

## 4. Untersuchungsmethoden

### 4.1 Allgemein

Den Hauptteil der Untersuchung bildet die neuropsychologische Testung. An den Meßzeitpunkten T1 und T3 kommt eine Testbatterie zur Anwendung, die neben dem FCSRT (Free and Cued Selective Reminding Test; Buschke, 1984; Grober et al., 1988) fünf weitere Tests enthält (s. Tab. 4). Neben der neuropsychologischen Testung werden einige eigenanamnestische Daten des Probanden erhoben, die zum Beispiel die soziale Situation, die Schulbildung und aktuelle Erkrankungen betreffen. Als Maß für Depressivität kommt der Euro-D Fragebogen zur Anwendung (Prince et al., 1999). Vom Untersucher werden nach der Befragung des Probanden eine Fremdbeurteilungsskala zu „Speziellen Merkmalen des Verhaltens und des Sprechens“ ausgefüllt (unveröffentlichte Untersuchungsmaterialien der Berliner Altersstudie), die auf eventuell vorhandene Besonderheiten oder Symptome psychischer Krankheiten hinweisen könnten, sowie eine Skala über Symptome des Delirs (Reischies, Diefenbacher & Reichwaldt, 2003). Den Abschluß der Untersuchung bildet die Fremdanamnese, die anhand der Aussagen eines guten Bekannten des Probanden oder einer pflegenden oder betreuenden Person erhoben wird. An beiden Meßzeitpunkten werden die Studienteilnehmer anhand der Kriterien des DSM-IV auf das Vorliegen einer Demenzerkrankung untersucht.

An dem Meßzeitpunkt T2 kommt der „7 minute screen“ von Solomon und Mitarbeitern zur Anwendung (Solomon et al., 1998). Die Testbatterie „7 minute screen“ umfaßt vier kurze Tests (s. Tab. 4). Der „7 minute screen“ wird innerhalb dieser Untersuchung vor allem angewendet, um den Verlauf des FCSRT und der Wortflüssigkeit (Fluency) erfassen zu können. Anschließend werden vom Untersucher zwei Fremdratings durchgeführt, die Fremdbeurteilungsskala „Spezielle Merkmale des Verhaltens und des Sprechens“ sowie das Fremdrating über Symptome eines Delirs.

An dem Meßzeitpunkt T4, nach ungefähr zwei Jahren, wurde eine telefonische Testung durchgeführt, die zusammen mit der Fremdanamnese eine Demenzdiagnostik ermöglichte. Zu diesem Zweck wurden zwei verschiedene Telefoninterviews zusammengeführt (s. Tab. 4). Im dem umfassenden Telefoninterview sind neben einigen Elementen des Mini Mental Status Tests (Folstein, Folstein & McHugh, 1975) unter anderem der Zehnwort-Test von

Reischies (Reischies, Kühl & Krebs, 2000) und die Wortflüssigkeit enthalten.

| Test                      | Referenz  | T1 | T2 | T3 | T4 |
|---------------------------|---|----|----|----|----|
| FCSRT                     | Buschke, 1984; Grober et al., 1988  | x  | x* | x  |    |
| Reitan Trailmaking Test A | Reitan, 1958  | x  |    | x  |    |
| Lector Test               | Reischies, 1994, unveröffentlicht;<br>Reischies, Wertenaue & Kühl,<br>eingereicht | x  |    | x  |    |
| Attention Test Levine     | Mandell, Knoefel & Albert, 1997   | x  |    | x  |    |
| Wortflüssigkeit           | Lezak, 1995; Monsch et al., 1992  | x  | x  | x  | x  |
| MMST                      | Folstein, Folstein & McHugh,<br>1975  | x  |    | x  |    |
| Orientierung**            | Folstein, Folstein & McHugh,<br>1975; Solomon et al., 1998                        | x  | x  | x  | x  |
| Clock Drawing Test        | Sunderland et al., 1989; Freedman<br>et al., 1994                                 |    | x  |    |    |
| TIND                      | Kormann, 1998   |    |    |    | x  |
| TICS-m                    | Welsh et al., 1993  |    |    |    | x  |

**Tab. 4:** Testinstrumente und Meßzeitpunkte T1 bis T4. FCSRT = Free and Cued Selective Reminding Test; MMST = Mini Mental Status Test; TIND = Telefon-Interview für neuropsychologische Defizite; TICS-m = Telephone Interview for Cognitive Status - modified. \*Bei T2 wurde eine verkürzte Version des FCSRT durchgeführt. \*\*zu T2 und T3 wurde eine um ein Item verlängerte Version der Orientierung durchgeführt, die zusätzlich eine Gewichtung der Defizite zuließ (Solomon et al., 1998).

## 4.2 Speziell

Im folgenden werden die einzelnen Untersuchungsmethoden näher erläutert, die Variablen sowie die Durchführung der Untersuchung beschrieben.

### 4.2.1 Interview

Das Interview und der erste Teil des Fragebogens Euro-D sind der neuropsychologischen Testung vorangestellt. Die Fragen können relativ zügig, im Sinne einer standardisierten

ärztlichen Anamneseerhebung, gestellt werden. Der Eindruck eines ärztlichen Untersuchungsgesprächs ist durchaus gewollt, da dieses für die meisten Probanden eine vertraute Situation darstellt. Der Einstieg in die neuropsychologische Testung wird hierdurch erleichtert.

Im Interview macht der Proband Angaben zu einigen persönlichen aktuellen und biographischen Daten. Die Wohnsituation des Probanden wird erfaßt, Schul- und Berufsausbildung, sowie ausgeübte Berufe und Tätigkeiten. Anschließend wird der Teilnehmer nach schwerwiegenden körperlichen Erkrankungen befragt. Die Angaben zur Wohnsituation dienen mehreren Zwecken: Einerseits kann man Aufschluß darüber gewinnen, wie selbständig der Proband noch ist. Wohnt er in einer privaten Wohnung, in einem Seniorenwohnhaus mit Betreuung oder in einem Seniorenheim? Lebt er mit jemandem zusammen dort, der ihm eventuell Hilfestellung geben kann? Andererseits läßt die Beobachtung der Wohnsituation über den Untersuchungszeitraum Rückschlüsse auf mögliche einschneidende Veränderungen im Leben des Probanden zu. Hat zum Beispiel ein Umzug von einer privaten Wohnung in ein Pflegeheim stattgefunden, so könnte eine drastische Verschlechterung des kognitiven oder körperlichen Zustandes der Grund sein.

