

# Depressionsdiagnostik bei geriatrischen Patienten

*Die Entwicklung der Depression-im-Alter-Skala (DIA-S)*

Dissertationsschrift von Sonja Heidenblut

# Depressionsdiagnostik bei geriatrischen Patienten -

## Die Entwicklung der Depression-im-Alter-Skala (DIA-S)

Inauguraldissertation zur Erlangung des Doktorgrades der Humanwissenschaftlichen Fakultät  
der Universität zu Köln nach der Promotionsordnung vom 10.05.2010 vorgelegt

von Sonja Heidenblut aus Heidelberg, März 2012

Diese Dissertation wurde von der Humanwissenschaftlichen Fakultät der Universität zu Köln  
im September 2012 angenommen.

### **Förderung des Forschungsprojektes**

Der empirische Teil dieser Arbeit ist Teil des Forschungsprojektes LEANDERIII, das von Prof. Dr. Susanne Zank und Prof. Dr. Claudia Schacke geleitet und vom Bundesministerium für Familie, Senioren, Frauen und Jugend finanziert wurde.

## **Eidesstattliche Erklärung**

Ich versichere eidesstattlich, dass ich die von mir vorgelegte Dissertation selbständig und ohne unzulässige Hilfe angefertigt, die benutzten Quellen und Hilfsmittel vollständig angegeben und die Stellen der Arbeit einschließlich Tabellen, Karten und Abbildungen, die anderen Werken im Wortlaut oder dem Sinn nach entnommen sind, in jedem Einzelfall als Entlehnung kenntlich gemacht habe; dass diese Dissertation noch keinem anderen Fachbereich zur Prüfung vorgelegen hat; dass sie noch nicht veröffentlicht worden ist sowie dass ich eine solche Veröffentlichung vor Abschluss des Promotionsverfahrens nicht vornehmen werde. Die Promotionsordnung ist mir bekannt. Die von mir vorgelegte Dissertation ist von Prof. Dr. Susanne Zank betreut worden.

Der empirische Teil dieser Arbeit ist Teil des Forschungsprojektes LEANDERIII. Entsprechend sind Teile der Ergebnisse bereits im Zwischen- und Abschlussbericht des Projektes sowie in einer ersten Publikation (Heidenblut & Zank, 2010) veröffentlicht. Die Arbeit an den entsprechenden Teilen der Berichte sowie an dem Zeitschriftenartikel lag dabei in der alleinigen Verantwortung der Autorin dieser Dissertation.

## **Danksagung**

Ich danke Prof. Dr. Susanne Zank für die Ermöglichung und Betreuung dieses Promotionsprojektes.

Ich danke Prof. Dr. Claudia Schacke, Dr. Gernot Lämmler, Dr. Lissy Stechl, Stefan Mix und Andrea Mihail für ihren Einsatz bei der inhaltlichen Diskussion des Itempools.

Ich danke Prof. Dr. Steinhagen-Thiessen vom evangelischen Geriatriezentrum Berlin, Dr. Dr. Michael Rapp vom gerontopsychiatrischen Zentrum des St. Hedwigkrankenhauses sowie Prof. Dr. Hans Gutzmann der gerontopsychiatrischen Abteilung des Krankenhauses Hedwigshöhe für die Erlaubnis, die Vorerhebung in Ihren Einrichtungen durchführen zu dürfen.

Ich danke Dr. Xaver Sünkeler vom Herz-Jesu-Krankenhaus in Lindlar, Dr. Ingo Reinecke vom St. Marien Hospital in Köln und Dr. Thomas Heimig vom Krankenhaus St. Barbara in Attendorn für die Erlaubnis, den Hauptdatensatz in ihren Einrichtungen erheben zu dürfen. Ich danke den MitarbeiterInnen der Ergotherapie sowie der Sozialdienste der entsprechenden Kliniken für Ihre Unterstützung bei der Datenerhebung. Mein besonderer Dank gilt Lee Feis, die mir durch Ihre Begeisterung für das Projekt sowie durch Ihre tatkräftige Hilfe bei der Erhebung immer wieder Mut gemacht hat. Ich danke meinen studentischen Mitarbeitern Moritz Hielscher, Stefan Kohlbach, Tobias Knabe, Cornelia Hofmann und ganz besonders Denise Stein für Ihre fleißige und engagierte Mitarbeit bei der Datenerhebung.

Für moralische Unterstützung und Freundschaft danke ich Irene Fuchs, Constanze Steinhusen, Inka Wilhelm, Denise Stein, Katy Schleicher und Daniel Steinert.

<b>FÖRDERUNG DES FORSCHUNGSPROJEKTES</b>	<b>II</b>
<b>EIDESSTATTLICHE ERKLÄRUNG</b>	<b>III</b>
<b>DANKSAGUNG</b>	<b>IV</b>
<b>ZUSAMMENFASSUNG</b>	<b>2</b>
<b>1. EINLEITUNG</b>	<b>3</b>
1.1 Gliederung der vorliegenden Arbeit	5
<b>2. THEORETISCHER HINTERGRUND</b>	<b>7</b>
<b>2.1 Das Krankheitsbild depressiver Störungen</b>	<b>7</b>
2.1.1 Begriffsbestimmung und historische Entwicklung	7
2.1.2 Aktuelle Definition und Symptomatik	9
2.1.3 Erkrankungsformen und Klassifikation	11
2.1.3.1 Klassifikation der verschiedenen Krankheitsbilder nach der ICD-10	11
2.1.4 Epidemiologie	18
2.1.4.1 Erkrankungsverläufe und Prävalenzen	19
2.1.4.2 Prognosen und Behandelbarkeit	22
2.1.5 Ätiologie	23
2.1.5.1 Risikofaktoren	23
2.1.5.2 Ätiologische Risikofaktorenmodelle	26
2.1.5.3 Psychologische und Psychiatrische Erklärungsansätze	29
2.1.6 Versorgung	36
2.1.6.1 Behandlungsempfehlungen	36
2.1.6.2 Aktuelle Entwicklung	37
2.1.7 Zusammenfassung	38

<b>2.2 Depressive Störungen im Alter</b>	<b>40</b>
2.2.1 Begriffsbestimmung und historische Entwicklung	40
2.2.2 Aktuelle Definition und Klassifikation	40
2.2.3 Symptomatik	41
2.2.4 Epidemiologie	44
2.2.4.1 Prävalenzen	44
2.2.5 Erkrankungsverläufe und Prognosen	50
2.2.6 Ätiologie	51
2.2.6.1 Risikofaktoren	51
2.2.6.2 Psychologische und Psychiatrische Erklärungsansätze	54
2.2.7 Versorgung depressiver Störungen im Alter	65
2.2.7.1 Aktuelle Entwicklung	66
2.2.8 Therapie depressiver Störungen im Alter	67
2.2.8.1 Psychodynamische Therapie	67
2.2.8.2 Kognitiv -behaviorale Therapie	68
2.2.9 Zusammenfassung	71
<b>2.3 Die Diagnostik depressiver Störungen</b>	<b>73</b>
2.3.1 Diagnostik in der klinischen Psychologie	73
2.3.2 Der diagnostische Prozess	74
2.3.3 Standardisierte Testverfahren in der Depressionsdiagnostik	78
2.3.4 Depressivität als Messgröße	78
2.3.5 Eigenschaften verschiedener psychometrischer Testverfahren	79
2.3.5.1 Standardisierte Fragebögen	79
2.3.5.2 Halbstrukturierte Interviews	80
2.3.6 Gütekriterien psychometrischer Testverfahren	81
2.3.6.1 Objektivität	83
2.3.6.2 Reliabilität	84
2.3.6.3 Validität	85
2.3.6.4 Normierung	89
2.3.6.5 Ökonomie	90
2.3.6.6 Die Gewichtung von Gütekriterien bezogen auf die Zielsetzung eines Tests	90
2.3.7 Gütekriterien einzelner Items	91
2.3.7.1 Schwierigkeit	91
2.3.7.2 Trennschärfe	92
2.3.7.3 Itemvalidität	93
2.3.8 Zusammenfassung	93
<b>2.4. Die Diagnostik depressiver Störungen im Alter</b>	<b>95</b>
2.4.1 Phänomene des Alters als „Störvariablen“ der Depressionsdiagnostik	95
2.4.2 Screeningverfahren in der geriatrischen Praxis	98
2.4.2.1 Definition des Screening-Begriffes	98
2.4.2.2 Pro- und Contra-Argumente zu flächendeckendem Screening	100
2.4.3 Die Entwicklung der Geriatrischen Depressionskala (GDS)	100
2.4.4 Depressionsscreening bei multimorbiden Patienten mit der GDS	107
2.4.4.1 Schwierigkeiten unter theoretischen Gesichtspunkten	108
2.4.4.2 Empirische Befunde	109
2.4.5 Kürzere Versionen der GDS	115
2.4.6 Zusammenfassung	115

<b>3. FRAGESTELLUNG UND HYPOTHESEN</b>	<b>118</b>
Hypothese 01	119
Hypothese 02	119
Hypothese 03	119
Hypothese 03a	119
Hypothese 03b	119
Hypothese 03c	119
<b>4. METHODEN</b>	<b>120</b>
<b>4.1 Die Entwicklung der Depression- im -Alter- Skala (DIA-S)</b>	<b>120</b>
4.1.1 Zielsetzung	120
4.1.2 Vorgehen	120
4.1.3 Zugrundeliegende Testtheorie	122
4.1.4 Erstellung des Itempools	122
4.1.4.1 Formale Itemgestaltung.	122
4.1.4.2 Inhaltliche Itemgestaltung	123
4.1.5 Pilotstudie: empirische Prüfung des Itempools	126
4.1.5.1 Rekrutierung und Untersuchungsablauf	126
4.1.5.2 Stichprobe	127
4.1.5.3 Itemanalysen	128
4.1.6 Zusammenstellung der Skala	132
4.1.6.1 Angestrebte Anzahl an Items	132
4.1.6.2 Auswahl der Items	132
4.1.6.3 Polung der Items	134
4.1.6.4 Anordnung der Skala	135
<b>4.2 Vergleichende Validierung der DIA-S und der GDS15</b>	<b>135</b>
4.2.1 Rekrutierung	135
4.2.2 Untersuchungsablauf	136
4.2.3 Messinstrumente	138
4.2.4 Stichprobe	140
4.2.5 Auswertungsstrategien	143
4.2.5.1 Umgang mit fehlenden Werten	143
4.2.5.2 Analyse der Testeffektivität	143
4.2.5.3 Einflussfaktoren auf die Testeffektivität	145

<b>5. ERGEBNISSE</b>	<b>147</b>
<b>5.1 Itemanalysen</b>	<b>147</b>
<b>5.2 Skalenanalysen</b>	<b>150</b>
5.2.1 Reliabilität	150
5.2.2 Validität	151
5.2.2.1 Konvergente Validität	152
5.2.2.2 Testeffektivität	153
5.2.2.3 Testeffektivität in einzelnen Trennwerten	154
5.2.3 Einflussfaktoren auf die Testvalidität	156
5.2.3.1 Umgang mit fehlenden Werten	156
5.2.3.2 Erstellung der Regressionsmodelle	157
5.2.3.3 Überprüfung der Multikolarität	157
5.2.3.4 Ergebnisse der Regressionsanalysen: Fehlklassifikation durch die GDS15	158
5.2.3.5 Ergebnisse der Regressionsanalysen: Fehlklassifikation durch die DIA-S	161
5.2.3.6 Überprüfung der Modellrelevanz anhand bivariater Zusammenhänge	162
5.2.4 Zusammenfassung der Ergebnisse und Hypothesenprüfung	170
<b>6. DISKUSSION</b>	<b>172</b>
<b>6.1 Interpretation der Ergebnisse</b>	<b>172</b>
6.1.1 Itemanalysen	172
6.1.2 Skalenreliabilität	174
6.1.3 Testeffektivität	176
6.1.3.1 Geschlechtsunterschiede in der Testeffektivität	179
6.1.4 Konvergente Validität	180
<b>6.2 Überlegungen zur Anwendung der Skala</b>	<b>180</b>
6.2.1 Klinische Anwendbarkeit der Skala.	180
6.2.2 Stellenwert des Instrumentes im diagnostischen Prozess.	181
<b>6.3 Grenzen der vorliegenden Arbeit und Empfehlungen für künftige Untersuchungen</b>	<b>181</b>
6.3.1 Entwicklung des Instrumentes	181
6.3.2 Einflussfaktoren auf die Validität der Testverfahren	183
6.3.4 Neutralität der Untersuchungssituation	184
6.3.4 Setting der Untersuchungssituation	185
<b>6.4 Ergänzende Überlegungen zur Depressionsdiagnostik</b>	<b>185</b>
6.4.1 Das Recht eines Patienten, die Testung zu verweigern	186
6.4.2 Das Verhältnis von Diagnostik und Behandlung	188
<b>6.5 Ausblick für Forschung und Praxis</b>	<b>188</b>
<b>LITERATUR</b>	<b>190</b>
<b>ANHANG</b>	

## ABBILDUNGSVERZEICHNIS

Abbildung 1: Das biopsychosoziale Krankheitsmodell .....	28
Abbildung 2: Das Diathese-Stress-Modell .....	29
Abbildung 3: Depressive Attributionsschemata .....	33
Abbildung 4: Die Theorie der erlernten Hilflosigkeit/Hoffnungslosigkeit.....	34
Abbildung 5: Entwicklung durch Akkomodation und Assimilation.....	57
Abbildung 6: Vereinfachte Darstellung des diagnostischen Prozesses .....	75
Abbildung 7: Entscheidungsbaum zur Differentialdiagnostik affektiver Störungen .....	77
Abbildung 8: Die kategoriale Diagnostik depressiver Störungen durch Trennwerte.....	87
Abbildung 9: Testeffektivität als Vergleich zweier dichotomisierter Merkmale.....	88
Abbildung 10: Die Geriatriische Depressionsskala nach Yesavage et al. (1983).....	103
Abbildung 11: Erhebungsbogen der GDS15 in der Übersetzung nach AGAST (1995) .....	108
Abbildung 12: Schritte der Skalenentwicklung im Überblick.....	121
Abbildung 13: monoton verlaufendes Lösungsmuster des Items „Ich habe Angst vor der Zukunft“ .....	129
Abbildung 14: Vergleich der Lösungsmuster kritischer Items (1) .....	133
Abbildung 15: Vergleich der Lösungsmuster kritischer Items (2) .....	134
Abbildung 16: Untersuchungsablauf der Validierungsstudie .....	137
Abbildung 17: Vier – Felder - Schema zur Ermittlung der Testeffektivität.....	144
Abbildung 18: Darstellung der Testeffektivität als Fläche unter der ROC-Kurve.....	144
Abbildung 19: Vergleich der ROC-Kurven von DIA-S und GDS.....	153
Abbildung 20: Vergleich der ROC-Kurven von DIA-S und GDS15 nach Geschlecht.....	154

## FORMELVERZEICHNIS

Formel (1): die Spearman-Brown-Formel.....	175
Formel (2), (3): Die Anwendung des Base-Theorems auf negative und positive predictive values .....	177

## TABELLENVERZEICHNIS

<i>Tabelle 1: Die Klassifikation depressiver Störungsbilder nach der ICD-10 (WHO, 2011).....</i>	<i>12</i>
<i>Tabelle 2: Haupt- und Nebensymptome depressiver Störungen nach der ICD-10 (WHO,2011) .....</i>	<i>14</i>
<i>Tabelle 3: 12 Monatsprävalenz depressiver Verstimmungen im weltweiten Vergleich.....</i>	<i>20</i>
<i>Tabelle 4: Prävalenzraten von depressiven Störungen bei stationären geriatrischen Patienten .....</i>	<i>47</i>
<i>Tabelle 5: Qualitätsmerkmale der vorgestellten Prävalenzstudien.....</i>	<i>49</i>
<i>Tabelle 5: Häufige Abwehrmechanismen des höheren Lebensalters .....</i>	<i>59</i>
<i>Tabelle 7: Beispiele für dysfunktionale Denkmuster .....</i>	<i>69</i>
<i>Tabelle 7: Fragetechniken zur Identifikation und Modifikation dysfunktionaler Gedanken .....</i>	<i>70</i>
<i>Tabelle 8: Fehlerquellen von Selbst- und Fremdbeurteilungsverfahren .....</i>	<i>81</i>
<i>Tabelle 9 : Kriterien zur Testbeurteilung des Testkuratoriums deutscher Psychologen-verbände.....</i>	<i>82</i>
<i>Tabelle 10: Differentialdiagnostische Unterscheidung zwischen Demenz und Depression.....</i>	<i>97</i>
<i>Tabelle 11: Vor – und Nachteile von Flächendeckendem Screening .....</i>	<i>100</i>
<i>Tabelle 12: Entwicklungsstudien zur Geriatrischen Depressionsskala .....</i>	<i>105</i>
<i>Tabelle 13: Zusammenfassung der Ergebnisse aus Validierungsstudien der GDS .....</i>	<i>111</i>
<i>Tabelle 14: Itempool für die Pilotstudie (N=31) – systematisch nach Symptomen geordnet.....</i>	<i>125</i>
<i>Tabelle 15 : Verteilung der Erkrankungsschweregrade der Stichprobe nach Einrichtungsform .....</i>	<i>127</i>
<i>Tabelle 16: weitere Stichprobenmerkmale.....</i>	<i>128</i>
<i>Tabelle 17: Itemkennwerte der Voruntersuchung, nach Symptombereichen sortiert.....</i>	<i>130</i>
<i>Tabelle 18 : Erhebungsinstrument für die Validierungsstudie.....</i>	<i>135</i>
<i>Tabelle 19: Stichprobenmerkmale des Hauptdatensatzes .....</i>	<i>142</i>
<i>Tabelle 20: Itemkennwerte des Hauptdatensatzes.....</i>	<i>149</i>
<i>Tabelle 21: Interne Konsistenz der Skalen .....</i>	<i>151</i>
<i>Tabelle 22: kriteriumsbezogene Validität der Skalen .....</i>	<i>152</i>
<i>Tabelle 23: Vergleich von Sensitivität und Spezifität für die effektivsten Trennwerte der Skalen.....</i>	<i>155</i>
<i>Tabelle 24: Interkorrelationen der Prädiktoren für die hierarchische logistische Regression .....</i>	<i>158</i>
<i>Tabelle 25: Hierarchisch logistische Regression zur Erklärung von Fehlklassifizierungen durch die GDS15 .....</i>	<i>159</i>
<i>Tabelle 26: Hierarchisch logistische Regression zur Erklärung von Fehlklassifizierungen durch die DIA-S .....</i>	<i>160</i>
<i>Tabelle 27: Bivariate Zusammenhänge zwischen Fehlklassifikationen von DIA-S und GDS15.....</i>	<i>163</i>
<i>Tabelle 28: Logistische Regression zur Erklärung von falsch- positiv- Klassifizierungen durch die GDS15.....</i>	<i>166</i>
<i>Tabelle 29: Logistische Regression zur Erklärung von falsch- positiv- Klassifizierungen durch die DIA-S.....</i>	<i>167</i>
<i>Tabelle 30: Logistische Regression zur Erklärung von falsch- negativ- Klassifizierungen durch die GDS15 .....</i>	<i>168</i>
<i>Tabelle 31: Logistische Regression zur Erklärung von falsch- negativ- Klassifizierungen durch die DIA-S .....</i>	<i>169</i>

## **Zusammenfassung**

Die vorliegende Arbeit stellt die Entwicklung und Validierung der Depression-im-Alter-Skala (DIA-S) vor. Die DIA-S ist ein kurzes Depressionsscreening, das in Durchführung und Auswertung ähnlich praktikabel und ökonomisch ist wie die Kurzversion der Geriatrischen Depressionsskala (GDS15), das jedoch inhaltlich besser auf den Einsatz in der Geriatrie zugeschnitten ist.

Die DIA-S und die GDS15 wurden an einer Stichprobe von N=331 geriatrischen Patienten (n=151 depressive Probanden und n=180 nicht depressive Probanden) validiert, wobei die Montgomery-Asberg Depressionsskala (MADRS) als Goldstandard-Kriterium zur Einschätzung der Depressivität verwendet wurde. Die Datenauswertung umfasst Itemanalysen (fehlende Werte, Schwierigkeiten, Trennschärfen) und Skalenanalysen (Interne Konsistenzen, Korrelationen mit den MADRS-Werten, Testeffektivitäten) sowie logistische Regressionsanalysen bezüglich verschiedener Einflussfaktoren auf die Testqualität.

Alle Ergebnisse sprechen für eine gute diagnostische Qualität der DIA-S. Dabei übertrifft die neue Skala die GDS15 vor allem in der Trennschärfe der Items, in der internen Konsistenz und in der Testspezifität.

## 1. Einleitung

Depressionen zählen zu den häufigsten psychiatrischen Erkrankungen. In der älteren Bevölkerung sind sie stark mit körperlich-funktionalen Einschränkungen und einem schlechten physischen Allgemeinzustand assoziiert. Entsprechend sind geriatrische Akutpatienten mit Prävalenzraten für Depressionen von bis zu 40% besonders gefährdet (McCusker et al., 2005). Depressionen sind unter geriatrischen Patienten nicht nur besonders verbreitet, sie haben auch einen negativen Einfluss auf den funktionalen Status der Patienten und führen zu schlechteren Rehabilitationsergebnissen, stärkeren kognitiven Einschränkungen, früherer Institutionalisierung und sogar zu früherer Mortalität (McCusker, Cole, Ciampi, Latimer, Windholz, & Belzile, 2007; Lasser, Siegel, Dukoff, & Sunderland, 1998; Schmeling-Kludas, Jäger & Niemann, 2000). Dennoch sind die Erkennungsraten depressiver Störungen in allgemeinmedizinischen Versorgungseinrichtungen gering. So zeigt sich bei leichter Depression ein Anteil von 28% korrekt diagnostizierter Patienten, während Patienten mit schwererer Symptomatik etwa zu 56% richtig erkannt werden (McCusker, Cole, Ciampi, Belzile, Cepoiu, & Sewitch, 2008).

Das Defizit an ausreichender Versorgung mag dabei teilweise an dem vorwiegend negativen Altersbild unserer Gesellschaft liegen, das depressive Symptome bei älteren Menschen vorzugsweise unter einem zwangsläufigen Stimmungsabfall im Alter subsummiert. Eine andere Ursache, die für die klinische Praxis von direkterer Bedeutung ist, hängt jedoch mit der Schwierigkeit zusammen, depressive Erkrankungen bei älteren Menschen zu diagnostizieren (Heuft, 1992; Schaller & Göbel, 1988). So treten beispielsweise somatische Symptome wie Schlaflosigkeit und Appetitmangel, die bei jüngeren Patienten wichtige Identifikationskriterien sein können, auch bei einem großen Teil der nicht an Depressionen erkrankten älteren Bevölkerung auf. Als weitere Gründe nennen Gauggel und Birkner (1998) verminderte Kooperationsfähigkeit aufgrund sensorischer und kognitiver Defizite,

Symptomüberlappungen mit demenziellen Erkrankungen, Multimedikation und daraus resultierende Nebenwirkungen, sowie ein stärkeres Misstrauen älterer Menschen gegenüber psychologischen Untersuchungen. Auch die Heterogenität der Zielgruppe erschwert es, allen Patienten durch dasselbe Instrument gleichermaßen gerecht zu werden. So kann ein sehr einfaches Testverfahren von einem Patienten mit einem hohen kognitiven Funktionsniveau abgelehnt werden, weil derjenige die Fragen als nicht differenziert genug empfindet, während dasselbe Verfahren für einen Patienten mit kognitiver Einschränkung bereits eine Überforderung darstellt.

Einen Lösungsversuch stellt die Geriatriische Depressionsskala dar, die sich hierzulande als Messinstrument für depressive Symptome bei geriatrischen Patienten durchgesetzt hat (Yesavage et al., 1983; Gauggel & Birkner, 1998; AGAST, 1997). Der Fragebogen besteht aus 30 bzw. in der Kurzversion 15 Fragen zu depressiven Störungen, die jeweils mit „ja“ oder „nein“ zu beantworten sind. Aufgrund seiner Einfachheit stellt dieses Screeningverfahren aktuell das gebräuchlichste Instrument zur Depressionsfrüherkennung bei älteren Patienten dar und wird sowohl im stationären als auch im ambulanten Setting eingesetzt. Dabei ist eine von der Forschungsgruppe AGAST (1997) vorgestellte deutsche Übersetzung der Skala in zahlreichen Kliniken das Standardinstrument zum Screening depressiver Symptome im deutschsprachigen Raum.

Die Durchführung der GDS ist jedoch gerade bei multimorbiden Patienten mitunter problematisch und führt häufig zu Unverständnis auf Seiten der Befragten. So richtet sich die Skala durch die Art und Weise, wie die Symptome durch die Fragen operationalisiert werden implizit an eine Subgruppe körperlich gesunder, kognitiv kaum beeinträchtigter älterer Menschen. Der weitaus größere Anteil des Klientels, das in der Praxis untersucht wird, besteht jedoch aus multimorbiden und /oder kognitiv eingeschränkten Patienten, deren Situation durch den Fragebogen nicht angemessen berücksichtigt wird (vgl. Heidenblut, 2005). Gerade für diesen immer größer werdenden Anteil an Patienten wäre jedoch ein

geeignetes Messinstrument wünschenswert, um eine optimale Versorgung zu gewährleisten, und somit einen ganzheitlichen Heilungsprozess zu begünstigen.

Die Ergebnisse einer Pilotstudie (Heidenblut & Zank, 2005) weisen darauf hin, dass bereits eine geringe Modifikation einiger Items die Geriatrischen Depressionsskala hinsichtlich wichtiger psychometrischer Gütekriterien positiv beeinflussen und für die Zielgruppe multimorbider Patienten kompatibel machen kann. Die Items wurden dabei dahingehend verändert, dass sie entweder im Hinblick auf die äußere Situation der Probanden neutral formuliert wurden, oder dass sie die spezielle Situation des stationären Settings berücksichtigten. Von den Ergebnissen dieser Pilotstudie ausgehend wurde in der vorliegenden Untersuchung die Depression-im-Alter-Skala (DIA-S) als Screeningverfahren entwickelt, das für die Zielgruppe multimorbider geriatrischer Patienten alternativ zur Geriatrischen Depressionsskala eingesetzt werden kann.

### *1.1 Gliederung der vorliegenden Arbeit*

Kapitel 2 der vorliegenden Arbeit beschäftigt sich mit den theoretischen Voraussetzungen der psychometrischen Diagnostik depressiver Störungen im Alter. In Abschnitt 2.1.1 wird zunächst kurz auf die Begriffsdefinition des Störungsbildes sowie auf die historische Entwicklung der Krankheitsauffassung eingegangen. Die Punkte 2.1.2 bis 2.1.6 beschäftigen sich schließlich mit dem aktuellen Wissensstand zu depressiven Erkrankungen im Erwachsenenalter. Kapitel 2.2 geht auf die Besonderheiten von depressiven Störungen im höheren Lebensalter ein. Teil 2.3 beschäftigt sich mit der Diagnostik depressiver Störungen unter besonderer Berücksichtigung der Vorteile und Beurteilungskriterien standardisierter psychometrischer Testverfahren. In Abschnitt 2.4 wird schließlich auf Phänomene des Alterns hingewiesen, die für die psychometrische Diagnostik depressiver Störungen sowie für darüber hinausgehende sogenannte Screeningprogramme relevant sind. Anhand der Geriatrischen Depressionsskala wird aufgezeigt, wie solche Phänomene bei der Konstruktion eines

Testverfahrens berücksichtigt werden können, wobei in Punkt 2.3.3 schließlich auf die Grenzen des Verfahrens hinsichtlich bestimmter Zielgruppen hingewiesen wird.

Anschließend werden die Fragestellungen und Hypothesen des empirischen Teils dieser Arbeit vorgestellt (Kapitel 3).

Kapitel 4, der Methodenteil dieser Arbeit, stellt zunächst die Entwicklung der Depression im Alter-Skala (DIA-S) vor, wobei einerseits auf die theoretischen Voraussetzungen und andererseits auf die Bedingungen des intendierten Erhebungssettings eingegangen wird. Anschließend werden die Untersuchungs- und Auswertungsplanung für eine vergleichende Validierung der DIA-S und der GDS15 vorgestellt.

Kapitel 5 enthält die Itemanalysen sowie die Analysen zu den Beschaffenheiten der Gesamtskalen. Dabei werden neben den gängigen Daten zur Reliabilität und Validität auch Störanfälligkeiten der Skalen im Hinblick auf bestimmte Eigenschaften der Stichprobe sowie des Erhebungssettings diskutiert.

Die Diskussion dieser Arbeit (Kapitel 6) enthält neben einer Diskussion der Ergebnisse im engeren Sinne auch Überlegungen über wissenschaftliche und klinische Rahmenbedingungen der Diagnostik und Behandlung depressiver Störungen im höheren Erwachsenenalter.

## 2. Theoretischer Hintergrund

### 2.1 Das Krankheitsbild depressiver Störungen

#### 2.1.1 Begriffsbestimmung und historische Entwicklung

Historiker gehen davon aus, dass das Auftreten psychischer Störungen die Menschheit seit ihren Ursprüngen begleitet. So deuten Schädelknochen aus prähistorischer Zeit darauf hin, dass bereits in sehr frühen Gesellschaften Versuche unternommen wurden, auffälliges oder krankhaftes Verhalten durch primitive Operationen am Gehirn, sogenannte Trepanationen, zu behandeln (Comer, 2001). Ein schriftliches Zeugnis über das Phänomen der depressiven Verstimmung findet sich erstmalig in den ältesten Schriften des *corpus hippocraticum*, die auf das Ende des 5. Jahrhunderts v. Chr. datiert werden. Seit dieser ersten Erwähnung findet sich in den wissenschaftlichen Schriften eine über die Zeit hinweg verhältnismäßig einheitliche Beschreibung der Krankheitssymptomatik; Vorstellungen über Ursachen des Leidens sowie über adäquate Behandlungsansätze sind dagegen stark durch das jeweils zugrunde liegende Menschenbild geprägt, d.h., diese sind einerseits einem historischen Wandel unterzogen, andererseits findet sich aber auch ein Nebeneinander divergierender Vorstellungen über die verschiedenen Epochen hinweg (Schott & Tölle, 2006). Im Folgenden wird in einem kurzen historischen Überblick auf die jeweils einflussreichsten Strömungen eingegangen.

In der Antike wurde die depressive Verstimmung im Rahmen einer allgemeinen Krankheitslehre auf somatische Ursachen zurückgeführt. Die Humoralpathologie oder Viersäftelehre geht davon aus, dass Gesundheit oder Krankheit des Menschen auf das Gleich- bzw. Ungleichgewicht der vier Körpersäfte, Blut, Schleim sowie gelbe und schwarze Galle zurückzuführen sind. Ausgehend von dieser Lehre bedeutet der Begriff Melancholie auch in der wörtlichen Übersetzung „schwarze Galle“ (Schott & Tölle, 2006). Die in diesem Zusammenhang einflussreichste Theorie stammt von Galen, der drei mögliche Ursachen einer Affizierung des Gehirns durch zu viel schwarze Galle unterscheidet. Durch ein generelles

Übermaß an schwarzer Galle im Blut kann diese ins Gehirn aufsteigen oder Schwarze Galle im Gehirn kann aufgrund von örtlicher Überhitzung durch gelbe Galle entstehen. Das Leiden kann aber auch von einer Verdauungsstörung ausgehen, bei der eine schwarzgallige, rußartige Ausdünstung aus dem Oberbauch ins Gehirn aufsteigt. Zur Therapie wurden entsprechend körperliche Maßnahmen wie Bäder, Aderlass oder Diäten empfohlen.

So bizarr entsprechende Theorien aus heutiger Sicht erscheinen mögen, so machen sie doch deutlich, dass bereits frühe Auseinandersetzungen mit der Melancholie einerseits eine Verschränkung zwischen körperlichen und seelischen Aspekten der Erkrankung annahmen und dass andererseits verschiedene mögliche Ursachen und Formen derselben Symptome berücksichtigt wurden. Das auf Galen beruhende vorwiegend somatische Modell der Melancholie prägte die medizinische Sichtweise der Erkrankung bis in die frühe Neuzeit. Dabei wurden die rein somatischen Vorstellungen während des Mittelalters durch metaphysisch anmutende Thesen ergänzt. So nahm man an, dass das Zuviel an schwarzer Galle die Abwehrkräfte gegen Dämonen schwäche, bzw. dass sie den im Gehirn ansässigen spiritus animalis, also den Lebensgeist, umkehre (vgl. Schott & Tölle, 2006).

Mit der Entstehungszeit der modernen Psychiatrie wurden erstmalig psychologische bzw. tiefenpsychologische Ansätze mit einbezogen um die Entstehung der Melancholie zu erklären. So beschreibt beispielsweise bereits Pinel zu Beginn des 19. Jahrhunderts zahlreiche, auf unterschiedlichen Ebenen angesiedelte Symptome und mögliche Ursachen der Erkrankung:

Traurigkeit, Schrecken, anhaltendes Studieren, die Unterbrechung eines thätigen Lebens, heftige Liebe, das Uebermaß in den Vergnügungen, Missbrauch betäubender und narkotischer Mittel, vorhergehende Krankheiten, die unrichtig behandelt wurden, die Unterdrückung des Hamorrhoidalflusses...“

(Pinel, 1800, zitiert nach Schott & Tölle, 2006, S.12)

Im Verlauf des 19. Jahrhunderts wurde der Begriff Melancholie als wissenschaftlicher Krankheitsbegriff allmählich durch den auch heute noch gebräuchlichen Terminus Depression ersetzt, da der Melancholie-Begriff durch zahlreiche damit verbundene Theorien und

Erscheinungsbilder in Kunst und Literatur im Laufe der Zeit immer mehr verwässert worden war. (vgl. Schott & Tölle, 2006).

### *2.1.2 Aktuelle Definition und Symptomatik*

Depressionen werden innerhalb der Klinischen Psychologie (Hautzinger, 1998) definiert als „psychische Störungen“, bei denen „Traurigkeit, Niedergeschlagenheit, Verstimmung, Energielosigkeit, Antriebsminderung, Selbstzweifel, Wertlosigkeit, Hoffnungslosigkeit, Sinnlosigkeit, häufig begleitet von Ängstlichkeit und Unruhe, Appetitstörungen, Gewichtsverlust, Libidoverlust, Schlafstörungen, Konzentrationsprobleme und Suizidideen ... typische Beschwerden und Auffälligkeiten“ sind (S.3).

Die hier aufgezählten Merkmale beziehen sich nicht allein auf eine traurige oder niedergeschlagene Stimmung, wie sie in der Alltagsbedeutung des Begriffes „Depression“ häufig gemeint ist, sondern sie sind vielgestaltig und betreffen nahezu alle Bereiche des menschlichen Erlebens und Verhaltens. Man unterscheidet zwischen Symptomen emotionaler, motivationaler, behavioraler, kognitiver und somatischer Art (Comer, 1995, 2008). Die meisten Beschwerden, die für depressive Störungen charakteristisch sind, sind dabei für sich genommen nicht pathologisch; bei vielen Menschen treten sie im Rahmen des alltäglichen Gefühlserlebens hin und wieder auf und ermöglichen einen angemessenen Umgang mit belastenden Situationen. Eine krankheitswertige Depression ist daher nicht gleichbedeutend mit dem Vorliegen einzelner Symptome; entscheidend sind vielmehr die Dauer und Intensität des negativen Erlebens, die Bandbreite der auftretenden Beschwerden sowie die Beeinflussbarkeit der Stimmung durch äußere Umstände. So sind Gefühle der Niedergeschlagenheit oder Hoffnungslosigkeit bei depressiven Erkrankungen nicht als Bestandteil gesunder emotionaler Schwingungsfähigkeit nachvollziehbar, sondern sie werden zu einem dauerhaften Zustand psychischen Leidens, der die Wahrnehmung des Betroffenen verzerrt und überschattet (Comer, 1995, 2008). Der Erkrankte kann kaum Freude empfinden, sich für nichts begeistern und er fühlt sich energie- und wertlos. Die Zukunft wird als

Bedrohung wahrgenommen und Gedanken an die Vergangenheit sind mit Schuldgefühlen behaftet. Häufig treten Zustände der Anspannung und des Unwohlseins auf, die durch körperliche Erkrankungen nicht zu erklären sind. Die Konzentrationsfähigkeit ist gestört. Das Leben erscheint unerträglich und die Anforderungen des Alltags werden als unerfüllbar erlebt. Lebensmüdigkeit und Suizidgedanken sind häufig. Das „normale“ Erleben und Verhalten ist durch eine anhaltende affektive Verstimmung beeinträchtigt. In der Klinischen Psychologie zählt man depressive Erkrankungen daher auch zu den affektiven Störungen.

### *2.1.3 Erkrankungsformen und Klassifikation*

Obwohl sich die wesentlichen Merkmale depressiver Erkrankungen relativ einheitlich beschreiben lassen, unterscheiden sich verschiedene Formen der Störung durch ihren Erkrankungsverlauf, das zugrundeliegende Ursachen- und Behandlungswissen sowie durch ihre Verbreitung in der Population. Um allgemein gültige Diagnosen zu gewährleisten, werden Klassifikationssysteme verwendet, die beruhend auf dem aktuellen wissenschaftlichen Konsens eine deskriptive Zuordnung der verschiedenen Krankheitsbilder ermöglichen. Allgemein anerkannt sind das amerikanische „Diagnostische und Statistische Manual psychischer Störungen“ (DSM – IV), sowie die 10. Überarbeitung der durch die Weltgesundheitsorganisation (WHO) entwickelten klinisch-diagnostischen Leitlinien zur Klassifikation psychischer Störungen (ICD-10, Kapitel V) (Hautzinger, 2002). Da letztere das hierzulande verbindliche Klassifikationsinstrument darstellt und sich die Systeme in den wesentlichen Punkten überschneiden, orientiert sich die folgende Beschreibung der Krankheitsbilder lediglich an der ICD-10.

#### *2.1.3.1 Klassifikation der verschiedenen Krankheitsbilder nach der ICD-10*

Die Krankheitsbilder mit depressiver Symptomatik, die in der ICD-10 beschrieben werden, sind hauptsächlich in Kapitel F3 unter den affektiven Störungen kodiert. Einige Ausnahmen finden sich unter anderen Kapiteln (siehe Tabelle 1).

*Tabelle 1: Die Klassifikation depressiver Störungsbilder nach der ICD-10 (WHO, 2011)*

Kapitel	Bezeichnung	Verschlüsselung
F3: Affektive Störungen	bipolare affektive Störung	F31
	depressive Episode	F32
	rezidivierende depressive Störung	F33
	anhaltende affektive Störungen	F34
	andere affektive Störungen	F38
	nicht näher bezeichnete affektive Störung	F39
F4: Reaktion auf schwere Belastung und Anpassungsstörungen	kurze depressive Reaktion	F43.20
	längere depressive Reaktion	F43.21
F6: Andere psychische Störungen aufgrund einer Schädigung oder Funktionsstörung des Gehirns oder einer körperlichen Krankheit	organische bipolare Störung	F06.31
	organische depressive Störung	F06.32
	organische gemischte affektive Störung	F06.33
F2: Schizophrenie, schizotype und wahnhaftige Störungen	postschizophrene Depression	F20.4

Die Einteilung in die verschiedenen Störungsgruppen geschieht nach Richtlinien, die die Dauer und Schwere der Erkrankung, die psychopathologische Krankengeschichte des Patienten und gegebenenfalls die Auslösung der Störung durch ein vorangehendes kritisches Lebensereignis berücksichtigen. Zudem ist bei der Diagnostik zu beachten, ob das depressive Syndrom ursächlich mit einem einzelnen körperlichen Faktor zusammenhängt oder substanzinduziert ist, da das Störungsbild in diesem Fall nicht als affektive Störung kodiert wird und eher über eine Beseitigung der ursächlichen Faktoren behandelt wird. Eine weitere Ausnahme stellt die postschizophrene Depression dar, die dem Krankheitsbild nach die Kriterien einer affektiven Störung erfüllt, aber aufgrund ihrer besonderen Verlaufsbedingung dem Kapitel 2 zugeordnet wird, da sie wahrscheinlich ursächlich mit der vorangegangenen Psychose in Verbindung steht (vgl. WHO, 2011). Bei Störungsformen, die das Kriterium einer depressiven Episode erfüllen, kann eine Einteilung in die Schweregrade leicht, mittelschwer und schwer erfolgen. Alle Störungsgruppen setzen eine bestimmte Anzahl und

Zusammensetzung aus einer festgelegten Menge von 10 Symptomen voraus, von denen die ersten drei als Hauptsymptome und die übrigen als Nebensymptome gelten (siehe Tabelle 2).

Tabelle 2: Haupt- und Nebensymptome depressiver Störungen nach der ICD-10 (WHO, 2011)

Hauptsymptome	Nebensymptome	Somatisches Syndrom
1. depressive Stimmung, in einem für die Betroffenen deutlich ungewöhnlichen Ausmaß, die meiste Zeit des Tages, fast jeden Tag, im Wesentlichen unbeeinflusst von den Umständen und mindestens zwei Wochen andauernd	1. Verlust des Selbstwertvertrauens und des Selbstwertgefühles  2. unbegründete Selbstvorwürfe oder ausgeprägte, unangemessene Schuldgefühle  3. wiederkehrende Gedanken an den Tod oder an Suizid oder suizidales Verhalten	1. deutlicher Interessenverlust oder Verlust der Freude an normalerweise angenehmen Aktivitäten  2. mangelnde Fähigkeit, auf Ereignisse oder Aktivitäten emotional zu reagieren, auf die normalerweise reagiert würde  3. Früherwachen, zwei Stunden oder mehr, vor der gewohnten Zeit
2. Interessen- oder Freudeverlust an Aktivitäten, die normalerweise angenehm waren	4. Klagen über oder Nachweis eines verminderten Denk- oder Konzentrationsvermögens, Unschlüssigkeit oder Unentschlossenheit (subjektiv oder objektiv)	4. Morgentief  5. objektiver Befund einer ausgeprägten psychomotorischen Hemmung oder Agitiertheit (beobachtet oder von anderen berichtet)
3. verminderter Antrieb oder gesteigerte Ermüdbarkeit	5. psychomotorische Agitiertheit oder Hemmung (subjektiv oder objektiv)  6. Schlafstörungen jeder Art  7. Appetitverlust oder gesteigerter Appetit mit entsprechender Gewichtsveränderung	6. deutlicher Appetitverlust  7. Gewichtsverlust (5% oder mehr des Körpergewichts im vergangenen Monat)  8. deutlicher Libidoverlust

*Depressive Episode.* Um das Kriterium für eine depressive Episode zu erfüllen, sollten die Symptome, unter denen der Patient leidet, mindestens zwei Wochen andauern. Die Schweregradeinteilung der Erkrankung erfolgt nach Anzahl und Ausprägung der Symptome. Die Diagnose einer leichten Episode setzt das Vorhandensein von mindestens zwei Hauptsymptomen und zusätzlich zwei weiteren Symptomen voraus. Dabei wird davon ausgegangen, dass der Patient bei der Aufrechterhaltung seines Alltags zwar Mühe hat, die meisten seiner Aktivitäten jedoch fortsetzen kann.

Bei der mittelschweren Episode sollten zwei Hauptsymptome und zusätzlich mindestens drei, besser jedoch vier der anderen Symptome vorliegen. Dabei sollten entweder einige der Symptome stärker ausgeprägt sein oder es sollte eine weite Bandbreite von Symptomen feststellbar sein. Die Fähigkeit des Patienten, seinen alltäglichen Anforderungen nachzukommen sowie berufliche und soziale Rollen auszufüllen sind bei einer mittelschweren Episode erheblich beeinträchtigt.

Im Fall einer schweren Episode wird davon ausgegangen, dass alle drei Hauptsymptome sowie mindestens vier weitere Symptome bei dem Patienten festzustellen sind, von denen einige besonders stark ausgeprägt sein sollten. In der Regel ist der Patient bei diesem Schweregrad nicht in der Lage, soziale, berufliche und häusliche Aktivitäten fortzusetzen.

*Psychotische Symptome.* Bei schweren depressiven Episoden können möglicherweise psychotische Symptome eintreten. Diese können Wahnideen, gewöhnlich Vorstellungen der Versündigung oder Verarmung, akustische oder Geruchshalluzinationen sowie eine schwere psychomotorische Hemmung, bis hin zu einem sogenannten depressiven Stupor sein.

*Somatisches Syndrom.* Unabhängig von den diagnostischen Informationen, die zur Klassifizierung der Störung dienen, kann bei depressiven Episoden das Vorhandensein eines sogenannten „somatischen Syndroms“ berücksichtigt werden. Dieses Syndrom bezeichnet eine Reihe gemeinsam auftretender Symptome, die für bestimmte Erkrankungsformen, früher

als endogene Depressionen bezeichnet, typisch sind (siehe Tabelle 2). Wie Tabelle 2 zu entnehmen ist, überschneiden sich einige Symptome des somatischen Syndroms dabei erheblich mit den allgemeinen Haupt- und Nebensymptomen depressiver Erkrankungen. Entsprechend ist das somatische Syndrom bei schwereren Erkrankungsformen schon per definitionem vorhanden.

