b): Prädiktionsanalysen

Neben den Analysen zur Effektivität des CWMT auf die WM-Leistung sollten Prädiktionsanalysen die Frage klären, welche Variablen das Ansprechen auf bzw. den Erfolg des CWMT bedingten. In allen elf berücksichtigten multiplen linearen Regressionen konnte gezeigt werden, dass ein niedriges Baselinenniveau in dem untersuchten WM-Test prädiktiv für einen größeren positiven Trainingseffekt war. Diese Ergebnisse lassen sich im Rahmen des „Magnification and Compensation“-Modells⁴² aufseiten des „Compensation“-Ansatzes einordnen und unterstützen die These, dass bei Probanden mit einem geringeren Ausgangsscore mehr Raum für Verbesserungen und somit das Potenzial für eine höhere Effektivität eines WMT gegeben ist. Diese Ergebnisse gehen einher mit den Beobachtungen von Weicker et al. und Zimmermann et al. und widersprechen denen von Nemmi et al., die eine größere Effektivität eines WMT bei Probanden mit einem besseren Ausgangsniveau sahen und dies mit überlegenen neurobiologischen Faktoren wie einer höheren neuronalen Plastizität erklärten^{27,43,48}. An dieser Stelle ist jedoch der Effekt des „Regression towards the mean“-Phänomens⁸⁴ nicht außer Acht zu lassen. Übertragen auf unsere Studie besagt dieses Phänomen, dass Probanden mit sehr guten bzw. sehr schlechten Ergebnissen in den WM-Tests bei der nächsten Durchführung natürlicherweise näher am Mittelwert aller Probanden lägen, ohne dass diese Verbesserung bzw. Verschlechterung durch eine Intervention, in diesem Fall durch das durchgeführte CWMT, beeinflusst worden wäre. Somit ist eine deutliche Verbesserung bei anfangs schwachen Scores zur Baseline nicht allein durch das CWMT zu erklären, sondern zudem durch den Umstand, dass deutlich vom Mittelwert abweichende Werte eine größere Tendenz aufweisen, sich dem Mittelwert anzunähern. An dieser Stelle lässt sich jedoch auf die Untersuchungen

von Ophey et al. verweisen: Das Baselineniveau als negativer Prädiktor behielt auch nach dem Vergleich zu den Ergebnissen der CG seine statistische Signifikanz⁸². Dies könnte hinweisend auf einen tatsächlichen Zusammenhang zwischen niedrigem Ausgangsniveau der Testscores und Trainingserfolg sein.

Der Score zur Baseline im UPDRS Part III⁴⁷ als Maß der motorischen Beeinträchtigung im Sinne der Krankheitsschwere wurde in drei Regressionen signifikant bzw. marginal signifikant, zeigte jedoch uneinheitliche Richtungen bei kurz- und mittelfristigen Trainingseffekten an. Bessere kurzfristige Ergebnisse ergaben sich in einer Regression bei einer geringeren motorischen Beeinträchtigung, wohingegen zwei Regressionen bessere mittelfristige Ergebnisse bei einer höheren motorischen Beeinträchtigung zeigten. Somit lässt sich keine eindeutige Tendenz der Richtung der Prädiktion zwischen Krankheitsschwere und Trainingsoutcome feststellen, was der bisher ebenfalls uneinheitlichen Literatur zu diesem Thema^{27,40} keinen neuen Trend geben kann. An dieser Stelle ist allerdings anzumerken, dass die Krankheitsschwere in den aufgeführten Studien anhand von kognitiver Beeinträchtigung⁴⁰ bzw. von Hirnverletzungen²⁷ gemessen wurde. Da das Maß der Krankheitsschwere in dieser Arbeit anhand der motorischen Beeinträchtigung festgemacht wurde, sind die Ergebnisse also nur bedingt vergleichbar. Daher lässt sich die Prädiktion der Krankheitsschwere anhand des Scores im UPDRS Part III am ehesten mit der Arbeit von Edwards et al. vergleichen, in der mit der H&Y-Skala ebenfalls ein Tool der UPDRS eingesetzt wurde, um die motorische Beeinträchtigung der Probanden mit PD anzugeben⁴⁶. Edwards und Kollegen konnten jedoch keinen prädiktiven Zusammenhang zwischen Einstufung auf der H&Y-Skala und Trainingserfolg finden. Dass der in unserer Studie verwendete Score des UPDRS Part III zumindest einige signifikante Prädiktion des Trainingsoutcomes aufdecken konnte, könnte eventuell darauf hindeuten, dass der UPDRS Part III sensitiver für die Bewertung der motorischen Beeinträchtigung ist als die H&Y-Skala (UPDRS Part V). Die Diskrepanz der Richtung zwischen kurz- und mittelfristiger Prädiktion könnte aber auch ein Hinweis auf die erst allmählich eintretende Auswirkung des CWMT auf die Leistung des WM sein, wie es insgesamt für die EG in unserer Studie beobachtbar war, da sie sich in der Tendenz zum FU im Vergleich zur Nachtestung in ihren Testscores noch verbesserte. Einschränkend muss hier jedoch erwähnt werden, dass die relative Homogenität der Probanden unserer Studie eine nur geringe Range in den UPDRS-Skalen und somit eine geringe Bandbreite an motorischer Beeinträchtigung bedingte. Ophey et al. konnten in ihrer Untersuchung des Einflusses der motorischen Beeinträchtigung, ebenfalls gemessen anhand des UPDRS Part III, auf das Trainingsoutcome keine statistische Signifikanz feststellen⁸².

Der Prädiktor Alter ergab in einer Regression signifikant bessere kurzfristige Trainingseffekte bei jüngeren Probanden, was mit den Ergebnissen von Dahlin, Nyberg et al. sowie Ophey et al. einhergeht^{41,82}. Da der Prädiktor aber in allen anderen der zehn durchgeführten Regressionen unterhalb des Signifikanzniveaus liegt, lässt sich der Einfluss des Alters der Probanden

auf das Trainingsoutcome infrage stellen, wie es auch schon an anderer Stelle geschehen ist²⁵.

Ebenfalls für die kurzfristigen Trainingseffekte zeigte sich die LEDD als positiver Prädiktor und deutete somit auf ein besseres Outcome bei höher dosierter Medikation hin, wenngleich das Ergebnis lediglich annähernd signifikant blieb. Dies widerspräche der „dopamine-overdose“-Hypothese⁴⁴, die wegen eines durch ein WMT verursachten Überschusses an Dopamin in frontostriatalen Netzwerken einen negativen prädiktiven Zusammenhang zwischen LEDD und Effektivität eines WMT vermutet, und bestätigte die Ergebnisse von Edwards et al., die in der LEDD ebenfalls einen tendenziellen positiven prädiktiven Charakter auf den Erfolg eines kognitiven Trainings bei Probanden mit PD sahen⁴⁶. Da die LEDD als Prädiktor in dieser Arbeit lediglich in einer der elf Regressionen annähernd signifikant wurde, lässt sich jedoch keine klare Aussage zum Einfluss der LEDD auf die Effektivität eines CWMT treffen und geht damit am ehesten mit den Ergebnissen von Fellman et al. einher, die ebenfalls keinen prädiktiven Zusammenhang zwischen LEDD und Trainingsoutcome bei Probanden mit PD feststellen konnten³⁹.

Der Prädiktor Bildung blieb ohne jegliche Signifikanz. Demnach konnten die Ergebnisse von Fernández-Ballesteros et al., die einen positiven prädiktiven Zusammenhang zwischen Bildung und Trainingserfolg sahen⁴⁰, nicht reproduziert werden. Auch Ophrey et al. konnten lediglich einen annähernd signifikanten positiven prädiktiven Charakter des Bildungsstands auf die Nahtransfereffekte des CWMT feststellen⁸².