Die Angaben des Probanden zu schwerwiegenden körperlichen Erkrankungen sollen es möglich machen, somatische Ursachen für eventuelle schlechte kognitive und Gedächtnisleistungen auszuschließen. Anhand dieser Daten läßt sich auch erkennen, ob ein Zusammenhang zwischen dem Maß an Morbidität und dem Leistungsniveau in den Tests besteht.

#### 4.2.2 Euro-D

Eine Einschätzung der Depressivität des Probanden ist mit dem Fragebogen Euro-D möglich (Prince et al., 1999). Dieser Fragebogen wurde als europaweit einheitliches Maß der Depressivität entwickelt. Euro-D wurde für diese Untersuchung in zwei Teile aufgeteilt: Die Fragen 1 bis 6 stehen vor der neuropsychologischen Testung, die restlichen Fragen folgen ihr. Die ersten Fragen betreffen Schlafstörungen, Interesseverlust, Appetit, Antrieb, Freude an Dingen und die Konzentration. Diese Fragen können, wie oben schon erwähnt, ähnlich wie bei einer standardisierten ärztlichen Anamneseerhebung gestellt werden, was für die meisten Probanden eine vertraute Situation darstellt. Ein weiterer Grund, die Fragen

voranzustellen, ist der, daß zum Beispiel Fragen zur Konzentrationsfähigkeit nach der Bearbeitung der Tests deutlich anders beantwortet werden könnten, als davor. Nach der Testung könnten die Probanden den aktuellen Zustand beschreiben, in dem sie aufgrund der Anstrengung nicht mehr Konzentrationsfähig sind. Dies ist aber nicht gewollt, da eine generelle Einschätzung der Konzentration benötigt wird.

Die Fragen 7 bis 12 werden erst nach der neuropsychologischen Untersuchung gestellt. Es wird in diesen Fragen zum Beispiel nach Depressivität und Suizidgedanken gefragt. Da die Testleistungen eventuell durch langes Grübeln und Hinwendung auf die eigene Depressivität, ausgelöst durch die Befragung, beeinträchtigt werden könnte, wurde dieser Teil von Euro-D an das Ende der Untersuchung gestellt. Der zweite Teil beginnt mit der Frage, ob der Proband in der letzten Zeit traurig war. Es folgen Fragen nach der Einstellung zur eigenen Zukunft, Suizidalität (im Verlauf des ganzen Lebens), Selbstvorwürfen oder Schuldgefühlen, Reizbarkeit und ob der Proband in der letzten Zeit einmal geweint hat. Als abschließende Frage an den Probanden wird gefragt, ob ihm diese Untersuchung große Mühe gemacht habe. Diese Frage ist nicht Bestandteil des Euro-D. Sie wird einerseits gestellt, um ein subjektives Maß dafür zu erhalten, wie sehr die Untersuchung den Probanden angestrengt hat. Andererseits ist hierdurch ein verbindlicher Abschluß der Testung gegeben.

#### 4.2.3 Free and Cued Selective Reminding Test

Der Free and Cued Selective Reminding Test (FCSRT) von Buschke, der auch unter dem Namen Enhanced Cued Recall Test (ECR) bekannt ist, dient zur Erfassung der Leistung des episodischen Gedächtnisses (Buschke, 1984; Grober, Buschke, Crystal, Bang & Dresner, 1988; Grober, Merling, Heimlich, & Lipton, 1997). Besonderer Wert wird auf die elaborierte und für alle Probanden möglichst gleiche Enkodierung der zu lernenden Begriffe gelegt. Dadurch werden mögliche Fehlerquellen, wie etwa mangelnde Konzentration beim Lernen der Begriffe oder Benutzung von Lernstrategien, weitgehend ausgeschaltet. Die Retest-Reliabilität ist mit  $r=0.92$  bis  $r=0.98$  hoch (Solomon et al., 1998, bzw. Reischies et al., 1997b).

Durchführung des Free and Cued Selective Reminding Tests: Die 16 zu lernenden Items sind Strichzeichnungen von leicht zu erkennenden Objekten, jedes davon aus einer anderen semantischen Kategorie. Um die Trefferwahrscheinlichkeit durch Raten zu minimieren, werden dominante Objekte aus den einzelnen Kategorien vermieden. Die 16 zu lernenden

Items werden, jeweils 4 auf einmal auf einer Tafel im Format DIN A4, präsentiert, wobei ein Item in jedem Quadranten gezeigt wird. Eine Tafel nach der anderen wird vor dem Teilnehmer plaziert, die Reihenfolge der Tafeln ist für alle Teilnehmer die gleiche. Der Teilnehmer wird aufgefordert, jedes einzelne Item auf der Tafel zu suchen, darauf zu zeigen und es zu benennen (z.B. „Zange“), nachdem ihm das Stichwort (in diesem Fall: „Werkzeug“) verbal gegeben wurde. Wenn der Teilnehmer ein Item nicht benennen kann, wird es vom Untersucher genannt. Nachdem alle 4 Items korrekt identifiziert wurden, wird die Tafel entfernt, und die unmittelbare Wiedergabe dieser 4 Items wird getestet. Für die Items, die nach freier Erinnerung nicht genannt werden konnten, werden die Kategorie-Stichworte zur Verfügung gestellt. Kann der Teilnehmer sich anhand des Kategorie-Stichworts auch nicht an das Item erinnern, wird es vom Untersucher genannt. Der Teilnehmer wird daraufhin an jedes Item, an welches er sich auch nicht mit Hilfestellung hat erinnern können, zusammen mit dem kategorialen Hinweisreiz erinnert (z.B. das Werkzeug war eine Zange). Sobald die unmittelbare Wiedergabe für eine Gruppe von 4 Items komplett ist, wird die nächste Tafel mit 4 Items zum Lernen präsentiert.