Da die Abgrenzung einer als endogene Depression bezeichneten Krankheitsform nicht nur unter theoretischen Gesichtspunkten schwierig ist, sondern sich auch in wissenschaftlichen Studien nicht als haltbar erwiesen hat (Comer, 1995), spielt das somatische Syndrom in der aktuellen Klassifikation lediglich eine untergeordnete Rolle und kann ohne Verlust diagnostischer Informationen vom Anwender des Klassifikationssystems vernachlässigt werden.

*Rezidivierende depressive Störungen.* Ist in der Krankengeschichte des Patienten mindestens eine weitere depressive Episode bekannt und lag zwischen der vorhergehenden und der erneuten Erkrankung ein wenigstens zwei Monate andauernder, relativ beschwerdefreier Zeitraum, so wird die Erkrankung als rezidivierende depressive Störung diagnostiziert; dabei kann mit einer Zusatzkennziffer der aktuelle Beschwerdegrad des Betroffenen von gegenwertig remittiert, leicht, mittelgradig oder schwer kodiert werden.

*Bipolare affektive Störung.* Bei Patienten, die im Laufe ihres Lebens neben der gegenwertigen depressiven Episode bereits Zustände manischer, hypomanischer oder gemischter affektiver Episoden erlebt haben, wird die Erkrankung als bipolare affektive Störung klassifiziert. Dabei sind manische oder hypomanische Episoden durch Zustände unverhältnismäßig gehobener oder gereizter Stimmung charakterisiert, während bei gemischten affektiven Episoden manisches und depressives Erleben entweder gleichzeitig vorliegt oder in schnellem Wechsel aufeinander folgt. Auch hier wird durch eine weitere Kennziffer Beschaffenheit und Schweregrad der gegenwertigen Episode näher bezeichnet.

*Anhaltende affektive Störungen.* Für die anhaltenden affektiven Störungen gilt, dass die Symptome über Jahre hinweg andauern, dabei aber nur selten so ausgeprägt sind, dass sie die Kriterien für eine depressive oder manische Episode erfüllen. Die unipolare Form dieser chronischen Verstimmung wird als Dysthymia, die bipolare Form als Zykllothymia bezeichnet. Die genauere Verlaufsform wird durch eine zusätzliche Kennziffer kodiert.

*Anpassungsstörungen.* Depressive Störungsbilder, die nachweislich nach einer besonderen Belastungssituation (in der Regel 1 Monat nach einem belastenden Ereignis) auftreten, und die nicht den Schweregrad einer depressiven Episode erreichen, werden als Anpassungsstörungen im Sinne von depressiven Reaktionen bezeichnet. Diese Störungen sind insofern eine Sonderform, als hier die Diagnose eine direkte Kenntnis der Störungsursache voraussetzt; so geht der Diagnostiker davon aus, dass die Störung ohne das belastende Lebensereignis nicht aufgetreten wäre. Als belastende Lebensereignisse können schwere körperliche Krankheit, eine Störung des sozialen Netzwerkes (bei Trennungserlebnissen oder Trauerfällen) oder eine Veränderung des weiteren Umfeldes (bei Emigration oder Flucht) gelten. Erfüllt die Störung die Kriterien einer depressiven Episode, wird sie jedoch unabhängig von einer möglichen Ursache als solche kodiert.

*Andere affektive Störungen.* In dieser Kategorie werden gemischte affektive Episoden und die rezidivierende kurze depressive Störungen beschrieben, bei der die Phasendauer der Verstimmung nur wenige Tage beträgt, die Symptome aber die Kriterien einer depressiven Episode erfüllen. Treten die depressive Verstimmung sowie die Remission der Symptome im Jahresrhythmus überwiegend innerhalb des gleichen Zeitraums auf, besteht in der aktuellen Klassifikation auch die Möglichkeit, eine sogenannte saisonale affektive Störung zu diagnostizieren. Die Herausgeber weisen allerdings darauf hin, dass eine entsprechende Abgrenzung durch weitere Forschungsergebnisse gesichert werden muss; entsprechend besteht auch die Möglichkeit, ein solches Krankheitsbild unter die Kategorien F30, F31, F32, oder F33 zu fassen (vgl. WHO, 2011, S.209).

*Nicht näher bezeichnete affektive Störung.* Eine „nicht näher bezeichnete“ Kodierung des Krankheitsbildes sollte nur erfolgen, sofern keine Zuordnung zu den anderen Kategorien möglich ist.

#### 2.1.4 Epidemiologie

Das öffentliche Bewusstsein für depressive Störungen und ihre Verbreitung hat in den letzten Jahren stark zugenommen. So widmet sich nicht nur der Boulevard anhand von prominenten Beispielen immer wieder diesem Krankheitsbild, sondern auch Wissenschaft und Politik fokussieren in den letzten Jahren verstärkt die sogenannte Volkskrankheit Depression<sup>1</sup>. Aktuelle methodisch fundierte Studien sprechen dafür, dass die Verbreitung ernsthafter depressiver Erkrankungen in den letzten Jahren innerhalb der Bevölkerung angestiegen ist (Murray & Lopez, 1997). So wird die Gruppe der depressiven Erkrankungen nach Schätzungen der WHO im Jahr 2020 die zweithäufigste Volkskrankheit nach der ischämischen Herzerkrankung stellen. Dabei liegt die Lebenszeitprävalenz für eine klinisch relevante depressive Erkrankung weltweit bereits jetzt bei einem Fünftel der Bevölkerung und unipolare Depressionen sind in allen Altersgruppen die häufigste Ursache für durch Behinderung beeinträchtigte Lebensjahre<sup>2</sup> (Lopez et al., 2006). Dabei ist unklar, ob die höheren Zahlen mit einer verbesserten Diagnostik verbunden sind, oder ob die Zahl der Grunderkrankungen tatsächlich angestiegen ist. So sind zwar aktuelle gesamtgesellschaftliche Entwicklungen wie etwa die Globalisierung mit Risikofaktoren für die Erkrankung assoziiert, es lässt sich jedoch kaum beurteilen, inwiefern die derzeitigen gesellschaftlichen Herausforderungen die Vulnerabilität für depressive Verstimmungen in der Bevölkerung

---

<sup>1</sup> Der vermeintlich verharmlosende Begriff Volkskrankheit bedeutet dabei nicht, dass es sich bei diesen Erkrankungen um ein verbreitetes und daher weniger ernstzunehmendes Krankheitsbild handelt; vielmehr ist die zunehmende Verbreitung der Störung mit erheblichen Folgen für die einzelnen Erkrankten sowie auch mit gesamtgesellschaftlichen Folgen wie etwa einer Schädigung des Arbeitsmarktes oder einer Überlastung des Gesundheitssystems verbunden.

<sup>2</sup> Durch Behinderung beeinträchtigte Lebensjahre, sind Zeiträume, in denen der Betroffene aufgrund seiner Erkrankung nicht in der Lage ist, seine beruflichen, familialen und sozialen Rollen vollständig auszufüllen bzw. in denen die Lebensqualität aufgrund der Erkrankung beeinträchtigt ist. Die Berechnung ist daher relativ komplex (vgl. Lopez et al., 2007)

stärker erhöhen, als es vorherige zeitgeschichtliche Phänomene wie etwa der Zweite Weltkrieg inklusive seiner Folgen getan haben. Unabhängig von den möglichen Ursachen für die größeren Erkrankungszahlen ist jedoch das Bewusstsein für einen erhöhten Handlungsbedarf in Diagnostik, Therapie und Prävention in den letzten Jahren angestiegen.

#### *2.1.4.1 Erkrankungsverläufe und Prävalenzen*

Die Abgrenzung verschiedener depressiver Störungsbilder mithilfe des Klassifikationssystems ICD-10 ist für die klinische Praxis insofern relevant, als die verschiedenen Störungsformen einerseits unterschiedliche Prognosen bezüglich des Erkrankungsverlaufs und –ausgangs zulassen und andererseits in unterschiedlicher Weise auf verschiedene Behandlungsformen ansprechen und somit verschiedene klinische Maßnahmen erfordern. Die Klassifikation eines depressiven Störungsbildes ist dabei in der Praxis häufig nicht einfach, da sich die interindividuellen Krankheitsverläufe auch innerhalb der einzelnen Klassen erheblich unterscheiden können. Dennoch lassen sich Aussagen über typische Erkrankungsmuster treffen, von denen einige innerhalb der Bevölkerung wesentlich häufiger auftreten als andere. In Deutschland wird der Anteil der Bevölkerung, der zu irgendeinem Zeitpunkt seines Lebens an einer depressiven Episode erkrankt, auf 9 % geschätzt, wobei Frauen etwa doppelt so oft betroffen sind wie Männer (Hautzinger & de Jong-Meyer, 2003).

Eine Studie zur 12 Monats-Prävalenz der depressiven Verstimmungen (Major Depression, Dysthymia und bipolare affektive Störung nach DSM) in insgesamt 14 Nationen in Europa, den USA, Asien und Afrika berichtet über eine große Bandbreite an Erkrankungshäufigkeiten (siehe Tabelle 3).

*Tabelle 3: 12 Monatsprävalenz depressiver Verstimmungen im weltweiten Vergleich*

Nation	Prozentsatz der Stichprobe mit depressiver Verstimmung innerhalb der letzten 12 Monate
<b>Amerika</b>	
Kolumbien	6.8
Mexiko	4.8
Vereinigte Staaten	9.6
<b>Europa</b>	
Belgien	6.2
Frankreich	8.5
Deutschland	3.6
Italien	3.8
Niederlande	6.9
Spanien	4.9
Ukraine	9.1
<b>Mittlerer Osten und Afrika</b>	
Libanon	6.6
Nigeria	0.8
<b>Asien</b>	
Japan	3.1
Volksrepublik China	
Peking	2.5
Shanghai	1.7

Anmerkung. Tabelle des WHO World Mental Health Survey Consortium (2004), zitiert nach Hammen & Watkins (2008).

*Depressive Episode /Majore depressive Störung.* Die depressive Episode oder Majore Depression (nach DSM IV) ist die häufigste Verlaufsform der depressiven Störung. Die Lebenszeitprävalenz wird nach aktuellen epidemiologischen Studien in den USA auf 16.2 % (Kessler et al., 2003) und in Europa auf 12,8% (ESEMED/MHEDEA 2000 Investigators, 2004) geschätzt.

Die Symptome dauern bei dieser Erkrankungsform durchschnittlich drei bis sechs Monate an. Zwischen 50% und 85% der Patienten erleben dabei mehr als nur eine depressive Phase, wobei die Anzahl der einzelnen Phasen im individuellen Krankheitsverlauf stark variiert (Hautzinger & de Jong-Meyer, 2003).

*Bipolare affektive Störung.* Von den bipolaren affektiven Störungen ist ein weitaus geringerer Anteil der Bevölkerung betroffen; die Lebenszeitprävalenz wird für Deutschland auf 0.2 % geschätzt (Hautzinger & de Jong-Meyer, 2003). Die Erkrankungsform tritt bei Männern und Frauen etwa gleich häufig auf (Davidson, Neale & Hautzinger, 2007). Bei bipolaren

Verläufen ist die Phasendauer durchschnittlich kürzer als bei rezidivierenden depressiven Störungen, die Phasenanzahl dagegen etwa doppelt so hoch (Hautzinger & de Jong-Meyer, 2003). Die obigen Prävalenzangaben der bipolaren Störungen stellen allerdings dann eine Unterschätzung dar, wenn man mildere Verlaufsformen dieser Erkrankung berücksichtigt, wie es in der ICD-10 zwar prinzipiell möglich ist, aber anders als im DSM-IV bisher nicht explizit berücksichtigt wird. So wird im Klassifikations-System DSM-IV zwischen der Bipolar-I- und der Bipolar-II-Störung unterschieden. Die Bipolar-I-Störung ist durch ausgeprägt manische sowie depressive Phasen gekennzeichnet und bei Männern und Frauen etwa gleich häufig. Die Bipolar-II-Störung bezeichnet dagegen einen etwas weniger schwerwiegenden Verlauf mit hypomanischen und depressiven Episoden. Dabei sind für die depressiven Episoden eine Gewichtszunahme und ein vermehrtes Schlafbedürfnis bei den Betroffenen charakteristisch. Für diese Form der Erkrankung werden höhere Prävalenzschätzungen von bis zu 10 % berichtet. Sie ist vor allem bei jüngeren Frauen stark verbreitet (Möller, Laux & Deiser, 2005). Heitz, Stewart & Dammann (2010) weisen in einem Literaturüberblick auf den Zusammenhang zwischen Bipolar-II-Störung und atypischer Depression hin. So ist die depressive Symptomatik bei bipolarem Verlauf häufig i.S. der oben genannten Symptome atypisch, das Syndrom der atypischen Depression hat jedoch nicht notwendiger Weise einen bipolaren Verlauf. Dabei schätzen die Autoren den Anteil depressiver Patienten, die ein atypisches Symptommuster aufweisen, auf bis zu 47%.

*Dysthymia.* In Europa wurde für die Dysthymie ein Morbiditätsrisiko von 4,1 % (ESEMeD/MHEDEA 2000 Investigators, 2004) ermittelt, wobei auch hier das Risiko für Frauen wie bei den anderen unipolaren depressiven Störungen etwa doppelt so hoch liegt. Dysthymien haben per definitionem einen chronifizierten Verlauf, bei dem die Symptome über Jahre hinweg andauern (Hautzinger & de Jong-Meyer, 2003).

#### *2.1.4.2 Prognosen und Behandelbarkeit*

Bei einem relativ hohen Anteil depressiver Patienten bessert sich die Symptomatik nach einer gewissen Zeit unabhängig davon, ob sie behandelt werden oder nicht. So tritt bei etwa der Hälfte bis zwei Drittel der Patienten, die an einer depressiven Episode erkranken, nach vier bis sechs Monaten eine so starke Besserung ein, dass sie wieder arbeitsfähig sind und zu ihrem alten Selbst zurückfinden. Diese beschwerdefreien Zeiten, die auch als Remissionszeiten bezeichnet werden, sind von unterschiedlich langer Dauer. Die Rückfallwahrscheinlichkeit ist innerhalb des ersten Jahres nach Abklingen der Episode am höchsten; sie wird für diesen Zeitraum auf 40 % geschätzt. Bei 42% der unipolaren und bei 30% der bipolaren Patienten wurden jedoch beschwerdefreie Zeiträume von über fünf Jahren festgestellt (Hautzinger & de Jong-Meyer, 2003; Davidson, Neale & Hautzinger, 2007). Die Forschungsliteratur berichtet allerdings auch eine beträchtliche Anzahl ungünstiger Krankheitsverläufe. Neben den chronifizierten Verläufen der Dysthymen Störungen besteht bei 20% bis 30% der depressiven Episoden die Symptomatik länger als ein Jahr (Hautzinger & de Jong-Meyer, 2003), bei 12% sogar bis zu zwei Jahren (Davidson, Neale & Hautzinger, 2007). Insbesondere bei Rückfällen besteht eine Chronifizierungsgefahr. Darüber hinaus begehen 10% bis 15% depressiver Patienten Suizid, ein Anteil, der deutlich über dem geschätzten Risiko der Normalbevölkerung liegt (Hautzinger & de Jong-Meyer, 2003).

*Komorbidität.* Man schätzt, dass über 70% aller depressiven Erkrankungen komorbid mit weiteren psychischen Störungen einhergehen (Kessler et al., 2003). Insofern ist die Heterogenität der Krankheitsbilder bei den Betroffenen nicht nur auf die Varianz der Krankheitsbilder affektiver Störungen selbst sondern auch auf die begleitenden psychischen Auffälligkeiten der Betroffenen zurückzuführen. Das bedeutet für die Praxis, dass eine Prognose des Krankheitsverlaufes unbedingt das gesamte Bild psychischer Vulnerabilität bzw. psychischer Ressourcen des Betroffenen einschließen muss. So erschwert beispielsweise eine komorbide Persönlichkeitsstörung wie etwa die Borderline-Störung dem Betroffenen

ganz erheblich den Umgang mit der depressiven Verstimmung während ein ansonsten psychisch wenig beeinträchtigter Patient, der aufgrund vorübergehender psychischer Belastung eine Anpassungsstörung erleidet, eine wesentlich bessere Verlaufsprognose aufweist.

*Behandlungsindikation.* Aufgrund des hohen Leidensdrucks sowie des erhöhten Mortalitätsrisikos depressiver Patienten ist eine Behandlung dieser Erkrankung in jedem Fall indiziert. Depressionen gelten dabei im Allgemeinen als gut therapierbar. Es wird angenommen, dass mehr als 60% der Betroffenen auf eine der gängigen Therapiemethoden ansprechen (Comer, 1995, 2008). Fragen, inwiefern der Krankheitsverlauf durch Therapien grundsätzlich beeinflusst werden kann und welche Kombination von Methoden bei welcher Symptomatik am besten geeignet ist, sind dabei Gegenstand zahlreicher aktueller Forschungsbemühungen (Davidson & Neale, 2002; Davidson, Neale & Hautzinger, 2007; Hautzinger & de Jong-Meyer, 2003).

### 2.1.5 Ätiologie

Depressive Störungen sind weit verbreitete psychische Erkrankungen, deren Symptome sich auf nahezu alle Bereiche menschlichen Erlebens und Verhaltens auswirken. Dabei konnte bisher keine eindeutige kausale Bedingung für die Entstehung dieser Störung identifiziert werden. Vielmehr sprechen empirische Studien für eine Reihe sogenannter Risikofaktoren, die aus der Beobachtung abgeleitet werden, dass Anteile der Bevölkerung, die bestimmte Merkmale aufweisen, häufiger an der Störung erkranken als andere (Hautzinger & de Jong-Meyer, 2003; Davidson, Neale & Hautzinger, 2007).

#### 2.1.5.1 Risikofaktoren

*Geschlecht.* Von den unipolaren depressiven Erkrankungen sind Frauen in allen Kulturen häufiger betroffen als Männer. Dabei berichten einige Studien von etwa doppelt so hohen Prävalenzraten (ESEMED/MHEDEA 2000 Investigators, 2004; Kessler, 2003; Kessler et al.,

2003; Paykel et al., 2005), während eine aktuelle europäische Studie, die sich durch eine sorgfältige Kontrolle verschiedener Messfehler auszeichnet (Vand de Velde, Bracke, Levecque & Meuleman, 2010), wesentlich kleinere und regional verschiedene Effekte berichtet. Dennoch stellt das weibliche Geschlecht nach wie vor einen der wichtigsten Risikofaktoren für die Ausbildung einer depressiven Erkrankung dar. Eine einfache Erklärung lässt sich hierfür anhand des komplexen Zusammenspiels hormoneller, neuronaler, biografischer sowie sozialer Faktoren, die das Geschlecht ausmachen, nicht finden. Dabei bedeutet weibliches Geschlecht mit dem Eintreten in die Adoleszenz in vielerlei Hinsicht eine erhöhte Stressexposition im Hinblick auf soziale Benachteiligung, dem Eintreten kritischer Lebensereignisse wie etwa sexueller Missbrauch sowie eine höhere Vulnerabilität im Hinblick auf negative Körperwahrnehmung. Zudem ist der Verlauf vieler weiblicher Biografien durch zahlreiche weitere langanhaltende Stressoren wie etwa alleinerziehende Elternschaft sowie die Übernahme der Pflege eines nahen Angehörigen gekennzeichnet. Darüber hinaus tendieren Frauen eher zu ungünstigen Coping-Strategien wie etwa einer Überpsychologisierung ihrer Probleme, während Männer durch mehr externale Coping-Stile weniger gefährdet für negative Gedanken sind (vgl. Hammen & Watkins, 2008).

*Affektive Erkrankungen in der Familienanamnese.* Familienangehörige von depressiven Patienten weisen abhängig vom Verwandtschaftsgrad und der Subgruppe der jeweiligen Störung ein höheres Morbiditätsrisiko auf (die Morbiditätsrisiken liegen zwischen 8% und 60 %, Hautzinger & de Jong-Meyer, 2003). Dabei liegen die Konkordanzraten für alle affektiven Störungen bei monozygotischen Zwillingen am höchsten: hier werden Erkrankungsrisiken von bis zu 80 % berichtet, während das Erkrankungsrisiko bereits bei dizygotischen Zwillingen wesentlich niedriger ist (0% - 30%). Die Beobachtung, dass Adoptivgeschwister nur geringe Konkordanzraten aufweisen (8%), unterstützt dabei die Annahme, dass hier eher genetische als psychosoziale Faktoren eine Rolle spielen.

Zudem ist das Erkrankungsrisiko der Angehörigen erheblich von der Störungsform des Betroffenen abhängig. So gibt es für die unipolaren depressiven Störungen kaum gesicherte Angaben darüber, ob sie bei nahem Verwandtschaftsgrad überhaupt vermehrt auftreten und in Studien, in denen sich entsprechende Effekte zeigten, wurden nur erhöhte Risiken für unipolare Depressionen berichtet. Für Verwandte ersten Grades von bipolaren Patienten steigt das Risiko für alle affektiven Störungen dagegen deutlich. Es werden Erkrankungsrisiken von 34% für die Eltern, 35-39% für Geschwister und 60% für Kinder der Betroffenen berichtet. Aufgrund dieser komplexen empirischen Daten werden heutzutage von den meisten Autoren polygenetische Übertragungsmodelle bzw. genetisch-faktorielle Schwellenmodelle zur Erklärung verwandtschaftlicher Konkordanzraten favorisiert (Hautzinger & de Jong-Meyer, 2003).

*Soziale Beziehungen.* Das Fehlen einer nahen Bezugsperson stellt ebenfalls einen relativ gesicherten Risikofaktor dar. Geschiedene oder getrennt lebende Personen erkranken ebenso wie Personen ohne vertraute Beziehungen häufiger an Depressionen (Comer, 2008). Eine aktuelle Studie zum Kommunikationsverhalten und nahen Bezugspersonen weist jedoch darauf hin, dass nur solche sozialen Beziehungen präventiv wirken, in denen ein liebevoller Umgangston und gegenseitige emotionale Unterstützung gepflegt werden, während nahe Beziehungen mit verstärkter negativer Emotionalität als unabhängiger Risikofaktor für die Entwicklung einer depressiven Verstimmung gelten (Stafford, McMunn, Zaninotto & Nazroo, 2011).

*Kulturelle Bedingungen.* Depressionen sind in den westlich geprägten Gesellschaften stärker verbreitet als in anderen Kulturen. Die Tatsache, dass jedoch auch die meisten Untersuchungen und Messungen westlich geprägt sind und dass darüber hinaus kulturelle Verschiedenheiten in der Wahrnehmung der Symptome berichtet werden, macht eine Einschätzung der Bedeutung des Risikofaktors Kultur schwierig. Dabei ist davon auszugehen, dass in Kulturräumen, deren Bewohner mit Risikofaktoren wie Armut, militärischen

Auseinandersetzungen etc. verbunden sind, ein erhöhtes Risiko für Depressivität gegeben ist. In der Forschungspraxis besteht allerdings die Schwierigkeit, dass sich die rein kulturellen Bedingungen jeweils kaum von den aktuellen politisch-gesellschaftlichen und makroökonomischen Gegebenheiten trennen lassen (vgl. Hammens & Watkins, 2008).

*Sozioökonomischer Status.* Die Mehrzahl der Forschungsbeiträge spricht für einen Zusammenhang zwischen niedrigem Einkommen, geringem sozialen Status und depressiven Erkrankungen. Dabei ließe sich ein kausaler Zusammenhang zwischen Armut und Depression in zweierlei Richtung konstruieren: Einerseits mindern depressive Verstimmungen die Erwerbsfähigkeit und führen somit langfristig zu einem niedrigeren Einkommen und andererseits ist Armut ein Lebensumstand, der mit zahlreichen Belastungen und Beeinträchtigungen verbunden ist, die das Entstehen einer depressiven Erkrankung stark begünstigen können. Eine Längsschnittuntersuchung mit finanziell benachteiligten aber psychisch gesunden Probanden spricht mit erhöhten Neuerkrankungsraten in dieser Bevölkerungsgruppe allerdings eher für die zweite These (vgl. Hammen & Watkins, 2008; Hautzinger & de Jong-Meyer, 2003).

*Urbanes Wohnumfeld.* Die lange Zeit vertretene These, dass depressive Verstimmungen in urbanen Zentren verbreiteter sind als in ländlichen Gebieten, wird durch neuere Studien immer wieder infrage gestellt. Insofern kann das Leben in einem städtischen Raum nicht mehr als gesicherter Risikofaktor gelten (vgl. Hammens & Watkins, 2008).

*Belastende Lebensereignisse.* Vor dem Ausbrechen depressiver Erkrankungen treten in vielen Fällen belastende Lebensereignisse auf. Dabei handelt es sich häufig um Verlusterlebnisse wie etwa Scheidung oder den Austritt aus dem Berufsleben (Davison & Neale, 2002).

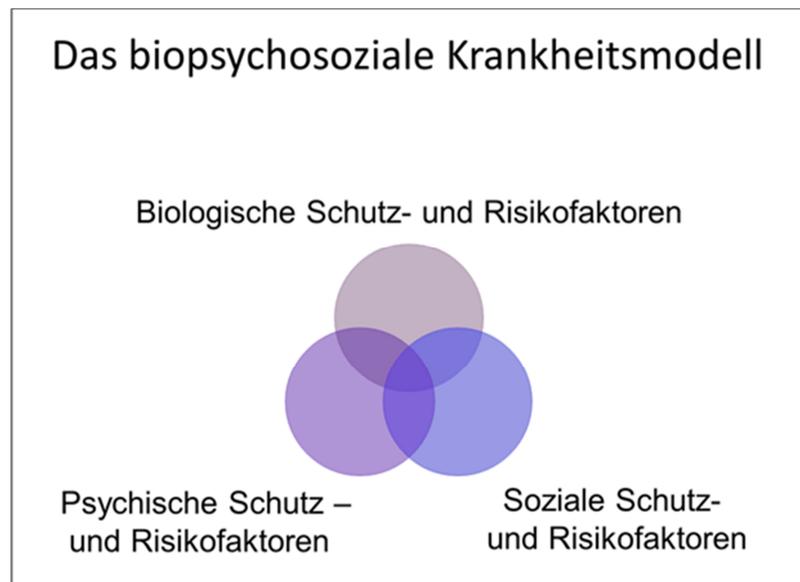
#### 2.1.5.2 Ätiologische Risikofaktorenmodelle

Risikofaktoren können für sich gesehen nicht als Erklärungsansatz für die Entstehung einer depressiven Störung gelten, da sie lediglich empirische Zusammenhänge aufzeigen. Die Kausalität der entsprechenden Effekte ist dabei häufig unklar. So kann durch empirische

Befunde über den Risikofaktor soziale Isolation beispielsweise kaum etwas darüber ausgesagt werden, ob Menschen depressiv werden, weil sie wenige soziale Kontakte haben, oder ob sie wenige soziale Kontakte haben, weil sie depressiv sind. Ein weiteres Problem von Risikofaktoren ist, dass sie häufig in Form von Aufzählungen präsentiert werden, ohne dass dabei eine logische Ordnung vorliegt. Weibliches Geschlecht etwa kann zwar als ein einziger hochwirksamer Risikofaktor betrachtet werden; weibliches Geschlecht ist aber gleichzeitig auch eine Ansammlung verschiedener Faktoren, die darunter liegen: eine entsprechende biologisch-genetische Veranlagung, soziale Phänomene wie z.B. die Geschlechtsrollen, lebensgeschichtlich wirksame Faktoren wie hormonelle Umstellungen oder geschlechtsspezifische biografische Vulnerabilitäten wie sexueller Missbrauch.

Um diesem Phänomen gerecht zu werden, gibt es Modelle, die dabei helfen, eine logisch-hierarchische Ordnung verschiedener Risiko- und Schutzfaktoren herzustellen. Am gebräuchlichsten sind das biopsychosoziale Krankheitsmodell sowie das Diathese-Stress-Modell. Dabei werden auch die verschiedenen möglichen Überlappungen der Bereiche, wie z.B. biosozial, psychosozial oder psychobiologisch mitberücksichtigt.

*Das biopsychosoziale Krankheitsmodell.* Das biopsychosoziale Krankheitsmodell ordnet Risikofaktoren thematisch nach biologischen, psychologischen und sozialen Aspekten.



*Abbildung 1: Das biopsychosoziale Krankheitsmodell*

Ein zu depressiven Reaktionen neigender Patient kann beispielsweise eine besondere, genetisch bedingte Vulnerabilität aufweisen (biologischer Risikofaktor), während ein gutes soziales Netzwerk (sozialer Schutzfaktor) sowie eine wenig neurotische Persönlichkeit (psychologischer Schutzfaktor) ihn gleichzeitig vor dem Ausbruch der Erkrankung bewahren (Hörning, 2004; Waller, 2002).

*Das Diathese-Stress-Modell.* Das Diathese-Stress-Modell beschreibt, wie Risiko- und Schutzfaktoren über die Zeit hinweg wirksam werden. So kann eine biologische Veranlagung seit der Geburt bestehen. Diese Veranlagung oder Diathese stellt jedoch nur dann ein Problem dar, wenn im weiteren Lebenslauf des Betroffenen zusätzlich Stressfaktoren hinzukommen, wie etwa der Verlust eines nahen Angehörigen oder eine belastende Arbeitssituation. Eine Depression entsteht dann, wenn das Zusammenspiel von Risiko- und Schutzfaktoren derartig mit der Veranlagung eines Patienten interagieren, dass ein psychisches Gleichgewicht vom Organismus nicht mehr hergestellt werden kann. Ein hilfreiches Bild zum Verständnis des Diathese-Stress-Modells ist die Vorstellung des menschlichen Organismus in Form eines

Bootes (vgl. Abbildung 2). Dieses Boot hat bestimmte physikalische Eigenschaften, wie z.B. einen gewissen Schwerpunkt, mit dem es im Wasser liegt. Diese physikalischen Eigenschaften entsprechen der Diathese des Organismus. Die Umgebung des Bootes i.S. der Windstärke und des Wellenganges entsprechen dem Stress, dem der Organismus ausgesetzt ist. Übersteigt der Wellengang ein Ausmaß, dem das Boot nach seinen physikalischen Gesetzmäßigkeiten standhalten kann, so wird ein kritischer Wasserpegel überschritten und das Boot sinkt. Entsprechend „versinkt“ der Organismus in einer depressiven Verstimmung, wenn er über längere Zeit hinweg Stressoren ausgesetzt ist, die das Ausmaß seiner Belastbarkeit überschreiten.

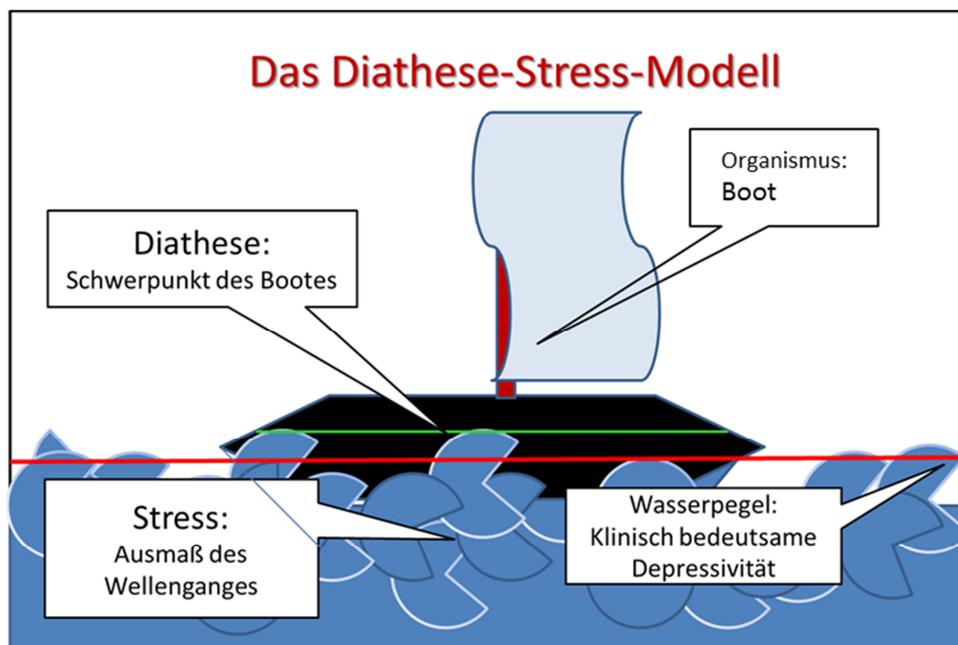


Abbildung 2: Das Diathese-Stress-Modell

### 2.1.5.3 Psychologische und Psychiatrische Erklärungsansätze

Erklärungsansätze depressiver Störungen versuchen das empirische Wissen über diese Erkrankungen theoretisch zu untermauern. Da dieses Wissen, wie aus den Risikofaktoren ersichtlich wird, in verschiedenen Bereichen angesiedelt ist, sind Modelle zur Erklärung dieser Erkrankungen ebenso zahlreich wie unterschiedlich. Medizinische Forschungsansätze weisen genetischen, physiologischen oder biochemischen Faktoren die entscheidende

Bedeutung für die Entwicklung der Störung zu. Psychologische Theorien beschäftigen sich mit psychodynamischen, kognitiven oder behavioralen Aspekten der Krankheitsgenese während soziologische Ansätze sich mit interpersonalen und gesellschaftlichen Bedingungen auseinandersetzen, die im Zusammenhang mit depressiven Erkrankungen relevant sind (vgl. Comer, 2008; Davison, Neale & Hautzinger 2007; Hautzinger & de Jong-Meyer, 2003; Pinel, 1997; Birbaumer & Schmidt, 2010). Da die Darstellung aller relevanten Theorien zu depressiven Erkrankungen an dieser Stelle nicht möglich ist, werden im Folgenden nur einige Ansätze vorgestellt, die zur Entwicklung wirksamer Therapiemethoden beigetragen haben und die daher für die klinische Praxis besonders bedeutsam sind.

*Der psychodynamische Ansatz.* Das psychoanalytische Krankheitsverständnis (Mertens, 2000) sieht die Ursachen psychischer Leiden in unbewussten inneren Konflikten, die nicht durch gegenwärtige belastende Umstände entstehen, sondern die auf frühe Erlebnisse in der Lebensgeschichte des Patienten zurückgehen. Im Fall der depressiven Störungen wird vermutet, dass diese inneren Konflikte mit real erfahrenen oder imaginären Verlusterlebnissen in der Entwicklung des Erkrankten zusammenhängen. Als Ursprung der psychodynamischen Theorien zur Entstehung der Depression gilt der Ansatz Freuds (1917) (Comer, 2008; Davison, Neale & Hautzinger 2007), der sich seitdem in zahlreichen Schulrichtungen verändert und weiterentwickelt hat, von denen einige zu nachweislich wirksamen Behandlungsmethoden geführt haben (Mertens 2000; Peters, 2008; Zank, Wilz & Peters, 2010). (vgl. hierzu auch die Abschnitte 2.2.6.1 und 2.2.8.1 zur psychodynamischen Psychotherapie mit älteren Patienten).

*Der behaviorale Ansatz.* Das behaviorale Konzept der Depression steht in der Tradition der operanten Lerntheorie. Es beschäftigt sich nicht mit innerpsychischen Vorgängen sondern fokussiert auf die nach außen hin wahrnehmbaren Umstände und Verhaltensmuster, die im Zusammenhang mit depressiven Erkrankungen relevant sind. Die einflussreichste Theorie, die auf dieser Grundlage entstanden ist, wurde von Lewinsohn (1975, 1981) entwickelt

(Blöschl, 1998; Hautzinger & de Jong-Meyer, 2003; Zank, Wilz & Peters, 2010). Lewinsohn geht davon aus, dass bei der Entwicklung der depressiven Störung eine massive Reduktion positiver Erlebnisse im Leben des Erkrankten die entscheidende Rolle spielt. Dabei hat das Ausbleiben von angenehmen zwischenmenschlichen Kontakten, die Lewinsohn als soziale Verstärker bezeichnet, einen besonderen Stellenwert. Die Ursachen dafür können sowohl widrige Umweltbedingungen als auch ein Fehlverhalten des Erkrankten aufgrund mangelhafter sozialer Kompetenzen sein. Im Verlauf der depressiven Erkrankung führt der Verstärkerverlust zur Ausprägung depressiver Verhaltensweisen wie Passivität, Antriebsmangel und Interessenverlust, die ihrerseits eine weitere Reduktion von verstärkenden Erlebnissen nach sich ziehen. Der Ansatz Lewinsohns hat zu Therapiemethoden geführt, die im Rahmen der kognitiven Verhaltenstherapie immer noch erfolgreich angewendet werden.

*Der kognitive Ansatz.* Kognitive Modelle der Depression sehen die Ursache der depressiven Erkrankung in einer Störung des Denkens, die sich der Patient aufgrund von belastenden Erfahrungen in seiner Entwicklung erworben hat. Das fehlangepasste, negative Denken des Depressiven begünstigt im Laufe der Erkrankung schließlich emotionale und motivationale Depressionssymptome (z.B. Niedergeschlagenheit, sozialer Rückzug etc.), die ihrerseits die depressive Weltsicht des Erkrankten bestätigen, und so zu einer Verfestigung des Krankheitsbildes führen. Die zwei allgemein anerkannten Depressionsmodelle mit kognitivem Schwerpunkt sind die kognitive Depressionstheorie von Beck und die Theorie der erlernten Hilflosigkeit von Seligmann (vergl. Blöschl, 1998, Davison & Neale, 2002, Hautzinger & de Jong-Mayer, 2003).

*Die kognitive oder Schema-Theorie nach Beck.*

Aaron Beck (1967, 1987) war der Auffassung, dass depressive Menschen ein negatives Bild über sich selbst, über ihre Umwelt und über ihre Zukunft haben. Dieses Bild, auch kognitive Triade der Depression genannt wird durch besonders belastende Lebensereignisse im Verlauf

der Entwicklung erworben und führt im Erwachsenenalter zu einer Reihe von Denkfehlern wie willkürlichen Schlussfolgerungen (Ich bin nichts wert, weil das Wetter an meinem Geburtstag schlecht ist.), Übergeneralisierung (Ich versage immer in Prüfungen.) sowie Über- oder Untertreibung (Ich bin dick und vollkommen hässlich.), die negative Stimmungen begünstigen. Entsprechend war Beck der Auffassung, dass es möglich ist, über eine Veränderung des Denkens eine Verbesserung der Stimmung zu erzielen (vgl. hierzu auch die Abschnitte 2.2.6.2 bzw. 2.2.8.2 zur kognitiven Verhaltenstherapie mit älteren Patienten). Die Theorie von Beck wird durch umfangreiches empirisches Datenmaterial unterstützt. So konnten Disner, Beevers, Haigh und Beck (2011) in einem aktuellen Review zahlreiche Zusammenhänge zwischen neuronalen Mechanismen der Depression und der kognitiven Triade herstellen.

*Die Theorie der erlernten Hilflosigkeit nach Seligman.* Der Theorie der erlernten Hilflosigkeit nach Seligman ist eine Beobachtung des Forschers in Lernexperimenten mit Hunden vorrausgegangen. So machte Seligman die Erfahrung (1974), dass Versuchstiere, die in Lernexperimenten über einen längeren Zeitraum hinweg aversiven Reizen ausgesetzt werden, die von ihnen als unkontrollierbar erlebt werden, weil sie kein vorhersagbares Muster aufweisen (z.B. ein wiederholtes Bestrafen durch Elektroschocks unabhängig vom Verhalten des Tieres), ähnliche Verhaltensweisen zeigen (Apathie, Rückzug) wie Menschen, die unter einer depressiven Verstimmung leiden. Diese Versuchstiere zeigten auch dann weiterhin apathisches resignatives Verhalten, wenn die Möglichkeit bestand, der Situation aktiv zu entfliehen, wenn also beispielsweise die Zwingertür vom Versuchsleiter offen gelassen wurde. Seligman und seine Nachfolger entwickelten daraufhin die These, dass der Entwicklung einer depressiven Verstimmung beim Menschen möglicherweise ebenfalls negative Ereignisse vorausgehen (wie z.B. der Verlust eines nahen Angehörigen, der Verlust der Arbeitsstelle oder das Eintreten einer schweren Erkrankung), die als unabhängig vom eigenen Verhalten und somit als unkontrollierbar erlebt werden. Habituiert sich eine

entsprechende Grundüberzeugung der eigenen Hilflosigkeit, so versucht der an Depressionen erkrankte Mensch auch nicht mehr, aversive Zustände zu beeinflussen (etwa Konflikte in Beziehungen oder eine unbefriedigende Arbeitssituation), die er eigentlich mitgestalten könnte. Die passive Haltung verstärkt in der Folge die negativen Auswirkungen der Situation, was zu weiterem Hilflosigkeitserleben führt.

Eine Erweiterung der Theorie um verschiedene Attributionsstile wird der Tatsache gerecht, dass aversive Lebensereignisse nur bei einigen Menschen zu depressiven Verstimmungen führen. Dabei ist die subjektive Theorie des einzelnen Menschen darüber, wie und warum bestimmte Ereignisse eintreten, ein entscheidender Moderator im Umgang mit schwierigen Lebenssituationen. Grundsätzlich werden drei Dimensionen der möglichen Attribuierung (Ursachenzuschreibung) von Ereignissen unterschieden, die auf unterschiedliche Weise kombiniert werden können (vgl. Abbildung 3).

	Internal (persönlich)		External (umweltbedingt)	
	Stabil	Variabel	Stabil	Variabel
Global	Ich bin dumm.	Ich bin erschöpft.	Die Tests sind alle unfair.	Heute ist Freitag der 13.
Spezifisch	Ich bin mathematisch unbegabt.	Ich habe Mathematik im Moment satt.	Die Mathe-Tests sind unfair.	Mein Aufgabenheft hatte die Nr. 13.

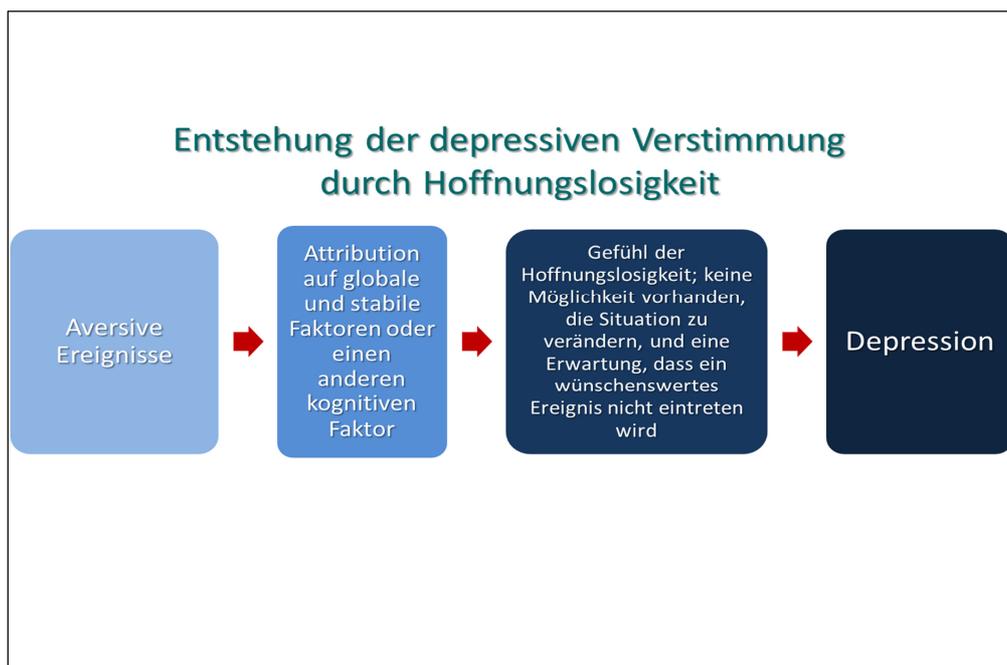
Anmerkungen. Attributionsschemata zur Beantwortung der Frage „Warum habe ich bei der Mathematik-Prüfung versagt?“, (vergl. Davidson, Neale & Hautziger, 2007)

*Abbildung 3: Depressive Attributionsschemata*

Auf kritische Lebensereignisse, wie z.B. das Versagen bei einer Prüfung wird nach der Attributionstheorie eher mit depressiver Verstimmung reagiert, je stärker diese auf globale (in allen Situationen relevante), internale (in der eigenen Person liegende) und zeitlich stabile Ursachen zurückgeführt werden. Entsprechend würde nach dem in der Abbildung gegebenen

Beispiel die Attribution „Ich bin dumm.“ zur Erklärung beim Scheitern im Mathematiktest am ehesten die Entstehung einer Depression begünstigen, während die Erklärung „Mein Aufgabenheft hatte die Nummer 13“ die Stimmung wesentlich weniger beeinträchtigen würde.

In der aktuellsten Version der Theorie der erlernten Hilflosigkeit (Abramson et al., 1989) wurde der Begriff der Hilflosigkeit durch den etwas differenzierteren Term „Hoffnungslosigkeit“ ersetzt. Hoffnungslosigkeit bezieht sich dabei darauf, dass das Individuum erwartet, dass unerwünschte Ereignisse eintreffen und erwünschte Ereignisse ausbleiben werden, wobei das Individuum jeweils keine Möglichkeit sieht, die Situation zu verändern (Hilflosigkeit).