Insgesamt lässt sich festhalten, dass einzig der Prädiktor Baselineniveau der WM-Tests über alle Regressionen hinweg das Signifikanzniveau erreichte. Die motorische Beeinträchtigung gemessen anhand des UPDRS Part III sowie das Alter der Probanden wurden zumindest in ein bis zwei Regressionen signifikant, die LEDD näherte sich der Signifikanz nur in einer der elf Regressionen. Die Anzahl der Bildungsjahre schien keinerlei Einfluss auf den Trainingserfolg zu haben. Somit konnte die eingangs aufgestellte Hypothese (b), dass diese fünf eben genannten Variablen eine Prädiktion des CWMT-Outcome zulassen, nur bedingt bestätigt werden. Die mangelnde prädiktive Aussagekraft auf den Erfolg des CWMT lässt sich eventuell damit erklären, dass die statistisch ohnehin schwachen Trainingseffekte die Möglichkeit schmälerten, durch die Prädiktionsanalysen Variablen mit prädiktivem Charakter aufzudecken.

5.2 Stärken und Schwächen

Im Verlauf der Studie und bei der Auswertung der Ergebnisse ergaben sich diverse Stärken und Schwächen. So lässt sich vermuten, dass die statistisch schwachen Trainingseffekte der EG in den WM-Aufgaben zumindest teilweise der mangelnden Sensitivität der verwendeten neuropsychologischen Testbatterie geschuldet sein könnten und kleine Veränderungen der WM-Leistung schlichtweg nicht sichtbar wurden. Die Möglichkeit, dass eine Steigerung der

WM-Leistung durch Deckeneffekte verschleiert wurde, ist ebenfalls in Betracht zu ziehen. Falls ein WM-Test so gestellt war, dass Probanden schon zum Zeitpunkt der Vortestung Höchstwerte erzielt haben, wären keine messbaren Verbesserungen in den nachfolgenden Testungen mehr möglich. Neben der Batterie an etablierten neuropsychologischen Tests kam in unserer Studie zusätzlich die WwW-Task^{64–66} an allen drei Testungszeitpunkten zur Erfassung der WM-Leistung zum Einsatz. Diese Form des WM-Tests ermöglichte nicht nur die Detektion der WM-Kapazität, sondern auch der Präzision und des Wiederholens bestimmter „Fehler“. Somit ließen sich auch diejenigen subtilen Verbesserungen der WM-Leistung messen, die durch die etablierten WM-Tests vielleicht nicht fassbar waren. Die Auswertung der WwW-Task wird an anderer Stelle beschrieben und diskutiert⁷⁹. Eine weitere Schwäche bezogen auf die Studienergebnisse ist der Umstand, dass die CG gegenüber der EG ein insgesamt höheres Ausgangsniveau zur Baseline zeigte. Diese Ungleichheit war jedoch durch die randomisierte Einteilung der Probanden zur CG und EG dem Zufall geschuldet und daher durch die Studienleitung weder voraussehbar noch beeinflussbar, insbesondere, da die Probandengruppe insgesamt durch die Anwendung der Level-II-Kriterien nach Litvan et al.⁵⁰ bezogen auf ihren kognitiven Status recht homogen war. In den Ergebnissen der WM-Tests spiegelte sich das ungleiche Baselineniveau insofern wider, dass die Verbesserungen der EG durch das CWMT zu einer Annäherung des Niveaus beider Gruppen in den sich der Interventionsphase anschließenden Testungen führten, was bei der Betrachtung der Ergebnisse nicht außer Acht gelassen werden sollte.

Zudem ließen sich besonders zur Nachtestung auch Verbesserungen der CG in den Tests feststellen, die wahrscheinlich auf Retest-Effekte zurückzuführen sind. Daher wurde nicht die reine Entwicklung der Testungsergebnisse über den Studienzeitraum hinweg als Trainingsoutcome ausgewertet, sondern die Interaktion zwischen der EG und CG über den Zeitraum, um die Testungsergebnisse der beiden Gruppen vergleichend auswerten zu können. In der Mehrheit der untersuchten WM-Tests ließen die Retest-Effekte zur FU-Testung wieder nach, was die im Durchschnitt schlechteren Ergebnisse der CG in der FU- gegenüber der Nachtestung offenlegte. Die EG hingegen steigerte sich im Trend zur FU-Testung in ihren WM-Leistungen noch, was für einen mittelfristigen Effekt des CWMT spricht. Um auch langfristige Effekte des CWMT aufdecken zu können, hätte der Zeitraum des FU länger gewählt werden müssen als drei Monate bzw. es hätte sich eine vierte Testung zu einem späteren Zeitpunkt anschließen müssen. In der vergleichenden Metaanalyse von Weicker et al. beispielsweise wurden langfristige Effekte der WMT der verschiedenen RCT als solche benannt, die in den FU-Testungen mit einem durchschnittlichen Zeitraum von fünf Monaten ermittelt wurden²⁷.

Als weitere Schwäche der Studie kann die Verwendung einer passiven anstatt einer aktiven CG angeführt werden, was dem Goldstandard in der klinischen Forschung entsprochen hätte. Da WMT und allgemeines kognitives Training jedoch bisher noch keinen definitiven Einzug in

den klinischen Alltag erhalten haben und noch keine nennenswerten Erfahrungswerte bestehen, ist es dennoch sinnvoll, WMT mit keinerlei Form der Intervention, das heißt eine EG mit einer passiven CG, zu vergleichen. So können beispielsweise reine Retest-Effekte detektiert werden²⁸. Außerdem stellen passive CG eine ökonomischere Alternative dar, die auch schon in vorangegangenen Studien angewendet wurde^{28,32}. Dennoch sollte in folgenden Studien erwogen werden, sowohl eine passive als auch eine aktive CG einzusetzen, um die Effekte eines WMT weiterführend analysieren zu können. In dem u.a. von uns angewendeten Studiendesign ist die Passivität der CG allerdings nicht vollständig kontrollierbar, da in einem solchen Rahmen niemals gänzlich ausgeschlossen werden kann, dass die Anweisung, während des Interventionszeitraums kein eigenständiges WMT anzuwenden, möglicherweise nicht von allen Probanden konsequent eingehalten wird.

Außerdem ließ sich die Verblindung der Testleiter bei den Nach- und FU-Testungen nicht immer aufrechterhalten. Zwar wurde darauf geachtet, dass die Nach- und FU-Testung eines Probanden nicht durch diejenige Person geleitet wurde, die die Vortestung durchgeführt und dem Probanden somit die Zuteilung zur EG oder CG mitgeteilt hatte. Jedoch nannten die Probanden der Testleitung trotz der Bitte um Verschwiegenheit teilweise ihre Gruppenzugehörigkeit vor bzw. während der Durchführung der Nach- und FU-Testungen, was nicht immer rechtzeitig unterbunden werden konnte.

Prädiktionsanalysen weisen in der Literatur ein recht heterogenes Bild auf. Neuere Studien empfehlen bei multiplen Regressionen den Einsatz von Veränderungen in den Testscores zu den unterschiedlichen Testungszeitpunkten als abhängige Variable^{85,86}, was in dieser Arbeit durch die Berechnung der Differenzen der Testscores Anwendung fand. Außerdem wird der Nutzen vom Einschluss des Interaktionsfaktors *Gruppenzugehörigkeit x Prädiktor* in die Analysen betont. Dies diene dem Aufdecken von Gruppenunterschieden bezogen auf den Einfluss der Prädiktoren sowie der Unterscheidung von Retest- und tatsächlichen Trainingseffekten⁸⁵. Da der Einschluss des jeweiligen Interaktionsfaktors die Anzahl der unabhängigen Variablen allerdings verdoppelt, kam diese Art der Prädiktionsanalyse aufgrund des hierfür zu kleinen *N* in dieser Arbeit nicht zum Einsatz, sondern die Prädiktionsanalysen erfolgten explorativ unter alleiniger Auswertung der Ergebnisse der EG. Eine Durchführung von Prädiktionsanalysen unter Berücksichtigung der CG ist in einer weiteren Arbeit zu unserer Studie nachzulesen⁸². Zudem veröffentlichten Riley et al. Empfehlungen zur Auswahl von Prognosestudien für die Durchführung von Metaanalysen⁸⁷. Zur Sicherstellung der Qualität der Studien und zur Detektion von möglicher Beeinflussung durch Bias sollten Tools wie PICOTS (Population, Index prognostic factor, Comparator prognostic factors, Outcome, Time und Setting) und QUIPS (Quality in prognostic factor studies) angewendet werden. Diese beziehen sich vor allem auf die Auswahl und Beschreibung von Stichprobe und Variablen, die Anwendung eines zum Stu-

diendesign passenden und in der Literatur validierten Prädiktionsmodells sowie die nachvollziehbare und vollständige Darstellung der Ergebnisse. Bei Anwendung dieser Tools würde diese Arbeit den Kriterien standhalten, wenngleich Riley et al. ein Augenmerk auf die potenzielle Problematik vermeintlich großer Effekte bei Prädiktionsanalysen von kleinen Stichproben legen⁸⁷.

