Aus der Klinik für Gerontopsychiatrie und -psychotherapie
des Alexianer Krankenhauses Maria Hilf Krefeld
Akademisches Lehrkrankenhaus
der Heinrich-Heine-Universität Düsseldorf
Chefarzt: Prof. Dr. med. Dipl.-Psych. Ralf Ihl

Entwicklung und Validierung einer Parallelversion des Tests zur Früherkennung von Demenzen mit Depressionsabgrenzung (TFDD)

Dissertation

zur Erlangung des Grades eines Doktors der Medizin
der Medizinischen Fakultät der Heinrich-Heine-Universität
Düsseldorf

vorgelegt von
Andreas Rigas
2022

Als Inauguraldissertation gedruckt mit der Genehmigung der Medizinischen Fakultät der Heinrich-Heine-Universität Düsseldorf

gez.:
Dekan: Prof. Dr. med. Nikolaj Klöcker
Erstgutachter: Prof. Dr. med. Dipl.-Psych. Ralf Ihl
Zweitgutachter: Prof. Dr. med. Stefan Wilm

Zusammenfassung

Für die Früherkennung von dementiellen Syndromen sind neuropsychologische Testverfahren unabdingbar. Die kurzfristige Wiederholung von Tests ist allerdings aufgrund des testspezifischen Lerneffekts problematisch. Um den Lerneffekt zu minimieren, können Paralleltests verwenden werden. Der Test zur Früherkennung von Demenzen mit Depressionsabgrenzung (TFDD) hat sich aufgrund der guten Testgütekriterien besonders bei der Frühdiagnostik der Demenz bei Alzheimer-Krankheit bewährt. Ziel dieser Arbeit war die Entwicklung und Validierung einer Parallelversion des TFDD.

Zunächst wurden Parallelitems auf der Grundlage der Theorie und Empirie entwickelt, mit denen die gleichen kognitiven Domänen untersucht werden und die gleichen Messeigenschaften wie die Originalitems aufweisen. Die Depressionsitems wurden bei nicht zu erwartendem Lerneffekt unverändert belassen. Die neu entwickelten Items wurden mit einer ersten Stichprobe von 30 Probanden mit einer Demenz bei Alzheimer-Krankheit (AG1) überprüft und mit den Originalitems verglichen, um gegebenenfalls Mängel aufzudecken. Nach der initialen Überprüfung gab es keinen Anlass für eine Überarbeitung des Tests, sodass weiterhin Daten für die erweiterte statistische Analyse erhoben wurden.

Mit den resultierenden Items wurden 60 Probanden mit einer Demenz bei Alzheimer-Krankheit (AG2) in einem leichten bis zu einem mittelschweren Stadium (*Global Deterioration Scale*: III-V/VII) mit einem Durchschnittsalter von 75,4 Jahren untersucht. Für die Kontrollgruppe (KG) wurden 40 kognitiv unauffällige Probanden mit einem Durchschnittsalter von 74 Jahren untersucht.

Durch eine Faktorenanalyse wurde die Eindimensionalität des Paralleltests bestätigt. Es fand sich ein Generalfaktor, durch den sich 63,3 % der Varianz erklären lässt. Der Schwierigkeitsindex jedes Items war für beide Versionen des Tests nahezu identisch und somit als positiv für die Annahme der Parallelität der Items zu werten. Bei der Reliabilitätsprüfung erzielte die Retest-Reliabilität einen Wert von $r_s = .985$, $p < .001$, $n = 40$. Für den gesamten Paralleltest fand sich ein Cronbachs-Alpha-Wert von .893 und somit eine sehr gute Item-Reliabilität. Alle Items korrelierten $r_{it} = .58$ oder höher. Bei der Objektivitätsprüfung betrug die Inter-Rater-Reliabilität $r_s = .995$, $p < .001$, $n = 20$. Bei der Validitätsprüfung konnte eine signifikante Korrelation zwischen den Punktwerten vom Parallel- und Originaltest ($r_s = .969$, $p < .001$, $n = 100$) nachgewiesen werden. Anhand des Youden-Index wurde ein optimaler *Cut-Off*-Wert von > 38 ermittelt. Bei diesem Wert erreichte der Paralleltest eine Sensitivität von 97 % bzw. eine Spezifität von 98 % und somit eine sehr gute Zuordnungsfähigkeit zwischen AG2 und KG.

Die Parallelversion des TFDD entspricht in allen Validierungskriterien der Originalversion. Somit stehen zwei Versionen des Tests zur Verfügung, die Verlaufsmessungen bzw. Testwiederholungen mit einem minimalen testspezifischen Lerneffekt ermöglichen, ohne Qualitätseinbuße gegenüber dem Originaltest.

Abstract

Neuropsychological test procedures are essential for the early detection of dementia syndromes. The short-term repetition of tests is problematic due to the test-specific practice effect. Parallel tests can be used to minimize this practice effect. The Test for the early Detection of Dementia with Differentiation of Depression (TE4D) has proven itself to be particularly effective in the early diagnosis of Alzheimer's dementia due to the good quality test criteria. The aim of this dissertation was to develop and validate a parallel version of the TE4D.

The parallel items were developed on the basis of theory and empiricism. They had to be able to investigate the same cognitive domains and have the same measurement properties as the original items. The depression items were left unchanged, because a learning effect was not to be expected. The newly developed items were then administrated to an initial sample of 30 people with a diagnosis of Alzheimer's disease (AG1) and compared to the original items in order to identify any deficiencies. After the initial review, there was no reason to revise the test. Consequently, data for the extended statistical analysis was collected.

Sixty people with dementia due to Alzheimer's disease (AG2) in mild to moderate stages (GDS-*Global Deterioration Scale*: III-V/VII) with an average age of 75.4 years were examined with the developed items. For the control group (KG), 40 cognitively intact people with an average age of 74 years were examined.

The one-dimensionality of the parallel test was confirmed by a factor analysis. A general factor was found to explain 63.3 % of the variance. The difficulty index of each item was almost identical for both versions of the test and was therefore rated as positive for the assumption that the items were parallel. In the reliability test, the retest reliability had values of $r_s = .985$, $p < .001$, $n = 40$. The Cronbach's alpha value was .893 for the entire parallel test, and therefore implied strong item reliability. All items had correlations of $r_{it} = .58$ or higher. In the objectivity test, the inter-rater reliability was $r_s = .995$, $p < .001$, $n = 20$. The validity check showed a significant correlation between the scores from the parallel and original tests ($r_s = .969$, $p < .001$, $n = 100$). Using the Youden index, an optimal cut-off value of > 38 was determined. With this value, the parallel test achieved a sensitivity of 97 % and specificity of 98 % and thus demonstrated a strong capacity to discriminate between AG2 and KG.

The parallel version of the TE4D corresponds to the original version in all validation criteria. Thus, two versions of the TE4D are available without any loss of quality compared to the original test, which enables progress measurements or test repetitions with minimal test-specific learning effect.

Abkürzungsverzeichnis

a	Signifikanzniveau
Ab	beta-Amyloid
ABCA7	*ATP Binding Cassette Subfamily A Member 7*
ADAS	*Alzheimer Disease Asessment Scale*
AG1	Demenz bei Alzheimer-Krankheit-Gruppe (Voruntersuchung)
AG2	Demenz bei Alzheimer-Krankheit-Gruppe (Validierungsstudie)
APA	*American Psychiatric Association*
APH-1	*Anterior pharynx-defective*
APP	Amyloid-Vorläufer-Protein
ApoE	Apolipoprotein E
AWMF	Arbeitsgemeinschaft Wissenschaftlich Medizinischer Fachgesellschaften
BPSD	*Behavioral and Psychological Symptoms of Dementia*
CCA	*Correct Classification Accuracy*
CDK5	*Cyclin-dependent cinase 5*
CERAD	*Consortium to Establish a Registry for Alzheimers' Disease*
DemTect	Demenz-Detection-Test
EEG	Elektroenzephalogramm
FDG	Fluorodesoxyglucose
GDS	*Global Deterioration Scale for Assessment of Primary Degenerative Dementia*
ICD-10	*10th revision of the International Statistical Classification of Diseases and Related Health Problems*
IWG	*International Working Group*
kD	*Kilodalton*
M	Mittelwert
MCI	*Mild Cognitive Impairment*
MIBG	Metaiodobenzylguanidin

MMST	Mini-Mental-Status-Test
MoCA	*Montreal Cognitive Assessment*
MRT	Magnetresonanztomographie
MS4A	*Membrane-spanning 4A genes*
n	Größe der Stichprobe
NIA-AA	*National Institute on Aging* and *Alzheimer's Association*
NINCDS-ADRDA	*National Institute of Neurological Disorders and Stroke - Alzheimer's Disease and Related Disorders Association*
PET	Positronenemissionstomographie
PICALM	*Phosphatidylinositol Binding Clathrin Assembly Protein*
r	Korrelationskoeffizient
r_{it}	Trennschärfekoeffizient
r_s	Spearman-Korrelationskoeffizient
ROC	*Receiver-Operator-Characteristics*
SD	*Standard Deviation*
SKT	Syndrom-Kurz-Test
SPECT	*Single-photon emission computed tomography*
SPSS	*Statistical Package for the Social Sciences*
SSRI	*Selective serotonin reuptake inhibitor*
Tau	Tau-Protein
TDP-43	*Transactive response DNA binding protein 43 kDa*
TFDD	Der Test zur Früherkennung von Demenzen mit Depressionsabgrenzung
TFDD-P	Der Test zur Früherkennung von Demenzen mit Depressionsabgrenzung-Parallelversion
Te4D	*Test for the early Detection of Dementia with Discrimination from Depression*
TREM2	*Triggering receptor expressed on myeloid cells 2*
WFSBP	*World Federation of Societies of Biological Psychiatry*

Inhaltsverzeichnis

1 Einleitung 1

1.1 Hintergrund 1

1.2 Validierung neuropsychologischer Verfahren 4

1.3 Demenzerkrankungen 8

1.4 Leichte Kognitive Störung 11

1.5 Alzheimer-Krankheit 13

1.5.1 Kognitive Einschränkungen 13

1.5.2 Affekt und Verhalten 15

1.5.3 Pathophysiologie 16

1.5.4 Genetik 19

1.5.5 Diagnosekriterien 21

1.5.6 Therapie 28

1.6 Neuropsychologische Screening-Verfahren 30

1.6.1 Mini-Mental-Status-Test (MMST) 30

1.6.2 *Montreal Cognitive Assessment* (MoCA) 30

1.6.3 Demenz-Detection-Test (DemTect) 31

1.6.4 Syndrom-Kurz-Test (SKT) 31

1.6.5 Test zur Früherkennung von Demenzen mit Depressionsabgrenzung (TFDD) 32

1.6.6 Ausführliche neuropsychologische Testverfahren 34

1.7 Ziele der Arbeit 36

2 Methodik 37

2.1 Stichprobe 37

2.2 Einschluss- und Ausschlusskriterien 38

2.3 Prozedere 38

2.3.1 Testkonstruktion 38

2.3.2 Voruntersuchung der Parallelitems 41

2.3.3 Itemanalyse 41

2.3.4 Reliabilitätsprüfung 42

2.3.5 Objektivitätsprüfung 42

2.3.6 Validitätsprüfung 42

2.4 Verwendete neuropsychologische Testverfahren 43

2.5 Datenerhebung 46

2.6 Statistische Auswertung 46

2.6.1 Voruntersuchung 46
2.6.2 Faktorenanalyse 47
2.6.3 Prüfung der Reliabilität und Objektivität 47
2.6.4 Validitätsprüfung 47
2.7 Ethikvotum 48
3 Ergebnisse 49
3.1 Zusammenstellung des Paralleltests 49
3.1.1 Unmittelbare und verzögerte Reproduktion (Item 1, Item 8) 49
3.1.2 Zeitliche Orientierung und Kategorisieren (Items 2–5) 50
3.1.3 Befolgen von Anweisungen (Item 6) 50
3.1.4 Konstruktive Praxis (Item 7) 50
3.1.5 Wortflüssigkeit (Item 9) 51
3.2 Untersuchung des Vortests 51
3.3 Faktorenanalyse 52
3.4 Itemschwierigkeit 54
3.5 Interne Konsistenz 59
3.6 Retest-Reliabilität 61
3.7 Inter-Rater-Reliabilität 61
3.8 Konstruktvalidität 61
3.9 Kriteriumsvalidität 62
3.10 Trennung zwischen kognitiv gesunden Personen und Menschen mit beginnender Demenz bei Alzheimer-Krankheit 64
3.11 Trennung zwischen kognitiv gesunden Personen und Menschen mit mittelschwerer Demenz bei Alzheimer-Krankheit 66
4 Diskussion 67
4.1 Parallelität der beiden Testversionen 67
4.2 Validierung des Paralleltests 69
4.3 Studieneinschränkungen 74
4.4 Ausblick 76
5 Literaturverzeichnis 77
6 Anhang 95

1 Einleitung

1.1 Hintergrund

Der Begriff ‚Demenz' lässt sich aus dem lateinischen Wort ‚dementia' ableiten, das wörtlich übersetzt ‚ohne Verstand' bedeutet. Hiermit wird direkt das Hauptmerkmal von Demenzerkrankungen beschrieben: Kennzeichnend für die Demenz ist eine Verschlechterung bis hin zum Verlust der kognitiven Funktionen.

Demenzerkrankungen und insbesondere die Demenz bei Alzheimer-Krankheit rücken in den letzten Jahrzehnten aufgrund der epidemiologischen Daten immer stärker in das öffentliche Interesse. Im Jahr 2013 belief sich die Prävalenzrate der Demenzerkrankungen der über 60-Jährigen weltweit auf 5–7 %. Die höchste Prävalenz fand sich mit 8,5 % in Südamerika (Prince et al., 2013).

Unsere Gesellschaft wird in Zukunft eine zunehmende Zahl von Demenzbetroffenen versorgen müssen. Im Jahr 2020 wurde geschätzt, dass weltweit 50 Millionen Menschen mit einer Demenz gelebt haben. Diese Zahl soll in den nächsten Jahrzehnten insbesondere in Entwicklungsländern rasant ansteigen und eventuell im Jahr 2050 die 150 Millionen erreichen (Livingston et al., 2020). Die Kosten, die infolge der Demenzerkrankungen entstehen, sind allein in den USA von 604 Milliarden Dollar im Jahr 2010 um 35,4 % auf 818 Milliarden Dollar im Jahr 2015 angestiegen (Wimo et al., 2017).

Aus den neuesten Studien gehen allerdings überraschende Ergebnisse bezüglich der Inzidenz und teilweise Prävalenz der Erkrankung hervor. Im Jahr 2016 wurde die Auswertung der über 40 Jahre zurückgehenden Framingham-Studie veröffentlicht. Diese zeigte einen kontinuierlichen Rückgang der Inzidenz an Demenzerkrankten bei den über 5000 Teilnehmern. Als Studieneinschränkung ist allerdings die bessere Ausbildung und europäische Herkunft aller Studienteilnehmer zu werten (Satizabal et al., 2016). Auch in weiteren Studien aus England, Schweden, Frankreich, Spanien und den USA finden sich Hinweise auf eine zurückgehende Inzidenz und Prävalenz. Diese Ergebnisse sind teilweise durch die bessere Einstellung der kardiovaskulären Risikofaktoren (arterielle Hypertonie, Hyperlipidämie, Nikotinkonsum) erklärbar, können aber nicht vollständig darauf zurückgeführt werden (Wu et al., 2017).

Die Früherkennung der Demenzerkrankung dient dazu, den Verlauf der Krankheit positiv zu beeinflussen und die Selbstständigkeit der Betroffenen länger aufrechtzuerhalten. Zwar können neurodegenerative Demenzerkrankungen noch nicht geheilt werden, es gibt allerdings Hinweise darauf, dass nicht medikamentöse Maßnahmen wie etwa gesunde Ernährung, körperliche und geistige Aktivität sowie soziale Interaktion und Integration die Neuroplastizität positiv beeinflussen können (Reimers et al., 2012; Lin et al., 2013). Somit können vorhandene Ressourcen aktiviert werden, um eine hohe Lebensqualität bis in fortgeschrittene Stadien hinein zu erhalten und eventuell das Eintreten von Demenz und Pflegebedürftigkeit hinauszuzögern. Je früher ein dementielles Syndrom erkannt und diagnostiziert wird, desto effektiver können diese Maßnahmen wirken.

Es existieren bis dato keine diagnostischen Instrumente zur zuverlässigen Vorhersage für das Auftreten einer Demenz. Demzufolge ist es von besonderer Relevanz, die ersten klinischen Anzeichen der Erkrankung frühzeitig zu erkennen. Allerdings weist nicht jeder Betroffene die gleichen Frühsymptome auf. Das klinische Bild kann von vielen Faktoren beeinflusst werden, wie z. B. der körperlichen Verfassung, vom Bildungsstand und den Lebensverhältnissen.

Nach der 10. Revision der Internationalen Klassifikation psychischer Störungen (ICD-10) bezeichnet der Begriff ‚Demenz' ein Syndrom, das mit einer fortschreitenden Beeinträchtigung des Gedächtnisses und weiteren Defiziten verschiedener höherer kortikaler Funktionen einhergeht, einschließlich Denken, Orientierung, Auffassung, Rechnen, Lernfähigkeit, Sprache und Urteilsvermögen. Die kognitiven Beeinträchtigungen werden oft von einer Verschlechterung der emotionalen Kontrolle, des Sozialverhaltens oder der Motivation begleitet. Die Symptomatik muss mindestens sechs Monaten bestanden haben, um eine Demenz mit Sicherheit diagnostizieren zu können. Chronische Erkrankungen, wie die Alzheimer-Krankheit, zerebrovaskuläre Schäden sowie andere Zustandsbilder, die primär oder sekundär das Gehirn betreffen, können zu einem dementiellen Syndrom führen (Dilling et al., 2015).

Aus dieser Definition lässt sich ableiten, dass zur Früherkennung eines dementiellen Syndroms die Eigen- bzw. Fremdanamnese, der psychopathologische Befund sowie eine Prüfung der kognitiven Funktionen unerlässlich bzw. die zentralen diagnostischen Bausteine sind. Weitere Untersuchungen wie Bildgebung,

Labor und ggf. Biomarker spielen erst bei der genaueren Klassifizierung der Demenzursache eine wesentliche Rolle. Es gibt eine Reihe von standardisierten und normierten neuropsychologischen Testverfahren, die für die Prüfung der kognitiven Funktionen herangezogen werden können. Mithilfe von Screening-Tests, die in der hausärztlichen Praxis eingesetzt werden können, lassen sich über das altersentsprechende Maß hinausgehende kognitive Defizite erfassen. Bei Patienten, die eine überdurchschnittliche prämorbide kognitive Leistung hatten oder noch minimale Einbußen aufweisen, kann eine beginnende Demenz dennoch übersehen werden. Darüber hinaus sind Screening-Tests nicht in der Lage, alle kognitiven Domänen zu erfassen. In solchen Fällen können ausführliche neuropsychologische Testverfahren eingesetzt werden (Förstl, 2011).

Im Rahmen der Demenzdiagnostik sind in der Regel Verlaufsmessungen bzw. Wiederholungen der neuropsychologischen Testungen erforderlich. Grund dafür können grenzwertige Ergebnisse, depressive Pseudodemenz oder kognitive Einbuße im Rahmen von anderen Erkrankungen sein.

Darüber hinaus ist oft nach der Behandlung der zugrundeliegenden Erkrankung eine Wiederholung der neuropsychologischen Testung notwendig, um sicher zu stellen, dass die kognitiven Einbußen sich ebenfalls zurückgebildet haben, oder ob weiterer Abklärungsbedarf besteht.

Eine zeitnahe Verlaufskontrolle erweist sich auch für Patienten mit einer diagnostizierten leichten kognitiven Störung und grenzwertigen Testergebnissen neben der weiteren apparativen Diagnostik als sinnvoll. Ebenso bedeutsam sind die wiederholten Testungen für die Verlaufsdokumentation des dementiellen Syndroms und regelmäßige Einschätzung des Schweregrades sowie der Progredienz der Erkrankung.

Der Mensch ist allerdings in der Lage über längere Zeit stabile Leistungen in den neuropsychologischen Testungen zu zeigen (Rubin et al., 1998). Es gibt einen Übungseffekt und mit der wiederholten Durchführung einer bestimmten Aufgabe können die Anweisungen besser verstanden und umgesetzt werden (Goldberg et al., 2015). Dieser Lerneffekt kann insbesondere auf nonverbale Aufgaben zur Prüfung der exekutiven Funktionen, des Gedächtnisses und der Lernfähigkeit einen Einfluss haben. Die Prüfung anderer kognitiver Bereiche, wie visuelle

Wahrnehmung bzw. Visuokonstruktion oder Aufmerksamkeit wird anscheinend weniger vom Lerneffekt beeinflusst (Pereira et al., 2015)

Die mehrfache Wiederholung von Tests in kurzen Zeitabständen bewirkt einen verstärkten Lerneffekt, wobei die Probanden die Möglichkeit erhalten, ihre Lösungsstrategien zu verbessern. Dies kann eine Verfälschung der Testergebnisse zur Folge haben, wenn die tatsächliche Zunahme der kognitiven Defizite durch den Lerneffekt kompensiert und in der Auswertung als Verbesserung interpretiert wird. Aus verschiedenen Studien geht hervor, dass in Abhängigkeit von der Itemschwierigkeit, den Vorerkrankungen, der Intelligenz und dem Alter der Probanden sowie dem Zeitintervall zwischen den Testungen der Lerneffekt unterschiedlich ausfallen kann (Falleti et al., 2006). Die kurzfristige Wiederholung eines Tests stellt also ein wesentliches Problem in der Objektivierung bzw. Messung der kognitiven Defizite dar.

Um den testspezifischen Lerneffekt bei Verlaufsmessungen gering zu halten, wird empfohlen, eine modifizierte Parallelversion der jeweils eingesetzten neuropsychologischen Testung zur Verfügung zu stellen (Collie et al., 2004). Dafür müssen mehrere Varianten eines Tests vorliegen, mit denen die kognitiven Leistungen trotz unterschiedlicher Aufgabenstellungen mit annähernd gleicher Reliabilität und Validität erfasst werden. Ziel dieser Arbeit ist die Entwicklung und Validierung einer Parallelversion eines solchen Testverfahrens, und zwar des Tests zur Früherkennung von Demenzen mit Depressionsabgrenzung (TFDD).

1.2 Validierung neuropsychologischer Verfahren

Aufgabe einer neuropsychologischen Testung ist es, die Leistung einer Person in unterschiedlichen kognitiven Domänen zu messen und mögliche Defizite zu erfassen. Dies sollte mit höchster Genauigkeit erfolgen, weshalb die neuropsychologische Diagnostik berufsethischen Geboten der Psychologie unterliegt. Die Messgenauigkeit eines neuropsychologischen diagnostischen Instruments wird an sogenannten Gütekriterien festgemacht. Die drei Hauptgütekriterien sind Objektivität, Reliabilität und Validität (Schmitt et al., 2014).

Mit Blick auf das Kriterium der Objektivität wird verlangt, dass die Durchführungsbedingungen eines Tests nicht vom jeweiligen Untersucher abhängen und die

Ergebnisse einheitlich ausgewertet und interpretiert werden. So wird Objektivität in Durchführungs-, Auswertungs- und Interpretationsobjektivität unterteilt (Moosbrugger, 2008).

Der Begriff ‚Reliabilität' bezeichnet die Messgenauigkeit eines Tests bzw. wie zuverlässig der Test das misst, was er misst, unabhängig davon, ob der Test auch das misst, was er messen soll. Bei guter Reliabilität sollen sich die Testergebnisse unter den gleichen Bedingungen reproduzieren lassen. Dabei existieren mehrere Methoden, um die Reliabilität eines Tests zu messen. Bei der Retest-Methode wird der Test zweimal mit dem gleichen Probanden durchgeführt und anschließend werden zur Schätzung der Reliabilität Korrelationskoeffizienten berechnet. Im Rahmen der internen Konsistenzanalyse kann mit dem Koeffizienten Cronbachs Alpha aus der Korrelation der einzelnen Items miteinander die Reliabilität des gesamten Tests geschätzt werden (Schmitt et al., 2014).

Die Gütekriterien Objektivität und Reliabilität sind als Voraussetzung für eine hohe Validität anzusehen. Bei letzterer steht die Frage im Zentrum, ob abhängig vom Verwendungszweck ein Test tatsächlich das Merkmal misst, was er messen soll. Die Validität stellt damit das bedeutendste Gütekriterium für eine neuropsychologische Testung dar, denn ohne eine hohe Validität ist die Generalisierbarkeit der Ergebnisse nicht gegeben und der Test ist dementsprechend unbrauchbar. Die Validitätsprüfung einer neuropsychologischen Testung ist erheblich komplexer und anspruchsvoller als die Objektivitäts- und Reabilitätsprüfung. In der quantitativen Forschung werden vier Typen der Validität unterschieden: Konstruktvalidität, interne Validität, externe Validität und statistische Validität. Die Konstruktvalidität wird während der Entwicklung und Operationalisierung einer Untersuchung geprüft und kann als Teil der externen Validität gesehen werden. Als externe Validität wird die Generalisierbarkeit der Ergebnisse bezeichnet, ob also die Schlussfolgerungen einer Studie auf andere Probanden, Variablen und Zeiten generalisierbar sind. Die interne Validität prüft kausale Zusammenhänge bzw. Relationen zwischen Ursache und Wirkung und setzt statistische Validität voraus, sodass diese als Unteraspekt der internen Validität aufgefasst werden kann. Mit der internen Validität wird folglich geprüft, ob die Veränderungen der abhängigen Variablen von der unabhängigen Variablen tatsächlich verursacht werden (Döring & Bortz, 2016).

Die vier Arten der Validität, die im Rahmen des quantitativen wissenschaftlichen Forschungsprozesses beschrieben werden, können bei der Entwicklung von wissenschaftlichen neuropsychologischen Testungen auf drei Aspekte reduziert werden: Inhalts-, Konstrukts- und Kriteriumsvalidität. Dementsprechend stehen drei unterschiedliche Methoden zur Validierung zur Verfügung.

Die Inhaltsvalidität wird nicht statistisch, sondern nach theoretischen und inhaltlichen Überlegungen durch Fachexperten geprüft. Unter Inhaltsvalidität ist zu verstehen, inwieweit die einzelnen Testitems das gesamte Zielkonstrukt inhaltlich widerspiegeln. Von der Inhaltsvalidität muss die Augenscheinvalidität abgegrenzt werden, die sich darauf bezieht, ob der Test einem Laien augenscheinlich valide erscheint (Moosbrugger, 2008).

Ein Test weist Konstruktvalidität auf, wenn aus den Testergebnissen gut fundierte Hypothesen und Rückschlüsse auf andere Variablen bzw. Konstrukte abgeleitet werden können. Im Sinne der konvergenten Validität lässt sich die Übereinstimmung mit konstruktnahen Verfahren ermitteln. So wird z. B. bei einem zu validierenden Demenz-Screening-Verfahren die Korrelation mit anderen etablierten Tests berechnet. Im Rahmen der Prüfung der faktoriellen Validität kann eine Faktorenanalyse erfolgen, ob sich bei eindimensionalen Tests alle Items auf einen einzigen Faktor vereinigen lassen.

Bei der Kriteriumsvalidität werden die Ergebnisse eines Tests mit einem Kriterium außerhalb der Testsituation korreliert (z. B. Demenzdiagnose). In Abhängigkeit vom Zeitpunkt der Erfassung des Außenkriteriums wird von retrospektiver, konkurrenter oder prognostischer Validität gesprochen (Döring & Bortz, 2016).

Normalerweise spielen Sensitivität und Spezifität für neuropsychologische Testungen keine wesentliche Rolle. Diese sonst relevanten Gütekriterien werden bei diagnostischen Verfahren zur Trennung eines dichotomen Kriteriums verwendet, was bei neuropsychologischen Testungen eher nicht der Fall ist. In neuropsychologischen Tests werden in der Regel quantitative Aussagen von der Testperson verlangt, wobei keine dichotome Lösung geboten wird (z. B. Ja-/Nein-Antwort). Die Ausnahme bilden Verfahren, die für die diagnostische Zuordnung durch die Erfassung von klinischen Symptomen eingesetzt werden. So ist es bei einem Demenz-Screening-Test essenziell, prüfen zu können, wie viele erkrankte

Menschen prozentual durch den Test als krank erkannt werden (Sensitivität) bzw. wie viele Gesunde entsprechend richtig erkannt werden (Spezifität). Je besser der Test zwischen den beiden Gruppen unterscheidet, desto höher kann seine Diskriminationsfähigkeit geschätzt werden, im Sinne einer diskriminativen Validität. Diese kann der Kriteriumsvalidität zugeordnet werden, wenn die Gruppenzugehörigkeit (dement, gesund) als Außenkriterium festgelegt wird (Rost, 2004).

Die Nebengütekriterien eines neuropsychologischen Verfahrens umfassen unter anderem Normierung, Testökonomie, Nützlichkeit, Akzeptanz und Testfairness (Schmitt et al., 2014). Diese Kriterien sind indes nicht nur für neuropsychologische Testungen bedeutsam, sondern für alle diagnostischen Verfahren.

Für die Parallelversion einer neuropsychologischen Testung müssen für die gleichen kognitiven Domänen Items entwickelt werden, die vergleichbare zentrale Tendenzen (Mittelwert, Median), eine ähnliche Trennschärfe und Schwierigkeit sowie eine hohe Korrelation zu den Originalitems aufweisen. Schwierigkeit und Trennschärfe werden im Rahmen der Itemanalyse ermittelt (Bühner, 2021).

Mit dem Begriff ‚Schwierigkeitsindex' wird die Schwierigkeit der verschiedenen Items bezeichnet; er entspricht der prozentualen Häufigkeit der richtig gelösten oder bejahten Aufgaben in einer Personenstichprobe. Die Itemschwierigkeit lässt sich bei dichotomen Antwortalternativen berechnen, indem die Anzahl der Probanden, die die Aufgabe richtig gelöst haben, durch die Anzahl der Probanden insgesamt geteilt wird. Um den Schwierigkeitsindex bei mehrstufigen Items zu berechnen, muss die tatsächlich erreichte Punktsumme bei einem Item durch die maximal erreichbare Punktzahl dividiert werden. So entsteht ein Wert zwischen 0 und 1. Bei einem Wert von 1 erreichen alle Probanden die maximale Punktzahl, was dem niedrigsten Schwierigkeitsgrad entspricht, während bei einem Wert von 0 keiner der Probanden einen Punkt erhält, wodurch dieses Item als äußerst schwer zu bewerten ist (Bortz & Döring, 2006).

Mit dem Begriff ‚Trennschärfe' bzw. ‚Trennschärfekoeffizient' wird die Korrelation eines einzelnen Items mit dem Gesamttest beschrieben. Dabei wird berechnet, wie gut das Item das Zielkonstrukt des Tests misst. Soll eine Person bei einem trennscharfen Item eine hohe Punktzahl erreichen, wird im Gesamtergebnis des Tests ebenfalls eine hohe Punktzahl erreicht. Umgekehrt wird bei einem

niedrigen Gesamtergebnis auch im einzelnen trennscharfen Item eine niedrige Punktzahl erreicht. So kann das Item hohe und niedrige Leistungen gut voneinander trennen (Döring & Bortz, 2016).

1.3 Demenzerkrankungen

Wesentliches Merkmal der Demenzerkrankungen ist eine Verschlechterung der kognitiven Fähigkeiten und insbesondere der Gedächtnisleistungen. Zu Beginn der Erkrankung liegen häufig Defizite des Kurzzeitgedächtnisses bzw. der Merkfähigkeit vor, wobei im weiteren Verlauf auch Defizite des Langzeitgedächtnisses entstehen (Fleischman & Gabrieli, 1999).

Eine Demenz ist allerdings mehr als eine reine Gedächtnisstörung. In späteren Stadien der Erkrankung kann es zu Beeinträchtigungen aller kognitiven Bereiche kommen. Aufmerksamkeit, Orientierung, Auffassung, Sprache, Denkvermögen können betroffen sein. Letztlich sind alle Demenzerkrankungen durch den Abbau bis hin zum Verlust der kognitiven Fähigkeiten und Alltagskompetenzen charakterisiert (AWMF, 2016).

Die häufigsten Demenzsymptome werden in der folgenden Tabelle dargestellt (Tabelle 1).

Symptom	Beispiel
Verstärkte Vergesslichkeit	Schlüssel oder andere Gegenstände werden häufiger als früher verlegt. Fragen wiederholen sich, die Haustür wird nicht verschlossen, der Herd wird angelassen.
Orientierungsstörungen	Tag, Monat und Jahr können nicht mehr benannt werden, Monate und Jahreszeiten können nicht mehr zugeordnet werden, die Lage früher bekannter Orte wird nicht mehr gewusst, die Zuordnung von Namen zu Personen geht verloren.
Sprachstörungen	Schwierigkeiten, die richtigen Worte zu finden, zu verstehen, was gemeint ist, und sich selbst verständlich zu machen. Umschreibungen häufen sich.
Gefühlsstörungen	Eine leichte Euphorie, Depressivität und auch Aggressivität können auftreten. Oft werden diese Gefühle als ‚Böswilligkeit' oder ‚wahre Gefühle' verkannt.
Wahnvorstellungen	Kranke können sich bestohlen fühlen, wenn sie vergessen, wohin sie ihre Gegenstände gelegt haben.
Erhaltene Fassade	Die vorhandenen Fähigkeiten werden genutzt, um die Beschwerden zu überspielen: die äußere Fassade ist sehr lange intakt.
Weitere Symptome: Motorische Unruhe, Umherlaufen, Halluzinationen, Unsicherheit, Interesselosigkeit, fehlende Organisation von Körperpflege und Kleidung, Blasen- und Darmentleerungsstörungen, Persönlichkeitsveränderungen.	