Anmerkungen. Die Abbildung nach Abramson et al. (1989) wurde zitiert nach Davidson, Neale & Hautzinger (2007)

Abbildung 4: Die Theorie der erlernten Hilflosigkeit/Hoffnungslosigkeit

Abbildung 4 Beschreibt schematisch den durch die Theorie der Hoffnungslosigkeit beschriebenen Zusammenhang zwischen dem Eintreten aversiver Ereignisse und dem Eintreten einer depressiven Erkrankung.

*Der neuro-chemische Ansatz.* Neuro-chemische Theorien zur Depression nehmen an, dass bei dieser Erkrankung die Kommunikation zwischen den Nervenzellen im Gehirn durch eine veränderte Konzentration von chemischen Botenstoffen, den sogenannten Neurotransmittern, gestört ist. Sie sind im Zusammenhang mit Medikamentenstudien entstanden, durch die gezeigt werden konnte, dass mit Substanzen, die Einfluss auf die Konzentration einiger im Gehirn wirksamer Neurotransmitter haben, depressive Symptome erfolgreich behandelt werden können (Birbaumer & Schmidt, 2010; Möller, 1997; Pinel, 1997). In diesem Zusammenhang ist vor allem die Monoamintheorie der Depression relevant. Sie beruht auf der Beobachtung, dass die meisten antidepressiv wirkenden Medikamente die Konzentration von Noradrenalin und Serotonin, zweier Neurotransmitter, die zur Gruppe der Monoamide gezählt werden, im Gehirn erhöhen. Die Kernaussage der Theorie ist, dass depressive Verstimmungen durch einen Mangel an Aktivität der Synapsen entstehen, deren Signalübertragung über diese Monoamine wirkt. Dabei wurde die Wirksamkeit von Antidepressiva vor allem bei schwereren akuten Depressionen durch zahlreiche Studien nachgewiesen (Möller, Laux & Kampfhammer, 2008).

*Integrative Ansätze.* In aktuellen theoretischen Arbeiten (Pössel & Knopf, 2010) gibt es Bestrebungen, das Depressionsmodell nach Beck, die Hoffnungslosigkeitstheorie und die Response-Style-Theorie in ein einheitliches kognitives Modell der Depression zu integrieren. Die Response-Style-Theorie (Nolen-Hoeksema & Morrow, 1991) beschäftigt sich in erster Linie mit der Reaktion des Betroffenen auf seine depressive Verstimmung und besagt, dass depressive Patienten, die sich vermehrt mit ihren Symptomen beschäftigen und viel grübeln, eine schlechtere Prognose aufweisen als Patienten, denen es gelingt, ihre Depressivität als eine vorübergehende Verstimmung zu akzeptieren. Dieser Ansatz ist theoretisch umstritten, da er vermeintliche Ursachen der depressiven Verstimmung nicht einwandfrei von depressiven Symptomen differenzieren kann; so ist beispielsweise Grübeln ein Symptom der Depression und kann daher nicht gleichzeitig als Ursache der Verstimmung gelten (Treyner,

2003). Der Versuch, die verschiedenen Modelle zu integrieren führte zu komplexeren Beobachtungen, als die Autoren zunächst vermutet hatten. Dabei ließ sich zunächst die These stützen, dass sich die verschiedenen Ansätze tatsächlich empirisch voneinander unterscheiden. Eine Überführung in ein einheitliches Modell konnte dabei aber nur teilweise gelingen. So konnte z.B. nicht bestätigt werden, dass eine negative Sichtweise der Umwelt mit einem bestimmten Attributionsstil korrespondiert. Entsprechend bedarf es weiterer Forschung zur Erstellung eines einheitlichen kognitiven Depressionsmodells.

## *2.1.6 Versorgung*

### *2.1.6.1 Behandlungsempfehlungen*

Die empfohlene Behandlung der Verschiedenen Erkrankungsformen der depressiven Störung ist in Deutschland durch den Gesetzgeber und Empfehlungen des Gesundheitsministeriums geregelt:

Als psychotherapeutische Verfahren stehen in Deutschland für die ambulante Behandlung von Patientinnen und Patienten mit depressiven Erkrankungen kognitive Verhaltenstherapie und psychodynamische Psychotherapieverfahren zur Verfügung, im stationären Bereich darüber hinaus unter anderem die Interpersonelle Psychotherapie und die Gesprächspsychotherapie... Die Wirksamkeit von Antidepressiva in der Akuttherapie von Depressionen Erwachsener ist ebenfalls durch zahlreiche Meta-Analysen belegt (z.B. Strossem et al. 2001; Anderson 2000; Smith et al. 2002)....

Bei länger andauernder leichter und bei mittelschwerer Depression liefern die vorhandenen Studien keine Kriterien für eine Präferenz für eine Psycho- oder Pharmakotherapie in der Akutphase (De Rubeis et al. 1999; Hollon et al. 2002). Deshalb entscheidet hier primär der Wunsch des oder der Betroffenen nach entsprechender Aufklärung über die Besonderheiten jedes Verfahrens. Bei schwerer Depression empfehlen derzeit die nationalen und internationalen Leitlinien die Präferenz der Pharmakotherapie oder eine Kombination von Psycho- und Pharmakotherapie (NICE 2004; de Jong-Meyer et al. (im Druck).

(Bundesgesundheitsministerium, 2006, S.11-12).

### *2.1.6.2 Aktuelle Entwicklung*

Die starke Verbreitung der Volkskrankheit Depression mit ihren entsprechenden sozialen und auch ökonomischen Folgen hat in letzter Zeit zu einer stärkeren gesellschaftlichen Sensibilisierung für dieses Thema geführt. So wurde die Prävention, Früherkennung und effektive Behandlung depressiver Störungen vom Bundesministerium für Gesundheit und der Gesellschaft für Versicherungswissenschaft und -gestaltung 2006 zum 6. Nationalen Gesundheitsziel erklärt (vgl. Heidenblut & Zank, 2009). Gesundheitsziele sind ein verhältnismäßig neues Instrument der Gesundheitspolitik, die zu einer effektiveren Lösung von komplexen Problemen beitragen sollen:

Gesundheitsziele fördern gemeinsame Problemwahrnehmungen und Problemlösungen aller Akteure. Darüber hinaus ermöglichen sie ein planvolles, abgestimmtes Handeln und fördern die sektorenübergreifende Zusammenarbeit. Sie tragen dazu bei, die vorhandenen Mittel möglichst effizient und effektiv einzusetzen. Gesundheitsziele, die gemeinsam und im Kontext entwickelt werden, bieten den verantwortlichen Akteuren eine tragfähige Grundlage für ihre Aktivitäten.

(Bundesministerium für Gesundheit, 2006, Seite V).

In diesem Zusammenhang werden vom ärztlichen Zentrum für Qualität in der Medizin zurzeit nationale Versorgungsrichtlinien zur Behandlung von Depressionen entwickelt. Das Ziel, die Versorgungslage der depressiven Störungen zu verbessern, wird vom Bundesministerium wie folgt präzisiert:

Ziele sind ..., bei den Behandelnden, in der Pflege und in den Beratungsstellen die Kompetenz zu stärken, depressive Erkrankungen differenziert zu erkennen und mehrdimensionale Behandlungen im Rahmen fachübergreifender Kooperationsformen sicherzustellen. Besondere Bedeutung misst die Arbeitsgruppe dem Ziel zu, dass Kinder und Jugendliche sowie ältere Menschen mit depressiven Erkrankungen eine umfassende und angemessene Versorgung erhalten.

(Bundesministerium für Gesundheit, 2006, S.12)

### 2.1.7 Zusammenfassung

Depressionen sind psychische Störungen, über die es bereits zur Zeit der Antike erste schriftliche Zeugnisse gab. Sie sind durch eine dauerhafte negative Verzerrung der Stimmung gekennzeichnet und manifestieren sich in zahlreichen emotionalen, kognitiven, behavioralen, motivationalen und somatischen Beschwerden.

Um möglichst viel Transparenz bei klinischen Diagnosen depressiver Störungen zu erreichen, bedient man sich moderner Klassifikationssysteme, von denen hierzulande die ICD-10 verbindlich ist. In der ICD-10 werden depressive Störungsbilder hauptsächlich hinsichtlich ihrer Verlaufsform und des Schweregrades der Symptomatik differenziert. Man unterscheidet zwischen depressiven Episoden, rezidivierenden depressiven Störungen, bipolaren affektiven Störungen, Dysthymien und Anpassungsstörungen.

Depressive Erkrankungen zählen zu den am weitesten verbreiteten psychischen Störungen. Das Morbiditätsrisiko der deutschen Bevölkerung wird auf 9% für depressive Episoden im Rahmen unipolarer Erkrankungen, 4% für Dysthymien und 0.2% für bipolare Störungen geschätzt. Von den unipolaren Störungen sind Frauen etwa doppelt so oft betroffen wie Männer. Neben dem Geschlecht gelten depressive Erkrankungen in der Familie, niedriger sozioökonomischer Status, Mangel an sozialen Beziehungen, sowie belastende Lebensereignisse in der Vorgeschichte als Risikofaktoren. Es gibt zahlreiche Erklärungsansätze für depressive Störungen. Zu den klinisch relevanten Theorien, die zur Entwicklung wirksamer Therapiemethoden beigetragen haben, zählen psychodynamische, behaviorale, kognitive und neuro-chemische Modellvorstellungen. Dabei vermutet man, dass etwa 60% der Patienten auf eine der gängigen Therapieformen anspricht. Die Versorgung depressiver Erkrankungen in Deutschland wird durch Empfehlungen des Gesundheitsministeriums geregelt. Im ambulanten Bereich sind Antidepressiva, kognitive Verhaltenstherapie, tiefenpsychologisch fundierte Psychotherapie und Psychoanalyse zur

Behandlung anerkannt. In der stationären Versorgung kommen mitunter auch andere Therapieformen wie die interpersonale Psychotherapie zum Einsatz.

## *2.2 Depressive Störungen im Alter*

### *2.2.1 Begriffsbestimmung und historische Entwicklung*

Depressionen im Alter stellen innerhalb der Klinischen Psychologie einen eigenen Begriff dar.

Die sogenannte Alters- oder Involutionen-depression galt bis zur Einführung der ICD-10 zu Beginn der 90er Jahre als eine eigene Erkrankungsform, die innerhalb der psychiatrischen Klassifikationssysteme gesondert aufgeführt wurde (Kurz, 1997; Müller-Spahn, 2002). Der Begriff Involutionen-depression wurde dabei erstmalig im 19. Jahrhundert durch Kraepelin geprägt, der den Kern dieser Erkrankung vor allem in einem gestörten Körpererleben sah, das mit zahlreichen hypochondrischen Befürchtungen einhergeht (Schwärzer, 2003). Ein wichtiger Faktor für das Erleben des eigenen körperlichen Verfalls war dabei das höhere Lebensalter der Betroffenen. Diese Beschreibung der depressiven Erkrankung stellte bis zur 8. Ausgabe von Kraepelins Klassifikationssystem (1909-1915) eine eigene Krankheitsentität dar (Berrios, 1991), die auch von zahlreichen anderen Wissenschaftlern aufgegriffen und beschrieben wurde (Häfner, 1991; Seidel, Pantel & Schröder, 2004). Neuere Untersuchungen, die vor allem auf einer statistischen Reanalyse einer von Dreyfus zu Beginn des 20. Jahrhunderts durchgeführten klinischen Studie mit 85 Fällen beruhen, berücksichtigen zwar das Lebensalter des Erkrankten als Einflussgröße für den Verlauf der depressiven Erkrankung; die Abgrenzung eines eigenen besonderen Krankheitsbildes der Altersdepression gilt aber aufgrund der Heterogenität der Subgruppen von Erkrankungen als nicht mehr haltbar und wurde mit Einführung der ICD-10 abgeschafft (Berrios, 1991).

### *2.2.2 Aktuelle Definition und Klassifikation*

In den aktuellen Klassifikationssystemen werden depressive Störungen, die im höheren Lebensalter auftreten, unter den in Kapitel 2.1.2 aufgeführten, allgemein für das Erwachsenenalter geltenden Erkrankungsformen klassifiziert. Die neutralere Bezeichnung

„Depressionen im Alter“, die neben der Klassifikation verwendet wird, bezieht sich auf depressive Erkrankungen, die nach dem 60. Lebensjahr auftreten, unabhängig davon, ob es sich um Erstmanifestationen oder um Rezidive früher einsetzender Störungen handelt.

Da die Grenze zwischen jüngeren und „alten“ Menschen aufgrund sozialpolitischer Überlegungen in der Regel jedoch erst beim Überschreiten des 65. Lebensjahres gezogen wird (Davison & Neale, 2002), betreffen viele Forschungsbeiträge, die sich mit depressiven Erkrankungen im Alter beschäftigen, häufig auch die Altersgruppe der über 65-Jährigen.

### 2.2.3 Symptomatik

Da depressive Störungen im Alter nicht mehr gesondert klassifiziert werden, geht man per definitionem davon aus, dass bei Patienten im höheren Lebensalter grundsätzlich die gleichen Symptome vorliegen wie bei jüngeren. Dagegen sprechen viele Forschungsarbeiten für ein verändertes Krankheitsbild in der älteren Bevölkerung. Reischies (2003) weist darauf hin, dass Depressionssymptome im Alter häufig leicht oder atypisch sind. Forschungsüberblicke und Lehrbücher (Davison, Neale & Hautzinger 2007; Hammen & Watkins, 2008; Heuft, 1992; Heuft, Kruse & Radebold, 2006; Katona, 1994) berichten in vielen Fällen von vermehrtem Auftreten somatischer Beschwerden, kognitiver Beeinträchtigungen sowie von Sorgen und Nervosität bei einer geringeren Ausprägung klassischer Symptome wie Traurigkeit, Schuldgefühlen und Suizidgedanken. Dabei lässt sich ein einheitliches, im Alter typisches Krankheitsbild nicht eindeutig abgrenzen; vielmehr muss in diesem Lebensabschnitt von einer größeren Bandbreite und Variabilität der Symptome ausgegangen werden als bei jüngeren Patienten.

Das offizielle Fallbuch der WHO (Dilling, 2000) beschreibt einen der typischen Verläufe für eine im Alter auftretende depressive Verstimmung wie folgt:

Eine trauernde Witwe

Frau Randall ist eine...englische Witwe. Nach dem Tod ihres Mannes vor zehn Monaten verkaufte sie ihr Haus und zog in eine gemietete Wohnung.

*Aktuelle Problematik.* Frau Randall wurde in die örtliche Psychiatrische Poliklinik überwiesen, weil sie nach dem Tode ihres Ehemannes zunehmend depressiv geworden war. Frau Randalls Mann war vor zehn Monaten nach dreimonatiger Krankheit an einem malignen Lymphom verstorben. Er war zu Strahlenbehandlung und Chemotherapie ins Krankenhaus aufgenommen worden. Diese erwiesen sich jedoch als nutzlos, und er kehrte niemals nach Hause zurück. Frau Randall betrauerte den Tod ihres Ehemannes. Es gelang ihr aber, die Begräbnisfeierlichkeiten zu organisieren, ebenso wie eine Reihe von persönlichen Angelegenheiten, so den Verkauf ihres Hauses und ihres Lebensmittelgeschäftes und den Umzug in eine Mietwohnung. Dieses Appartement war in einer ihr unvertrauten Vorstadt, und Frau Randall fühlte sich dort völlig alleine. Sie versuchte zunächst, ihre Kontakte mit den Kindern und den Freunden aufrechtzuerhalten, aber schließlich ging sie nicht mehr aus dem Hause. Sie wurde immer depressiver, besonders in den vier Wochen vor der Überweisung. Frau Randall fühlte sich völlig freudlos und konnte sich für nichts mehr interessieren. Ihre gewohnten Alltagstätigkeiten wurden zu unüberwindlichen Hindernissen und sie war unfähig, noch irgendeine Hausarbeit auszuführen. Sie fühlte sich krank und hatte ein brennendes Gefühl in den Augen. Sie war ruhelos und konnte sich nicht entspannen. Sie spürte etwas wie einen engen Reifen um ihren Kopf. Ihr Schlaf war gestört, sie wachte früh auf, aß wenig und verlor in drei Monaten 6,8 kg Gewicht. Der Tod ihres Mannes war ständig in ihren Gedanken. Sie konnte sich mit seinem Tod nicht abfinden, und mehr als einmal hatte sie das unangenehme Gefühl, dass er in ihrer Wohnung anwesend sei. Sie dachte zwar schon, dass ihr Tod ... (eine) Erleichterung sein würde, aber wegen der Belastung, die dieses für die Kinder bedeuten würde, erwog sie einen Selbstmord nicht ernsthaft.

Dilling (2000, S.185-186).

Für die Komplexität verschiedener protektiver und risikobehafteter Faktoren, die bei älteren Patienten über die gesamte Lebensspanne hinweg wirksam werden, ist der folgende Abschnitt über die Vorgeschichte der Patientin besonders aufschlussreich:

Vorgeschichte. Frau Randall wuchs in einem Dorf im Süden Englands auf. Ihr Vater war Landwirt, und sie war die jüngere von zwei Schwestern. Nach der High-School arbeitete sie eine Reihe von Jahren in einem Textilgeschäft. Mit 24 heiratete sie einen Angestellten in einem Lebensmittelgeschäft, der zehn Jahre älter als sie war. Nach einigen Jahren eröffnete das Ehepaar in einer Großstadt

einen eigenen Lebensmittelladen. Sie hatten zwei Söhne, und als die Kinder groß genug waren, um in die Schule zu gehen, begann Frau Randall, auch mit im Lebensmittelgeschäft zu arbeiten. Nachdem vor zehn Jahren ein Supermarkt in der Nachbarschaft eröffnet worden war, ging es mit dem Laden der Familie Randall abwärts. Kurz vor Herrn Randalls Tod zeichnete sich ab, dass das Geschäft aufgegeben werden musste.

Die Mutter von Frau Randall und der Vater ihrer Mutter hatten unter depressiven Episoden gelitten, aber keine Behandlung erhalten.

Frau Randall beschrieb sich als aktive und extravertierte Frau, allerdings mit einer Neigung zu unerklärten Phasen von Ermüdung während der Winterszeit. Nach der Entbindung von ihrem jüngeren Sohn erlebte sie mit 28 Jahren eine zwei Monate dauernde Episode eines Stimmungstiefs und starker Ermüdbarkeit mit Schlafstörungen und Gewichtsverlust. Sie erholte sich davon schließlich ohne besondere Behandlung.

Mit 36 Jahren wurde Frau Randall drei Wochen lang wegen einer Depression in das örtliche Psychiatrische Krankenhaus aufgenommen. Dieser Behandlung war eine unharmonische Zeit in ihrer Ehe vorausgegangen, denn ihr Ehemann war dem Alkohol zugeneigt. Damals fühlte sie sich verzweifelt und dachte an Selbstmord. Sie war erschöpft und konnte gleichzeitig doch nicht zur Ruhe kommen, schlief schlecht und verlor Gewicht. Nach dreiwöchiger antidepressiver Behandlung besserte sich Frau Randalls Stimmung. In den zwei Jahren danach nahm sie eine Erhaltungsdosis desselben Medikaments. Seither fühlte sie sich ziemlich gut, wenngleich sie sich über vieles aufregte. Niemals erlebte sie Episoden von gehobener Stimmung oder euphorische Zustände.

Mit 40 Jahren wurde Frau Randall wegen variköser Venen zur Behandlung in ein Allgemeinkrankenhaus aufgenommen. Ihre Menopause trat mit 45 Jahren ein und sie erhielt seither gegen klimakterische Beschwerden eine hormonelle Substitutionsbehandlung.

(Dilling , 2000, S.186-187).

Die Diskussion dieses prototypischen Falls durch die Herausgeber veranschaulicht, dass nach der gegenwärtigen Klassifikation Depressionen im Alter keinesfalls mehr als eigene Krankheitskategorie angesehen werden.

Unter Berücksichtigung der Tatsache, dass ältere Patienten heterogene Krankheitsbilder mit sehr unterschiedlichen lebensgeschichtlichen Verläufen aufweisen, stellt dies eine sinnvolle Neuerung der Klassifikation nach der ICD-10 dar:

Während der gegenwärtigen Krankheitsphase fanden sich bei Frau Randall die Kriterien für eine mindestens mittelgradige depressive Episode (F32.1) mit allen drei typischen und vier zusätzlichen depressiven Symptomen, ebenso wurden die Kriterien für ein somatisches Syndrom erfüllt. Sie hatte bereits früher depressive Episoden von mindestens leichter Ausprägung; manische oder hypomanische Episoden waren nicht aufgetreten. Somit lautet ihre Diagnose: rezidivierende depressive Störung, gegenwärtig mittelgradige Episode mit somatischem Syndrom (F33.11). Eine Anpassungsstörung mit längerer depressiver Reaktion ist nicht ernsthaft zu erwägen, weil die Symptome nicht innerhalb eines Monats nach der Belastung eintraten, nämlich dem Tod des Ehemannes, sondern nach drei bis sechs Monaten. So treffen die Kriterien für eine depressive Episode zu, der Vorrang zukommt gegenüber der Anpassungsstörung. Im Übrigen ist wegen des rezidivierenden Bildes der depressiven Störung zusammen mit dem somatischen (melancholischen) Syndrom und der Familienanamnese unwahrscheinlich, dass ihr gegenwärtiges Krankheitsbild ausschließlich als Reaktion auf den Verlust ihres Ehemannes aufzufassen ist. Diese Belastung kann aber zusammen mit dem Umzug und dem Gefühl des Verlassenseins als ein mitauslösender Faktor der gegenwärtigen depressiven Episode verstanden werden.

(Dilling , 2000, S.186-187).

## *2.2.4 Epidemiologie*

### *2.2.4.1 Prävalenzen*

Die meisten Forschungsarbeiten vertreten die These, dass depressive Störungen in der Altenpopulation etwa ebenso häufig vorkommen wie in anderen Altersgruppen. Bickel (1997; Weyerer & Bickel, 2007) schätzt in einer Zusammenfassung zahlreicher Feldstudien aus Westeuropa und den USA von 1964-1995 die Prävalenzzahlen depressiver Erkrankungen, die die Kriterien für eine klinische Diagnose erfüllen, auf etwa 10%, wobei er für nur circa 1% die schwerste Erkrankungsform annimmt. Die Berliner Altersstudie (Helmchen et al., 1996, 2010), die sich auf DSM-III Diagnosen bezieht und somit eine Diagnostik gewährleistet, die

sich an den festgeschriebenen Krankheitsnormen orientiert, nennt Gesamtprävalenzraten von 9.1% für depressive Störungen in der Berliner Altenbevölkerung. Davon beziehen sich 5.4% auf die Diagnose Majore Depression (nach ICD-10 depressive Episode), 2.0% auf dysthyme Störungen, 0.7% auf Anpassungsstörungen und 1.0% auf Demenzen mit Depression. Eine besondere Rolle scheint in der älteren Bevölkerungsgruppe der sogenannten subdiagnostischen Symptomatik zuzukommen. Als subdiagnostische Symptomatik werden depressive Verstimmungen bezeichnet, die durch eine geringe Anzahl von Symptomen gekennzeichnet sind (Hautzinger, 2002). Die Depressivität ist bei diesen Krankheitsbildern nicht stark genug ausgeprägt, um das Vergeben einer klinischen Diagnose zu rechtfertigen. Man geht jedoch davon aus, dass die Lebensqualität der Betroffenen bereits deutlich eingeschränkt ist. Dabei gilt subdiagnostische Depressivität als ein Prädiktor für eine spätere depressive Erkrankung (Heuft, Kruse & Radebold, 2000). Helmchen et al. (1996) geben einen Anteil der Studienteilnehmer von 22% an, der eine subdiagnostische Depressivität unterschiedlichen Ausmaßes bejahte.

Eine spezielle Risikopopulation in der älteren Bevölkerung sind chronisch oder akut körperlich erkrankte stationäre Patienten in der Geriatrie. Dabei handelt es sich hierbei um eine Subgruppe älterer Patienten, über deren psychische Symptomatik es nur wenige methodisch fundierte Prävalenzuntersuchungen gibt. So beruhen Schätzungen über die Häufigkeit depressiver Erkrankungen meistens auf den Ergebnissen verschiedener Screening-Instrumente und sind nicht vergleichbar mit Surveys über den Erkrankungsstand der Bevölkerung. Einer Recherche der Datenbanken PsychInfo, Medline und Psyn dex nach Publikationen in englischer oder deutscher Sprache von 1990-2012 zu Stichwortkombinationen, die die depressive Erkrankung (depression, depr\*), das Setting (medical inpatients, geriatric inpatients, hospital), das Alter der Zielgruppe (older, geriatric, elderly) und den Typus der Studie (prevalence) berücksichtigten, ergab die in Tabelle 4 zusammengefassten Prävalenzangaben. Dabei wurden nur Studien einbezogen, die auf

Stichproben aus körperlich erkrankten stationären Patienten mit einem Mindestalter von 60 Jahren beruhen und Kriterien für depressive Erkrankungen zugrunde legen, die auf durch die Diagnosesysteme DSM, ICD oder AGE-CAT<sup>3</sup> (Copeland, Dewey & Griffith-Jones, 1986, 1990) zurückzuführen sind. Abhängig von den zugrundeliegenden diagnostischen Kriterien zeigen sich hier Prävalenzangaben von bis zu 36.0 % für ältere Menschen in der stationären geriatrischen Versorgung<sup>4</sup>. Die Heterogenität der in der Tabelle aufgeführten Erkrankungsdaten ist dabei einerseits auf die verschiedenen Diagnose-Systeme und Erhebungstechniken zurückzuführen<sup>5</sup>, andererseits werden körperliche Symptome der Depression, die auch durch andere Krankheitsbilder entstehen können, in den Studien jeweils unterschiedlich bewertet. Werden diese als Depressionssymptome nicht berücksichtigt, so fallen die Erkrankungsdaten wesentlich geringer aus. Zudem ergeben sich Unterschiede je nach körperlichen Komorbiditäten der Patienten sowie nach dem Prozentsatz der für depressive Erkrankungen anfälligeren weiblichen Teilnehmer der Studie. Darüber hinaus können weitere Faktoren eine Rolle spielen. So bezieht sich beispielsweise eine Studie von König (1991) ausschließlich auf männliche Kriegsveteranen, die aufgrund von Kriegstraumatisierungen eine besondere Risikogruppe bilden; somit ergibt sich hier ein hoher Prozentsatz an Betroffenen trotz der rein männlichen Stichprobe. Für eine genauere Beschreibung von Faktoren, die Depressionen begünstigen, sei auf Kapitel 2.2.6.1 zu den Risikofaktoren depressiver Störungen bei älteren Patienten verwiesen.

---

<sup>3</sup> AGE-CAT ist ein Computer-gestütztes Diagnose-System das Vergleichbare Kriterien zur Diagnostik depressiver Störungen verwendet, wie die ICD-10 oder das DSM-IV.

<sup>4</sup> Die in Tabelle 4 angegebenen Jahreszahlen beziehen sich jeweils auf den Forschungsartikel der jeweiligen Arbeitsgruppe, in der der Hauptdatensatz publiziert wurde. So hat beispielsweise die Forschungsgruppe McCusker et al... (2005, 2007 zahlreiche weitere Artikel publiziert, die weitere Analysen des entsprechenden Datensatzes vorstellen) (Cole, McCusker et al..., 2005, 2006, 2007, 2008; Han, McCusker et al..., 2006, 2008).

<sup>5</sup> Tabelle 5 zeigt die Qualitätsbeschreibung der einzelnen Studien. Das hierfür von der Autorin erstellte Qualitätsprotokoll ist im Anhang dieser Arbeit abgedruckt.

Tabelle 4: Prävalenzraten von depressiven Störungen bei stationären geriatrischen Patienten

Studie	Land	Population	N(Stichprobe)	Alter bei Einschluss (MW)	% weibliche Teilnehmer	Klassifikation	Prävalenz
McCusker et al. (2005)	Kanada	Akutkrankenhaus	380	65+ (79.4)	64.30	DSM-IV	MD 19.7 mD 9.13
Fenton et al. (1994)	Kanada	Akutkrankenhaus	215	65+ (76.0)	54.00	DSM-III	MD 27.4
Gantner et al. (2003)	USA	Akutgeriatrie/ geriatrische Frührehabilitation	60	(79.8)	66.66	DSM-III-R	Case 13.3
Kok et al. (1995)	Niederlande	Akutkrankenhaus	198	65+ (74.0)	42.00	AGECAT	Case 5.9
König et al. (1997)	USA	Akutkrankenhaus	542	60+ (70.0)	52.00	DSM-III-R	MD 21.58 mD 27.85
Linka et al. (2000)	Ungarn	Innere Medizin Akutgeriatrie/ geriatrische Frührehabilitation	100	65+ (73.0)	64.00	DSM-IV	MD 11.0
Ryan et al. (1995)	UK	Akutgeriatrie/ geriatrische Frührehabilitation	50	k. A. (82.1)	70.00	ICD-9	MD 36.0
Schmeling-Kludas et al. (2000)	Deutschland	Akutgeriatrie/ geriatrische Frührehabilitation	74	60+	68.00	ICD-10	Case 17.56
Burn et al. (1993)	UK	Akutgeriatrie/ geriatrische Frührehabilitation	100	65+ (82.0)	64.00	AGECAT	Case 23.0
Evans (1993)	UK	Akutgeriatrie/ geriatrische Frührehabilitation	72	65+ (80.2)	62.50	AGECAT	Case 31.94

Anmerkungen. MD = Major Depression, mD = minor Depression, D = Depressionsdiagnose nach ICD, Case = vergebene Depressionsdiagnose nach dem jeweils verwendeten Diagnosesystem.

*Fortsetzung Tabelle 4*

Studie	Land	Population	N(Stichprobe)	Alter, Einschluss (MW)	%weiblicher Teilnehmer	Klassifikation	Prävalenz
Schneider et al. (2000)	Deutschland	Akutkrankenhaus Akutgeriatrie/ geriatrische	262	60+ (75.1)	67.90	ICD-10	Case 14.1
Müller-Thomsen et al. (1999)	Deutschland	Frührehabilitation	104	65+ (78.19)	72.11	ICD-10	D. 22.11
Wancata et al. (2000)	Österreich	Innere Medizin	150	60+	62.70	DSM-III-R	MD 2.7 mD 7.3
Wancata et al. (2000)	Österreich	Chirurgie	94	60+	44.70	DSM-III-R	MD 3.2 mD 6.4
König et al. (1991)	USA	Akutkrankenhaus	332	70+	0.00	DSM-III	MD 13.2 mD 29.2

Anmerkungen. MD = Major Depression, mD = minor Depression, D = Depressionsdiagnose nach ICD, Case = vergebene Depressionsdiagnose nach dem jeweils verwendeten Diagnosesystem.

Tabelle 5: Qualitätsmerkmale der vorgestellten Prävalenzstudien

Studie	Rekrutierung	Einschlusskriterien		Respons-Rate		Stichproben- größe		Diagnostisches Interview				Ergebnisse	
	Vollerhebung oder Zufalls- auswahl	Kognition	Andere	Beschrie- ben	Hin- reichend: ≥ 75%	>200	>300	Zeitpunkt definiert	Strukturier- tes Interview	Geschul- ter Inter- viewer	Mit allen Proban- den	Ansatz beschrie- ben	Konfidenz- intervall für Effekt
McCusker et al. (2005)	+	++	+	+	+	+	+	+/-	+	+	+	+	+
Gantner et al. (2003)	-	++	+	+	+	-	-	+	-	-	+	-	-
Linka et al. (2000)	-	+/-	+/-	+	+	-	-	+/-	+	-	+	-	-
Schmeling-Kludas et al. (2000)	+	++	+	-	-	-	-	+/-	-	-	-	-	-
Schneider et al. (2000)	+	+	+	+	+	+	-	+/-	+	+	+	-	-
Wancata et al. (2000)	+	+	+	+	+	-	-	+	+	-	+	-	-
Müller-Thomson et al. (1999)	+	-	-	-	-	-	-	+	-	--	+	-	-
König et al. (1997)	-	+	+	+	+	+	+	+	+	-	+	+	-
Kok et al. (1995)	+	+	+	+	+	-	-	+/-	+	+	+/-	-	+
Ryan et al. (1995)	+	-	+/-	+	+	-	-	-	-	-	+	-	-
Fenton et al. (1994)	+	++	+	+	+	+	-	+	+	-	+	+	-
Burn et al. (1993)	-	+/-	+/-	+	+	-	-	+/-	+	+	+	-	-
Evans (1993)	+	+/-	+/-	+	+	-	-	+	+	+	+	+	-
König et al. (1991)	+	++	+	+	+	+	+	+	+	+	+	+	-

Anmerkung. ++ Kriterium gut erfüllt, + Kriterium befriedigend erfüllt, +/- Kriterium ausreichend erfüllt, - Kriterium nicht erfüllt. Vgl. Qualitätsprotokoll im Anhang.

### *2.2.5 Erkrankungsverläufe und Prognosen*

Die Angaben der Fachliteratur über einen grundsätzlich unterschiedlichen Verlauf von Depressionen im Alter sind widersprüchlich (vgl. Snowdon, 2001). Während einige Forschungsüberblicke befürworten, dass Erkrankungen im höheren Lebensalter durch längere Phasendauer, häufigere Therapieresistenz sowie erhöhte Rückfallhäufigkeit (Müller-Spahn, 2002) oder Chronifizierung (Katona, 1994; Kurz, 1997; Wetterling, 1999) gekennzeichnet sind, sprechen andere Autoren sich gegen grundsätzlich schlechtere Prognosen aus (Blazer, 1982; Hautzinger 2002). Ein weiteres Phänomen der Depression im höheren Lebensalter ist die Tatsache, dass trotz tendenziell zurückgehender Suizidraten sowie vermehrt milderer Erkrankungsformen die Anzahl der erfolgreichen Suizide mit dem Alter deutlich ansteigt (vgl. Nordhus, 2007). So sind nach einer Erhebung aus dem Jahr 2000 bei den unter 25-Jährigen 3,4 % von Suizidalität betroffen, bei den über 75-Jährigen steigt der Anteil auf 30,5% (Weyerer & Bickel, 2007).

Angesichts dieser heterogenen Forschungsergebnisse liegt die Vermutung nahe, dass nicht ein höheres Alter zwangsläufig mit einem ungünstigen Erkrankungsverlauf einhergeht, sondern dass bestimmte Gruppen älterer Patienten, die zusätzlich durch andere ungünstige Lebensbedingungen benachteiligt sind, eine höhere Chronifizierungswahrscheinlichkeit sowie ungünstigere Krankheitsverläufe aufweisen. Roose und Sackeim (2004) weisen darauf hin, dass es für ältere Populationen drei typische Erkrankungsmuster gibt, von denen jeweils circa ein Drittel der Erkrankten betroffen ist: Ein Drittel der Betroffenen ist nach einem Jahr remittiert und hat keine weiteren Beschwerden, ein Drittel ist remittiert, wird aber mindestens einen Rückfall erleiden und ein weiteres Drittel wird sich nach einem Jahr nicht von der Erkrankung erholen. In diesem Zusammenhang sei auch auf die im nächsten Abschnitt dargestellten Risikofaktoren für Depressionen im Alter verwiesen.

## 2.2.6 Ätiologie

Da sich die in Abschnitt 2.1.5 dargestellten Erklärungsansätze und Risikofaktoren depressiver Störungen auch auf Erkrankungen im höheren Lebensalter beziehen, wird an dieser Stelle ausschließlich auf Erklärungsansätze und Risikofaktoren eingegangen, die für die ältere Bevölkerungsschicht besonders bedeutsam sind.

### 2.2.6.1 Risikofaktoren

Relevante Faktoren für das Auftreten depressiver Erkrankungen im Alter sind vor allem Merkmale, die auf der Multimorbidität eines Teils der Altenpopulation beruhen, das heißt dem gleichzeitigen Vorliegen vieler verschiedenartiger Erkrankungen.

*Körperliche Erkrankungen.* Die Forschungsliteratur berichtet übereinstimmend (Bickel, 1997; Blazer, 2003; Djernes, 2006; Fiest, Currie, Williams & Wang, 2010; Hautzinger & Radebold, 1997; Meller et al., 1997) von einem ausgeprägten Zusammenhang zwischen chronischen körperlichen Erkrankungen, vermehrten Hausarztbesuchen oder Krankenhausaufenthalten und einem erhöhten Maß an Depressivität. Schmeling-Kludas, Jäger & Nieman (2000) fanden in einer Untersuchung körperlich erkrankter geriatrischer Patienten bei 18% der Stichprobe eine depressive Störung von Krankheitswert. Müller-Thomsen, Tabrizian & Mittermeier (2003) berichten bei 38.5% einer untersuchten Stichprobe mit hüftgelenksnahen Frakturen ein erhöhtes Maß an Depressivität. Arolt & Driessen (1997) geben für eine im Rahmen der Lübecker Allgemeinkrankenhausstudie untersuchten Stichprobe über 65-jähriger Patienten eine Prävalenzrate von 32.7% für die gesamte Gruppe depressiver Störungen an. Dabei ist die Richtung des Zusammenhangs zwischen körperlichen Erkrankungen und depressiven Verstimmungen kaum eindeutig festzustellen, da Depressionen ihrerseits einen Risikofaktor für zahlreiche körperlicher Erkrankungen darstellen. Eine prospektive Studie über einen Zeitraum von zwölf Jahren konnte so beispielsweise nachweisen, dass depressive Patienten im Alter von 50 bis 62 erhöhte Risiken für Diabetes, Herzprobleme und Arthritis aufweisen

(Karakus & Patton, 2011). Eine Metaanalyse über die Art des Zusammenhangs von Depressionen bei körperlich erkrankten Patienten und ihren pflegenden Angehörigen legt nahe, dass der Zusammenhang zu einem bedeutsamen Anteil durch einen Rückgang angenehmer Aktivitäten bedingt ist (Mausbach, Chattillion, Moore, Roepke, Depp & Roesch, 2011).

*Kognitive Beeinträchtigung.* Der Verlust kognitiver Fähigkeiten kann depressive Reaktionen hervorrufen, und zwar insbesondere dann, wenn die Urteilsfähigkeit des Betroffenen noch relativ unbeeinträchtigt ist. So fanden Helmchen et al. (1996) einen signifikanten Anstieg depressiver Symptomatik bei leichter kognitiver Beeinträchtigung; jedoch deutlich weniger Symptome bei schwerer Beeinträchtigung. Dabei weisen ältere Probanden mit leichter Depression eine besondere Sensitivität gegenüber negativem Feedback nach dem Bearbeiten kognitiver Testaufgaben auf (von Gunten, Herrmann, Elliott & Duc, 2011).

Im Hinblick auf kognitive Beeinträchtigung als Risikofaktor ist vor allem die Komorbidität von Depressionen im Alter mit demenziellen Erkrankungen empirisch bedeutsam. So vermutet man, dass etwa 30% der Patienten mit Demenzdiagnose zusätzlich an einer Depression leiden (Gallagher-Thomsen & Wolford-Thomsen, 1995).

*Neurobiologische Faktoren.* Medizinische Studien (Müller-Spahn & Hock, 1997) vermuten einen Zusammenhang zwischen der Verletzung bestimmter Hirnstrukturen und der Entwicklung einer depressiven Symptomatik. Diese Studien beruhen auf Beobachtungen, dass Erkrankungen, die einen negativen Einfluss auf die zerebrale Funktionsfähigkeit in bestimmten Hirnregionen haben, wie die Alzheimerische Erkrankung, Morbus Parkinson sowie Multiple Sklerose, zerebrale Gefäßerkrankungen und Hirninfarkte gehäuft mit depressiver Symptomatik einhergehen.

*Medikation.* Einige Medikamente können als Nebenwirkung depressive Symptome auslösen. Hierzu zählen hauptsächlich Antikonvulsiva, einige Antihypertonika, Antiparkinsonmittel und Hormone (Hautzinger, 2002; Heuft, Kruse & Radebold, 2000). Dabei stellen

Medikamente keinen Risikofaktor im klassischen Sinne dar, da depressive Krankheitsbilder, die nachweislich durch bestimmte Substanzen ausgelöst werden, nicht zu den affektiven Störungen zählen (Hautzinger & Meyer, 2002), sondern in der ICD-10 in Kapitel F1 bei den durch psychotrope Substanzen ausgelösten psychischen Störungen klassifiziert werden. Entscheidend für eine solche Kodierung ist die Annahme eines Kausalzusammenhanges zwischen der Substanzeinnahme und dem Ausbruch der psychischen Störung. Affektive Auffälligkeiten stellen dabei nur eine mögliche Form der Reaktion auf psychotrope Substanzen dar und werden durch hintere Kennziffern kodiert (WHO, 2005, 2010). Stellen medikamentöse Nebenwirkungen dagegen nur einen möglichen Faktor für den Ausbruch der Störung dar, wird das Krankheitsbild wie beschrieben als affektive Erkrankung kodiert.

*Institutionalisierung/ Pflegebedürftigkeit.* Der Verlust der Selbstständigkeit scheint stark mit Depressivität assoziiert zu sein. Die Münchener Hochbetagten Studie (Meller et al., 1997) fand unter den pflegebedürftigen Probanden einen etwa doppelt so hohen Anteil Depressiver wie bei den Teilnehmern mit selbstständiger Haushaltsführung. Bei Untersuchungen mit Heimbewohnern sind Depressionsraten zwischen 15 und 25% ermittelt worden (Hautzinger, 2002), einige Studien berichten sogar Werte bis zu 40% (Bickel, 1997). Kruse (1992) weist darauf hin, dass Depressionen bei Pflegebedürftigkeit vermehrt auftreten, wenn die Bedürftigen ein zu hohes Maß an Unterstützung erfahren und dadurch noch mehr in die Passivität gedrängt werden. Dieser Mechanismus scheint auch in umgekehrter Richtung wirksam zu sein. So berichtet Osterman (1997) über statistisch hoch signifikante Korrelationen zwischen der Verminderung der Pflegebedürftigkeit im Rahmen eines Rehabilitationsaufenthaltes und einer Verminderung der Depressivität bei geriatrischen Patienten.

Aus den Forschungsbeiträgen wird deutlich, dass sich eine ätiologische Richtung des Zusammenhanges zwischen Multimorbidität und Depressivität nicht eindeutig bestimmen lässt. So können verschiedenartige Erkrankungen, wie Katona (1994) zusammenfasst, je nach

Interpretation der empirischen Befunde als Ursache oder als Folge depressiver Erkrankungen vermutet werden, der Zusammenhang kann zufällig sein oder nur unter bestimmten Bedingungen wie etwa Pflegebedürftigkeit eine Rolle spielen. Eine sinnvolle Klärung des Verhältnisses von Ursache und Wirkung ist daher eher in der qualitativen Interpretation individueller Krankheitsverläufe möglich und kaum durch quantitative Erhebungen zu erreichen.

Dabei stellt eine angemessene psychologische und psychiatrische Versorgung depressiver geriatrischer Patienten eine große Herausforderung für die klinische Praxis dar. So gilt zwar inzwischen als relativ gesichert, dass auch Patienten im höheren Lebensalter gut auf antidepressive Medikation ansprechen (Möller, 1997), über die Möglichkeiten psychoanalytischer Therapie bei älteren Patienten (Radebold, 1997; Zank, Peters & Wilz, 2010) ist dagegen noch wenig bekannt. Erste Studien hinsichtlich der Wirksamkeit kognitiver Verhaltenstherapie bei Depressionen im Alter (Hautzinger & Welz, 2004) zeigen ermutigende Ergebnisse für das Ansprechen geriatrischer Patienten auf psychologische Intervention. Zudem konnte Zank (2002) in einer Studie zeigen, dass sich Vorbehalte bei Psychoanalytikern gegenüber älteren Patienten hauptsächlich bei denjenigen Therapeuten zeigen, die keine Erfahrung mit dieser Klientel haben.

#### *2.2.6.2 Psychologische und Psychiatrische Erklärungsansätze*

Die in Kapitel 2.1.5.2 beschriebenen Erklärungsansätze für das Entstehen depressiver Störungen haben selbstverständlich auch bei Patienten im höheren Lebensalter ihre Gültigkeit. Darüber hinaus gibt es innerhalb der gerontologischen Forschung einige theoretische Modelle, die einen Beitrag dazu leisten können, das Phänomen depressiver Störungen bei geriatrischen Patienten besser zu verstehen. Diese Ansätze, wie z.B. das Modell der Selektiven Optimierung durch Kompensation von Baltes und Baltes (1989) oder das Modell von Assimilation und Akkommodation nach Brandstädter und Renner (1992) sind keine Theorien zur Ätiologie von Depression im Alter im engeren Sinne, können aber einen

Rahmen zur genaueren Beschreibung und Erklärung des Krankheitsbildes im höheren Lebensalter bilden.