Trotz der genannten Schwächen sind auch einige Stärken der Studie vorzuweisen. Fellman et al. waren die ersten, die ihre Studienergebnisse zu einem WMT bei Probanden mit PD veröffentlichten³⁹. Allerdings wurden die Probanden weder kognitiv noch motorisch ausführlich charakterisiert, noch wurde ein FU durchgeführt, um mittel- oder langfristige Effekte des WMT feststellen zu können. Mit der Anwendung der Level-II-Kriterien nach Litvan et al.⁵⁰ zur Auswahl von Probanden ohne PD-MCI oder PD-Demenz und mit der Durchführung einer FU-Testung drei Monate nach Abschluss des Interventionszeitraums führten wir somit die erste RCT nach CONSORT-Richtlinien⁸⁸ zu einem WMT bei Probanden mit PD durch, welche auf einer elaborierten Testbatterie sowohl zur kognitiven als auch motorisch-klinischen Charakterisierung der Probanden basierte und neben der ersten Nachtestung ein weiteres FU mit einschloss.

Außerdem trug die Gestaltung des CWMT zu einer hohen Motivation der Probanden bei, das Training täglich durchzuführen. Die Oberfläche des Anbieters NeuroNation (Synaptikon GmbH, Berlin, Germany, <https://www.neuronation.com/>) war sowohl ansprechend als auch benutzerfreundlich gestaltet. Außerdem ließ die Durchführung am eigenen Gerät zu Hause eine hohe Flexibilität zu. Die Adaptation der durchzuführenden Übungen an die eigene Leistung machte die Fortschritte für die Probanden plastisch sichtbar, was ebenfalls motivierend wirkte. Die genannten Punkte resultierten in einer sehr hohen Compliance. So führte jeder der 37 Probanden der EG mindestens 75% des WMT durch, womit 100% der Ergebnisse der EG in der Nachtestung berücksichtigt werden konnten.

5.3 Offene Fragen und Perspektiven

Die Ergebnisse unserer Studie legen nahe, dass die Implementierung eines am eigenen Gerät durchzuführenden, adaptiven CWMT in den Alltag von Patienten mit PD durchaus sinnvoll sein könnte, um die bereits etablierte medikamentöse Therapie der Erkrankung begleitend zu unterstützen. Da die Auswertung nur teilweise signifikante Verbesserungen der WM-Leistung durch das CWMT ergab, gilt es, diese Ergebnisse in weiteren Studien zu reproduzieren. Um signifikantere Ergebnisse zu erzielen, wäre es vielleicht sinnvoll, sensitivere Tests als die elaborierten und standardmäßig verwendeten WM-Tests wie die Span- oder N-back-Aufgaben zum Messen der Veränderungen der WM-Leistung einzusetzen. Interessant wäre es außerdem, den Zeitraum des FU zu verlängern, um auch langfristige Trainingseffekte aufdecken zu können.

Die Probanden unserer Studie wurden dahingehend ausgewählt, dass sie nach Level-II-Kriterien⁵⁰ keine bzw. noch keine PD-MCI oder PD-Demenz aufwiesen. Die Frage, ob ein CWMT für diese Personengruppe auch präventiv wirken, das heißt die möglicherweise durch die PD-Erkrankung eintretenden kognitiven Beeinträchtigungen hinauszögern bzw. abmildern kann, kann anhand der Durchführung nur einer Studie noch nicht beantwortet werden. Auch im Hinblick auf diese Fragestellung bleiben weiterführende Studien mit längeren Beobachtungszeiträumen abzuwarten.

Zudem lassen die durchgeführten Prädiktionsanalysen weiterhin Raum für die Frage, welche Personengruppen den größten Nutzen aus einem CWMT ziehen könnten. Die Ergebnisse der gerechneten Analysen deuten darauf hin, dass ein geringes Baselinenniveau in WM-Tests den größten Raum für Verbesserungen durch ein CWMT bietet. Da in der Literatur diesbezüglich jedoch uneinheitliche Ergebnisse zu finden sind, braucht es auch hier weitere RCT mit größeren Probandenzahlen, um die Richtung der Prädiktion zwischen Baselinenniveau und Trainingsoutcome abschließend klären zu können. Auch im Hinblick auf die anderen untersuchten Prädiktoren wäre es wünschenswert, eindeutiger Ergebnisse finden zu können, um diejenigen Personen zu ermitteln, bei denen die Anwendung eines CWMT den größten Effekt hätte. Der Einsatz einer vergleichenden Methodik zwischen EG und CG wäre zudem sinnvoll, um die Prädiktoren sicher den Trainingseffekten zuordnen zu können.

5.4 Fazit

Unsere Studie konnte zeigen, dass ein selbstständig durchführbares, adaptives CWMT eine sinnvolle und im Alltag praktikable Maßnahme darstellt, um die WM-Leistung von Probanden mit PD ohne kognitive Beeinträchtigungen kurz- und auch mittelfristig zu steigern, wenngleich die Effekte insgesamt eher klein bis mittel und nicht immer signifikant waren. Von einem CWMT scheinen insbesondere diejenigen Personen zu profitieren, die eine geringe WM-Leistung zur Baseline erbringen. Wir hoffen, mit den Ergebnissen der Studie einen Anstoß zur weiterführenden Forschung in diesem Gebiet gegeben und langfristig zur Etablierung von CWMT als begleitende präventive oder therapeutische Maßnahme bei Patienten mit PD beigetragen zu haben.

6. Literaturverzeichnis

- 1 Lee A, Gilbert RM. Epidemiology of Parkinson Disease. *Neurol Clin* 2016; **34**: 955–65. <https://doi.org/10.1016/j.ncl.2016.06.012>.
- 2 Zeng X-S, Geng W-S, Jia J-J, Chen L, Zhang P-P. Cellular and Molecular Basis of Neurodegeneration in Parkinson Disease. *Front Aging Neurosci* 2018; **10**: 109. <https://doi.org/10.3389/fnagi.2018.00109>.
- 3 Schäffer E, Berg D. Neudefinition der Parkinson-Erkrankung. *Aktuelle Neurologie* 2017; **44**: 260–66. <https://doi.org/10.1055/s-0043-103081>.
- 4 Aarsland D, Brønnick K, Larsen JP, Tysnes OB, Alves G. Cognitive impairment in incident, untreated Parkinson disease. The Norwegian ParkWest Study. *Neurology* 2009; **72**: 1121–26. <https://doi.org/10.1212/01.wnl.0000338632.00552.cb>.
- 5 Baiano C, Barone P, Trojano L, Santangelo G. Prevalence and clinical aspects of mild cognitive impairment in Parkinson's disease. A meta-analysis. *Mov Disord* 2020; **35**. <https://doi.org/10.1002/mds.27902>.
- 6 Litvan I, Aarsland D, Adler CH, et al. MDS Task Force on mild cognitive impairment in Parkinson's disease. Critical review of PD-MCI. *Mov Disord* 2011; **26**: 1814–24. <https://doi.org/10.1002/mds.23823>.
- 7 Hely MA, Reid WGJ, Adena MA, Halliday GM, Morris JGL. The Sydney multicenter study of Parkinson's disease. The inevitability of dementia at 20 years. *Mov Disord* 2008; **23**: 837–44. <https://doi.org/10.1002/mds.21956>.
- 8 Kehagia AA, Barker RA, Robbins TW. Cognitive impairment in Parkinson's disease. The dual syndrome hypothesis. *Neurodegener Dis* 2013; **11**: 79–92. <https://doi.org/10.1159/000341998>.
- 9 Kalbe E, Rehberg SP, Heber I, et al. Subtypes of mild cognitive impairment in patients with Parkinson's disease. Evidence from the LANDSCAPE study. *J Neurol Neurosurg Psychiatry* 2016; **87**. <https://doi.org/10.1136/jnnp-2016-313838>.
- 10 Lawrence BJ, Gasson N, Loftus AM. Prevalence and Subtypes of Mild Cognitive Impairment in Parkinson's Disease. *Sci Rep* 2016; **6**: 33929. <https://doi.org/10.1038/srep33929>.
- 11 Baddeley AD, Hitch G. Working Memory. Bower, G.H. (Hg.), *The Psychology of Learning and Motivation: Advances in Research and Theory* 1974: 47–89. [https://doi.org/10.1016/S0079-7421\(08\)60452-1](https://doi.org/10.1016/S0079-7421(08)60452-1).
- 12 Baddeley AD. Is working memory still working? *Am Psychol* 2001; **56**: 851–64.
- 13 Hautzel H, Mottaghy FM, Schmidt D, et al. Topographic segregation and convergence of verbal, object, shape and spatial working memory in humans. *Neurosci Lett* 2002; **323**: 156–60.