Tabelle 1: Symptome der Demenz *(Ihl & Müller, 2020, S. 247)*

Die Demenz ist nicht eine einzelne Störung. Es gibt eine Vielzahl von verschiedenen Ursachen, die zu langsam fortschreitenden kognitiven Leistungseinschränkungen und Verhaltensänderungen führen. Abhängig von der Lokalisation der Demenzursache kann zwischen primären Demenzen (Ursache direkt im Gehirn) und sekundären Demenzen (Folge anderer körperlicher Erkrankungen) unterschieden werden (Möller, 2015).

Die vaskuläre Demenz ist nach der Alzheimer-Krankheit die zweithäufigste Ursache einer Demenz (O'Brien & Thomas, 2015). Hierunter wird eine Demenz bedingt durch vaskuläre Schäden des Gehirns verstanden. Ischämische, hypoxische oder hämorrhagische Läsionen des Gehirns können zu kognitiven Einschränkungen führen, wobei das Fortschreiten sowohl von vaskulären als auch von neurodegenerativen Läsionen für den kognitiven Rückgang verantwortlich ist (Bordet et al., 2017). Patienten, die an einem dementiellen Syndrom leiden und bildmorphologisch zerebrovaskuläre Läsionen zeigen, werden häufig mit einer reinen vaskulären Demenz diagnostiziert. Oft liegt allerdings zusätzlich eine Alzheimer-Pathologie vor (Kosunen et al., 1996). Eine reine vaskuläre Demenz findet sich bei bis zu 11 % aller Demenzerkrankten (Jellinger et al., 2007).

Nach der Alzheimer-Krankheit stellt Morbus Parkinson die zweithäufigste neurodegenerative Erkrankung dar (Lee et al., 2016). Es wird geschätzt, dass 24–31 % aller Parkinsonpatienten an einer Demenz leiden und 3–4 % aller Demenzerkrankungen durch Morbus Parkinson entstehen (Aarsland et al., 2005).

Die Definitionen der primären neurodegenerativen Demenzen ändern sich im Laufe der Zeit kontinuierlich, sodass Studien über die Jahre hinweg kaum miteinander verglichen werden können. Dies stellt ein entscheidendes Hindernis für den wissenschaftlichen Fortschritt dar. Um validere Aussagen für die Praxis treffen zu können, ist die Bildung homogener Gruppen durch neue Klassifizierungskriterien notwendig (Ihl & Müller, 2020).

Eine differenzierte Auflistung der Demenzursachen mit deren bekannten Diagnosekriterien findet sich in Tabelle 2.

	Alzheimer-Demenz	Vaskuläre Demenz	Lewy-Körperchen-Demenz	Frontotemporale Degeneration	Creutzfeldt-Jakob-Krankheit
Spezielle Symptome neben dem dementiellen Syndrom	Nur 50 % zeigen zu Beginn eine Gedächtnisstörung, erhaltene Fassade	Frühe Gangstörung, Blasenstörung ohne urologische Ursache, Stürze, fokale neurologische Zeichen	Fluktuierende Vigilanz, paranoide halluzinatorische Symptome, Parkinson-ähnliche Rigidität, Übersensibilität gegenüber Neuroleptika, REM-Schlafstörungen	Euphorie, emotionale Verflachung, Vergröberung des Sozialverhaltens, visuelle und räumliche Funktionen anfangs gut erhalten	Visuelle und zerebelläre Störungen, pyramidale und extrapyramidale Symptome, Myoklonus, akinetischer Mutismus
Verlauf	Langsame Progression	Schrittweise Progression, partielle Kompensation nach einem Schritt möglich	Wie AD	Wie AD nur im Vergleich schneller	Schnelle Progression, häufig weniger als 1 Jahr Dauer
EEG	Vom Schweregrad abhängige Verlangsamung, verringerte Alpha-Aktivität assoziiert mit schnellerer Progression, keine Veränderung auch möglich	Öfter fokale Veränderungen	Prominente verlangsamte posteriore Aktivität, EEG mit periodischem Fluktuieren im Prä-Alpha/Theta-Bereich; polysomnographische Bestätigung von REM-Schlaf ohne Atonie-Veränderungen im EMG	Keine charakteristischen Veränderungen	Periodische scharfe Wellen (triphasische Wellen oft präsent), nicht bei der neuen Variante
Biomarker	Erhöhtes Tau und Phospho-Tau; erniedrigtes Ab im Liquor	Keine	Auffällig geringe Aufnahme von ^{123}Iod-MIBG in der myokardialen Szintigraphie	Keine	Erhöhtes Protein 14-3-3 im Liquor
Strukturelle Bildgebung CT/MRT	Atrophie (*medial* temporal am Anfang, später temporo-parietal, frontal und am Ende generalisiert; hippokampale vom Schweregrad abhängige Atrophie)	Multiple Infarkte, einzelne Infarkte an strategisch wichtiger Stelle, extensive Läsionen der weißen Substanz	Medio-temporale Strukturen im CT/MRI-Scan meist ausgespart	Frontale und/oder temporale Atrophie, oft asymmetrisch	Unspezifisch
Funktioneller Glucose-Hypometabolismus im PET	Anfangs temporo-parietal und posteriorer Cingulus, später frontal, zuletzt generalisiert	In ischämischen Gebieten	Generalisiert niedrige Aufnahme im SPECT/PET Perfusions/Metabolismus-Scan mit reduzierter okzipitaler Aktivität, anfangs mit dem Zeichen der cingulären Insel im FDG-PET	Frontaler und temporaler Kortex, oft asymmetrisch	Variabel
Amyloid-PET positiv in %	Bis 100	7–64	29–84	3–43	Unzureichende Datenlage
Neuropathologie	Plaques, fibrilläre *Tangles*, kongophile Angiopathie	Ischämische Läsionen	Lewy-Körperchen	Astrozytose, Atrophie der Lamina I-III im frontalen Kortex, Neuropil-Mikrovaskularisierung, Veränderungen in Tau und TDP43	Spongiforme Enzephalopathie

Tabelle 2: Vergleich diagnostischer Kriterien der verschiedenen Demenzformen *(Ihl & Müller, 2020, S. 248)*

Ab: beta-Amyloid, EEG: Elektroenzephalogramm, FDG: Fluorodesoxyglucose, MIBG: Metaiodobenzylguanidin, MRT: Magnetresonanztomographie, PET: Positronenemissionstomographie, SPECT: Einzelphotonen-Emissionscomputertomographie, TDP-43: Transactive response DNA binding protein 43 kDa

1.4 Leichte kognitive Störung

Unter dem Begriff ‚Leichte kognitive Störung' oder *‚Mild Cognitive Impairment'* (MCI) werden kognitive Defizite subsumiert, die über das altersentsprechende Maß hinausgehen, ohne die Kriterien einer Demenz zu erfüllen. Es handelt sich hierbei allerdings nicht um einen einheitlichen Begriff, da eine eindeutige Unterscheidung zwischen kognitiven Veränderungen im Rahmen des Alterungsprozesses oder im Rahmen pathologischer Prozesse nicht immer möglich ist. Im Laufe der Zeit wurden viele unterschiedliche Konzepte, Begriffe und Kriterien für die operationalisierte Diagnose der leichten kognitiven Beeinträchtigungen verwendet.

Bereits im Jahr 1962 beschrieb Kral mit dem Begriff *‚Benign Senescent Forgetfulness'* kognitive Defizite und insbesondere Gedächtnisstörungen im Alter, die von einer dementiellen Erkrankung abzugrenzen waren (Kral, 1962). Aktuell wird am häufigsten der von Petersen eingeführte Begriff ‚*Mild Cognitive Impairment'* (MCI) verwendet. Damit werden Gedächtnisdefizite bezeichnet, die zwar über die Altersnorm hinausgehen, aber dennoch keine Beeinträchtigungen der alltäglichen Funktionen mit sich bringen (Petersen et al., 2001). Petersen und seine Kollegen von den Mayo-Kliniken in den USA empfahlen folgende klinische Kriterien für die Diagnose eines *Mild Cognitive Impairment* (Petersen et al., 2009):

- Angegebene Gedächtnisdefizite, vorzugsweise bestätigt durch einen Informanten
- Objektivierbare Gedächtnisdefizite nach entsprechenden Referenzwerten
- Normale Leistung in restlichen kognitiven Bereichen (außer Gedächtnis)
- Keine Defizite im Bereich der alltäglichen Funktionen
- Keine Demenz diagnostiziert

Nach den diagnostischen Leitlinien von ICD-10 (Internationale statistische Klassifikation der Krankheiten und verwandter Gesundheitsprobleme) sind die Hauptmerkmale der Diagnose ‚Leichte kognitive Störung' Klagen über Gedächtnisstörungen, Vergesslichkeit sowie Lern- oder Konzentrationsschwierigkeiten, also eine zunehmende Beeinträchtigung kognitiver Funktionen. Objektive Testuntersuchungen zeigen gewöhnlich normwidrige Werte und keines dieser Symptome

ist so schwerwiegend, dass die Diagnose ‚Demenz', ‚organisches amnestisches Syndrom' oder ‚Delir' gestellt werden kann (Dilling et al., 2015). Die Diagnose kann zusätzlich in Verbindung mit einer systemischen Erkrankung klassifiziert werden.

Die Diagnose ‚Leichte kognitive Störung' als eigenständige, einheitliche Entität ist allerdings umstritten. Sie beinhaltet keine ätiologische Zuordnung und kann nicht prädiktiv für die Betroffenen verwendet werden. Die jährliche Konversionsrate zu einer Alzheimerdemenz schwankt zwischen 5 und 15 % (Petersen et al., 2009), weshalb unklar ist, ob es sich um ein Prodromalstadium der Alzheimer-Krankheit handelt. In diesem Fall würde der Begriff fälschlicherweise suggerieren, dass eine beginnende Alzheimer-Demenz eine andere Krankheitsentität als die fortgeschrittenen Stadien der Erkrankung darstellt. Dementsprechend ist es sinnvoll, eine weitere Differenzierung der MCI-Patienten vorzunehmen, um diejenigen mit einem erhöhten Demenzrisiko zu erkennen.

Nach den Kriterien und Empfehlungen des *National Institute on Aging and Alzheimer's Association* (NIA-AA) von 2011 können bei Patienten, die die Kriterien für eine leichte kognitive Störung erfüllen, Biomarker die Diagnosesicherheit einer Alzheimer-Krankheit erhöhen. Dabei wird der Begriff ‚Leichte kognitive Störung aufgrund einer Alzheimer-Krankheit' empfohlen und abhängig von den vorliegenden Biomarkern die Wahrscheinlichkeit einer Alzheimer-Krankheit geschätzt (Tabelle 3) (Albert et al., 2011).

	Wahrscheinlichkeit einer Alzheimer-Krankheit	**Hirn-Amyloid-Marker (PET oder Liquor)**	**Neurodegenerationsmarker (Tau, FDG-PET, MRT)**
Klinische Kriterien für eine leichte kognitive Störung erfüllt	Nicht informativ	Nicht verfügbar, widersprüchlich oder nicht bestimmt	Nicht verfügbar, widersprüchlich oder nicht bestimmt
	Geringe Wahrscheinlichkeit	Negativ	Negativ
	Mittlere Wahrscheinlichkeit	Nicht untersucht	Positiv
	Mittlere Wahrscheinlichkeit	Positiv	Nicht untersucht
	Hohe Wahrscheinlichkeit	Positiv	Positiv

Tabelle 3: Wahrscheinlichkeiten für das Vorliegen einer Alzheimer-Krankheit bei MCI *(In Anlehnung an Albert et al., 2011, S. 17)*
PET:Positronenemissionstomographie, Tau:Tau-Protein in Liquor, FDG:Fluordesoxyglucose, MRT: Magnetresonanztomographie des Schädels

Bei gleichzeitig positiven Amyloid- und Tau-Markern erhöht sich das Risiko einer Alzheimer Demenz auf 59 % für die nächsten drei Jahre (Vos et al. 2015). Ist nur ein Biomarker positiv, liegt die Konversionsrate zu einer Alzheimerdemenz für die kommenden drei Jahre zwischen 22 und 49 %. So können durch die biomarkerbasierte Feststellung der Alzheimerpathologie genauere und individuelle Prognosen zum Krankheitsverlauf gestellt werden, woraufhin wiederum MCI-Patienten und Angehörige besser informiert, beraten und weiterhin mit Präventions- und Therapiemaßnahmen begleitet werden können (Rostamzadeh & Jessen, 2020).

1.5 Alzheimer-Krankheit

Nach der ICD-10-Definition wird die Alzheimer-Krankheit als eine primär degenerative zerebrale Erkrankung mit unbekannter Ätiologie sowie charakteristischen neuropathologischen und neurochemischen Merkmalen beschrieben. In der Regel beginnt sie schleichend und zeigt eine langsame, aber stetige Progredienz über mehrere Jahre (Dilling et al., 2015). Das Krankheitsbild wurde erstmalig 1906 vom deutschen Psychiater und Pathologen Alois Alzheimer als ein „eigenartiger schwerer Erkrankungsprozeß der Hirnrinde“ beschrieben (Alzheimer, 1906). Die Erkrankung wurde 1913 vom deutschen Psychiater Kraepelin als ‚Alzheimer‘sche Krankheit‘ benannt (Keuck, 2018).

Die Alzheimer-Krankheit ist bei weitem die häufigste unter den Demenzerkrankungen. Ungefähr 70 % aller Demenzerkrankten leiden an einer Alzheimer-Krankheit (Reitz et al., 2011). Allein in Deutschland leben über 1,5 Millionen Menschen mit einer Alzheimer-Demenz (Georges et al., 2020).

1.5.1 Kognitive Einschränkungen

Es gibt eine Reihe kognitiver und nicht kognitiver Symptome, die das Krankheitsbild einer Alzheimer-Demenz kennzeichnen. In der Regel werden zuerst die kognitiven Leistungsdefizite beobachtet. Bei 80 % der Patienten wird als Erstsymptom eine Verschlechterung der Gedächtnisleistungen festgestellt (Bature et al., 2017). Unter dem Begriff ‚Gedächtnis‘ werden die Aufnahme, das Speichern und das Abrufen von verbalen und nonverbalen Informationen verstanden. Das Gedächtnis kann anhand von zeitlichen oder inhaltlichen Kriterien unterteilt werden.

In Bezug auf die zeitliche Gliederung werden das sensorische oder Ultrakurzzeitgedächtnis, das Kurzzeit- oder Arbeitsgedächtnis und das Langzeitgedächtnis voneinander abgegrenzt (Zoelch et al., 2019). Inhaltlich wird zwischen dem deklarativen (expliziten) und dem nicht deklarativen (impliziten) Gedächtnis unterschieden (Smith et al., 2018). Die Inhalte des deklarativen Gedächtnisses können bewusst abgerufen und verbal berichtet werden. Das implizite Gedächtnis dagegen wirkt auf Verhaltensdispositionen und das Erleben des Menschen, ohne dabei ins Bewusstsein zu treten. Das deklarative Gedächtnis kann weiter in das episodische Gedächtnis, das Erlebnisse in einem räumlichen bzw. zeitlichen Kontext enthält, und das semantische Gedächtnis, das Wissen über allgemeine Fakten und Begriffe beinhaltet, unterteilt werden (Goschke & Dreisbach, 2011).

Zu Beginn der Erkrankung zeigen sich zunächst Beeinträchtigungen des episodischen Gedächtnisses (Bature et al., 2017). Die Patienten sind nicht mehr fähig, sich an kürzlich erlebte Ereignisse zu erinnern; es wird immer schwieriger, neue Informationen zu lernen und zu behalten. Solche Defizite sind durch neuropsychologische Testungen wie z. B. das Lernen einer Wortliste (verbales Gedächtnis) bereits in einem sehr frühen Stadium der Erkrankung erkennbar (Moscoso et al., 2019). Bei den Betroffenen selbst besteht dagegen häufig eine Anosognosie im Sinne von fehlender Krankheitseinsicht bzw. eines fehlenden Bewusstseins, dass solche Symptome aufgetreten sind und eine Krankheit vorliegt (Hanseeuw et al., 2020). Auch das semantische Gedächtnis kann relativ früh betroffen sein. Diese Defizite zeigen sich anfänglich in Form von Wortfindungsstörungen und können mit Aufgaben wie z. B. Wortgenerieren überprüft werden (Vogel et al., 2005).

Die Alzheimer-Demenz wird in den meisten Fällen von Aufmerksamkeitsstörungen begleitet (Hennawy et al., 2019). Daraufhin sind auch weitere kognitive Funktionen beeinträchtigt, die von der Aufmerksamkeit abhängig sind, wie z. B. Rechnen, Schreiben oder Zeichnen. Im Rahmen der Aufmerksamkeitsprüfung kann zunächst die Orientierung getestet werden; darüber hinaus sind Aufgaben zur Prüfung der Informationsverarbeitungsgeschwindigkeit geeignet (Gegenstände benennen, Zahlen lesen, Zahlen ordnen usw.).

Weiterhin sind Störungen der Visuokonstruktion und des räumlichen Denkens häufig in den beginnenden Demenzstadien zu beobachten und gelten als

prognostische Faktoren für die rasche Progredienz der Erkrankung (Ahmed et al., 2016; Müller et al., 2019). Dies führt zu Einschränkungen der alltäglichen Funktionen (Benutzung von Gegenständen, örtliche Orientierung, Teilnahme am Straßenverkehr usw.). Zur Erfassung von Störungen der Visuokonstruktion hat sich der Uhren-Zeichen-Test (*Clock Drawing Test)* nach Shulman bewährt (Shulman, 2000). Hierbei handelt es sich um ein einfaches Screening-Verfahren, womit visuelles Gedächtnis, exekutive Funktionen und visuokonstruktive Ausführung geprüft werden.

Im weiteren Verlauf der Erkrankung können apraktische Störungen auftreten. Die Patienten sind dadurch unfähig, zielgerichtete und geordnete Handlungen korrekt auszuführen, obwohl die motorischen Funktionen unbeeinträchtigt sind (Johnen et al., 2016). Zur Überprüfung der Fähigkeit zu komplexeren, nicht automatisierten Handlungsabläufen können Aufgaben zum Befolgen von Anweisungen verwendet werden (Grass-Kapanke et al., 2005).

In fortgeschrittenen Stadien der Erkrankung nehmen die Gedächtnisdefizite zu und letztlich sind alle kognitiven Funktionen beeinträchtigt.

1.5.2 Affekt und Verhalten

Neben den kognitiven Leistungseinbußen entwickeln die Alzheimer-Patienten im Verlauf der Erkrankung Veränderungen des Affektes bzw. der Stimmungslage sowie Störungen des formalen Denkens und inhaltliche Denkstörungen. Die Mehrheit der betroffenen Menschen zeigt in späteren Stadien der Erkrankung Symptome wie Depressivität, Apathie, Unruhe und Aggressivität (Lanctôt et al., 2017). Im englischen Sprachraum wird das Akronym BPSD (*Behavioral and Psychological Symptoms of Dementia*) für diese Symptome verwendet (Finkel et al., 1997).

Häufig gibt es organische Komorbiditäten, durch die die Verhaltensstörungen bei dementen Patienten verursacht werden. So kann ein Schmerzsyndrom wie bei nicht diagnostizierten Frakturen nach Stürzen Aggression und Enthemmung auslösen. Hirninfarkte führen oft zu organischen affektiven Störungen, mit depressiven Symptomen und Manie (Villa et al., 2018). Darüber hinaus können Polypharmazie, Infekte, metabolische und Elektrolytstörungen in organischen

wahnhaften Störungen und Halluzinosen resultieren. Aus diesem Grund ist es besonders relevant, bei neu aufgetretenen Verhaltensstörungen eine differentialdiagnostische Abklärung durchzuführen.

Auch depressive Syndrome gehen mit kognitiven Einschränkungen und Defiziten im Bereich der alltäglichen Funktionen einher, sodass die differentialdiagnostische Abklärung zwischen Pseudodemenz im Rahmen einer depressiven Störung und Depressivität bei Alzheimer-Demenz von großer Bedeutung ist.

1.5.3 Pathophysiologie

Die Alzheimer-Demenz ist als progrediente neurodegenerative Erkrankung durch den Verlust von Neuronen charakterisiert. Bereits 1906 wurden die beiden histopathologischen Merkmale der Erkrankung von Alzheimer beschrieben (Alzheimer, 1906). Es handelt sich um die sogenannten senilen Plaques und die intrazellulären neurofibrillären Bündel. Die senilen Plaques befinden sich im extrazellulären Raum und bestehen im Wesentlichen aus den Abbauprodukten des Amyloid-Vorläufer-Proteins (APP). Dieses bildet durch proteolytische Spaltung ein 4-kDA-Peptid, das beta-Amyloid (Ab). Das beta-Amyloid wurde im Jahr 1984 als Hauptbestandteil der Plaques erkannt (Glenner & Wong, 1984). Im Zentrum der senilen Plaques befindet sich in der grauen Substanz des Gehirns ein verdichteter Amyloid-Kern, der von degenerierten Neuriten umgeben wird (Nakajima et al., 2016).

Als zweites histopathologisches Merkmal bewirkt das intrazellulär lokalisierte Tau-Protein die Ausbildung von Neurofibrillen. Das Tau-Protein lagert sich an Mikrotubuli an und ist Bestandteil des neuronalen Zytoskeletts. Im Gehirn der Alzheimer-Patienten befindet sich das Tau-Protein in einem hyperphosphorylierten Zustand (Nizynski et al., 2017). Dadurch wird die Bindung vom Tau-Protein an Mikrotubuli deutlich beeinträchtigt, was zur Folge hat, dass der axonale Zytoskelettstruktur-Transport destabilisiert wird. Allerdings sind die neurofibrillären Bündel nicht spezifisch für die Alzheimer-Demenz. Eine Akkumulation von Tau-Protein wird ebenfalls bei den sogenannten Taupathien beobachtet, wie etwa der progressiven supranukleären Blickparese, der kortikobasalen Degeneration oder der frontotemporalen Demenz (Orr et al., 2017).

Seit langem besteht die Hypothese, dass diese Proteine die Ursache der Alzheimer-Krankheit sind. Durch die Entdeckung der autosominal-dominant erblichen Form der Erkrankung sowie weiterer Mutationen des Gens für das Amyloid-Vorläufer-Protein (APP) wird diese Hypothese unterstützt (Tcw & Goate, 2017).

Seit den 1980er Jahren wird demzufolge überwiegend im Bereich der Amyloid-Hypothese geforscht. In den meisten Studien wird versucht, die Bildung von neuen Plaques zu verhindern und deren Abbau zu induzieren. Es wird versucht, in die bis jetzt angenommenen pathophysiologischen Mechanismen einzugreifen und somit die Erkrankung aufzuhalten. Allerdings sind bis jetzt nahezu alle Versuche gescheitert. In 99 % der durchgeführten Studien zwischen 2002 und 2014 konnte kein Unterschied zwischen Placebo und Medikament festgestellt werden (Cummings et al., 2018). Im Jahr 2003 war Memantin das letzte zugelassene Medikament im Bereich der antidementiven Behandlung bis zur US-Zulassung von Aducanumab. Die letzte negative Entwicklung in der symptomatischen Behandlung der Alzheimer-Demenz gab es im Jahr 2019 für den Serotonin-Rezeptor-Antagonisten Idalopirdin. In drei Phase-3-Studien mit mehr als 2500 Patienten konnte keine signifikante Wirkung bezüglich der kognitiven Leistung nachgewiesen werden (Atri et al., 2018).

Auch in der gegen die Plaques gerichteten Antikörper-Forschung konnten bis jetzt keine eindeutigen Erfolge erzielt werden. Zuletzt sind die Substanzen Solanezumab und Crenezumab gescheitert (Sacks et al., 2017; Loureiro et al., 2020). Nur für den Wirkstoff Aducanumab wurde im Jahr 2021 von der US-Arzneimittelbehörde trotz widersprüchlicher Studienergebnisse und negativen Expertenvotums eine Zulassung erteilt. Die Entscheidung basierte auf einem Rückgang der beta-Amyloid-Plaques, der sich in zwei großen Studien zeigte, obwohl eine klinische Relevanz nicht nachgewiesen wurde und noch eine Phase-4-Studie erfolgen muss (Knopman et al., 2021; Sevigny et al., 2016).

Obwohl Antikörper in der Lage sind, Amyloid-Ablagerungen zu entfernen, bremst dies nicht die Progredienz der Erkrankung. Selbst in den frühen Stadien der Erkrankung sind Amyloid-Behandlungen nicht effektiv, was letztlich die Vermutung nahelegt, dass die Plaques nicht die Ursache, sondern eine Begleiterscheinung der Alzheimer-Pathologie darstellen.

Die chronische Neuroinflammation rückt als mögliche Ursache für die Entstehung der Alzheimer-Pathologie immer mehr in den Mittelpunkt. Bereits in den 1980er Jahren wurden Immunproteine in der Nähe von Amyloid-Plaques beobachtet (Griffin et al., 1989). Die Neuroinflammation scheint eine doppelte Funktion in der Entstehung der Alzheimer-Demenz zu besitzen. Während der akuten Immunreaktion spielt sie eine neuroprotektive Rolle mittels Aktivierung von Hirn-Makrophagen, der sogenannten Mikrogliazellen. Diese sind in der Lage, die Amyloid-Plaques im Gehirn zu bereinigen. Die chronische Inflammation kann allerdings nachteilig wirken. Die chronisch aktivierte Mikrogliazellen setzen inflammatorische und toxische Substanzen wie Sauerstoffradikale, Stickstoffmonoxid und Zytokine frei. Es hat sich gezeigt, dass ein erhöhter Spiegel von Zytokinen wie Interleukin-18 zu einer gesteigerten Produktion des Amyloid-Vorläufer-Proteins (APP) und demzufolge beta-Amyloids führt (Alasmari et al., 2018). Dazu wird die Freisetzung von anderen Zytokinen einschließlich des Interleukin-6 durch Interleukin-1b stimuliert. Das Interleukin-6 wiederum stimuliert die Aktivierung der Kinase CDK5, die für die Hyperphosphorylierung des Tau-Proteins mitverantwortlich ist (Saito et al., 2019). Die chronische Neuroinflammation scheint also eine zentrale Rolle in der Entstehung sowohl der Amyloid-Plaques als auch der neurofibrillären Bündel zu spielen. Demzufolge stehen inzwischen die Entzündungsmechanismen im Gehirn und insbesondere die Mikrogliazellen im Mittelpunkt der aktuellen Alzheimerforschung.

Seit Anfang der 1990er Jahre wird die Hypothese der infektiösen Genese der Alzheimer-Pathologie intensiv erforscht. In zahlreichen Studien konnte gezeigt werden, dass Herpes-Viren im Gehirn von Alzheimerpatienten und insbesondere in den Amyloid-Plaques vorkommen (Jamieson et al., 1991; Carbone et al., 2014; Harris & Harris 2015). Es wurde stets davon ausgegangen, dass Amyloid-Plaques rein pathologisch im Rahmen der Alzheimer-Pathophysiologie wirken und keine weitere Funktionalität besitzen. In den letzten Jahren konnte allerdings durch Studien belegt werden, dass beta-Amyloid die Eigenschaften eines antimikrobiellen bzw. antiviralen Peptides besitzt (Eimer et al., 2018), was die Entstehung der antimikrobiellen Hypothese der Alzheimer-Demenz zur Folge hatte. Diese besagt, dass eine intrazerebrale Infektion die Produktion von beta-Amyloid im Rahmen einer immunologischen Reaktion zur Verteidigung erhöhen und somit

die Amyloid-Kaskade initiieren kann (Long & Holtzman, 2019). In zwei epidemiologischen Studien aus Taiwan konnte ein erhöhtes Risiko für die spätere Demenz-Erkrankung nachgewiesen werden, sowohl nach Herpes-simplex-Virusinfektionen (Tzeng et al., 2018) als auch nach Varizella-Zoster-Virusinfektionen (Tsai et al., 2017).

In den letzten Jahren wurden schlechte Mundhygiene und die damit einhergehende chronische Inflammation mit der Alzheimer-Demenz in Verbindung gebracht (Stewart et al., 2015). Menschen mit Serumantikörpern gegen Mundbakterien scheinen ein höheres Risiko zu haben, an einer Demenz zu erkranken (Sparks Stein et al., 2012). Eine zentrale Rolle spielt dabei das Bakterium Porphyromonas gingivalis. Hierbei handelt es sich um einen Erreger von Parodontitis, der hauptsächlich die subgingivalen Bereiche des Zahnfleisches besiedelt. Im Gehirn vom Alzheimer-Patienten fanden sich DNA von Porphyromonas gingivalis sowie erhöhte Mengen seiner Proteasen, der sogenannten Gingipainen. Die orale Infektion von Mäusen mit Porphyromonas gingivalis führte zu einem erhöhten Spiegel von intrazerebralem beta-Amyloid und durch den Einsatz von Wirkstoffen, die Gingipaine blockieren, konnte sogar der Spiegel von beta-Amyloid gesenkt und der entzündliche Prozess gehemmt werden (Dominy et al., 2019).

1.5.4 Genetik

Verwandte ersten Grades von Menschen mit einer Demenz bei Alzheimer-Krankheit haben ein deutlich höheres Risiko, an einer Alzheimer-Demenz zu erkranken. Für Menschen mit einem erkrankten Verwandten ersten Grades ist das Risiko fast zweifach erhöht, für Menschen mit zwei Verwandten ersten Grades vierfach (Cannon-Albright et al., 2019). Eine ursächliche genetische Mutation konnte allerdings bis jetzt nur bei einem kleinen Anteil der Alzheimer-Patienten bestätigt werden. Hierbei handelt sich in weniger als 1 % aller Fälle um Einzelgenmutationen, die eine autosomal-dominal erbliche Form der Erkrankung auslösen. Drei Gene scheinen von solchen Einzelgenmutationen betroffen zu sein. Die erste Mutation findet im Gen für das Amyloid-Vorläufer-Protein (APP) auf dem Chromosom 21 statt. Weitere Mutationen finden sich auf dem Präsenilin-1-Gen

(Chromosom 14) und Präsenilin-2-Gen (Chromosom 1) (Tcw & Goate, 2017). Dabei handelt sich um eine Familie von Transmembranproteinen, die zusammen mit drei anderen Proteinen (Nicastrin, Stabilisationfaktor APH-1, Präsenilin-*Enhancer* 2) den Protein-Komplex Gamma-Sekretase bilden (Johnson et al., 2017). Die Gamma-Sekretase ist neben der Beta-Sekretase für die enzymatischen Spaltungen im Amyloid-Vorläufer-Protein (APP) und die daraus resultierenden beta-Amyloid-Plaques verantwortlich.

Für die meisten Fälle der spätmanifestierten Alzheimer-Demenz liegen allerdings offensichtlich polygenetische Faktoren vor, die bis jetzt nicht ausreichend aufgeklärt sind. Es sind inzwischen hunderte von Genen auf eine Assoziation mit der spätmanifestierten Alzheimer-Demenz untersucht und ca. 150 identifiziert worden (Naj et al., 2017). Dabei ist das Allel des Gens für das Apolipoprotein E (ApoE) von großer Bedeutung. Das ApoE ist in einen komplexen Regulationsmechanismus bezüglich des Serumlipidspiegels eingebunden. Es gibt drei Allele E2, E3 und E4 als Normvarianten auf dem Chromosom 19. Bei Patienten mit Alzheimer-Demenz zeigt sich eine erhebliche Häufung des ApoE-4-Allels, während dies in der Allgemeinbevölkerung deutlich seltener vorkommt. Das ApoE-4-Allel scheint eine Rolle bei der Hemmung des Wachstums von Neuriten, Störung des neuronalen Zytoskeletts, Stimulation der Tau-Phosphorylierung und Neurodegeneration zu spielen (Misra et al., 2018).

Die Bedeutung der Mikroglia in der Pathophysiologie der Alzheimer-Demenz wurde bereits erläutert. Die Aktivierung dieser Hirn-Makrophagen wird durch das TREM2-Gen *(Triggering Receptor Expressed on Myeloid Cells 2)* reguliert. In Mausmodellen zeigte sich, dass sich in den Frühphasen der Alzheimer-Demenz Mikrogliazellen um die neubildenden Amyloid-Plaques sammeln und deren Vergrößerung und Ausbreitung verhindern. Die Mäuse mit defizitärem oder fehlendem TREM2-Gen wiesen eine deutlich geringere Anzahl der plaqueassoziierten Mikrogliazellen auf, was vermuten lässt, dass Mutationen des TREM2-Gens mit einem erhöhten Risiko für Alzheimer-Demenz einhergehen (Jay et al., 2015).