### *Das SOK-Modell*

Das SOK Modell (S = Selektion, O = Optimierung, K = Kompensation) wurde von Baltes (1989) als Metatheorie zur Erklärung von erfolgreichem<sup>6</sup> Altern entwickelt. Die Grundthese dieses Ansatzes besagt, dass Individuen sich verschiedener Strategien bedienen können, um ihre Ziele trotz verschiedener Einschränkungen zu erreichen. Ein erfolgreicher Anpassungsprozess an einen Rückgang gesundheitlicher und sozialer Ressourcen im Alter besteht in einer sorgfältigeren Auswahl der eigenen Lebensziele (Selektion), in der Optimierung der hierfür benötigten Ressourcen (Optimierung), sodass eine Kompensation in Bezug auf das Weniger der absolut vorhandenen Ressourcen erreicht wird. Beispielsweise kann eine ältere Patientin nach dem Erleiden eines Oberschenkelhalsbruchs andere Ziele zurückstellen, um sich zunächst dem Ziel „Die Rehabilitation erfolgreich bewältigen“ zu widmen. Darauf kann sie im Anschluss sowohl ihre gesamte körperliche Kraft als auch ihre sozialen Ressourcen verwenden, indem sie beispielsweise täglich mit ihrer Tochter telefoniert, um Trost und Hilfe zu erhalten (soziale Unterstützung & instrumentelle Unterstützung).

Ältere Menschen, die entsprechende Strategien nicht oder nicht ausreichend anwenden können, leiden häufiger unter Zielkonflikten, etwa wenn die körperliche Kraft nicht ausreicht, sowohl die Rehabilitation zu bewältigen, als auch den eigenen Haushalt ohne fremde Hilfe weiterzuführen. Die daraus resultierende Überlastung kann schließlich zu depressiven Verstimmungen führen.

---

<sup>6</sup> Der Terminus „erfolgreiches Altern“ ist innerhalb der Forschungsgemeinschaft umstritten. Während der von Baltes geprägte Begriff immer noch in zahlreichen Studien verwendet wird, verwenden einige Autoren eher den neutraleren Begriff „gelingendes Altern“, um den Fokus weniger auf die vermeintliche Eigenleistung des positiv verlaufenden Alternsprozesses zu legen.

### *Assimilation und Akkommodation*

Das Modell von Assimilation und Akkommodation nach Brandstädter und Renner (1992) ist ebenfalls ein Ansatz, der den individuell unterschiedlichen Umgang mit Zielkonflikten beschreibt. Dabei sieht Brandstädter grundsätzlich zwei Möglichkeiten des Umgangs mit einer Situation, in der die ursprünglichen persönlichen Ziele aufgrund eines Verlustes vorhandener Ressourcen nicht ohne weiteres erreicht werden können, die er als Assimilation und Akkommodation bezeichnet. Mit Assimilation ist das Bestreben eines handelnden Individuums gemeint, die Umweltbedingungen so zu verändern, dass sie mit dem Organismus zu einer optimalen Passung kommen. Akkommodation bezeichnet dagegen die Anstrengung des Organismus, sich seinerseits seinen Umweltbedingungen anzupassen. Übertragen auf den Zielkonflikt eines geriatrischen Patienten nach einer Oberschenkelfraktur kann dies bedeuten, dass der Betroffene weiterhin versucht, zuvor festgelegte Ziele, wie z.B. die Teilnahme an einem Wanderurlaub trotz der erschwerten Bedingungen zu verfolgen. Hierfür müssten die Umweltbedingungen etwa derartig angepasst werden, dass der Patient sich für das doppelte Maß an Physiotherapie einsetzt und versucht, den Termin für den Urlaub um eine Woche zu verschieben etc. Je nach Diskrepanz des ursprünglichen Ziels und der vorhandenen Ressourcen ist dieser Umgang mit dem inneren Konflikt mehr oder weniger konstruktiv. Erfordert das Erreichen des Ziels eine geringe bis mittlere Assimilationsleistung, so kann sie den Patienten motivieren und somit seinen Heilungsprozess fördern. Ist die geforderte Anpassungsleistung jedoch zu hoch, so können Erschöpfung, erhebliche Frustration und Depression die Folge sein.



Abbildung 5: Entwicklung durch Akkommodation und Assimilation

Entscheidet sich der Patient dagegen für die Strategie der Akkommodation, so bedeutet dies einen inneren Prozess der Neubewertung und Modifikation des ursprünglich gesetzten Zieles, im Sinne einer Lösung der Zielbindung. So kann der Betroffene etwa zu dem Schluss kommen, dass er den Wanderurlaub solange verschieben möchte, bis der Bruch vollkommen abgeheilt ist. Statt des ursprünglich geplanten Wanderns hätte er dann beispielsweise Ressourcen frei, um in der Zwischenzeit eine neue, gelenkschonende Sportart zu erlernen. Diese Form der Anpassung kann dann schädlich sein, wenn ursprünglich gesetzte Ziele bei leichteren Schwierigkeiten bereits bereitwillig aufgegeben werden. In diesem Fall kann es zu einer Grundhaltung der Resignation kommen, die die Entwicklung depressiver Erkrankungen fördert.

*Psychodynamische Erklärungsansätze.* Die psychoanalytische Schule war aufgrund der Tatsache, dass ihr Begründer Sigmund Freud Vertreter eines ausgesprochen defizitorientierten Altersbildes war, lange Zeit der Psychotherapie von Patienten mittleren und höheren Lebensalters grundsätzlich abgeneigt. Eine Beschäftigung mit den allgemeinen und

spezifischen Bedingungen psychodynamischer Vorgänge im höheren Lebensalter fand daher im Wesentlichen während der letzten 30 Jahre statt (vgl. Zank, Peters & Wilz, 2010).

Dabei war es Freud selbst, der bereits einige wichtige Argumente für die Sinnhaftigkeit psychodynamischer Therapie im höheren Lebensalter in seinen Schriften erwähnte. So ist das Unbewusste nach Freud nicht den gleichen zeitlichen Bedingungen unterworfen wie bewusste oder vorbewusste Gedächtnisinhalte; vielmehr bleiben alte Bilder nach der Sicht Freuds in ihrer Bedeutung erhalten, unabhängig davon, wie viel Zeit seit ihrer Speicherung verstrichen ist. Diese „Zeitlosigkeit des Unbewussten“ bedeutet, dass die entsprechenden inneren Konflikte auch im höheren Lebensalter virulent sind und entsprechend bearbeitet werden können.

Neben dem Argument, dass die Bedingungen des psychischen Organismus in jedem Lebensalter ähnlich sind, birgt das höhere Lebensalter einige spezifische Aspekte, die bei der Theoriebildung psychodynamischer Ansätze besonders berücksichtigt wurden. So wird das Altern mit seinen zahlreichen Entwicklungsverlusten (Verlust der körperlichen Attraktivität und Stärke, Verlust sozialer Rollen, Verlust sozialer Beziehungen, immer stärkeres Abweichen vom Ich-Ideal) vom Selbst als eine narzisstische Krise erfahren, die es zu verarbeiten gilt. Zudem wird der alternde Mensch verstärkt auf die existenziellen Dimensionen des Daseins, nämlich auf das Bewusstsein der eigenen Sterblichkeit, zurückgeworfen, was mit zahlreichen Ängsten verbunden ist. Als Möglichkeiten, den eigenen Selbstwert zu stabilisieren und peinliche und unangenehme Gefühle abzuwehren, stehen dem alternden Menschen dabei grundsätzlich dieselben Abwehrmechanismen zur Verfügung wie jüngeren Menschen. Dabei nennt Peters (Zank, Peters & Wilz, 2010) eine Reihe reifer und früher Abwehrmechanismen, deren Verwendung in diesem Lebensabschnitt besonders charakteristisch sind (vgl. Tabelle 5).

*Tabelle 5: Häufige Abwehrmechanismen des höheren Lebensalters*

Frühe Abwehrmechanismen	Reife Abwehrmechanismen
<p><b>Verleugnung:</b> Den Verleugneten innerpsychischen Aspekten oder Merkmalen in der realen Welt wird eine scheinbare Indifferenz entgegengebracht.</p>	<p><b>Unterdrückung:</b> Zeitweises Ausblenden von zugänglichen Bewusstseinsinhalten, die als belastend erlebt werden i.S. von Selbstdisziplin.</p>
<p><b>Spaltung:</b> Zwei sich gegenseitig widersprechende Bewusstseinsinhalte existieren parallel, ohne dass sich das handelnde Subjekt dieses Widerspruchs bewusst wird.</p>	<p><b>Humor:</b> Schaffung einer scherzhaften Distanz zu den unvermeidlichen Aspekten des Daseins.</p>
<p><b>Projektion:</b> Als bedrohlich empfundene altersbedingte Einschränkungen werden nicht bewusst als solche erlebt, sondern nach außen getragen, indem z.B. Objekte in der äußeren Welt als unzulänglich, bössartig oder verfolgend abgewertet werden.</p>	
<p><b>Verschiebung:</b> Ein valenter psychischer Inhalt wird auf einen anderen weniger bedeutenden Bereich verlagert. Beispielsweise die Kontrolle über den eigenen Körper in Richtung Kontrolle über die Umwelt.</p>	<p><b>Sublimation:</b> Modifikation ursprünglicher Triebwünsche im Sinne einer Verschiebung des ursprünglichen Zieles auf höhere kulturelle Inhalte.</p>
<p><b>Somatisierung:</b> Verschiebung psychischer Inhalte auf die somatische Ebene.</p>	
<p><b>Omnipotenz:</b> Leugnung der Endlichkeit des Lebens und Festhalten an einem Größen-Selbst.</p>	<p><b>Altruismus:</b> Partielles Sich- Abwenden von der eigenen Person und aktive, unterstützende Anteilnahme am Leben anderer Menschen.</p>
<p><b>Reaktionsbildung:</b> Unbewusstes, überzogenes Festhalten an Gewohnheiten, Werten und Einstellungen, um gegen die Ungewissheit des Lebens besser gewappnet zu sein.</p>	<p><b>Planende Vorwegnahme:</b> Ein vorwegnehmendes sich Einstellen auf die kommenden unausweichlichen Verluste des Alters.</p>

Ein sukzessives Durcharbeiten der inneren Konflikte des höheren Lebensalters kann den alternden Menschen dabei unterstützen, die frühen Abwehrmechanismen immer mehr durch reifere Verarbeitungsstrategien zu ersetzen und dadurch aversive Affekte wie Enttäuschung, Traurigkeit, Neid und Eifersucht, Ärger, Schuld und Scham sowie Verzweiflung

einzudämmen. Gelingt dieser Prozess, so ist das Ergebnis Weisheit i.S. einer Weltsicht, die über das alleinige Lernen von Wissensinhalten weit hinausgeht und die ein starker protektiver Faktor gegen die Entwicklung depressiver Verstimmungen sein kann. So kann dann eine „Art erfüllte und gelöste Anteilnahme am Leben im Angesicht des Todes“ (Erikson, 1980, zitiert nach Zank, Peters & Wilz, 2010, S.119), entstehen (vgl. zu diesem Abschnitt Zank, Peters & Wilz, 2010).

#### *Die Entwicklungstheorie nach Erikson.*

Der Psychoanalytiker und Entwicklungspsychologe E.H. Erikson war einer der ersten Forscher, der den Verlauf der Lebensspanne als einen Prozess des lebenslangen Lernens beschrieben hat. Dabei hat er für verschiedene Lebensabschnitte verschiedene zentrale normative Themen und Krisen beschrieben, die es in den entsprechenden Entwicklungsstadien zu bewältigen gilt; bzw., mit denen sich das sich entwickelnde Individuum auseinandersetzen muss. Das Scheitern an der entsprechenden Krise bzw. die Unfähigkeit diese zu überwinden, kann jeweils die Entwicklung einer Depression zur Folge haben. Für das höhere Lebensalter postuliert Erikson eine Reihe von Herausforderungen, die mit den sozialen Rollen bzw. den biologischen Bedingungen des Alterns verknüpft sind und mit denen der alternde Mensch einen Umgang finden muss (vgl. Peters, 2004):

*Sorge – Generativität mit Stagnation.* Die Fürsorge für die nachfolgende Generation, die eine zentrale Aufgabe des mittleren Lebensalters darstellt und einen Großteil der sozialen, beruflichen und familialen Rollen in diesem Lebensabschnitt ausmacht, ist eine Aufgabe, die im höheren Lebensalter nur noch in abgewandelter Form ausgeführt werden kann und muss. So ist es in der Regel nicht die Aufgabe der Großeltern, das direkte Überleben ihrer Enkelkinder zu sichern und durch die Verrentung sind viele Aspekte des Berufslebens, bei denen es um die Vermittlung von Wissen und Fertigkeiten an nachfolgende Generationen geht, nicht mehr Teil des alltäglichen Lebens. Nach Erikson muss der alternde Mensch in dieser Phase eine neuere, abstraktere und auch zugleich allgemeinere Form der Generativität

für sich entdecken, die von den Autoren auch als Grand Generativität bezeichnet wird. Dabei geht es um das Sichern des Überlebens der Menschheit in einem allgemeineren Sinne und nicht um eine direkte Unterstützung der nachfolgenden Generation.

Wenn ein alternder Mensch diese Form von Generativität nicht für sich als Alternative zu der direkteren Form entdecken kann, ist ein Gefühl der Stagnation und damit verbunden ein Frustrationsgefühl die Folge, das beispielsweise zur Entstehung einer depressiven Anpassungsstörung führen kann.

*Liebe – Intimität mit Isolation.* Der Verlust nahestehender Bezugspersonen oder eine räumliche Isolation von der eigenen Familie, stellen eine weitere Herausforderung dar. Die Gefahr der empfundenen Isolation im höheren Lebensalter kann jedoch nach Erikson durch eine Integration von Erinnerungen an frühere Bindungserfahrungen sowie durch ein sich Zuwenden zu der äußeren Welt mit den darin noch enthaltenen Möglichkeiten, Beziehungen mit anderen Menschen einzugehen, überwunden werden. Gelingt dies nicht, so ist ein Gefühl der Einsamkeit die Folge, das einen der wichtigsten Risikofaktoren für die Ausbildung einer Depression im höheren Lebensalter ausmacht.

*Kontinuität – Identität mit Identitätsdiffusion.* Viele Rollen, die eng mit dem Identitätsempfinden jüngerer Menschen verbunden sind, z.B. die Mutterrolle, oder die Rolle des Arbeitnehmers, müssen im höheren Lebensalter nicht mehr ausgefüllt werden. Dabei besteht für das Individuum die Gefahr einer Identitätskrise, die nur durch das Auffinden von Elementen der Identität, die nicht von den Rollen des mittleren Erwachsenenalters abhängig sind, überwunden werden kann. Gelingt diese Verschiebung des eigenen Identitätsempfindens auf allgemeinere Aspekte des eigenen Daseins nicht, können Gefühle von Sinnlosigkeit und Hoffnungslosigkeit entstehen, die ihrerseits Symptome der depressiven Erkrankung sind.

*Kompetenz – Werksinn mit Minderwertigkeitsgefühl.* Das Nachlassen der eigenen Leistungsfähigkeit, das im Alter zu Minderwertigkeitsgefühlen führen kann, kann in dieser Lebensphase durch die Besinnung auf frühere Erfolge und das damit verbundene Gefühl von

Stolz, durch Stolz auf die Leistung der nachfolgenden Generation sowie durch ein Aufrechterhalten weiterhin möglicher Tätigkeiten erreicht werden.

*Zugang zur Welt – Initiative mit Schuld.* Die Konfrontation mit der eigenen Endlichkeit, und ein möglicherweise daraus resultierendes Gefühl, sich selbst und der Welt etwas schuldig geblieben zu sein, kann durch einen Prozess der Trauer um verpasste Chancen und vertane Gelegenheiten sowie durch eine weiterhin aufgeschlossene Haltung der äußeren Welt gegenüber überwunden werden. Gelingt der Prozess der Trauer nicht, so kann die Folge eine pathologische Trauerreaktion mit ambivalenten Gefühlen sich selbst und der eigenen Lebensgeschichte gegenüber sein. Zudem kann bezüglich der eigenen Fehlleistungen in der Vergangenheit ein starkes Schuld- und Schamempfinden entstehen, das mit depressivem Empfinden einhergeht.

*Willenskraft – Autonomie mit Zweifel und Scham.* Der Verlust der Kontrolle über den eigenen Körper, der zunächst Gefühle von Zweifel und Scham auslöst, kann durch eine Integration der extremen Pole möglicher Reaktionen wie totaler Resignation auf der einen Seite bzw. der Entwicklung einer gewissen Starrsinnigkeit auf der anderen Seite, verarbeitet werden. Hier zeigt sich eine Parallele zum Modell der Assimilation und Akkommodation von Brandstädter. Der Verlust wichtiger Bereiche der Autonomie, wie etwa der Fähigkeit, selbst auf die Toilette gehen zu können, muss zunächst durch die betroffene Person akzeptiert werden, damit ein positives Selbstbild wiederhergestellt werden kann. Gleichzeitig würde bei einem Zuviel an Akzeptanz und einem Zuwenig an Aktivität die Gefahr entstehen, auch noch weitere wichtige Funktionen zu verlieren, wenn z.B. das Bett überhaupt nicht mehr selbstständig verlassen wird. Sowohl das zu starre Festhalten an verlorengegangenen Funktionen als auch das zu bereitwillige Aufgeben derselben können zu einer depressiven Verstimmung führen.

*Hoffnung – Vertrauen mit Misstrauen.* Im hohen Lebensalter stellt die Konfrontation mit der eigenen Sterblichkeit eine große Herausforderung dar. Die Besinnung auf ein Gefühl des Urvertrauens, das idealiter in der frühen Kindheit erworben wurde, kann dem alternden

Individuum dazu dienen, diesen Konflikt durch eine Transzendierung der eigenen Welt- und Selbstsicht sowie durch eine Hinwendung zur Spiritualität zu verarbeiten. Alternde Menschen, die das Vertrauen in die Sinnhaftigkeit des Lebens- und Sterbe-Prozesses nicht erleben können, sind in diesem Lebensabschnitt sehr anfällig für Gefühle der Verzweiflung, Hoffnungslosigkeit und Depression.

*Weisheit – Integration mit Verzweiflung.* Im Alter müssen alle Lebensthemen, die im Verlauf der Entwicklung relevant waren, erneut durchlebt und die dazu gefundene Haltung muss modifiziert werden, um einer Verzweiflung über den Verlust der realen Welt vorzubeugen. Wenn das gelingt, so kann dies als Weisheit im Sinne von Erikson bezeichnet werden. Wenn nicht, so besteht die Gefahr, den zahlreichen Verlusten, die dieser Lebensabschnitt mit sich bringt, nichts entgegensetzen zu können, was vor einem Gefühl von Leere und Verzweiflung schützt.

#### *Kognitiv behaviorale Erklärungsansätze*

Kognitiv behaviorale Erklärungsansätze betrachten psychische Störungen im Alter allgemein als dysfunktionales Verhalten sowie als dysfunktionale Gedankenmuster (vgl. 2.1.5). Dabei sind Annahmen über die Entstehung psychischer Störungen eng am Diathese-Stress-Modell orientiert, wobei das Hauptaugenmerk in der KVT allerdings weniger auf den ursächlichen Faktoren der Störung liegt, sondern stärker auf die aufrechterhaltenden Bedingungen der gestörten Interaktion zwischen Organismus und Umwelt fokussiert. Spezifische Aspekte der KVT im Alter beziehen sich einerseits auf besondere Lebensumstände älterer Patienten, die ursächlich bzw. aufrechterhaltend im Hinblick auf eine psychische Störung wirken können.

Hier werden von Rabiolo-Fischer (2008, zitiert nach Zank, Peters & Wilz, 2010, S. 187)

beispielsweise einige Aspekte genannt, die besonders auf ältere Patientinnen zutreffen:

- Fürsorge für andere als wichtiges Lebensthema und eine hohe subjektive Beeinträchtigung, wenn diese wegfällt
- Ängste vor psychischer und physischer Wehrlosigkeit; Verunsicherung, wenn die klassische Rollenaufteilung in der Ehe wegfällt (z.B. bei Aufgaben, die technische oder handwerkliche Fähigkeiten erfordern),
- Langes Alleinsein (Witwenschaft: hochsignifikante Anstiege für Depressionen und Hospitalisierungen),
- Existenzängste
- Traumatisierung durch Überfälle (ältere Frauen erleben sich als besonders gefährdet).

Dabei spielt bei der Arbeit mit älteren Patienten die Aktivierung vorhandener Ressourcen, die protektiv auf die Situation einwirken und Entwicklungsverluste kompensieren können, eine besondere Rolle. Die Zielformulierung des therapeutischen Prozesses im Sinne eines gelingenden Anpassungsprozesses für diese Zielgruppe kann somit auch als eine Förderung der Prozesse Zielselektion, Optimierung von Ressourcen und Kompensation von Verlusten, wie im SOK-Modell nach Baltes & Baltes (1990, vgl. Abschnitt 2.2.6.2) beschrieben, dargestellt werden (Zank, Wilz & Peters, 2010). Hautzinger (2000, zitiert nach Zank, Wilz & Peters, 2010, S. 193) präzisiert diese Schritte in einem altersspezifischen Erklärungsmodell zur Entstehung und Behandlung depressiver Störungen wie folgt:

- Selektion: Ziele, Ansprüche und Wünsche den vorhandenen Möglichkeiten und der körperlichen Verfassung anpassen
- Bearbeiten und Aufgeben alter Enttäuschungen und Hoffnungen
- Optimierung: Depressionsfördernde Bedingungen im Leben der Patienten verringern (z.B. sozialer Rückzug, belastende Wohnbedingungen); wichtige Bezugspersonen mit in die Behandlung integrieren.
- Kompensation: Aktives, nicht depressives und die Umwelt kontrollierendes Verhalten aufbauen oder wieder aktivieren; dysfunktionale, wenig hilfreiche, resignative Kognitionen reduzieren und durch hilfreiche, selbstwertförderliche ersetzen.

Ziel der Therapie ist es, die genannten Ziele durch konkrete Techniken zu erreichen (vgl. Abschnitt 2.2.8.2).

### *2.2.7 Versorgung depressiver Störungen im Alter*

Die Versorgungslage hinsichtlich der Diagnostik und Behandlung von depressiven Störungen im Alter ist derzeit in einigen Aspekten noch deutlich defizitär (Heidenblut & Zank, 2009). Eine mögliche Ursache hierfür mag die im Alter deutlich komplexere Verschränkung körperlicher Erkrankungen und depressiver Verstimmungen sein, die von Häfner (1986, S.34-35) wie folgt auf den Punkt gebracht wird:

Ein passiver Lebensstil begünstigt Bewegungsmangel und Überernährung. In deren Folge kann sich ein Diabetes mellitus entwickeln, der nach langjährigem Bestehen das Auftreten einer Depression und die Entwicklung eines zerebralen Gefäßprozesses begünstigt. Die depressive Krankheit hat eine Vernachlässigung diätischer Regeln und medizinischer Verordnungen zur Folge, die zur Dekompensation des Diabetes und, in Zusammenhang damit, zu einem Hirninfarkt führt. Als Folge des Hirninfarktes verbleibt eine Demenz mit Pflegebedürftigkeit.

Das hier beschriebene Szenario stellt einerseits einen besonders ungünstigen Erkrankungsverlauf dar: Aufgrund der Begleiterkrankungen ist es hier sehr schwierig, eine angemessene Behandlung für die depressive Erkrankung zu finden, da die zerebrale Vulnerabilität des Patienten den Einsatz zentral wirksamer Psychopharmaka zu einem Risiko werden lässt und gleichzeitig die geringe körperliche und kognitive Belastbarkeit die Durchführung psychotherapeutischer Maßnahmen nahezu unmöglich machen dürfte.

Andererseits stellen entsprechende Fälle in der geriatrischen Praxis keineswegs eine Ausnahme dar. Dabei fehlt es „dem Hausarzt, der für diese Patientengruppe die erste Anlaufstelle darstellt, häufig [an] ...Fachwissen, um zwischen primär psychogenen [und] somatischen Auffälligkeiten im Sinne depressiver Symptome und körperlich begründeten Beschwerden sicher zu unterscheiden“. In einer Metastudie mit 633 Hausarztpraxen wird beispielsweise eine Rate korrekter Diagnosen nach ICD-10 von lediglich 50% berichtet

(Jacobi et al., 2002)“, (Heidenblut & Zank, 2009, S. 17.). Statistiken zur Behandlung älterer Patienten sprechen eine ähnlich deutliche Sprache:

Die Berliner Altersstudie berichtet nur bei 40 der 133 von Depressionen betroffenen Probanden eine Versorgung durch Psychopharmaka. Keiner der Patienten wurde an einen Facharzt überwiesen oder psychotherapeutisch behandelt (Wernicke & Linden, 1997). Dieser Befund überrascht nicht, wenn man weitere Studien zur Versorgungslage älterer Depressiver hinzuzieht: So machen die über 60-Jährigen, die in der Allgemeinbevölkerung einen Prozentsatz von 20% stellen, nur 5% der Patienten in ambulanten psycho-therapeutischen Praxen aus; dabei geben 90% der Therapeuten sogar explizit an, Patienten dieser Altersgruppe nicht behandeln zu wollen (Richter, 2006). Dieser (psycho)-therapeutische Nihilismus gegenüber der Behandlung älterer Patienten scheint sich in der praktischen Erfahrung nicht zu bestätigen; so war in einer Befragung von Berliner Psychotherapeuten Vorerfahrung mit älterem Klientel der Hauptprädiktor für die Bereitschaft zur Übernahme weiterer Patienten im höheren Lebensalter (Zank, 2002).<sup>7</sup>

#### *2.2.7.1 Aktuelle Entwicklung*

Einen innovativen Umgang mit den zahlreichen Komplikationen bei der Versorgung älterer depressiver Patienten stellt das sogenannte Case Management dar.

Case Management ist eine Behandlungsform, die sich durch ein hohes Maß an Interaktion der professionell beteiligten medizinischen Heilberufe wie Hausarzt, Facharzt, ggfs. Psychotherapeuten untereinander und mit dem Patienten auszeichnet. Koordiniert wird diese Form der Zusammenarbeit von einer klinisch geschulten Pflegekraft oder Psychologin, der Case ManagerIN. Diese entwickelt nach sorgfältiger Diagnostik gemeinsam mit Patient und Hausarzt einen Therapieplan, überprüft regelmäßig die Wirksamkeit der Behandlung und regt ggfs. Änderungen an. Eine wichtige Behandlungskomponente ist bei diesem Programm

---

<sup>7</sup> Dieses Kapitel ist in einigen Abschnitten ein indirektes Zitat aus einer eigenen, selbstständig erstellten Arbeit an einem Buchkapitel zur Versorgung depressiver Störungen im Alter (Heidenblut & Zank, 2009)

auch der stetige Kontakt zwischen der Koordinatorin und der oder dem Erkrankten, um neben dem psychoedukativen Austausch auch psychotherapeutische Elemente einer vertrauensvollen, kontinuierlichen Beziehung einzuschließen. Dieses Modell wurde in den USA durch das Programm IMPACT an einer Stichprobe von 1800 über 60-jährigen Probanden bereits erfolgreich evaluiert (Unnützer et al., 2002). In der Nachfolge wurde ein BMBF-Projekt (Gensichen et al., 2004) mit dem Ziel initiiert, ähnliche Effekte im deutschsprachigen Raum nachzuweisen.

## *2.2.8 Therapie depressiver Störungen im Alter*

### *2.2.8.1 Psychodynamische Therapie*

Die psychodynamische Therapie beruht im höheren Lebensalter in gleichem Maß wie bei anderen Altersgruppen auf einer tragfähigen therapeutischen Beziehung. Aufgrund der veränderten Altersverhältnisse zwischen Therapeut und Patient, entsteht in dieser Konstellation zusätzlich zu den allgemeinen Übertragungsphänomenen die sogenannte Umkehrübertragung. Diese beruht darauf, dass der jüngere Therapeut für den Patienten symbolisch die Rolle eines realen oder imaginierten Kindes annimmt, bzw., dass der Therapeut sich ebenfalls in diese Rolle hineinversetzt. Da das klassische analytische Szenario jedoch in der Regel mit einer Elternübertragung arbeitet, sind Patient und Therapeut bei der dynamischen psychotherapeutischen Behandlung Älterer vor besondere Herausforderungen gestellt. Peters (Zank, Peters & Wilz, 2010) fasst die Wirkweise der psychodynamischen Therapie im Alter wie folgt zusammen:

Die psychodynamische Therapie zielt auf ein besseres Verstehen und damit Verarbeiten unbewusster Zusammenhänge und baut dabei auf Erfahrungen im Rahmen einer tragfähigen therapeutischen Beziehung auf. Diese zu entwickeln erfordert jedoch in der Regel bei der Behandlung Älterer durch jüngere Therapeuten mehr Aufmerksamkeit. Sowohl auf Seiten der Patienten wie auf Seiten der Therapeuten sind zahlreiche Widerstände zu erwarten, die es zu reflektieren gilt, wobei die Altersdifferenz von ausschlaggebender Bedeutung ist. Besonders zu beachten ist das Konzept der umgekehrten Übertragung, bei der der

Therapeut vom Patienten in einer Sohn- oder Tochterposition erlebt wird. Aber auch narzisstische Aspekte sind in der Übertragungsbeziehung von Bedeutung. Die zentrale psychodynamische Interventionsmethode stellt die Deutung dar, die sich sowohl auf die Übertragungsbeziehung wie auch auf die Abwehr bzw. auf Widerstand oder lebensgeschichtliche Zusammenhänge beziehen kann. Gerade Letztere gewinnen in der Arbeit mit Älteren aufgrund der Aufgabe der Lebensbilanzierung eine besondere Bedeutung. Das Interventionsspektrum hat sich jedoch im Laufe der Zeit erweitert und umfasst heute ein breites Methodenarsenal, wodurch es dem Therapeuten erleichtert ist, sich flexibel auf die Möglichkeiten eines Patienten einzulassen (Zank, Peters & Wilz, 2010, S. 167-168).

#### *2.2.8.2 Kognitiv -behaviorale Therapie*

Die Kognitive Verhaltenstherapie oder kognitiv-behaviorale Therapie (KVT) hat das Ziel, eine Veränderung der depressiven Symptomatik durch einen Aufbau von selbstwertförderlichen Aktivitäten sowie über ein Bemerken und Verändern dysfunktionaler Gedankenmuster auf Seiten des Patienten zu erreichen.

*Tabelle 7: Beispiele für dysfunktionale Denkmuster*

Dysfunktionales Denkmuster	Beispiel
Absolute Forderungen	Die Ärzte müssen uns besser behandeln.
Globale negative Selbstbeurteilungen	Wenn ich die Betreuung nicht alleine schaffe, bin ich eine schlechte Ehefrau/Tochter.
Katastrophendenken	Es wäre absolut schrecklich, wenn ...
Niedrige Frustrationstoleranz	Ich könnte es nicht aushalten, wenn ...
Alles-oder-Nichts-Denken	Wenn ich versage, ist das ein totaler Fehler...
Fokus auf negative Aspekte	Ich habe immer nur Pech.
Ignorieren positiver Aspekte	Ich kann an ... gar nichts Positives sehen.
Moralisch-absolutistisches Denken	Man darf nicht krank werden.
Willkürliche Schlussfolgerungen	Es gibt einfach niemanden, der mir in meiner Situation helfen kann.
Unsinnige Prognosen	Das habe ich schon mal ausprobiert, das wird nie klappen.
Suche nach Anerkennung	Es ist schrecklich, wenn andere über mich und meinen Mann schlecht reden oder Denken.

Anmerkung. Tabelle zitiert nach Zank, Wilz & Peters (2010, S. 196).

Das Verfahren, das strukturierter und direkter vorgeht als psychodynamische Therapien, orientiert sich an den Grundprinzipien „Problemorientierung, Zielorientierung, Handlungsorientierung, Transparenz, Hilfe zur Selbsthilfe, Überschreiten des therapeutischen Settings und ... empirisch fundierte Weiterentwicklung“, (Margraf, 2003, S. 4). Während der Aufbau angenehmer Aktivitäten direkt durch den jeweiligen Therapeuten angeregt und in der Therapie immer wieder evaluiert wird, erfordert die Identifikation negativer Gedanken zunächst eine Erkennensleistung auf Seiten des Patienten, die durch hilfreiche Fragen des Therapeuten angeleitet werden kann (Siehe Tabelle 7).

*Tabelle 7: Fragetechniken zur Identifikation und Modifikation dysfunktionaler Gedanken*

Technik zur Überprüfung dysfunktionaler Gedanken	Beispielfragen
Realitätstestung	<ul style="list-style-type: none"> <li>• Ist es wirklich so?</li> <li>• Erinnern Sie sich an Erfahrungen, die zeigten, dass dieser Gedanke nicht immer total zutraf?</li> <li>• Welche Beweise/Tatsachen sprechen für Ihre Sichtweise?</li> <li>• Welche anderen Möglichkeiten gibt es, die Situation zu erklären?</li> <li>• Das sind die negativen Seiten der Situation? Gibt es auch positive Seiten?</li> <li>• Verallgemeinern Sie zu stark?</li> <li>• Übernehmen Sie in der Bewertung der Situation für Dinge Verantwortung, auf die Sie gar keinen Einfluss haben?</li> <li>• Gibt es in der Situation vielleicht kleine, leicht zu übersehende Einzelheiten, die gegen Ihren Gedanken sprechen oder zu einer anderen Bewertung führen würden?</li> </ul>
Hedonistisches Kalkül	<ul style="list-style-type: none"> <li>• Wie fühlen Sie sich, wenn Sie diesen Gedanken haben?</li> <li>• Hilft der Gedanke Ihnen, sich so zu fühlen, wie Sie möchten?</li> <li>• Hilft der Gedanke Ihnen, die Situation gut zu bewältigen?</li> </ul>
Distanzierung durch Rollentausch	<ul style="list-style-type: none"> <li>• Was würden Sie einem Freund, den Sie mögen, sagen, wenn er diesen Gedanken hätte?</li> <li>• Was würde Ihnen ein guter Freund in dieser Situation sagen, wenn er von diesem Gedanken wüsste und selbst nicht vollständig von seiner Richtigkeit überzeugt wäre?</li> <li>• Kennen Sie jemanden, der mit dieser Situation leichter fertig wird als Sie? Was sagt diese Person möglicherweise zu sich selbst?</li> </ul>
Temporale Relativierung	<ul style="list-style-type: none"> <li>• In einer anderen Gefühlslage, wenn es Ihnen z.B. gut geht, denken Sie dann über diese Situation anders? Wie?</li> <li>• Wie werden Sie später, in einem Monat oder einem Jahr, darüber denken?</li> <li>• Stellen Sie sich vor, Sie seien zehn Jahre weiter. Wie werden Sie rückblickend die jetzige Situation betrachten? Stehen aus dieser Perspektive dann andere Aspekte im Mittelpunkt?</li> </ul>
Entkatastrophisieren	<ul style="list-style-type: none"> <li>• Was würde schlimmstenfalls geschehen? Wie schlimm wäre das wirklich? Wie wahrscheinlich ist das?</li> <li>• Was wäre schlimmer als diese Situation?</li> <li>• Wie wichtig ist diese Sache wirklich für Sie?</li> </ul>
Fokussieren auf Ressourcen und Kompetenzen	<ul style="list-style-type: none"> <li>• Haben Sie schon einmal eine ähnlich schwierige Situation geschafft? Was hat Ihnen geholfen?</li> <li>• Gibt es etwas, das Ihnen sehr wichtig ist, an das Sie sich in dieser Situation erinnern könnten und das Ihnen Mut und Sicherheit geben kann?</li> <li>• Worauf können Sie vertrauen?</li> </ul>
Sinnorientierung	<ul style="list-style-type: none"> <li>• Was können Sie in dieser Situation lernen?</li> <li>• Welchen Sinn finden Sie in dieser Situation?</li> </ul>

Anmerkung. Tabelle zitiert nach Zank, Wilz & Peters, 2010, Sokratischer Dialog nach Kaluza (2004).

Dabei können neben individuell erworbenen negativen Bewertungen des einzelnen Patienten auch kohortenspezifische dysfunktionale Denkmuster eine Rolle spielen, die vom Therapeuten jeweils frühzeitig erkannt und berücksichtigt werden sollten. Für die Nachkriegsgeneration könnten das etwa eine Aussage sein wie „außerhalb der Familie (sollten) keine persönlichen Probleme besprochen werden“ (Laidlaw, Thompson, Dich-Siskin & Gallager-Thomson, 2003, zitiert nach Zank, Wilz & Peters, 2010, S. 199). Generell gilt für die KVT im höheren Lebensalter wie für alle Therapie-Techniken mit älteren Patienten, dass einerseits altersspezifische Themen wie Entwicklungsverluste nicht tabuisiert werden, jedoch auch nicht im Sinne eines negativen Altersstereotyps als selbstverständlich hingenommen werden sollten. Dabei bietet die KVT durch ihre differenzierte Verhaltensdiagnostik jedes individuellen Patienten am Anfang des therapeutischen Prozesses die Möglichkeit, den jeweiligen Bedürfnissen einzelner Klienten aus dieser sehr heterogenen Zielgruppe gerecht zu werden. Der Ansatz der KVT wurde in Deutschland durch Hautzinger (Hautzinger & Welz, 2004) im Rahmen des Gruppentherapiekonzeptes DiA (Depression im Alter) für ältere Patienten adaptiert, deren Wirksamkeit bereits nachgewiesen werden konnte.

### *2.2.9 Zusammenfassung*

Als depressive Störungen im Alter werden Erkrankungen bezeichnet, die nach Vollendung des 60. Lebensjahres auftreten. Dabei zeigt sich im höheren Lebensalter eine größere Variabilität der Symptome, die Erkrankungsformen werden jedoch im ICD-10 nach den gleichen Kriterien kodiert, die auch für jüngere Patienten gelten.

Depressionen scheinen nach dem 60. Lebensjahr nicht häufiger aufzutreten als in anderen Lebensabschnitten. Die Berliner Altersstudie gibt Prävalenzraten von insgesamt 9.1% für krankheitswertige depressive Störungen in der älteren Bevölkerung an. Dabei leidet ein weitaus größerer Anteil der Altenpopulation, insgesamt 22% an sub-diagnostischer Depressivität, also depressiven Symptomen die weniger zahlreich und ausgeprägt sind als bei

klinisch relevanten Störungen, die jedoch die Lebensqualität der Betroffenen beeinträchtigen und das Risiko für spätere Erkrankungen erhöhen.

Es gibt keine einheitlichen Forschungsbefunde zu der Frage, ob bei Depressionen im Alter grundsätzlich von schlechteren Prognosen auszugehen ist als in früheren Lebensabschnitten; so scheinen im Alter auch die Krankheitsverläufe durch ein hohes Maß an Variabilität gekennzeichnet zu sein und erheblich von den körperlichen, sozialen, kognitiven und materiellen Ressourcen des Betroffenen abzuhängen. Risikofaktoren, die bei der Entwicklung depressiver Störungen im Alter eine besondere Rolle spielen, sind körperliche Erkrankung, kognitive Beeinträchtigungen, pathologische zerebrale Veränderungen, Pflegebedürftigkeit und Multimedikation.

Im höheren Lebensalter kann eine depressive Verstimmung durch Umstände ausgelöst werden, die für diese Lebensphase besonders spezifisch sind, beispielsweise durch die Erfahrung nachlassender körperlicher und geistiger Kraft, die Erfahrung des Verlustes sozialer Rollen oder durch Tod oder schwere Erkrankungen naher Bezugspersonen.

Auch ältere depressive Patienten scheinen im Allgemeinen gut auf medikamentöse sowie auf psychologische Intervention anzusprechen. Die Versorgungslage älterer depressiver Patienten stellt allerdings nach wie vor eine Herausforderung für unser Gesundheitssystem dar. Besonders vielversprechende Therapieansätze für ältere Patienten sind die psychodynamische Therapie und die kognitive Verhaltenstherapie. Beide Ansätze wurden unter Berücksichtigung gerontologischer Erklärungsmodelle wie dem SOK-Modell oder dem Entwicklungsmodell nach Erikson für Depressionen im Alter im Hinblick auf eine ältere Zielgruppe angepasst.

## 2.3 Die Diagnostik depressiver Störungen

### 2.3.1 Diagnostik in der klinischen Psychologie

Im Rahmen der klinischen Psychologie bezeichnet Diagnostik „eine Gesamtheit der Aktivitäten, die darauf abzielen, eine (psychologische oder psychiatrische) Aussage über einen Patienten zu formulieren“ (Bouman, 1993, S.45). Mit dem Erstellen einer Diagnose trifft der Kliniker eine Entscheidung über den Krankheitswert bzw. die Ursachen einer Symptomatik. Die Diagnose einer bestimmten depressiven Störung schließt dabei auf der einen Seite eine Prognose über den weiteren Erkrankungsverlauf ein, auf der anderen Seite wird damit eine Aussage über den Behandlungsbedarf bzw. die Behandlungsmöglichkeiten der Störung getroffen. Nach Fydrich (1997) bestehen die wichtigsten Aufgaben der klinisch-psychologischen Diagnostik in der qualitativen und quantitativen Beschreibung einer psychischen Störung, ihrer Klassifikation sowie in der Beobachtung von Krankheitsverlauf, Intervention und Veränderung der Symptomatik. Für die Diagnostik depressiver Störungsbilder ergeben sich somit für den Diagnostiker folgende Aufgaben: Er muss die Symptomatik klassifizieren, d.h. er muss dem Krankheitsbild und seinen Entstehungsbedingungen gegebenenfalls eine der in Abschnitt 2.1.3 beschriebenen ICD-10-Diagnosen zuordnen, er muss den Schweregrad der Störung bestimmen, einen der Erkrankung angemessenen Therapievorschlagn machen und er muss überprüfen, ob sich die Symptomatik unter einer bestimmten Intervention, z.B. einer antidepressiven Medikation, verändert.

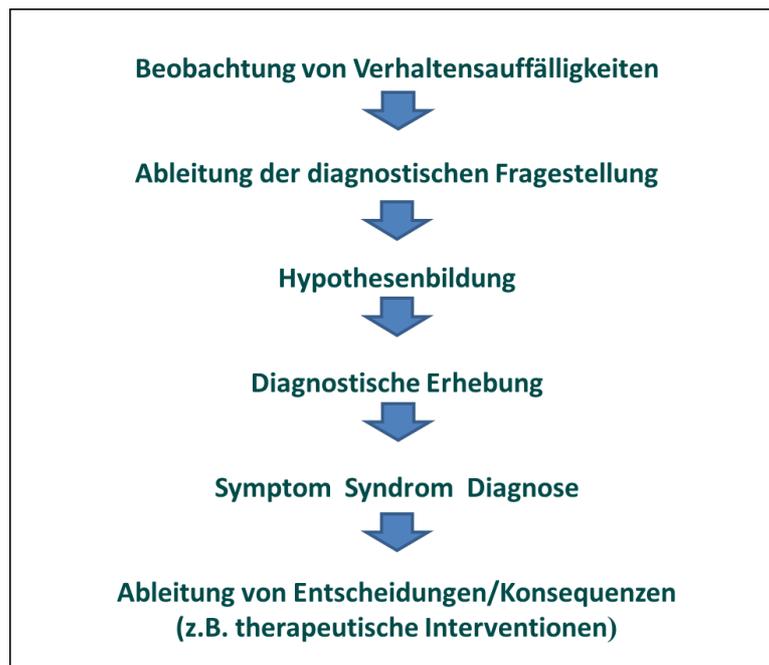
Die Vergabe der korrekten Diagnose stellt somit die Voraussetzung für eine angemessene Behandlung des Erkrankten dar und ist daher für die klinische Praxis von erheblicher Relevanz. Aus diesem Grund bleibt die Durchführung der Diagnostik nicht allein den subjektiven Entscheidungen des jeweiligen Klinikern überlassen, sondern sie ist an „vielfältige Regeln und Vorschriften“ gebunden (Fisseni, 1997, S.4, S.8-9), die Objektivität, Zuverlässigkeit und Gültigkeit diagnostischer Urteilsbildungen gewährleisten sollen. Diese

Regeln und Vorschriften stammen aus Erkenntnissen der wissenschaftlichen Psychologie, insbesondere der Methodenlehre und leiten die sachgemäße Durchführung des sogenannten „diagnostischen Prozesses“ an.

### 2.3.2 *Der diagnostische Prozess*

Der diagnostische Prozess ist „ die Abfolge von Maßnahmen zur Gewinnung diagnostisch relevanter Informationen“ (Amelang & Zielinski, 1997, S.342). Dabei existieren zahlreiche messtheoretische Forderungen bzw. praktische Vorschläge hinsichtlich des genauen Ablaufs einer diagnostischen Untersuchung, die sich zum Teil erheblich widersprechen (Amelang & Zielinski, 1997, Kap.5). Eine vereinfachte Darstellung des Prozesses bezogen auf eine klinische Fragestellung (Stieglitz, 2000) wird in Abbildung 6 dargestellt.

Die Fragestellung des diagnostischen Prozesses ergibt sich im Bereich der klinischen Psychologie in der Regel dadurch, dass bei einem Patienten bestimmte Verhaltensweisen beobachtet werden, die in irgendeiner Form „auffällig“ erscheinen, beispielsweise eine anhaltende Niedergeschlagenheit oder deutlicher Gewichtsverlust. Aufgabe des Diagnostikers ist, zu entscheiden ob diesen Verhaltensauffälligkeiten möglicherweise eine bestimmte psychische Störung, etwa eine depressive Verstimmung, zugrunde liegt und gegebenenfalls eine entsprechende Hypothese zu formulieren.