- 14 Klingberg T. Limitations in information processing in the human brain. Neuroimaging of dual task performance and working memory tasks. *Prog Brain Res* 2000; **126**: 95–102. [https://doi.org/10.1016/S0079-6123\(00\)26009-3](https://doi.org/10.1016/S0079-6123(00)26009-3).
- 15 Chai WJ, Abd Hamid AI, Abdullah JM. Working Memory From the Psychological and Neurosciences Perspectives. A Review. *Front Psychol* 2018; **9**: 401. <https://doi.org/10.3389/fpsyg.2018.00401>.
- 16 McNab F, Klingberg T. Prefrontal cortex and basal ganglia control access to working memory. *Nat Neurosci* 2008; **11**: 103–07. <https://doi.org/10.1038/nn2024>.
- 17 Klingberg T. Training and plasticity of working memory. *Trends Cogn Sci (Regul Ed)* 2010; **14**: 317–24. <https://doi.org/10.1016/j.tics.2010.05.002>.
- 18 Dahlin E, Neely AS, Larsson A, Bäckman L, Nyberg L. Transfer of learning after updating training mediated by the striatum. *Science* 2008; **320**: 1510–12. <https://doi.org/10.1126/science.1155466>.
- 19 Edin F, Klingberg T, Johansson P, McNab F, Tegnér J, Compte A. Mechanism for top-down control of working memory capacity. *Proc Natl Acad Sci U S A* 2009; **106**: 6802–07. <https://doi.org/10.1073/pnas.0901894106>.
- 20 Miller GA. The magical number seven plus or minus two. Some limits on our capacity for processing information. *Psychol Rev* 1956; **63**: 81–97.
- 21 Cowan N. The Magical Mystery Four. How is Working Memory Capacity Limited, and Why? *Curr Dir Psychol Sci* 2010; **19**: 51–57. <https://doi.org/10.1177/0963721409359277>.
- 22 Melby-Lervåg M, Redick TS, Hulme C. Working Memory Training Does Not Improve Performance on Measures of Intelligence or Other Measures of „Far Transfer“. Evidence From a Meta-Analytic Review. *Perspect Psychol Sci* 2016; **11**: 512–34. <https://doi.org/10.1177/1745691616635612>.
- 23 Sala G, Aksayli ND, Tatlidil KS, Gondo Y, Gobet F. Working memory training does not enhance older adults' cognitive skills. A comprehensive meta-analysis. *Intelligence* 2019; **77**: 101386. <https://doi.org/10.1016/j.intell.2019.101386>.
- 24 Soveri A, Antfolk J, Karlsson L, Salo B, Laine M. Working memory training revisited. A multi-level meta-analysis of n-back training studies. *Psychon Bull Rev* 2017; **24**: 1077–96. <https://doi.org/10.3758/s13423-016-1217-0>.
- 25 Karbach J, Verhaeghen P. Making working memory work. A meta-analysis of executive-control and working memory training in older adults. *Psychol Sci* 2014; **25**: 2027–37. <https://doi.org/10.1177/0956797614548725>.
- 26 Teixeira-Santos AC, Moreira CS, Magalhães R, et al. Reviewing working memory training gains in healthy older adults. A meta-analytic review of transfer for cognitive outcomes. *Neurosci Biobehav Rev* 2019; **103**: 163–77. <https://doi.org/10.1016/j.neubiorev.2019.05.009>.

- 27 Weicker J, Villringer A, Thöne-Otto A. Can impaired working memory functioning be improved by training? A meta-analysis with a special focus on brain injured patients. *Neuropsychology* 2016; **30**: 190–212. <https://doi.org/10.1037/neu0000227>.
- 28 Westerberg H, Jacobaeus H, Hirvikoski T, et al. Computerized working memory training after stroke—a pilot study. *Brain Inj* 2007; **21**: 21–29. <https://doi.org/10.1080/02699050601148726>.
- 29 Klingberg T, Forssberg H, Westerberg H. Training of working memory in children with ADHD. *J Clin Exp Neuropsychol* 2002; **24**: 781–91. <https://doi.org/10.1076/jcen.24.6.781.8395>.
- 30 Klingberg T, Fernell E, Olesen PJ, et al. Computerized training of working memory in children with ADHD—a randomized, controlled trial. *J Am Acad Child Adolesc Psychiatry* 2005; **44**: 177–86. <https://doi.org/10.1097/00004583-200502000-00010>.
- 31 Green CT, Long DL, Green D, et al. Will working memory training generalize to improve off-task behavior in children with attention-deficit/hyperactivity disorder? *Neurotherapeutics* 2012; **9**: 639–48. <https://doi.org/10.1007/s13311-012-0124-y>.
- 32 Olesen PJ, Westerberg H, Klingberg T. Increased prefrontal and parietal activity after training of working memory. *Nat Neurosci* 2004; **7**: 75–79. <https://doi.org/10.1038/nn1165>.
- 33 Brehmer Y, Rieckmann A, Bellander M, Westerberg H, Fischer H, Bäckman L. Neural correlates of training-related working-memory gains in old age. *Neuroimage* 2011; **58**: 1110–20. <https://doi.org/10.1016/j.neuroimage.2011.06.079>.
- 34 Constantinidis C, Klingberg T. The neuroscience of working memory capacity and training. *Nat Rev Neurosci* 2016; **17**: 438–49. <https://doi.org/10.1038/nrn.2016.43>.
- 35 Salmi J, Nyberg L, Laine M. Working memory training mostly engages general-purpose large-scale networks for learning. *Neurosci Biobehav Rev* 2018; **93**: 108–22. <https://doi.org/10.1016/j.neubiorev.2018.03.019>.
- 36 McNab F, Varrone A, Farde L, et al. Changes in cortical dopamine D1 receptor binding associated with cognitive training. *Science* 2009; **323**: 800–02. <https://doi.org/10.1126/science.1166102>.
- 37 Bäckman L, Nyberg L, Soveri A, et al. Effects of working-memory training on striatal dopamine release. *Science* 2011; **333**: 718. <https://doi.org/10.1126/science.1204978>.
- 38 Trujillo JP, Gerrits NJHM, Veltman DJ, Berendse HW, van der Werf YD, van den Heuvel OA. Reduced neural connectivity but increased task-related activity during working memory in de novo Parkinson patients. *Hum Brain Mapp* 2015; **36**: 1554–66. <https://doi.org/10.1002/hbm.22723>.