In genomweiten Assoziationsstudien wurden wesentlich mehr Einzelnukleotid-Polymorphismen in verschiedenen Genen identifiziert, die mit der Alzheimer-Krankheit assoziiert sind. Einige der wesentlichen Gene lauten wie folgt: ABCA7, B1N1, CD33, Clusterin-Gen, EPHA1, MS4A und PICALM. Die von diesen Genen

regulierten Proteine sind an einer Vielzahl von für die Alzheimer-Pathologie relevanten zellulären Prozessen beteiligt, wie etwa dem oxidativen Gleichgewicht, dem Proteinstoffwechsel, dem Cholesterinstoffwechsel und der synaptischen Funktion (Misra et al., 2018).

1.5.5 Diagnosekriterien

Das weltweit dominierende Klassifikationssystem für psychiatrische Erkrankungen ist der Diagnostische und Statistische Leitfaden Psychischer Störungen (*Diagnostic and Statistical Manual of Mental Disorders,* DSM). Die fünfte Auflage wurde im Mai 2013 von der Amerikanischen Psychiatrischen Gesellschaft (APA) herausgegeben und löste damit das DSM-IV von 1994 ab.

Um der zunehmenden Stigmatisierung der Demenzkranken in unserer Bevölkerung vorzubeugen, wird der Begriff ‚Demenz' im DSM-V nicht mehr direkt verwendet und stattdessen unter der Kategorie ‚Neurokognitive Störungen' subsummiert. Die neurokognitiven Störungen werden nach Schweregrad und Ätiologie unterteilt, wobei zur Diagnosestellung einer neurokognitiven Störung in einem oder mehreren der folgenden kognitiven Domänen Beeinträchtigungen vorliegen müssen: komplexe Aufmerksamkeit, exekutive Funktionen, Lernen und Gedächtnis, Sprache, perzeptuell-motorische Fähigkeiten, soziale Kognition. Der Leistungsabfall soll zudem nicht im Rahmen eines Delirs oder einer anderen psychiatrischen Störung aufgetreten sein. Je nach Ausprägung der Defizite und ob dadurch die alltäglichen Funktionen beeinträchtigt werden, kann zwischen einer leichten oder schweren neurokognitiven Störung differenziert werden (Oedekoven & Dodel, 2019).

Weiterhin werden die neurokognitiven Störungen hinsichtlich der Ätiologie klassifiziert. Folgende Ursachen werden dabei erfasst:

- Alzheimer-Krankheit
- Vaskuläre Demenz
- HIV-Infektion
- Schädel-Hirn-Trauma
- Frontotemporale Degeneration

- Prionenkrankheiten
- Substanz- bzw. Medikationsinduziert
- Morbus Parkinson
- Huntington-Krankheit
- Multiple Ätiologien
- Unspezifische Ätiologie

Zur Diagnosestellung einer wahrscheinlichen Alzheimer-Krankheit müssen gemäß DSM-V die Kriterien einer neurokognitiven Störung und darüber hinaus weitere für die Alzheimer-Krankheit spezifische Kriterien erfüllt sein. Diese Voraussetzungen sind in den Tabellen 4 und 5 dargestellt.

Für die Diagnose einer leichten neurokognitiven Störung werden gefordert:

• Nachweis einer mäßigen Abnahme kognitiver Leistung relativ zum vorherigen Leistungsniveau in einem oder mehreren kognitiven Bereichen auf der Basis von
  ○ Besorgtheit des Patienten oder eines sachkundigen Informanten oder des Klinikers
  ○ eine mäßige Beeinträchtigung der kognitiven Leistungsfähigkeit, belegt durch eine standardisierte neuropsychologische Testung (im Bereich von 1–2 Standardabweichungen unter Normwerten) bzw. durch eine sonstige qualifizierte klinische Bewertung
• Die Fähigkeit zur selbstständigen Verrichtung alltäglicher Aktivitäten ist erhalten
• Kein Delir
• Die kognitiven Einschränkungen können nicht besser durch eine andere psychische Störung erklärt werden

Für die Diagnose einer schweren neurokognitiven Störung werden gefordert:

• Nachweis einer erheblichen Abnahme kognitiver Leistung relativ zum vorherigen Leistungsniveau in einem oder mehreren kognitiven Bereichen auf der Basis von
  ○ Besorgtheit des Patienten oder eines sachkundigen Informanten oder des Klinikers
  ○ eine erhebliche Beeinträchtigung der kognitiven Leistungsfähigkeit, belegt durch eine standardisierte neuropsychologische Testung (im Bereich von $\geqslant$ 2 Standardabweichungen unter Normwerten) bzw. durch eine sonstige qualifizierte klinische Bewertung
• Die Fähigkeit zur selbstständigen Verrichtung alltäglicher Aktivitäten ist beeinträchtigt.
• Kein Delir
• Die kognitiven Einschränkungen können nicht besser durch eine andere psychische Störung erklärt werden

Tabelle 4: DSM-V-Kriterien einer leichten oder schweren neurokognitiven Störung *(Oedekoven & Dodel, 2019, S. 92)*

A. Die Kriterien für das Vorliegen einer leichten oder schweren neurokognitiven Störung sind erfüllt.

B. Schleichender Beginn und allmähliche Progression der Defizite in einem oder mehreren kognitiven Bereichen (bei einer schweren neurokognitiven Störung mindestens zwei).

C. Die Kriterien für eine wahrscheinliche oder mögliche Alzheimer-Demenz müssen wie folgt erfüllt sein:

Bei einer schweren neurokognitiven Störung:

Eine **wahrscheinliche Alzheimer-Krankheit** wird diagnostiziert, wenn mindestens eines der folgenden Kriterien erfüllt ist, ansonsten soll eine **mögliche Alzheimer-Krankheit** diagnostiziert werden.

1. Nachweis einer für die Alzheimer-Pathologie ursächlichen genetischen Mutation

2. Alle drei der folgenden Voraussetzungen liegen vor:

- Klarer Nachweis von Defiziten der Gedächtnisleistung und Lernfähigkeit sowie in mindestens einem weiteren kognitiven Bereich (basierend auf ausführlicher Anamnese oder neuropsychologischen Testungen)
- Fortlaufend progrediente Verschlechterung der kognitiven Leistungen
- Kein Hinweis auf eine gemischte Ätiologie (z. B. andere neurodegenerative, vaskuläre oder systemische Erkrankung, die zu einer kognitiven Verschlechterung führen könnte)

Bei einer leichten neurokognitiven Störung:

Eine **wahrscheinliche Alzheimer-Krankheit** wird diagnostiziert, wenn eine für die Alzheimer-Pathologie ursächliche genetische Mutation festgestellt wird.

Eine **mögliche Alzheimer-Krankheit** wird diagnostiziert, wenn keine genetische Mutation, aber alle drei der folgenden Voraussetzungen vorliegen:

- Klarer Nachweis von Defiziten der Gedächtnisleistung und Lernfähigkeit sowie in mindestens einem weiteren kognitiven Bereich (basierend auf ausführlicher Anamnese oder neuropsychologischen Testungen)
- Fortlaufend progrediente Verschlechterung der kognitiven Leistungen
- Kein Hinweis auf eine gemischte Ätiologie (z. B. andere neurodegenerative, vaskuläre oder systemische Erkrankung, die zu einer kognitiven Verschlechterung führen könnte)

D. Die neurokognitive Störung kann nicht durch das Vorliegen anderer Erkrankungen besser erklärt werden (zerebrovaskuläre, andere neurodegenerative Erkrankungen, Effekte einer Substanz oder andere neurologische, psychiatrische und systemische Störungen)

Tabelle 5: DSM-V-Kriterien einer leichten oder schweren neurokognitiven Störung im Rahmen der Alzheimer-Krankheit *(In Anlehnung an American Psychiatric Association, 2013, S. 611)*

Bei der ICD-10-Klassifikation handelt es sich um die Kriterien der 10. Revision der Internationalen Statistischen Klassifikation der Krankheiten und verwandter Gesundheitsprobleme. Diese wird von der Weltgesundheitsorganisation erfasst und stellt das weltweit bedeutendste Klassifikationssystem für medizinische Diagnosen dar (Dilling et al., 2015). Nach der ICD-10-Klassifikation wird zwischen der Alzheimer-Demenz mit spätem und frühem Beginn differenziert. Dabei müssen die allgemeinen Kriterien eines dementiellen Syndroms erfüllt sein:

- Progrediente irreversible kognitive Störung
- Keine Bewusstseinsstörung
- Symptomatik seit mindestens sechs Monaten bestehend

Zusätzlich zu diesen Kriterien gibt es weitere Kriterien für die Einordung der Alzheimer-Krankheit. Bei der Alzheimer-Krankheit mit frühem Beginn wird die Diagnose vor dem 65. Lebensjahr gestellt und der Verlauf weist eine vergleichsweise rasche Verschlechterung mit deutlichen und vielfältigen Störungen der höheren kortikalen Funktionen auf. Bei der Alzheimer-Krankheit mit spätem Beginn dagegen wird die Diagnose nach dem 65. Lebensjahr gestellt, die Entwicklung der klinischen Symptome ist langsam progredient und das Hauptmerkmal ist die Gedächtnisstörung.

Für die Forschung existieren darüber hinaus die NIA-AA-Kriterien (*National Institute on Aging and Alzheimer‘s Association)* und die IWG-Kriterien (*International Working Group)*. Die NIA-AA-Kriterien wurden von einem vom *National Institut of Aging* (NIA) und der *Alzheimer‘s Assotiation* (AA) zusammengestellten Komitee vorgestellt, das die Kriterien des *National Institute of Neurological Disorders and Stroke – Alzheimer‘s Disease and Related Disorders Association* (NINCDS-ADRDA) übernommen und ergänzt hat (Tabellen 6, 7, 8 und 9). Die NIA-AA-Kriterien unterscheiden sich von der reinen syndromalen ICD-10- und DSM-V-Klassifikation insofern, als für die Diagnose einer Alzheimer-Demenz neben dem klinischen Syndrom auch Biomarker eingebunden werden (McKhann et al., 2011). Sind die klinischen Kriterien der wahrscheinlichen oder möglichen Alzheimer-Demenz erfüllt, können durch die Anwendung von Biomarkern verschiedene Konstellationen für das Vorliegen einer Alzheimer-Krankheit beschrieben werden (Tabelle 10).

Eine Demenz wird diagnostiziert, wenn kognitive oder verhaltensbezogene (neuropsychiatrische) Symptome vorliegen, die:

- die Funktionsfähigkeit bei der Arbeit oder bei üblichen Aktivitäten beeinträchtigen und
- eine Verschlechterung gegenüber dem früheren Funktions- und Leistungsniveau darstellen und
- nicht durch ein Delir oder eine schwere psychiatrische Störung erklärbar sind

Kognitive Beeinträchtigungen werden durch eine Kombination aus einer Anamneseerhebung und einer objektiven Bewertung der kognitiven Funktionen, entweder durch klinische Untersuchung oder eine neuropsychologische Testung, erkannt und diagnostiziert. Neuropsychologische Testungen sollten durchgeführt werden, wenn Anamnese und Untersuchung der Kognition keine sichere Diagnose liefern können. Es müssen mindestens zwei der folgenden Bereiche beeinträchtigt sein:

a. Fähigkeit, neue Informationen zu erlangen und sich daran zu erinnern

b. Verstehen und Durchführung von komplexen Aufgaben, Urteilsvermögen

c. Räumlich-visuelle Funktionen

d. Sprachfunktionen (Sprechen, Lesen, Schreiben)

e. Veränderungen der Persönlichkeit und des Verhaltens

Die Abgrenzung der Demenz zur leichten kognitiven Störung *(Mild Cognitive Impairment,* MCI) beruht auf der Feststellung, ob eine signifikante Beeinträchtigung der Arbeitsfähigkeit oder der alltäglichen Funktionen vorliegt. Diese werden auf der Grundlage der individuellen Patientenkonstellation und der Beschreibung der Defizite durch den Patienten selbst und einen sachkundigen Informanten beurteilt

Tabelle 6: Allgemeine NIA-AA-Kriterien einer Demenz *(in Anlehnung an McKhann et al., 2011, S. 3–4)*

- Die allgemeinen klinischen Demenzkriterien sind erfüllt (Tabelle 6)
- Die Symptome beginnen allmählich über Monate bis Jahre
- Eindeutige Verschlechterung der kognitiven Leistungen durch Anamnese und Beobachtung
- Die ersten und eindeutigen Symptome zeigen sich in einer der folgenden Kategorien:
    a. Amnestische Variante: Die Defizite sollten Beeinträchtigungen beim Lernen und beim Abrufen kürzlich erlernter Informationen umfassen. Es sollte auch Hinweise auf eine kognitive Dysfunktion in mindestens einem anderen kognitiven Bereich geben

b. Nichtamnestische Varianten:

-Sprachvariante: Die auffälligsten Defizite liegen in der Wortfindung

-Visuell-räumliche Präsentation: Die auffälligsten Defizite liegen in der räumlichen Wahrnehmung, einschließlich Objektagnosie, beeinträchtigter Gesichtserkennung, Simultanagnosie und Alexie.

-Exekutive Dysfunktion: Die häufigsten Defizite sind beeinträchtigtes Denken, Urteilsvermögen und Problemlösung

Ausschlusskriterien für die Diagnose einer wahrscheinlichen Alzheimer-Demenz sind:

- Erhebliche zerebrovaskuläre Erkrankung, definiert durch einen Schlaganfall zum Beginn der kognitiven Beeinträchtigungen oder das Vorhandensein von multiplen bzw. ausgedehnten Infarkten
- Eindeutige Symptome einer Lewy-Körperchen-Demenz
- Eindeutige Symptome der behavioralen Variante einer frontotemporalen Demenz
- Eindeutige Symptome einer primär progressiven Aphasie (nicht flüssige, agrammatische oder semantische Variante)
- Hinweis für andere Erkrankungen oder Medikamente, die die Kognition beeinträchtigen könnten

Tabelle 7: NIA-AA-Kriterien für eine wahrscheinliche Demenz bei Alzheimer-Krankheit *(In Anlehnung an McKhann et al., 2011, S. 4–5)*

- Beim atypischen Verlauf sind die typischen Symptome der Demenz bei Alzheimer-Krankheit erfüllt (Tabelle 7), diese beginnen allerdings plötzlich oder es gibt keine ausreichenden anamnestischen Angaben bezüglich des klinischen Verlaufs oder es konnte keine Progredienz der Symptomatik dokumentiert werden
- Bei der ätiologisch gemischten klinischen Variante sind ebenfalls die typischen Symptome der Demenz bei Alzheimer-Krankheit erfüllt (Tabelle 7), aber es gibt gleichzeitig Hinweise auf:
    a. Erhebliche zerebrovaskuläre Erkrankung, definiert durch einen Schlaganfall zum Beginn der kognitiven Beeinträchtigungen oder das Vorhandensein von multiplen bzw. ausgedehnten Infarkten
    b. Eindeutige Symptome einer Lewy-Körperchen-Demenz
    c. Andere Erkrankungen oder Medikamente, die die Kognition beeinträchtigen könnten

Tabelle 8: NIA-AA-Kriterien für eine mögliche Demenz bei Alzheimer-Krankheit *(In Anlehnung an McKhann et al., 2011, S. 6)*

Entweder sind die klinischen Kriterien für eine Demenz bei Alzheimer Krankheit nicht erfüllt oder:

- Unabhängig davon, ob die klinischen Kriterien für eine wahrscheinliche oder mögliche Alzheimer-Demenz erfüllt sind, gibt es genügend Hinweise für eine alternative Diagnose wie HIV-Demenz, Demenz bei Chorea Huntington oder andere Ursachen, die sich selten mit der Alzheimer-Demenz überschneiden
- Unabhängig davon, ob die klinischen Kriterien für eine wahrscheinliche oder mögliche Alzheimer-Demenz erfüllt sind, sind sowohl die Hirn-Amyloid-Marker als auch die Neurodegenerationsmarker negativ

Tabelle 9: NIA-AA-Kriterien für eine Demenz, wahrscheinlich nicht durch Alzheimer-Krankheit bedingt *(In Anlehnung an McKhann et al., 2011, S. 8)*

Klinische Diagnose-Kategorie	**Wahrscheinlichkeit einer Alzheimer-Krankheit**	**Hirn-Amyloid-Marker (PET oder Liquor)**	**Neurodegenerationsmarker (Tau, FDG-PET, MRT)**
Klinische Kriterien für eine wahrscheinliche Demenz bei Alzheimer-Krankheit erfüllt	Nicht informativ	Nicht verfügbar, widersprüchlich oder nicht bestimmbar	Nicht verfügbar, widersprüchlich oder nicht bestimmbar
	Mittlere Wahrscheinlichkeit	Nicht verfügbar oder widersprüchlich	Positiv
	Mittlere Wahrscheinlichkeit	Positiv	Nicht verfügbar oder widersprüchlich
	Hohe Wahrscheinlichkeit	Positiv	Positiv
Klinische Kriterien für eine mögliche Demenz bei Alzheimer-Krankheit erfüllt (atypische klinische Präsentation)	Nicht informativ	Nicht verfügbar, widersprüchlich oder nicht bestimmbar	Nicht verfügbar, widersprüchlich oder nicht bestimmbar
	Hohe Wahrscheinlichkeit, schließt aber eine zweite Ätiologie nicht aus	Positiv	Positiv
Klinische Kriterien für eine Demenz erfüllt, wahrscheinlich nicht durch Alzheimer-Krankheit bedingt	Geringste Wahrscheinlichkeit	Negativ	Negativ

Tabelle 10: Wahrscheinlichkeiten einer Alzheimer-Krankheit unter Einbeziehung von Biomarkern *(In Anlehnung an McKhann et al., 2011, S. 11)*
PET:Positronenemissionstomographie, Tau:Tau-Protein in Liquor, FDG:Fluordesoxyglucose, MRT: Magnetresonanztomographie des Schädels

Die NINCDS-ADRDA-Kriterien weisen mit 95 % eine hohe Sensitivität auf, allerdings mit 79 % ebenfalls eine relativ geringe Spezifität (Lopez et al., 1999).

1.5.6 Therapie

Die medikamentöse Therapie der Demenz richtet sich auf die Behandlung einerseits der Primärsymptomatik bzw. der kognitiven Einschränkungen und anderseits der begleitenden affektiven und Verhaltensstörungen. Es kommen hauptsächlich drei Medikamentengruppen zum Einsatz: Antidementiva, Antipsychotika und Antidepressiva.

Zu den Antidementiva zählen die Acetylcholinesterase-Inhibitoren Donepezil, Rivastigmin und Galantamin, der NMDA-Rezeptorantagonist Memantin und der Ginkgo-Biloba-Extrakt EGb 761. Die Acetylcholinesterase-Inhibitoren sind zur Behandlung der leichten bis mittelschweren Alzheimer-Demenz zugelassen, das Memantin wird bei Patienten mit mittelschwerer bis schwerer Alzheimer-Demenz eingesetzt und der Ginkgo-Extrakt ist zur symptomatischen Behandlung von hirnorganisch bedingten geistigen Leistungseinbußen im Rahmen eines dementiellen Syndroms zugelassen. Nach der Demenzleitlinie der Weltföderation der Fachgesellschaften für Biologische Psychiatrie (WFSBP) von 2011 ist keine der Substanzen gegenüber den anderen überlegen und alle Substanzen sind in der Lage, einen moderaten Effekt für einen begrenzten Zeitraum zu ermöglichen, allerdings nur für einen Teil der behandelten Menschen (Ihl et al., 2011).

Symptome im Rahmen von affektiven und Vehaltensstörungen lassen sich in der Regel mithilfe von nicht medikamentösen Maßnahmen im Sinne von psychosozialen Interventionen gut beeinflussen. Darunter werden Interventionen wie Psychoedukation (Hepburn et al., 2007), Sozialberatung und kognitives Training zur Stärkung der betreuenden Angehörigen (Schoenmakers et al., 2010), Aromatherapie (Thorgrimsen et al., 2003), basale Stimulation, psychotherapeutische kognitive Verfahren (Livingston et al., 2005) sowie Musiktherapie und Physio- bzw. Aktivierungstherapie (Heyn et al., 2004) verstanden.

Sind die oben genannten Maßnahmen jedoch nicht ausreichend wirksam,

können zusätzlich Medikamente verabreicht werden. Häufig bestehen bei älteren Patienten eine Polypharmazie und eine verstärkte Sensitivität gegen Psychopharmaka. Aus diesem Grund ist es essenziell, mit einer niedrigen Dosis zu beginnen, eine Monotherapie anzustreben und die Indikation für die medikamentöse Behandlung in regelmäßigen Abständen zu überprüfen. Zur Behandlung von depressiven bzw. Angstsymptomen spricht die aktuelle Studienlage für den Einsatz einer medikamentösen antidepressiven Therapie. Insbesondere der Serotonin-Wiederaufnahmehemmer (SSRI) Citalopram scheint eine gute Wirksamkeit zu haben (Aga, 2019). Eine Behandlung mit trizyklischen Antidepressiva ist aufgrund des Nebenwirkungsprofils dagegen nicht zu empfehlen (Hessmann et al., 2019).

Zur Behandlung von Verhaltensstörungen und psychotischen Symptomen können Antipsychotika eingesetzt werden. Im Vergleich zum Placebo zeigen z. B. Risperidon, Aripiprazol und Olanzapin eine Besserung der neuropsychiatrischen Symptome der Demenz. Gleichzeitig sind allerdings dieselben Substanzen mit einem erhöhten Risiko für kardiovaskuläre Ereignisse wie Myokardinfarkt oder Stürze und Frakturen assoziiert (Farlow & Shamliyan, 2017). Insgesamt ist die Einnahme von Antipsychotika mit einem erhöhten Mortalitätsrisiko und einer Verschlechterung der kognitiven Funktionen assoziiert (Gill et al., 2007). Zur Behandlung der schweren chronischen Aggressivität ist nach der S3-Leitlinie für Demenzen Risperidon zugelassen und kann dementsprechend für diese Indikation verabreicht werden (AWMF, 2016).

Darüber hinaus gibt es Hinweise für die Wirksamkeit des Antikonvulsivums Carbamazepin gegen Agitation und Aggressivität bei dementen Patienten (AWMF, 2016).

In den letzten Jahren hat sich die Behandlung von Demenzen insgesamt nur wenig verändert. Alle Substanzen haben einen ähnlichen moderaten Effekt bezüglich der kognitiven Symptome und neuropsychiatrischen Störungen. Dagegen zeigen die verschiedenen Medikamente ein unterschiedliches Nebenwirkungsprofil. Demzufolge steht bei der medikamentösen Einstellung die Schaden-Nutzen-Abwägung an erster Stelle (Ihl & Müller, 2020).

1.6 Neuropsychologische Screening-Verfahren

1.6.1 Mini-Mental-Status-Test (MMST)

Als gebräuchlichster und bekanntester Demenz-Screening-Test gilt der Mini-Mental-Status-Test (MMST). Er war der erste Screening-Test, der entwickelt und veröffentlicht wurde, und beinhaltet standardisierte Fragen zu Orientierung, Aufmerksamkeit, Merkfähigkeit, Rechnen, Sprache, Erinnerungsfähigkeit, Praxis, Erkennen und räumlichem Denken (Folstein et al., 1975). Der MMST ist ökonomisch und leicht anwendbar. Er wird innerhalb von 10 bis 15 Minuten durchgeführt und die Antworten des Probanden werden mit insgesamt bis zu 30 Punkten bewertet. Eine Leistung zwischen 26 und 19 Punkten spricht dabei für ein leichtes dementielles Syndrom, eine Leistung zwischen 10 und 18 für ein mittelschweres. Eine Punktzahl von weniger als 10 Punkten ergibt einen deutlichen Hinweis für ein schweres dementielles Syndrom. Als Nachteile sind indes die geringe Sensitivität und Spezifität des MMST im Vergleich zu anderen neuropsychologischen Verfahren zu nennen, insbesondere bei beginnenden Demenzen (Kohn et al., 2007; Ihl et al., 2005). Fragen wie das unmittelbare Erinnern von drei einfachen Worten zeigen erst bei einer fortgeschrittenen Demenz Auffälligkeiten. Neben den begrenzten Aussagen beim Screening wird er auch aufgrund der hohen Variabilität des Ergebnisses für Verlaufsuntersuchungen als ungeeignet angesehen.

1.6.2 *Montreal Cognitive Assessment* (MoCA)

Ein neueres Testverfahren ist das *Montreal Cognitive Assessment* (MoCA), (Nasreddine et al., 2005). Dieses weist eine gute Spezifität und Sensitivität auf und schnitt verglichen mit dem MMST deutlich besser ab (Ciesielska et al, 2016). Ähnlich wie beim MMST beträgt die Durchführungsdauer 10 bis 15 Minuten und es können maximal 30 Punkte erreicht werden. Ergebnisse über 26 Punkte gelten dabei als normal. Der MoCA beinhaltet insgesamt neun Aufgaben. Geprüft werden räumliches Denken, Aufmerksamkeit, Merkfähigkeit, kurz- und mittelfristiges Behalten, Abstraktion, Sprache und Orientierung. Zur Durchführung des Tests ist ein unbeeinträchtigtes Gehör notwendig, was die Testung für Patienten mit Presbyakusis erheblich erschwert.

1.6.3 Demenz-Detektions-Test (DemTect)

Der Demenz-Detektions-Test (DemTect) wurde im Jahr 2000 zur Früherkennung von beginnenden Demenzen entwickelt (Kalbe et al. 2004). Er weist eine kurze Durchführungsdauer von 5 bis 10 Minuten auf und beinhaltet Aufgaben zur Prüfung von der unmittelbaren Wiedergabe einer Wortliste, verbaler Flüssigkeit, der Zahlenspanne sowie des verzögerten Abrufs der Wortliste. Ein Nachteil besteht indes darin, dass wesentliche kognitive Bereiche wie z. B. exekutive Funktionen, räumliches Denken und visuell-konstruktive Fähigkeiten nicht erfasst werden. Bei der Erinnerungsaufgabe wird nicht geprüft, ob der Untersuchte das zu Erinnernde überhaupt wahrgenommen hat, wodurch die Aussagekraft eingeschränkt ist.

1.6.4 Syndrom-Kurz-Test (SKT)

Der Syndrom-Kurz-Test (SKT) wurde bereits im Jahr 1977 entwickelt und publiziert (Erzigkeit, 1977). Zuletzt wurde das Material des SKT im Jahr 2015 überarbeitet (Stemmler et al., 2015).

Ähnlich wie die meisten Screening-Verfahren beträgt die Durchführungsdauer des SKT ca. 10 bis 15 Minuten. Dabei werden die Gedächtnisleistungen durch drei Items und die Aufmerksamkeit bzw. Informationsverarbeitungsgeschwindigkeit durch sechs Items geprüft. Es gibt fünf Parallelversionen des SKT, die verwendet werden können.

Auch in diesem Test werden allerdings kognitive Funktionen wie räumliches Denken und visuell-konstruktive Fähigkeiten nicht erfasst. Darüber hinaus gelingt es nicht, hinreichend unauffällige von erkrankten Menschen zu trennen. Ab einem mittleren Schweregrad der Demenz wird die Höchstpunktzahl des Tests bereits erreicht, sodass eine weitere Differenzierung im späteren Verlauf nicht mehr möglich ist. Es gibt also einen Deckeneffekt. Im Bereich der leichten bis zu den mittelschweren kognitiven Störungen lässt sich aber aufgrund der exakten Aufgabendefinition eine Verlaufsmessung sehr gut durchführen.

1.6.5 Test zur Früherkennung von Demenzen mit Depressionsabgrenzung (TFDD)

Der Test zur Früherkennung von Demenzen mit Depressionsabgrenzung (TFDD) wurde im Jahr 2000 veröffentlicht und berücksichtigt neben dem Demenz-Screening auch differentialdiagnostische Fragen. Die häufigste Differenzialdiagnose eines dementiellen Syndroms ist die Pseudodemenz im Rahmen einer depressiven Störung. Hier beinhaltet der TFDD zwei Items mit hoher Sensitivität und Spezifität zur Abgrenzung der Demenz von der depressiven Pseudodemenz. Darüber hinaus besteht der Test aus neun Items, die zuvor auf ihre Sensitivität für eine beginnende Demenz geprüft wurden. Die Durchführungsdauer beträgt 5 bis 10 Minuten. In der Validierungsarbeit erreichte der TFDD eine Sensitivität von 98 % und eine Spezifität von 97 %. Dabei zeigte sich eine Inter-Rater-Reliabilität von r_s = .99 und eine Retest-Reliabilität von r_s = .99. Darüber hinaus korrelierten die Punktwerte des Tests signifikant mit der GDS-*Global Deterioration Scale*-Stufe (r_s = -.91), was auf die Möglichkeit der Einschätzung der Krankheitsschwere durch das TFDD-Ergebnis hinweist. Durch die guten Testgütekriterien sowie die schnelle und kosteneffiziente Durchführung ist der TFDD besonders gut für einen Einsatz im niedergelassenen Bereich geeignet (Ihl et al., 2000). Die englische Version des TFDD wurde bereits im Jahr 2005 validiert (Mahoney et al., 2005) und aktuell wird eine türkische Version untersucht. Ein Paralleltest stand bis zur aktuellen Arbeit nicht zur Verfügung.

Mit der unmittelbaren und verzögerten Reproduktion einer Wortliste wird das Kurzzeitgedächtnis geprüft. Dieses ist bereits im Frühstadium eines dementiellen Syndroms beeinträchtigt, weswegen dessen Prüfung eine wesentliche Aufgabe in der Frühdiagnostik der Demenz darstellt (Chaves & Camozzato 2007).

Bereits bei der initialen Beschreibung der Demenzerkrankung von Alois Alzheimer wurde die Störung der zeitlichen Orientierung als Kardinalsyndrom der beginnenden Demenz erkannt (Alzheimer, 1907). Diese wird im TFDD mit Fragen nach dem aktuellen Datum, den vier Jahreszeiten und der aktuellen Jahreszeit überprüft.

Das Kategorisieren bzw. die Zuordnung von Begriffen zu übergeordneten Kategorien ist ebenfalls im Frühstadium einer Demenz beeinträchtigt (Monsch et al.,

1997), sodass die Frage nach zur aktuellen Jahreszeit gehörenden Monaten ergänzt wurde.

Mit dem Befolgen von einfachen Anweisungen wird die Fähigkeit zu komplexeren nicht automatisierten Handlungsabläufen getestet.

Mit dem Uhren-Zeichen-Test (*Clock Drawing Test*) wird die Zeichenfähigkeit bzw. konstruktive Praxis geprüft. Hierbei handelt sich um ein eigenständiges und etabliertes Screening-Verfahren, das seit vielen Jahrzehnten für die Erfassung von kognitiven Einschränkungen verwendet wird (Freedman et al., 1994). Die Zeichenfähigkeit ist bereits bei Patienten mit einer leichten Demenz erheblich beeinträchtigt und sogar häufig bei gesunden alten Menschen defizitär (Ihl et al., 2000). Es gibt viele verschiedene Uhrentestversionen und Auswertungsschemata (Shulman, 2000). Im TFDD wurde die Methode nach Sunderland verwendet (Sunderland et al., 1989), bei der der Proband selbst das Zifferblatt einer Uhr mit allen Zahlen zeichnen und Stunden- und Minutenzeiger auf 11:10 Uhr platzieren muss.

Von vielen Autoren wird diskutiert, welche Uhrzeit und ob überhaupt eine konkrete einheitliche Uhrzeit verwendet werden soll. Die drei häufigsten verwendeten Uhrzeiten sind 11:10 Uhr, 8:20 Uhr und 3:00 Uhr (Freedman et al., 1994). Die Uhrzeit 11:10 hat den Vorteil, dass dabei zum einen bilaterale hemisphärische Prozesse überprüft werden (Zeiger auf unterschiedlichen Seiten des Zifferblattes) und zum anderen die exekutiven Funktionen abverlangt werden, um *Stimulus-Bound Responses* zu vermeiden. Diese beschreiben die Neigung von kognitiv beeinträchtigen Probanden, das Zeichnen der Zeiger auf einen einzelnen Reiz zu basieren. Bei der Angabe ‚10 nach' muss z. B. der Zeiger auf die Ziffer 2 gestellt werden. Weil allerdings die Ziffer 10 neben der 11 auf dem Zifferblatt zu finden ist, neigen viele demente Patienten dazu, den Minutenzeiger auf die Ziffer 10 zu stellen (Eknoyan et al., 2012).