Diese Versuchsphase wird gefolgt von drei Wiedergabedurchgängen, vor denen als Ablenkung jeweils eine halbe Minute rückwärts gezählt wird, um die Erinnerung schließlich aus dem sekundären Gedächtnis zu erhalten. Jeder Wiedergabedurchgang besteht aus zwei Teilen. Zuerst wird dem Teilnehmer eine ausgedehnte Zeitspanne von bis zu zwei Minuten Zeit gelassen, um sich an möglichst viele Items frei zu erinnern. Die Anzahl der erinnerten Begriffe bildet den Punktwert „freie Wiedergabe“. Als nächstes werden Kategorie-Stichwörter für die Items zur Verfügung gestellt, die nicht frei erinnert wurden. Die nun wiedergegebenen Begriffe werden unter „Wiedergabe mit Hilfestellung“ notiert. Wenn der Teilnehmer sich auch mit dem Kategorie-Stichwort nicht an das Item erinnern kann, wird ihm eine Hilfestellung gegeben, die nur dazu dient, dem Probanden nochmals den Begriff zu nennen, damit er im nächsten Durchgang die gleichen Voraussetzungen hat wie diejenigen Studienteilnehmer, die den Begriff erinnern (Grober & Kawas, 1997).

Anschließend werden mit dem Probanden die anderen Aufgaben der Testbatterie bearbeitet. Nach Abschluß dieser Aufgaben, ungefähr nach zwanzig Minuten, folgt nochmals eine Wiedergabe. Auch dieser Durchgang ist zweigeteilt. Zuerst wird die freie Wiedergabe getestet, danach die Wiedergabe mit Hilfestellung für die restlichen Begriffe. Die Wortlisten

der verschiedenen FCSRT-Versionen sind im Anhang abgedruckt (Anhang, S. 3).

Man erhält in dem FCSRT 12 Einzel-Scores von jedem Probanden. Dies sind im einzelnen: Benennen, Zeigen und unmittelbare Wiedergabe der Items, freie Wiedergabe und Wiedergabe mit Hilfestellung der drei Lerndurchgänge und der verzögerten Wiedergabe, sowie Wiedererkennen. In verschiedenen Studien, die den FCSRT verwendeten, wurden unterschiedliche Scores zur Analyse der Ergebnisse benutzt. Grober et al. (1988) benutzten den FCSRT als Screening-Instrument für Demenz und nahmen den Score der gesamten Wiedergabe über drei Wiedergabedurchgänge als Kriterium. Im 7-minute-screen von Solomon et al. (Solomon et al., 1998) kommt die gesamte Wiedergabe aus nur einem Lerndurchgang zur Anwendung. In den beiden oben genannten Studien wird nach bereits dementen Probanden gescreent. Grober und Mitarbeiter (Grober, Lipton, Hall & Crystal, 2000) benutzten zur Prädiktion der Demenz einen Cut-Off ausschließlich in der freien Wiedergabe von drei aufeinanderfolgenden Wiedergabedurchgängen. Wir testeten die genannten Scores und außerdem einen Summenscore des FCSRT, der die unmittelbare Wiedergabe beinhaltet, einen Durchgang der freien Wiedergabe und einen Durchgang der Wiedergabe mit Hilfestellung. Diesen Score bezeichneten wir als "3pi" (**3 points per item**), da für jedes Item 3 Punkte (maximal 48 Punkte) erreicht werden können.

Angemerkt sei noch, daß der FCSRT nicht mit dem SRT verwechselt werden sollte, dem Selective Reminding Test, den Buschke 1973 entwickelte. Bei der Durchführung des SRT werden nur Wortlisten gelernt (keine Bilder), es wird keine Förderung der Einspeicherung vorgenommen, außerdem können keine Hilfestellungen gegeben werden.

#### 4.2.4 Reitan Trailmaking Test A

Mit dem Reitan Trailmaking Test A mißt man die psychomotorische Geschwindigkeit des Probanden. Bei diesem Test sollen Zahlen möglichst schnell miteinander verbunden werden (Reitan, 1958).

Durchführung des Reitan Trailmaking Tests A: Dem Probanden wird zuerst ein Übungsblatt vorgelegt, auf dem die Zahlen 1 bis 8, jeweils in einem kleinen Kreis, zu sehen sind. Der Proband wird angewiesen, die Zahlen mit einem Bleistift zu verbinden, und zwar in der richtigen Reihenfolge, also die Eins mit der Zwei, die Zwei mit der Drei und so fort. Sollte der Proband hierbei Fehler machen, so wird er darauf aufmerksam gemacht. Somit kann

sichergestellt werden, daß er die Aufgabe verstanden hat und praktische Probleme bei der Durchführung können im voraus beseitigt werden.

Nachdem die acht Zahlen erfolgreich miteinander verbunden wurden wird der Proband darauf aufmerksam gemacht, daß nun die Aufgabe nochmals mit Zeitmessung durchgeführt wird. Das Blatt mit der richtigen Aufgabe wird vorgelegt, die Eins wird dem Probanden gezeigt und es wird gesagt, daß ab nun die Zeit läuft. Auf dem Aufgabenblatt sind die Zahlen 1 bis 25 zu sehen, in der selben Art wie zuvor auf dem Übungsblatt. Die Zahlen sind über die ganze Fläche des DIN A4-Blattes verteilt. Wenn der Proband einen Fehler macht, zum Beispiel wenn er eine Zahl überspringt, wird er darauf hingewiesen. Er muß zur letzten richtigen Zahl zurückgehen, die Zeit läuft dabei weiter. Einziges Kriterium ist bei dieser Aufgabe die Zeit, die der Proband für die Durchführung benötigt hat. Fehler gehen somit von der Zeit ab. Die Zeit wird angehalten, sobald der Proband bei der Zahl 25 angekommen ist. Benötigt der Proband jedoch länger als 240 Sekunden für die Durchführung der Aufgabe, wird die Aufgabe abgebrochen.