- 39 Fellman D, Salmi J, Ritakallio L, Ellfolk U, Rinne JO, Laine M. Training working memory updating in Parkinson's disease. A randomised controlled trial. *Neuropsychol Rehabil* 2018; 1–36. <https://doi.org/10.1080/09602011.2018.1489860>.
- 40 Fernández-Ballesteros R, Botella J, Zamarrón MD, et al. Cognitive plasticity in normal and pathological aging. *Clin Interv Aging* 2012; **7**: 15–25. <https://doi.org/10.2147/CIA.S27008>.
- 41 Dahlin E, Nyberg L, Bäckman L, Neely AS. Plasticity of executive functioning in young and older adults. Immediate training gains, transfer, and long-term maintenance. *Psychol Aging* 2008; **23**: 720–30. <https://doi.org/10.1037/a0014296>.
- 42 Lövdén M, Brehmer Y, Li S-C, Lindenberger U. Training-induced compensation versus magnification of individual differences in memory performance. *Front Hum Neurosci* 2012; **6**: 141. <https://doi.org/10.3389/fnhum.2012.00141>.
- 43 Nemmi F, Helander E, Helenius O, et al. Behavior and neuroimaging at baseline predict individual response to combined mathematical and working memory training in children. *Dev Cogn Neurosci* 2016; **20**: 43–51. <https://doi.org/10.1016/j.dcn.2016.06.004>.
- 44 Vaillancourt DE, Schonfeld D, Kwak Y, Bohnen NI, Seidler R. Dopamine overdose hypothesis. Evidence and clinical implications. *Mov Disord* 2013; **28**: 1920–29. <https://doi.org/10.1002/mds.25687>.
- 45 Opey A, Roheger M, Folkerts A-K, Skoetz N, Kalbe E. A Systematic Review on Predictors of Working Memory Training Responsiveness in Healthy Older Adults. Methodological Challenges and Future Directions. *Front Aging Neurosci* 2020; **12**: 575804. <https://doi.org/10.3389/fnagi.2020.575804>.
- 46 Edwards JD, Hauser RA, O'Connor ML, Valdés EG, Zesiewicz TA, Uc EY. Randomized trial of cognitive speed of processing training in Parkinson disease. *Neurology* 2013; **81**: 1284–90. <https://doi.org/10.1212/WNL.0b013e3182a823ba>.
- 47 Fahn, S. & Elton, R. Members of the UPDRS development committee. In: (Eds.), Florham Park, NJ: Macmillan Healthcare Information, pp. 153–163, 293–304. *Recent Developments in Parkinson's Disease* 1987: 153–163, 293–304.
- 48 Zimmermann R, Gschwandtner U, Benz N, et al. Cognitive training in Parkinson disease. Cognition-specific vs nonspecific computer training. *Neurology* 2014; **82**: 1219–26. <https://doi.org/10.1212/WNL.0000000000000287>.
- 49 Opey A, Giehl K, Rehberg S, et al. Effects of working memory training in patients with Parkinson's disease without cognitive impairment. A randomized controlled trial. *Parkinsonism Relat Disord* 2020; **72**: 13–22. <https://doi.org/10.1016/j.parkreldis.2020.02.002>.

- 50 Litvan I, Goldman JG, Tröster AI, et al. Diagnostic criteria for mild cognitive impairment in Parkinson's disease: Movement Disorder Society Task Force guidelines. *Mov Disord* 2012; **27**: 349–56. <https://doi.org/10.1002/mds.24893>.
- 51 Giehl K, Ophey A, Hammes J, et al. Working memory training increases neural efficiency in Parkinson's disease. A randomized controlled trial. *Brain Commun* 2020; **2**. <https://doi.org/10.1093/braincomms/fcaa115>.
- 52 Faul F, Erdfelder E, Lang A-G, Buchner A. G*Power 3. A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods* 2007; **39**: 175–91.
- 53 Leung IHK, Walton CC, Hallock H, Lewis SJG, Valenzuela M, Lampit A. Cognitive training in Parkinson disease. A systematic review and meta-analysis. *Neurology* 2015; **85**: 1843–51. <https://doi.org/10.1212/WNL.0000000000002145>.
- 54 Lenhard W, Lenhard A. Berechnung von Effektstärken. Dettelbach: Psychometrica, 2016. <https://www.psychometrica.de/effektstaerke.html> (Zuletzt abgerufen am 13.03.2021)
- 55 Ellis PD. Effect size equations. http://www.polyu.edu.hk/mm/effectsizafaqs/effect_size_equations2.html (Zuletzt abgerufen am 13.03.2021)
- 56 Hughes AJ, Daniel SE, Kilford L, Lees AJ. Accuracy of clinical diagnosis of idiopathic Parkinson's disease. A clinico-pathological study of 100 cases. *J Neurol Neurosurg Psychiatry* 1992; **55**: 181–84.
- 57 Emre M, Aarsland D, Brown R, et al. Clinical diagnostic criteria for dementia associated with Parkinson's disease. *Mov Disord* 2007; **22**: 1689-707; quiz 1837. <https://doi.org/10.1002/mds.21507>.
- 58 Yesavage JA, Brink TL, Rose TL, et al. Development and validation of a geriatric depression screening scale: a preliminary report. *J Psychiatr Res* 1982; **17**: 37–49.
- 59 Nasreddine ZS, Phillips NA, Bédirian V, et al. The Montreal Cognitive Assessment, MoCA. A brief screening tool for mild cognitive impairment. *J Am Geriatr Soc* 2005; **53**: 695–99. <https://doi.org/10.1111/j.1532-5415.2005.53221.x>.
- 60 Schellig D, Schuri U. NBV N-Back verbal. <https://www.schuhfried.at/test/NBV> (Zuletzt abgerufen am 13.03.2021)
- 61 Schellig D, Schuri U, Arendasy M. NBN N-Back nonverbal. <https://www.schuhfried.at/test/NBN> (Zuletzt abgerufen am 13.03.2021)
- 62 Schellig D. CORSI Corsi-Block-Tapping-Test. <https://www.schuhfried.at/test/CORSI> (Zuletzt abgerufen am 13.03.2021)

- 63 Härting C, Markowitsch HJ, Neufeld H, Calabrese P, Deisinger K, Kessler J. WMS-R, Wechsler Gedächtnistest - revidierte Fassung. deutsche Adaptation der revidierten Fassung der Wechsler Memory Scale. Bern, CH: Verlag Hans Huber, 2000.
- 64 Pertzov Y, Dong MY, Peich M-C, Husain M. Forgetting what was where. The fragility of object-location binding. *PLoS ONE* 2012; **7**: e48214. <https://doi.org/10.1371/journal.pone.0048214>.
- 65 Pertzov Y, Heider M, Liang Y, Husain M. Effects of healthy ageing on precision and binding of object location in visual short term memory. *Psychol Aging* 2015; **30**: 26–35. <https://doi.org/10.1037/a0038396>.
- 66 Liang Y, Pertzov Y, Nicholas JM, et al. Visual short-term memory binding deficit in familial Alzheimer's disease. *Cortex* 2016; **78**: 150–64. <https://doi.org/10.1016/j.cortex.2016.01.015>.
- 67 Horn W. Leistungsprüfsystem: L-P-S. Göttingen: Verl. für Psychologie Hogrefe, 1983.
- 68 Sturm, W., Willmes, K., & Horn, W. Leistungsprüfsystem für 50- bis 90-Jährige (LPS 50+). Göttingen: Hogrefe, 2015.
- 69 Chandler MJ, Lacritz LH, Hynan LS, et al. A total score for the CERAD neuropsychological battery. *Neurology* 2005; **65**: 102–06. <https://doi.org/10.1212/01.wnl.0000167607.63000.38>.
- 70 Bäumler G. Farb-Wort-Interferenztest (FWIT) nach J.R. Stroop. Göttingen, D.: Hogrefe, 1985.
- 71 Schretlen D, Bobholz JH, Brandt J. Development and psychometric properties of the brief test of attention. *The Clinical Neuropsychologist* 1996; **10**: 80–89. <https://doi.org/10.1080/13854049608406666>.
- 72 Kalbe, E., Reinhold, N., Ender, U., & Kessler, J. Aphasie-Check-Liste (ACL). Köln: Pro-Log, Therapie- und Lernmittel, 2002.
- 73 Schwarzer R, ed. Skalen zur Erfassung von Lehrer- und Schülermerkmalen. Dokumentation der psychometrischen Verfahren im Rahmen der wissenschaftlichen Begleitung des Modellversuchs Selbstwirksame Schulen. Berlin: R. Schwarzer, 1999.
- 74 Vogler A, Vanbellingen T. Das Einfrieren beurteilen – Freezing of Gait Questionnaire. *ergopraxis* 2015; **8**: 36–37. <https://doi.org/10.1055/s-0035-1552966>.
- 75 Buysse DJ, Reynolds CF, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh sleep quality index. A new instrument for psychiatric practice and research. *Psychiatry Research* 1989; **28**: 193–213. [https://doi.org/10.1016/0165-1781\(89\)90047-4](https://doi.org/10.1016/0165-1781(89)90047-4).
- 76 Jessen F, Wiese B, Bickel H, et al. Prediction of dementia in primary care patients. *PLoS ONE* 2011; **6**: e16852. <https://doi.org/10.1371/journal.pone.0016852>.
- 77 Tomaszewski Farias S, Mungas D, Harvey DJ, Simmons A, Reed BR, Decarli C. The measurement of everyday cognition. Development and validation of a short form of the