Auch die Sprache bzw. Wortflüssigkeit ist bei einer beginnenden Demenz häufig defizitär (Rascovsky et al., 2007). Wortflüssigkeitsaufgaben werden im neuropsychologischen Bereich für die Überprüfung exekutiver Funktionen und semantischer Leistungen eingesetzt (Henry et al., 2015). Mögliche Defizite können durch zwei alternative Vorgehensweisen festgestellt werden. Bei einem phonologischen Wortflüssigkeitstest müssen die Probanden Wörter abrufen, die mit einem

bestimmten Buchstaben beginnen. Bei einem semantischen Wortflüssigkeitstest müssen so viele Wörter einer Kategorie (z. B. Tiere oder Supermarktprodukte) wie möglich genannt werden. Bei Patienten mit einer Alzheimer-Krankheit ist insbesondere das semantische kategoriale Wortgenerieren gestört (Kessler et al. 1997; Taylor et al., 2005). Demzufolge wurde eine entsprechende Aufgabe für den TFDD gewählt (*Animal Naming Fluency*).

Die maximale Punktzahl, die bei den neun Demenz-Items erreicht werden kann, beträgt 50 Punkte. Der *Cut-Off*-Wert liegt bei 35, sodass ein Score von 35 Punkten oder darunter auf ein dementielles Syndrom hinweist.

Die zwei letzten Items dienen zur Abgrenzung der Demenz von der depressiven Pseudodemenz. Dabei wird eine globale Fremd- und Selbsteinschätzung der Stimmungslage auf einer elfstufigen Likert-Skala (0 = ausgeglichen, 10 = schwer depressiv) erhoben. Der *Cut-Off*-Wert für das Vorliegen einer depressiven Symptomatik liegt bei 8/9 (Summenscore aus Selbst- und Fremdrating). Somit ergibt ein Punktwert von über 8 von 20 möglichen Punkten in der Depressionsskala einen Hinweis auf eine relevante depressive Störung, die eine zusätzliche Diagnostik und fachärztliche Behandlung erforderlich macht.

1.6.6 Ausführliche neuropsychologische Testverfahren

Der Wunsch nach einer umfassenderen und genaueren Erfassung von kognitiven Einschränkungen ließ zwei weitere, längere Testverfahren erstehen. Ergänzend zu den Screening-Verfahren besteht mit diesen Tests die Möglichkeit, verschiedene kognitive Funktionen genauer zu erfassen und die festgestellten Defizite zu quantifizieren. Unter anderem werden kognitive Bereiche wie Gedächtnis, Orientierung, Fähigkeit zu Handlungsabläufen, Sprache und Konstruktive Praxis geprüft.

Die Arbeitsgruppe um Mohs untersuchte dazu retrospektiv die Symptome von Patienten, bei denen eine Demenz neuropathologisch bestätigt worden war, und reduzierte statistisch die entsprechenden Aufgaben auf eine Anzahl kognitiver und nichtkognitiver Items, die schließlich als *Alzheimer Disease Asessment Scale* (ADAS) publiziert wurde (Mohs et al.,1983).

Einen anderen Weg schlug das *Consortium to Establish a Registry of Alzheimer's Disease* ein (Morris et al., 1988). Die Autoren suchten aus allen zu diesem Zeitpunkt existierenden Tests Items heraus, die unterschiedliche Aspekte abdeckten, und fassten sie in einer umfangreichen Skala mit dem Namen CERAD (*Consortium to Establish a Registry for Alzheimers' Disease*) zusammen. Diese beinhaltet den MMST sowie sieben weitere Aufgaben, die episodische Gedächtnisleistungen, Sprachfertigkeiten, Orientierungsvermögen und Handlungspraxis umfassen.

Beide Verfahren sind jedoch bei mittelschweren und schweren Demenzen nicht mehr einsatzfähig, denn ab einem gewissen Schweregrad der Erkrankung können die Aufgaben nicht mehr gelöst werden, im Sinne eines Deckeneffektes. Darüber hinaus ist der CERAD bildungsabhängig.

Die Durchführungsdauer beider Testverfahren beträgt 30 bis 45 Minuten und aufgrund des hohen Zeitaufwandes sind sie für den Einsatz im niedergelassenen Bereich nicht geeignet.

In Tabelle 11 werden die verschiedenen Testverfahren mit ihren Vor- und Nachteilen sowie ihrer Eignung aufgelistet.

Testverfahren	Anwendbarkeit	Vorteile	Nachteile
MMST	Demenz-Screening, Schweregradmessung ab mittelschwerem Stadium	Zeitökonomisch und leicht anwendbar	Geringe Sensitivität und Spezifität, geringe Reliabilität
MoCA	Demenz-Screening	Gute Sensitivität und Spezifität	Presbyakusis erheblicher Störfaktor
DemTect	Demenz-Screening	Schnellstes Screening-Verfahren	Einige kognitive Domänen werden nicht geprüft
SKT	Demenz-Screening, Schweregrad- und Verlaufsmessung	Aufgrund der exakten Aufgabendefinition sehr gut für Verlaufsmessungen geeignet	Weniger sensitiv bei höhergradigen Einschränkungen
TFDD	Demenz-Screening, Abgrenzung von depressiven Syndromen	Sehr gute Sensitivität und Spezifität, zwei Depressionsitems	Trotz eines linearen Zusammenhangs nicht für Schweregradeinstufung entwickelt
CERAD	Ausführliche Diagnostik, Schweregrad- und Verlaufsmessung	Umfassende und genaue Erfassung der kognitiven Leistungen	Deckeneffekt, Bildungsabhängig
ADAS	Ausführliche Diagnostik, Schweregrad- und Verlaufsmessung	Umfassende und genaue Erfassung der kognitiven Leistungen	Boden- bzw. Deckeneffekt

Tabelle 11: Vergleich neuropsychologischer Testverfahren

1.7 Ziele der Arbeit

In den ersten Abschnitten der vorliegenden Arbeit wurde deutlich, dass für die Frühdiagnose der Demenz die klinische Symptomatik und insbesondere die kognitiven Defizite entscheidend sind. Aktuell existieren keine zuverlässigen diagnostischen Instrumente zur Vorhersage für das Auftreten einer Demenz. Dementsprechend sind die ICD-10-, DSM-V- und NIA-AA-Kriterien für die Diagnosestellung einer allgemeinen Demenz rein klinisch, im Sinne von kognitiven Beeinträchtigungen und verhaltensbezogenen Symptomen, die das Funktionieren im alltäglichen Bereich beeinträchtigen. Daher spielen neuropsychologische Testungen eine zentrale Rolle für die Beurteilung und Objektivierung dieser Symptome.

Mit einer Sensitivität von 98 % und einer Spezifität von 97 % hat sich der TFDD gerade bei der Frühdiagnostik einer Demenz bewährt und erfüllt aufgrund der einfachen und schnellen Durchführbarkeit bzw. Auswertung sowie des geringen Materialaufwandes die Validierungskriterien für die Anwendung in der ärztlichen Praxis am besten.

Um grenzwertige Ergebnisse zu kontrollieren, nachgewiesene kognitive Defizite zu bestätigen, oder auch den Verlauf einer diagnostizierten Demenz zu beurteilen sind in der Regel Wiederholungen der neuropsychologischen Testungen erforderlich. Diese gehen allerdings mit einem Übungs- bzw. Lerneffekt einher. Insbesondere zu Beginn dementieller Erkrankungen könnte der Lerneffekt einen Einfluss auf das Testergebnis haben und damit Verlaufsuntersuchungen erschweren.

Um diesen Lerneffekt zu minimieren und eine kurzfristige Wiederholung des Tests zu ermöglichen, soll im Rahmen dieser Arbeit eine Parallelversion des TFDD entwickelt und validiert werden.

2 Methodik

2.1 Stichprobe

Die Stichprobe für die vorliegende Arbeit bestand aus Menschen mit einer bereits diagnostizierten Demenz bei Alzheimer-Krankheit in einem leichten bis zu einem mittelschweren Stadium (GDS-*Global Deterioration Scale*: III-V/VII) sowie neurokognitiv gesunden, altersgematchten Kontrollpersonen (KG). Alle an der Studie teilnehmenden Probanden sind im Rahmen der Gedächtnissprechstunde der Alexianer Krefeld GmbH (Krankenhaus Maria Hilf, akademisches Lehrkrankenhaus der Heinrich-Heine-Universität Düsseldorf) diagnostiziert und behandelt worden. Die Diagnose wurde von einem der Fachärzte für Neurologie, Psychiatrie und Psychotherapie gestellt, die im Rahmen der Gedächtnissprechstunde tätig sind, auf der Grundlage einer ausführlichen Eigen- und Fremdanamnese, neurologischen und psychiatrischen Untersuchung, Laboruntersuchung, zerebralen Bildgebung sowie eines Elektroenzephalogramms und eventuell einer Lumbalpunktion.

Den Probanden wurde die Teilnahme an der Studie freiwillig angeboten, es erfolgte eine ausführliche mündliche und schriftliche Aufklärung, alle Probanden unterschrieben vor Durchführung der Tests eine entsprechende Einverständniserklärung. Durch die Teilnahme an der Studie entstanden den Probanden keine Nachteile.

Die Voruntersuchung der Parallelitems erfolgte mit einer Stichprobe von n = 30 Menschen mit einer Demenz bei Alzheimer-Krankheit (AG1) in einem leichten bis hin zu einem mittelschweren Stadium (GDS-*Global Deterioration Scale*: III-V/VII). Für die anschließende Validierungsstudie wurden n = 60 Patienten mit einer bereits diagnostizierten Demenz bei Alzheimer-Krankheit (AG2) in einem leichten bis zu einem mittelschweren Stadium (GDS-*Global Detorioration Scale*: III-V/VII) untersucht. Davon waren 60 % weiblich, 40 % männlich und das Durchschnittsalter betrug 75,4 Jahre (Median = 77).

Für die Kontrollgruppe wurden n = 40 kognitiv unauffällige Angehörige der Patienten der Gedächtnissprechstunde getestet. Davon waren 58 % weiblich, 42 % männlich und das Durchschnittsalter betrug 74 Jahre (Median = 74).

2.2 Einschluss- und Ausschlusskriterien

Für AG1 und AG2 wurde die Diagnose einer wahrscheinlichen Demenz bei Alzheimer-Krankheit nach den NIA-AA-Kriterien (*National Institute on Aging und die Alzheimer's Association*) gestellt. Zusätzlich waren bei allen AG1- und AG2-Probanden die DSM-V-Kriterien einer neurokognitiven Störung bei einer wahrscheinlichen bzw. möglichen Alzheimer-Krankheit sowie die ICD-10-Kriterien für das Vorliegen einer Alzheimer-Demenz erfüllt.

Um sprachliche Barrieren bei der Durchführung des Tests zu vermeiden, wurden nur Menschen untersucht, deren Muttersprache Deutsch ist.

Ausschlusskriterien waren das Vorliegen anderer psychiatrischer oder organischer Erkrankungen, die kognitive Störungen verursachen, sowie fehlende Einwilligungserklärungen. Darüber hinaus wurden Menschen mit einer Demenz in einem fortgeschrittenen Stadium ausgeschlossen, da der TFDD als Screening-Verfahren gezielt ermöglichen soll, Patienten mit einer beginnenden Demenz zu identifizieren.

Ausschlusskriterien für die Kontrollgruppe waren Beeinträchtigungen der alltäglichen Funktionen, beschriebene Defizite der kognitiven Leistungen und andere angegebene psychische Auffälligkeiten.

Aufgrund der höheren Prävalenz der Alzheimer-Krankheit im fortgeschrittenen Alter wurde eine Altersbegrenzung vorgenommen und sich in diesem Zuge für das zusätzliche Ausschlusskriterium Alter < 50 Jahre entschieden. Dadurch war das Durchschnittalter beider Gruppen vergleichbar.

2.3 Prozedere

2.3.1 Testkonstruktion

Für die Parallelversion des TFDD mussten zunächst neue Parallelitems entwickelt werden, mit denen die gleichen kognitiven Domänen untersucht werden und die die gleichen Messeigenschaften wie die Originalitems aufweisen, aber nicht identisch sind. Die Formulierung der Items erfolgte inhaltlich bzw. theoretisch auf

der Basis von Praxiserfahrungen und empirischen Befunden mit Anregungen aus der aktuellen Literatur. Bei den Depressionsitems, mit denen nur eine Einschätzung der gegenwärtigen Stimmungslage vorgenommen wird, sind keine Übungseffekte zu erwarten und somit sollte sich auch kein Lerneffekt einstellen. Demzufolge wurden für die Depressionsitems keine Parallelversionen entwickelt und sie konnten unverändert in den Paralleltest übernommen werden.

So wurde ein Test mit neun Parallelitems entwickelt, für die Prüfung der folgenden kognitiven Bereiche:

- Gedächtnis durch die unmittelbare und verzögerte Reproduktion von sieben Wörtern
- Orientierung mit anschließenden Kategorisierungsaufgaben
- Exekutive Funktionen und insbesondere die Fähigkeit zu komplexeren Handlungsabläufen durch Befolgen von Anweisungen
- Konstruktive Praxis durch den Uhren-Zeichen-Test (*Clock Drawing Test*)
- Wortgenerieren mittels einer kategorialen Wortflüssigkeitsaufgabe

Für die Items 1 und 8 (unmittelbare und verzögerte Reproduktion) wurden entsprechend den originalen Aufgaben sieben Substantive mit unterschiedlicher Häufigkeit in der gesprochenen Sprache gewählt. Als Quelle wurde das Häufigkeitswörterbuch gesprochener Sprache des deutschen Sprachwissenschaftlers Arno Ruoff (Ruoff, 1981) verwendet. Weiterhin wurden Wörter ausgesucht, die keine starke assoziative Beziehung untereinander haben, um Verknüpfungen zu bereits aktivierten Denkinhalten zu vermeiden.

Zur Prüfung der zeitlichen Orientierung wurde für das Item 2 die Frage nach dem aktuellen Datum unverändert belassen, da beim wiederholten Einsatz des Items an unterschiedlichen Tagen kein Lerneffekt zu erwarten ist.

Für die Items 3 bis 5 (zeitliche Orientierung und Kategorisieren) wurden drei neue Aufgaben entwickelt. Dabei wurde darauf geachtet, dass Schweregrad und Punktaufteilung den Aufgaben des Originaltests entsprechen und das Untersuchungsziel exakt abbilden. So wurde bei der Entwicklung der Items für das Fragen nach den Tageszeiten statt nach den Jahreszeiten entschieden. Morgen und Vormittag sowie Mittag und Nachmittag wurden wegen der starken zeitlichen

Überlappung als jeweils eine Zeitspanne festgelegt. Bei der Frage nach der aktuellen Tageszeit wurde eine Toleranz von einer Stunde festgelegt, weil die Tageszeiten relativ unverbindlich sind und von Region und Jahreszeit abhängen (Thompson, 1967). Die Stundeneinteilung der Tageszeiten fußt auf der gesetzlichen Definition von Nachtruhe bzw. Mittagsruhe und wird in Tabelle 12 dargestellt. Demzufolge ist diese Einordung der Tageszeiten nur für den deutschen Kulturraum gültig.

	Sommer	Winter
Morgen/Vormittag	7:00–13:00	6:00–12:00
Mittag/Nachmittag	13:00–18:00	12:00–17:00
Abend	18:00–22:00	17:00– 21:00
Nacht	22:00–7:00	21:00– 6:00

Tabelle 12: Stundeneinteilung der Tageszeiten im Sommer und Winter

Für das Item 6 wurde zur Prüfung der Fähigkeit zu nicht automatisierten komplexeren Handlungsabläufen eine Aufgabe zum Befolgen von Anweisungen entwickelt, wobei die Anweisung im Ganzen gegeben wird und die einzelnen Schritte ähnliche Handlungsabläufe wie bei der originalen Aufgabe darstellen.

Für das Item 7 wurde der Uhren-Zeichen-Test übernommen, da dieser Test zur Erfassung von Störungen der Visuokonstruktion seit vielen Jahrzehnten sich bewährt hat (Shulman, 2000). Um den Lerneffekt zu minimieren, wurde dabei nach einer neuen Uhrzeit gesucht. Die Uhrzeit 8:20 gehört zu den drei häufigsten verwendeten Uhrzeiten und zeigt nach der Uhrzeit 11:10 die beste Sensitivität zur Entdeckung von kognitiven Einschränkungen (Freedman et al., 1994). Ähnlich wie die Uhrzeit 11:10 hat die Uhrzeit 8:20 darüber hinaus den Vorteil, dass mit ihr bilaterale hemisphärische Prozesse überprüft werden (Zeiger auf unterschiedlichen Seiten des Zifferblattes).

Für das Item 9 wurde wie beim Originaltest auch für die Parallelversion des Tests eine semantische Wortflüssigkeitsaufgabe ausgewählt. Neben dem Aufzählen von Tieren ist das Benennen von Supermarkt-Produkten eine der am häufigsten untersuchten Aufgaben zur Erfassung der semantischen Wortgenerierungsfähigkeit (Clark et al., 2009).

2.3.2 Voruntersuchung der Parallelitems

Die neu formulierten Parallelitems wurden zu einem vorläufigen Vortest zusammengestellt. Um zu bestimmen, ob die neu entwickelten Aufgaben als Parallelitems geeignet sind, erfolgte eine Voruntersuchung mit einer ersten Stichprobe von *n* = 30 Probanden mit einer bereits diagnostizierten Demenz bei Alzheimer-Krankheit (AG1). Alle Probanden wurden mit Original- und Parallelitems untersucht. Die Untersuchungen wurden in zwei separaten Sitzungen innerhalb eines Monats durchgeführt und die Reihenfolge war nicht vorgegeben, sondern erfolgte nach dem Zufallsprinzip. Es wurden die zentralen Tendenzen (Mittelwert, Median) für Originalitems und Parallelversionen verglichen, bevor anschließend ermittelt werden konnte, wie hoch die beiden Versionen miteinander korrelieren.

2.3.3 Itemanalyse

Nach der initialen Voruntersuchung der Items erfolgte mit einer umfangreicheren separaten Stichprobe von *n* = 60 Probanden mit einer Demenz bei Alzheimer-Krankheit (AG2) und *n* = 40 neurokognitiv gesunden Kontrollpersonen (KG) die anschließende Validierungsstudie.

Im Rahmen der Itemanalyse wurde zum einen die Dimensionalität des Tests geprüft und zum anderen erfolgte eine Berechnung der Schwierigkeit und der Trennschärfe der verschiedenen Items.

Um die dimensionale Teststruktur zu überprüfen und zu entscheiden, ob die Items zu einem einzigen oder zu mehreren Faktoren zusammengefasst werden sollen, wurde eine explorative Faktorenanalyse durchgeführt. Diese erfolgte sowohl für den Originaltest als auch für die Parallelversion, um die Faktorenladungen der beiden Versionen vergleichen zu können.

Die Itemschwierigkeit bzw. der Schwierigkeitsindex wurden ermittelt, indem für jedes Item die Summe der erreichten Punkte durch die maximal erreichbare Punktzahl dividiert wurde. Diese Berechnung erfolgte für Original- und Parallelversion jedes Items für die gesamte Stichprobe (AG2 und KG).

Ebenfalls für die gesamte Stichprobe wurde der Trennschärfekoeffizient bzw. die Korrelation jedes einzelnen Items mit dem Gesamttest berechnet.

2.3.4 Reliabilitätsprüfung

Zur Prüfung der Reliabilität wurden Testwiederholungsreliabilität (Retest-Reliabilität) und interne Konsistenz berechnet.

Um die Retest-Reliabilität zu prüfen, wurden 40 Probanden von der AG2 bei einem zweiten Untersuchungstermin mit dem Paralleltest erneut getestet. Der zweite Termin fand zwei Wochen nach der ersten Untersuchung statt. Berechnet wurde die Korrelation der beiden Testergebnissen miteinander.

Bei der Konsistenzanalyse wurde für die gesamte Stichprobe (AG2 und KG) die Korrelation aller einzelnen Items miteinander berechnet.

2.3.5 Objektivitätsprüfung

Durch das standardisierte Testmanual und die genauen Angaben zur Durchführung und Auswertung des Tests können diese Aspekte der Objektivität als erfüllt betrachtet werden. Zur weiteren Prüfung der Interpretationsobjektivität wurde die Inter-Rater-Reliabilität berechnet. Hierfür wurden 20 Probanden von der AG2 durch zwei Untersucher (einen Facharzt für Neurologie und eine trainierte Arzthelferin) getestet. Um mögliche Interpretationsdifferenzen der beiden Untersucher zu erfassen, erfolgte die Testung wie folgt: Abwechselnd wurde die Hälfte der Probanden durch den ersten bzw. den zweiten Untersucher getestet. Dabei führte der eine die Untersuchung durch, während der andere diese beobachtete. Am Ende nahmen beide Untersucher die Auswertung parallel und selbstständig vor. Um die Messgenauigkeit bzw. Beurteiler-Übereinstimmung zu überprüfen, wurde die vergebene Punktzahl der beiden Untersucher verglichen und miteinander korreliert.

2.3.6 Validitätsprüfung

Zunächst wurde eine Konstruktvalidierung vorgenommen. Der Paralleltest sollte das angestrebte Konstrukt erfassen, wofür er entwickelt wurde, und zwar identisch zum Originaltest. Hierfür wurde die gesamte Stichprobe (AG2 und KG) mit

Original- und Paralleltest getestet und die Korrelation zwischen den beiden Testergebnissen berechnet. Darüber hinaus wurden deskriptiv die zentralen Tendenzen der Testergebnisse im Hinblick auf ihre Übereinstimmung geprüft. Durch die Korrelation der beiden Testergebnisse konnte, als Teilaspekt der Konstruktvalidität, die konvergente Validität geprüft werden. Des Weiteren wurde im Rahmen der Konstruktvalidierung die Korrelation des Paralleltests mit dem GDS (*Global Deterioration Scale for Assessment of Primary Degenerative Dementia*) berechnet.

In einem nächsten Schritt erfolgte eine Kriteriumsvalidierung. Als Außenkriterium wurde die Diagnose ‚Demenz' bzw. die Zugehörigkeit zur Gruppe AG2 oder KG festgelegt und anschließend die Diskriminationsfähigkeit des Paralleltests zwischen den beiden Gruppen anhand der Testergebnisse geprüft. Als Ausdruck der Zuordnungsgenauigkeit des Tests wurden hier Sensitivität und Spezifität ermittelt. Das identische Verfahren erfolgte für den Originaltest, um die Diskriminationsfähigkeit der beiden Versionen zu vergleichen.

Um die Eignung des Paralleltests als Screening-Verfahren zu untersuchen, wurden separat die Leistungen der Probanden mit einer beginnenden Demenz (GDS III-IV, n = 38) mit diesen der KG verglichen. Anhand der ermittelten Sensitivität und Spezifität wurde die Diskriminationsfähigkeit des Paralleltests zwischen den beiden Gruppen geprüft. Die Gruppenunterschiede und die Diskriminationsfähigkeit des Paralleltests (Sensitivität, Spezifität) wurden auch für n = 22 Menschen mit einer mittelschweren Demenz (GDS V) und KG ermittelt.

2.4 Verwendete neuropsychologische Testverfahren

Alle Patienten und Kontrollpersonen wurden mit der originalen und der Parallelversion des TFDD getestet. Die Untersuchungen erfolgten im Rahmen der Verlaufstermine innerhalb eines Monats. Die Reihenfolge war nicht vorgegeben, sondern ergab sich nach dem Zufallsprinzip.

Die genauen Aufgabenstellungen des Originaltests werden in der folgenden Tabelle dargestellt (Tabelle 13).

Item	Durchführung	Anweisung	Auswertung
Item 1 – Unmittelbare Reproduktion	Es werden sieben Wortkarten dargeboten, die der Proband laut vorliest und unmittelbar danach reproduziert. Zur Vorbereitung für das Item 8 werden die Wortkarten im Anschluss nochmal vorgelesen. Die Worte werden vom Probanden selbst laut gelesen, um den Störfaktor der Presbyakusis zu eliminieren.	*Lesen Sie bitte jedes Wort laut vor und prägen Sie es sich gut ein!* *An welche Wörter erinnern Sie sich?* Ggf. nach einer Minute: *Bitte lesen Sie jetzt die Wörter noch einmal!*	Für jedes richtig wiedergegebene Wort wird ein Punkt vergeben (maximal sieben Punkte)
Item 2 – Zeitliche Orientierung	Es wird nach dem aktuellen Datum gefragt und für die Beantwortung der Frage stehen 15 Sekunden zur Verfügung.	*Welches Datum ist heute?*	Für richtigen Tag, Monat und richtiges Jahr jeweils ein Punkt (maximal drei Punkte). Abweichungen vom aktuellen Datum um einen Tag erlaubt.
Item 3 – Zeitliche Orientierung	Es werden die Jahreszeiten befragt und für die Beantwortung der Frage sind 15 Sekunden vorgesehen.	*Welche Jahreszeiten gibt es?*	Für jede richtige Jahreszeit wird ein Punkt vergeben (maximal vier Punkte).
Item 4 – Zeitliche Orientierung	Es wird nach der aktuellen Jahreszeit gefragt und für die Beantwortung der Frage stehen 15 Sekunden zur Verfügung.	*Welche Jahreszeit haben wir jetzt?*	Für die richtige Antwort wird ein Punkt vergeben. 14 Tage vor oder nach dem Wechsel der Jahreszeiten werden beide Jahreszeiten als richtig gewertet.
Item 5 – Zeitliche Orientierung/Kategorisieren	Für diese Aufgabe wird die in Frage 3 genannte Jahreszeit zugrunde gelegt. Es werden für diese Jahreszeit die zugehörigen Monate erfragt. Es stehen für diese Aufgabe 30 Sekunden zu Verfügung.	*Welche Monate gehören zu dieser Jahreszeit?*	Für jeden richtig genannten Monat wird ein Punkt vergeben. Werden auch falsche Monate genannt, so wird für jeden falsch genannten Monat ein Punkt von der Anzahl richtig genannter Monate abgezogen (maximal vier Punkte).
Item 6 – Anweisungen befolgen	Für diese Aufgabe wird die Anweisung im Ganzen gegeben. Bei Bedarf ist eine Wiederholung zulässig. Auch bei der Wiederholung wird die Anweisung im Ganzen gegeben. Für jeden der beiden Versuche beträgt die zulässige Zeit 15 Sekunden.	*Greifen Sie sich erst mit der linken Hand an das rechte Ohr, dann mit der rechten an das linke Ohr und klatschen Sie danach in die Hände!*	Für jeden der drei Teilschritte wird ein Punkt vergeben. Die vollkommen richtige Reihenfolge der Ausführung wird mit einem weiteren Punkt bewertet (maximal vier Punkte).
Item 7 – Konstruktive Praxis	Bei dieser Aufgabe soll das Zifferblatt einer Uhr gezeichnet werden. Dem Probanden wird ein Blatt mit der Anweisung übergeben. Für die Lösung der Aufgabe stehen 60 Sekunden zur Verfügung.	*Bitte zeichnen Sie das Zifferblatt einer Uhr mit allen Zahlen und stellen Sie die Zeiger auf 11.10 Uhr ein!*	Die Auswertung der Uhr erfolgt nach dem Schema von Sunderland. Die Auswertung umfasst eine Skala von zehn Punkten, die von 10 ‚korrekte Darstellung' bis hin zu 1 ‚nicht mehr als Uhr erkennbar' reicht. Beurteilungsleitlinien mit illustrierten Beispielen helfen dem Beurteiler bei der Zuordnung des Punktwertes.
Item 8 – Verzögerte Reproduktion	Bei der verzögerten Reproduktion werden die Wörter abgefragt, die bei Item 1 zu lernen waren. Für die Beantwortung dieser Aufgabe stehen 60 Sekunden zur Verfügung.	*Vorhin haben Sie Wörter gelesen, die Sie sich einprägen sollten. An welche dieser Wörter können Sie sich noch erinnern?* Wird die Reproduktion für mehr als zehn Sekunden unterbrochen, wird der Patient aufgefordert: *Fällt Ihnen noch ein weiteres Wort ein?*	Für jedes erinnerte Wort wird ein Punkt vergeben (maximal sieben Punkte).
Item 9 – Wortflüssigkeit	*Hier werden die Probanden* aufgefordert, so viele Tiere wie möglich aufzuzählen und es gilt eine Zeitdauer von 60 Sekunden.	*Sie haben jetzt eine Minute Zeit, mir so viele Tiere zu nennen, wie Sie können!*	Für jedes genannte Tier wird bis zur Anzahl von zehn jeweils ein Punkt vergeben. Repetitionen und Eigennamen werden vom Gesamtscore abgezogen.

Tabelle 13: Aufgabenstellungen des Originaltests *(In Anlehnung an Ihl & Grass-Kapanke, 2000, S. 24–28)*

Darüber hinaus kamen zwei weitere neuropsychologische Testverfahren zum Einsatz. Zur Einschätzung des Demenzschweregrades wurde die *Global Deterioration Scale for Assessment of Primary Degenerative Dementia* (GDS) eingesetzt (Reisberg, 1982). Die GDS unterteilt 7 Stufen der Krankheit, welche die Veränderungen der kognitiven Fähigkeiten im Verlauf dementieller Erkrankungen abbilden (Tabelle 14). Hier wurde die deutsche Übersetzung verwendet (Ihl & Fröhlich, 1991).

Stadium	Klinische Merkmale
Stufe I–Keine kognitiven Einbußen	Im ersten Stadium zeigt sich klinisch ein normaler Befund. Es liegen weder subjektive noch objektivierbare Gedächtnisprobleme vor.
Stufe II–Zweifelhafte kognitive Einbußen	Erste Gedächtnisdefizite werden wahrgenommen. Am häufigsten werden Namen vergessen und Gegenstände verlegt. Die Defizite sind klinisch nicht objektivierbar.
Stufe III–Geringe kognitive Einbußen	Erste objektivierbare kognitive Defizite. Diese lassen sich in einem klinischen Interview nachweisen. Konzentrationsstörungen sowie Schwierigkeiten der räumlichen Orientierung. Reduzierte Leistung im Beruf und im sozialen Umfeld. Diese ersten Einschränkungen werden von Familienmitgliedern, Freunden oder Bekannten bemerkt.
Stufe IV–Mäßige kognitive Einbußen	Eindeutige Defizite in verschiedenen kognitiven Bereichen. Aktuelle oder kurz zurückliegende Ereignisse werden vergessen und das Erinnern des eigenen Lebenslaufs ist beeinträchtigt. Komplexe Aufgaben können nicht mehr selbstständig durchgeführt werden. Konzentrationsschwierigkeiten. Die betroffenen Personen ziehen sich oft zurück, isolieren sich und versuchen die Defizite zu leugnen bzw. verheimlichen.
Stufe V–Mittelschwere kognitive Einbußen	Zunehmende Hilfe bei den alltäglichen Aktivitäten erforderlich. Wichtige persönliche Daten (Adresse, Telefonnummer) werden vergessen. Einige Fakten von sich selbst und der Familie sind noch abrufbar (Namen). Häufig zeitliche und örtliche Desorientierung. Toilettengang und Einnahme von Mahlzeiten noch selbstständig.
Stufe VI–Schwere kognitive Einbußen	Der Alltag kann ohne fremde Hilfe nicht mehr bewältigt werden. Die Namen der engen Angehörigen werden vergessen und der Tag-Nacht-Rhythmus ist oft gestört. Verhaltensstörungen im Sinne von Misstrauen, Zwangshandlungen, Agitiertheit, Apathie oder Aggressivität werden beobachtet. Auch Halluzinationen und Wahnideen können auftreten. Die Patienten verlieren allmählich die Fähigkeit, in vollständigen Sätzen zu sprechen. Harninkontinenz tritt auf.
Stufe VII–Sehr schwere kognitive Einbußen	Spätstadium der Alzheimer-Krankheit. Erhebliche Sprachstörungen bis zum totalen Sprachverlust treten auf, motorische Fähigkeiten sind oft schwer beeinträchtigt, die Betroffenen können nicht mehr laufen oder selbstständig essen. Fokale neurologische Ausfälle können häufig in der neurologischen Untersuchung festgestellt werden.

Tabelle 14: The Global Deterioration Scale for Assessment of Primary Degenerative Dementia *(In Anlehnung an Ihl & Fröhlich, 1991, S.10–15)*

2.5 Datenerhebung

Die Datenerhebung erfolgte im Rahmen der Gedächtnissprechstunde der Alexianer Krefeld GmbH, Krankenhaus Maria Hilf, akademisches Lehrkrankenhaus der Heinrich-Heine-Universität Düsseldorf. Patienten, die für eine Demenzabklärung oder Verlaufsuntersuchung bei bereits diagnostizierter Demenz vorstellig wurden, wurde die Teilnahme an der Validierungsstudie angeboten. Wenn der Patient oder seine bevollmächtigte Person sein Einverständnis gab, wurden Original- und Parallelversion des TFDD in zwei separaten Sitzungen innerhalb eines Monats durchgeführt. Die Reihenfolge der Untersuchung ergab sich dabei durch das Zufallsprinzip. Nur die zweite Durchführung des Paralleltests erfolgte zwei Wochen nach der ersten Untersuchung zur Prüfung der Retest-Reliabilität.

Für die Kontrollgruppe wurden die Angehörigen der Patienten angefragt, ob sie einverstanden sind, an der Validierungsstudie teilzunehmen. Erklärten sie ihr Einverständnis, erfolgte die Teilnahme an der Studie genau wie bei den Alzheimer-Patienten. Die Untersuchungen wurden auch hier im Rahmen der Verlaufstermine innerhalb eines Monats durchgeführt.