#### 4.2.5 Lector Test

Anhand der Leistung im Lector Test (Reischies, 1994, unveröffentlicht; Reischies, Wertenaue & Kühl, eingereicht; Wortliste s. Anhang, S.4) kann der frühere oder prämorbid Intelligenzquotient des Probanden abgeschätzt werden. Es liegen umfangreiche Erfahrungen mit diesem Test aus der Berliner Gedächtnissprechstunde vor. Mit dem Lector Test wird beurteilt, inwieweit der Proband selbst gelesene Wörter richtig aussprechen kann. Die Wörter sind Fremdwörter oder andere Wörter, die nach den Ausspracheregeln des Deutschen anders ausgesprochen werden müßten, als sie tatsächlich ausgesprochen werden, wie z.B. „Niveau“. Kennt der Proband das Wort nicht, so wird er es fälschlicherweise „eindeutschen“. Die Fähigkeit zur richtigen Aussprache ist, war sie einmal vorhanden, auch bei Demenzkranken nicht gestört. In sehr späten Stadien der Demenz sind natürlich auch hier Defizite vorhanden. In unserer Stichprobe werden jedoch keine Probanden mit derart schwerer Demenz untersucht.

Durchführung des Lector Tests: Dem Probanden wird ein Blatt vorgelegt, auf dem insgesamt 48 Wörter in drei Spalten zu je 16 Wörtern abgedruckt sind. Die 3 Spalten sind jeweils in aufsteigender Schwierigkeit sortiert. Der Proband bekommt die Anweisung, diese Wörter

vorzulesen. Ihm wird gesagt, daß es bei dieser Aufgabe nicht auf Geschwindigkeit ankommt, sondern nur auf eine richtige Aussprache. Es könnten unter den 48 Wörtern einige Wörter sein, die er nicht kennt, er solle sie aber trotzdem versuchen auszusprechen. Die Aussprache jedes einzelnen Wortes wird auf Richtigkeit überprüft; jedes richtig ausgesprochene Wort gibt einen Punkt.

#### 4.2.6 Attention Test Levine

Der Attention Test Levine ist ein kurzer Test der Aufmerksamkeit (Mandell, Knoefel & Albert, 1997). Er ist eine Variation eines Continuous Performance Tests.

Durchführung des Attention Tests Levine: Dem Probanden werden bei dieser Aufgabe 60 einstellige Zahlen, jeweils eine pro Sekunde, vorgelesen. Wenn zwei ungerade Zahlen aufeinander folgen, soll er mit dem Zeigefinger ein Zeichen geben. Gibt der Proband kein Zeichen, obwohl zwei ungerade Zahlen hintereinander vorgelesen wurden, wird dies als falsch-negativ gewertet. Wird ein Zeichen gegeben, obwohl die Voraussetzungen hierfür nicht gegeben sind, wird dies als falsch-positiv notiert. Die Anzahl der falsch-negativen und der falsch-positiven Zeichen wird getrennt voneinander notiert und ausgewertet. Falsch positive Fehler zeigen impulsives, ablenkbares und reizgebundenes Verhalten an, wogegen falsch negative Fehler auf eine zerebrale Unteraktivierung, durch verschiedene Ursachen, hinweisen. Das pathologische Substrat, welches diesen Fehlern zugrunde liegt, ist meist eine Schädigung des Frontallappens oder frontaler Systeme, obwohl die Befunde relativ unspezifisch sind. Statistische Validität muß für diesen Test noch gezeigt werden. Jedoch zeigen Befunde an einer großen Zahl normaler und zerebral geschädigter Personen, daß ältere Menschen in der Altersgruppe vom 6. bis 9. Lebensjahrzehnt nicht mehr als zwei Fehler machen sollten (Mandell, Knoefel & Albert, 1997).

#### 4.2.7 Wortflüssigkeit

Mit dem Test der Wortflüssigkeit wird überprüft, ob der Proband Zeichen einer Aphasie aufweist. In der Demenzforschung wird hauptsächlich kategoriale Wortflüssigkeit eingesetzt (z.B. Tiere), da die Sensitivität und Spezifität hierbei deutlich höher liegen als bei Wortflüssigkeitsaufgaben nach Anfangsbuchstaben (Lezak, 1995; Monsch et al., 1992).

Durchführung des Tests der Wortflüssigkeit: Zur Einleitung dieser Aufgabe wird dem

Probanden gesagt, er habe 90 Sekunden Zeit, um so viele Tiere zu nennen wie möglich und er könne sofort damit beginnen. Dann wird die Uhr gestartet. Die vom Probanden genannten Tiere werden mitgeschrieben, um Mehrfachnennungen zu erkennen. Nach jeweils 30 Sekunden wird im Mitschrieb ein Zeichen gesetzt, um die Verteilung der Wortnennungen über die Zeit verfolgen zu können.

#### 4.2.8 Mini Mental Status Test und MMblind

Der Mini Mental Status Test erlaubt anhand eines Fragebogens eine Abschätzung verschiedener kognitiver und Gedächtnisleistungen innerhalb von 10 bis 15 Minuten (Folstein, Folstein & McHugh, 1975). Bei diesem Test sind insgesamt 30 Punkte zu vergeben. Der Test überprüft folgende Fähigkeiten: Zeitliche und örtliche Orientierung (jeweils 5 Punkte), Aufnahmefähigkeit und Gedächtnis (jeweils 3 Punkte), Aufmerksamkeit und Rechnen (5 Punkte), Sprache (3 Punkte), Lesen (1 Punkt), Schreiben (1 Punkt), konstruktive Praxis (oder Kopieren) (1 Punkt) sowie das Ausführen eines Befehls (3 Punkte). Üblicherweise wird ein Cut-off für Hirnleistungsstörung bei 23/24 Punkten angenommen (Folstein, Folstein & McHugh, 1975). Daten aus BASE zeigen Durchschnittswerte für die Altersgruppen 80-84 Jahre bei 27,33 Punkten (SD = 2,28), 85-89 Jahre bei 25,85 Punkten (SD = 2,85), 90-94 Jahre bei 25,07 Punkten (SD = 2,77) und älter als 95 Jahre bei 24,64 Punkten (SD = 2,67). In der Altersgruppe von 70-84 Jahren wird ein cut-off für Hirnleistungsstörungen von 26-27 Punkten angegeben, für die Altersgruppe 85 bis 95 Jahre und älter liegt der cut-off bei 23-24 Punkten. Die cut-off-Werte sind bewußt hoch gelegen, damit möglichst früh der Verdacht auf eine Hirnleistungsstörung geäußert werden kann (Reischies & Geiselmann, 1997a). Der Mini Mental Status Test ist jedoch für frühe Demenzformen nicht empfindlich. Ein normaler Wert kann folglich eine beginnende Demenz nicht ausschließen. Außerdem bedeuten schlechte Testergebnisse nicht, daß der Proband dement ist. Der MMST ist als grobe Screening Methode gedacht, wobei bei auffälligen Ergebnissen genauer nachuntersucht werden muß.