- Everyday Cognition scales. *Alzheimers Dement* 2011; **7**: 593–601.
<https://doi.org/10.1016/j.jalz.2011.02.007>.
- 78 Tomlinson CL, Stowe R, Patel S, Rick C, Gray R, Clarke CE. Systematic review of levodopa dose equivalency reporting in Parkinson's disease. *Mov Disord* 2010; **25**: 2649–53.
<https://doi.org/10.1002/mds.23429>.
- 79 Giehl K, Ophay A, Reker P, et al. Effects of Home-Based Working Memory Training on Visuo-Spatial Working Memory in Parkinson's Disease. A Randomized Controlled Trial. *J Cent Nerv Syst Dis* 2020; **12**: 1179573519899469.
<https://doi.org/10.1177/1179573519899469>.
- 80 Ellis PD. *The Essential Guide to Effect Sizes*. Cambridge: Cambridge University Press, 2010.
- 81 Berkovits I, Hancock GR, Nevitt J. Bootstrap Resampling Approaches for Repeated Measure Designs. Relative Robustness to Sphericity and Normality Violations. *Educational and Psychological Measurement* 2016; **60**: 877–92.
<https://doi.org/10.1177/00131640021970961>.
- 82 Ophay A, Rehberg S, Giehl K, et al. Predicting Working Memory Training Responsiveness in Parkinson's Disease. Both „System Hardware“ and Room for Improvement Are Needed. *Neurorehabil Neural Repair* 2021; **35**: 117–30.
<https://doi.org/10.1177/1545968320981956>.
- 83 Taya F, Sun Y, Babiloni F, Thakor N, Bezerianos A. Brain enhancement through cognitive training. A new insight from brain connectome. *Front Syst Neurosci* 2015; **9**: 44.
<https://doi.org/10.3389/fnsys.2015.00044>.
- 84 Bland JM, Altman DG. Some examples of regression towards the mean. *BMJ* 1994; **309**: 780. <https://doi.org/10.1136/bmj.309.6957.780>.
- 85 Roheger M, Folkerts A-K, Krohm F, Skoetz N, Kalbe E. Prognostic Models for Changes in Memory Performance After Memory Training in Healthy Older Adults. A Systematic Review. *J Cogn Enhanc* 2020; **25**: 97. <https://doi.org/10.1007/s41465-020-00194-0>.
- 86 Mattes A, Roheger M. Nothing wrong about change. The adequate choice of the dependent variable and design in prediction of cognitive training success. *BMC Med Res Methodol* 2020; **20**: 296. <https://doi.org/10.1186/s12874-020-01176-8>.
- 87 Riley RD, Moons KGM, Snell KIE, et al. A guide to systematic review and meta-analysis of prognostic factor studies. *BMJ* 2019; **364**: k4597. <https://doi.org/10.1136/bmj.k4597>.
- 88 Schulz KF, Altman DG, Moher D, Group tC. CONSORT 2010 Statement. Updated guidelines for reporting parallel group randomised trials. *BMC Medicine* 2010; **8**: 18.
<https://doi.org/10.1186/1741-7015-8-18>.

7. Anhang

7.1 Tabellenverzeichnis

Tabelle 1. Studiendesign	16
Tabelle 2. Zur Studienteilnahme erforderliche Ein- und Ausschlusskriterien.....	17
Tabelle 3. Übersicht aller klinischen und neuropsychologischen Verfahren	20
Tabelle 4. Übersicht aller Fragebögen	21
Tabelle 5. Gruppenvergleich der Interventions- und Kontrollgruppe – Baseline- Charakteristika	30
Tabelle 6. Gruppenvergleich der Interventions- und Kontrollgruppe im Ausgangsniveau der Arbeitsgedächtnistests	31
Tabelle 7. Mixed 2x2-ANOVAs: Deskriptive Statistik der Interventionsgruppe	32
Tabelle 8. Mixed 2x2-ANOVAs: Deskriptive Statistik der Kontrollgruppe	32
Tabelle 9. Trainingseffekte: Haupteffekt Zeit	33
Tabelle 10. Trainingseffekte: Haupteffekt Gruppenzugehörigkeit	34
Tabelle 11. Trainingseffekte: Interaktionseffekt Zeit x Gruppenzugehörigkeit	34
Tabelle 12. Trainingseffekte: t-Tests für abhängige Stichproben: Interventionsgruppe	35
Tabelle 13. Trainingseffekte: t-Tests für abhängige Stichproben: Kontrollgruppe	35
Tabelle 14. Prädiktionsanalyse: t2 – t1	38
Tabelle 15. Prädiktionsanalyse: t3 – t1	39
Tabelle 16. Mixed 2x2-ANOVAs: Deskriptive Statistik der Interventionsgruppe	58
Tabelle 17. Mixed 2x2-ANOVAs: Deskriptive Statistik der Kontrollgruppe	58
Tabelle 18. Prädiktionsanalyse: t2 – t1	61
Tabelle 19. Prädiktionsanalyse: t3 – t1	62

7.2 Abbildungsverzeichnis

Abbildung 1. Studienverlauf bezüglich der Probandenzahl.....	29
Abbildung 2. Vergleich der Mittelwerte von Interventions- und Kontrollgruppe im Wechsler Memory Scale-Revised vorwärts (Prozentränge).....	36
Abbildung 3. Vergleich der Mittelwerte von Interventions- und Kontrollgruppe im N-back verbal (Prozentränge der Richtigen).....	37
Abbildung 4. Vergleich der Mittelwerte von Interventions- und Kontrollgruppe im Corsi-Block-Tapping-Test rückwärts (Prozentränge der unmittelbaren Blockspanne)	37
Abbildung 5. Vergleich der Mittelwerte von Interventions- und Kontrollgruppe im Wechsler Memory Scale-Revised rückwärts (Prozentränge).....	59
Abbildung 6. Vergleich der Mittelwerte von Interventions- und Kontrollgruppe im N-back nonverbal (Prozentränge der Richtigen).....	60
Abbildung 7. Vergleich der Mittelwerte von Interventions- und Kontrollgruppe im Corsi-Block-Tapping-Test vorwärts (Prozentränge der unmittelbaren Blockspanne)	60

Zu 4.3.1 Mixed ANOVAs

Tabelle 16. Mixed 2x2-ANOVAs: Deskriptive Statistik der Interventionsgruppe

Domäne	t1		t2		t3	
	M	SD	M	SD	M	SD
Verbales Arbeitsgedächtnis						
WMS-R vw.	56.63	26.43	57.16	27.87	64.54	27.30
WMS-R rw.	49.03	27.75	52.62	28.02	54.11	30.79
NBV (WTS)	39.57	27.13	50.22	23.62	57.20	24.85
Non-verbales Arbeitsgedächtnis						
NBN (WTS)	47.63	26.38	54.11	22.44	54.11	28.63
CORSI vw. (WTS)	42.80	27.24	47.16	24.18	44.37	23.84
CORSI rw. (WTS)	42.74	24.55	47.41	25.86	49.97	24.43

Notiz. CORSI, Corsi-Block-Tapping-Test; M, Mittelwert; NBN, N-back nonverbal; NBV, N-back verbal; rw., rückwärts; SD, Standardabweichung; vw., vorwärts; WMS-R, Wechsler Memory Scale-Revised; WTS, Wiener Testsystem.