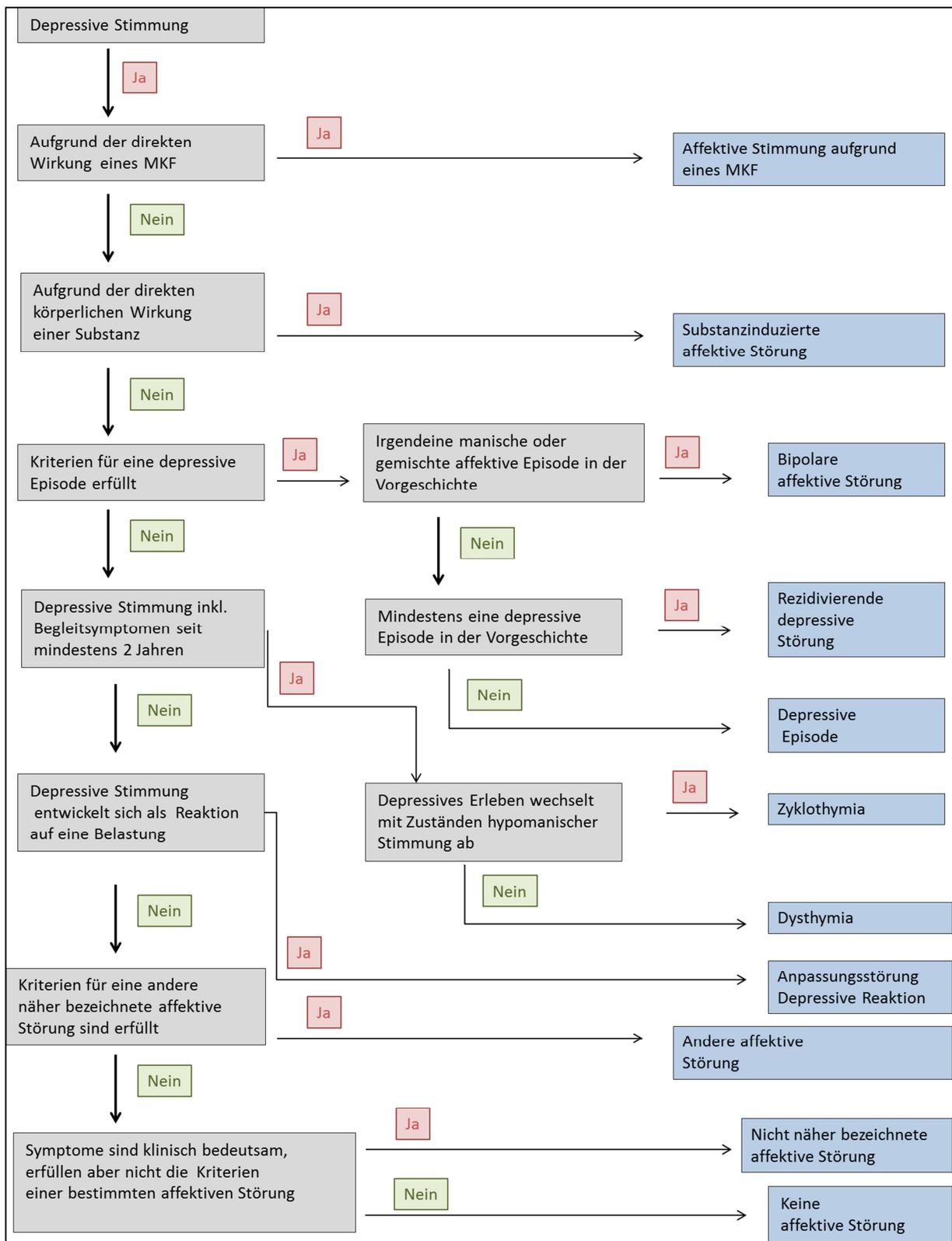


Anmerkungen. Vgl. Stieglitz (2000)

*Abbildung 6: Vereinfachte Darstellung des diagnostischen Prozesses*

Zur Prüfung der Hypothese „Verdacht auf Depression“ müssen dann im Rahmen der diagnostischen Erhebung Informationen über den Patienten gewonnen werden, die auf der einen Seite möglichst objektiv und auf der anderen Seite für die Fragestellung relevant sind. Für die Diagnostik depressiver Störungen beziehen sich relevante Informationen auf das Vorliegen bzw. die Ausprägung der in der ICD-10 benannten Haupt- und Nebensymptome, auf den bisherigen Verlauf der Störung bzw. auf die psychiatrische Vorgeschichte des Patienten sowie auf eventuelle Ursachen für das Zustandekommen der Erkrankung. Differentialdiagnostisch spielt darüber hinaus der Ausschluss von durch „körperliche Prozesse (MKF) provozierten bzw. durch Substanzen induzierten“ depressiven Erscheinungsbildern eine wichtige Rolle (Hautzinger & Meyer, 2002, S.35). Hat der Diagnostiker ein ausreichendes Maß solcher Informationen erhoben, kann er sie zu einer Diagnose kombinieren, die das Krankheitsbild des Patienten möglichst genau abbildet und gegebenenfalls einen Therapievorschlagn machen. Hautzinger und Meyer (2002) stellen den diagnostischen Weg von Verhaltensbeobachtungen hin zu einer differenzierten Diagnose

(nach DSM IV) durch einen Entscheidungsbaum dar, der in Abbildung 1 auf die ICD-10 Klassifizierung übertragen dargestellt wird.



Anmerkungen. Vgl. Hautzinger & Meyer (2002).

Abbildung 7: Entscheidungsbaum zur Differentialdiagnostik affektiver Störungen

### *2.3.3 Standardisierte Testverfahren in der Depressionsdiagnostik*

Die Erhebung möglichst objektiver und für die Fragestellung relevanter Informationen ist für die korrekte diagnostische Urteilsbildung von maßgeblicher Bedeutung. Um den Kliniker bei diesem Vorgang nicht seinen subjektiven und möglicherweise fehlerhaften Entscheidungen zu überlassen, wurden in der Psychologie zahlreiche standardisierte Testverfahren entwickelt, deren Anwendung in der modernen Diagnostik mehr oder weniger obligatorisch ist

(De Bruyn, 2003). Der Vorteil dieser Verfahren besteht darin, dass sie die Erhebung von Informationen gewährleisten, die nach dem aktuellen Wissensstand am besten zur Beantwortung der Fragestellung beitragen. Darüber hinaus ermöglichen sie, die Verhaltensmerkmale oder Symptome eines Patienten durch einen bestimmten Testwert oder Score abzubilden, der mit einer an der Mehrheit orientierten Norm verglichen werden kann (Pawlik, 2003). Die Ausprägung des Merkmals Depressivität kann so bei einer bestimmten Person gemessen und mit einer Durchschnittsgröße verglichen werden.

### *2.3.4 Depressivität als Messgröße*

In den gebräuchlichen Messverfahren wird Depressivität entlang eines Kontinuums konzipiert, d.h. als ein Merkmal verstanden, das alle Menschen mehr oder weniger ausgeprägt aufweisen. Wird mit einem solchen dimensional Messinstrument der Testwert einer bestimmten Person ermittelt, so beinhaltet das keine Entscheidung darüber, ob diese Person unter einer krankheitswertigen depressiven Störung leidet oder nicht (Bouman,1993).

Ziel klinischer Diagnostik ist es jedoch, diese Frage zu beantworten. So ist die Klassifizierung eines Störungsbildes als kategoriale Diagnostik zu verstehen, da „bei Vorliegen gewisser Symptome sowie einer bestimmten Anzahl von Symptomen über eine festgelegte Zeit und bei gleichzeitigem Ausschluss definierter Krankheiten, eine Diagnose vergeben (wird) oder eben nicht (Ja/Nein-Entscheidung)“ (Hautzinger & Meyer, 2002, S.36).

Um die Diskrepanz zwischen dimensional Messinstrumenten und kategorialen

diagnostischen Urteilen in der Praxis zu überwinden, werden für Depressionsskalen in der Regel sogenannte Cutoff-Scores festgelegt: d.h., bestimmte Messwerte, die auf der Grundlage empirischer Analysen die Grenze zwischen unbedenklichen und als klinisch auffällig geltenden Testergebnissen markieren (Bouman,1993). Überschreitet ein Messwert einen solchen Cutoff-Score, ist eine erhöhte Wahrscheinlichkeit für eine depressive Störung und somit die Indikation für weitere Diagnostik gegeben.

### *2.3.5 Eigenschaften verschiedener psychometrischer Testverfahren*

Es gibt zahlreiche Methoden, um Depressivität psychometrisch abzubilden, darunter strukturierte bzw. standardisierte Interviews, psychologische Tests, Fragebögen, Beobachtungsprotokolle, Diagnosechecklisten und vieles mehr (Hautzinger & Meyer, 2002). Alle diese Verfahren formen die Erlebens- und Verhaltensausprägung eines Patienten in einen numerischen Testwert um - sie quantifizieren sie. Dabei gibt es grundsätzlich zwei Möglichkeiten, die Verhaltensmerkmale einer Person in Erfahrung zu bringen. Einerseits kann der Patient sein Erleben und Verhalten selbst beurteilen und darüber Auskunft geben; andererseits kann es durch eine andere Person, in der Regel durch den Diagnostiker, beobachtet und eingeschätzt werden (Stieglitz, 1994). Diese beiden Datenquellen gehen abhängig von dem verwendeten Messinstrument in unterschiedlicher Gewichtung in den Gesamtscore ein. So beruhen standardisierte Fragebögen in der Regel ausschließlich auf der Selbstbeurteilung des Patienten, während Interviews sich sowohl auf Selbst- als auch auf Fremdbeurteilung beziehen. Im Folgenden wird lediglich der Aufbau dieser beiden Verfahrensmöglichkeiten genauer dargestellt, da sie in der Diagnostik depressiver Störungen den größten Stellenwert einnehmen und für die Fragestellung dieser Arbeit relevant sind.

#### *2.3.5.1 Standardisierte Fragebögen*

Standardisierte Fragebögen sind die am häufigsten verwendete Methode zur Diagnostik depressiver Störungen. Sie haben den Vorteil, dass sie leicht vorzugeben und auszuwerten

sind, in der Durchführung wenig Zeit erfordern und für sie in der Regel gut erforschte Normwerte existieren, die die Interpretation der Testwerte erleichtern (Stieglitz, 1994). Charakteristisch für den Aufbau standardisierter Fragebögen ist, dass sie ein stark strukturiertes Item- Responseschema vorgeben. Die Items bestehen aus fest formulierten Fragen, die so gestellt sind, dass sie nur Ja/Nein- Antworten oder Antwortabstufungen wie etwa „selten“, „manchmal“ oder „oft“ zulassen. Der Beantwortung einer Einzelfrage kann daraufhin ein bestimmter Wert zugeordnet werden, der auch als Item-Score bezeichnet wird. Addiert man die einzelnen Itemscores, die sich aus dem Antwortverhalten einer Person ergeben, so erhält man für diese Person einen Gesamtscore, der ihren „Depressivitätsmesswert“ darstellt und mit dem für das Verfahren festgelegten Cutoff-Score verglichen werden kann. Durch ihre stark strukturierte Form sind standardisierte Fragebögen anfällig für verschiedene Fehlerquellen (Stieglitz, 1994). So kann ein Patient etwa unklar formulierte Fragen missverstehen, er kann bedingt durch das Antwortschema unabsichtlich Fehler in seiner Selbsteinschätzung begehen, oder er kann das Testergebnis absichtlich durch unwahre Angaben verfälschen (siehe Tabelle 8).

#### *2.3.5.2 Halbstrukturierte Interviews*

Ein halbstrukturiertes Interview besteht aus einem Gespräch zwischen dem Diagnostiker und dem Patienten, wobei die Inhalte durch bestimmte Leitthemen, etwa bestimmte depressive Symptome vorgegeben sind, die Formulierung der einzelnen Fragen jedoch dem Untersucher überlassen bleibt. Bei dieser Methode geht in den Testwert sowohl die Selbstbeurteilung des Patienten ein, der innerhalb des Gespräches Angaben über seine Stimmung macht, als auch die Fremdeinschätzung des Diagnostikers, der die Antworten und das Verhalten des Patienten beurteilt, auf einer quantitativen Skala kodiert und zu einem Gesamtmesswert addiert. Halbstandardisierte Interviews haben den Vorteil, dass Missverständnisse bei der Formulierung von Fragen oder bestimmte Antworttendenzen durch den Diagnostiker eher aufgeklärt und vermieden werden können, als bei Fragebögen; so kann der Untersucher

gegebenenfalls unklare Fragen genauer erklären. In der Auswertung kann er darüber hinaus fehlerhafte Antworttendenzen des Patienten identifizieren und in seiner Kodierung berücksichtigen. Durch den größeren Spielraum stellt dieses Instrument an den Kliniker jedoch höhere Anforderungen und er kann selbst zu einer möglichen Fehlerquelle werden. Stieglitz (1994) fasst die verschiedenen Fehlerquellen von Selbst- und Fremdbeurteilungsverfahren in einer tabellarischen Gegenüberstellung zusammen (siehe Tabelle 8).

*Tabelle 8: Fehlerquellen von Selbst- und Fremdbeurteilungsverfahren*

Fehlerquellen	S	F
1. Fehlerquellen bedingt durch Testkonstruktion		
• z.B. unklare Formulierungen	X	(X)
• Interpretationsspielraum	X	(X)
2. Unwissende Fehler bedingt durch Erinnerungs-, Selbstbeobachtungs- oder Selbstdarstellungsfehler		
• z.B. Selbsttäuschung	X	
• Erinnerungs- und Gedächtnisfehler	X	
3. Absichtliche Verfälschungen		
• z.B. Simulation	X	
• Bagatellisierung	X	
4. Response Sets		
• Z.B. soziale Erwünschtheit	X	
• Tendenz zu alternativen oder extremen Antworten	X	X
• Tendenz zur Mitte	X	X
• Seriale Positionseffekte	X	X
5. Fehler durch falsche Schlussfolgerungen		
• Z.B. logische Fehler		X
• Halo-Effekt		X
• Kontrastfehler		X

*Anmerkungen.* S = Selbstbeurteilungsverfahren; F = Fremdbeurteilungsverfahren. X = die jeweilige Fehlerquelle (Zeile) ist für das Verfahren (Spalte) relevant. Tabelle zitiert nach Stieglitz (1994).

Da vor allem bei Selbstbeurteilungsverfahren Fehlerquellen durch eine sorgfältige Testkonstruktion erheblich reduziert werden können, hat die Evaluation von Testverfahren nach bestimmten Gütekriterien einen hohen Stellenwert in der psychologischen Diagnostik.

### 2.3.6 Gütekriterien psychometrischer Testverfahren

Es gibt vielfältige Kriterien, nach denen die Qualität von Messinstrumenten beurteilt werden kann. Diese sind im Hinblick auf besondere Umstände, Bedingungen und Zielsetzungen eines Verfahrens in unterschiedlichem Maß verbindlich. So wurde durch das Testkuratorium

Deutscher Psychologieverbände ein „breites Spektrum von Bewertungsmaßstäben erarbeitet, (deren) Anforderungen. kaum jemals von einem einzigen Verfahren in gleichem Ausmaße erfüllt werden (können), zumal sie nicht alle miteinander vereinbar sind.“ (Amelang & Zielinski, 2002, S.144).

*Tabelle 9 : Kriterien zur Testbeurteilung des Testkuratoriums deutscher Psychologerverbände*

<b>Beurteilungsbereich</b>	<b>Beurteilungskriterien</b>
<b>1. Testgrundlage</b>	1.1 Diagnostische Zielsetzung 1.2 Theoretische Grundlagen 1.3 Nachvollziehbarkeit der Testkonstruktion
<b>2. Testdurchführung</b>	2.1 Durchführungsobjektivität 2.2 Transparenz 2.3 Zumutbarkeit 2.4 Verfälschbarkeit 2.5 Störanfälligkeit
<b>3. Testverwertung</b>	3.1 Auswertungsobjektivität 3.2 Zuverlässigkeit 3.3 Gültigkeit 3.4 Normierung 3.5 Bandbreite 3.6 Informationsausschöpfung 3.7 Änderungssensitivität
<b>4. Testevaluation</b>	4.1 Ökonomie 4.2 Fairness 4.3 Akzeptanz 4.4 Vergleichbarkeit 4.5 Bewährung
<b>5. Äußere Testgestaltung</b>	

Anmerkungen. Tabelle aus Amelang et al., 1986, zitiert nach Stieglitz (2000)

Die in Tabelle 9 beschriebenen Kriterien sind hierarchisch nicht auf einer Ebene angesiedelt. So dürfen einige Richtlinien, die sich eher auf den Kontext eines Verfahrens beziehen, wie etwa die Forderung nach einer breiten Anwendbarkeit, in der Praxis mitunter vernachlässigt werden. Die sogenannten Hauptgütekriterien Objektivität, Reliabilität und Validität beziehen sich dagegen unmittelbar auf die technischen Eigenschaften des Messinstrumentes und sind

daher von „zentraler Bedeutung für die Qualität von Tests“ (Amelang & Zielinski, 2002, S.147). Im folgenden Abschnitt werden die Hauptgütekriterien nach der Klassischen Testtheorie (KTT) sowie einige Nebengütekriterien dargestellt, die im Hinblick auf die Fragestellung dieser Arbeit relevant sind.

#### *2.3.6.1 Objektivität*

„Objektivität steht für das Ausmaß, in dem die Ergebnisse eines Tests unabhängig von der Person des Untersuchungsleiters sind“ (Amelang & Zielinski, 2002, S.148). Sie betrifft die Durchführung eines Messverfahrens sowie die Auswertung und Interpretation der erhobenen Informationen. Durchführungsobjektivität kann in der diagnostischen Praxis durch eine möglichst hohe Neutralität im Verhalten verschiedener Tester gewährleistet werden. Bezogen auf die Diagnostik depressiver Störungen kann dies etwa bedeuten, dass der Untersucher die Stimmung des Patienten während der Testung so wenig wie möglich durch Aufmunterungsversuche oder Ähnliches beeinflussen sollte. Durchführungsobjektivität lässt sich quantitativ kaum bestimmen. Sie gilt jedoch nach rationalen Überlegungen als in hohem Maß gegeben, wenn alle Bedingungen, die sich auf das Testverhalten auswirken können, konstant gehalten werden.

Auswertungsobjektivität bezieht sich auf die Zuordnung diagnostischer Informationen zu bestimmten Messwerten. Für Selbstbeurteilungsverfahren, bei denen die Kodierung des Messwertes nicht durch einen Untersucher vermittelt wird, kann sie in absolutem Umfang unterstellt werden. Bei Messinstrumenten, die sich auch auf Fremdbeurteilung beziehen, kann sie durch entsprechende Schulungen der Untersucher sowie durch eine genaue Definition der Messskalen optimiert werden. Mit Interpretationsobjektivität ist das Maß an Übereinstimmung gemeint, mit der verschiedene Untersucher einen bestimmten Messwert interpretieren; ob zum Beispiel der gleiche Score auf einer bestimmten Skala zu denselben diagnostischen Schlüssen führt. Depressionsskalen, für die in der Regel Cutoff-Scores oder

Normwertetabellen existieren, gewährleisten per se einen hohen Grad an Interpretationsobjektivität.

### *2.3.6.2 Reliabilität*

“Die Reliabilität beschreibt die Genauigkeit, mit der ein Test eine Merkmalsdimension erfasst, und zwar unter der Vernachlässigung des Umstandes, ob es sich dabei auch um die Merkmalsdimension handelt, deren Erfassung intendiert ist“ (Amelang & Zielinski, 2002, S.152). Gefragt ist also die Zuverlässigkeit einer Messung unabhängig davon, ob auch das „Richtige“ gemessen wird. Dabei gibt es verschiedene Möglichkeiten, ein mathematisches Maß der Reliabilität zu ermitteln, die je nach Eigenschaft des zu messenden Merkmals angewendet werden können. Die einfachste Methode, mit der die sogenannte Re-Test-Reliabilität ermittelt wird, überprüft, ob eine Person bei wiederholtem Testen mit derselben Skala vergleichbare Werte erzielt, oder ob die Ergebnisse schwanken. Diese Methode kann jedoch nur auf Tests angewendet werden, die vergleichsweise konstante Eigenschaften messen, da ansonsten nicht zwischen wahren Wert und Messfehler unterschieden werden kann. Die Zuverlässigkeit von Verfahren, die sogenannte State-Merkmale erfassen, d.h. Zustände, die sich über die Zeit oder nach einer bestimmten Behandlung verändern, wird in der Regel auf andere Weise ermittelt. So besteht z.B. die Möglichkeit, einen Test in zwei gleichgroße Teile zu teilen, deren Werte man in einer Analyse miteinander vergleicht. Ist die Übereinstimmung der beiden Testteile hoch, so erzielt das Messinstrument ein zufriedenstellendes Maß an „Split-half-Reliabilität“. Die Split-half-Reliabilität bezieht sich also auf die Frage, inwiefern durch einzelne Teile eines Testverfahrens etwas Vergleichbares abgefragt wird. Sie ist somit die Verfahrensweise, die am ehesten die tatsächliche Messgenauigkeit eines Tests ermittelt. In der Praxis kann sie jedoch oft nicht angewendet werden, da sie für die Durchführung eine gerade Anzahl homogener Items erfordert, eine Voraussetzung, die viele kürzere Testverfahren nicht erfüllen. In solchen Fällen wird häufig die interne Konsistenz einer Skala als eine Sonderform der Reliabilität ermittelt. Zu diesem

Zweck wird der Test in so viele Hälften zerlegt, wie Items vorhanden sind. Daraufhin werden alle Iteminterkorrelationen ermittelt, der Gesamtlänge der Skala angepasst und zu einem Koeffizienten zusammengefasst.

Zur Ermittlung der Reliabilität einer Depressionsskala muss in der Praxis überprüft werden, ob die entsprechenden Bedingungen für das Verfahren im Einzelnen erfüllt sind. Da Depressivität beispielsweise ein State-Merkmal darstellt, ist die Bestimmung der Re-Test-Reliabilität hier nur bei Einhaltung strenger Zeitkriterien sinnvoll. Dagegen eignet sich die Bestimmung der internen Konsistenz nur bei Messinstrumenten, die unter rationalen Gesichtspunkten als eindimensional konstruiert sind (vgl. Abschnitt 2.31). In vielen Untersuchungen, die die Reliabilität von Messinstrumenten prüfen, werden mehrere Verfahren angewendet, um die Störanfälligkeit der einzelnen Methoden zu kontrollieren.

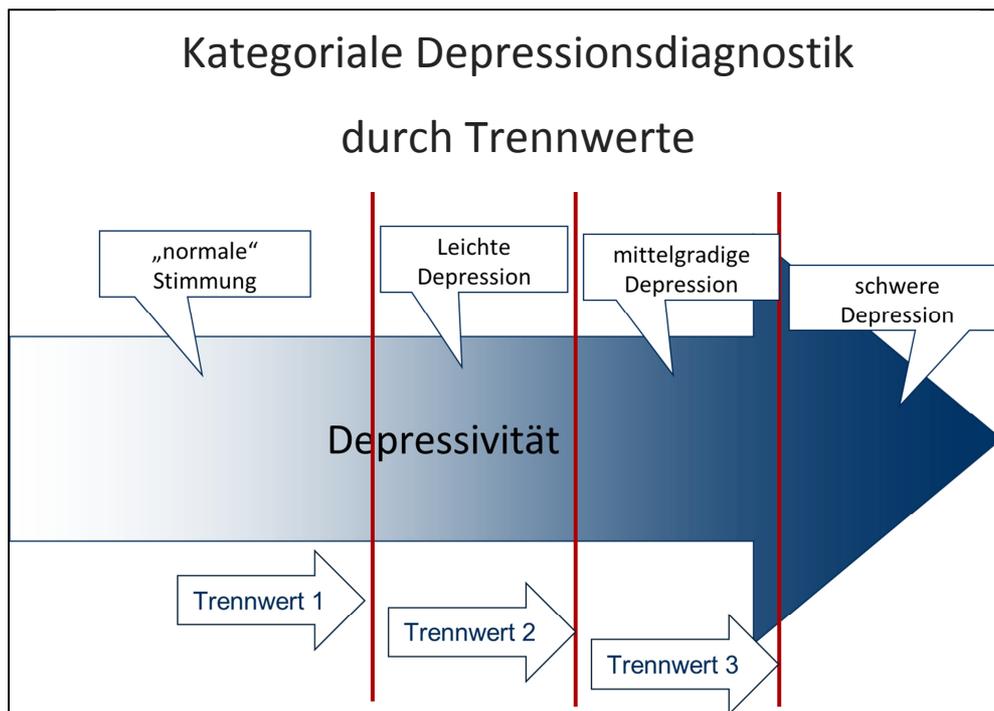
#### *2.3.6.3 Validität*

„Unter Validität wird das Maß an Genauigkeit verstanden, mit dem der Test dasjenige Persönlichkeits- oder Verhaltensmerkmal misst, das er messen soll oder zu erfassen vorgibt“ (Amelang & Zielinski, 2002, S.60). Gefragt ist also nach der Gültigkeit eines Testverfahrens, der Übereinstimmung zwischen dem, was gemessen wird und dem, was gemessen werden soll. Die Validität stellt das wichtigste aller Güterkriterien dar; so kann ein Test hohe Werte an Reliabilität und Objektivität erreichen und trotzdem als Messinstrument ungeeignet sein, wenn er eben nicht das Merkmal erfasst, für das er konzipiert wurde. Zur Ermittlung der Validität existieren, ähnlich wie bei der Reliabilität, zahlreiche Methoden; so unterscheidet man grundsätzlich zwischen Inhaltsvalidität, kriteriumsbezogener Validität und Konstruktvalidität. Im Folgenden wird in erster Linie auf die kriteriumsbezogene Validität näher eingegangen, da sie die Größe darstellt, auf die innerhalb der KTT fast ausschließlich Bezug genommen wird (Fisseni, 2004), und die daher auch die Grundlage für den empirischen Teil dieser Arbeit bildet. Das Prinzip zur Ermittlung der kriteriumsbezogenen Validität besteht darin, die Ergebnisse eines Testverfahrens mit einem sogenannten

Außenkriterium zu vergleichen, d.h. mit einer Variablen, die entweder dasselbe oder ein anderes Merkmal messen soll. Dabei gilt die Gültigkeit eines Verfahrens dann als gewährleistet, wenn es hoch mit Variablen korreliert, die das gleiche Merkmal messen (konvergente Validität) und niedrig mit solchen, deren Erfassung durch den Test nicht intendiert ist (diskriminative Validität) (Bouman, 1993). Da die so ermittelten Korrelationskoeffizienten abhängig von der jeweiligen, möglicherweise fehlerhaften Vergleichsgröße sind, existiert nicht „das“ Maß an Validität, sondern die Gültigkeit eines Tests kann immer nur relativ, bezogen auf ein bestimmtes Außenkriterium, gesehen werden. Die Wahl eines möglichst aussagekräftigen Kriteriums ist daher von erheblicher Bedeutung für die Beurteilung des Verfahrens.

Soll die Gültigkeit von Depressionsskalen bestimmt werden, werden als Vergleichskriterien häufig ICD-10- oder DSM-IV- Diagnosen verwendet, nach denen erfahrene Kliniker die Störungsbilder von Patienten klassifizieren. Solche auch als Gold-Standard bezeichneten Diagnosen gelten als „echte“ Außenkriterien, da sie in der Praxis die ausschlaggebende Informationsquelle für therapeutische Maßnahmen vorstellen (Amelang & Zielinski, 2002). Eine andere gängige Methode besteht darin, die Werte von Skalen mit denen von Testverfahren zu vergleichen, die sich in der Diagnostik depressiver Störungen bereits bewährt haben. Dabei stellen andere Messinstrumente Vergleichsvariablen dar, deren Qualität ihrerseits immer wieder geprüft werden muss; aus diesem Grund werden sie als Quasi-Kriterien bezeichnet.

*Effektivität.* Für Depressionsskalen, die mithilfe von Cutoff-Scores lediglich eine Aussage darüber treffen sollen, ob eine Symptomatik einen bestimmten Störungscharakter aufweist oder nicht, wird in der Regel die Effektivität des Messinstrumentes als eine Sonderform der Validität ermittelt.



*Abbildung 8: Die kategoriale Diagnostik depressiver Störungen durch Trennwerte*

Die Effektivität gibt dabei immer die Qualität eines Verfahrens bezogen auf einen bestimmten Trennwert an. In Abbildung 8 ist Depressivität als kontinuierliches Merkmal (blauer Pfeil) dargestellt. Dabei können in der Praxis jeweils unterschiedliche Trennwerte für unterschiedliche Fragestellungen festgelegt werden. So ist wäre für die Fragestellung „Liegt eine depressive Verstimmung vor?“ Trennwert 1 nützlich, während Trennwert 3 eher erforderlich wäre, um zu entscheiden, ob der Patient stationär behandelt werden sollte.

Das Vergleichskriterium für einen jeweiligen Trennwert bildet in allen Fällen eine Target-Variable mit lediglich zwei Ausprägungen, etwa der Ja/Nein-Entscheidung des Diagnostikers über den Krankheitswert einer depressiven Verstimmung. D.h., die Ergebnisse des Screenings und der Diagnose können nur zwei Ausprägungen annehmen: Der Patient ist depressiv oder der Patient ist nicht depressiv. Die Effektivität des Screenings wird dann definiert als Anzahl der Übereinstimmung der Testergebnisse mit den Ergebnissen der Diagnose. Abbildung 9 stellt den Vorgang des Messens analog zur Gewichtsmessung als Waage dar. Sind beide Waagschalen auf einer Ebene, so stimmt das Ergebnis des Screenings

mit dem der Diagnose überein und der Test ist für die jeweilige diagnostische Entscheidung valide oder gültig.



Abbildung 9: Testeffektivität als Vergleich zweier dichotomisierter Merkmale

Die durch den Vergleich mit dem klinischen Urteil ermittelte Effektivität des Messinstrumentes, depressive Personen von nicht depressiven Personen zu unterscheiden, besteht aus zwei Komponenten, die als Sensitivität und Spezifität eines Verfahrens bezeichnet werden:

„Unter Sensitivität versteht man das Ansprechen des Tests auf (eine Erkrankung), die später als solche diagnostiziert wird (Bortz & Lienert, 1998, S.223).“ Erzielen alle Personen, die nach klinischem Urteil als depressiv eingestuft werden, in einem Depressionstest einen entsprechenden Wert über dem Cutoff-Score, so ermittelt dieser Test keine falsch-negativen Ergebnisse und erreicht ein optimales Maß an Sensitivität. Spezifität bezeichnet dagegen „die Qualität eines Tests, bei negativem Ausfall auch eine negative Diagnose vorauszusagen“, (Bortz & Lienert, 1998, S.224). Damit ist die Fähigkeit des Verfahrens gemeint, Personen

nicht fälschlicherweise als depressiv einzustufen, wenn sie keine depressive Störung von Krankheitswert aufweisen - also keine falsch-positiven Testergebnisse zu produzieren.

#### *2.3.6.4 Normierung*

Betrachtet man ein Ergebnis, das durch einen objektiven, reliablen und validen Test ermittelt wurde isoliert, liefert es kaum hilfreiche diagnostische Informationen. Hat beispielsweise ein Patient in einer Summenskala zur Erfassung von Depressivität den Wert „15“ erzielt, so beantwortet dies für sich genommen nicht die Frage, ob dieser Proband als depressiv einzustufen sei oder nicht. Es fehlt ein Bezugsrahmen, der eine Aussage darüber macht, welches Testergebnis als normal gelten kann. Aus diesem Grund werden psychometrische Tests an sogenannten Eichstichproben *normiert*, d.h. sie werden einer großen Anzahl an Probanden vorgelegt, von denen angenommen wird, dass sie für den Anwendungsbereich des Tests repräsentativ sind. Zur Normierung eines Messinstrumentes, das Depressive von Nicht-Depressiven unterscheiden soll, empfiehlt es sich etwa, den Test einer Gruppe bereits als depressiv diagnostizierter Patienten vorzulegen, und einer Gruppe, die keine affektive Erkrankung aufweist. Analysiert man die so gewonnenen Testergebnisse, erhält man relative Häufigkeiten, die Rückschlüsse von einem Testergebnis auf das gemessene Merkmal erlauben. Wenn beispielsweise alle depressiven Probanden einer Normstichprobe einen Testwert über „11“ in einer Summenskala zur Erfassung von Depressivität erzielt haben, so stellt ein von einem einzelnen Patienten erzielter Wert „15“ ein Ergebnis dar, das mit hoher Wahrscheinlichkeit auf eine depressive Störung hinweist. Auf dieser Grundlage können auch die bereits erwähnten Cutoff-Scores ermittelt werden, die bei der Erfassung affektiver Störungen häufig zum Einsatz kommen.

Dabei sei darauf hingewiesen, dass die Kennwerte eines psychometrischen Tests in Bezug auf die Ergebnisse ermittelt werden, die durch die Eichstichprobe erzielt wurden, d.h. streng genommen ist ein Test nur auf Personen anwendbar, die nach rationalen Überlegungen mit der Normstichprobe vergleichbar sind. In diesem Zusammenhang spielen häufig gewisse

Altersbereiche eine Rolle. Wenn beispielsweise für einen Depressionstest nur Vergleichswerte einer Stichprobe junger Erwachsener existieren, so kann das Testergebnis eines 85-jährigen Patienten auf dieser Grundlage nicht vernünftig interpretiert werden; es ist schließlich nicht bekannt, ob alte Menschen bei dem Testverfahren nicht generell völlig anders abschnitten als jüngere.

#### *2.3.6.5 Ökonomie*

Die Ökonomie stellt ein Gütekriterium dar, das unter theoretischen Gesichtspunkten relativ trivial ist, dem aber in der Praxis häufig ein hoher Stellenwert zukommt. Sie wird durch das Testkuratorium der Föderation deutscher Psychologieverbände definiert als das „Ausmaß, in dem ein Test bei der Durchführung, Auswertung und Anschaffung als kostengünstig zu bewerten ist (Amelang und Zielinski, 2002, S.146). So stellt sich bei einem Messinstrument beispielsweise die Frage, welchen Zeitaufwand es erfordert, wie aufwendig die Schulungsmaßnahmen für den Untersucher sind, wie viel Material benötigt wird etc. Vor allem für Verfahren, die auf eine breite Anzahl von Personen angewendet werden sollen, sind solche Kriterien bei der Auswahl eines Tests in hohem Maße relevant.

#### *2.3.6.6 Die Gewichtung von Gütekriterien bezogen auf die Zielsetzung eines Tests*

Wie zu Beginn dieses Kapitels erwähnt, müssen zur angemessenen Beurteilung eines Testverfahrens immer auch Rahmenbedingungen einbezogen werden, wie etwa die diagnostische Zielsetzung des Tests, seine Position innerhalb des diagnostischen Prozesses sowie seine Anwendungsbandbreite. Rahmenbedingungen bestimmen dabei nicht nur, welche Nebengütekriterien für das Testverfahren relevant sind, sondern sie spielen auch bei der Auswahl der Methoden zur Ermittlung von Reliabilität und Validität sowie bei der Gewichtung der Kennwerte eine Rolle. Am Beispiel von Screeningverfahren sei dies erläutert: Screeningtests sind Verfahren, die zur Früherkennung relevanter diagnostischer Probleme konzipiert werden. Sie stehen am Anfang des diagnostischen Prozesses und werden

in der klinischen Psychologie zur Selektion von Personen eingesetzt, bei denen das Vorliegen einer psychischen Störung wahrscheinlich ist. In der Diagnostik depressiver Störungen sollen sie etwa die Frage beantworten, ob eine Person ein Maß an Depressivität aufweist, das als klinisch auffällig erscheint, ob also weitere Diagnostik erforderlich ist.

Solche Verfahren müssen somit eine hohe Anwendungsbreite aufweisen, in der Durchführung möglichst ökonomisch sein und Personen zuverlässig als depressiv bzw. nicht-depressiv klassifizieren, d.h. eine hohe Effektivität innehaben. Die Sensitivität ist für diese Tests das wichtigste Gütekriterium. So sollten möglichst alle Risikopersonen erfasst werden. Dabei darf jedoch die Spezifität des Verfahrens nicht vollkommen vernachlässigt werden, da eine zu hohe Anzahl falsch-positiver Ergebnisse ihrerseits die Gesamteffektivität vermindern würde.

### *2.3.7 Gütekriterien einzelner Items*

Ebenso, wie die Qualität von Testverfahren im Ganzen ermittelt werden kann, gibt es Methoden, die die Messeigenschaften einzelner Items erfassen. Eine solche Itemanalyse findet in der Regel hauptsächlich in der Entwicklungsphase oder bei der Modifikation von Testverfahren statt. Sie dient bei einem Fragebogen beispielsweise dazu, diejenigen Fragen auszuwählen, die am besten zur Merkmalerfassung geeignet sind. Im Folgenden werden die Kennwerte vorgestellt, die bei der Beurteilung von Einzelitems am aussagekräftigsten sind.

#### *2.3.7.1 Schwierigkeit*

Der Schwierigkeitsindex ist ein Maß, das angibt, „wie groß der relative Anteil von Probanden ist, die ein Item „richtig“ (also im Sinne höherer Merkmalsausprägung) beantworten“ (Amelang & Zielinski, 2002, S.121). Ist der numerische Wert des Schwierigkeitsindex hoch, so bedeutet dies, dass das Item von den meisten Probanden gelöst wird und somit eine geringe Schwierigkeit aufweist. D.h., eine niedrige Merkmalsausprägung ist ausreichend um das Item zu lösen (im Sinne des Merkmals zu beantworten). Ist der Schwierigkeitsindex des

Items dagegen niedrig, so muss der Patient eine hohe Merkmalsausprägung aufweisen, um des Item entsprechend zu beantworten. In diesem Fall handelt es sich um ein Item mit hoher Schwierigkeit. Am Beispiel von Items zur Selbstbeurteilung von Depressivität, sei dies veranschaulicht: Um eine Frage wie „Sind sie manchmal traurig?“ zu bejahen, reicht bereits ein geringes Maß an Depressivität aus; somit wird das Item von den meisten Probanden gelöst und es hat eine niedrige Schwierigkeit. Stellt man dagegen die Frage „Fühlen sie sich die meiste Zeit hoffnungslos und verzweifelt?“, kann man davon ausgehen, dass nur wenige Personen das Item lösen: nämlich nur solche, die in hohem Maße depressiv sind. Es hat also eine hohe Schwierigkeit.

#### *2.3.7.2 Trennschärfe*

„Die Trennschärfe einer Aufgabe ist definiert als die Korrelation der Itembeantwortung mit dem Summenwert der Skala, zu der das betreffende Item gehört“ (Amelang & Zielinski, 2002, S.127). Sie gibt das Ausmaß an, in dem die Beantwortung eines einzelnen Items mit dem Gesamtergebnis des Testverfahrens übereinstimmt. Erreicht das Item einen hohen Korrelationskoeffizienten, so handelt es sich um ein trennscharfes Item. Ein trennscharfes Item einer Depressionsskala, ist demnach gut dazu geeignet, Probanden in Depressive und Nicht -Depressive zu unterteilen.

*Der Zusammenhang zwischen Schwierigkeit und Trennschärfe.* Da Items mit einer mittleren Schwierigkeit per definitionem von etwa der Hälfte der Probanden gelöst werden, sind sie in der Regel besonders gut dazu geeignet, Probanden im mittleren Bereich der Merkmalsausprägung zu trennen. Items mit besonders hoher oder niedriger Schwierigkeit trennen Probanden dagegen in anderen Bereichen. In einem Depressionstest wären beispielsweise Items mit mittlerer Schwierigkeit gut dazu geeignet, den Krankheitswert einer Symptomatik zu bestimmen, während Items mit hoher Schwierigkeit den Schweregrad einer depressiven Erkrankung genauer festlegen könnten; Items mit niedriger Schwierigkeit könnten dagegen bereits Probanden mit subdiagnostischer Depressivität erfassen. Dabei gibt

es keine festen Richtwerte, welchen Schwierigkeitsindex die Items eines Tests idealerweise haben sollten; diese Entscheidung ist eher von der Zielsetzung des Verfahrens abhängig. Allerdings wird in als homogen konzipierten Verfahren für die meisten Items ein mittlerer Wert zwischen 0.4 und 0.6 angestrebt. Um Merkmalsabstufungen zu erlauben, sollten jedoch auch einige Items mit hoher und geringer Schwierigkeit enthalten sein (Fisseni, 2004).

#### *2.3.7.3 Itemvalidität*

Bei der Neukonstruktion oder Modifikation eines Testverfahrens ist die Trennschärfe als skaleninternes Kriterium zur Beurteilung der Items mitunter weniger geeignet, da die Validität des gesamten Verfahrens nicht als gewährleistet vorausgesetzt werden kann. In solchen Fällen besteht die Möglichkeit, die einzelnen Items selbst an einem Außenkriterium zu validieren. Die Methoden die dabei angewendet werden, sind mit den Verfahren vergleichbar, mit denen die Validität des gesamten Tests ermittelt wird. Man behandelt dabei in der Analyse jedes einzelne Item wie einen Gesamttest und geht je nach Skalenniveau nach den in Abschnitt 2.3.6 beschriebenen Schritten vor.

#### *2.3.8 Zusammenfassung*

Aufgabe der klinisch-psychologischen Diagnostik depressiver Störungen ist die qualitative und quantitative Beschreibung des Krankheitsbildes, die Klassifizierung der Symptomatik nach den Richtlinien der ICD-10 sowie gegebenenfalls die Einleitung und Evaluation entsprechender Therapieverfahren. Da eine korrekte Diagnostik von hoher Praxisrelevanz ist, ist die Gewinnung diagnostischer Informationen an zahlreiche Regeln und Vorschriften gebunden, die den sogenannten diagnostischen Prozess anleiten. Dieser Prozess bezeichnet eine Abfolge von Maßnahmen, die von der Beobachtung von Verhaltensauffälligkeiten über die Ableitung einer diagnostischen Fragestellung und Hypothesenbildung zur Erhebung diagnostisch relevanter Informationen und schließlich zur Formulierung einer Diagnose bzw. eines Therapievorschlages führen. Im Hinblick auf depressive Störungen beziehen sich

relevante diagnostische Informationen auf Anzahl und Ausprägung der Symptome, den bisherigen Krankheitsverlauf bzw. die psychiatrische Vorgeschichte des Patienten sowie gegebenenfalls auf Ursachen oder Ausschlusskriterien für die Erkrankung.

Die Erhebung relevanter, objektiver Informationen wird in der klinischen Diagnostik durch zahlreiche standardisierte, psychometrische Testverfahren erleichtert, die die Abbildung des Merkmals Depressivität durch einen Messwert ermöglichen, der durch einen Vergleich mit entsprechenden Normstichproben interpretiert werden kann. Zu den am häufigsten verwendeten Verfahren zählen auf der Selbstbeurteilung des Patienten beruhende standardisierte Fragebögen sowie halbstrukturierte Interviews, die neben der Selbstbeurteilung auch die Fremdbeobachtung des Diagnostikers mit einbeziehen.

Da mögliche Fehlerquellen standardisierter Messverfahren durch eine sorgfältige Testkonstruktion weitgehend vermieden werden können, existieren zahlreiche durch das Testkuratorium deutscher Psychologieverbände festgelegte Gütekriterien, anhand derer die Qualität eines Tests beurteilt werden kann. Neben den Hauptgütekriterien Objektivität, Reliabilität sowie Validität eines Testverfahrens, die sich auf messtheoretische Überlegungen der Klassischen Testtheorie beziehen, sind für die Anwendung in der Praxis häufig Nebengütekriterien wie Ökonomie und Normierung des Tests relevant. Zur Beurteilung von in der Depressionsdiagnostik häufig verwendeten Screeningverfahren ist die Effektivität des Testverfahrens das wichtigste Qualitätsmerkmal.

Bei der Entwicklung oder Modifikation von Testverfahren werden in einer Itemanalyse auch die psychometrischen Eigenschaften der einzelnen Items ermittelt. Die wichtigsten Kennwerte sind dabei die Schwierigkeit, die Trennschärfe sowie die Validität des Items.

#### *2.4. Die Diagnostik depressiver Störungen im Alter*

Die Diagnostik depressiver Störungen im Alter stellt für die klinische Praxis eine besondere Herausforderung dar. So existieren zahlreiche Studien, die belegen, dass vor allem Hausärzte, die in der Regel der erste Ansprechpartner für ältere Menschen mit psychischen Problemen sind, depressive Symptome in vielen Fällen nicht als solche erkennen (Geiger-Kabisch & Weyerer, 1991; Katona, 1994). Aber auch in Pflegeheimen und Krankenhäusern werden Depressionen bei geriatrischen Patienten häufig übersehen (Bach et al., 1994). Heuft, Kruse & Radebold (2000) vermuten, dass insgesamt etwa 40% der depressiven Erkrankungen im höheren Lebensalter nicht korrekt diagnostiziert werden.

##### *2.4.1 Phänomene des Alters als „Störvariablen“ der Depressionsdiagnostik*

Die Umstände, die das Erkennen depressiver Störungen im Alter erschweren, sind vielfältig. Auf der einen Seite zeigt sich in diesem Lebensabschnitt häufig ein verändertes Krankheitsbild, sodass von einer hohen Variabilität der Symptome ausgegangen werden muss (vgl. Abschnitt 2.2.3). Andererseits führt ein negatives gesellschaftliches Altersbild mitunter dazu, dass die Symptome vom Umfeld, aber auch von dem Erkrankten selbst als „normale“ Alterserscheinung interpretiert werden, was zu großer Zurückhaltung gegenüber psychologischer oder medizinischer Intervention führt (Gauggel & Birkner, 1998; Reischies, 2003). Zudem ist die Kooperationsfähigkeit älterer Patienten gegenüber aufwendigen oder komplizierten diagnostischen Verfahren in vielen Fällen ohnehin durch sensorische, motorische oder kognitive Defizite eingeschränkt (Gauggel & Birkner, 1998).