Für Teilnehmer, die einzelne Teilaufgaben aufgrund visueller oder motorischer Beeinträchtigungen nicht lösen konnten, wurde ein von Reischies und Geiselmann (1997a) eingeführter Score des MMST, der MMblind errechnet. Dieser Score beruht ausschließlich auf den Items des MMST, die nicht von dem Visus oder der Motorik der Untersuchten abhängen. Dies sind im einzelnen: Die Orientierungsitems (10 Punkte), Nachsprechen und

Erinnern (6 Punkte), Rückwärtsrechnen (5 Punkte) und einen Satz nachsprechen (1 Punkt). Insgesamt sind maximal 22 Punkte erreichbar. Der Cut-off für Hirnleistungsstörung liegt bei 16/17.

Durchführung des Mini Mental Status Tests: Zu Beginn wird die zeitliche Orientierung des Probanden überprüft. Er wird nach dem genauen Datum gefragt, nach dem Wochentag und der Jahreszeit. Danach folgt die Prüfung der Aufnahmefähigkeit. Dem Proband wird gesagt, daß er sich drei Wörter merken soll, da er später danach gefragt werden wird. Zunächst soll er diese jedoch nachsprechen. Als nächstes wird der Proband nach seiner Orientierung zum Ort befragt. Er soll den Staat, das Bundesland, die Stadt, den Stadtteil sowie das Stockwerk nennen, in dem er sich gerade befindet. Anschließend wird er gebeten, einen beliebigen, vollständigen Satz auf ein leeres Blatt Papier zu schreiben. Als Test der konstruktiven Praxis folgt eine Kopierprobe: Zwei sich überschneidende Fünfecke sollen abgezeichnet werden. Den Punkt gibt es, wenn sich die Fünfecke an der richtigen Stelle überschneiden. Anhand der drei Begriffe, mit denen vorher schon die Aufnahmefähigkeit geprüft wurde, wird nun das Kurzzeitgedächtnis getestet. Für jeden richtig erinnerten Begriff gibt es einen Punkt. Die Sprache wird anhand verschiedener Fähigkeiten überprüft. Zuerst soll der Proband zwei Gegenstände benennen. Anschließend soll der Proband eine Anweisung lesen und dann ausführen („Schließen Sie die Augen“) und einen Satz nachsprechen. Nun folgt das Ausführen eines dreiteiligen Befehls. Der Proband soll ein Blatt Papier in die rechte Hand nehmen, in der Mitte falten und in seinen Schoß legen. Für jede Teilaufgabe gibt es hier einen Punkt. Des weiteren wird noch eine Rechenaufgabe durchgeführt. Hierbei soll der Proband von hundert sieben abziehen, das Ergebnis nennen, wieder sieben abziehen usw., insgesamt fünf Mal. Für jedes richtig genannte Ergebnis gibt es einen Punkt. Verrechnet sich der Proband, so gilt dies als Fehler, er darf jedoch das falsche Zwischenergebnis als Ausgangspunkt für die nächste Rechnung benutzen. Bei den verschiedenen Meßzeitpunkten wurden in den Merkwörtern unterschiedliche Versionen des MMST durchgeführt, um Verfälschungen der Testergebnisse durch dieselben Merkwörter zu vermeiden (s. Anhang, ab S. 5).

#### 4.2.9 Fremdbeurteilung: Spezielle Merkmale des Verhaltens und Sprechens

Im Anschluß an die Testung werden vom Untersucher zwei Fremdbeurteilungsbögen ausgefüllt. Der erste davon ist der Bogen: „Spezielle Merkmale des Verhaltens und

Sprechens“ (unveröffentlichte Untersuchungsmaterialien der Berliner Altersstudie). Hiermit werden verschiedene besondere Verhaltensmerkmale und Besonderheiten in der Sprache des Probanden erfaßt, die zum Beispiel Hinweise auf psychische Erkrankungen geben könnten. Die Merkmale sind im Einzelnen: Artikulationsprobleme, umständliche Sprache, verminderte Menge der Sprachäußerungen, vermehrte Menge der Sprachäußerungen, Verständnisstörungen im Gespräch, Bewegungsunruhe, Erinnerungsstörungen im Gesprächsverlauf, Mißtrauen, bedrückte Stimmung, vermehrte Stimmungsschwankungen, Verlangsamung, Ängstlichkeit und Erschöpfung nach der Untersuchung. Der letzte Punkt wurde von uns hinzugefügt, um für die Erschöpfung des Probanden eine Fremdbeurteilung zu erhalten.

#### 4.2.10 Fremdbeurteilung: Störung der Bewußtheit der Umgebung als Symptom des Delirs

Dieser Fremdbeurteilungsbogen deckt das zentrale Merkmal des Delirs, die Bewußtheitsstörung, ab (Reischies, Diefenbacher & Reichwaldt, 2003). Ein Delir soll aufgrund seiner Beeinflussung der Kognition und des Gedächtnisses ausgeschlossen werden. Es werden auf diesem Bogen des weiteren folgende Merkmale beurteilt: Ausmaß der Reagibilität, Geschwindigkeit der Reaktionen, Zuwendung der Aufmerksamkeit zur Person des Untersuchers und zur Umgebung, Vergegenwärtigung der Interviewsituation und Augenbewegungen.