Die Zeitpunkte *t1*, *t2* und *t3* entsprechen der Vor-, Nach- und Follow-Up-Testung. Angegeben sind die Mittelwerte und deren Standardabweichung der Prozentränge der erreichten Punktzahlen in den beiden Subtests der WMS-R, der Anzahl der Richtigen in den Subtests NBV und NBN des WTS sowie der unmittelbaren Blockspanne in den Subtests des CORSI des WTS. Die Werte von *t1* entsprechen den Werten beim Vergleich *t1* x *t3*.

Tabelle 17. Mixed 2x2-ANOVAs: Deskriptive Statistik der Kontrollgruppe

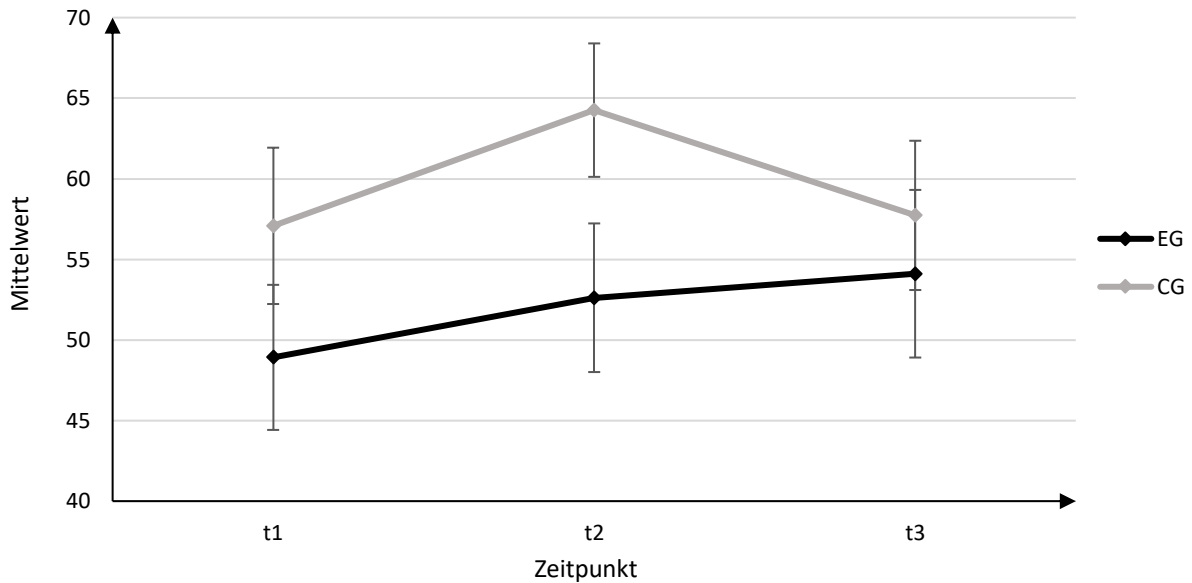
Domäne	t1		t2		t3	
	M	SD	M	SD	M	SD
Verbales Arbeitsgedächtnis						
WMS-R vw.	68.65	24.23	70.82	25.15	67.46	27.14
WMS-R rw.	57.89	29.91	64.26	25.50	57.73	28.15
NBV (WTS)	49.95	28.10	56.27	26.78	56.89	27.19
Non-verbales Arbeitsgedächtnis						
NBN (WTS)	53.49	29.12	61.35	23.31	60.16	26.10
CORSI vw. (WTS)	46.22	28.00	49.73	26.80	51.05	28.55
CORSI rw. (WTS)	45.51	24.64	48.22	25.85	46.57	27.61

Notiz. CORSI, Corsi-Block-Tapping-Test; M, Mittelwert; NBN, N-back nonverbal; NBV, N-back verbal; rw., rückwärts; SD, Standardabweichung; vw., vorwärts; WMS-R, Wechsler Memory Scale-Revised; WTS, Wiener Testsystem.

Die Zeitpunkte *t1*, *t2* und *t3* entsprechen der Vor-, Nach- und Follow-Up-Testung. Angegeben sind die Mittelwerte und deren Standardabweichung der Prozentränge der erreichten Punktzahlen in den beiden Subtests der WMS-R, der Anzahl der Richtigen in den Subtests NBV und NBN des WTS sowie der unmittelbaren Blockspanne in den Subtests des CORSI des WTS. Die Werte von *t1* entsprechen den Werten beim Vergleich *t1* x *t3*.

Zu 4.3.2 t-Tests für abhängige Stichproben

Abbildung 5. Vergleich der Mittelwerte von Interventions- und Kontrollgruppe im Wechsler Memory Scale-Revised rückwärts (Prozentränge)

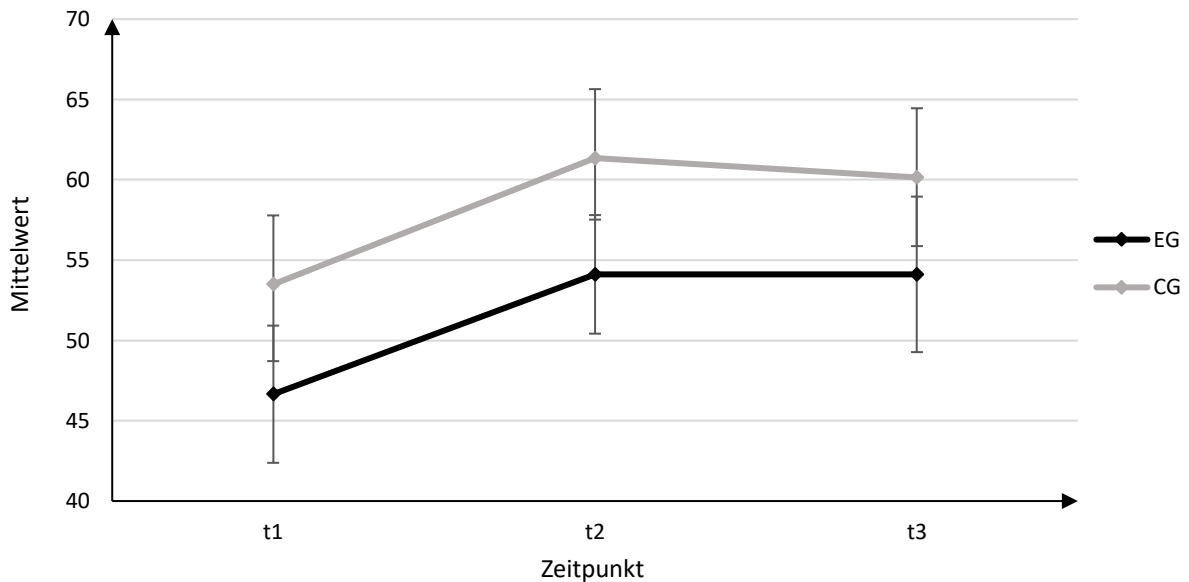


Notiz. CG, Kontrollgruppe (engl. control group); EG, Interventionsgruppe (engl. experimental group); WMS-R, Wechsler Memory Scale-Revised.

Die Zeitpunkte t_1 , t_2 und t_3 entsprechen der Vor-, Nach- und Follow-Up-Testung.

Verwendet wurden die errechneten Mittelwerte und Standardfehler aus den t -Tests für abhängige Stichproben. Für die Mittelwerte und Standardfehler des Zeitpunkts t_1 wurden die Werte des Vergleichs $t_1 \times t_2$ verwendet.