2.6 Statistische Auswertung

Die statistische Analyse wurde mithilfe der Statistiksoftware SPSS 25.0 durchgeführt (IBM, 2017). Ziel der statistischen Analyse war die Itemanalyse sowie Prüfung von Objektivität, Reliabilität und Validität.

2.6.1 Voruntersuchung

Die erhobenen Daten wurden mit dem Kolmogorov-Smirnov- und dem Shapiro-Wilk-Test auf Normalverteilung geprüft. Für jedes einzelnes Item sowie Gesamtpunktzahl fand sich ein Wert von $p < .05$. Daraufhin musste die Nullhypothese abgelehnt werden, was bedeutet, dass die Daten nicht normalverteilt sind. Demzufolge erfolgte die Analyse mit non-parametrischen Verfahren.

Um die Eignung der neu entwickelten Items als Parallelitems zu prüfen, wurde in der Voruntersuchung mit dem Wilcoxon-Test ermittelt, ob beide Itemversionen vergleichbare zentrale Tendenzen aufweisen. Darüber hinaus wurde der

Spearman-Korrelationskoeffizient ermittelt, um herauszufinden, ob beide Itemversionen signifikant miteinander korrelieren.

2.6.2 Faktorenanalyse

Die Eignung der Daten für die Faktorenanalyse wurde nach Kaiser-Meyer-Olki überprüft. Die Sphärizität wurde mit dem Bartlett-Test geprüft. Die optimale Faktorenzahl, die extrahiert werden sollte, wurde mit dem graphischen Verfahren Scree-Plot ermittelt. In der Faktorenmatrix wurden die Faktorenladungen der Parallelitems dargestellt. Die gleiche statistische Auswertung wurde auch für den Originaltest vorgenommen.

2.6.3 Prüfung der Reliabilität und Objektivität

Zur Reliabilitätsprüfung wurde der Spearman-Korrelationskoeffizient zwischen den Testergebnissen der beiden Untersuchungen mit dem Paralleltest (n = 40, AG2) berechnet.

Zur Prüfung der internen Konsistenz (Konsistenzanalyse) wurden Cronbachs Alpha und Item-Trennschärfe für jedes Item des Paralleltests ermittelt (Korrelation des Items mit dem Gesamttest) und mit dem Ergebnis für die Originalversion verglichen.

Zur Prüfung der Objektivität wurde der Spearman-Korrelationskoeffizient zwischen der vergebenen Punktzahl der beiden Untersucher (n = 20, AG2) berechnet.

2.6.4 Validitätsprüfung

Zur Prüfung der konvergenten Validität wurde der Spearman-Korrelationskoeffizient zwischen den Testergebnissen von Parallel- und Originaltest sowie zwischen Paralleltest und GDS berechnet.

Im Rahmen der Kriteriumsvalidierung wurde für Parallel- und Originaltest anhand des Gesamtscores für die beiden diagnostischen Gruppen (AG2, KG) mit dem Mann-Whitney-U-Test auf Gruppenunterschiede geprüft. Weiterhin wurden mit

dem Verfahren der *Receiver-Operator-Characteristics* (ROC) die maximale Spezifität und Sensitivität berechnet, um festzustellen, wie gut Parallel- und Originaltest zwischen den beiden diagnostischen Gruppen (AG2, KG) unterscheiden können.

Als weiteres Maß für die diagnostische Zuordnungsfähigkeit des Tests wurde der Youden-Index [J=Sensitivität-(1-Spezifität)] berechnet (Youden, 1950). Ein wesentlicher Zugewinn zum Youden-Index durch Verwendung von komplexeren Formeln, wie der *Correct Classification Accuracy* (CCA), ist fraglich, da ihre Berechnung den gleichen Einschränkungen wie die Berechnung des Youden-Index unterliegt. Durch einen komplizierteren Rechenweg ergeben sich keine wesentlichen Änderungen an der Beeinflussung durch unterschiedliche Gruppen (Ihl, 2020).

Zur grafischen Darstellung der Ergebnisse mit dem Youden-Index und der maximalen Distanz wurden die Berechnungen in *Microsoft Excel* 2019 durchgeführt.

Zusätzlich wurden für die Parallelversion separat die Probanden mit einer leichten Demenz (GDS III-IV, n = 38) mit der gesunden Kontrollgruppe (n = 40) verglichen. Es folgte eine deskriptive Analyse, Prüfung auf Gruppenunterschiede mittels Mann-Whitney-U-Test sowie Ermittlung der maximalen Spezifität und Sensitivität mittels ROC-Verfahren.

Die gleiche statistische Auswertung wurde mit dem Paralleltest für die Gruppen der mittelschweren Demenz (GDS V, n = 22) und der gesunden Kontrollgruppe (n = *40*) durchgeführt.

Als Signifikanzniveau wurde ein Wert von a = .05 für alle induktiven statistischen Verfahren festgelegt.

2.7 Ethikvotum

Im Rahmen der ethischen und rechtlichen Beratung von der Ethikkommission der Medizinischen Fakultät der Heinrich-Heine-Universität Düsseldorf bestanden keine Bedenken gegen die Durchführung der Studie (Studiennummer: 5133).

3 Ergebnisse

3.1 Zusammenstellung des Paralleltests

3.1.1 Unmittelbare und verzögerte Reproduktion (Item 1, Item 8)

Beim ersten Item werden 7 Wortkarten dargeboten, die der Proband laut vorliest und unmittelbar danach reproduziert. Zur Vorbereitung für das Item 8 werden die Wortkarten im Anschluss nochmal vorgelesen.

Beim Item 8 werden bei der verzögerten Reproduktion die Wörter erneut abgefragt. Für die Beantwortung dieser Aufgabe stehen 60 Sekunden zur Verfügung.

Die gewählten Worte und deren prozentualer Anteil am Gesamten der betreffenden Wortart (Substantive) sind in Tabelle 15 aufgelistet.

Für beide Items sind für die untersuchte Person jeweils maximal sieben Punkte zu erreichen.

Originalbegriffe	Neue Substantive	Prozentualer Anteil am Gesamt der Substantive
Verkäufer	Händler	0,06%
Komet	Mond	0,02%
Nachricht	Lehre	0,08%
Spiegel	Auto	0,23%
Märchen	Film	0,03%
Dampf	Rauch	0,01%
Abenteuer	Prüfung	0,05%

Tabelle 15: Originalbegriffe und neu gewählte Substantive mit deren Häufigkeit in der gesprochenen Sprache

3.1.2 Zeitliche Orientierung und Kategorisieren (Items 2–5)

Beim Item 2 wird unverändert zur originalen Aufgabe nach dem aktuellen Datum gefragt. Für die Beantwortung der Frage stehen 15 Sekunden zur Verfügung. Hier sind maximal drei Punkte zu erreichen, jeweils ein Punkt für den richtigen Tag, Monat und das richtige Jahr. Eine Abweichung vom aktuellen Datum um einen Tag ist gestattet.

Beim Item 3 werden die Tageszeiten abgefragt. Für die Beantwortung dieser Frage stehen ebenfalls 15 Sekunden zur Verfügung und für jede richtige Tageszeit wird ein Punkt vergeben (maximal vier Punkte).

Beim Item 4 wird nach der aktuellen Tageszeit gefragt und es wird ein Punkt für die richtige Antwort vergeben. Abweichungen von der aktuellen Tageszeit um eine Stunde sind erlaubt.

Beim Item 5 müssen vier konkrete Uhrzeiten (19:00, 13:00, 0:00, 7:00) zu den jeweiligen Tageszeiten zugeordnet werden. Für diese Aufgabe stehen 30 Sekunden zur Verfügung und für jede korrekte Antwort wird ein Punkt vergeben.

3.1.3 Befolgen von Anweisungen (Item 6)

Beim Item 6 wird eine Anweisung im Ganzen gegeben und bei Bedarf wiederholt. Für die Durchführung der angewiesenen Handlungen stehen 15 Sekunden zur Verfügung. Die neue Anweisung des Parallelitems lautet:

Greifen Sie sich erst mit der linken Hand an den rechten Ellenbogen, dann mit der rechten Hand an den linken Ellenbogen und legen Sie danach beide Hände auf den Tisch.

Für jede richtig ausgeführte Handlung wird ein Punkt vergeben. Wenn alle drei Teilschritte in der richtigen Reihenfolge erfolgen, wird ein weiterer Punkt vergeben. Somit beträgt die Maximalpunktzahl für diese Aufgabe vier Punkte.

3.1.4 Konstruktive Praxis (Item 7)

Beim Item 7 wird der Proband aufgefordert, das Zifferblatt einer Uhr zu zeichnen und die Zeiger auf 8:20 einzustellen. Dabei wird ein Blatt mit der Anweisung

ausgehändigt. Für die Lösung der Aufgabe stehen 60 Sekunden zur Verfügung. Die Auswertung erfolgt wie beim Originaltest nach dem Schema von Sunderland und umfasst eine Skala von 10 Punkten. Beurteilungsleitlinien mit illustrierten Beispielen helfen dem Untersucher bei der Zuordnung des Punktwertes.

3.1.5 Wortflüssigkeit (Item 9)

Beim Item 8 wird der Proband aufgefordert, möglichst viele Dinge zu nennen, die in einem Supermarkt gekauft werden können. Im genauen Wortlaut lautet die Aufgabe folgendermaßen:

Für die nächste Aufgabe haben Sie eine Minute Zeit. Bitten nennen Sie mir so viele Dinge wie möglich, die man im Supermarkt kaufen kann!

Für jeden genannten Artikel wird ein Punkt vergeben. Die Maximalpunktzahl beträgt zehn Punkte.

3.2 Untersuchung des Vortests

Die Probanden der AG1 wurden mit Original- und Parallelitems untersucht. In den Tabellen 16 und 17 werden Mittelwerte und Medianwerte beider Versionen der Items dargestellt. Original- und Parallelversion wiesen für jedes Item vergleichbare Mittelwerte und nahezu identische Medianwerte auf.

		Item 1 TFDD	Item 2 TFDD	Item 3 TFDD	Item 4 TFDD	Item 5 TFDD	Item 6 TFDD	Item 7 TFDD	Item 8 TFDD	Item 9 TFDD
n		30	30	30	30	30	30	30	30	30
M		2.07	1.97	2.87	.63	2.53	3.20	5.77	.77	7.37
Median		2.00	2.00	3.00	1.00	3.00	4.00	5.50	.00	8.00
SD		1.230	.964	1.196	.490	1.074	1.064	2.431	1.040	2.189
Perzentile	25	1.00	1.00	2.00	.00	2.00	2.00	4.00	.00	5.75
	50	2.00	2.00	3.00	1.00	3.00	4.00	5.50	.00	8.00
	75	3.00	3.00	4.00	1.00	3.00	4.00	8.00	1.00	9.25

Tabelle 16: Mittelwert und Median für Originalitems
TFDD: Originaltest, n: Größe der Stichprobe, M: Mittelwert, SD: Standardabweichung

		Item 1 TFDD-P	Item 2 TFDD-P	Item 3 TFDD-P	Item 4 TFDD-P	Item 5 TFDD-P	Item 6 TFDD-P	Item 7 TFDD-P	Item 8 TFDD-P	Item 9 TFDD-P
n		30	30	30	30	30	30	30	30	30
M		1.97	2.03	2.97	.60	2.70	3.10	5.87	.80	7.50
Median		2.00	2.00	3.00	1.00	3.00	4.00	5.00	.00	8.00
SD		1.326	1.098	1.217	.498	1.393	1.185	2.543	1.215	2.286
Perzentile	25	1.00	1.00	2.00	.00	1.75	2.00	4.00	.00	5.00
	50	2.00	2.00	3.00	1.00	3.00	4.00	5.00	.00	8.00
	75	3.00	3.00	4.00	1.00	4.00	4.00	8.00	1.00	10.00

Tabelle 17: Mittelwert und Median für Parallelitems
TFDD-P: Paralleltest, n: Größe der Stichprobe, M: Mittelwert, SD: Standardabweichung

Mithilfe des Wilcoxon-Tests konnte ebenfalls für alle Items kein signifikanter Unterschied zwischen den beiden Versionen nachgewiesen werden (Tabelle 18).

	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7	Item 8	Item 9
Z	-.832[a]	-.632[b]	-1.342[b]	-1.000[a]	-1.291[b]	-1.342[a]	-.632[a]	-.378[b]	-.447[a]
Asymptotische Signifikanz (2-seitig)	.405	.527	.180	.317	.197	.180	.527	.705	.655

Tabelle 18: Wilcoxon-Test für Original- und Parallelitems
a. Basiert auf negativen Rängen, b. Basiert auf positiven Rängen, p< .05

Um zu überprüfen, wie hoch die beiden Versionen miteinander korrelieren, wurde für alle Items der Spearman-Korrelationskoeffizient ermittelt.

Für jedes Item fand sich eine signifikante Korrelation zwischen Original- und Parallelversion (Item 1: $r_s = .872$, Item 2: $r_s = .810$, Item 3: $r_s = .882$, Item 4: $r_s = .932$, Item 5: $r_s = .863$, Item 6: $r_s = .925$, Item 7: $r_s = .924$, Item 8: $r_s = .866$, Item 9: $r_s = .844$, $p < .001$, $n = 30$).

3.3 Faktorenanalyse

Die Struktur des Paralleltests wurde mittels einer explorativen Faktorenanalyse getestet. Der Kaiser-Meyer-Olkin-Wert betrug .890 und lag damit deutlich über dem Minimum von 0.5 oder 0.6, was häufig empfohlen wird (Yong & Pearce, 2013; Kaiser, 1974). Die Sphärizität wurde mit dem Bartlett-Test geprüft (Chi-

Quadrat (120) = 635.049, $p < .001$). Somit gab es keinen Anlass, die Analyse zu unterbrechen.

Anhand des Scree-Plots (Abbildung 1) und der Faktorenmatrix wurde ein Generalfaktor gewählt. Dabei handelt es sich um den Faktor Demenz und dieser erklärt 63,3 % der Varianz (Tabelle 19).

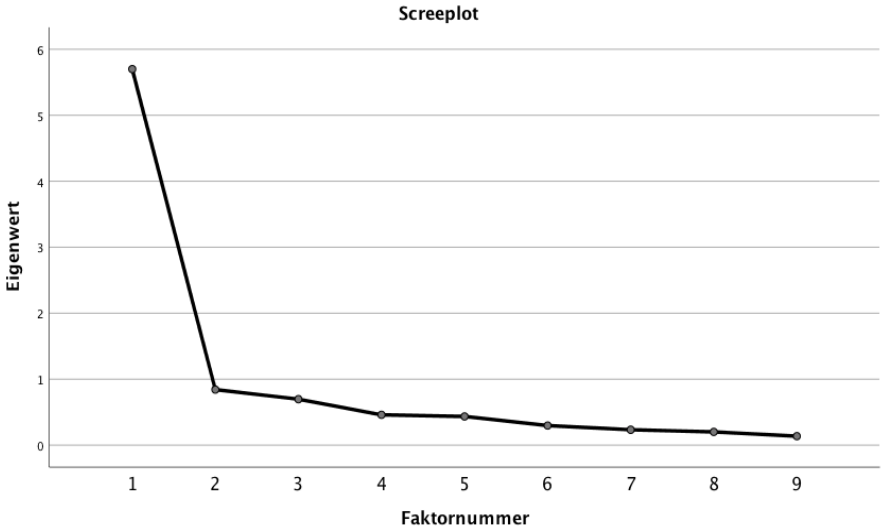


Abb. 1: SPSS-Output-Scree-Plot für den Paralleltest

	Generalfaktor
Item 1	.791
Item 2	.733
Item 3	.794
Item 4	.758
Item 5	.849
Item 6	.629
Item 7	.853
Item 8	.723
Item 9	.753

Tabelle 19: Faktorenladungen der Parallelitems auf dem Generalfaktor
Es wurde nur 1 Faktor extrahiert, Extraktionsmethode: Hauptkomponentenanalyse

Für den Originaltest wurde ebenfalls eine erneute Faktorenanalyse vorgenommen. Auch hier wurde die Stichprobeneignung geprüft und sowohl der Bartlett-Test (Chi-Quadrat (120) = 601.388, $p < .001$) als auch das Kaiser-Meyer-Olkin-Wert (KMO = .888) wiesen darauf hin, dass sich die Variablen für eine Faktoranalyse eignen.

Der Scree-Plot war identisch zu dem der Parallelversion und bei der Analyse des Scree-Plots und der Faktorenmatrix zeigte sich ein Generalfaktor, mit dem sich 61,1 % der Varianz erklären lässt.

3.4 Itemschwierigkeit

In Abbildung 2 ist der Anteil an vollständig richtigen Lösungen für jedes einzelne Item ersichtlich. Die Probanden der KG wiesen in allen Aufgaben einen höheren prozentualen Anteil vollständig richtiger Lösungen im Vergleich zur AG2 auf. Die größten Unterschiede fanden sich in den Items 3 (Tageszeiten nennen), 7 (Uhren-Zeichen-Test) und 9 (Wortflüssigkeit).

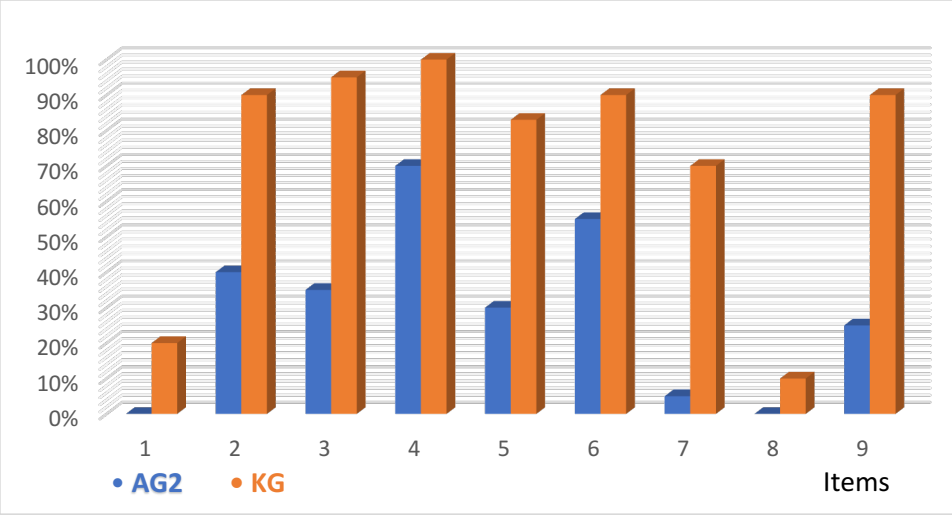


Abb. 2: Prozentualer Anteil vollständig richtiger Lösungen pro Item
AG2: Menschen mit einer Demenz bei Alzheimer-Krankheit, KG: Kontrollgruppe

Es folgte die Berechnung des Schwierigkeitsindexes für mehrstufige Items bei AG2 und KG. Dieser wurde für beide Versionen aller Items berechnet und beide Versionen wurden anschließend miteinander verglichen. Die Ergebnisse sind in der Tabelle 20 zu erkennen.

	Item 1	**Item 2**	**Item 3**	**Item 4**	**Item 5**	**Item 6**	**Item 7**	**Item 8**	**Item 9**
Paralleltest (AG2)	.36	.63	.69	.72	.66	.79	.57	.1	.75
Originaltest (AG2)	.37	.62	.67	.74	.67	.79	.57	.1	.74
Paralleltest (KG)	.82	.96	.98	1	.94	.97	.94	.65	.98
Originaltest (KG)	.83	.97	.97	1	.93	.98	.94	.64	.98
Paralleltest (AG2+KG)	.54	.76	.81	.83	.76	.85	.71	.31	.85
Originaltest (AG2+KG)	.56	.76	.79	.86	.76	.86	.72	.29	.84

Tabelle 20: Schwierigkeitsindex für Original- und Parallelitems
AG2: Menschen mit einer Demenz bei Alzheimer-Krankheit, KG: Kontrollgruppe

Bei der Analyse der genauen Leistungsverteilung für die einzelnen Items der Parallelversion lassen sich beim ersten Item erhebliche Unterschiede mit geringer Überlappung der Punktwerte zwischen AG2 und KG feststellen (Abbildung 3).

85 % der AG2-Probanden erreichten zwischen einem und vier Punkten (Nennungen). 95 % der KG-Probanden erzielten hingegen Werte zwischen fünf und sieben Punkten, sodass die Verteilung im Vergleich deutlich nach rechts verschoben ist.

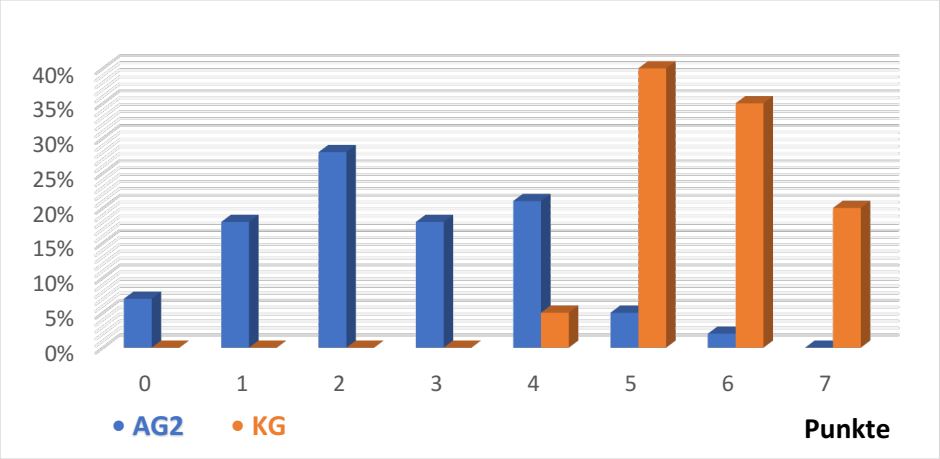


Abb. 3: Leistungsverteilung-Item 1
AG2: Alzheimer-Demenz-Gruppe, KG: Kontrollgruppe

Beim zweiten Item (Frage nach aktuellem Datum) erreichten 90 % der KG-Probanden die volle Punktzahl. Vier gesunde Personen erzielten nur zwei Punkte, da sie den aktuellen Tag nicht richtig genannt hatten. 24 AG2-Probanden bzw. 40 % lösten die Aufgabe ohne Probleme, die anderen zeigten Defizite bis hin zu einer vollständigen zeitlichen Desorientierung (Abbildung 4).

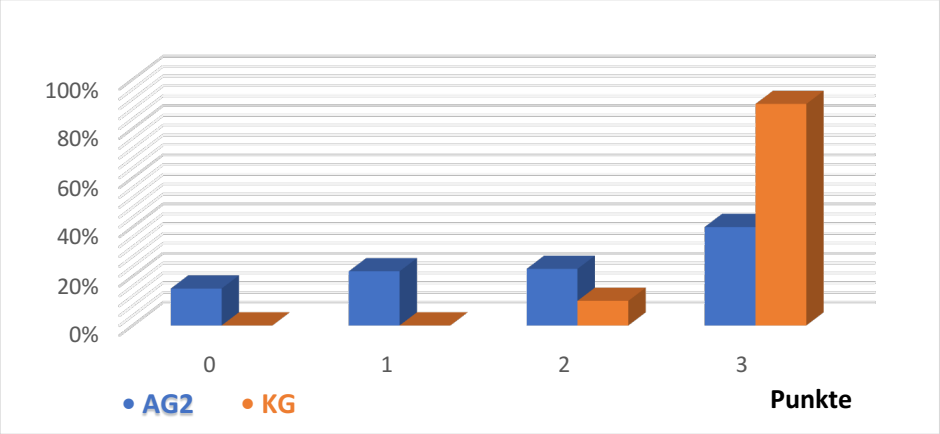


Abb. 4: Leistungsverteilung-Item 2
AG2: Alzheimer-Demenz-Gruppe, KG: Kontrollgruppe

Das Nennen der Tageszeiten war für die KG-Probanden ebenfalls unproblematisch, zwei gesunde Personen kamen auf drei Nennungen und erreichten damit drei Punkte. 21 Probanden der AG2 bzw. 35 % erreichten die volle Punktzahl und die anderen zeigten Defizite unterschiedlicher Ausprägung (Abbildung 5).

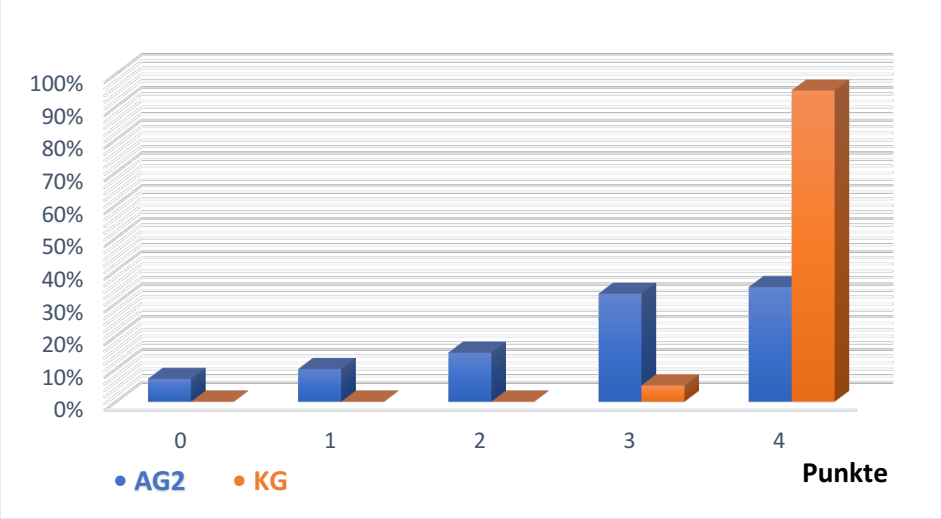


Abb. 5: Leistungsverteilung-Item 3
AG2: Alzheimer-Demenz-Gruppe, KG: Kontrollgruppe

Im Item 4 konnten 30 % der AG2-Probanden die aktuelle Tageszeit nicht nennen. Die Probanden der KG hatten bei dieser Aufgabe keine Schwierigkeiten (Abbildung 6).

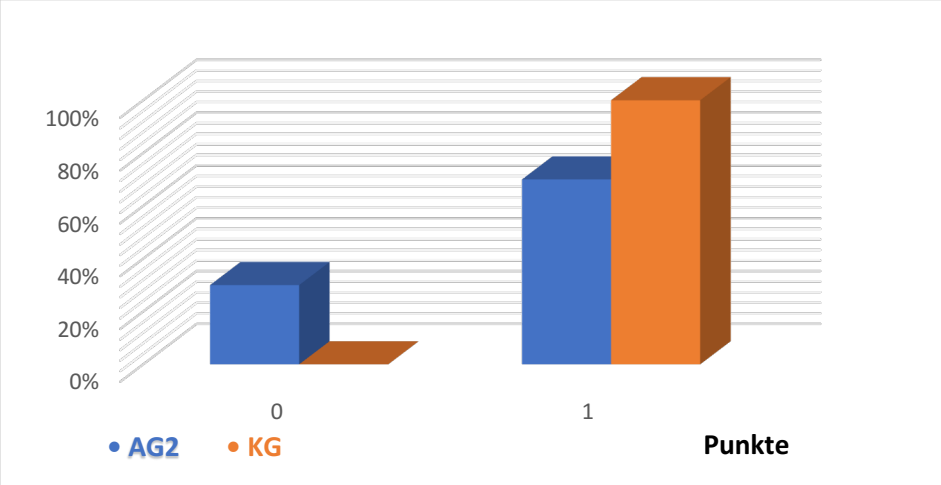


Abb. 6: Leistungsverteilung-Item 4
AG2: Alzheimer-Demenz-Gruppe, KG: Kontrollgruppe

95 % der KG-Probanden konnten im Item 5 die Uhrzeiten den entsprechenden Tageszeiten zuordnen. Nur zwei davon konnten eine Uhrzeit nicht richtig zuordnen und erreichten damit drei Punkte. 65 % der AG2-Probanden hatten Schwierigkeiten mit dieser Aufgabe und erreichten null bis drei Punkte (Abbildung 7).

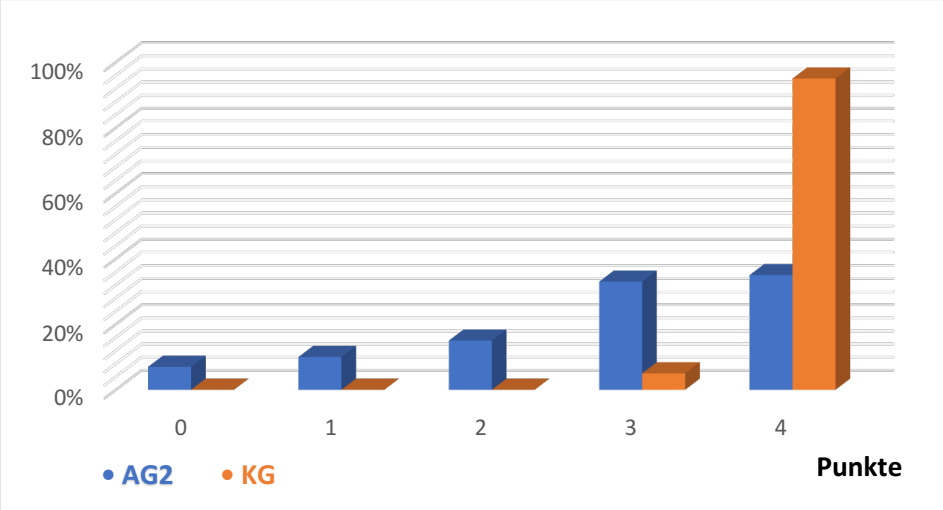


Abb. 7: Leistungsverteilung-Item 5
AG2: Alzheimer-Demenz-Gruppe, KG: Kontrollgruppe

Beim Befolgen von Anweisungen (Item 6) hatten vier KG-Probanden bzw. 10 % der Gruppe Schwierigkeiten und erzielten wegen der falschen Reihenfolge der Durchführung nur drei Punkte. 70 % der AG2-Probanden führten die Aufgabe unzureichend aus und erreichten null bis drei Punkte (Abbildung 8).

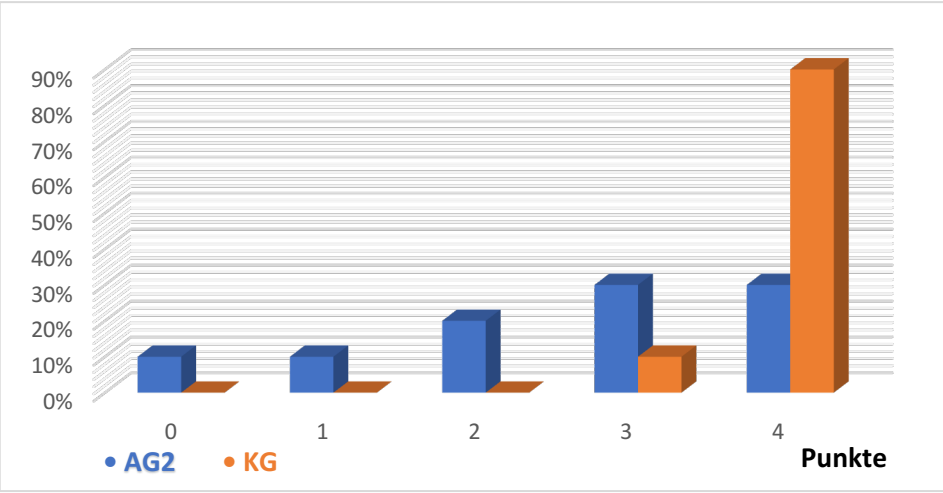


Abb. 8: Leistungsverteilung-Item 6
AG2: Alzheimer-Demenz-Gruppe, KG: Kontrollgruppe

Im Uhren-Zeichen-Test erreichten 70 % der KG-Probanden die volle Punktzahl, 18 % erhielten neun Punkte und der Rest erreichte sieben oder acht Punkte. Die AG2-Probanden zeigten hingegen sehr unterschiedliche Leistungen mit Werten von einem bis zehn Punkten (Abbildung 9).

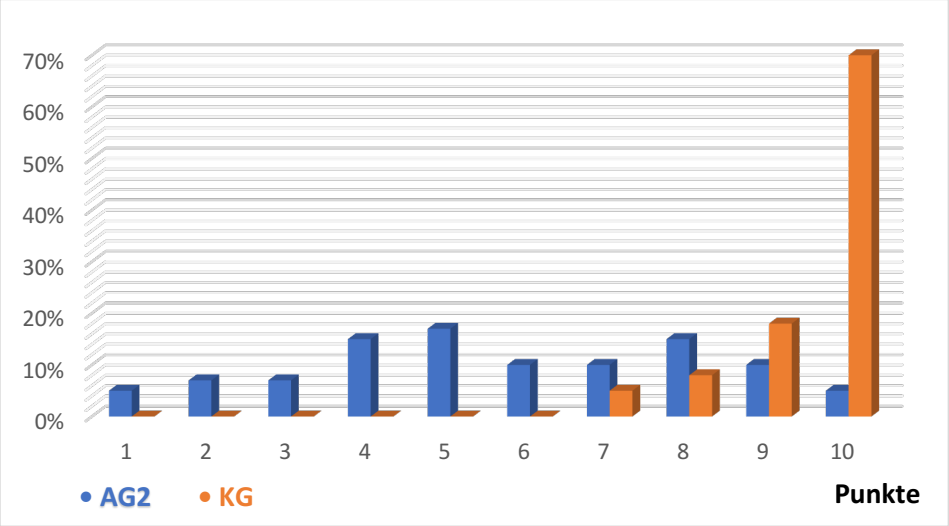


Abb. 9: Leistungsverteilung-Item 7
AG2: Alzheimer-Demenz-Gruppe, KG: Kontrollgruppe

Im Item 8 (Verzögerte Reproduktion) fand sich eine minimale Überlappung zwischen den beiden Gruppen. 95 % der AG2-Probanden erreichten zwischen null und zwei Punkten. Alle KG-Probanden erzielten zwischen zwei und sieben Punkten (Abbildung 10).