#### 4.2.11 Fremdanamnese

Nach der Untersuchung wird eine Fremdanamnese von einer Person erhoben, die den Probanden gut kennt. Dies kann ein Verwandter, ein guter Bekannter oder ein Betreuer sein. Bei Studienteilnehmern, die in einer Einrichtung wie zum Beispiel einem Seniorenwohnhaus wohnen, ist dies meist der Leiter der Einrichtung oder ein Betreuer, der häufigen Umgang mit dem Probanden hat. In der Fremdanamnese werden Daten zu folgenden Punkten erhoben: körperliche Erkrankungen, bleibende Schädigungen oder Behinderungen, Schlaganfälle, Schlafprobleme, Störungen des Gedächtnisses, der Orientierung und anderer kognitiver Fähigkeiten, geschäftliche und soziale Kompetenzen, Befähigung zu Hobbys und privaten Interessen, Körperpflege, soziale Einbindung und psychische Erkrankungen. Es wird auch danach gefragt, wie häufig die betreffende Person mit dem Probanden Kontakt hat, um die Verlässlichkeit der Informationen abschätzen zu können. Abschließend wird die

Möglichkeit eingeräumt, über Dinge zu berichten, die zur Charakterisierung des Probanden wichtig erscheinen, aber nicht explizit gefragt wurden.

#### 4.2.12 Untersuchung nach einem halben Jahr (T2); „7 minute screen“

Der „7 minute screen“ ist für die Alzheimer-Demenz hoch spezifisch (96%) und hoch sensitiv (92%; Solomon et al., 1998). Er besteht aus vier einzelnen Aufgaben, die im Schnitt innerhalb von ungefähr 7 Minuten absolviert werden können. Die einzelnen Aufgaben sind: zeitliche Orientierung aus dem Benton Test, ein Lerndurchgang des Enhanced Cued Recall Tests, der Clock Drawing Test (CDT) und Wortflüssigkeit. Die einzelnen Bestandteile des „7 minute screen“ sind in den folgenden Abschnitten dargestellt.

##### 4.2.12.1 Zeitliche Orientierung aus dem Benton Test

Die Messung der Orientierung anhand des Benton Schemas läßt eine unterschiedliche Gewichtung der Fehler zu (Solomon et al., 1998). Z.B. werden Fehler bei der Jahreszahl im Verhältnis zu Fehlern beim Wochentag 10fach bewertet.

Durchführung des Tests zur zeitlichen Orientierung: Der Studienteilnehmer soll fünf Fragen zur zeitlichen Orientierung beantworten. Er wird nach dem aktuellen Monat, dem Datum, dem Jahr, dem Wochentag und der genauen Uhrzeit gefragt, wie dies im MMST auch der Fall ist. Es gibt Fehlerpunkte für falsch gemachte Angaben. Wenn alles richtig genannt wird, bekommt der Proband null Punkte. Bei falschen Angaben gibt es je nach Schwere der Fehlleistung unterschiedlich viele Fehlerpunkte: Jahre zählen mit 10 Fehlerpunkten pro Jahr Abweichung mehr als z.B. Tage, bei denen nur ein Fehlerpunkt pro abweichendem Tag vergeben wird. Das ausführlich Bewertungsschema ist im Anhang dargestellt (Anhang, S. 9).

##### 4.2.12.2 Free and Cued Selective Reminding Test

Speziell für die Durchführung des „7 minute screen“ wurden von uns neue Kategorien und Items gesucht, da in der Arbeit von Solomon et al. (1998) Items verwendet wurden, die in unseren beiden Parallelversionen für T1 und T3 auch vorkommen. Wir gingen bei der Auswahl so vor wie Grober und Mitarbeiter, indem wir dominante Objekte aus den entsprechenden Kategorien nach Möglichkeit vermieden, um die Ratewahrscheinlichkeit zu minimieren (Grober et al., 1988). Die Wortliste des FCSRT von T2 ist im Anhang

abgedruckt (Anhang, S. 3).

Durchführung des FCSRT: Der Durchführung des Tests ist bis auf zwei Punkte identisch mit der Durchführung des FCSRT zu T1, wie sie oben beschrieben ist. Ein Unterschied liegt darin, daß nach der Lernphase nicht rückwärts gerechnet wird, sondern der Proband die Monate des Jahres rückwärts aufsagen soll. Der zweite wichtige Unterschied ist die Länge des Tests. In dem „7ms“ wird nur einmal die Wiedergabe der Items getestet, nicht wie in der langen Version des FCSRT drei Mal. Der  $\beta$ pi-Score (s. Methoden, S. 29) kann auch in dieser kurzen FCSRT-Version berechnet werden.

#### 4.2.12.3 Clock Drawing Test

Bei der hier verwendeten Version des Clock Drawing Tests soll der Proband das Ziffernblatt einer Uhr zeichnen, auf dem alle Zahlen präsent sind. Die zweite Aufgabe besteht darin, Zeiger einzuzeichnen, die auf zwanzig Minuten vor vier stehen. Diese Zeit wurde von Sunderland et al. empfohlen (Sunderland et al., 1989). In Anlehnung an Freedman et al. (Freedman et al., 1994) wird ein vereinfachte Bewertungsschema benutzt, in welchem 7 Merkmale bewertet werden (siehe unten). Die Bewertung kann von dem Untersucher vorgenommen werden, während der Proband die Uhr zeichnet.

Durchführung des Clock Drawing Tests: Der Proband bekommt vom Untersucher ein leeres Blatt Papier und einen Stift ausgehändigt. Anschließend bekommt er die Anweisung: „Bitte zeichnen Sie eine Uhr! Beginnen Sie bitte mit dem Ziffernblatt. Zeichnen Sie es groß“. Während der Proband die Anweisungen befolgt, können die ersten drei Kriterien überprüft werden: 1. Die Zahlen von 1 bis 12 sollten präsent sein. Der Punkt wird nicht vergeben, wenn eine Zahl fehlt, oder wenn andere Zahlen außer 1 bis 12 vorhanden sind. 2. Die Nummern sind in der richtigen Reihenfolge. Die Nummern müssen immer ansteigen, jedoch nicht unbedingt bis zur 12 reichen. Der Punkt wird auch vergeben, wenn Nummern fehlen. 3. Die Nummern sind in der richtigen Position. Es sollten in jedem Quadranten die zugehörigen Nummern stehen. Nun folgt die zweite Anweisung: „Zeichnen Sie jetzt bitte Zeiger ein, die auf zwanzig Minuten vor vier zeigen“. Die weiteren Punkte werden wie folgt vergeben: 4. Zwei Zeiger sind präsent. Nur für Zeiger gibt es den Punkt, nicht, wenn die Uhrzeit durch etwas anderes, z.B. Kreise oder Striche markiert ist. 5. Die richtige Stunde ist angezeigt. Diesen Punkt bekommt der Proband auch, wenn die Uhrzeit durch kleine Striche