Abbildung 6. Vergleich der Mittelwerte von Interventions- und Kontrollgruppe im N-back nonverbal (Prozentränge der Richtigen)

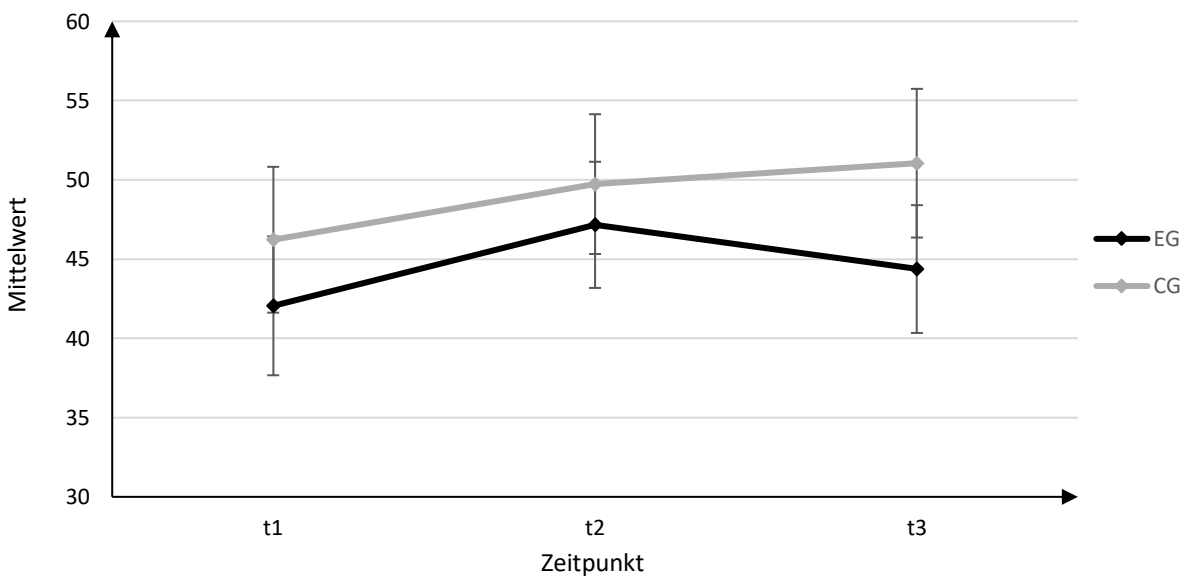


Notiz. CG, Kontrollgruppe (engl. control group); EG, Interventionsgruppe (engl. experimental group); NBN, N-back nonverbal.

Die Zeitpunkte t_1 , t_2 und t_3 entsprechen der Vor-, Nach- und Follow-Up-Testung.

Verwendet wurden die errechneten Mittelwerte und Standardfehler aus den t -Tests für abhängige Stichproben. Für die Mittelwerte und Standardfehler des Zeitpunkts t_1 wurden die Werte des Vergleichs $t_1 \times t_2$ verwendet.

Abbildung 7. Vergleich der Mittelwerte von Interventions- und Kontrollgruppe im Corsi-Block-Tapping-Test vorwärts (Prozentränge der unmittelbaren Blockspanne)



Notiz. CG, Kontrollgruppe (engl. control group); CORSI, Corsi-Block-Tapping-Test; EG, Interventionsgruppe (engl. experimental group).

Die Zeitpunkte t_1 , t_2 und t_3 entsprechen der Vor-, Nach- und Follow-Up-Testung.

Verwendet wurden die errechneten Mittelwerte und Standardfehler aus den t -Tests für abhängige Stichproben. Für die Mittelwerte und Standardfehler des Zeitpunkts t_1 wurden die Werte des Vergleichs $t_1 \times t_2$ verwendet.

Zu 4.4 Prädiktionsanalysen

Tabelle 18. Prädiktionsanalyse: $t_2 - t_1$

Domäne	R^2 (korr. R^2)	F-Test	nicht standardisierter Regressionskoeffizient B (Standardfehler)				
			Alter	Bildung	LEDD	UPDRS Part III	Baseline- niveau
Verbales Arbeitsgedächtnis							
WMS-R vw.		$F(5, 31)$					
	.262 (.143)	= 2.201, $p = .080$	0.593 (0.462)	-1.651 (1.395)	0.004 (0.010)	-0.464 (0.502)	-0.335* (0.156)
WMS-R rw.		$F(5, 31)$					
	.325 (.216)	= 2.980, $p = .026$	-0.030 (0.504)	0.247 (1.465)	0.018+ (0.011)	0.244 (0.551)	-0.449** (0.164)
NBV (WTS)		$F(5, 31)$					
	.463 (.376)	= 5.332, $p = .001$	-0.013 (0.448)	1.677 (1.320)	0.009 (0.009)	0.056 (0.477)	-0.695*** (0.145)
Non-verbales Arbeitsgedächtnis							
NBN (WTS)		$F(5, 31)$					
	.541 (.467)	= 7.307, $p < .001$	-0.441 (0.407)	-0.974 (1.183)	-0.014 (0.008)	0.599 (0.436)	-0.708*** (0.133)
CORSI vw. (WTS)		$F(5, 31)$					
	.431 (.340)	= 4.702, $p = .003$	-0.914* (0.396)	-1.271 (1.140)	0.004 (0.008)	-0.203 (0.425)	-0.550*** (0.128)
CORSI rw. (WTS)		$F(5, 31)$					
	.406 (.311)	= 4.245, $p = .005$	0.336 (0.416)	-0.322 (1.147)	-0.013 (0.008)	-1.227** (0.421)	-0.399** (0.141)

Notiz. CORSI, Corsi-Block-Tapping-Test; LEDD, Levodopa equivalent daily dose; NBN, N-back nonverbal; NBV, N-back verbal; rw., rückwärts; UPDRS, Unified Parkinson's Disease Rating Scale; vw., vorwärts; WMS-R, Wechsler Memory Scale-Revised; WTS, Wiener Testsystem.

Die Zeitpunkte t_1 und t_2 entsprechen der Vor- bzw. Nachtstung.

Für die Prädiktoren angegeben sind die nicht standardisierten Regressionskoeffizienten B mit Standardfehler und dem Signifikanzniveau α , eingeteilt wie folgt:

*** $\leq .001$; ** $\leq .010$; * $\leq .050$; + $\leq .100$.

Tabelle 19. Prädiktionsanalyse: $t3 - t1$

Domäne	R^2 (korr. R^2)	F-Test	nicht standardisierter Regressionskoeffizient B (Standardfehler)				
			Alter	Bildung	LEDD	UPDRS Part III	Baseline- niveau
Verbales Arbeitsgedächtnis							
WMS-R vw.		$F(5, 29)$					
	.323 (.206)	= 2.762, $p = .037$	-0.596 (0.473)	1.119 (1.366)	-0.011 (0.010)	-0.588 (0.495)	-0.442** (0.154)
WMS-R rw.		$F(5, 29)$					
	.339 (.225)	= 2.980, $p = .027$	-0.518 (0.577)	1.145 (1.619)	0.014 (0.012)	1.119+ (0.615)	-0.432* (0.181)
NBV (WTS)		$F(5, 29)$					
	.361 (.251)	= 3.277, $p = .018$	-0.047 (0.476)	1.223 (1.350)	-0.005 (0.010)	0.449 (0.492)	-0.579*** (0.150)
Non-verbales Arbeitsgedächtnis							
NBN (WTS)		$F(5, 29)$					
	.367 (.258)	= 3.362, $p = .016$	-0.007 (0.485)	1.859 (1.357)	0.002 (0.010)	1.320* (0.505)	-0.392* (0.153)
CORSI vw. (WTS)		$F(5, 29)$					
	.523 (.441)	= 6.372, $p < .001$	-0.715 (0.482)	0.563 (1.337)	-0.004 (0.010)	-0.624 (0.500)	-0.746*** (0.149)
CORSI rw. (WTS)		$F(5, 29)$					
	.199 (.060)	= 1.437, $p = .241$					

Notiz. CORSI, Corsi-Block-Tapping-Test; LEDD, Levodopa equivalent daily dose; NBN, N-back nonverbal; NBV, N-back verbal; rw., rückwärts; UPDRS, Unified Parkinson's Disease Rating Scale; vw., vorwärts; WMS-R, Wechsler Memory Scale-Revised; WTS, Wiener Testsystem.

Die Zeitpunkte $t1$ und $t3$ entsprechen der Vor- bzw. Follow-Up-Testung.

Für die Prädiktoren angegeben sind die nicht standardisierten Regressionskoeffizienten B mit Standardfehler und dem Signifikanzniveau α , eingeteilt wie folgt:

*** $\leq .001$; ** $\leq .010$; * $\leq .050$; + $\leq .100$.