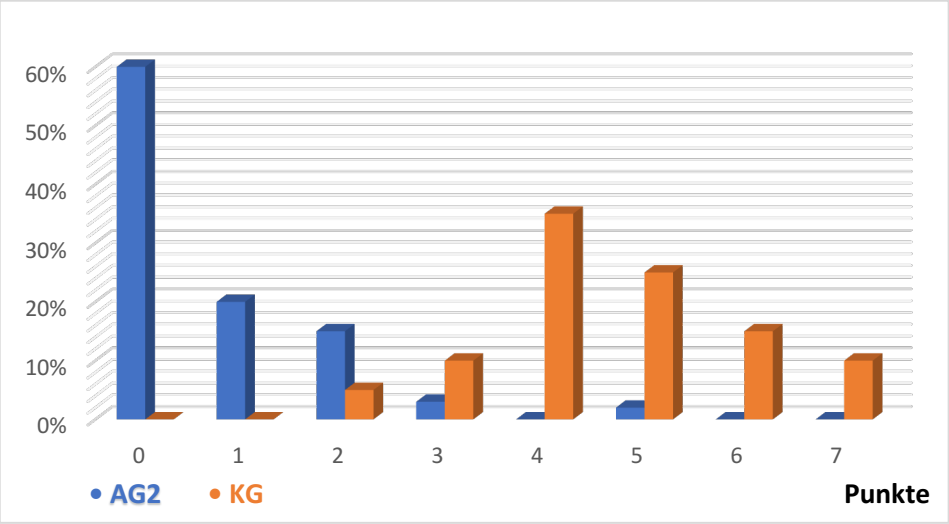


Abb. 10: Leistungsverteilung-Item 8
AG2: Alzheimer-Demenz-Gruppe, KG: Kontrollgruppe

Im Item ‚Wortflüssigkeit–Supermarktaufgabe' konnten 90 % der KG-Probanden die volle Punktzahl erreichen. Die Leistung der AG2-Probanden war in dieser Aufgabe wieder unterschiedlich, mit Werten zwischen drei und zehn Punkten (Abbildung 11).

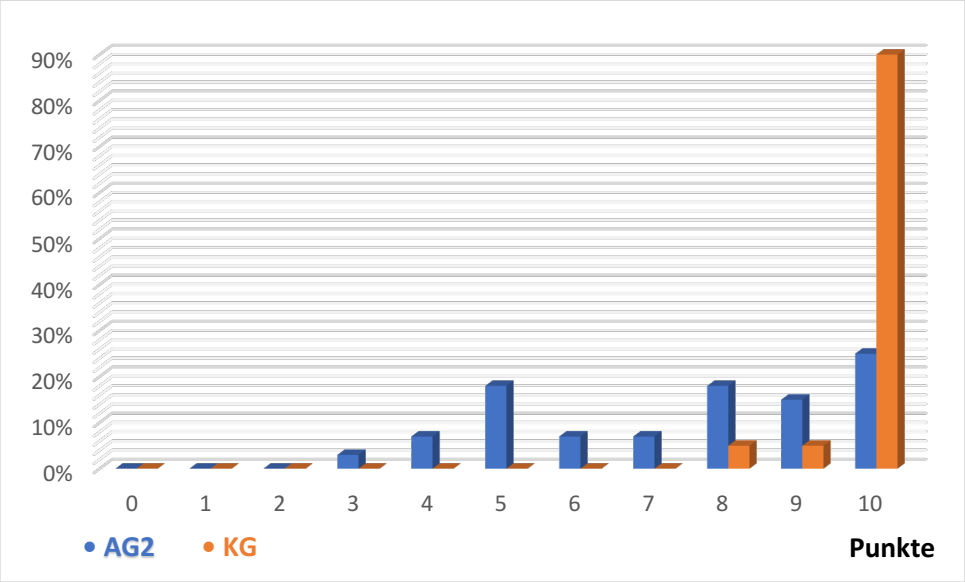


Abb. 11: Leistungsverteilung-Item 9
AG2: Alzheimer-Demenz-Gruppe, KG: Kontrollgruppe

Die Punktverteilungen wiesen also in allen Items deutliche Unterschiede zwischen den beiden Gruppen (AG2, KG) auf. Die Leistungen der AG2-Probanden waren durchgehend schlechter und sind demzufolge im linken, schwächeren Bereich des Leistungsdiagramms abgebildet.

3.5 Interne Konsistenz

Um die Stärke der Item-Reliabilität zu prüfen, wurde als Maß für die interne Konsistenz Cronbachs Alpha berechnet. Für den gesamten Test zeigte sich ein Wert von .893. Alle Items korrelierten r_{it} = .58 oder höher (Tabelle 21).

	Skalenmittelwert, wenn Item weggelassen	Skalenvarianz, wenn Item weggelassen	Korrigierte Item-Skala-Korrelation	Cronbachs Alpha, wenn Item weggelassen
Item 1	30.83	89.637	.806	.867
Item 2	32.35	107.543	.716	.883
Item 3	31.38	105.470	.740	.880
Item 4	33.82	117.543	.702	.896
Item 5	31.55	103.280	.778	.876
Item 6	31.20	110.828	.577	.889
Item 7	27.40	77.475	.804	.876
Item 8	32.37	87.589	.729	.877
Item 9	26.22	90.820	.714	.877

Tabelle 21: Cronbachs Alpha Item-Skala-Statistiken für den Paralleltest

Für den Originaltest wurde ebenfalls Cronbachs Alpha berechnet. Hier zeigte sich für den gesamten Test ein Wert von .890. Alle Items korrelierten r_{it} = .512 oder höher (Tabelle 22).

	Skalenmittelwert, wenn Item weggelassen	Skalenvarianz, wenn Item weggelassen	Korrigierte Item-Skala-Korrelation	Cronbachs Alpha, wenn Item weggelassen
Item 1	30.53	79.242	.799	.864
Item 2	32.34	93.237	.777	.878
Item 3	31.57	90.288	.791	.874
Item 4	33.74	103.245	.512	.896
Item 5	31.57	94.490	.665	.882
Item 6	31.17	96.648	.546	.888
Item 7	27.31	65.852	.840	.868
Item 8	32.65	75.402	.737	.873
Item 9	26.24	76.124	.746	.871

Tabelle 22: Cronbachs Alpha Item-Skala-Statistiken für den Originaltest

3.6 Retest-Reliabilität

Zur Prüfung der Retest-Reliabilität wurden n = 40 AG2-Probanden mit dem Paralleltest zweimal getestet.

Für die Retest-Reliabilität zeigte sich ein Wert von r_s = .985, p < .001, n = 40.

Mittels Wilcoxon-Tests wurden noch die zentralen Tendenzen der Leistungen für die beiden Zeitpunkte verglichen. In der Kontrolluntersuchung war die Leistung der Patienten höher (M = 29.90, SD = 9.018, Median = 32.50) als bei der Erstuntersuchung (M = 29.25, SD = 8.995, Median = 31.50). Der Wilcoxon-Test war signifikant (z = -2.592, p < .01, n = 40).

3.7 Inter-Rater-Reliabilität

Im Rahmen der Prüfung der Inter-Rater-Reliabilität wurden n = 20 AG2-Probanden mit einer Alzheimer-Demenz in einem leichten bis zu einem mittelschweren Stadium (GDS III-V) durch zwei Untersucher getestet.

Die Inter-Rater-Reliabilität betrug r_s = .995, p < .001, n = 20.

3.8 Konstruktvalidität

In der deskriptiven Analyse wurden die zentralen Tendenzen für Parallel- und Originaltest in der gesamten Stichprobe miteinander verglichen (AG2, KG). Hier ergaben sich nahezu identische Werte für beide Versionen des Tests (Originaltest: M = 34.60, SD = 11.114, Median = 35, Paralleltest: M = 34.64, SD = 10.374, Median = 36).

Bei der Prüfung der konvergenten Validität zeigte sich eine signifikante Korrelation zwischen den Punktwerten der gesamten Stichprobe (AG2, KG) im Original- und Paralleltest (r_s = .969, p < .001, n = 100).

Weiterhin wurde die Korrelation des Paralleltests mit der GDS-Stufe nach Reisberg ermittelt. Hier korrelieren die Punktwerte des Paralleltests signifikant mit der GDS-Stufe, r_s = -.958, p < .001, n = 100.

In Tabelle 23 sind die zentralen Tendenzen der erreichten Punktzahl im Paralleltest für jedes GDS-Stadium dargestellt. Deutlich erkennbar ist der Abfall der Mittelwerte bei ansteigendem GDS-Schweregrad.

GDS-Stadium	*n*	*M*	*SD*	Median
I	31	45.58	2.233	46.00
II	9	44.44	3.358	44.00
III	24	35.33	2.390	35.50
IV	14	29.50	3.459	29.50
V	22	17.73	4.641	17.50

Tabelle 23: Deskriptive Analyse der erreichten Punktzahl im Paralleltest für die verschiedenen GDS-Stadien
n: Größe der Stichprobe, M: Mittelwert, SD: Standardabweichung

3.9 Kriteriumsvalidität

In der deskriptiven Analyse wurden die Leistungen von beiden diagnostischen Gruppen (AG2, KG) für den Paralleltest verglichen. Die KG-Probanden wiesen signifikant höhere Scores auf (M = 45.32, SD = 2.526, Median = 46), im Vergleich zu den AG2-Probanden (M = 27.52, SD = 8.605, Median = 30).

Die Leistungen der beiden Gruppen wurden mit dem Mann-Whitney-U-Test geprüft. Die Gruppenunterschiede erwiesen sich dabei als signifikant, U = 6.500, n = 100, p < .001 (Tabelle 24). Die Effektstärke lag bei r = .84.

	AG2	KG	AG2 gegen KG
n	60	40	100
Mittlerer Rang	30.61	80.34	
Rangsumme	1836.50	3213.50	
Mann-Whitney-U			6.500
Z			-8.408
Asymptotische Signifikanz (2-seitig)			.000***

Tabelle 24: Mann-Whitney-U-Test für den Paralleltest
AG2: Alzheimer-Demenz-Gruppe, KG: Kontrollgruppe, n: Größe der Stichprobe.
****p < .001*

Um weiterhin Sensitivität und Spezifität zu ermitteln, wurde eine ROC-Kurve erstellt. Dabei wurde eine Fläche unter der Kurve von .997 ermittelt (Abbildung 12).

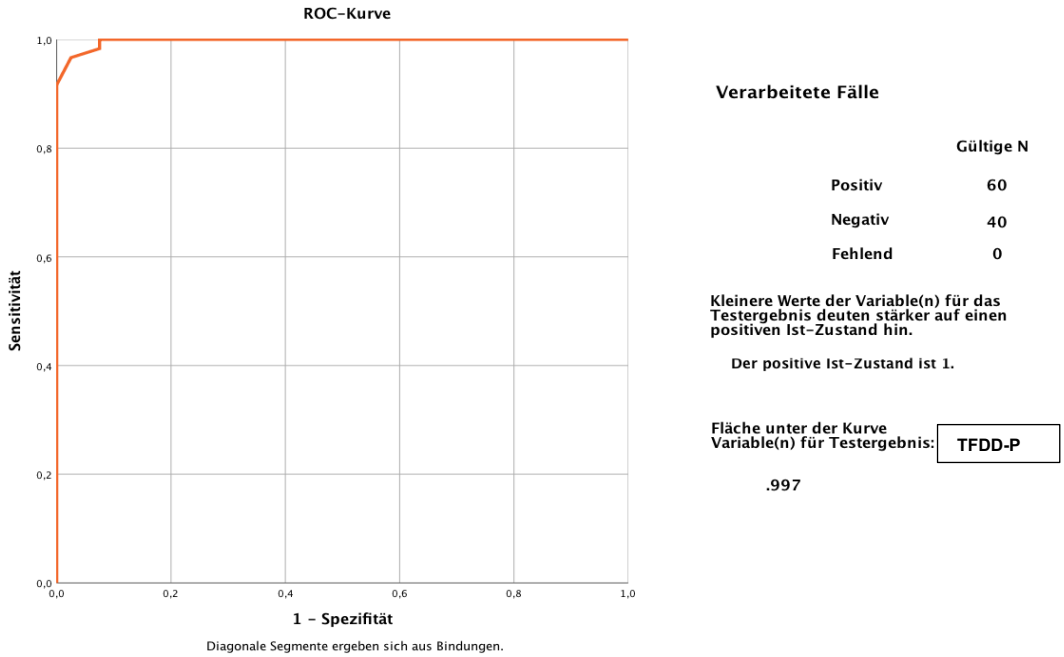


Abb. 12: ROC-Kurve für den Paralleltest bei der gesamten AG2 gegen KG
TFDD-P: Paralleltest

Für den zuvor definierten *Cut-Off*-Wert von > 35 Punkten ergab sich eine Spezifität von 100 % bei einer Sensitivität von 80 %.

Im Rahmen der ROC-Analyse wurde der optimale *Cut-Off*-Wert über den maximalen Youden-Index ermittelt. Bei einem *Cut-Off*-Wert von > 38 Punkten erreichte der Paralleltest eine Sensitivität von 97 % bzw. eine Spezifität von 98 % und somit den maximalen Youden-Index.

Zwei AG2-Probanden erzielten eine Punktzahl von über 38 und wurden somit fälschlicherweise als gesund eingestuft. Von der KG erzielte ein Proband eine Punktzahl von 38 Punkten und wurde fälschlicherweise der AG2 zugeordnet.

Für den Originaltest wurde ebenfalls die Zuordnungsfähigkeit zwischen AG2 und KG ermittelt.

In der deskriptiven Analyse wurden die Leistungen der beiden diagnostischen Gruppen (AG2, KG) für den Originaltest verglichen. Wie in der Analyse des Paralleltests wiesen die KG-Probanden signifikant höhere Scores auf (*M* = 44.78,

SD = 2.178, Median = 45), im Vergleich zu den AG2-Probanden (*M* = 27.82, *SD* = 7.775, Median = 30).

Im Mann-Whitney-U-Test erwiesen sich die Gruppenunterschiede zwischen AG2 und KG als signifikant, *U* = 2.500, *n* = 100, *p* < .001 (Tabelle 25). Die Effektstärke lag bei *r* = .84.

	AG2	KG	AG2 gegen KG
n	60	40	100
Mittlerer Rang	30.54	80.44	
Rangsumme	1832.50	3217.50	
Mann-Whitney-U			2.500
Z			-8.438
Asymptotische Signifikanz (2-seitig)			.000***

Tabelle 25: Mann-Whitney-U-Test für den Originaltest
AG2: Alzheimer-Demenz-Gruppe, KG: Kontrollgruppe, n: Größe der Stichprobe.
****p < .001*

In der ROC-Analyse fand sich für den Originaltest eine Fläche unter der Kurve von .998.

Für den zuvor definierten *Cut-Off*-Wert von > 35 Punkten wurde eine Spezifität von 100 % bei einer Sensitivität von 82 % ermittelt.

Der optimale *Cut-Off*-Wert für den Originaltest lag in der hier untersuchten Stichprobe ebenfalls bei > 38 Punkten. Dabei wurden eine Sensitivität und eine Spezifität von jeweils 98 % und somit der maximale Youden-Index erreicht. Ein Proband mit einer Demenz bei Alzheimer-Krankheit wurde nicht erkannt und eine gesunde Person wurde fälschlicherweise zur AG2 zugeordnet.

3.10 Trennung zwischen kognitiv gesunden Personen und Menschen mit beginnender Demenz bei Alzheimer-Krankheit

In der deskriptiven Analyse wurden für den Paralleltest die Leistungen von *n* = 38 Menschen mit einer beginnenden Demenz bei Alzheimer-Krankheit (GDS III-IV) der AG2 mit den Leistungen der KG-Probanden verglichen. Die KG-Probanden wiesen dabei höhere Scores auf (*M* = 45.32, *SD* = 2.526, Median = 46), im

Vergleich zu den Menschen mit einer beginnenden Demenz bei Alzheimer-Krankheit (*M* = 33.18, *SD* = 3.985, Median = 33.5).

Die Scores der beiden Gruppen wurden mit dem Mann-Whitney-U-Test geprüft. Die Gruppenunterschiede erwiesen sich dabei als signifikant, *U* = 6.500, *n* = 78, *p* < .001 (Tabelle 26). Die Effektstärke lag bei *r* = .85.

	AG2 (GDS III-IV)	KG	AG2 (GDS III-IV) gegen KG
n	38	40	78
Mittlerer Rang	19.67	58.34	
Rangsumme	747.50	2333.50	
Mann-Whitney-U			6.500
Z			-7.551
Asymptotische Signifikanz (2-seitig)			.000***

Tabelle 26: Mann-Whitney-U-Test für den Paralleltest bei leichter Demenz gegen KG
AG2: Alzheimer-Demenz-Gruppe, KG: Kontrollgruppe, n: Größe der Stichprobe.
****p < .001*

In der ROC-Analyse fand sich eine Fläche unter der Kurve von .996. (Abbildung 13).

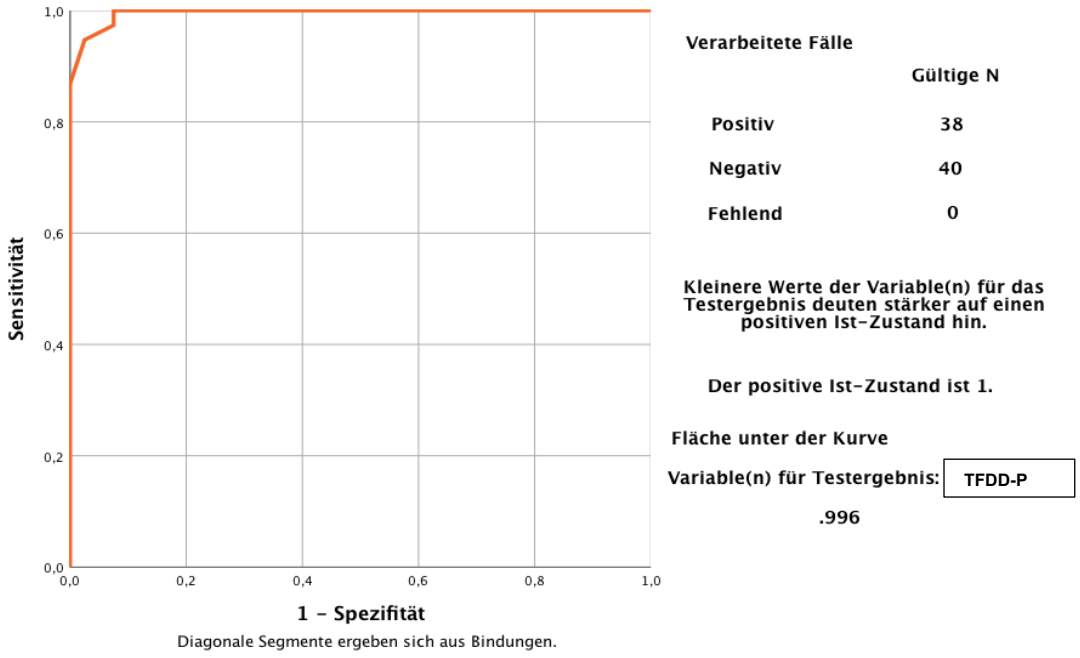


Abb. 13: ROC-Kurve für Paralleltest bei leichter Demenz gegen KG
TFDD-P: Paralleltest

Mit dem für die gesamte Stichprobe zuvor definierten *Cut-Off*-Wert von > 38 Punkten erreichte der Test eine Sensitivität von 95 % bzw. eine Spezifität von 97 % und somit einen Youden-Index von .92.

Zwei Menschen mit einer beginnenden Demenz erreichten eine Punktzahl von über 38 und wurden somit falsch als gesund eingestuft. Ein KG-Proband erzielte eine Punktzahl von 38 Punkten und wurde fälschlicherweise zur Gruppe der beginnenden Demenz zugeordnet.

Mit einem *Cut-Off*-Wert von > 39 Punkten erreichte der Test eine Sensitivität von 97%, eine Spezifität von 93% und einen Youden-Index von .90.

Mit einem *Cut-Off*-Wert von > 37 Punkten erreichte der Test eine 100-prozentige Spezifität bei einer Sensitivität von 87% und einem Youden-Index von .87.

3.11 Trennung zwischen kognitiv gesunden Personen und Menschen mit mittelschwerer Demenz bei Alzheimer-Krankheit

In der deskriptiven Analyse zeigten sich schlechtere Scores für n = 22 Probanden mit mittelschwerer Demenz (GDS V) der AG2 (Median = 17.5, M = 17.73, SD = 4.641) im Vergleich zur KG (Median = 46, M = 45.32, SD = 2.526).

Die Gruppenunterschiede erwiesen sich als signifikant, exakter Mann-Whitney-U-Test: U = .000, n = 62, p < .001 (Tabelle 27). Die Effektstärke lag bei r = .83.

	AG2 (GDS V)	**KG**	**AG2 (GDS V) gegen KG**
n	22	40	78
Mittlerer Rang	11.50	42.50	
Rangsumme	253.00	1700.00	
Mann-Whitney-U			.000
Z			-6.500
Asymptotische Signifikanz (2-seitig)			.000***

Tabelle 27: Mann-Whitney-U-Test für den Paralleltest bei mittelschwerer Demenz gegen KG
AG2: Alzheimer-Demenz-Gruppe, KG: Kontrollgruppe, n: Größe der Stichprobe.
****p < .001*

In der ROC-Analyse fand sich eine Fläche unter der Kurve von 1 und somit erreichte der Paralleltest hier eine 100-prozentige Sensitivität und Spezifität.

4 Diskussion

Das Ziel der vorliegenden Arbeit bestand in der Entwicklung und Validierung einer Parallelversion des Tests zur Früherkennung von Demenzen mit Depressionsabgrenzung (TFDD).

Der TFDD eignet sich aufgrund der guten Testgütekriterien (Velayudhan et al., 2014) sowie der schnellen und kosteneffizienten Durchführung besonders für die Prüfung von kognitiven Funktionen im Rahmen der Frühdiagnostik von dementiellen Erkrankungen (Ihl et al., 2005). Problematisch ist allerdings der Übungs- bzw. Lerneffekt bei Verlaufsmessungen bzw. Wiederholungen der neuropsychologischen Testungen, der insbesondere zu Beginn einer Demenzerkrankung die Testergebnisse beeinflussen kann. Durch die Verwendung von Paralleltests lässt sich der testspezifische Lerneffekt jedoch hinreichend gering halten.

4.1 Parallelität der beiden Testversionen

Für jedes Item des Originaltests wurden neue Aufgaben formuliert, um daraus Itemzwillinge zu bilden. Ziel war es, Parallelitems zu entwickeln, die zwar nicht identisch zu den Originalitems sind, aber dennoch die gleichen kognitiven Domänen untersuchen und die gleichen Messeigenschaften aufweisen. Durchführungsdauer, Schwierigkeit, Punktaufteilung und Auswertung sollten zwischen den beiden Versionen nicht unterscheidbar sein, um die Gleichwertigkeit in der Erhebung zu ermöglichen. Die Parallelitems wurden auf der Grundlage der Theorie und Empirie unter Berücksichtigung der Literatur und der klinischen Erfahrung im Rahmen der Gedächtnissprechstunde der Alexianer Krefeld GmbH, Krankenhaus Maria Hilf entwickelt.

In einer ersten Voruntersuchung wurden die neu entwickelten Items überprüft und mit den Originalitems verglichen, um gegebenenfalls Mängel aufzudecken und diese dann zu überarbeiten.

Original- und Parallelitems wiesen in der deskriptiven Analyse vergleichbare Mittelwerte und nahezu identische Medianwerte auf. Mithilfe des Wilcoxon-Tests konnte für keines der Items ein signifikanter Unterschied zwischen beiden Versionen nachgewiesen werden und der Spearman-Korrelationskoeffizient ergab für

jedes Item eine signifikante Korrelation zwischen Original- und Parallelversion (r_s ≥ .810, p < .001, n = 30). Somit konnte das erste Kriterium erfüllt werden, das bei der Entwicklung der Parallelitems definiert wurde: Grundvoraussetzung für die Parallelität der Items sind gleiche zentrale Tendenzen und eine hohe Korrelation zwischen den beiden Versionen.

Nach der initialen Überprüfung der Parallelitems gab es also keinen Anlass für die Überarbeitung des Tests, sodass weiterhin Daten für die erweiterte statistische Analyse erhoben wurden.

Als nächstes erfolgte die Itemanalyse, um die die Qualität der einzelnen Items und die Dimensionalität des Tests zu überprüfen. Durch die explorative Faktorenanalyse wurde die Teststruktur geprüft und anhand des Scree-Plots und der Faktorenmatrix konnte ein Generalfaktor gefunden werden, mit dem sich 63,3 % der Varianz erklären lässt. Dabei handelt sich um den Faktor Demenz. Die Faktorenanalyse erfolgte auch für den Originaltest, für den ebenfalls ein Generalfaktor gefunden werden konnte, der 61,1 % der Varianz erklärt. In der ursprünglichen Validierungsarbeit betrug dieser Wert 59,9 % (Ihl, 2000). Durch dieses Ergebnis bestätigt sich die Annahme der Eindimensionalität, denn alle Items wurden mit dem Ziel der Früherkennung einer Demenz entwickelt. So gibt es für alle Items des Original- und Paralleltests ein theoretisches Merkmal, das gemessen wird, und dementsprechend war die hier festgestellte Teststruktur mit nur einem Generalfaktor zu erwarten. Die Homogenität des Paralleltests wurde auch im Rahmen der Konsistenzanalyse bestätigt. Die Korrelation jeden Items mit dem Gesamtergebnis des Tests lag bei r_{it} = .58 oder höher, was eine gute Trennschärfe der einzelnen Items belegt. Für den Originaltest wurde auch die interne Konsistenz geprüft und hier fand sich ebenfalls eine sehr gute Trennschärfe jeden einzelnen Items mit einem Trennschärfekoeffizient von r_{it} = .512 oder höher.

Zur weiteren Prüfung der Itemqualität wurde getrennt für Parallel- und Originaltest der Schwierigkeitsindex für mehrstufige Items bei AG2 und KG ermittelt. Dieser wurde für beide Versionen jeden Items berechnet und die beiden Versionen wurden anschließend miteinander verglichen. Der Schwierigkeitsindex jeden Items war für beide Versionen nahezu identisch (Tabelle 20) und somit als positiv für die Annahme der Parallelität der Items zu werten. Entsprechend der guten Trennschärfe der einzelnen Items, die bei der Konsistenzanalyse festgestellt

werden konnte, zeigten sich in der Analyse der Itemschwierigkeit deutliche Unterschiede zwischen AG2 und KG. Es gelang mit jedem Item einen Unterschied im prozentualen Anteil vollständig richtiger Lösungen zwischen den beiden diagnostischen Gruppen aufzuzeigen.

Insgesamt wurde auf Itemebene anhand der Ergebnisse keine Verletzung der Annahme der Parallelität der beiden Testversionen festgestellt, sodass mit der Validierung des Paralleltests fortgefahren werden konnte.

4.2 Validierung des Paralleltests

Um die Qualität des Paralleltests zu beurteilen, wurden die Testgütekriterien geprüft, bzw. wurde ermittelt, welche Aspekte der verschiedenen Gütekriterien in welcher Weise erfüllt sind. Die Hauptgütekriterien der Reliabilität, Objektivität und Validität stehen in einer engen Relation zueinander und bauen aufeinander auf. Die Objektivität bildet die grundsätzliche Voraussetzung für die Reliabilität und beide sind eine Bedingung für die Validität, denn ohne gute Objektivität und Reliabilität kann ein Test keine hohe Validität aufweisen. Andersherum kann ein Test eine sehr hohe Reliabilität und Objektivität haben und dennoch nicht das messen, was er messen soll bzw. für die Diagnostik unbrauchbar sein. Damit befasst sich die Validität und stellt deswegen für die diagnostische Praxis das wesentliche Gütekriterium dar.

Durch die präzisen Anleitungen zur Durchführung und Auswertung des TFDD besteht für den Untersucher kaum Raum für Abweichungen, sodass die Objektivität als erfüllt betrachtet werden kann. Zusätzlich wurde durch die Berechnung der Inter-Rater-Reliabilität ein sehr hoher Spearman-Korrelationskoeffizient zwischen der vergebenen Punktzahl der beiden Untersucher bestätigt. Dabei handelt es sich um das Ausmaß der Übereinstimmung der verschiedenen Untersucher bzw. um ein Objektivitätsmaß.

Um die Verlässlichkeit bzw. Reliabilität der Parallelversion des TFDD zu prüfen, wurden die interne Konsistenz und die Testwiederholungsreliabilität (Retest-Reliabilität) berechnet. Entsprechend der Faktorenanalyse zeigte sich eine sehr gute interne Konsistenz mit einem Cronbachs-Alpha-Wert von .893 für den gesamten Test. Durch das Weglassen des vierten Items (aktuelle Tageszeit) war

eine numerische Verbesserung im Promillebereich des Wertes (Cronbachs Alpha = .896) erzielbar. Das Ziel der vorliegenden Arbeit lag allerdings in der Entwicklung einer Parallelversion mit der gleichen Anzahl von Items, sodass dieses Ergebnis vernachlässigt werden konnte. Darüber hinaus würde das Weglassen eines Items einen Normvergleich nahezu verunmöglichen. Demzufolge wurde davon abgesehen. Der Cronbachs-Alpha-Wert für den Originaltest betrug .890 und war somit mit dem Wert für die Parallelversion vergleichbar.

Bei der Prüfung der Retest-Reliabilität zeigte sich ein Spearman-Korrelationskoeffizient zwischen den Leistungen der beiden Zeitpunkte von $r_s = .985$ ($p < .001$), was auf eine hohe Reliabilität hindeutet. Darüber hinaus wurden die zentralen Tendenzen der Ergebnisse miteinander verglichen. In der Kontrolluntersuchung war die Leistung der Probanden höher ($M = 29.90$, $SD = 9.018$, Median = 32.50) als bei der Erstuntersuchung ($M = 29.25$, $SD = 8.995$, Median = 31.50). Mithilfe des Wilcoxon-Tests konnte ein signifikanter Unterschied in der Leistung der Probanden zwischen den beiden Zeitpunkten nachgewiesen werden ($z = -2.592$, $p < .01$, $n = 40$). Aufgrund des progredienten Charakters der Alzheimer-Demenz wäre normalerweise zu erwarten, dass die Leistungen bei der Erstuntersuchung besser sind, als bei der Kontrolluntersuchung. Grund und Voraussetzung der vorliegenden Arbeit war allerdings die Annahme, dass die Patienten in der Lage sind, durch Übungseffekte bei den Verlaufsuntersuchungen ihre Leistungen zu verbessern. Die Ergebnisse entsprechen dieser Hypothese, sodass die statistisch dargestellten verbesserten Leistungen der Patienten am ehesten auf den Lerneffekt zurückzuführen sind.

Zur Beurteilung der Validität des Paralleltests wurden die drei Kernaspekte der Inhalts-, Konstrukts- und Kriteriumsvalidität geprüft. Die Inhaltsvalidität wird nicht durch ein objektives statistisches Verfahren gemessen, sondern aufgrund von theoretischen und fachlichen Überlegungen bestimmt, sodass sie bereits zum Zeitpunkt der Konstruktion des Tests ansetzt. Dabei geht es um die Frage, ob die neu entwickelten Items den Inhalt des Konzeptes eines Demenz-Screening-Verfahrens adäquat repräsentieren. Zu diesem Zweck wurden bei der Entwicklung des Originaltests durch Fachexperten die kognitiven Domänen festgelegt, die bei einer beginnenden Demenz erfasst werden sollten, wobei zudem ein entsprechender Item-Pool definiert wurde. Daraus wurden ökonomische und sensitive

Items ermittelt und zu einem Demenz-Screening-Verfahren zusammengestellt.

In der vorliegenden Arbeit wurden nach logischen und fachlichen Überlegungen mit entsprechender Literaturrecherche und Auswertung der vorhandenen Fachliteratur Parallelitems ausgewählt bzw. entwickelt. Bei sonst inhaltlich unveränderter Teststruktur ist die Inhaltsvalidität beim Paralleltest ebenso wie beim Originaltest als gegeben zu betrachten.