oder Kreise angezeigt ist, vorausgesetzt, die Markierung ist näher an der vier als an jeder anderen Ziffer. 6. Die richtigen Minuten sind angezeigt. Die Bewertung erfolgt entsprechend der Bewertung der Stunde, in diesem Fall muß die Markierung näher an der acht als an jeder anderen Ziffer sein. 7. Den letzten Punkt gibt es, wenn die Zeiger in der korrekten Proportion zueinander sind. Der Proband bekommt diesen Punkt auch, wenn er von selbst äußert, daß der Stundenzeiger kleiner als der Minutenzeiger sein sollte.

#### 4.2.12.4 Wortflüssigkeit

Die Untersuchung der Wortflüssigkeit in dem „7 minute screen“ wird genauso durchgeführt, wie oben beschrieben, außer daß zu diesem Meßzeitpunkt eine andere semantische Kategorie gewählt wurde, nämlich „Nahrungsmittel“. Auch wenn für die Auswertung des „7 minute screen“ nur die Anzahl der genannten Wörter innerhalb einer Minute gewertet wird, werden zu T2, wie bereits zu T1, die innerhalb von 90 Sekunden genannten Wörter gezählt. Die Verlängerung der Meßzeit im Vergleich zur Version von Solomon und Mitarbeitern (Solomon et al., 1998) ist möglich, da die Messung der Wortflüssigkeit sowieso der letzte Punkt des „7 minute screen“ ist. Somit wird das Ergebnis des „7 minute screen“ nicht beeinflusst.

#### 4.2.13 Telefoninterview und Fremdanamnese zu T4

Das von uns durchgeführte Telefoninterview besteht aus zwei verschiedenen Instrumenten: Einerseits dem TICS-m (Telephone Interview for Cognitive Status - modified; Welsh et al., 1993), das an den Mini Mental Status Test angelehnt ist, und andererseits aus dem TIND (Telefon-Interview für neuropsychologische Defizite), das von Reischies und Mitarbeitern neu entwickelt wurde (Kormann, 1998). Es gibt mehrere Gründe, die uns zu dieser Kombination bewogen haben: Das TICS-m ging aus dem TICS hervor, das 1988 von Brandt, Spencer und Folstein vorgestellt wurde (Brandt, Spencer & Folstein, 1988). Es ist somit seit einigen Jahren verfügbar und bereits von verschiedenen Gruppen in Studien verwendet worden, jedoch nur in der englischen Version (z.B. Mangione et al., 1993; Gallo & Breitner, 1995; Lines et al., 2003). Für unsere Zwecke mußte es ins Deutsche übersetzt werden, für das noch keine Validierung vorliegt. Da durch das Telefonsetting bestimmte Umstände nicht kontrolliert werden können, war es ein Nebenziel der deutschen Version, Items auszuwählen, die nicht durch Hilfsmittel erschlossen werden konnten, wie z.B. das

Datum durch einen Kalender in der Wohnung des Untersuchten. Das TICS hat eine hohe Sensitivität (94%) und hohe Spezifität (100%) für die Demenzdiagnose (Brandt, Spencer & Folstein, 1988). Das TIND ist in deutscher Sprache in Untersuchungen von Reischies und Mitarbeitern bereits validiert; es liegt in Form der Dissertation von Sabine Kormann vor (Kormann, 1998). Durch das TIND wird somit sichergestellt, daß die Sprache keinen Einfluß auf das Testergebnis hat. Außerdem enthält das TIND Elemente, die in den bisherigen Testbatterien (von T1 bis T3) enthalten waren, wie z.B. die Wortflüssigkeit und die Orientierung. Im TIND enthalten ist des weiteren der Zehnwort-Test von Reischies, der ein genaueres Maß der Gedächtnisleistung darstellt, als die im TICS-m enthaltene Wortliste (Reischies, Kühl & Krebs, 2000). Das TIND erreicht in der Untersuchung von Kormann eine Sensitivität von 97% und eine Spezifität von 100% für die Demenzdiagnose. Durch Fragen, die in beiden Instrumenten vorkommen und folglich nicht doppelt gestellt werden mußten, hielt sich die Dauer des Telefoninterviews mit ca. 15 Minuten in einem vertretbaren Rahmen. Das Telefoninterview ist im Anhang abgedruckt (Anhang, S. 10). Die für die Demenzdiagnostik notwendige Fremdanamnese wurde anhand des selben Anamnesebogens abgefragt, wie bereits zu T1 und T3 (siehe dort).

### **4.3 Statistische Auswertung der Ergebnisse**

Zum Vergleich verschiedener Gruppengrößen wurden  $\chi^2$ -Tests durchgeführt (bei einzelnen Gruppengrößen unter 5 pro Feld wurde der exakte Test nach Fisher berechnet), Mittelwerte zwischen verschiedenen Gruppen wurden mit T-Tests untersucht (Armitage, Berry & Matthews, 2002; Guggenmoos-Holzmann & Wernecke, 1995). Verschiedene Tests, bzw. Testscores eines Tests, wurden mit Receiver-Operator-Characteristics-Kurven im Hinblick auf die Demenzprädiktion verglichen (ROC-Kurven) und einige Prädiktionen wurden durch logistische Regressionen genauer geprüft (Armitage, Berry & Matthews, 2002). Um longitudinal Unterschiede zwischen verschiedenen Diagnosegruppen über die drei Meßzeitpunkte hinweg erkennen zu können, wurde eine multivariate Varianzanalyse mit Meßwiederholungen durchgeführt; bei Verletzung der Homogenitätsannahme (gemessen durch den Box's M Test) wurden die Werte vorab Probit-transformiert (Tabachnick & Fidell, 1989). Die Probit-Werte ( $1/[1-P]$ ) reflektieren die empirische Wahrscheinlichkeit, daß ein Individuum die volle Punktzahl im jeweiligen Test erzielte. Für die Transformation wurden die empirischen Testwerte zunächst in Prozentwerte (volle Punktzahl entspricht 100%)

umgerechnet. Für einige Testscores, die miteinander in Verbindung stehen (wie zum Beispiel der MMST und der MMblind), wurden Korrelationsanalysen durchgeführt (Armitage, Berry & Matthews, 2002). Als Signifikanzniveau wurde  $p < .05$  festgelegt.