Und nicht zuletzt erschweren Symptomüberlappungen mit im Alter häufig auftretenden Erkrankungen die Differentialdiagnostik erheblich. In diesem Kontext stellt vor allem die Abgrenzung depressiver Störungsbilder von demenziellen Erkrankungen bei älteren Menschen ein Problem dar. So sind Symptome wie etwa verminderter Antrieb, Interessenverlust, Konzentrations- und Merkfähigkeitsstörungen, die häufig als

Begleitsymptomatik bei Depressionen im Alter auftreten, andererseits typische Charakteristika für eine beginnende Demenz (Reischies, 2003). Dabei sind durch depressive Störungen verursachte Konzentrations- und Merkfähigkeitsstörungen in der Regel reversibel und werden von dem Erkrankten in der subjektiven Wahrnehmung eher überschätzt, während bei Demenzerkrankungen die kognitiven Defizite progredient und durch neuropsychologische Testungen objektivierbar sind (Reischies, 2003). Da jedoch viele Patienten mit einer Demenzdiagnose zusätzlich an einer Depression leiden (Davison & Neale, 2002), stellt die Ermittlung der korrekten Diagnose für weniger erfahrene Kliniker häufig eine Überforderung dar, sodass in der Praxis von zahlreichen Fehldiagnosen ausgegangen werden muss. Tabelle 10 zeigt einen von Beblo et al. (2005) erstellten Überblick zur differentialdiagnostischen Unterscheidung zwischen depressiven Störungen und demenziellen Erkrankungen. Abgesehen von diesem speziellen Problem der Differentialdiagnostik erschweren zahlreiche durch körperliche Erkrankungen bedingte Beschwerden bei älteren Menschen eine korrekte Klassifikation des klinischen Bildes. So sind vor allem somatische Symptome wie etwa Schlafstörungen, Appetitlosigkeit und Energielosigkeit zwar typische Merkmale für Depressionen im Alter, jedoch auch unter nicht depressiven Älteren verbreitet, und weisen daher nur wenig Spezifität bei der Identifikation affektiver Erkrankungen auf (Brink, Yesavage et al., 1982). Angesichts dieser Schwierigkeiten, Depressionen im Alter zu erkennen, besteht eine große Nachfrage nach Messinstrumenten, die auf die speziellen Bedingungen dieses Lebensabschnittes zugeschnitten sind (Katona, 1994).

Tabelle 10: Differentialdiagnostische Unterscheidung zwischen Demenz und Depression

Krankheitsmerkmale	Depression	Demenz
<b>Neuropsychologisches Profil</b>		
Allgemeines Profil	<ul style="list-style-type: none"> <li>- Keine eindeutig imponierenden Defizite, am ehesten Flexibilitätsdefizit</li> <li>- Defizite eher bei Aufgaben mit Zeitlimit</li> <li>- Defizite eher bei emotional valentem Material</li> <li>- Verstärkte Defizite nach Misserfolg</li> <li>- Variierende Testleistungen</li> </ul>	<ul style="list-style-type: none"> <li>- Primär Gedächtnisstörungen</li> <li>- Defizite auch bei „Speed-unabhängigen“ Aufgaben</li> <li>- Defizite treten valenzunabhängig auf</li> <li>- Defizite unabhängig von Misserfolg</li> <li>- Konstant herabgesetzte Leistung</li> </ul>
Neugedächtnis	<ul style="list-style-type: none"> <li>- Besserer Erhalt von Wiedererkennensleistung</li> <li>- Auslassungsfehler</li> </ul>	<ul style="list-style-type: none"> <li>- Wiedererkennen ist deutlich Gestört</li> <li>- Intrusionen, Zufallsfehler, Rateverhalten</li> </ul>
Zeichnen	Eher ungenaues Arbeiten	Konstruktiv apraktische Fehler
Orientierung	Orientierung ist weitgehend intakt	Orientierungsstörungen
Sprache	Verlangsamung, Hypophonie	Aphasische Symptome
Räumlich-konstruktive Leistungen, Praxie	Eher ungenaues Arbeiten Nachlässigkeiten	Apraktische Symptome
<b>Defizitstärke</b>	Eher geringgradig	Eher deutlich
<b>Verhalten</b>	<ul style="list-style-type: none"> <li>- Klare Schilderung kognitiver Defizite</li> <li>- Misstrauen gegen sich (mangelnde Selbstwirksamkeitserwartung)</li> </ul>	<ul style="list-style-type: none"> <li>- Ungenaue Schilderung oder Verbergen kognitiver Defizite</li> <li>- Misstrauen gegen andere</li> </ul>
<b>Sonstige affektive Symptomatik</b>	<ul style="list-style-type: none"> <li>- Dysphorisch depressiv oder Ängstlich hilflos</li> <li>- Konsistent depressiv</li> <li>- zirkadiane Schwankungen („Morgentief“)</li> <li>- Häufig morgendliches Früherwachen, Ein- und Durchschlafstörungen</li> <li>- Antriebsstörungen</li> <li>- Appetitstörungen</li> <li>- Suizidgedanken</li> <li>- Pessimistisch- Grüblerisch</li> </ul>	<ul style="list-style-type: none"> <li>- Flacher Affekt oder emotional Labil</li> <li>- Fluktuierend depressiv</li> <li>- Nicht zirkadian gebundene Stimmungsschwankungen</li> <li>- Umkehr des Schlaf-Wachrhythmus</li> <li>- Verminderte Wachheit</li> </ul>
<b>Neurologische Auffälligkeiten</b>	Eher nicht vorhanden	Häufig vorhanden
<b>Krankheitsverlauf</b>	Abhängig von depressiven Episoden	Schleichende Verschlechterung

Anmerkungen. Apraxie = Defizit in Ausführung von Handlungen. Aphasie = Defizit in Sprachverständnis oder Sprachproduktion. Zirkadian = gebunden an den Schlaf-Wach-Zyklus.

## 2.4.2 Screeningverfahren in der geriatrischen Praxis

### 2.4.2.1 Definition des Screening-Begriffes

In einem Übersichtsartikel der WHO (Perel-Levin, 2008) werden Voraussetzungen für Screeningprogramme genannt, die gewährleisten sollen, dass von dem Verfahren sowohl der Patient als auch das Gesundheitssystem profitieren kann. Mit dem Begriff „Screening“ ist an dieser Stelle weitaus mehr gemeint, als die Entwicklung und Durchführung kurzer diagnostischer Verfahren. Der gesamte Prozess schließt nicht nur die Identifikation von Risikogruppen und erkrankten Personen ein, sondern bezieht sich auch auf gegebenenfalls notwendige Behandlungsimplicationen:

- The condition should be an important health problem, well understood and with a known risk factor, or indicator
- The test should be simple, safe and validated
- The screening test should be acceptable to the population
- There should be available effective interventions to follow up
- There must be evidence from reliable randomized controlled trials that the screening programme reduces mortality or morbidity and is cost effective
- There are adequate staff available
- There should be evidence that the complete screening programme (from test to intervention) is “clinically, socially and ethically acceptable to health professionals and the public.”

(Perel-Levin, 2008, S. 13)

In den o.g. Punkten werden Aspekte genannt, die über das Erstellen von Diagnose-Instrumenten hinausgehen, wie z.B. Akzeptanz der Zielgruppe und adäquates Personal zur Durchführung des Tests. Darüber hinaus wird auch der Anspruch gestellt, dass es für den Fall eines positiven Testergebnisses eine entsprechende Intervention geben muss. Der Anspruch, im Falle einer vorhandenen Erkrankung eine angemessene Behandlung des Betroffenen gewährleisten zu können, mag auf den ersten Blick banal erscheinen; die Versorgung depressiver Erkrankungen in der geriatrischen Praxis spricht allerdings, trotz eines verhältnismäßig flächendeckenden Screenings häufig eine andere Sprache. So sind

beispielsweise die Behandlungsmöglichkeiten für depressive ältere Patienten, die über die rein pharmakologische Versorgung hinausgehen, stark eingeschränkt (vgl. Heidenblut & Zank, 2009). Auch die Forderung nach adäquatem Personal, welches das jeweilige Screeningverfahren durchführen soll, kann in der Praxis keineswegs immer als erfüllt gelten. Kurzskalen werden häufig von Klinikmitarbeitern durchgeführt, denen die Grundlagen der Testdiagnostik nicht geläufig sind, was zu einer unzuverlässigeren Durchführung und Interpretation der Verfahren führen kann. So kommt beispielsweise in der Praxis häufiger vor, dass Fragen vom Anwender der Skala spontan umformuliert werden, dass Aussagen der Patienten, die sich nicht für eine Antwortkategorie entscheiden können, in unzulässiger Weise interpretiert werden oder dass in der Auswertung Missings als Indikatoren für das Vorliegen einer depressiven Verstimmung gewertet werden, ohne dass andere mögliche Gründe für das Ausbleiben einer eindeutigen Antwort berücksichtigt werden. Auch die Voraussetzungen für die Durchführbarkeit des Screenings, wie z.B. eine hinreichende kognitive Leistungsfähigkeit bei der Anwendung von Selbstbeurteilungsfragebögen, werden nicht immer fachgerecht geprüft. Entsprechende Vorgehensweisen mögen zwar die Akzeptanz und Durchführbarkeit der Verfahren in der Praxis erleichtern, sie mindern allerdings zugleich Objektivität, Reliabilität und Validität des entsprechenden Tests. Für die Praxis von Screeningverfahren bedeutet dies einerseits, dass das dafür vorgesehen Personal idealiter vorher geschult werden sollte. Andererseits kann man davon ausgehen, dass Durchführungs- und Interpretationsfehler der Anwender auch bei denjenigen Verfahren erhöht sind, die nur über eine geringe Augenscheinvalidität verfügen. D.h., ein Anwender, der von der Gültigkeit einer Frage nicht überzeugt ist, wird bei mangelndem Wissen eher dazu tendieren, diese Frage spontan umzuformulieren, als er es mit einer Frage tun würde, die ihm zur Erfassung des Krankheitsbildes geeignet erscheint. Testverfahren, die den „gesunden Menschenverstand“ der Anwender überzeugen, weisen daher meist auch bessere Werte in den Gütekriterien der KTT auf.

#### 2.4.2.2 Pro- und Contra-Argumente zu flächendeckendem Screening

Zur den Für- und Gegenargumenten des flächendeckenden Screenings lassen sich jeweils gut nachvollziehbare Argumente formulieren (siehe Tabelle 11).

*Tabelle 11: Vor – und Nachteile von Flächendeckendem Screening*

Pro	Contra
Objektivität der Untersuchungen wird erhöht.	
Depressive Erkrankungen werden häufiger erkannt und können entsprechend behandelt werden.	Die Anzahl Falsch positiver Befunde wird durch Screening möglicherweise stark erhöht; das kann zu folgenreichen Fehlbehandlungen führen.
Bei guter Anwendung: Stigmatisierung von Hochrisikogruppen wird verhindert; unerwünschte Verhaltensweisen von Patienten werden weniger auf das Vorliegen einer depressiven Erkrankung attribuiert.	Bei schlechter Anwendung: Scheinbare Objektivierung von diagnostischen Fehlteilen; fehlerhafte Screening-Befunde werden in eine bestimmte, beim Diagnostiker bereits angelegten Sichtweise interpretiert und nicht hinreichend hinterfragt.

So erhöht Screening einerseits die Wahrscheinlichkeit, dass tatsächlich vorhandene Erkrankungen erfasst werden, andererseits besteht jedoch auch immer die Gefahr von vermehrten falsch positiven Test-Befunden. Dabei lassen sich die meisten Argumente, die gegen flächendeckendes Screening sprechen, durch den Hinweis entkräften, dass das Screening verantwortungsvoll und nach wissenschaftlichen und ethischen Kriterien durchgeführt werden sollte, wobei die Teilnahme an dem Verfahren nach Möglichkeit dem einzelnen Patienten überlassen werden sollte. Denn, sofern eine adäquate Möglichkeit der Behandlung für die Erkrankung, die Gegenstand des Screenings ist, existiert, kann durch ein solches Verfahren für den Patienten und sein Umfeld einiges getan werden, um Folgeerscheinungen zu verhindern oder zumindest abzumildern.

#### 2.4.3 Die Entwicklung der Geriatrischen Depressionsskala (GDS)

Trotz der bekannten Problematik der Depressionsdiagnostik im Alter existieren nicht viele Diagnoseinstrumente, die für geriatrische Patienten entwickelt bzw. an entsprechenden

Stichproben normiert oder validiert wurden und die in Abschnitt 2.4.1 genannten Überlegungen mit einbeziehen (Hautzinger & Meyer, 2002). So stellt die 1983 von Yesavage et al. vorgestellte Geriatriische Depressionsskala (GDS) das einzige standardisierte Screeninginstrument dar, das sich in der Praktikabilität der Durchführung sowie in der Formulierung der Items speziell auf diese Patientengruppe bezieht (Katona, 1994). Dabei war es Ziel der Autoren, einen einfachen, ökonomischen und leicht zu beantwortenden Selbstbeurteilungsfragebogen zu entwickeln, der sich an den Belangen und dem Leistungsvermögen älterer Menschen orientiert, eine gute Diskriminationsfähigkeit aufweist, möglichst vielfältig einsetzbar ist und eine hohe Akzeptanz bei Patienten und Klinikern besitzt (Gauggel & Birkner, 1998). Der Fragebogen soll dabei in erster Linie dazu geeignet sein, Probanden als depressiv oder nicht- depressiv zu klassifizieren und orientiert sich weniger an formalen diagnostischen Richtlinien (Yesavage, Brink et al., 1983). Durch die Nichtbeachtung somatischer Symptome soll eine mögliche Fehlerquelle falsch-positiver Diagnosen ausgeschlossen werden. Zudem wird die Tendenz depressiver alter Menschen, ihre kognitiven Defizite zu überschätzen, berücksichtigt.

In der ersten Entwicklungsstudie wurden von im geriatrischen Bereich erfahrenen Klinikern 100 Items mit einem dichotomen Ja-/Nein- Antwortschema zur Erfragung von Depressivität konstruiert und einer Stichprobe von 47 älteren Probanden vorgelegt, die entweder psychisch gesund oder aufgrund depressiver Erkrankungen in stationärer Behandlung waren. In den Fragebogen wurden diejenigen 30 Items miteinbezogen, die am höchsten mit dem Gesamtskalenwert der 100 Items korrelierten, die also die höchste Trennschärfe aufwiesen. Die so ermittelten 30 Items der Skala beziehen sich bei bewusster Vernachlässigung somatischer Beschwerden auf affektive, kognitive und motivationale Aspekte von Depressivität bei älteren Menschen. Dabei sind 2/3 der Items so formuliert, dass sie mit einer positiven Antwort im Sinne der Merkmalsausprägung gelöst werden, während bei 1/3 der Fragen mit einer negativen Beantwortung das Item gelöst wird. Für jedes

symptomatisch beantwortete Item wird jeweils ein Punkt vergeben. In der Auswertung werden die Punkte zu einem Summenscore addiert, wobei ein höherer Testscore für eine höhere Merkmalsausprägung steht.

In der anschließenden Validierungsstudie wurde die GDS von zwei Probandengruppen beantwortet, von denen die erste (N= 40) aus psychisch gesunden älteren Menschen mit selbstständiger Haushaltsführung bestand und die zweite (N= 60) aus Patienten, die aufgrund depressiver Störungen ambulant oder stationär in Behandlung waren. Alle Probanden wurden nach den diagnostischen Forschungsrichtlinien für psychische Störungen klassifiziert und es wurden neben dem GDS-Wert die Werte auf zwei weiteren bewährten Depressionsskalen, der Hamilton Rating Scale for Depression (HRS-D) und der Self-Rating Depression Scale (SDS) ermittelt. Die Ergebnisse sprachen für eine hohe Reliabilität der neuen Skala sowie für eine hohe konvergente Validität mit den bewährten Messinstrumenten.

	Choose the best answer for how you felt over the past week	
1.	<b><i>Are you basically satisfied with your life?</i></b>	yes/ no
2.	<b>Have you dropped many of your activities and interests?</b>	yes/ no
3.	<b>Do you feel that your life is empty?</b>	yes/ no
4.	<b>Do you often get bored?</b>	yes/ no
5.	<i>Are you hopeful about the future?</i>	yes/ no
6.	Are you bothered by thoughts you can't get out of your head?	yes/ no
7.	<b><i>Are you in good spirits most of the time?</i></b>	yes/ no
8.	<b>Are you afraid that something bad is going to happen to you?</b>	yes/ no
9.	<b><i>Do you feel happy most of the time?</i></b>	yes/ no
10.	<b>Do you often feel helpless?</b>	yes/ no
11.	Do you often get restless and fidgety?	yes/ no
12.	<b>Do you prefer to stay at home rather than going out and doing new things?</b>	yes/ no
13.	Do you frequently worry about the future?	yes/ no
14.	<b>Do you feel you have more problems with your memory than most?</b>	yes/ no
15.	<b><i>Do you think it is wonderful to be alive now?</i></b>	yes/ no
16.	Do you often feel downhearted and blue?	yes/ no
17.	<b>Do you feel pretty worthless the way you are now?</b>	yes/ no
18.	Do you worry a lot about the past?	yes/ no
19.	<i>Do you find life very exciting?</i>	yes/ no
20.	Is it hard for you to get started on new projects?	yes/ no
21.	<b><i>Do you feel full of energy?</i></b>	yes/ no
22.	<b>Do you feel that your situation is hopeless?</b>	yes/ no
23.	<b>Do you think that most people are better off than you are?</b>	yes/ no
24.	Do you frequently get upset over little things?	yes/ no
25.	Do you frequently feel like crying?	yes/ no
26.	Do you have trouble concentrating?	yes/ no
27.	<i>Do you enjoy getting up in the morning?</i>	yes/ no
28.	Do you prefer to avoid social gatherings?	yes/ no
29.	<i>Is it easy for you to make decisions?</i>	yes/ no
30.	<i>Is your mind as clear as it used to be?</i>	yes/ no

*Anmerkungen.* Items, die auch Bestandteil der Kurzversion sind, sind fett gedruckt. Items, die durch negative Antwort gelöst werden, sind kursiv dargestellt.

*Abbildung 10: Die Geriatriische Depressionsskala nach Yesavage et al. (1983)*

Aus Gründen größerer Ökonomie und Zumutbarkeit wurde von den Autoren einige Zeit später eine Kurzform der Skala vorgestellt, die aus den 15 Items bestand, die die höchsten

Korrelationen mit depressiven Symptomen in den bis dahin von den Autoren durchgeführten Validierungsstudien aufwiesen. Anschließende Untersuchungen ergaben gute bis befriedigende Übereinstimmungen mit der ursprünglichen Fassung (Gauggel & Birkner, 1998). Sowohl für die ursprüngliche Version als auch für die Kurzversion der GDS wurden zahlreiche Testkennwerte ermittelt. Die Ergebnisse zeigten, dass die GDS ein zuverlässiges Messinstrument mit hoher interner Konsistenz ist, das wenig störanfällig in Bezug auf Darreichungsform und positive Erinnerungseffekte ist (Yesavage et al., 1983; Brink et al., 1985; Morishita et al., 1995). An zahlreichen Stichproben ambulanter geriatrischer Patienten hat sich die GDS darüber hinaus als effektives Testverfahren für die Bestimmung krankheitswertiger depressiver Störungen bewährt (vergleiche Tabelle 12), wobei für die Originalversion ein Score von mehr als 10 Punkten und für die Kurzversion ein Score von mehr als 5 Punkten als kritische Werte vorgeschlagen wurden (Brink et al., 1982; D'ath et al., 1994; Herrmann et al., 1996). Dabei sind die Ergebnisse der Skala bei Patienten mit schwerer kognitiver Funktionseinschränkung nicht verwendbar. So wurde als Voraussetzung für die Durchführung des Verfahrens ein Mindestmaß an kognitiver Leistungsfähigkeit festgesetzt, das durch ein basales kognitives Testverfahren wie etwa den Mini-Mental-State-Exam (MMSE) (Folstein et al., 1975) vorab geprüft werden sollte. So sind die Testergebnisse der GDS nur ab einem MMSE-Score von mehr als 15 Punkten verwendbar (Bach et al., 1995). Eine von Gauggel und Schmidt (1995) vorgestellte deutsche Version der Skala ergab befriedigende Ergebnisse für Reliabilität und konvergente Validität und erwies sich bei einer kleineren Stichprobe stationärer und chirurgischer Patienten als effektives Screeningverfahren (Gauggel & Birkner, 1995).

Tabelle 12: Entwicklungsstudien zur Geriatrischen Depressionsskala

Studie	Fragestellung	N	Stichprobe	Alter	Untersuchungsverfahren	zentrale Ergebnisse
Yesavage et al. (1983) 1. Studie	Selektion von 30 Fragen aus einem 100-Item Pool	47	Personen im eigenen Haushalt o. psychische Erkrankung; depressive Pat. in stat. Behandlung	55+	100 von Experten entwickelte Ja/Nein Fragen zur Depressivität	Auswahl der 30 Items, die am höchsten mit dem Gesamtwert korrelierten ( $r = .47-.83$ , Median: $r = .68$ )
Yesavage et al. (1983) 2. Studie	Ermittlung von Reliabilität und Validität	40 60	Personen im eigenen Haushalt ohne psych. Erkrankung; depressive Pat. in ambul. und stat. Behandlung	k. A.	RDC (Außenkriterium) HRS-D SDS GDS30	<u>Reliabilität:</u> Interitemkorrelation = .36; Cronbach $\alpha = .94$ „split-half“ = .94; Test-Retest = .85 <u>konvergente Validität</u> GDS-RDC:(Varianzanalyse) $F(2, 97) = 99.48$ $P < 0.001$ GDS-SDS: $r = 0.84$ ; GDS- HRS-D: $r = 0.83$
Brink et al. (1982)	Ermittlung der Effektivität im Vergleich zu bewährten Messinstrumenten; Ermittlung von Cut-off-Scores	20 51	„normale“ ältere Personen; depressive geriatrische Pat. in Behandlung	k. A.	GDS30 HRS-D SDS	<u>Effektivität:</u> Cut-off 9+: Sensitiv.= 90%, Spezif. = 80 %; Cut-off 11+: Sensitiv. = 84%,Spezif. = 95 % Cut-off 14 + : Sensitiv.= 80%, Spezif.= 100%
Brink et al. (1985)	Ermittlung der Reliabilität: Versuchsleiter-Anordnungs- und Positiv. Erinnerungseffekte	54	alte Menschen in Pflegeheimen	k. A.	GDS30 MSQ 5 Min.-Gespräch über positive Erinnerung	<u>Reliabilität:</u> Test-retest = .86 Inter-examiner = .85
D'ath et al. (1994)	Validierung der GDS15 und Entwicklung kürzerer Versionen	194	Geriatrische Patienten in Hausarztpraxen	65+	GDS15 GMS Diagnosen	<u>Effektivität:</u> Cut-off 4/5: Sensitiv.= 91%, Spezif. = 72% Cut-off 5/6: Sensitiv. = 78%, Spezif. = 82%
Morishita et al. (1995)	Validierung von durch Telefon-Interviews gewonnenen GDS-Scores	31	Ambulante ger. Patienten	60+ m 77,3	GDS30 GDS30 am Telefon	<u>Validität:</u> GDS30-GDS30 am Telefon: $r=.90$

*Anmerkungen.* BDI: Beck Depression Inventory, CES-D: Center of Epidemiological Studies – Depression Scale, GDS: Geriatrische Depressionsskala, GMS: Geriatric Mental Status Schedule, DSMIV:Diagnostic Manual of Mental Diseases, HRS-D: Hamilton Rating Scale for Depression, ICD-10: International Klassifikation psychischer Störungen, MADRS: Montgomery and Asberg Depression-Rating-Scale, MMSE: Mini Mental State Examination, MSQ: Mantal Status Questionnaire, RDC: Research Diagnostic Criteria, SDS: Self-Rating Depression Scale,

*Fortsetzung Tabelle 12*

Studie	Fragestellung	N	Stichprobe	Alter	Untersuchungsverfahren	zentrale Ergebnisse
Gauggel & Schmidt (1995)	Ermittlung von Gütekriterien der deutschen Version der GDS	87	Personen mit selbständiger Haushaltsführung	60+ m. 72,3	GDS(deutsche Ü.) BDI CES-D	<u>Reliabilität:</u> Interne Konsistenz : r=0.8 <u>Konvergente Validität:</u> GDS-BDI : r=0,74 GDS-CES-D :r=0,67
Herrmann et al. (1996)	Validierung der GDS15 an ambul. Patienten mit affektiven Störungen	116	ambul. Pat. mit affektiver Störung (DSMIII- Diagnosen)	65+	GDS15 MADRS	<u>Konvergente Validität:</u> GDS-MADRS: r = 0.78 <u>Effektivität:</u> Cut-off 5/6: Sensitiv.= 85 %, Spezifit.= 74%
Gauggel & Birkner (1999)	Ermittlung von Validität und Reliabilität einer deutschen Version der GDS	43	stat. ger. Patienten auf psychiatrischen und chirurgischen Stationen	60+	GDS30(deutsche Ü.) GDS15 MMSE Mini-Dips DSMIV ICD-10	<u>Effektivität:</u> GDS30: Cut-off 13+: Sensitiv.= 84%, Spezifit.= 89%; GDS15 : Cut-off 5+: Sensitiv.= 84%, Spezifit.= 89%

*Anmerkungen* .BDI: Beck Depression Inventory, CES-D: Center of Epidemiological Studies – Depression Scale, GDS: Geriatriische Depressionsskala, GMS: Geriatric Mental Status Schedule, DSMIV:Diagnostic Manual of Mantal Diseases, HRS-D:Hamilton Rating Scale for Depression, ICD-10: International Klassifikation psychischer Störungen, MADRS: Montgomery and Asberg Depression-Rating-Scale, MMSE: Mini Mental State Examination, MSQ: Mantal Status Questionnaire, RDC: Research Diagnostic Criteria, SDS:Self-Rating Depression Scale.

#### *2.4.4 Depressionsscreening bei multimorbiden Patienten mit der GDS*

Die GDS ist ein speziell für ältere Menschen entwickeltes Depressions-Früherkennungsverfahren, das in der Durchführung wenig Zeit erfordert und durch ein einfaches Antwortschema geringe Anforderungen an die kognitive Leistungsfähigkeit der Patienten stellt. Sie kann daher sowohl in schriftlicher als auch in mündlicher Form angewendet werden. Zudem kann sie auch von wenig erfahrenen Klinikern problemlos eingesetzt werden, da sie in Anwendung und Auswertung keine speziellen Kenntnisse erfordert. Im angloamerikanischen Raum ist sie ein wichtiger Bestandteil renommierter Assessmentprogramme für ältere Patienten und wird in vielen geriatrischen Kliniken routinemäßig durchgeführt (Bach et al., 1995).

Auch in Deutschland stellt die GDS aktuell ein Messinstrument dar, das in der Diagnostik depressiver Störungen bei geriatrischen Patienten häufig zum Einsatz kommt. So wird sie nicht nur in epidemiologisch orientierten Studien als Indikator für Depressivität verwendet (Lieberman et al., 1999; Elsen & Müller-Thomsen, 2003; Renteln-Kruse & Ebert, 2003), sie wird auch vor der Arbeitsgruppe geriatrisches Assessment (AGAST) als Teil eines Testprogramms international bewährter Instrumente zur Optimierung der Therapie- und Versorgungsplanung geriatrischer Patienten empfohlen (AGAST, 1997). Dabei stellt die von AGAST vorgestellte Skala eine möglichst wortgetreue Übersetzung der von Yesavage et al. entwickelten Kurzform mit 15 Items dar (Abbildung 11), die in den beteiligten Kliniken seit geraumer Zeit im Rahmen des geriatrischen Basisassessments zur Diagnostik eingesetzt werden. Die Durchführung der Skala erfolgt aufgrund der motorischen sowie kognitiven Einschränkung vieler Patienten mithilfe eines Untersuchers mündlich.

GDS	Frage				entf.	Kommentar
01	Sind Sie grundsätzlich mit Ihrem Leben zufrieden?	Ja	Nein	*		
02	Haben Sie viele Ihrer Aktivitäten und Interessen aufgegeben?	Ja	Nein			
03	Haben Sie das Gefühl, Ihr Leben sei unausgefüllt?	Ja	Nein			
04	Ist Ihnen oft langweilig?	Ja	Nein			
05	Sind Sie die meiste Zeit guter Laune?	Ja	Nein	*		
06	Haben Sie Angst, dass Ihnen etwas Schlimmes zustoßen wird?	Ja	Nein			
07	Fühlen Sie sich die meiste Zeit glücklich?	Ja	Nein	*		
08	Fühlen Sie sich oft hilflos?	Ja	Nein			
09	Bleiben Sie lieber zu Hause, anstatt auszugehen und Neues zu unternehmen?	Ja	Nein			
10	Glauben Sie, mehr Probleme mit dem Gedächtnis zu haben, als die meisten anderen?	Ja	Nein			
11	Finden Sie, es sei schön, jetzt zu leben?	Ja	Nein	*		
12	Kommen Sie sich in Ihrem jetzigen Zustand ziemlich wertlos vor?	Ja	Nein			
13	Fühlen Sie sich voller Energie?	Ja	Nein	*		
14	Finden Sie, dass Ihre Situation hoffnungslos ist?	Ja	Nein			
15	Glauben Sie, dass es den meisten Leuten besser geht als Ihnen?	Ja	Nein			
16	<b>Summe</b>					

*Anmerkungen.* Items die durch negative Beantwortung gelöst werden, sind mit einem „\*“ versehen. Im Kommentarfeld kann der Untersucher Anmerkungen machen, sofern eine Frage von dem Patienten nicht eindeutig beantwortet wurde.

*Abbildung 11: Erhebungsbogen der GDS15 in der Übersetzung nach AGAST (1995)*

#### *2.4.4.1 Schwierigkeiten unter theoretischen Gesichtspunkten*

Obwohl die GDS gerade im deutschsprachigen Raum in erster Linie zur Diagnostik stationärer geriatrischer Patienten angewendet wird, gibt es keine eindeutigen Forschungsbefunde darüber, ob das Messinstrument zur Depressionsfrüherkennung bei funktionell eingeschränkten, multimorbiden älteren Menschen geeignet ist. So wurden die ersten in Tabelle 12 dargestellten Entwicklungs- und Validierungsstudien fast ausschließlich an Stichproben körperlich gesunder, im eigenen Haushalt lebender Menschen durchgeführt

bzw. an Patienten, die aufgrund ihrer depressiven Störungen in Behandlung waren. Dass die GDS sich implizit eher an eine körperlich gesunde, funktionell wenig beeinträchtigte Zielgruppe richtet, zeigt sich auch in der Formulierung einiger Items. So setzt beispielsweise Item 9 mit der Frage „Bleiben sie lieber zu Hause anstatt auszugehen und Neues zu unternehmen“, einerseits ein Setting voraus, bei dem sich der Befragte in seiner gewohnten Umgebung befindet; andererseits erfordert die nicht-symptomatische Beantwortung der Frage ein relativ hohes Maß sensorischer, motorischer und kognitiver Leistungsfähigkeit von Seiten des Probanden. Gerade diese Bedingungen sind jedoch bei institutionalisierten oder körperlich eingeschränkten Stichproben in der Regel nicht gegeben.

#### *2.4.4.2 Empirische Befunde*

Empirisches Datenmaterial zur Nützlichkeit der GDS15 für stationäre geriatrische Patienten ist widersprüchlich. Eine große Studie mit insgesamt 2032 stationären physisch kranken Patienten ergab für die GDS15 eine niedrige interne Konsistenz (Cronbachs  $\alpha = .46$ ), sodass das Instrument von den Autoren nicht als eindimensional eingestuft wurde und somit auch nicht als Gesamtskala für depressive Verstimmung nach den bekannten Normwerten verwendet werden konnte (Incalzi, Cesari, Pedone, & Carbonin, 2003). Ein systematischer Review bezüglich der Validität der GDS (Wancata et al., 2006) berichtete zwar zufriedenstellende Ergebnisse für stationäre geriatrische Patienten, allerdings bemängeln die Autoren, dass die Rekrutierungstechniken der Originalstudien kaum berichtet werden, sodass erhebliche Überschätzungen der diagnostischen Qualität des Instrumentes nicht auszuschließen sind.

Weitere Empirische Untersuchungen zur Effektivität der GDS bei der Depressionserkennung multimorbider Patienten sind widersprüchlich und schwer vergleichbar (siehe Tabelle 13). So sprechen zwar einige Studien (Gauggel & Birkner, 1999; König, 1988c; Rinaldi, 2003; Shah et al., 1997) dafür, dass Sensitivität und Spezifität der Skala für hospitalisierte Patienten mit den von Yesavage et al. (1983) angenommenen

Kennwerten vergleichbar sind, andere Untersuchungen (Harper et al., 1990; Jackson & Baldwin, 1993; Ramsay et al., 1991) zeigten jedoch eine eher mittelmäßige bis bescheidene Diskriminationsfähigkeit des Messinstrumentes bei stationären Patienten. In Studien mit stark sehgeminderten (Rovner & Shmueli-Dulitzki, 1997) oder kognitiv leicht eingeschränkten Probanden (Burke et al., 1991) unterschied sich die Treffsicherheit der GDS bei der Ermittlung krankheitswertiger depressiver Störungen nicht signifikant von der Zufallswahrscheinlichkeit. Auch in Validierungsstudien an Schlaganfallpatienten (Agrell & Dehlin, 1989) geronto-psychiatrischen Patienten (Leshner & Berryhill, 1994), hochaltrigen Probanden (Craen et al., 2003) und einer institutionalisierten Stichprobe (Parmelee, Lawton & Katz) erreichte die GDS bei weitem nicht so gute Testkennwerte in der Effektivität wie bei ambulanten Stichproben. In einer Studie mit deutschsprachigen Probanden von Raasch (1996) erbrachte die Kurzversion der GDS keine signifikanten Ergebnisse zur Depressionserfassung in der geriatrischen Rehabilitation. Eine Untersuchung von Hammond (2004) zeigte darüber hinaus, dass die GDS auch bei Klinikern im stationären Setting nicht das von den Autoren angestrebte Maß an Akzeptanz erreicht, was einen entsprechend negativen Einfluss auf die Nützlichkeit des Messinstrumentes haben dürfte.

Es spricht also einiges dafür, dass die GDS durch die spezielle Formulierung ihrer Items weniger gut dazu geeignet ist, den psychischen Zustand multimorbider Patienten zu erfassen. So schaltet sie zwar durch die Ausklammerung somatischer Symptome einen wichtigen Störfaktor in der Depressionsdiagnostik älterer Menschen aus, durch die implizite Voraussetzung eines bestimmten Settings in der Fragestellung kann sie der besonderen Situation der in vielen Funktionen eingeschränkten, hospitalisierten Patienten dennoch nicht in befriedigendem Maß gerecht werden.

Tabelle 13: Zusammenfassung der Ergebnisse aus Validierungsstudien der GDS an multimorbiden geriatrischen Patienten

Studie	Fragestellung	N	Stichprobe	Alter	Screening	zentrale Ergebnisse
Koenig et al. (1988c)	Validierung der GDS an einer Stichprobe hospitalisierter geriatrischer Patienten	128	Männliche geriatrische Patienten, aufgrund med. Erkrankungen in stationärer Behandlung	70+	GDS30 BCDRS DSMIII- Diagn. HRS-D	<u>Effektivität:</u> Cut-off 11: Sensitiv.= 92%; Spezif.=89%
Parmelee, Lawton & Katz (1989)	Validierung GDS an einer institutionalisierten Stichprobe	806	Patienten in Pflegeheimen und betreuten Wohneinrichtungen	61+ m. 84	GDS30 SADS RDS DSMIII Versch. kognitive Tests	<u>konvergente Validität:</u> GDS30- Raskin Depression Skale: r=.34 <u>Effektivität:</u> GDS30-DSMII-Klassifik. Sensitiv. für MDD: 78,3% Sensitiv. für Minor DD: 64% Spezifität: 81,9% GDS30- klinische Diagnosen Sensitiv. für Depr. gesamt: 69%; Spezif. 86% Sensitiv. für MDD: 54%; Spezifität 91%
Agrell & Dehlin (1989)	Vergleich der Effektivität von 6 Depressionsskalen bei geriatrischen Schlaganfallpatienten	40	Schlaganfallpatienten (Akutereignis = 4 –24 Mon. vor Untersuch.) darunter Tagesklinikpatienten, Pflegeheimbewohner und Patienten in stat. ger. Rehabilitation Ausschluss: MMSE <20	61+ m. 80	GDS30 SDS CES-D HRS CPRS-D Cornell Scale MMSE Barthel Index psychiatrisches Interview	<u>Effektivität</u> GDS30:Cut-off 10+: Sensitiv.= 88%, Spezifit.= 64%

Anmerkungen. ADL: Activities of Daily Living, AGECAT: short community version of GMS, BCDRS: Brief Carroll Depression Rating Scale, BAS: Brief Assessment Schedule, BDI: Beck Depression Inventory, BSI: Brief Symptom Inventory, CDR: Clinical Dementia Rating, CES-D: Center of Epidemiological Studies – Depression Scale, CPRS-D: Comprehensive Psychopathological Rating Scale-Depression, GDS: Geriatrische Depressionsskala, GMS: Geriatric Mental Status Schedule, DSMIV: Diagnostic Manual of Mental Diseases, HRS-D: Hamilton Rating Scale for Depression, IADL: Instrumental Activities of Daily Living, ICD-10: Internationale Klassifikation psychischer Störungen, MADRS: Montgomery and Asberg Depression-Rating-Scale, MMPI: Multistruktural Multiphasic Personal Inventory; MMSE: Mini Mental State Examination, MSQ: Mental Status Questionnaire, RDC: Research Diagnostic Criteria, SADS: Schizophrenia and Affective Disorders Schedule, SDS: Self-Rating Depression Scale

Fortsetzung Tabelle 13

Studie	Ziel	N	Stichprobe	Alter	Screening	zentrale Ergebnisse
Burke et al. (1989)	Validierung der GDS15 an multimorbiden geriatrischen Patienten	69 72	Geriatrische Patienten ohne kognitive Beeinträchtigung (CDR=0.; MMSE >24) Patienten mit leichter kognitiver Beeinträchtigung (CDR=1; MMSE <23)	m. 77,6 m. 78,3	GDS30 GDS15 MMSE CDR ADL IADL	<u>Effektivität</u> Probanden ohne k.B.: GDS15: Cut-off 6+: Sensitiv.= 67%, Spezif.= 78%; ROC: AUC=.86 Probanden mit l.k.B-: GDS15: Cut-off 5+: Sensitiv.=60%; Spezif.=63% ROC: AUC=.65:n.s. größer als Zufall
Harper, Kotik Harper & Kirby (1990)	Validierung der GDS an stationären geriatrischen Patienten	211	Stat. geriatrische Patienten mit körperlichen Erkrankungen	60+ m. 74,9	GDS30 RDS BSI MMPI neuropsychologische Testbatterie	<u>Effektivität:</u> Falsch-negativ-Rate für minor DD Cut-off 10+ = 58% Falsch-negativ-Rate für MDD Cut-off 20+=59% Cut-off 10+=14,5% Gesamteffektivität=63% (ermittelt durch Diskriminanzanalyse)
Ramsay et al. (1991)	Validierung der GDS an stat. Patienten in der Akutgeriatrie	88	Geriatrische Patienten, aufgrund akuter körperl. Erkrankung in stat. Behandlung	k.A.	GDS	<u>Effektivität:</u> Depressive Symptome: Sensitiv.=74%; Spezif.=72%
Jackson & Baldwin (1993)	Validierung und Vergleich mit ärztl. und pflegerischen Urteilen	59	Geriatrische Patienten, aufgrund med. Erkrankungen in stat. Behandlung	65+	GDS30 GDS15 GMS	<u>Effektivität:</u> GDS30: Cut-off 11+: Sensitiv.: 86%, Spezif.:76% GDS15: Cut-off 5+: Sensitiv.= 86%, Spezif.= 66%; Cut-off 6+: Sensitiv.= 67%; Spezif.= 79%

Anmerkungen. ADL: Activities of Daily Living, AGECAT: short community version of GMS, BCDRS: Brief Carroll Depression Rating Scale, BAS: Brief Assessment Schedule, BDI: Beck Depression Inventory, BSI: Brief Symptom Inventory, CDR: Clinical Dementia Rating, CES-D: Center of Epidemiological Studies – Depression Scale, CPRS-D: Comprehensive Psychopathological Rating Scale-Depression, GDS: Geriatrische Depressionsskala, GMS: Geriatric Mental Status Schedule, DSMIV: Diagnostic Manual of Mental Diseases, HRS-D: Hamilton Rating Scale for Depression, IADL: Instrumental Activities of Daily Living, ICD-10: International Klassifikation psychischer Störungen, MADRS: Montgomery and Asberg Depression-Rating-Scale, MMPI: Multistructural Multiphasic Personal Inventory; MMSE: Mini Mental State Examination, MSQ: Mental Status Questionnaire, RDC: Research Diagnostic Criteria, SADS: Schizophrenia and Affective Disorders Schedule, SDS: Self-Rating Depression Scale

Fortsetzung Tabelle 13

Studie	Ziel	N	Stichprobe	Alter	Screening	zentrale Ergebnisse
Leshner & Berryhill (1994)	Validierung der GDS15 an stat. psychiatrischen Pat.	13 46 13	Stat. geronto-psychiatrische Patienten mit Demenzen, Depressionen o. Denkstörungen	m. 71,2	GDS30 GDS15 DSMIII	<u>Effektivität:</u> Gesamtsample: GDS30: Cut-off 11+: Sensitiv.= 91%, Spezif.=42% GDS15: Cut-off 5+: Sensitiv.= 91%, Spezif.= 54%; Cut-off 6+: Sensitiv.= 83%, Spezif.= 68%
Raasch (1996)	Verleich von 2 Screeninginstrumenten (GDS und ADL) zur Depressionsmessung in der geriatrischen Rehabilitation	62	Geriatrische Rehabilitationspatienten in stat. o. teil-stat. Behandlung	k.A.	GDS15 (d.Ü.) ADS MMSE Barthel-Index	Im Gegensatz zur ADS liefert die GDS keine signifikanten Ergebnisse zur Depressionserfassung in der geriatrischen Reha
Rovner & Shmueli-Dulitzki (1997)	Validierung der GDS an sehgeminderten geriatrischen Patienten	70	Ambulante Patienten einer Augenklinik mit stark eingeschränktem Visus	65+	GDS30 DSMIII- D.	<u>Effektivität:</u> Cut-off 11+:Sensitiv.= 63%; Spezif.= 77% ROC: AUC = 0.44 (95% CI 0.36, 0,51): Diskriminationsfähigkeit der GDS zwischen Pat. mit und ohne MDD n.s. größer als Zufall
Shah et al. (1997)	Validierung von GDS-Kurzformen an stationären geriatrischen Patienten	53	Stat. geriatrische Pat.	80+	GDS10 BAS	<u>Effektivität:</u> GDS10: Cut-off 2/3 Sensitiv.= 88%, Spezifit.= 75%
Gauggel & Birkner (1999)	Ermittlung von Validität und Reliabilität einer deutschen Version der GDS	43	Stat. ger. Patienten auf psychiatrischen und chirurgischen Stationen	60+	GDS30(d. Ü.) GDS15 MMSE Mini-Dips DSMIV/ICD-10	<u>Effektivität:</u> GDS30: Cut-off 13+: Sensitiv.= 84%, Spezifit.= 89%; GDS15 : Cut-off 5+: Sensitiv.= 84%, Spezifit.= 89%

Anmerkungen. ADL: Activities of Daily Living, AGE-CAT: short community version of GMS, BCDRS: Brief Carroll Depression Rating Scale, BAS: Brief Assessment Schedule, BDI: Beck Depression Inventory, BSI: Brief Symptom Inventory, CDR: Clinical Dementia Rating, CES-D: Center of Epidemiological Studies – Depression Scale, CPRS-D: Comprehensive Psychopathological Rating Scale-Depression, GDS: Geriatrische Depressionsskala, GMS: Geriatric Mental Status Schedule, DSMIV:Diagnostic Manual of Mental Diseases, HRS-D:Hamilton Rating Scale for Depression, IADL: Instrumental Activities of Daily Living, ICD-10: International Klassifikation psychischer Störungen, MADRS: Montgomery and Asberg Depression-Rating-Scale, MMPI: Multistructural Multiphasic Personal Inventory; MMSE: Mini Mental State Examination, MSQ: Mental Status Questionnaire, RDC: Research Diagnostic Criteria, SADS: Schizophrenia and Affective Disorders Schedule, SDS: Self-Rating Depression Scale

Fortsetzung Tabelle 13

Studie	Ziel	N	Stichprobe	Alter	Screening	zentrale Ergebnisse
Rinaldi et al. (2003)	Validierung der GDS15 und GDS5 in drei verschiedenen Settings	61 50 70	Stat. geriatrische Pat., Tagesklinikpatienten und Pflegeheimbewohner	m. 79,4	GDS15 GDS5 DSMIV-D.	<u>Effektivität</u> GDS15: Gesamtsample: Sensitiv.=92%, Spezifit.= 83% Stat. Patienten: Sensitiv.= 90%, Spezifit.= 81%
Craen, Heeren & Gussekloo (2003)	Validierung der GDS15 an hochaltrigen Probanden einer Gemeindestichprobe	56 23	Probanden mit selbständiger Haushaltsführung, institutionalisierte Probanden	85+	GDS15 MMSE AGECAT GMS	<u>Effektivität</u> GDS15: Gesamtsample: Cut-off 3/4: Sensitiv.= 88%, Spezifit.= 76%; Cut-off 5/6: Sensitiv.= 38%, Spezifit.= 90%

Anmerkungen. ADL: Activities of Daily Living, AGECAT: short community version of GMS, BCDRS: Brief Carroll Depression Rating Scale, BAS: Brief Assessment Schedule, BDI: Beck Depression Inventory, BSI: Brief Symptom Inventory, CDR: Clinical Dementia Rating, CES-D: Center of Epidemiological Studies – Depression Scale, CPRS-D: Comprehensive Psychopathological Rating Scale-Depression, GDS: Geriatrische Depressionsskala, GMS: Geriatric Mental Status Schedule, DSMIV: Diagnostic Manual of Mental Diseases, HRS-D: Hamilton Rating Scale for Depression, IADL: Instrumental Activities of Daily Living, ICD-10: International Klassifikation psychischer Störungen, MADRS: Montgomery and Asberg Depression-Rating-Scale, MMPI: Multistructural Multiphasic Personal Inventory; MMSE: Mini Mental State Examination, MSQ: Mental Status Questionnaire, RDC: Research Diagnostic Criteria, SADS: Schizophrenia and Affective Disorders Schedule, SDS: Self-Rating Depression Scale

#### *2.4.5 Kürzere Versionen der GDS*

Die Forschungsliteratur berichtet einige Bemühungen, die GDS weiter zu kürzen. So existieren sehr kurze Versionen mit zehn, (Shah et al., 1996), fünf (Kenny Weeks, McGann, Michaels & Pennix, 2003) oder nur vier Items (Goring, Baldwin, Marriott, Pratt & Roberts, 2004). Die Validierungsstudien dieser Skalen sind allerdings nach methodischen Gesichtspunkten als problematisch zu bewerten, da sie entweder auf sehr kleinen oder speziellen Stichproben basieren (Rinaldi et al., 2003), die Skala lediglich an längeren Versionen der GDS validieren (Goring et al., 2004; Kenny Weeks et al., 2003), oder niedrige Spezifitätskennwerte berichten (Pomeroy, Clark, & Philp, 2001; Shah et al., 1996). Entsprechend können diese Kurzverfahren basierend auf dem bisherigen Datenmaterial nicht als Alternative für die klinische Praxis gelten und unter Berücksichtigung der theoretischen und methodischen Überlegungen dieser Arbeit sind entsprechende Bemühungen auch kritisch zu sehen.