Zur Prüfung der Konstruktvalidität wurde die Korrelation des Parallel- mit dem Originaltest berechnet, um die konvergente Validität als Teilaspekt der Konstruktvalidität zu prüfen. Mit einem Korrelationskoeffizient von r_s = .969 zeigte sich eine signifikante Korrelation zwischen den Punktwerten im Original- und Paralleltest. Auch deskriptiv wurden die zentralen Tendenzen der beiden Testversionen in der gesamten Stichprobe verglichen (AG2, KG) und nahezu identische Werte für beide Versionen des Tests ermittelt (Originaltest: *M* = 34.60, *SD* = 11.114, Median = 35, Paralleltest: *M* = 34.64, *SD* = 10.374, Median = 36). Durch diese Ergebnisse wird die Annahme unterstützt, dass der Paralleltest genau wie der Originaltest das angestrebte Konstrukt der Demenz bei Alzheimer-Krankheit erfasst. Darüber hinaus wurde die Korrelation des Paralleltests mit der GDS ermittelt und für die beiden Instrumente eine signifikante Korrelation festgestellt (r_s = -.958, p < .001). Die negative Korrelation bedeutet, dass die Leistung bzw. Punktwerte der getesteten Probanden mit zunehmender GDS-Stufe abfallend sind, was auf die Fähigkeit zur Schweregradbestimmung des Paralleltests hinweist. Der Score des Paralleltests stimmt also mit dem Schweregrad der Demenz bzw. der GDS-Stufe überein. Dies konnte auch in der ursprünglichen Validierungsstudie des Originaltests nachgewiesen werden (Ihl et al., 2000). Allerdings wurden in der vorliegenden Arbeit nur Patienten in einem leichten bis zu einem mittelschweren Stadium eingeschlossen (GDS III-V), sodass eine genaue Analyse der Fähigkeit zur Schweregradbestimmung nicht vorgenommen wurde.

Bereits in der Faktorenanalyse wurde die Teststruktur des Paralleltests geprüft, wobei festgestellt werden konnte, dass sich alle Items im Sinne eines eindimensionalen Tests auf einem einzigen Faktor vereinigen lassen. Für eine gute Konstruktvalidität sprechen damit sowohl die konvergente als auch die faktorielle Validität.

Zur Prüfung der Kriteriumsvalidität wurde die Diskriminationsfähigkeit des Parallel- und Originaltests ermittelt und miteinander verglichen. Als Außenkriterium wurde die Diagnose ‚Demenz' bzw. die Zugehörigkeit zur Gruppe AG2 oder KG festgelegt, bevor die Leistung der beiden diagnostischen Gruppen (AG2, KG) mittels Mann-Whitney-U-Test und deskriptiver Analyse verglichen wurde. Anschließend wurden die Sensitivität und die Sensibilität als Ausdruck der Zuordnungsgenauigkeit des Paralleltests im Sinne der diskriminativen Validität berechnet. Das identische Verfahren erfolgte für den Originaltest, um die Diskriminationsfähigkeit der beiden Versionen zu vergleichen.

Im Mann-Whitney-U-Test konnten signifikante Unterschiede in der Leistung zwischen den beiden diagnostischen Gruppen (AG2, KG) für den Paralleltest nachgewiesen werden (U = 6.500, n = 100, p < .001). Die Effektstärke lag bei r = .84, was einem starken Effekt entspricht.

In der deskriptiven Analyse hatten die KG-Probanden zwar deutlich bessere Leistungen als die AG2-Probanden, es fanden sich allerdings Überschneidungen zwischen den Testleistungen der beiden diagnostischen Gruppen. Mit dem zuvor beim Originaltest definierten *Cut-Off*-Wert von > 35 Punkten wurden zwölf AG2-Probanden fälschlicherweise als gesund eingestuft, da sie eine Leistung von über 35 Punkten hatten. Alle KG-Probanden wurden dagegen korrekt zugeordnet. Der Paralleltest erreichte mit diesem *Cut-Off*-Wert eine 100-prozentige Spezifität, bei allerdings einer Sensitivität von nur 82 %, und maß somit der Spezifität eine höhere Bedeutung zu. Hier wurde der Sensitivität der Vorzug gegeben, um möglichst alle Kranken zu erkennen, sodass anhand der ROC-Kurve und des maximalen Youden-Index-Wertes der optimale *Cut-Off*-Wert ermittelt wurde. Dieser lag bei 38/39. Bei einem *Cut-Off*-Wert von > 38 Punkten erreichte der Test eine Sensitivität von 97 % bei einer Spezifität von 98 %. Der Youden-Index lag damit bei .95.

Der Originaltest wurde ebenfalls im Hinblick auf seine Diskriminationsfähigkeit geprüft. Im Mann-Whitney-U-Test konnten signifikante Unterschiede zwischen AG2 und KG festgestellt werden und in der ROC-Analyse erreichte der Test mit einer Fläche von .998 eine nahezu ideale Spezifität und Sensitivität. Allerdings musste anhand der ROC-Kurve und des Youden-Indexes auch beim Originaltest der *Cut-Off*-Wert höher angesetzt werden. Bei einem *Cut-Off*-Wert von

> 38 Punkten erreichte auch der Originaltest einen Youden-Index-Wert von .96 bei einer Spezifität und Sensitivität von jeweils 98 %. Ein höherer Youden-Index bei somatischen oder testpsychologischen Untersuchungen zur Diagnose der Demenz wurde in der Literatur nicht gefunden.

Eine mögliche Erklärung für den höheren *Cut-Off*-Wert liegt in der Zusammensetzung der Stichprobe. In der vorliegenden Arbeit wurden viele Menschen mit einer Alzheimer-Krankheit rekrutiert, die nur geringe kognitive Defizite aufwiesen. Diese Probanden befanden sich zum Zeitpunkt der Untersuchung im klinischen Stadium einer beginnenden Demenz und die Diagnose der wahrscheinlichen Alzheimer-Krankheit wurde anhand der NIA-AA-Kriterien nach dem Einbeziehen von Alzheimer-Biomarkern (Hirn-Amyloid- und Neurodegenerationsmarker) gestellt. Neuropsychologische Testungen weisen häufig bei Patienten mit leichten kognitiven Beeinträchtigungen eine niedrigere Sensitivität und Spezifität auf (Arevalo-Rodriguez et al., 2015; Weissberger et al., 2017). Diese Patienten weisen in der Regel in den Testungen bessere Leistungen als die fortgeschrittenen Demenzpatienten auf und erreichen erwartungsgemäß höhere Punktzahlen. Dies führt mit einer höheren Wahrscheinlichkeit zu Überschneidungen mit den Leistungen der kognitiv gesunden Personen und dementsprechend zu einer niedrigeren Sensitivität und Spezifität. Abhängig vom *Cut-Off*-Wert gibt es Patienten mit einer beginnenden Demenz, die falsch als gesund klassifiziert werden, und andersherum gesunde Personen, die falsch als dement identifiziert werden.

Sowohl für die Parallelversion des TFDD als auch für den Originaltest war die durchschnittliche Leistung der Stichprobe in der vorliegenden Arbeit höher als jene von der originalen Validierungsstudie, sodass letztlich, um den maximalen Youden-Index-Wert sowie die optimale Spezifität und Sensitivität zu erreichen, ein höherer *Cut-Off*-Wert angesetzt werden musste.

Der TFDD wurde als Screening-Verfahren zur Früherkennung beginnender Demenzen entwickelt. Um die Zuordnungsfähigkeit des Paralleltests weiter zu prüfen, wurden separat nur die Leistungen der Menschen mit einer beginnenden Demenz (GDS III-IV, $n = 38$) mit diesen der KG verglichen. Auch hier konnten signifikante Unterschiede im Mann-Whitney-U-Test bezüglich der Leistung der beiden Gruppen beobachtet werden. Dabei handelte es sich um einen starken Effekt mit einer Effektstärke von $r = .85$. Deskriptiv ließen sich ebenfalls

erhebliche Leistungsunterschiede erkennen, und mit dem *Cut-off*-Wert von > 38 Punkten erreichte der Test eine Sensitivität von 95 % bzw. eine Spezifität von 97 % und somit einen maximalen Youden-Index von .90. So konnte auch in dieser separaten statistischen Analyse nachgewiesen werden, dass der Paralleltest als Demenz-Screening-Verfahren gut geeignet ist, da er eine sehr hohe Zuordnungsfähigkeit zwischen Menschen mit einer beginnenden Demenz und gesunden Personen aufweist.

In der erweiterten Analyse der Zuordnungsfähigkeit des Paralleltests zwischen Menschen mit einer mittelschweren Demenz (GDS V) und gesunden Personen zeigten sich erwartungsgemäß optimale Ergebnisse bezüglich deskriptiver Analyse, Gruppenunterschiede sowie ermittelter Sensitivität und Spezifität.

Zusammenfassend hat sich der Paralleltest genauso wie der Originaltest basierend auf der hier erfolgten Prüfung der Hauptgütekriterien als objektiv, reliabel und valide erwiesen. Die beiden Versionen des Tests unterscheiden sich kaum in ihrer Güte.

Auch die Nebengütekriterien der Testökonomie, Nützlichkeit und Akzeptanz sind aufgrund der kurzen Durchführungsdauer, Einfachheit und hohen Relevanz in der alltäglichen diagnostischen Praxis als erfüllt zu betrachten.

4.3 Studieneinschränkungen

Insgesamt wird anhand der Ergebnisse deutlich, dass die Parallelversion des TFDD ähnlich wie die Originalversion ein einfaches, schnelles und gleichzeitig verlässliches Instrument für die neuropsychologische Abklärung und Frühdiagnose einer Demenz bei Alzheimer-Krankheit darstellt.

Allerdings müssen einige Einschränkungen der vorliegenden Studie berücksichtigt werden, die die Interpretation der Ergebnisse beeinträchtigen.

Der TFDD wurde gezielt als Screening-Instrument für die Anwendung durch niedergelassene Ärzte und insbesondere in der Hausarztpraxis entwickelt. Die diagnostischen Gruppen der Studie inkludierten allerdings Menschen mit einer Demenz bei Alzheimer-Krankheit, die durch den Hausarzt bzw. den niedergelassenen Neurologen oder Psychiater zur Gedächtnissprechstunde der Alexianer

Krefeld GmbH überwiesen wurden. Diese Menschen waren zum Zeitpunkt der Datenerhebung bereits an das Demenznetzwerk Krefeld angebunden und profitierten daher von einer frühestmöglichen und umfassenden Versorgung. Aufgrund dessen sind die Ergebnisse der vorliegenden Arbeit nicht bedingungslos auf die hausärztliche Praxis anwendbar. Eine weitere Untersuchung zur Durchführung des Tests in diesem Bereich könnte hierüber Aufschluss geben.

Darüber hinaus wurden in die Studie nur Menschen mit einer Alzheimer-Krankheit einbezogen. Die Alzheimer-Krankheit ist mit Abstand die häufigste Demenzursache und dementsprechend litten die in der Gedächtnissprechstunde diagnostizierten und behandelten Demenzerkrankten bis auf wenige Ausnahmen an der Alzheimer-Krankheit. Daher wurden aus pragmatischen Gründen nur Menschen mit einer Demenz bei Alzheimer-Krankheit untersucht und demzufolge kann die Zuordnungsfähigkeit des Paralleltests bezüglich anderer Demenzerkrankungen durch die vorliegende Arbeit nicht beurteilt werden. Wie hoch die Diskriminationsfähigkeit des Paralleltest bei anderen Demenzerkrankungen ist, müsste in separaten Studien untersucht werden.

Eine weitere Einschränkung der Studie stellt die relativ kleine Stichprobe dar. Für die Voruntersuchung wurden 30 Menschen mit einer Demenz bei Alzheimer-Krankheit (AG1) untersucht, wobei in der anschließenden Validierungsstudie 60 Menschen mit einer Demenz bei Alzheimer-Krankheit (AG2) und 40 gesunde Kontrollpersonen (KG) getestet wurden. Es handelt sich zwar um eine Zufallsstichprobe, dennoch kann sich der relativ kleine Stichprobenumfang nachteilig auf die Repräsentativität auswirken.

Des Weiteren gaben neun Personen der Kontrollgruppe nach der Datenerhebung zum späteren Zeitpunkt eine subjektive Verschlechterung der mnestischen Funktionen an, diese konnte allerdings im Rahmen von einem ausführlichen *Interview* sowie nach einer durchgeführten CERAD-Testbatterie nicht objektiviert werden.

Schließlich wurden in der Studie nicht alle demographischen Variablen erfasst. Neben dem Alter kann der Bildungsgrad Einfluss auf die Ergebnisse eines neuropsychologischen Tests nehmen. So ist z.B. zu erwarten, dass das schulische Bildungsniveau einen Effekt auf den Umfang des Wortschatzes und das verbale Denken haben kann (Schmitz, 2004). Eine weitere Untersuchung unter

Berücksichtigung aller relevanten demographischen Kovarianten (Alter, Beruf, Bildungsgrad, Ausbildung) musste noch ergänzt werden.

4.4 Ausblick

Die Frühdiagnostik, Behandlung und Prävention der Demenz bei Alzheimer-Krankheit stellt aufgrund der demographischen Entwicklung und der epidemiologischen Daten weltweit eine große Herausforderung von herausragender Bedeutung dar. Da die Alzheimer-Krankheit durch die progredienten kognitiven Einschränkungen gekennzeichnet wird, spielen neuropsychologische Testungen nach wie vor eine zentrale Rolle zur Erfassung der Kernsymptome der Krankheit und somit zur Diagnosestellung. Mit der bereits in der ursprünglichen Validierungsstudie nachgewiesenen Testgüte erbringt der TFDD überlegene Validierungsparameter im Sinne von Testgütekriterien gegenüber anderen Demenz-Screening-Verfahren.

Mit der vorliegenden Arbeit ist es gelungen, eine Parallelversion des TFDD zu entwickeln, die in allen Aspekten und ohne Einschränkung wie der Originaltest eingesetzt werden kann. Der Paralleltest beinhaltet genau die gleiche Anzahl von Items für die identischen kognitiven Domänen mit der exakt gleichen Durchführungsdauer und Schwierigkeit. Alle wesentlichen Testgütekriterien wurden erfolgreich geprüft. Trotz Studieneinschränkungen konnte gezeigt werden, dass mit der Parallelversion des TFDD genau wie mit dem Originaltest eine Demenz bei Alzheimer-Krankheit bereits in einem sehr frühen Stadium der Erkrankung erkannt werden kann, noch bevor überhaupt erhebliche Beeinträchtigungen der alltäglichen Funktionen auftreten. Durch seine einfache Handhabung und zeitökonomische sowie kosteneffiziente Anwendbarkeit eignet sich der Test besonders gut für die alltägliche ärztliche Praxis.

Somit steht eine Parallelversion des TFDD zur Verfügung, durch die Verlaufsmessungen bzw. Testwiederholungen ohne Qualitätseinbußen gegenüber dem Originaltest ermöglicht werden. Dabei entfällt der Faktor eines Lerneffekts bei der Lösung von Aufgaben, der bei einem singulären Test durch Wiedervorlage bekannter Items entstehen kann.

5 Literaturverzeichnis

Aarsland, D., Zaccai, J., & Brayne, C. (2005). A systematic review of prevalence studies of dementia in Parkinson's disease. *Movement Disorders*, *20*(10), 1255–1263. https://doi.org/10.1002/mds.20527

Aga, V. M. (2019). When and How to Treat Agitation in Alzheimer's Disease Dementia with Citalopram and Escitalopram. *The American Journal of Geriatric Psychiatry*, *27*(10), 1099–1107. https://doi.org/10.1016/j.jagp.2019.04.016

Ahmed, S., Brennan, L., Eppig, J., Price, C. C., Lamar, M., Delano-Wood, L., Bangen, K. J., Edmonds, E. C., Clark, L., Nation, D. A., Jak, A., Au, R., Swenson, R., Bondi, M. W., & Libon, D. J. (2016). Visuoconstructional Impairment in Subtypes of Mild Cognitive Impairment. *Applied Neuropsychology. Adult*, *23*(1), 43–52. https://doi.org/10.1080/23279095.2014.1003067

Alasmari, F., Alshammari, M. A., Alasmari, A. F., Alanazi, W. A., & Alhazzani, K. (2018). Neuroinflammatory Cytokines Induce Amyloid Beta Neurotoxicity through Modulating Amyloid Precursor Protein Levels/Metabolism. *BioMed Research International*, *2018*, 1–8. https://doi.org/10.1155/2018/3087475

Albert, M. S., DeKosky, S. T., Dickson, D., Dubois, B., Feldman, H. H., Fox, N. C., Gamst, A., Holtzman, D. M., Jagust, W. J., Petersen, R. C., Snyder, P. J., Carrillo, M. C., Thies, B., & Phelps, C. H. (2011). The diagnosis of mild cognitive impairment due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association*, *7*(3), 270–279. https://doi.org/10.1016/j.jalz.2011.03.008

Alzheimer, A. (1906). Über einen eigenartigen schweren Erkrankungsprozeß der Hirnrinde. Neurologisches *Centralblatt*, 25: 1129-1136.

American Psychiatric Association & DSM-5 Task Force. (2013). *Diagnostic and statistical manual of mental disorders: DSM-5* (5th ed). American

Psychiatric Association.

Arbeitsgemeinschaft Wissenschaftlich Medizinischer Fachgesellschaften (AWMF). (2016). S3- Leitlinie „Demenzen“. AWMF-Register-Nr.: 038-013 www.awmf.org. Zugriff am 10.01.2020

Arevalo-Rodriguez, I., Smailagic, N., Roqué i Figuls, M., Ciapponi, A., Sanchez-Perez, E., Giannakou, A., Pedraza, O. L., Bonfill Cosp, X., & Cullum, S. (2015). Mini-Mental State Examination (MMSE) for the detection of Alzheimer’s disease and other dementias in people with mild cognitive impairment (MCI). *Cochrane Database of Systematic Reviews*. https://doi.org/10.1002/14651858.CD010783.pub2

Atri, A., Frölich, L., Ballard, C., Tariot, P. N., Molinuevo, J. L., Boneva, N., Windfeld, K., Raket, L. L., & Cummings, J. L. (2018). Effect of Idalopirdine as Adjunct to Cholinesterase Inhibitors on Change in Cognition in Patients with Alzheimer Disease: Three Randomized Clinical Trials. *JAMA*, *319*(2), 130–142. https://doi.org/10.1001/jama.2017.20373

Bature, F., Guinn, B., Pang, D., & Pappas, Y. (2017). Signs and symptoms preceding the diagnosis of Alzheimer’s disease: A systematic scoping review of literature from 1937 to 2016. *BMJ Open*, *7*(8), e015746. https://doi.org/10.1136/bmjopen-2016-015746

Bordet, R., Ihl, R., Korczyn, A. D., Lanza, G., Jansa, J., Hoerr, R., & Guekht, A. (2017). Towards the concept of disease-modifier in post-stroke or vascular cognitive impairment: A consensus report. *BMC Medicine*, *15*(1), 107. https://doi.org/10.1186/s12916-017-0869-6

Bortz, J., & Döring, N. (2006). *Forschungsmethoden und Evaluation: Für Human- und Sozialwissenschaftler ; mit 87 Tabellen* (4., überarb. Aufl., [Nachdr.]). Springer.

Bühner, M. (2021). *Einführung in die Test- und Fragebogenkonstruktion* (4., korrigierte und erweiterte Auflage). Pearson.

Cannon-Albright, L. A., Foster, N. L., Schliep, K., Farnham, J. M., Teerlink, C. C., Kaddas, H., Tschanz, J., Corcoran, C., & Kauwe, J. S. K. (2019). Relative risk for Alzheimer disease based on complete family history.

Neurology, *92*(15), e1745–e1753. https://doi.org/10.1212/WNL.0000000000007231

Carbone, I., Lazzarotto, T., Ianni, M., Porcellini, E., Forti, P., Masliah, E., Gabrielli, L., & Licastro, F. (2014). Herpes virus in Alzheimer's disease: Relation to progression of the disease. *Neurobiology of Aging*, *35*(1), 122–129. https://doi.org/10.1016/j.neurobiolaging.2013.06.024

Chaves, M. L. F., & Camozzato, A. L. (2007). How many items from a word list can Alzheimer's disease patients and normal controls recall? Do they recall in a similar way? *Dementia & Neuropsychologia*, *1*(1), 52–58. https://doi.org/10.1590/S1980-57642008DN10100009

Ciesielska, N., Sokołowski, R., Mazur, E., Podhorecka, M., Polak-Szabela, A., & Kędziora-Kornatowska, K. (2016). Is the Montreal Cognitive Assessment (MoCA) test better suited than the Mini-Mental State Examination (MMSE) in mild cognitive impairment (MCI) detection among people aged over 60? Meta-analysis. *Psychiatria Polska*, *50*(5), 1039–1052. https://doi.org/10.12740/PP/45368

Clark, L. J., Gatz, M., Zheng, L., Chen, Y.-L., McCleary, C., & Mack, W. J. (2009). Longitudinal Verbal Fluency in Normal Aging, Preclinical, and Prevalent Alzheimer's Disease. *American Journal of Alzheimer's Disease & Other Dementiasr*, *24*(6), 461–468. https://doi.org/10.1177/1533317509345154

Collie, A., Maruff, P., Makdissi, M., McStephen, M., Darby, D. G., & McCrory, P. (2004). Statistical procedures for determining the extent of cognitive change following concussion. *British Journal of Sports Medicine*, *38*(3), 273–278. https://doi.org/10.1136/bjsm.2003.000293

Cummings, J., Ritter, A., & Zhong, K. (2018). Clinical Trials for Disease-Modifying Therapies in Alzheimer's Disease: A Primer, Lessons Learned, and a Blueprint for the Future1. *Journal of Alzheimer's Disease*, *64*(s1), S3–S22. https://doi.org/10.3233/JAD-179901

Dilling, H., & Weltgesundheitsorganisation (Hrsg.). (2015). *Internationale Klassifikation psychischer Störungen: ICD–10 Kapitel V (F) klinisch–*

diagnostische Leitlinien (10. Aufl., unter Berücksichtigung der Änderungen entsprechend ICD-10-GM 2015). Hogrefe.

Dominy, S. S., Lynch, C., Ermini, F., Benedyk, M., Marczyk, A., Konradi, A., Nguyen, M., Haditsch, U., Raha, D., Griffin, C., Holsinger, L. J., Arastu-Kapur, S., Kaba, S., Lee, A., Ryder, M. I., Potempa, B., Mydel, P., Hellvard, A., Adamowicz, K., … Potempa, J. (2019). Porphyromonas gingivalis in Alzheimer's disease brains: Evidence for disease causation and treatment with small-molecule inhibitors. *Science Advances*, *5*(1), eaau3333. https://doi.org/10.1126/sciadv.aau3333

Döring, N., & Bortz, J. (2016). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften*. Springer. https://doi.org/10.1007/978-3-642-41089-5

Eimer, W. A., Vijaya Kumar, D. K., Navalpur Shanmugam, N. K., Rodriguez, A. S., Mitchell, T., Washicosky, K. J., György, B., Breakefield, X. O., Tanzi, R. E., & Moir, R. D. (2018). Alzheimer's Disease-Associated β-Amyloid Is Rapidly Seeded by Herpesviridae to Protect against Brain Infection. *Neuron*, *99*(1), 56-63.e3. https://doi.org/10.1016/j.neuron.2018.06.030

Eknoyan, D., Hurley, R. A., & Taber, K. H. (2012). The clock drawing task: Common errors and functional neuroanatomy. *The Journal of Neuropsychiatry and Clinical Neurosciences*, *24*(3), 260–265. https://doi.org/10.1176/appi.neuropsych.12070180

Erzigkeit, H. (1977). *Der Syndrom-Kurztest zur Erfassung von Aufmerksamkeits- und Gedächtnisstörungen: Vorläufiges Manual*. Vless.

Falleti, M. G., Maruff, P., Collie, A., & Darby, D. G. (2006). Practice effects associated with the repeated assessment of cognitive function using the CogState battery at 10-minute, one week and one month test-retest intervals. *Journal of Clinical and Experimental Neuropsychology*, *28*(7), 1095–1112. https://doi.org/10.1080/13803390500205718

Farlow, M. R., & Shamliyan, T. A. (2017). Benefits and harms of atypical antipsychotics for agitation in adults with dementia. *European Neuropsychopharmacology: The Journal of the European College of Neuropsychopharmacology*, *27*(3), 217–231.

https://doi.org/10.1016/j.euroneuro.2017.01.002

Finkel, S. I., Costa e Silva, J., Cohen, G., Miller, S., & Sartorius, N. (1997). Behavioral and Psychological Signs and Symptoms of Dementia: A Consensus Statement on Current Knowledge and Implications for Research and Treatment. *International Psychogeriatrics*, *8*(S3), 497–500. https://doi.org/10.1017/S1041610297003943

Fleischman, D. A., & Gabrieli, J. (1999). Long-term memory in Alzheimer's disease. *Current Opinion in Neurobiology*, *9*(2), 240–244. https://doi.org/10.1016/s0959-4388(99)80034-8

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state". *Journal of Psychiatric Research*, *12*(3), 189–198. https://doi.org/10.1016/0022-3956(75)90026-6

Förstl, H. (Hrsg.). (2011). *Demenzen in Theorie und Praxis: Mit 52 Tabellen* (3., aktualisierte und überarb. Aufl). Springer.

Freedman, M. (Hrsg.). (1994). *Clock drawing: A neuropsychological analysis*. Oxford University Press.

Georges, J., Miller, O., & Bintener, C. (2020). *Estimating the prevalence of dementia in Europe*. https://doi.org/10.13140/RG.2.2.16880.81923

Gill, S. S., & Seitz, D. P. (2007). Association of antipsychotics with mortality among elderly patients with dementia. *The American Journal of Geriatric Psychiatry: Official Journal of the American Association for Geriatric Psychiatry*, *15*(11), 983–984; author reply 984-985. https://doi.org/10.1097/JGP.0b013e31814622a5

Glenner, G. G., & Wong, C. W. (1984). Alzheimer's disease and Down's syndrome: Sharing of a unique cerebrovascular amyloid fibril protein. *Biochemical and Biophysical Research Communications*, *122*(3), 1131–1135. https://doi.org/10.1016/0006-291x(84)91209-9

Goldberg, T. E., Harvey, P. D., Wesnes, K. A., Snyder, P. J., & Schneider, L. S. (2015). Practice effects due to serial cognitive assessment: Implications for preclinical Alzheimer's disease randomized controlled trials. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*,

1(1), 103–111. https://doi.org/10.1016/j.dadm.2014.11.003

Goschke, Th., & Dreisbach, G. (2011). Kognitiv-affektive Neurowissenschaft: Emotionale Modulation des Erinnerns, Entscheidens und Handelns. In H.-U. Wittchen & J. Hoyer (Hrsg.), *Klinische Psychologie & Psychotherapie* (S. 129–168). Springer. https://doi.org/10.1007/978-3-642-13018-2_6

Grass-Kapanke, B., Brieber, S., Pentzek, M., & Ihl, R. (2005). Der TFDD: Test zur Früherkennung von Demenzen mit Depressionsabgrenzung. *Zeitschrift für Gerontopsychologie & -psychiatrie*, *18*(3), 155–167. https://doi.org/10.1024/1011-6877.18.3.155

Griffin, W. S., Stanley, L. C., Ling, C., White, L., MacLeod, V., Perrot, L. J., White, C. L., & Araoz, C. (1989). Brain interleukin 1 and S-100 immunoreactivity are elevated in Down syndrome and Alzheimer disease. *Proceedings of the National Academy of Sciences of the United States of America*, *86*(19), 7611–7615. https://doi.org/10.1073/pnas.86.19.7611

Hanseeuw, B. J., Scott, M. R., Sikkes, S. A. M., Properzi, M., Gatchel, J. R., Salmon, E., Marshall, G. A., Vannini, P., & Alzheimer's Disease Neuroimaging Initiative. (2020). Evolution of anosognosia in alzheimer's disease and its relationship to amyloid. *Annals of Neurology*, *87*(2), 267–280. https://doi.org/10.1002/ana.25649

Harris, S. A., & Harris, E. A. (2015). Herpes Simplex Virus Type 1 and Other Pathogens are Key Causative Factors in Sporadic Alzheimer's Disease. *Journal of Alzheimer's Disease: JAD*, *48*(2), 319–353. https://doi.org/10.3233/JAD-142853

Hennawy, M., Sabovich, S., Liu, C. S., Herrmann, N., & Lanctôt, K. L. (2019). Sleep and Attention in Alzheimer's Disease. *The Yale Journal of Biology and Medicine*, *92*(1), 53–61.

Henry, L. A., Messer, D. J., & Nash, G. (2015). Executive functioning and verbal fluency in children with language difficulties. *Learning and Instruction*, *39*, 137–147. https://doi.org/10.1016/j.learninstruc.2015.06.001

Hepburn, K., Lewis, M., Tornatore, J., Sherman, C. W., & Bremer, K. L. (2007). The Savvy Caregiver program: The demonstrated effectiveness of a

transportable dementia caregiver psychoeducation program. *Journal of Gerontological Nursing*, *33*(3), 30–36. https://doi.org/10.3928/00989134-20070301-06

Hessmann, P., Zeidler, J., Stahmeyer, J., Eberhard, S., Vogelgsang, J., Abdel-Hamid, M., Wolff-Menzler, C., Wiltfang, J., & Kis, B. (2019). Claims Data Analysis on the Dispensing of Tricyclic Antidepressants Among Patients With Dementia in Germany. *Frontiers in Pharmacology*, *10*, 841. https://doi.org/10.3389/fphar.2019.00841

Heyn, P., Abreu, B. C., & Ottenbacher, K. J. (2004). The effects of exercise training on elderly persons with cognitive impairment and dementia: A meta-analysis. *Archives of Physical Medicine and Rehabilitation*, *85*(10), 1694–1704. https://doi.org/10.1016/j.apmr.2004.03.019

IBM Corp. Released 2017. IBM SPSS Statistics for Windows, Version 25.0. Armonk, NY: IBM Corp. (o. J.).

Ihl, R., Biesenbach A., Brieber S., Grass-Kapanke B., Salamon T. (2005). A head-to-head comparison of the sensitivity of two Screening tests for dementia Mini-Mental-State-Examination (MMSE) and the Test for the Early Detection of Dementia with Discrimination from Depression (TE4D). Psychogeriatria Polska 2, 263–272.

Ihl R., Frölich L (1991) Die Reisberg-Skalen (GDS, BCRS, FAST) (dt.). Beltz, Weinheim

Ihl R., Frölich, L., Winblad, B., Schneider, L., Burns, A., Möller, H.-J., & WFSBP Task Force on Treatment Guidelines for Alzheimer's Disease and other Dementias. (2011). World Federation of Societies of Biological Psychiatry (WFSBP) Guidelines for the Biological Treatment of Alzheimer's disease and other dementias. *The World Journal of Biological Psychiatry*, *12*(1), 2–32. https://doi.org/10.3109/15622975.2010.538083

Ihl, R. & Grass-Kapanke, B. (2000). Manual - Test zur Früherkennung von Demenzen mit Depressionsabgrenzung. Libri Books.

Ihl, R., Grass-Kapanke, B., Lahrem, P., Brinkmeyer, J., Fischer, S., Gaab, N., & Kaupmannsennecke, C. (2000). Entwicklung und Validierung eines Tests

zur Früherkennung der Demenz mit Depressionsabgrenzung (TFDD). *Fortschritte der Neurologie · Psychiatrie*, *68*(9), 413–422. https://doi.org/10.1055/s-2000-11799

Ihl, R., & Müller, W. E. (2020). Weiterbildungs-Curriculum Psychopharmakologie/Pharmakotherapie, Folge 14: Antidementiva. *Psychopharmakotherapie, Arzneimitteltherapie psychischer und neurologischer Errkrankungen*, 27, 247-254.