#### **4.4 Ethische Erwägungen**

Wichtigstes ethisches Problem dieser Studie ist das Problem der Einwilligungsfähigkeit der Probanden in die geplante Untersuchung. Einwilligungsfähigkeit liegt vor, wenn die betreffende Person in der Lage ist, „Wesen, Bedeutung und Tragweite der klinischen Prüfung zu erkennen und sich danach zu bestimmen“ (Arzneimittelgesetz §40 Abs. 2 Nr. 1, nach Helmchen & Lauter, 1995). Diese Einwilligungsfähigkeit ist im Einzelfall zu bestimmen; es gibt keine generelle Einwilligungsfähigkeit, sondern nur Einwilligungsfähigkeit in bezug auf eine konkrete Maßnahme. Sie kann gleichzeitig für bestimmte Bereiche vorhanden und für andere Bereiche nicht vorhanden sein. In diesem Zusammenhang ist zu betonen, daß eine Demenzdiagnose nicht zwangsläufig fehlende Einwilligungsfähigkeit in bestimmte Forschungsvorhaben bedeutet (Marson, Schmitt, Ingram & Harrell, 1994). Zur Feststellung der Einwilligungsfähigkeit ist vor der Untersuchung zu prüfen, ob der Studienteilnehmer die Informationen zur geplanten Untersuchung verstanden hat und sie rational verarbeiten, subjektiv bewerten und schließlich seinen Willen kundtun kann (Helmchen & Lauter, 1995). In einer neueren Untersuchung wurde festgestellt, daß momentan noch kein Goldstandard zur Feststellung der Einwilligungsfähigkeit, zum Beispiel mit kognitiven Tests, zur Verfügung steht (Vollmann et al., 2004).

Es liegt bei dieser Untersuchung keine körperliche Intervention am Probanden vor. Jedoch kann die Konfrontation des Probanden mit seinen kognitiven und Gedächtnisdefiziten eine Belastung für ihn darstellen. Deshalb ist es erforderlich, den Probanden vor der Untersuchung über den Untersuchungsplan und die Art der Untersuchung inklusive der Wiederholungsuntersuchungen aufzuklären. Ihm wird erklärt, daß er jederzeit das Recht habe, die Untersuchung ohne Angabe von Gründen und ohne Nachteile für ihn abbrechen. Die Daten würden nur zu wissenschaftlichen Zwecken erhoben und vertraulich behandelt. Abschließend wird durch eigene Wiedergabe dieser Informationen überprüft, ob der potentielle Proband diese Informationen verstanden hat, sie bewerten kann und an der Untersuchung teilnehmen möchte. Ist dies der Fall wird mit der Befragung begonnen. Dieser letzte Punkt ist im Sinne einer Überprüfung der Einwilligungsfähigkeit der

Studienteilnehmer zu sehen.

Bei einem Nutzen-Risiko-Verhältnis, bei dem der Nutzen bei weitem überwiegt, ist die Schwelle zur Annahme der Einwilligungsfähigkeit niedriger anzusetzen als bei größerem Risiko (Helmchen & Lauter, 1995). Das Nutzen-Risiko-Verhältnis dieser Untersuchung liegt aufgrund des sehr geringen Risikos deutlich auf Seiten des Nutzens. Großer Nutzen könnte für die Allgemeinheit und besonders für zukünftige Patienten mit beginnender Demenz entstehen, wenn es gelänge, eine Möglichkeit der frühzeitigen Diagnostizierung von Demenzen zu finden. Dieser Nutzen kommt jedoch nicht dem Studienteilnehmer zugute und darf somit nicht für die Begründung der Untersuchung herangezogen werden. Für die Teilnehmer bedeutsam ist der Nutzen, den sie aus der Beschäftigung mit kognitiven Aufgaben im Sinne eines Trainings ziehen können. Des weiteren ist der persönliche soziale Kontakt, der in der Untersuchungssituation entsteht, für viele der älteren Menschen eine willkommene Abwechslung zu dem ansonsten recht eintönigen Alltag. Es hat sich gezeigt, daß die meisten der Studienteilnehmer sehr gerne an den Untersuchungen teilgenommen haben. Sie wünschten sich sogar mehr Untersuchungszeitpunkte, als in dem Studiendesign vorgesehen sind. Dadurch, daß sich die Untersuchung über einen Zeitraum von zwei Jahren erstreckte, konnte in vielen Fällen eine persönliche Beziehung zu den Studienteilnehmern aufgebaut werden, die für beide Seiten sehr positiv war. Das geringe Risiko der Untersuchung besteht in eventuell auftretenden Verstimmungen oder Kränkungen der Probanden durch schlechte Testergebnisse oder persönliche Erinnerungen. Hier liegt die Verantwortung im Bereich der Befragung und der Testdurchführung bei dem Untersucher. Reagiert ein Studienteilnehmer auf Fragen zu persönlichen Belangen oder auf Probleme in einzelnen Leistungsbereichen verstimmt oder gekränkt, so wird dies vom Untersucher in angemessener Weise aufgefangen (Helmchen & Lauter, 1995): Dies kann bedeuten, daß über die Probleme geredet, eine Aufgabe abgebrochen und mit einer leichteren Aufgabe fortgefahren oder die Testsitzung beendet wird. In den meisten Fällen kann durch Beruhigen des Probanden die Stimmung deutlich gehoben und die Testung fortgeführt werden.

Die Studie wurde am 13.4.1999 von der Ethikkommission des Universitätsklinikums Benjamin Franklin der Freien Universität Berlin genehmigt.