Ein Versuch, die GDS um Items zu bereinigen, die mit dem Geschlecht, dem Alter, dem ethnischen Hintergrund sowie dem Gesundheitszustand des Patienten assoziiert sind, resultiert in einer Skala mit 7 Items (Broekman, Niti, Nyunt, Meng Ko, Kumar & Pin, 2011). Dabei enthält jedoch auch diese Skala Items (z.B. „Ist Ihnen oft langweilig?“, „Glauben Sie, dass es den meisten Leuten besser geht als Ihnen?“), die in einem stationären Setting mit Umgebungsfaktoren konfundiert sind. Insofern stellt auch diese theoretisch besser fundierte, kürzere Version der Skala keine echte Alternative für ein Screeninginstrument in der stationären Geriatrie dar.

#### *2.4.6 Zusammenfassung*

Bei etwa 40% der älteren Bevölkerung werden depressive Störungen nicht erkannt und behandelt. Dabei erschweren zahlreiche im höheren Lebensalter wirksame Faktoren wie ein negatives gesellschaftliches Altersbild, die Zurückhaltung älterer Patienten gegenüber psychologischen oder psychiatrischen Untersuchungen, kognitive und sensorische

Einschränkungen sowie die Komorbidität mit demenziellen und körperlichen Erkrankungen die Depressionsdiagnostik in diesem Lebensabschnitt erheblich. Trotz dieser Problematik existieren nur wenige Messinstrumente, die speziell für die Zielgruppe geriatrischer Patienten entwickelt und an entsprechenden Stichproben normiert wurden.

Die Geriatrische Depressionsskala (GDS) ist ein standardisierter Selbstbeurteilungs-Fragebogen zur Depressivität mit dichotomem Ja/Nein- Antwortschema, der mit der Zielsetzung konstruiert wurde, ein ökonomisches, einfach zu handhabendes Screening-Instrument zu schaffen, das von Patienten und Untersuchern gleichermaßen akzeptiert wird. Durch die Nichtbeachtung somatischer Symptome wird in der Skala eine mögliche Fehlerquelle falsch-positiver Diagnosen ausgeschlossen. Zudem wird die Tendenz depressiver alter Menschen, ihre kognitiven Defizite zu überschätzen, berücksichtigt. Neben der ursprünglichen Form mit 30 Items wurde aufgrund besserer Praktikabilität eine Kurzform entwickelt, die aus den 15 trennschärfsten Items der ursprünglichen Skala besteht. Validierungsstudien an gesunden Probanden bzw. an Patienten, die sich aufgrund depressiver Erkrankungen in Behandlung befanden, ergaben gute Ergebnisse für Reliabilität, konvergente Validität und Effektivität für beide Versionen der Skala. Erste Untersuchungen mit einer deutschen Übersetzung erbrachten ebenfalls befriedigende Kennwerte.

Die GDS15 ist Bestandteil zahlreicher Assessmentprogramme für ältere Menschen im anglo-amerikanischen Raum und wird in Deutschland von der Arbeitsgruppe Geriatrisches Assessment (AGAST) als Depressionsscreening geriatrischer Patienten empfohlen. In den an AGAST beteiligten Kliniken wird eine möglichst wortgetreue Übersetzung der GDS-Kurzform seit geraumer Zeit routinemäßig durchgeführt. Dabei ist nicht gesichert, ob das Messinstrument für eine Zielgruppe funktionell eingeschränkter, multimorbider Patienten im stationären Setting geeignet ist. So wurden in die Entwicklungsstudien lediglich Probanden mit selbstständiger Haushaltsführung bzw. depressive Probanden miteinbezogen. Dass sich

der Fragebogen eher an funktionell gesunde, nicht institutionalisierte ältere Menschen richtet, zeigt sich auch implizit an der Formulierung der Items.

Empirische Befunde zur Effektivität der GDS15 im stationären Setting bzw. bei multimorbiden Patienten sind widersprüchlich; bei zahlreichen Stichproben multimorbider stationärer bzw. institutionalisierter Patienten war die Treffsicherheit der GDS15 zur Ermittlung krankheitswertiger depressiver Störungen jedoch bei weitem nicht so hoch wie bei Stichproben körperlich weitgehend gesunder Probanden.

Kürzere Versionen der GDS mit 10, 5 oder 4 Items haben in bisherigen Validierungsstudien keine überzeugenden Qualitätsmerkmale erzielt.

### 3. Fragestellung und Hypothesen

Ausgangsüberlegung der vorliegenden Arbeit ist es, die Diagnostik depressiver Störungen in geriatrischen Einrichtungen durch die Konstruktion eines neuen Screeningverfahrens zu verbessern, das in Aspekten der Praktikabilität weitgehend mit der GDS15 übereinstimmt, inhaltlich jedoch besser auf die besonderen Aspekte der Zielgruppe zugeschnitten ist. So enthält die GDS15 einige Items, die unter rationalen Gesichtspunkten wenig aussagekräftig für das Vorliegen einer depressiven Erkrankung zu sein scheinen:

Item 02: Haben Sie viele Aktivitäten und Interessen aufgegeben?

Item 04: ist Ihnen oft langweilig?

Item 09: Bleiben Sie lieber zu Hause anstatt auszugehen und Neues zu erleben?

Item 13: Fühlen Sie sich voller Energie?

Diese Fragen beziehen sich zwar inhaltlich auf Hauptsymptome depressiver Erkrankungen (Verlust von Freude/ Interessenlosigkeit, Verminderung des Antriebs sowie erhöhte Ermüdbarkeit), sie gehen jedoch implizit von Umständen aus, bei denen Aktivitäten, Interessen und die Energie eines Befragten nicht aus anderen Gründen beeinträchtigt sein können. In einem stationären Erhebungssetting mit multimorbiden Patienten muss jedoch fast davon ausgegangen werden, dass nahezu alle Probanden die entsprechenden Items lösen, da ihre Energie durch die Symptome zahlreicher körperlicher Erkrankungen beeinträchtigt ist, oder weil es Ihnen aufgrund langer Krankenhausaufenthalte nicht möglich ist, ihren üblichen Aktivitäten nachzugehen.

Bei der Entwicklung der Depression im Alter-Skala (DIA-S) wurden die oben genannten Aspekte für die inhaltliche Gestaltung der Items besonders berücksichtigt (vgl. Abschnitt 4.1). Die Hypothesen dieser Arbeit beziehen sich entsprechend auf den empirischen Vergleich der diagnostischen Qualität der beiden Verfahren anhand der Gütekriterien der klassischen Testtheorie. Dabei wird sowohl auf Itemebene als auch auf Skalenebene erwartet, dass die DIA-S für die Population stationärer geriatrischer Patienten bessere Ergebnisse erzielt. Dabei werden folgende Effekte erwartet:

### *Hypothese 01*

Die Items der DIA-S weisen bessere Kennwerte in Schwierigkeit, Effektivität und Trennschärfe auf als die Items der GDS15. Dabei sind besonders für die Items 02, 04, 09 und 13 der GDS15 geringe Schwierigkeiten sowie geringe Übereinstimmungen mit dem Außenkriterium zu erwarten.

### *Hypothese 02*

Die genauere Berücksichtigung inhaltlicher Aspekte bei der Itemgestaltung führt zu einer höheren Messgenauigkeit der DIA-S im Vergleich zur GDS15 i.S. einer höheren Internen Konsistenz.

### *Hypothese 03*

Die DIA-S weist eine höhere Validität auf als die GDS15, wobei in folgenden Aspekten eine Verbesserung zu erwarten ist:

#### *Hypothese 03a*

Die DIA-S weist insgesamt eine höhere Übereinstimmung mit dem Außenkriterium i.S. einer höheren Korrelation mit der Depressivität des Probanden auf.

#### *Hypothese 03b*

Durch die neue Skala wird eine signifikante Verbesserung bei der Klassifikation der Patienten in depressiv bzw. nicht depressiv erzielt. Die Verbesserung der Testqualität ist dabei auf eine signifikant höhere Spezifität bei vergleichbarer oder höherer Sensitivität zurückzuführen.

#### *Hypothese 03c*

Die Effektivität der DIA-S wird weniger durch Eigenschaften des Probanden oder des Erhebungssettings beeinflusst als die Effektivität der GDS15.

## 4. Methoden

### 4.1 Die Entwicklung der Depression- im -Alter- Skala (DIA-S)

#### 4.1.1 Zielsetzung

Die DIA-S wurde als Selbstbeurteilungsverfahren zum Screening depressiver Störungen konzipiert, das alternativ zur GDS in der Geriatrie eingesetzt werden kann. Die Skala soll dabei einerseits an der aktuellen Definition depressiver Störungen durch die WHO ausgerichtet sein, andererseits sollen die Anforderungen des klinischen Alltags geriatrischer Einrichtungen berücksichtigt werden. D.h., Durchführung und Auswertung sollen wenig Zeit in Anspruch nehmen und nur geringe Anforderungen an Untersucher und Patienten stellen, die Testergebnisse sollen wenig durch das potentielle Vorliegen körperlicher oder kognitiver Einschränkungen beeinflusst werden, und das Verfahren soll in verschiedenen Erhebungskontexten gleichermaßen eingesetzt werden können. Im Hinblick auf die ersten beiden Aspekte hat sich die GDS15 durch ihre äußere Beschaffenheit (geringe Anzahl an vollstandardisierten Items, einfaches, vollstandardisiertes Antwortschema) als praktikabel erwiesen und dient somit als Vorbild für das neue Instrument, das jedoch inhaltlich besser auf die Besonderheiten der Zielgruppe zugeschnitten sein soll (vgl. Zank, Schacke, Opterbeck & Heidenblut, 2008).

#### 4.1.2 Vorgehen

Abbildung 12 gibt einen Überblick über die einzelnen Schritte der Skalenentwicklung. Zunächst wurde unter Berücksichtigung der o.g. Kriterien ein Itempool mit 79 Selbstaussagen zu depressiven Symptomen zusammengestellt. Dieser Itempool wurde im nächsten Schritt im Rahmen einer Expertendiskussion kritisch geprüft, wobei besonders auf die Eignung der Aussagen für multimorbide Patienten in geriatrischen Erhebungskontexten geachtet wurde. Dabei wurden die 31 Items mit der größten Inhaltsvalidität zur weiteren empirischen Erprobung ausgewählt. Das Expertenteam setzte sich aus klinisch-diagnostisch tätigen

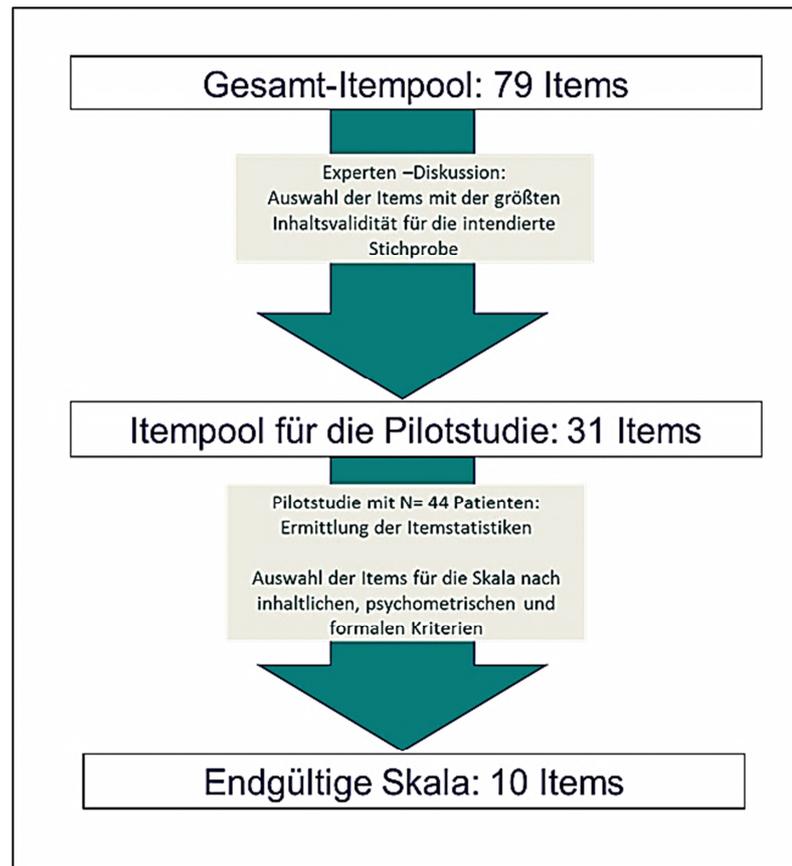


Abbildung 12: Schritte der Skalenentwicklung im Überblick

Psychologen mit jahrelanger Berufserfahrung in der stationären und teilstationären Geriatrie zusammen.

In einer Pilotstudie mit  $N= 44$  Studienteilnehmern wurden die ausgewählten Items schließlich geriatrischen und gerontopsychiatrischer Patienten vorgelegt, von denen 52% ( $N=23$ ) zum Zeitpunkt der Befragung unter einer depressiven Störung litten. Die so ermittelten psychometrischen Eigenschaften der Items bildeten neben den formalen und inhaltlichen Überlegungen zur Gesamtbeschaffenheit des Messinstrumentes die Entscheidungsgrundlage für die Zusammenstellung der aus 10 Items bestehenden Depression-im-Alter-Skala (DIA-S).

### *4.1.3 Zugrundeliegende Testtheorie*

Innerhalb der psychologischen Forschungstradition gibt es verschiedene Ansätze, psychometrische Messverfahren theoretisch zu fundieren (Amelang & Schmidt-Atzert, 2006). Dabei sind mehr oder weniger alle aktuell gebräuchlichen Ansätze in Auseinandersetzung mit der so genannten klassischen Testtheorie (KTT) entstanden, deren Grundlage bereits Mitte des letzten Jahrhunderts gelegt wurde (Gullikson, 1950), und deren Axiome seitdem weiterentwickelt und präzisiert wurden (Amelang & Schmidt-Atzert, 2006). Lassen sich die Axiome dieses Theoriegebäudes, das auf Überlegungen über den Anteil des Messfehlers an einem empirisch gewonnenen Testwert beruht, unter theoretischen Gesichtspunkten relativ leicht kritisieren (Fisseni, 2004), so ist dieser Ansatz in der Forschungspraxis jedoch auch heute noch der gebräuchlichste. Die theoretisch wesentlich aufwendigeren Verfahren der probabilistischen Testtheorie sowie die Ansätze zur kriteriumsorientierten Leistungsmessung können zwar nach der KTT konstruierte Testverfahren hinsichtlich der Messgenauigkeit und Gültigkeit geringfügig übertreffen; der so entstandene Nutzen steht jedoch in keinem Verhältnis zu dem erheblich größeren und riskanteren Konstruktionsaufwand (Amelang & Schmidt-Atzert, 2006). Zudem ist fraglich, ob bei diesen Verfahren tatsächlich von eigenen Ansätzen gesprochen werden kann oder ob sie nicht vielmehr theoretische Erweiterungen der KTT darstellen, welche nur unter speziellen Bedingungen berücksichtigt werden sollten.

In Anbetracht der obigen Überlegungen und im Sinne eines eher forschungspragmatischen Ansatzes orientiert sich die Testkonstruktion für die vorliegende Untersuchung in erster Linie an den Konstruktionsprinzipien der KTT.

### *4.1.4 Erstellung des Itempools*

#### *4.1.4.1 Formale Itemgestaltung.*

Da das Erhebungsinstrument wenig Anforderungen an Untersucher und Patienten stellen soll, wurde bei der formalen Gestaltung besonders auf leichte Verständlichkeit geachtet. Die Items wurden als einfache Selbstaussagen zu depressiven Symptomen formuliert (z.B. „ Ich fühle

mich bedrückt“), die vom Probanden jeweils als zutreffend oder nicht zutreffend eingeschätzt werden können. Die Beantwortung erfolgt mithilfe eines vollstandardisierten Ja/Nein-Antwortschemas.

Bei der sprachlichen Itemgestaltung wurden besonders die von Porst (2000) zusammengestellten allgemeinen Kriterien zur Fragebogengestaltung berücksichtigt, die hier für die Gestaltung von Fragen formuliert sind, sich aber ebenso gut auf die Gestaltung von Aussage Items beziehen lassen:

1. . . . Einfache, unzweideutige Begriffe verwenden, die von allen Befragten in gleicher Weise verstanden werden!
2. . . . Lange und komplexe Fragen vermeiden!
3. . . . Hypothetische Fragen vermeiden!
4. . . . Doppelte Stimuli und Verneinungen vermeiden!
5. . . . Unterstellungen und suggestive Fragen vermeiden!
6. . . . Fragen vermeiden, die auf Informationen abzielen, über die viele Befragte mutmaßlich nicht verfügen!
7. . . . Fragen mit eindeutigem zeitlichen Bezug verwenden
8. . . . Antwortkategorien verwenden, die eindeutig und disjunkt sind
9. . . . Der Kontext einer Frage [sollte] sich nicht auf deren Beantwortung [auswirken] . . . !
10. . . . Unklare Begriffe definieren!

Porst (2000, S.2)

#### *4.1.4.2 Inhaltliche Itemgestaltung*

Inhaltlicher Ausgangspunkt zur Formulierung der Items waren die in den klinisch-diagnostischen Leitlinien der WHO benannten Haupt – und Nebensymptome depressiver Störungen (Siehe Abschnitt 2.2.2).

Dabei erfordern die Eigenschaften der Zielpopulation im Hinblick auf einige der in der ICD-10 aufgeführten Symptome eine besondere Berücksichtigung. So können Beschwerden wie erhöhte Ermüdbarkeit, verminderte Konzentration und Aufmerksamkeit, Schlafstörungen und verminderter Appetit bei geriatrischen Patienten durch zahlreiche Begleiterscheinungen der Multimorbidität ausgelöst werden und sind nicht spezifisch für das Vorliegen einer

depressiven Störung. Eine standardisierte Abfrage dieser Symptome, bei der die Ursache des Symptoms nicht abgeklärt werden kann, ist daher nur bedingt sinnvoll.

*Tabelle 14: Itempool für die Pilotstudie (N=31) – systematisch nach Symptomen geordnet*

<b>Gedrückte Stimmung</b>
Ich bin grundsätzlich mit meinem Leben zufrieden
Ich bin ohne Grund gereizt
Ich fühle mich bedrückt
Ich fühle mich leicht gekränkt
Ich fühle mich innerlich wie versteinert
<b>Verlust von Freude/ Interessenlosigkeit</b>
Mich interessiert, was um mich herum vorgeht
In meinem Alltag gibt es etwas, das mir Freude macht
Ich kann mein Leben genießen (auch wenn mir manches schwerer fällt)
Ich bin gerne am Leben
<b>Verminderung des Antriebs und erhöhte Ermüdbarkeit</b>
Es fällt mir schwer, mich aufzuraffen
Ich möchte nur noch meine Ruhe haben
Das Leben ist nur noch anstrengend für mich
<b>Grübeltendenz (statt Verminderter Konzentration und Aufmerksamkeit)</b>
Ich muss viel grübeln
Mir gehen immer die gleichen Gedanken im Kopf herum
<b>Vermindertes Selbstwertgefühl und Selbstvertrauen</b>
Ich habe Angst, dass ich etwas falsches sagen oder tun könnte
Ich fühle mich durch Schwierigkeiten leicht überfordert
Ich fühle mich meinen Problemen gewachsen
Es macht mir sehr zu schaffen, wenn etwas schief geht
Ich traue mir nichts mehr zu
Ich fühle mich hilflos
<b>Schuldgefühle und Gefühle der Wertlosigkeit</b>
Ich bin für andere Menschen eine Last
Mein Leben kommt mir sinnlos vor
Ich fühle mich wertlos
<b>Negative und pessimistische Zukunftsperspektiven</b>
Ich habe Angst vor der Zukunft
Ich kann mir nicht vorstellen, dass es mir je wieder besser gehen wird
Ich denke mit Zuversicht an die Zukunft
Ich glaube, dass meine Situation hoffnungslos ist
<b>Suizidalität</b>
Ich wäre froh, wenn alles vorbei wäre
Ich wünschte, ich könnte einfach einschlafen und nicht wieder aufwachen
<b>Psychomotorik/Fähigkeit zu entspannen (statt Schlafstörungen)</b>
Ich kann mich gut entspannen
Ich bin die ganze Zeit innerlich angespannt

So wurden zur Berücksichtigung der Symptome Antriebsminderung, Schlafstörung und verminderte Konzentration und Aufmerksamkeit für den Itempool Aussagen konstruiert, die über das eigentliche Symptom hinausgehen und stärker auf die psychische Dimension der Beschwerde abzielen. Es werden daher nicht Schlafstörungen im engeren Sinne abgefragt, sondern die selbst eingeschätzte Fähigkeit, entspannen zu können. Die Items, die sich auf die Antriebsminderung beziehen, fokussieren weniger auf ein körperliches Erschöpfungsgefühl als auf die Motivation, sich zu etwas aufzuraffen zu können. Und anstatt von

Konzentrationsproblemen wird eine allgemeine Grübeltendenz als initiales kognitives Phänomen berücksichtigt. Auf Items, die sich auf verminderten Appetit beziehen, wurde gänzlich verzichtet, da eine entsprechende Differenzierung für dieses Symptom durch eine einfache Selbstaussage nicht gewährleistet werden konnte.

Tabelle 14 zeigt eine Auflistung der 31 Items, die auf der Grundlage von Experteneinschätzungen von in der Geriatrie tätigen Neuropsychologen als inhaltssvalide und für die Zielpopulation geeignet aus dem ursprünglichen Itempool zur empirischen Erprobung ausgewählt wurden.

#### *4.1.5 Pilotstudie: empirische Prüfung des Itempools*

##### *4.1.5.1 Rekrutierung und Untersuchungsablauf*

Die orientierende empirische Einschätzung des Itempools erfolgte an einer Stichprobe von Patienten mit bereits diagnostiziertem Erkrankungsstatus, wobei für die Teilstichprobe depressiver Patienten eine möglichst ausgewogene Bandbreite verschiedener Erkrankungsschweregrade angestrebt wurde. Die Rekrutierung der Teilnehmer erfolgte im Juni 2007 in stationären und teilstationären Bereichen einer geriatrischen Klinik, dem evangelischen Geriatriezentrum Berlin, sowie an zwei gerontopsychiatrischer Einrichtungen, dem gerontopsychiatrischen Zentrum des St. Hedwigkrankenhauses sowie der gerontopsychiatrischen Abteilung des Krankenhauses Hedwigshöhe.

Die Auswahl geeigneter Patienten erfolgte durch die Kontaktpersonen der jeweiligen Klinik, wobei der Erkrankungsstatus der Teilnehmer der Untersucherin erst nach der Befragung bekannt gegeben wurde. Die Teilnehmer wurden über das Ziel der Studie aufgeklärt und schriftlich um Ihre Einwilligung gebeten.

Die Items wurden den Teilnehmern von der Untersucherin mündlich vorgegeben, wobei jeweils der genaue Wortlaut der Aussagen eingehalten wurde. Es erfolgten keine Begriffserklärungen bei der Untersuchung, bei akustischen Verständnisschwierigkeiten wurde das entsprechende Item erneut vorgelesen. Die Teilnehmer wurden gebeten, die Items jeweils

mit „ja“ oder „nein“ zu beantworten und dabei ihre Stimmung der letzten zwei Wochen zugrunde zu legen.

#### 4.1.5.2 Stichprobe

Die Stichprobe der Voruntersuchung (N = 44 Probanden) besteht zu etwa gleichen Teilen aus depressiven (N=23) und nicht depressiven Teilnehmern (N=21).

*Tabelle 15 : Verteilung der Erkrankungsschweregrade der Stichprobe nach Einrichtungsform*

Erkrankungsform	Patienten (Anzahl (Prozent))				Gesamt
	Gerontopsychiatrie		Geriatric		
Depressionsdiagnose	stationär	teilstationär	stationär	teilstationär	
<b>nicht vergeben</b>	-	-	<b>21 (100%)</b>	-	<b>21 (48%)</b>
<b>Vergeben</b>	<b>9 (39%)</b>	<b>3 (13%)</b>	<b>7 (30%)</b>	<b>4 (17%)</b>	<b>23 (52%)</b>
Anpassungsstörung	-	1	1	3	5 (22%)
leichte depressive Episode	2	-	4	1	7 (30%)
mittelgrad. depressive Episode	4	2	1	-	7 (30%)
schwere depressive Episode	2	-	-	-	2 (9%)
atypische Depression	-	-	1	-	1 (4%)
n.n.bez. Depression	1	-	-	-	1 (4%)
<b>Gesamt</b>					<b>44 (100%)</b>

Tabelle 15 gibt einen Überblick über die Zusammensetzung der Stichprobe nach Einrichtungsart sowie über die Verteilung der Krankheitsbilder. Durch die Einbeziehung gerontopsychiatrischer Patienten sind bei den depressiven Probanden mittelgradige und schwere Erkrankungsformen Patienten tendenziell häufiger vertreten als es in einer rein geriatricen Stichprobe zu erwarten wäre. Hinsichtlich der weiteren Stichprobenmerkmale (siehe Tabelle 16) zeigt sich ein dem Durchschnittsalter der Stichprobe (76.6 Jahre) entsprechender, überwiegender Anteil an weiblichen Probanden (77%). 27% der Gesamtstichprobe zeigen eine leichte oder mittelgradige kognitive Einschränkung. Patienten mit schwerer kognitiver Funktionseinschränkung (MMSE<16 Punkte) wurden nicht in die Erhebung einbezogen, da für dieses Klientel Selbstbeurteilungsverfahren nicht mehr valide angewendet werden können.

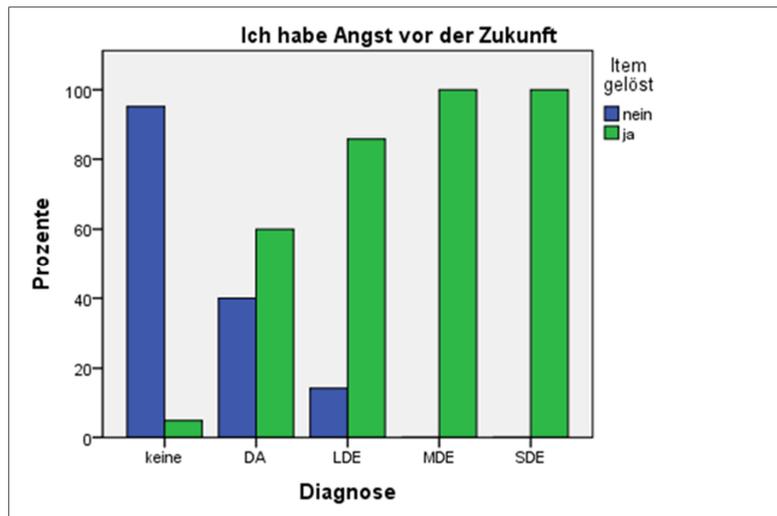
*Tabelle 16: weitere Stichprobenmerkmale*

Stichprobenmerkmale	Probanden (Anzahl (Prozent)/ MW (SD))		
	Depressionsdiagnose		
	nicht vergeben	vergeben	Gesamt
<b>Geschlecht</b>			
Männlich	4 (19%)	6 (26%)	10 (23%)
Weiblich	17 (81%)	17 (74%)	34 (77%)
<b>Alter</b>	79.7 (7.0)	73.8 (9.0)	76.6 (8.6)
<b>Kognition</b>			
Einschränkung	7 (33%)	5 (21%)	12 (27%)
keine Einschränkung	14 (67%)	18 (78%)	32 (73%)

#### 4.1.5.3 Itemanalysen

Zur Erprobung der Items wurden Schwierigkeit und Validität der Items bezogen auf das Außenkriterium (Depressionsdiagnose vergeben vs. Depressionsdiagnose nicht vergeben) berechnet (siehe Tabelle 17). Die Validität der Items wurde dabei nochmals genauer als Itemsensitivität, die Eigenschaft des Items korrekt auf depressive Erkrankungen zu reagieren, und Itemspezifität, die Eigenschaft des Items, korrekt auf nicht gegebene Erkrankungen zu reagieren, aufgeschlüsselt. Auf die Ermittlung der Trennschärfe als internes Qualitätskriterium wurde in der Voruntersuchung aufgrund der unbekanntenen Eigenschaften der Gesamtskala verzichtet.

Darüber hinaus wurden die Lösungsmuster der Items über Patientengruppen mit verschiedenem Erkrankungsschweregrad erstellt und bei der Auswahl von Aussagen mit geringeren Sensitivitätswerten als zusätzliches Qualitätsmerkmal berücksichtigt. Nach diesem Kriterium zur Itemselektion sollte die Lösungswahrscheinlichkeit eines Items mit zunehmender Merkmalsausprägung monoton ansteigen, d.h. von Patienten mit schwerem Erkrankungsgrad sollte das Item häufiger gelöst werden, als von Patienten mit geringer Symptomatik (siehe Abbildung 13 als Beispiel für einen monoton ansteigenden Lösungsverlauf).



Anmerkungen. Keine = keine depressive Erkrankung; DA = depressive Anpassungsstörung; LDE = leichte depressive Episode; MDE = mittelgradige depressive Episode; SDE = schwere depressive Episode;

*Abbildung 13: monoton verlaufendes Lösungsmuster des Items „Ich habe Angst vor der Zukunft“*

Die Schätzung der Merkmalsausprägung der Probanden wird bei diesem Verfahren in der Regel anhand der Beantwortung des Gesamtitempools vorgenommen und orientiert sich somit eher an der Trennschärfe des Items (Lienert & Raatz, 1989); für die vorliegenden Itemvergleiche erfolgte die Unterteilung der Stichprobe in verschiedene Merkmalsbereiche dagegen nach dem Außenkriterium, d.h. nach dem durch die Diagnose eingeschätzten Erkrankungsschweregrad, sodass die Lösungsmuster hier weitere Rückschlüsse auf die Validität des Items zulassen. Aufgrund der geringen Anzahl an Probanden pro Zelle wurden die Lösungsmuster für die vorliegende Untersuchung eher qualitativ betrachtet und nur in wenigen Fällen als zusätzliches Entscheidungskriterium berücksichtigt.

*Tabelle 17: Itemkennwerte der Voruntersuchung, nach Symptombereichen sortiert*

Item	Polung	Validität	Sensitivität	Spezifität	Schwierigkeit
<b>Gedrückte Stimmung</b>					
Ich fühle mich bedrückt	positiv	.814	.773	.857	.465
Ich bin grundsätzlich mit meinem Leben zufrieden	negativ	.727	.652	.810	.432
Ich fühle mich innerlich wie versteinert	negativ	.667	.333	1.00	.167
Ich fühle mich leicht gekränkt	negativ	.651	.364	.950	.209
Ich bin ohne Grund gereizt	negativ	.569	.217	.952	.136
<b>Verlust von Freude/ Interessenlosigkeit</b>					
Ich kann mein Leben genießen (auch wenn mir manches schwerer fällt)	negativ	.698	.544	.857	.349
In meinem Alltag gibt es etwas, das mir Freude macht	positiv	.651	.500	.810	.349
Mich interessiert, was um mich herum vorgeht	negativ	.614	.348	.905	.227
Ich bin gerne am Leben	negativ	.643	.333	.952	.190
<b>Verminderung des Antriebs und erhöhte Ermüdbarkeit</b>					
Es fällt mir schwer mich aufzuraffen	positiv	.744	.727	.762	.488
Das Leben ist nur noch anstrengend für mich	positiv	.721	.591	.857	.372
Ich möchte nur noch meine Ruhe haben	negativ	.558	.318	.810	.256
<b>Grübeltendenz</b>					
Ich muss viel grübeln	positiv	.886	.957	.810	.591
Mir gehen immer die gleiche Gedanken im Kopf herum	positiv	.791	.818	.762	.535

Fortsetzung Tabelle 17

Item	Polung	Validität	Sensitivität	Spezifität	Schwierigkeit
<b>Vermindertes Selbstwertgefühl und Selbstvertrauen</b>					
Ich fühle mich durch Schwierigkeiten leicht überfordert	positiv	.907	.909	.905	.512
Ich fühle mich meinen Problemen gewachsen	negativ	.837	.773	.905	.442
Ich habe Angst, dass ich etwas falsches sagen oder tun könnte	positiv	.773	.652	.905	.386
Es macht mit sehr zu schaffen, wenn etwas schief geht	negativ	.697	.955	.429	.767
Ich fühle mich hilflos	negativ	.659	.652	.660	.500
Ich traue mir nichts mehr zu	negativ	.604	.227	1.00	.116
<b>Schuldgefühle und Gefühle der Wertlosigkeit</b>					
Ich bin für andere Menschen eine Last	positiv	.772	.565	1.00	.295
Mein Leben kommt mir sinnlos vor	positiv	.767	.591	.952	.326
Ich fühle mich wertlos	positiv	.714	.429	1.00	.214
<b>Negative und pessimistische Zukunftsperspektiven</b>					
Ich habe Angst vor der Zukunft	positiv	.932	.870	.952	.477
Ich glaube, dass meine Situation hoffnungslos ist	negativ	.674	.364	1.00	.186
Ich denke mit Zuversicht an die Zukunft	positiv	.651	.500	.810	.349
Ich kann mir nicht vorstellen, dass es mir je wieder besser gehen wird	negativ	.628	.273	1.00	.140
<b>Psychomotorik /Fähigkeit zu Entspannen (statt Schlafstörungen)</b>					
Ich kann mich gut entspannen	negativ	.884	.818	.952	.442
Ich bin die ganze Zeit innerlich angespannt	positiv	.767	.773	.762	.512
<b>Suizidalität</b>					
Ich wäre froh, wenn alles vorbei wäre	negativ	.674	.545	.810	.372
Ich wünschte, ich könnte einfach einschlafen und nicht wieder aufwachen	negativ	.651	.364	.952	.209

Anmerkung: Items, die in das Erhebungsinstrument aufgenommen wurden, sind durch grauen Hintergrund markiert.

#### 4.1.6 Zusammenstellung der Skala

##### 4.1.6.1 Angestrebte Anzahl an Items

Da die Ökonomie eines Verfahrens im klinischen Alltag gerade bei Screeninginstrumenten eine große Rolle spielt und es zahlreiche Bemühungen gibt, die Geriatrische Depressionsskala für die Praxis auf Versionen mit 10 (Shah et al., 1996), 5 (Kenny Weeks et al., 2003) oder 4 (Goring et al., 2004) Items weiter zu verkürzen, wurde für die DIA-S eine etwas geringere Anzahl als 15 Items angestrebt. Um trotzdem eine hohe inhaltliche Validität der Skala zu gewährleisten wurden für die Itemselektion rationale und externale Strategien kombiniert.

##### 4.1.6.2 Auswahl der Items

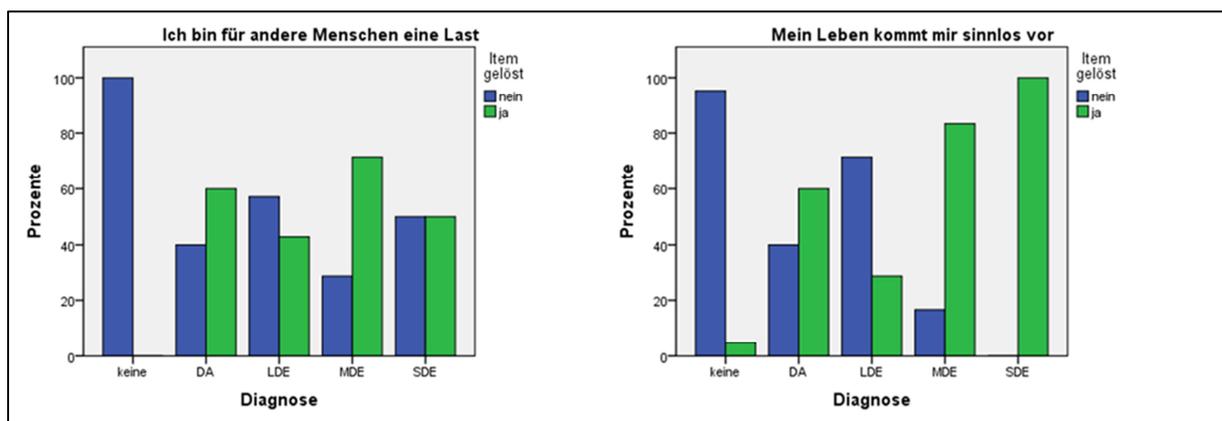
Ausgangsüberlegung zur Auswahl der Items war es, die nach ICD-10 beschriebenen Symptome depressiver Erkrankungen durch die Skala unter Berücksichtigung der entsprechenden Modifikationen für die Zielgruppe so weit als möglich abzubilden. Dabei wurde ein leichter Schwerpunkt auf die affektiven und kognitiven Aspekte depressiver Verstimmung gelegt, da diese sich sowohl in einer Voruntersuchung zu dieser Studie (Heidenblut & Zank, 2005) als auch bei faktorenanalytischen Untersuchungen zur GDS30 (Parmalieu, Lawton & Katz, 1989, Adams, Mattro & Sanders, 2004) für die richtige Klassifizierung von Patienten als besonders effektiv erwiesen haben.

Konkret wurde angestrebt, aus jedem der Symptombereiche ein bzw. aus den Schwerpunkt-Symptombereichen zwei Items in die Skala zu integrieren. Von den inhaltlich verwandten Aussagen wurden dabei diejenigen bevorzugt, die die besten empirischen Kennwerte aufwiesen, d.h. die bei hohen Itemvaliditäten einen mittleren Schwierigkeitsgrad aufwiesen. In einigen Ausnahmefällen, die im Folgenden kurz aufgeführt werden, wurden darüber hinaus weitere Kriterien zu Auswahl berücksichtigt.

Aus dem Symptombereich „Vermindertes Selbstwertgefühl und Selbstvertrauen“ wurde das Item „Ich fühle mich meinen Problemen gewachsen“ trotz guter empirischer Kennwerte nicht in die Skala integriert, um inhaltliche Redundanz mit der Aussage „Ich fühle

mich durch Schwierigkeiten leicht überfordert“ zu vermeiden. Stattdessen wurde mit „Ich habe Angst, dass ich etwas falsches sagen oder tun könnte“ ein Item ausgewählt, das noch einen weiteren Aspekt der Symptomdimension berücksichtigt.

Aus dem Bereich „Schuldgefühle und Gefühle der Wertlosigkeit“ wurde anstatt des Items „Ich bin für andere Menschen eine Last“ das Item „Mein Leben kommt mir sinnlos vor“ für die Skala ausgewählt, da dieses bei geringerer Schwierigkeit ein insgesamt schlüssigeres Lösungsmuster aufweist (siehe Abbildung 14).



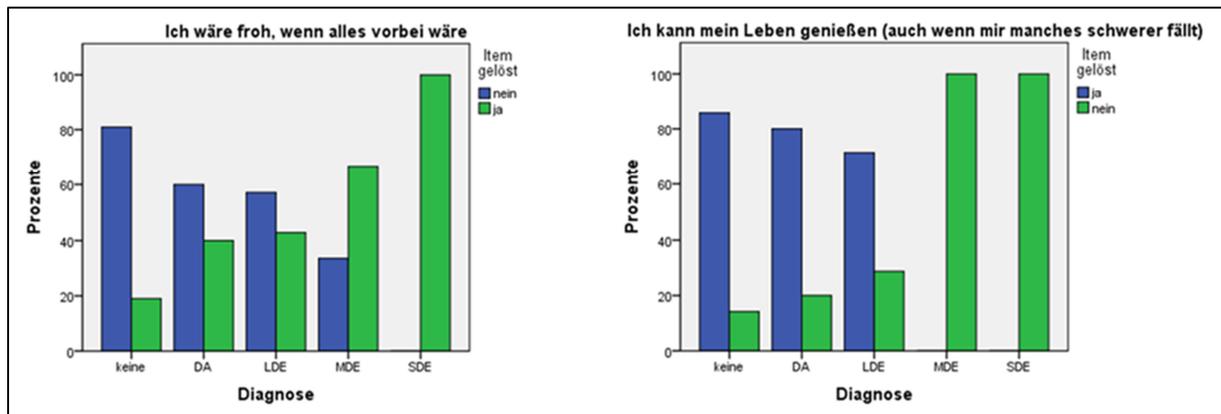
Anmerkungen. Keine = keine depressive Erkrankung; DA = depressive Anpassungsstörung; LDE = leichte depressive Episode; MDE = mittelgradige depressive Episode; SDE = schwere depressive Episode;

Abbildung 14: Vergleich der Lösungsmuster kritischer Items (1)

So ergibt sich bei der Lösungsverteilung des Items der Eindruck, als sei die Beantwortung durch die depressiven Probanden nicht in erster Linie vom Schweregrad der Symptomatik abhängig, sondern auch durch andere Faktoren wie möglicherweise soziale Erwünschtheit bestimmt.

Aus dem Bereich Suizidalität wurde kein Item in die Skala integriert. So weist das Item „Ich wäre froh wenn alles vorbei wäre“ zwar bei grenzwertigen psychometrischen Kennwerten ein monoton ansteigendes Lösungsmuster auf, es ergeben sich aber gerade bei den Probanden mit leichterer Symptomatik zu hohe Überschneidungen in den Antworttendenzen (siehe Abbildung 8). Zudem weist das Item zwar, wie aufgrund des Inhalts zu erwarten ist, eine verhältnismäßig hohe Schwierigkeit auf, es wird jedoch trotzdem von

einem Anteil von Probanden ohne depressive Symptomatik gelöst. Inhaltlich ergab sich hier bei der Erhebung der Eindruck, dass das Item von einem Teil der Probanden anders verstanden wurde als es gemeint war. So kommentierten einige Probanden das Item so, als würden sie unter „alles vorbei“ nicht ihr gesamtes Leben sondern ihre gegenwärtige Krankenhaussituation verstehen.



Anmerkungen. Keine = keine depressive Erkrankung; DA = depressive Anpassungsstörung; LDE = leichte depressive Episode; MDE = mittelgradige depressive Episode; SDE = schwere depressive Episode;

Abbildung 15: Vergleich der Lösungsmuster kritischer Items (2)

Das Item „Ich kann mein Leben genießen (auch wenn mir manches schwerer fällt)“ aus dem Bereich Verlust von Freude/ Interessenlosigkeit wurde dagegen in die Skala integriert. So weist zwar auch dieses Item eine verhältnismäßig geringe Itemsensitivität auf, erfasst aber zuverlässig Probanden mit ausgeprägter Symptomatik.