Jamieson, G. A., Maitland, N. J., Wilcock, G. K., Craske, J., & Itzhaki, R. F. (1991). Latent herpes simplex virus type 1 in normal and Alzheimer's disease brains. *Journal of Medical Virology*, *33*(4), 224–227. https://doi.org/10.1002/jmv.1890330403

Jay, T. R., Miller, C. M., Cheng, P. J., Graham, L. C., Bemiller, S., Broihier, M. L., Xu, G., Margevicius, D., Karlo, J. C., Sousa, G. L., Cotleur, A. C., Butovsky, O., Bekris, L., Staugaitis, S. M., Leverenz, J. B., Pimplikar, S. W., Landreth, G. E., Howell, G. R., Ransohoff, R. M., & Lamb, B. T. (2015). TREM2 deficiency eliminates TREM2+ inflammatory macrophages and ameliorates pathology in Alzheimer's disease mouse models. The Journal of Experimental Medicine, 212(3), 287–295. https://doi.org/10.1084/jem.20142322

Jellinger, K. A., & Attems, J. (2007). Neuropathological evaluation of mixed dementia. Journal of the Neurological Sciences, 257(1–2), 80–87. https://doi.org/10.1016/j.jns.2007.01.045

Johnen, A., Frommeyer, J., Modes, F., Wiendl, H., Duning, T., & Lohmann, H. (2016). Dementia Apraxia Test (DATE): A Brief Tool to Differentiate Behavioral Variant Frontotemporal Dementia from Alzheimer's Dementia Based on Apraxia Profiles. *Journal of Alzheimer's Disease: JAD*, *49*(3), 593–605. https://doi.org/10.3233/JAD-150447

Johnson, D. S., Li, Y.-M., Pettersson, M., & St George-Hyslop, P. H. (2017). Structural and Chemical Biology of Presenilin Complexes. *Cold Spring Harbor Perspectives in Medicine*, *7*(12). https://doi.org/10.1101/cshperspect.a024067

Kaiser, H.F. An index of factorial simplicity. *Psychometrika* 39**,** 31–36 (1974). https://doi.org/10.1007/BF02291575

Kalbe, E., Kessler, J., Calabrese, P., Smith, R., Passmore, A. P., Brand, M., & Bullock, R. (2004). DemTect: A new, sensitive cognitive screening test to support the diagnosis of mild cognitive impairment and early dementia. *International Journal of Geriatric Psychiatry*, *19*(2), 136–143. https://doi.org/10.1002/gps.1042

Kessler, J., Bley, M., Mielke, R., & Kalbe, E. (1997). Strategies and structures in verbal fluency tasks in patients with Alzheimer's disease. *Behavioural Neurology*, *10*(4), 133–135. https://doi.org/10.3233/BEN-1997-10406

Keuck, L. (2018). Diagnosing Alzheimer's disease in Kraepelin's clinic, 1909–1912. *History of the Human Sciences*, *31*(2), 42–64. https://doi.org/10.1177/0952695118758879

Knopman, D. S., Jones, D. T., & Greicius, M. D. (2021). Failure to demonstrate efficacy of aducanumab: An analysis of the EMERGE and ENGAGE trials as reported by Biogen, December 2019. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association*, *17*(4), 696–701. https://doi.org/10.1002/alz.12213

Kohn, N., Kalbe, E., Georg, H., & Kessler, J. (2007). Vergleich MMST und DemTect: Spezifität und Sensitivität bei primär kognitiven Störungen. *Aktuelle Neurologie*, *34*(S 2), s-2007-987943. https://doi.org/10.1055/s-2007-987943

Kosunen, O., Soininen, H., Paljärvi, L., Heinonen, O., Talasniemi, S., & Riekkinen, P. J. (1996). Diagnostic accuracy of Alzheimer's disease: A neuropathological study. *Acta Neuropathologica*, *91*(2), 185–193. https://doi.org/10.1007/s004010050412

Kral, V. A. (1962). Senescent forgetfulness: Benign and malignant. *Canadian Medical Association Journal*, *86*, 257–260.

Lanctôt, K. L., Amatniek, J., Ancoli-Israel, S., Arnold, S. E., Ballard, C., Cohen-Mansfield, J., Ismail, Z., Lyketsos, C., Miller, D. S., Musiek, E., Osorio, R. S., Rosenberg, P. B., Satlin, A., Steffens, D., Tariot, P., Bain, L. J.,

Carrillo, M. C., Hendrix, J. A., Jurgens, H., & Boot, B. (2017). Neuropsychiatric signs and symptoms of Alzheimer's disease: New treatment paradigms. *Alzheimer's & Dementia (New York, N. Y.)*, *3*(3), 440–449. https://doi.org/10.1016/j.trci.2017.07.001

Lee, A., & Gilbert, R. M. (2016). Epidemiology of Parkinson Disease. *Neurologic Clinics*, *34*(4), 955–965. https://doi.org/10.1016/j.ncl.2016.06.012

Lin, J. S., O'Connor, E., Rossom, R. C., Perdue, L. A., Burda, B. U., Thompson, M., & Eckstrom, E. (2013). *Screening for Cognitive Impairment in Older Adults: An Evidence Update for the U.S. Preventive Services Task Force*. Agency for Healthcare Research and Quality (US). http://www.ncbi.nlm.nih.gov/books/NBK174643/

Livingston, G., Johnston, K., Katona, C., Paton, J., Lyketsos, C. G., & Old Age Task Force of the World Federation of Biological Psychiatry. (2005). Systematic review of psychological approaches to the management of neuropsychiatric symptoms of dementia. *The American Journal of Psychiatry*, *162*(11), 1996–2021. https://doi.org/10.1176/appi.ajp.162.11.1996

Livingston, G., Huntley, J., Sommerlad, A., Ames, D., Ballard, C., Banerjee, S., Brayne, C., Burns, A., Cohen-Mansfield, J., Cooper, C., Costafreda, S. G., Dias, A., Fox, N., Gitlin, L. N., Howard, R., Kales, H. C., Kivimäki, M., Larson, E. B., Ogunniyi, A., … Mukadam, N. (2020). Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. *The Lancet*, *396*(10248), 413–446. https://doi.org/10.1016/S0140-6736(20)30367-6

Long, J. M., & Holtzman, D. M. (2019). Alzheimer Disease: An Update on Pathobiology and Treatment Strategies. *Cell*, *179*(2), 312–339. https://doi.org/10.1016/j.cell.2019.09.001

Lopez, O. L., Litvan, I., Catt, K. E., Stowe, R., Klunk, W., Kaufer, D. I., Becker, J. T., & DeKosky, S. T. (1999). Accuracy of four clinical diagnostic criteria for the diagnosis of neurodegenerative dementias. *Neurology*, *53*(6), 1292–1292. https://doi.org/10.1212/WNL.53.6.1292

Loureiro, J. C., Pais, M. V., Stella, F., Radanovic, M., Teixeira, A. L., Forlenza, O. V., & de Souza, L. C. (2020). Passive antiamyloid immunotherapy for

Alzheimer's disease: *Current Opinion in Psychiatry*, *33*(3), 284–291. https://doi.org/10.1097/YCO.0000000000000587

Mahoney, R., Johnston, K., Katona, C., Maxmin, K., & Livingston, G. (2005). The TE4D-Cog: A new test for detecting early dementia in English-speaking populations. *International Journal of Geriatric Psychiatry*, *20*(12), 1172–1179. https://doi.org/10.1002/gps.1412

McKhann, G. M., Knopman, D. S., Chertkow, H., Hyman, B. T., Jack, C. R., Kawas, C. H., Klunk, W. E., Koroshetz, W. J., Manly, J. J., Mayeux, R., Mohs, R. C., Morris, J. C., Rossor, M. N., Scheltens, P., Carrillo, M. C., Thies, B., Weintraub, S., & Phelps, C. H. (2011). The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & Dementia*, *7*(3), 263–269. https://doi.org/10.1016/j.jalz.2011.03.005

Misra, A., Chakrabarti, S. S., & Gambhir, I. S. (2018). New genetic players in late-onset Alzheimer's disease: Findings of genome-wide association studies. *The Indian Journal of Medical Research*, *148*(2), 135–144. https://doi.org/10.4103/ijmr.IJMR_473_17

Mohs, R. C., Rosen, W. G., & Davis, K. L. (1983). The Alzheimer's disease assessment scale: An instrument for assessing treatment efficacy. *Psychopharmacology Bulletin*, *19*(3), 448–450.

Monsch, A. U., Seifritz, E., Taylor, K. I., Ermini-Fünfschilling, D., Stähelin, H. B., & Spiegel, R. (1997). Category fluency is also predominantly affected in Swiss Alzheimer's disease patients. *Acta Neurologica Scandinavica*, *95*(2), 81–84. https://doi.org/10.1111/j.1600-0404.1997.tb00073.x

Morris, J. C., Mohs, R. C., Rogers, H., Fillenbaum, G., & Heyman, A. (1988). Consortium to establish a registry for Alzheimer's disease (CERAD) clinical and neuropsychological assessment of Alzheimer's disease. *Psychopharmacology Bulletin*, *24*(4), 641–652.

Moscoso, A., Silva-Rodríguez, J., Aldrey, J. M., Cortés, J., Fernández-Ferreiro, A., Gómez-Lado, N., Ruibal, Á., & Aguiar, P. (2019). Staging the cognitive continuum in prodromal Alzheimer's disease with episodic memory.

Neurobiology of Aging, *84*, 1–8. https://doi.org/10.1016/j.neurobiolaging.2019.07.014

Moosbrugger, H. (2008). *Testtheorie und Fragebogenkonstruktion: Mit ... 43 Tabellen*. Springer. https://doi.org/10.1007/978-3-540-71635-8

Möller, H.-J., Laux, G., & Deister, A. (Hrsg.). (2015). *Psychiatrie, Psychosomatik und Psychotherapie* (6. Aufl., S. b-003-120842). Georg Thieme. https://doi.org/10.1055/b-003-120842

Müller, S., Herde, L., Preische, O., Zeller, A., Heymann, P., Robens, S., Elbing, U., & Laske, C. (2019). Diagnostic value of digital clock drawing test in comparison with CERAD neuropsychological battery total score for discrimination of patients in the early course of Alzheimer's disease from healthy individuals. *Scientific Reports*, *9*(1), 3543. https://doi.org/10.1038/s41598-019-40010-0

Naj, A. C., Schellenberg, G. D., & Alzheimer's Disease Genetics Consortium (ADGC). (2017). Genomic variants, genes, and pathways of Alzheimer's disease: An overview. *American Journal of Medical Genetics. Part B, Neuropsychiatric Genetics: The Official Publication of the International Society of Psychiatric Genetics*, *174*(1), 5–26. https://doi.org/10.1002/ajmg.b.32499

Nakajima, K., Ogi, H., Adachi, K., Noi, K., Hirao, M., Yagi, H., & Goto, Y. (2016). Nucleus factory on cavitation bubble for amyloid β fibril. *Scientific Reports*, *6*(1), 22015. https://doi.org/10.1038/srep22015

Nasreddine, Z. S., Phillips, N. A., Bédirian V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L., & Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: A Brief Screening Tool For Mild Cognitive Impairment: MOCA: A BRIEF SCREENING TOOL FOR MCI. *Journal of the American Geriatrics Society*, *53*(4), 695–699. https://doi.org/10.1111/j.1532-5415.2005.53221.x

Nizynski, B., Dzwolak, W., & Nieznanski, K. (2017). Amyloidogenesis of Tau protein: Tau Aggregation. *Protein Science*, *26*(11), 2126–2150. https://doi.org/10.1002/pro.3275

O'Brien, J. T., & Thomas, A. (2015). Vascular dementia. *Lancet (London, England)*, *386*(10004), 1698–1706. https://doi.org/10.1016/S0140-6736(15)00463-8

Oedekoven, C., & Dodel, R. (2019). Diagnostische Kriterien und Diagnose der Demenz vom Alzheimer-Typ. *Neurologie up2date*, *2*(01), 91–105. https://doi.org/10.1055/a-0803-4606

Orr, M. E., Sullivan, A. C., & Frost, B. (2017). A Brief Overview of Tauopathy: Causes, Consequences, and Therapeutic Strategies. *Trends in Pharmacological Sciences*, *38*(7), 637–648. https://doi.org/10.1016/j.tips.2017.03.011

Pereira, D. R., Costa, P., & Cerqueira, J. J. (2015). Repeated Assessment and Practice Effects of the Written Symbol Digit Modalities Test Using a Short Inter-Test Interval. *Archives of Clinical Neuropsychology*, *30*(5), 424–434. https://doi.org/10.1093/arclin/acv028

Petersen, R. C., Doody, R., Kurz, A., Mohs, R. C., Morris, J. C., Rabins, P. V., Ritchie, K., Rossor, M., Thal, L., & Winblad, B. (2001). Current concepts in mild cognitive impairment. *Archives of Neurology*, *58*(12), 1985–1992. https://doi.org/10.1001/archneur.58.12.1985

Petersen, R. C., Roberts, R. O., Knopman, D. S., Boeve, B. F., Geda, Y. E., Ivnik, R. J., Smith, G. E., & Jack, C. R. (2009). Mild cognitive impairment: Ten years later. *Archives of Neurology*, *66*(12), 1447–1455. https://doi.org/10.1001/archneurol.2009.266

Prince, M., Bryce, R., Albanese, E., Wimo, A., Ribeiro, W., & Ferri, C. P. (2013). The global prevalence of dementia: A systematic review and metaanalysis. *Alzheimer's & Dementia*, *9*(1), 63-75.e2. https://doi.org/10.1016/j.jalz.2012.11.007

Rascovsky, K., Salmon, D. P., Hansen, L. A., Thal, L. J., & Galasko, D. (2007). Disparate letter and semantic category fluency deficits in autopsy-confirmed frontotemporal dementia and Alzheimer's disease. *Neuropsychology*, *21*(1), 20–30. https://doi.org/10.1037/0894-4105.21.1.20

Reimers, C., Knapp, G., & Tettenborn, B. (2012). Einfluss körperliche Aktivität

auf die Kognition. Ist körperliche Aktivität Demenz-präventiv? *Aktuelle Neurologie*, *39*(06), 276–291. https://doi.org/10.1055/s-0032-1316354

Reisberg, B., Ferris, S. H., de Leon, M. J., & Crook, T. (1982). The Global Deterioration Scale for assessment of primary degenerative dementia. *The American Journal of Psychiatry*, *139*(9), 1136–1139. https://doi.org/10.1176/ajp.139.9.1136

Reitz, C., Brayne, C., & Mayeux, R. (2011). Epidemiology of Alzheimer disease. *Nature Reviews. Neurology*, *7*(3), 137–152. https://doi.org/10.1038/nrneurol.2011.2

Rost, J. (2004). *Lehrbuch Testtheorie—Testkonstruktion* (2., vollst. überarb. und erw. Aufl). Huber.

Rostamzadeh, A., & Jessen, F. (2020). Früherkennung der Alzheimer-Krankheit und Demenzprädiktion bei Patienten mit leichter kognitiver Störung: Zusammenfassung aktueller Empfehlungen. *Der Nervenarzt*, *91*(9), 832–842. https://doi.org/10.1007/s00115-020-00907-y

Rubin, E. H., Storandt, M., Miller, J. P., Kinscherf, D. A., Grant, E. A., Morris, J. C., & Berg, L. (1998). A prospective study of cognitive function and onset of dementia in cognitively healthy elders. *Archives of Neurology*, *55*(3), 395–401. https://doi.org/10.1001/archneur.55.3.395

Ruoff, A. (1981). *Häufigkeitswörterbuch gesprochener Sprache: Gesondert nach Wortarten, alphabetisch, rückläufig alphabetisch und nach Häufigkeit geordnet*. Niemeyer.

Sacks, C. A., Avorn, J., & Kesselheim, A. S. (2017). The Failure of Solanezumab—How the FDA Saved Taxpayers Billions. *The New England Journal of Medicine*, *376*(18), 1706–1708. https://doi.org/10.1056/NEJMp1701047

Saito, T., Oba, T., Shimizu, S., Asada, A., Iijima, K. M., & Ando, K. (2019). Cdk5 increases MARK4 activity and augments pathological tau accumulation and toxicity through tau phosphorylation at Ser262. *Human Molecular Genetics*, *28*(18), 3062–3071. https://doi.org/10.1093/hmg/ddz120

Satizabal, C. L., Beiser, A. S., Chouraki, V., Chêne, G., Dufouil, C., & Seshadri,

S. (2016). Incidence of Dementia over Three Decades in the Framingham Heart Study. *The New England Journal of Medicine*, *374*(6), 523–532. https://doi.org/10.1056/NEJMoa1504327

Schmitt, M., Gerstenberg, F., & Heim, A.-C. (2014). *Psychologische Diagnostik kompakt: Mit Arbeitsmaterial zum Download* (1. Aufl). Beltz.

Schmitz, K. (2004): Standardisierung psychometrischer Diagnostik in der neurologischen Rehabilitation: Eine computergestützte neuropsychologische Testbatterie. Dissertation, Universität Konstanz.

Schoenmakers, B., Buntinx, F., & DeLepeleire, J. (2010). Supporting the dementia family caregiver: The effect of home care intervention on general well-being. *Aging & Mental Health*, *14*(1), 44–56. https://doi.org/10.1080/13607860902845533

Sevigny, J., Chiao, P., Bussière, T., Weinreb, P. H., Williams, L., Maier, M., Dunstan, R., Salloway, S., Chen, T., Ling, Y., O'Gorman, J., Qian, F., Arastu, M., Li, M., Chollate, S., Brennan, M. S., Quintero-Monzon, O., Scannevin, R. H., Arnold, H. M., … Sandrock, A. (2016). The antibody aducanumab reduces Aβ plaques in Alzheimer's disease. *Nature*, *537*(7618), 50–56. https://doi.org/10.1038/nature19323

Shulman, K. I. (2000). Clock-drawing: Is it the ideal cognitive Screening test? *International Journal of Geriatric Psychiatry*, *15*(6), 548–561. https://doi.org/10.1002/1099-1166(200006)15:6<548::aid-gps242>3.0.co;2-u

Smith, J. D., Jamani, S., Boomer, J., & Church, B. A. (2018). One-back reinforcement dissociates implicit-procedural and explicit-declarative category learning. *Memory & Cognition*, *46*(2), 261–273. https://doi.org/10.3758/s13421-017-0762-8

Sparks Stein, P., Steffen, M. J., Smith, C., Jicha, G., Ebersole, J. L., Abner, E., & Dawson, D. (2012). Serum antibodies to periodontal pathogens are a risk factor for Alzheimer's disease. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association*, *8*(3), 196–203. https://doi.org/10.1016/j.jalz.2011.04.006

Stemmler, M., Lehfeld, H., Horn, R., Erzigkeit, H., & Erzigkeit, A. (2015). *SKT nach Erzigkeit: SKT Manual* (Edition 2015). Universität Erlangen-Nürnberg.

Stewart, R., Stenman, U., Hakeberg, M., Hägglin, C., Gustafson, D., & Skoog, I. (2015). Associations between oral health and risk of dementia in a 37-year follow-up study: The prospective population study of women in Gothenburg. *Journal of the American Geriatrics Society*, *63*(1), 100–105. https://doi.org/10.1111/jgs.13194

Sunderland, T., Hill, J. L., Mellow, A. M., Lawlor, B. A., Gundersheimer, J., Newhouse, P. A., & Grafman, J. H. (1989). Clock drawing in Alzheimer's disease. A novel measure of dementia severity. *Journal of the American Geriatrics Society*, *37*(8), 725–729. https://doi.org/10.1111/j.1532-5415.1989.tb02233.x

Taylor, K. I., Salmon, D. P., Monsch, A. U., & Brugger, P. (2005). Semantic and phonemic sequence effects in random word generation: A dissociation between Alzheimer's and Huntington's disease patients. *Journal of the International Neuropsychological Society*, *11*(3), 303–310. https://doi.org/10.1017/S1355617705050356

Tcw, J., & Goate, A. M. (2017). Genetics of β-Amyloid Precursor Protein in Alzheimer's Disease. *Cold Spring Harbor Perspectives in Medicine*, *7*(6), a024539. https://doi.org/10.1101/cshperspect.a024539

Thompson, E. P. (1967). Time, Work-Discipline, and Industrial Capitalism. *Past & Present*, *38*, 56–97.

Thorgrimsen, L., Spector, A., Wiles, A., & Orrell, M. (2003). Aroma therapy for dementia. *The Cochrane Database of Systematic Reviews*, *3*, CD003150. https://doi.org/10.1002/14651858.CD003150

Tsai, M.-C., Cheng, W.-L., Sheu, J.-J., Huang, C.-C., Shia, B.-C., Kao, L.-T., & Lin, H.-C. (2017). Increased risk of dementia following herpes zoster ophthalmicus. *PLOS ONE*, *12*(11), e0188490. https://doi.org/10.1371/journal.pone.0188490

Tzeng, N.-S., Chung, C.-H., Lin, F.-H., Chiang, C.-P., Yeh, C.-B., Huang, S.-Y.,

Lu, R.-B., Chang, H.-A., Kao, Y.-C., Yeh, H.-W., Chiang, W.-S., Chou, Y.-C., Tsao, C.-H., Wu, Y.-F., & Chien, W.-C. (2018). Anti-herpetic Medications and Reduced Risk of Dementia in Patients with Herpes Simplex Virus Infections-a Nationwide, Population-Based Cohort Study in Taiwan. *Neurotherapeutics: The Journal of the American Society for Experimental NeuroTherapeutics*, *15*(2), 417–429. https://doi.org/10.1007/s13311-018-0611-x

Villa, R. F., Ferrari, F., & Moretti, A. (2018). Post-stroke depression: Mechanisms and pharmacological treatment. *Pharmacology & Therapeutics*, *184*, 131–144. https://doi.org/10.1016/j.pharmthera.2017.11.005

Velayudhan, L., Ryu, S.-H., Raczek, M., Philpot, M., Lindesay, J., Critchfield, M., & Livingston, G. (2014). Review of brief cognitive tests for patients with suspected dementia. *International Psychogeriatrics*, *26*(8), 1247–1262. https://doi.org/10.1017/S1041610214000416

Vogel, A., Gade, A., Stokholm, J., & Waldemar, G. (2005). Semantic Memory Impairment in the Earliest Phases of Alzheimer's Disease. *Dementia and Geriatric Cognitive Disorders*, *19*(2–3), 75–81. https://doi.org/10.1159/000082352

Vos, S. J. B., Verhey, F., Frölich, L., Kornhuber, J., Wiltfang, J., Maier, W., Peters, O., Rüther, E., Nobili, F., Morbelli, S., Frisoni, G. B., Drzezga, A., Didic, M., van Berckel, B. N. M., Simmons, A., Soininen, H., Kłoszewska, I., Mecocci, P., Tsolaki, M., … Visser, P. J. (2015). Prevalence and prognosis of Alzheimer's disease at the mild cognitive impairment stage. *Brain*, *138*(5), 1327–1338. https://doi.org/10.1093/brain/awv029

Weissberger, G. H., Strong, J. V., Stefanidis, K. B., Summers, M. J., Bondi, M. W., & Stricker, N. H. (2017). Diagnostic Accuracy of Memory Measures in Alzheimer's Dementia and Mild Cognitive Impairment: A Systematic Review and Meta-Analysis. *Neuropsychology Review*, *27*(4), 354–388. https://doi.org/10.1007/s11065-017-9360-6

Wimo, A., Guerchet, M., Ali, G.-C., Wu, Y.-T., Prina, A. M., Winblad, B., Jönsson, L., Liu, Z., & Prince, M. (2017). The worldwide costs of dementia

2015 and comparisons with 2010. *Alzheimer's & Dementia*, *13*(1), 1–7. https://doi.org/10.1016/j.jalz.2016.07.150

Wu, Y.-T., Beiser, A. S., Breteler, M. M. B., Fratiglioni, L., Helmer, C., Hendrie, H. C., Honda, H., Ikram, M. A., Langa, K. M., Lobo, A., Matthews, F. E., Ohara, T., Pérès, K., Qiu, C., Seshadri, S., Sjölund, B.-M., Skoog, I., & Brayne, C. (2017). The changing prevalence and incidence of dementia over time—Current evidence. *Nature Reviews. Neurology*, *13*(6), 327–339. https://doi.org/10.1038/nrneurol.2017.63

Yong, A. G., & Pearce, S. (2013). A Beginner's Guide to Factor Analysis: Focusing on Exploratory Factor Analysis. *Tutorials in Quantitative Methods for Psychology*, *9*(2), 79–94. https://doi.org/10.20982/tqmp.09.2.p079

Youden, W. J. (1950). Index for rating diagnostic tests. *Cancer*, *3*(1), 32–35. https://doi.org/10.1002/1097-0142(1950)3:1<32::aid-cncr2820030106>3.0.co;2-3

Zoelch, C., Berner, V.-D., & Thomas, J. (2019). Gedächtnis und Wissenserwerb. In D. Urhahne, M. Dresel, & F. Fischer (Hrsg.), *Psychologie für den Lehrberuf* (S. 23–52). Springer. https://doi.org/10.1007/978-3-662-55754-9_2

6 Anhang

TFDD, Originalversion

TFDD-P, Parallelversion

Auswertung des *„Clock Drawing Test“* nach Sunderland

Name

Alter Jahre

Datum

Demenz-Score ☐ Depressions-Score ☐

Test zur Früherkennung von Demenzen mit Depressionsabgrenzung

TFDD®

© Schwabe

Unmittelbare Reproduktion

1 *„Lesen Sie bitte jedes Wort laut vor und prägen Sie es sich gut ein!"*

„An welche Wörter erinnern Sie sich?"

- ☐ Verkäufer
- ☐ Komet
- ☐ Nachricht
- ☐ Spiegel
- ☐ Märchen
- ☐ Dampf
- ☐ Abenteuer

„Bitte lesen Sie jetzt die Worte noch einmal!"

Erreichte Punktzahl (max. 7) ☐

Zeitliche Orientierung

2 *„Welches Datum ist heute?"*

Tag _____

Monat _____

Jahr _____

Erreichte Punktzahl (max. 3) ☐

3 *„Welche Jahreszeiten gibt es?"*

- ☐ Frühling
- ☐ Sommer
- ☐ Herbst
- ☐ Winter

Erreichte Punktzahl (max. 4) ☐

4 *„Welche Jahreszeit haben wir jetzt?"*
Toleranz ± 14 Tage

- ☐ falsch
- ☐ richtig

Erreichte Punktzahl (max. 1) ☐

5 *„Welche Monate gehören zu dieser Jahreszeit?"*

Frühling	Sommer	Herbst	Winter
☐ März	☐ Juni	☐ September	☐ Dezember
☐ April	☐ Juli	☐ Oktober	☐ Januar
☐ Mai	☐ August	☐ November	☐ Februar
☐ Juni	☐ September	☐ Dezember	☐ März

Erreichte Punktzahl (max. 4) ☐

Anweisungen befolgen

6 *„Greifen Sie sich erst mit der linken Hand an das rechte Ohr,*
dann mit der rechten Hand an das linke Ohr
und klatschen Sie danach in die Hände!"

☐
☐
☐

Vollständig richtige Reihenfolge der Durchführung ☐

Erreichte Punktzahl (max. 4) ☐

Konstruktive Praxis

7 *„Bitte zeichnen Sie das Zifferblatt einer Uhr mit allen Zahlen*
und stellen Sie die Zeiger auf 11:10 Uhr ein!"
Auswertung nach Schema Sunderland, (siehe Anlage)

Erreichte Punktzahl (max. 10) ☐

Verzögerte Reproduktion

8 *„Vorhin haben Sie Wörter gelesen, die Sie sich einprägen sollten.*
An welche dieser Wörter können Sie sich noch erinnern?"

Verkäufer ☐
Komet ☐
Nachricht ☐
Spiegel ☐
Märchen ☐
Dampf ☐
Abenteuer ☐

Erreichte Punktzahl (max. 7) ☐

Wortflüssigkeit

9 *„Für die nächste Aufgabe haben Sie eine Minute Zeit.*
Bitte nennen Sie mir soviele Tiere, wie sie können!"

Erreichte Punktzahl (max. 10) ☐

Erreichte Punktzahl Teil 1, Demenzscore: ☐

Fremdbeurteilung Depression

10 wirkt: ausgeglichen ☐ schwer depressiv

0 1 2 3 4 5 6 7 8 9 10

Selbstbeurteilung Depression

11 gibt an: ausgeglichen … ☐ … schwer depressiv …

0 1 2 3 4 5 6 7 8 9 10

… zu sein

Erreichte Punktzahl Teil 2, Depressionsscore: ☐

Name

Alter Jahre

Datum

„Bitte zeichnen Sie das Zifferblatt einer Uhr mit allen Zahlen und stellen Sie die Zeiger auf 11.10 Uhr ein!“

Name: _____
Alter: _____ Jahre
Datum: _____

Test zur Früherkennung von Demenzen mit Depressionsabgrenzung-P

Demenz-Score: _____ Depressions-Score: _____

Unmittelbare Reproduktion

1 *„Lesen Sie bitte jedes Wort laut vor und prägen Sie es sich gut ein!“*
„An welche Wörter erinnern Sie sich“

Händler ___
Mond ___
Lehre ___
Auto ___
Film ___
Rauch ___
Prüfung ___

„Bitte lesen Sie jetzt die Wörter noch einmal!“

Erreichte Punktzahl (max. 7) ____

Zeitliche Orientierung

2 „Welches Datum ist Heute“

Tag _____
Monat _____
Jahr _____

Erreichte Punktzahl (max. 3) ____

3 „Welche Tageszeiten gibt es?“

Morgen/Vormittag ___
Mittag/Nachmittag ___
Abend ___
Nacht ___

Erreichte Punktzahl (max. 4) ____

4 „Welche Tageszeit haben wir jetzt?“
Toleranz ±1Stunde

falsch ____
richtig ____

Erreichte Punktzahl (max. 1) ____

5 „Bitte ordnen Sie folgende Uhrzeiten den entsprechenden Tageszeiten zu“

19:00 → Abend ___
13:00 → Mittag ___
0:00 → Nacht ___
7:00 → Morgen ___

Erreichte Punktzahl (max. 4) ____

Anweisungen befolgen

6 *„Greifen Sie sich erst mit der linken Hand an den rechten Ellenbogen,* ___
dann mit der rechten Hand an den linken Ellenbogen ___
und legen Sie danach beide Hände auf den Tisch“ ___
Vollständig richtige Reihenfolge der Durchführung ___
Erreichte Punktzahl (max. 4) ____

Konstruktive Praxis

7 *„Bitte zeichnen Sie das Ziffernblatt einer Uhr mit allen Zahlen*
und stellen Sie die Zeiger auf 8:20 Uhr ein!“
Auswertung nach Schema Sunderland, (siehe Anlage)
Erreichte Punktzahl (max. 10) ____

Verzögerte Reproduktion

8 *„Vorhin haben Sie Wörter gelesen, die Sie sich einprägen sollten.*
An welche dieser Wörter können Sie sich noch erinnern?“

Händler ___
Mond ___
Lehre ___
Auto ___
Film ___
Rauch ___
Prüfung ___
Erreichte Punktzahl (max. 7) ____

Wortflüssigkeit

9 *„Für die nächste Aufgabe haben Sie eine Minute Zeit.*
Bitten nennen Sie mir so viele Dinge wie möglich,
die man im Supermarkt kaufen kann!“
Erreichte Punktzahl (max. 10) ____

Erreichte Punktzahl Teil 1, Demenzscore (max. 50): ____

Fremdbeurteilung Depression

10 wirkt: ausgeglichen schwer depressiv

___	___	___	___	___	___	___	___	___	___	___
0	1	2	3	4	5	6	7	8	9	10

Selbstbeurteilung Depression

11 gibt an: ausgeglichen… schwer depressiv…

___	___	___	___	___	___	___	___	___	___	___
0	1	2	3	4	5	6	7	8	9	10

…zu sein

Erreichte Punktzahl Teil 2, Depressionsscore (max. 20): ____

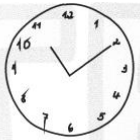
Name: _____
Alter: _____ Jahre
Datum: _____

„Bitte zeichnen Sie das Zifferblatt einer Uhr mit allen Zahlen und stellen Sie die Zeiger auf 8:20 Uhr ein!“

Die Auswertung des Clock Drawing Test

Die Zeichnung des Zifferblattes (Kreis und Zahlen) ist richtig

10 Die Zeiger sind in der richtigen Position

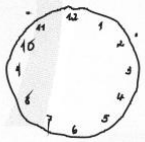


9 Leichte Fehler in der Zeigerposition

8 Auffällige Fehler in der Platzierung von Stunden- und Minutenzeiger.

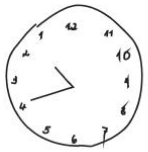
7 Die Zeigerplatzierung weicht deutlich von der Zeitvorgabe ab.

6 Statt der Verwendung von Zeigern andere Kennzeichnung der Uhrzeit.

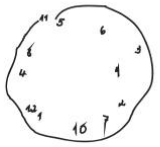


Die Zeichnung des Zifferblattes (Kreis und Zahlen) ist nicht richtig

5 Häufung von Zahlen in einem Bereich des Zifferblattes oder seitenverkehrte Plazierung der Zahlen. Möglicherweise werden noch Zeiger eingezeichnet.

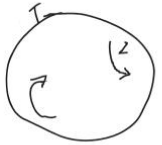


4 Deutliche Entstellung der Zahlenreihenfolge. Die Einheit des Zifferblattes ist aufgehoben (Zahlen fehlen oder sind außerhalb des Kreises plaziert).



3 Die Zahlen und das Zifferblatt bilden in der Zeichnung keine erkennbare Einheit. Zeiger werden nicht mehr eingezeichnet.

2 Die Zeichnung läßt erkennen, dass der Versuch unternommen wurde, eine Uhr zu zeichnen. Die Ähnlichkeit mit einer Uhr ist jedoch vage.



1 Entweder es wurde kein Versuch unternommen, die Aufgabe zu erfüllen, oder der Versuch ist nicht interpretierbar.

Bei allen Beispielen lautete die Zeitvorgabe 11.10 Uhr

Verändert nach Sunderland T, Hill J, Mellow A, Lawler BA, Gundersheimer J, Newhouse PA, Grafman JH (1989)
Clock drawing in Alzheimer´s disease: A novel measure of dementia severity. J Am Geriatr Soc 37:725-729