#### 4.1.6.3 Polung der Items

Formal wurde ein Anteil von etwa 1/3 der Aussagen angestrebt, die durch eine „Nein“-Antwort gelöst werden, also eine negative Polung aufweisen. Eine entsprechende Zusammensetzung wurde durch die in Tabelle 18 genannte Auswahl mit 3 von 10 Items erreicht.

#### 4.1.6.4 Anordnung der Skala

Für das Erhebungsinstrument wurden die Aussagen so kombiniert, dass das Verfahren einerseits auf die Patienten möglichst transparent wirkt und andererseits durch Formulierungen bedingte Antwortmuster weitgehend kontrolliert werden. Als Einstiegssitem wurde die Aussage „Ich fühle mich bedrückt“ gewählt, da durch diese Aussage die Intention des Fragebogens, die Stimmung des Patienten abzufragen, besonders deutlich wird. Um eine Ja-Sage-Tendenz zu kontrollieren, wechseln sich positiv und negativ gepolte Items so ab, dass nicht mehr als 3 Items gleicher Polung aufeinander folgen. Außerdem wurde in der Zusammenstellung darauf geachtet, dass sich aufeinander folgende Aussagen möglichst semantisch voneinander unterscheiden. Tabelle 18 zeigt die Skala, wie sie in der Hauptuntersuchung als Teil des Erhebungsinstrumentes verwendet wird.

*Tabelle 18 : Erhebungsinstrument für die Validierungsstudie*

Item		Bewertung	
1	Ich fühle mich bedrückt	<b>Ja</b>	Nein
2	Ich habe Angst, dass ich etwas falsches sagen oder tun könnte	<b>Ja</b>	Nein
3	Ich kann mich gut entspannen	Ja	<b>Nein</b>
4	Mein Leben kommt mir sinnlos vor	<b>Ja</b>	Nein
5	Es fällt mir schwer, mich aufzuraffen	<b>Ja</b>	Nein
6	Ich habe Angst vor der Zukunft	<b>Ja</b>	Nein
7	Ich kann mein Leben genießen, auch wenn mir manches schwerer fällt	Ja	<b>Nein</b>
8	Ich fühle mich durch Schwierigkeiten leicht überfordert	<b>Ja</b>	Nein
9	Ich muss viel grübeln	<b>Ja</b>	Nein
10	Ich bin grundsätzlich mit meinem Leben zufrieden	Ja	<b>Nein</b>

Anmerkung. Die Bewertung, die das Item löst, ist jeweils fett gedruckt

## 4.2 Vergleichende Validierung der DIA-S und der GDS15

### 4.2.1 Rekrutierung

Für die Validierung der Skalen wurde eine Stichprobe von N = 331 stationären geriatrischer Patienten erhoben. Die Erhebung erfolgte von 2007 bis 2009 in 3 geriatrischen Abteilungen in Nordrheinwestfalen. Kooperierende Kliniken waren das St. Marien Hospital in Köln, das Herz-Jesu-Krankenhaus in Lindlar sowie das Krankenhaus St. Barbara in Attendorn. Die

Probanden wurden jeweils über die Mitarbeiter der Stations-Teams, die routinemäßig für die Erhebung der Geriatrischen Depressionsskala zuständig waren, rekrutiert. Es wurden zunächst Patienten aus den stationären und teilstationären Bereichen der Kliniken in die Erhebung eingeschlossen; aus Gründen größerer Stichprobenhomogenität wurden nach der ersten Testphase jedoch nur noch stationäre Patienten untersucht. Von der Teilnahme ausgeschlossen wurden psychotische, delirante und aphasische Patienten sowie Patienten mit einer schweren kognitiven Funktionseinschränkung (operationalisiert über einen MMSE-Wert  $\leq 15$ ). In der ersten Erhebungsphase wurden Probanden ungeachtet ihres Depressions-Status eingeschlossen, um einem Spektrum-Bias in den Daten vorzubeugen. Nach Erreichen einer Fallzahl von N=150 affektiv unauffälliger Patienten wurden die Kliniken gebeten, speziell Patienten zu rekrutieren, die nach klinischem Eindruck eine affektive Verstimmung aufwiesen, bis eine Anzahl von N = 151 depressiver Patienten erreicht war – aus Gründen der Verbindung der Interviewer sollten jedoch weiterhin auch wenige unauffällige Patienten vorgeschlagen werden – die Fallzahl der Gruppe ohne Diagnose erhöhte sich auf diese Weise auf N = 180. Als Kriterium für das Vorliegen einer krankheitswertigen depressiven Verstimmung diente der Trennwert der Montgomery-Asberg-Depressionsskala.

#### *4.2.2 Untersuchungsablauf*

Abbildung 16 gibt einen Überblick über den Untersuchungsablauf der Validierungsstudie.

Diejenigen Patienten, die die Einschlusskriterien erfüllten, wurden mithilfe von schriftlichem Informationsmaterial über die Bedeutung der Studie sowie über ihre diesbezüglichen Rechte (Freiwilligkeit, Vertraulichkeit, Wahrung der Anonymität) aufgeklärt und um die Unterschrift auf einer Einwilligungserklärung (siehe Anhang) gebeten.

Die Erhebungen erfolgten durch geschultes Forschungs- und Klinikpersonal. Die teilnehmenden Probanden beantworteten in ein bis zwei Untersuchungseinheiten jeweils mündlich die DIA-S sowie die GDS15; in einer weiteren Erhebungseinheit wurde von einem

unabhängigen, gegenüber den Testergebnissen verblindeten Untersucher ein halbstandardisiertes Interview als Grundlage für die Einschätzung der Symptomatik des

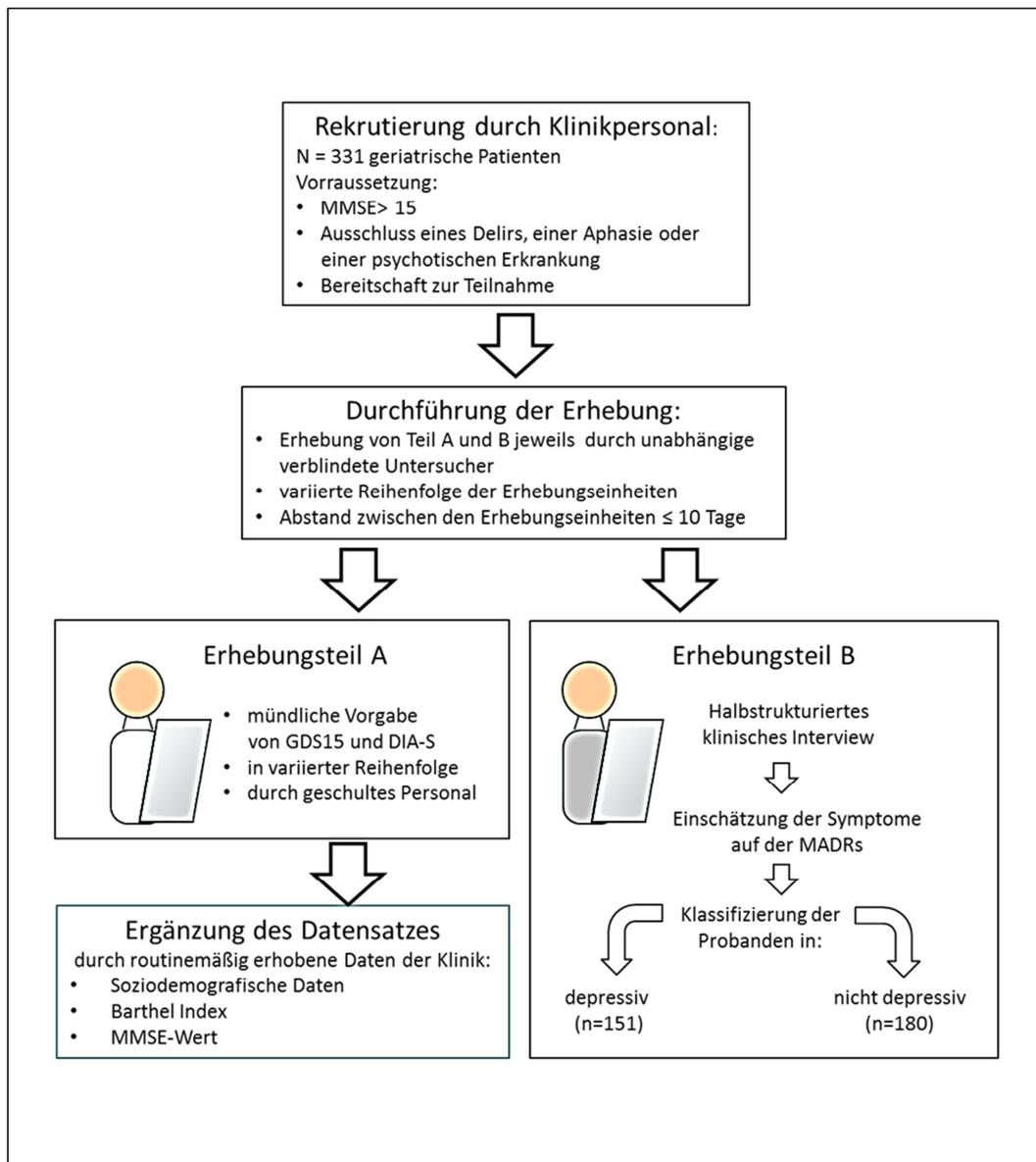


Abbildung 16: Untersuchungsablauf der Validierungsstudie

Patienten auf der MADRS durchgeführt. Die einzelnen Erhebungsschritte mussten aus praktischen Gründen teilweise an verschiedenen Tagen erfolgen, dabei wurden nur Datensätze in die Untersuchung eingeschlossen, bei denen eine Dauer von 10 Tagen zwischen den Untersuchungsteilen nicht überschritten wurde. Die Reihenfolge der Erhebungseinheiten wurde variiert.

Die Datensätze der teilnehmenden Patienten wurden durch von den Kliniken routinemäßig erhobene Variablen zu Soziodemographie (Alter, Geschlecht) sowie zu klinisch relevanten Merkmalen (Barthel-Index, MMSE-Wert) ergänzt.

#### 4.2.3 Messinstrumente

*DIA-S.* Eine ausführliche Beschreibung des neu entwickelten Messinstrumentes erfolgte in Abschnitt 4.1. Die DIA-S ist eine Selbstbeurteilungsskala mit 10 Aussagen zur Depressivität, die nach einem Ja-Nein-Schema beantwortet werden. Die theoretische Ausprägung liegt zwischen 0 (keine Symptomatik) und 10 (maximale Symptomatik) Punkten. Die psychometrischen Eigenschaften des Instrumentes werden im Ergebnisteil vorgestellt.

*MADR-Skala.* Die MADRS (Montgomery & Asberg, 1979; Neumann & Schulte, 1988) ist ein Fremdbeurteilungsverfahren zur Erfassung depressiver Verstimmungen bei Erwachsenen, deren Anwendung sich in der Population geriatrischer Patienten bewährt hat (Leentjens, Verhey, Lousberg, Spitsbergen & Wilmlink, 2000; Mottram, Wilson & Copeland, 2000). Die Skala misst die Ausprägung eines Patienten auf 10 Depressionssymptomen (sichtbare Traurigkeit, mitgeteilte Traurigkeit, innere Anspannung, reduzierter Schlaf, reduzierter Appetit, Konzentrationsschwierigkeiten, Antriebsmangel, Gefühl der Gefühllosigkeit, pessimistische Gedanken, Suizidgedanken), wobei der Schweregrad jeweils mithilfe einer 7-stufigen Lickert-Skala bestimmt wird. Eine Einstufung von „0“ entspricht dabei einer unauffälligen Stimmungslage, eine Einstufung von „6“ einer sehr starken Ausprägung des entsprechenden Symptoms. Die MADRS hat somit eine theoretische Ausprägung von 0-60 Punkten, wobei ein Testwert ab 13 Punkten einer leichten Depression, ein Testwert ab 22 Punkten einer mittelgradigen und ein Testwert ab 29 Punkten einer schweren Depression entspricht. Kriterium für das Vorliegen einer klinisch relevanten depressiven Störung ist für die vorliegende Untersuchung der erste Trennwert von 13 Punkten. Grundlage für die Einschätzung der Symptomatik ist ein diagnostisches Interview, das durch die Verwendung eines Leitfadens (siehe Anhang) teilstandardisiert wurde (Zank & Schacke, 2001).

*GDS15.* Die Kurzversion der Geriatrischen Depressionsskala (Yesavage et. al., 1983) in der Fassung nach AGAST (1997) ist hierzulande das gängige Screeningverfahren für geriatrische Patienten. Sie besteht aus 15 Items, die auf einem Ja/Nein- Antwortschema beantwortet und zu einem Summenscore zusammengefasst werden. Ein Ergebnis von 15 Punkten entspricht der maximalen Symptomatik, bei einem Score von über 5 Punkten wird von einer klinisch relevanten depressiven Verstimmung ausgegangen.

*Mini-Mental-State-Examination (MMSE).* Die kognitive Leistungsfähigkeit der Patienten wurde mithilfe des MMSE (Folstein et al., 1975) in der deutschen Bearbeitung nach AGAST (1997) erhoben. Der MMSE ist das international gebräuchlichste kognitive Screeningverfahren und wird von den kooperierenden Kliniken routinemäßig durchgeführt. Getestet werden örtliche und zeitliche Orientierung (Was für ein Wochentag ist heute?; In welcher Stadt befinden wir uns?) Konzentrationsfähigkeit und Gedächtnisleistung (Bitte merken sie sich: Zitrone, Schlüssel, Ball), sprachliche Leistungen (Bitte schreiben Sie einen Satz) sowie visuell-konstruktive und praktische Fähigkeiten (Bitte kopieren Sie die Zeichnung). Der Test besteht insgesamt aus 30 Items, wobei für jede korrekt durchgeführte Aufgabe ein Punkt vergeben wird. Liegt der Testwert bei 15 Punkten oder darunter, liegt eine schwere kognitive Funktionseinschränkung vor, sodass von der Durchführung von Selbstbeurteilungsverfahren abgesehen werden muss.

*Kognitiver Status.* In Einzelfällen, in denen der Proband für eine Studienteilnahme geeignet war, in denen aber aus Gründen des klinischen Ablaufs keine MMSE-Testung erfolgen konnte, wurde der Patient von den Klinik-Mitarbeitern grob hinsichtlich der kognitiven Leistungsfähigkeit nach der Ausprägung „unauffällig“, „leichte Störung“ oder „mittelgradige Störung“ eingeschätzt.

*Barthel-Index.* Der Barthel Index (Mahoney & Barthel, 1965) ist ein weiterer Bestandteil des Geriatrischen Basisassessments (AGAST, 1997). Er dient dazu, die funktionale Selbstständigkeit eines Patienten im Hinblick auf basale alltägliche Aktivitäten (Essen und

Trinken, Baden/Duschen, Körperpflege, An- und Ausziehen, Stuhlkontrolle, Harnkontrolle, Benutzung der Toilette, Bett- Stuhltransfer, Mobilität (selbstständiges Gehen/Fahren mit Rollstuhl), Treppe steigen) abzubilden. Die theoretische Ausprägung variiert in den einzelnen Bereichen von 5 bis zu 15 maximal erreichbaren Punkten. Der Gesamtwert des Barthel-Indexes kann Werte von 0 Punkten (keine Selbstständigkeit in basalen Aktivitäten) bis zu 100 Punkten (Vollständige Selbstständigkeit) annehmen. Der Barthel-Index wurde als Gesamtwert in die Variablenliste aufgenommen, da verschiedene Einflüsse des Funktionstatus der Probanden einerseits auf die Depressivität und andererseits auf das Antwortverhalten möglich sind.

*Saisonale Einflüsse.* Die Datenerhebung für die vorliegende Studie fand in einem Zeitraum von insgesamt über einem Jahr statt, sodass die einzelnen Testungen jeweils zu verschiedenen Jahreszeiten erfolgten. Um gegebenenfalls einen Einfluss saisonaler Bedingungen auf das Antwortverhalten der Probanden erfassen zu können, wie er beispielsweise bei De Craen et al. (2005) beschrieben wird, wurde der Zeitpunkt der Erhebung für einige der Analysen mitberücksichtigt. Der Einfluss saisonaler Bedingungen auf die Stimmung kann dabei theoretisch sehr komplex sein, lässt sich aber nur sehr bedingt empirisch erfassen. Entsprechende Analysen beschränken sich daher in der vorliegenden Studie auf die Erfassung der Tageslichtdauer zum Zeitpunkt des diagnostischen Interviews, operationalisiert über die Messung des Abstandes zwischen dem Erhebungsdatum und der Sommersonnenwende.

#### 4.2.4 Stichprobe

Tabelle 19 gibt einen Überblick über die Eigenschaften der Stichprobe. Erwartungsgemäß sind die weiblichen Teilnehmer mit 73% häufiger vertreten als die männlichen. Die Altersverteilung entspricht mit einem Mittelwert von 80.7 und einer Standardabweichung von 7.6 den aus epidemiologischen Befunden bekannten Werten in der stationären Akutgeriatrie (Renteln-Kruse & Ebert, 2003). Die Verteilung depressiver Störungen nach Geschlecht ist der Schichtung der Stichprobe nach Depressivität geschuldet und weist daher nicht die aus der

Literatur bekannten Proportionen auf. Kognitive Leistungsfähigkeit sowie alltagsrelevante Funktionalität der Patienten fallen etwas besser aus als bei entsprechenden Referenzstichproben (Renteln-Kruse & Ebert, 2003), da die Befragung ein Mindestmaß an kognitiver und physischer Belastbarkeit erfordert. Dabei liegen beide Mittelwerte für die nicht erkrankte Teilstichprobe höher, was mit den in der Literatur gut belegten Zusammenhängen zwischen depressiven Erkrankungen und Multimorbidität bzw. depressiven Erkrankungen und kognitiven Beeinträchtigungen vereinbar ist (Helmchen et al., 1996; Weyerer & Bickel, 2007). Die erhobenen Tagesklinikpatienten (N = 9) präsentieren sich hinsichtlich ihrer funktionalen Selbstständigkeit sowie in ihrer kognitiven Leistungsfähigkeit etwas besser als der Durchschnitt der stationären Patienten, liegen jedoch innerhalb des Streuungsbereiches und wurden daher in den Hauptdatensatz einbezogen. Insgesamt liegt bei 29,6% der Probanden zum Zeitpunkt der Befragung eine kognitive Beeinträchtigung vor. Somit können die Eigenschaften der zu prüfenden Testverfahren auch unter der Voraussetzung eingeschränkter Verständnismöglichkeiten geprüft werden. Der Schweregrad der depressiven Symptomatik verteilt sich über die Stichprobe weitgehend erwartungsgemäß; so weist der Großteil der Erkrankten (84%) eine leichte Störung auf und nur 16% leiden an einer mittelgradigen oder schweren Störung.

Tabelle 19: Stichprobenmerkmale des Hauptdatensatzes

Stichprobenmerkmale	Probanden (Anzahl (Prozent)/ MW (SD))		
	Gesamt N = 331	nicht Depressive n = 180	Depressive n=151
<b>Geschlecht</b>			
Männlich	88 (27%)	45 (25%)	43 (29%)
Weiblich	243 (73%)	135 (75%)	108 (72%)
<b>Alter</b>	80.7 (7.6) (n=321)	81.0 (7.8) (n=175)	80.3 (7.4) (n=146)
<b>Assessment</b>			
MMSE	25.0(3.7) (n=311)	25.5(3.5) (n=172)	24.2(4.0) (n=139)
Barthel Index	51.8 (24.7) (n=211)	56 .3(26.8) (n=100)	47.7 (22.3) (n=111)
<b>Schweregrad</b>			
Unauffällig	133 (74%)	-	-
Subdiagnostisch	47 (26%)	-	-
leichte Störung	127 (84%)	-	-
mittelgradige Störung	19 (6%)	-	-
schwere Störung	5 (2%)	-	-

Anmerkung. Aufgrund des Untersuchungsdesigns ließen sich in einigen Variablen fehlende Werte nicht vermeiden. Die entsprechenden Zellen enthalten jeweils die zugrundeliegende Substichprobe für den angegebenen Mittelwert.

In der nicht erkrankten Stichprobe ist mit 26% ein Anteil subdiagnostisch verstimmter Probanden zu verzeichnen, der etwas über dem Wert der Berliner Altersstudie (Helmchen et al., 1996) liegt (operationalisiert über einer MADR-Wert von 10-12 Punkten); dieser etwas erhöhte Wert ist vermutlich auf die 2. Rekrutierungsphase zurückzuführen, in der besonders Probanden in die Erhebung eingeschlossen wurden, deren Erkrankungsstatus kritisch war. Durch den insgesamt hohen Anteil leicht erkrankter sowie subdiagnostisch verstimmter Probanden kann davon ausgegangen werden, dass in der Stichprobe kein sogenannter Spektrum-Bias vorliegt. Ein entsprechender Effekt der systematisch positiven Verzerrung in der Testeffektivität ergibt sich, wenn besonders eindeutig affektiv unauffällige und schwerer erkrankte Probanden in eine Studie einbezogen werden, da diese Patienten für ein Screeningverfahren in der Regel keine Herausforderung darstellen. Eine Einbeziehung von einem hohen Anteil Probanden, deren Depressionsstatus sich eng um den Trennwert zu einer klinisch relevanten Verstimmung verteilt, entspricht dagegen einer methodisch strengen

Validitätsprüfung (Sullivan Pepe, 2004). Insgesamt weist die erhobene Stichprobe hinsichtlich derjenigen Variablen, durch die ein Einfluss auf die Testperformanz zu erwarten ist, eine hohe Übereinstimmung mit den Eigenschaften der Zielgruppe multimorbider Patienten auf, für die das neu entwickelte Verfahren intendiert ist.

#### *4.2.5 Auswertungsstrategien*

Die Datenauswertung wurde mit dem Statistikprogramm SPSS 19 für Windows durchgeführt. Als Software für die ROC-Analysen wurde das MS-Excel-Tool (ROC-Tool2) der Firma ACOMED Statistik verwendet.

##### *4.2.5.1 Umgang mit fehlenden Werten*

Einige Variablen der Hauptuntersuchung enthalten in unterschiedlichem Umfang fehlende Werte, wobei diese jeweils mit unterschiedlichen Aspekten des Untersuchungsdesigns zusammenhängen und hinsichtlich der Zufälligkeit ihres Entstehens unterschiedlich zu klassifizieren sind (vgl. Baltes-Görtz, 2008b). Dabei sind Schätzungen fehlender Werte nach den derzeit empfohlenen Methoden der Maximum-Likelihood-Schätzung bzw. der multiplen Imputation für den vorliegenden Datensatz grundsätzlich nicht möglich, da die für diese Methoden erforderliche Voraussetzung multivariater Normalverteilung nicht erfüllt ist.

Die Strategien im Umgang mit fehlenden Werten wurden daher im Einzelnen den betroffenen Variablen sowie der jeweiligen Auswertungsstrategie angepasst und werden an den entsprechenden Stellen im Ergebnisteil diskutiert.

##### *4.2.5.2 Analyse der Testeffektivität*

Die Effektivität der Verfahren in der Klassifikation der Probanden als depressiv oder nicht depressiv wurde anhand der Relative-Operating-Characteristic-Curves (ROC-Kurven) verglichen, wobei die dichotomisierten Werte der MADR-S als Außenkriterium dienten.

Bei der Erstellung einer ROC-Kurve wird nach einer einfachen Vier-Felder-Übereinstimmungslogik die Anzahl korrekt klassifizierter Probanden eines Verfahrens mit der Anzahl der Fehlklassifikationen in Beziehung gesetzt (siehe Abbildung 17). Dabei werden für

jeden möglichen Trennwert eines Verfahrens die Gütekriterien Sensitivität und Spezifität ermittelt. Sensitivität ist die Fähigkeit des Tests, korrekt auf gegebene Erkrankungen zu reagieren und wird ermittelt als Anteil wahr positiver Testergebnisse an der Gesamtzahl positiver Diagnose-Ergebnisse. Die Spezifität wird analog als Anteil wahr negativer Ergebnisse an der Gesamtzahl negativer Diagnose-Ergebnisse ermittelt und beschreibt die Eigenschaft des Tests, Patienten ohne Erkrankungsrisiko richtig zu erkennen.

	Test positive	Test negative
Diagnose Positive	Wahr Positive	Falsch Negative
Diagnose Negative	Falsch Positive	Wahr Negative

Abbildung 17 : 4 – Felder - Schema zur Ermittlung der Testeffektivität

Die ROC- Kurve wird in ein Koordinatensystem mit der Sensitivität als Ordinate und der Falsch-Positiven-Rate als Abzisse eingetragen.

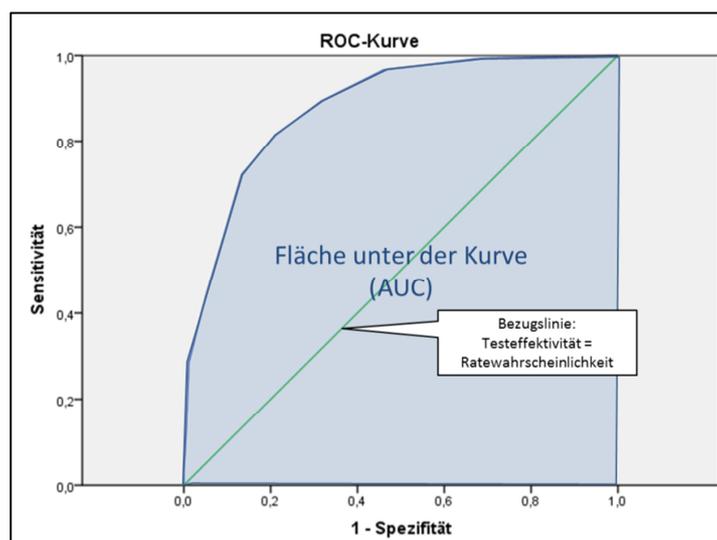


Abbildung 18: Darstellung der Testeffektivität als Fläche unter der ROC-Kurve

Die Punkte der Kurve sind Sensitivität und Falsch-Positiven-Rate (1-Spezifität) der einzelnen Trennwerte eines Tests (Beck & Shultz, 1986). Als Maß für die Testeffektivität über alle Trennwerte kann die Fläche unter der Kurve (AUC-Maß) berechnet werden (siehe Abbildung 17). Dieses Maß kann theoretisch Werte zwischen 0 und 1 annehmen und gibt die Wahrscheinlichkeit an, eine beliebige Person aus der Stichprobe durch den Test korrekt zu klassifizieren (Hanley & Mc Neil, 1982). Ein AUC-Maß von 0,5 entspricht genau der Ratewahrscheinlichkeit. Ein Test, der alle Personen der Stichprobe richtig zuordnen würde, hätte ein AUC-Maß von 1.

Der Vergleich der AUC-Werte von DIA-S und GDS15 auf signifikante Unterschiede erfolgte unter nonparametrischer Verteilungsannahme (De Long, De Long & Clark-Pearson, 1988). Zudem wurden Sensitivität und Spezifität der Tests für den jeweils besten Kennwert mithilfe des Mc-Nemar-Tests auf signifikante Unterschiede überprüft.

#### *4.2.5.3 Einflussfaktoren auf die Testeffektivität*

Die Klassifikationsraten von Selbstbeurteilungsskalen können durch verschiedene Umstände beeinflusst werden. Wenn eine Skala an Probanden getestet wird, deren Depressionsstatus besonders offensichtlich ist, besteht eine hohe Wahrscheinlichkeit, die Testeffektivität des Verfahrens zu überschätzen. Während dieser Effekt, der auch als Spektrums-, oder Extrapolations-Bias bezeichnet wird, durch das Ziehen einer für die Zielgruppe repräsentativen Stichprobe vermieden werden kann (Sullivan Pepe, 2003), gibt es darüber hinaus zahlreiche Effekte hinsichtlich der Klassifikationsraten unterschiedlicher Subpopulationen in der intendierten Zielgruppe (Henkel et al., 2004). So wurde beschrieben, dass Frauen eher dazu tendieren, einige Depressionssymptome zu berichten als Männer (Sigmon et al., 2005), und einige Studien berichten genderspezifische Ergebnisse in der Testeffektivität von Selbstbeurteilungsskalen (Semmler & Klumb, 2004; Camozzato et al., 2007). Für die Population geriatrischer Patienten sind außerdem Alter sowie funktionaler und kognitiver Status des Betroffenen relevant, wie sich anhand der großen Variabilität der

Ergebnisse für die Testqualität der GDS in einer aktuellen Meta-Analyse zeigt, bei der verschiedene Subpopulationen einander gegenübergestellt werden (stationäre Patienten, Tagesklinikpatienten, ambulante Patienten, Pflegeheimbewohner, Patienten mit oder ohne Demenz und Probanden, die zuhause befragt werden (Wancata et al., 2008). Darüber hinaus können sich Umstände des Erhebungssettings auf die Ergebnisse der Befragung auswirken. So zeigt sich beispielsweise bei gesunden Patienten ein unterschiedliches Beschwerdeverhalten hinsichtlich depressiver Symptome je nach Anzahl der Tageslichtstunden zum Erhebungszeitpunkt (Schlager, Schwartz & Bromet, 1993). Obwohl saisonale Effekte auf die Depressionsdiagnostik bei älteren Probanden noch nicht eindeutig belegt sind (de Craen, Gussekloo, van der Mast, le Cessie, Lemkes & Westendorp, 2005), sind sie wahrscheinlich i.S. eines Moderatoreffektes bei Subgruppen mit bestimmten Vulnerabilitäten relevant.

Einflüsse dieser Faktoren auf die Ergebnisse der Testeffektivitäten werden anhand multivariater binär-logistischer Regressionen getestet, wobei die Gültigkeit der multivariaten Modelle mithilfe zusätzlicher univariater Analysen überprüft wird.

## 5. Ergebnisse

### 5.1 Itemanalysen

Auf der Grundlage des Hauptdatensatzes wurden sowohl für die Items der DIA-S als auch für die der GDS15 Schwierigkeiten und Itemvaliditäten ermittelt. Darüber hinaus wurde die Trennschäfte ermittelt, da in der Hauptuntersuchung ein Vergleich des jeweiligen Items mit dem Ergebnis der Gesamtskala möglich war.

Fehlende Werte in den Items der DIA-S sowie der GDS15 wurden bei der Datenerhebung der Hauptuntersuchung explizit in Kauf genommen, wenn es einem Probanden nicht möglich war, eine „ja“/„nein“-Entscheidung zu treffen. Diese Strategie wurde gewählt, um die externe Validität der Untersuchung zu erhöhen, da fehlende Werte bei Screeninginstrumenten in der Praxis regelmäßig vorkommen und dabei jeweils die Aussagekraft der Gesamtskala mindern. Entsprechend wird die Anzahl fehlender Werte pro Item als Qualitätsmerkmal in den Itemanalysen mitberücksichtigt.

Bei der Ermittlung der Itemschwierigkeiten wurden fehlende Werte fallweise ausgeschlossen, da sich diese Kennwerte nur für die Population von Probanden, die ein Item beantworten, sinnvoll ermitteln lassen. Zur Ermittlung der Trennschärfen als korrigierte Korrelationen des Items mit der jeweiligen Gesamtskala wurden Schätzungen für fehlende Werte wie in Abschnitt 5.2.1 vorgenommen, um eine Ermittlung der Trennschärfe der einzelnen Items unabhängig von fehlenden Werten in anderen Items zu gewährleisten. Für die Analysen zur Itemvalidität wurden fehlende Werte jeweils als Werte behandelt, die nicht mit dem Kriterium übereinstimmen, da nicht beantwortete Fragen die Wahrscheinlichkeit falsch negativer Testergebnisse erhöhen und somit die Validität des Verfahrens mindern.

In der DIA-S liegt der Anteil fehlender Werte bei einem Item über 5%, in der GDS15 ist dies bei sechs Items der Fall, davon wurden drei Items von etwa 10 % der Befragten nicht beantwortet. In den Itemschwierigkeiten ergibt sich für die DIA-S eine Spanne an Lösungswahrscheinlichkeiten von .27 – .55 für die GDS15 von .20 - .75. Sowohl die

Kennwerte für die Itemvaliditäten als auch die für die Trennschärfe fallen für die DIA-S-Items (.44-.62 bzw. .41-.62) insgesamt deutlich höher aus als für die GDS15-Items (.17-.53 bzw. .08-.51).

Tabelle 20: Itemkennwerte des Hauptdatensatzes

Item	Fehlende Werte			Schwierig- keit (MW)	Validi- tät	Trenn- schärfe Rpb
	Gesamt (N(%))	D (n(%))	ND (n(%))			
<b>DIA-S</b>						
Bedrückt	8 (2.4)	5 (3.3)	3 (1.7)	.48	.57	.55
Angst falsches tun	5 (1.5)	3 (2.0)	2 (1.1)	.29	.41	.44
Gut entspannen	10 (3.0)	6 (4.0)	4 (2.2)	.36	.54	.50
Leben sinnlos	12 (3.6)	6 (4.0)	6 (3.3)	.27	.55	.55
Schwer aufzuraffen	7 (2.1)	3 (2.0)	4 (2.2)	.48	.55	.52
Angst vor der Zukunft	8 (2.4)	4 (2.6)	4 (2.2)	.37	.59	.61
Leben genießen	18 (5.4)	14 (9.3)	4 (2.2)	.32	.62	.62
Leicht überfordert	12 (3.6)	9 (6.0)	3 (1.7)	.46	.46	.47
Viel grübeln	3 (.9)	1 (.7)	2 (1.1)	.55	.56	.56
Grundsätzlich zufrieden	10 (3.0)	6 (4.0)	4 (2.2)	.26	.56	.58
<b>GDS15</b>						
Grundsätzlich zufrieden?	9 (2.7)	5 (3.3)	4 (2.2)	.29	.43	.51
Aktivitäten aufgegeben?	6 (1.8)	4 (3.6)	2 (1.1)	.69	.19	.29
Leben unausgefüllt?	13 (3.9)	7 (4.6)	6 (3.3)	.29	.50	.48
Oft langweilig?	2 (.6)	1 (.7)	1 (.6)	.35	.37	.34
Guter Laune?	20 (6.0)	11 (7.3)	9 (5.0)	.21	.49	.42
Angst Schlimmes zustoßen?	7 (2.1)	4 (2.6)	3 (1.7)	.35	.08	.17
Meiste Zeit glücklich?	31 (9.4)	14 (9.3)	17 (9.4)	.39	.51	.53
Oft hilflos?	10 (3.0)	6 (4.0)	4 (2.2)	.54	.45	.47
Zu Hause, anstatt auszugehen?	10 (3.0)	4 (2.6)	6 (3.3)	.75	.11	.14
Probleme mit dem Gedächtnis?	23 (6.9)	12 (7.9)	11 (6.1)	.20	.14	.23
Schön, jetzt zu leben?	36 (1.9)	18 (11.9)	18 (1.0)	.39	.40	.44
ziemlich wertlos?	11 (3.3)	7 (4.6)	4 (2.2)	.47	.48	.49
Voller Energie?	16 (4.8)	5 (3.3)	11 (6.1)	.62	.29	.28
Situation hoffnungslos ist?	18 (5.4)	11 (7.3)	7 (3.9)	.25	.42	.43
Meisten Leuten geht es besser ?	33 (1.0)	12 (7.9)	21 (11.4)	.33	.41	.44

Anmerkungen. D = depressive Patienten. ND = nicht depressive Patienten. Kritische Werte in den Itemstatistiken (fehlende Werte  $\geq 5\%$ , Validität  $\leq .3$ , Trennschärfe  $\leq .3$ ) wurden grau markiert.

Dabei weisen erwartungsgemäß besonders diejenigen Items der GDS15 bei geringer Schwierigkeit kritische Werte in Trennschärfe und Validität auf, deren Eignung für die vorliegende Stichprobe auch unter inhaltlichen Aspekten infrage gestellt wurde.

## 5.2 Skalenanalysen

### 5.2.1 Reliabilität

Als Kennwert für die Zuverlässigkeit der Testverfahren wurde die Interne Konsistenz ermittelt (siehe Theorieteil). Die Berechnung der Internen Konsistenz kann nur bei Skalen erfolgen, deren Items jeweils keine fehlenden Werte aufweisen, d.h. für die Berechnung des Kennwertes müssen entweder Fälle mit fehlenden Werten in einzelnen Items ausgeschlossen werden, oder fehlende Werte müssen durch Schätzwerte ersetzt werden. Bei der ersten Variante wird nur eine Teilstichprobe für die Analyse berücksichtigt (für die DIA-S 77,3% und für die GDS15 56,2% der Fälle), was zu einer Verzerrung des Ergebnisses in unbekannte Richtung führen kann, während bei der zweiten Variante der alpha-Koeffizient tendenziell überschätzt wird. Um Informationsverlust zu vermeiden, werden beide Kennwerte zum Vergleich der Skalen vorgestellt.

Als Schätzwerte für fehlenden Daten zur Berechnung des zweiten Kennwertes wurden dabei nicht die Mittelwerte der Items eingesetzt, da aufgrund der Schichtung der Stichprobe nach Depressivität die Lösungswahrscheinlichkeiten der beiden Substichproben deutlich unterschiedlich ausfallen und daher der Mittelwert nicht als guter Schätzwert für fehlende Daten angenommen werden kann.

*Tabelle 21: Interne Konsistenz der Skalen*

Skala	Items (N)	Fallausschluss bei fehlenden Werten		Schätzung von fehlenden Werten	
		Fälle (n/%)	Cronbachs Alpha	Fälle (n/%)	Cronbachs Alpha
DIA-S	10	256 (77.3%)	.84	328 (99.1%)	.84
GDS15	15	186 (56.2%)	.75	323 (97.6%)	.78

Stattdessen wurde der fehlende Wert jeweils fallweise über die durchschnittliche Beantwortung der übrigen Items der entsprechenden Skala geschätzt. Eine Schätzung des Wertes wurde dabei nur vorgenommen, wenn mindestens 80% der Skala beantwortet waren.

Nach beiden Rechenprozeduren weist die DIA-S eine Interne Konsistenz ( $\alpha = .84$ ) auf, die als hoch zu bewerten ist, während die GDS trotz einer größeren Itemzahl mit ( $\alpha = .75$ ) bzw. mit geschätzten Werten ( $\alpha = .78$ ) in einem mittleren Bereich liegt.

### 5.2.2 Validität

Zur Beurteilung der Validität der Skalen wurden jeweils die Summenwerte von DIA-S und GDS15 und MADR-Skala gebildet. Der Umgang mit fehlenden Werten bei DIA-S und GDS15 orientierte sich dabei am gängigen Vorgehen in der geriatrischen Praxis. Da das Ergebnis beider Skalen jeweils der Summenwert der beantworteten Items ist und dieser nicht ins Verhältnis zur Anzahl insgesamt beantworteter Items gesetzt wird, werden nicht beantwortete Items bei der Ermittlung des Skalenscores jeweils wie nicht gelöste Items behandelt.

Tabelle 22 gibt einen Überblick über die kriteriumsbezogene Validität der Skalen. Neben den Kennwerten zur Gesamtstichprobe wurden die Testeigenschaften auch jeweils für Männer und Frauen getrennt analysiert, um gegebenenfalls Hinweise auf geschlechtsspezifisches Antwortverhalten der Probanden zu erhalten.

Tabelle 22: Kriteriumsbezogene Validität der Skalen

Stichprobe	Skala	Korrelation		ROC-Kurve		
		$R_s$	AUC	OG	UG	D
Gesamt (N=331)	DIA-S	.737**	.882	.914	.841	
	GDS15	.662**	.828	.867	.783	.053*
Frauen (n=243)	DIA-S	.753**	.895	.931	.849	
	GDS15	.662**	.841	.885	.789	.054*
Männer (n=88)	DIA-S	.711**	.870	.933	.780	
	GDS15	.675**	.808	.885	.709	.062

Anmerkungen. \*\* $p < .01$ , \* $p < .05$ . Außenkriterium: MADRS.

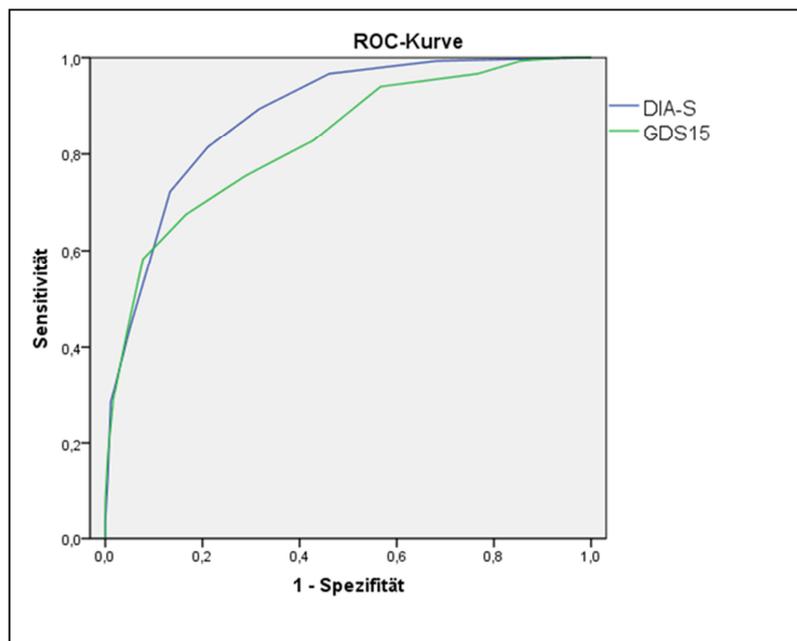
#### 5.2.2.1 Konvergente Validität

Als Kennwert für die konvergente Validität der Tests wurde jeweils die Korrelation zwischen DIA-S, GDS15 und der MADR-Skala berechnet. Aufgrund der nicht gegebenen Normalverteilung der Daten wurde dabei auf die Rangkorrelationskoeffizienten nach Spearman (Spearman's Rho) zurückgegriffen.

Die neu entwickelte Skala erzielt mit einem  $r_s$ -Wert von .737 ( $p < .001$ ) insgesamt eine höhere Übereinstimmung mit dem Außenkriterium als die GDS ( $r_s = .662$ ,  $p < .001$ ). Die Korrelation zwischen DIA-Skala und GDS beträgt  $r_s = .673$  ( $p < .000$ ). Die Korrelationskoeffizienten der männlichen und weiblichen Teilstichprobe unterscheiden sich für beide Skalen nur geringfügig.

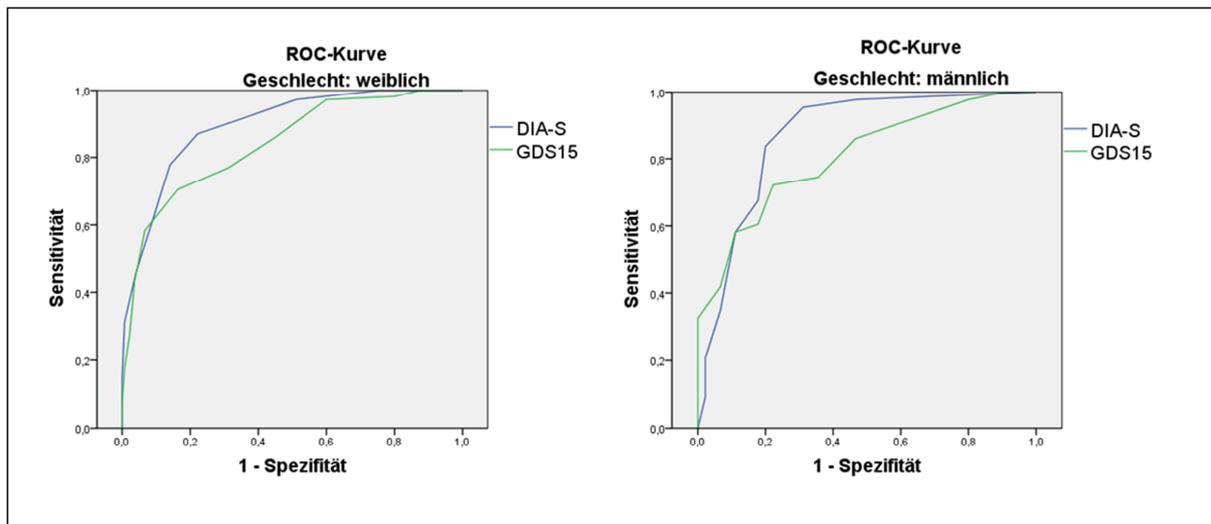
### 5.2.2.2 Testeffektivität

Der Vergleich der Testeffektivitäten von DIA-S und GDS15 wurde über den Vergleich der ROC-Kurven realisiert. Wie aus Abbildung 19 ersichtlich wird, deckt der Graph der neuen Skala eine größere Fläche ab (AUC = .88; (95% KI .84, .91)) als der Graph der GDS15 (AUC = .83, (95% KI .78, .87)), was einer höheren Wahrscheinlichkeit entspricht, einen Patienten durch die neue Skala korrekt als depressiv oder nicht depressiv zu klassifizieren. Dieser Unterschied zwischen den beiden Skalen ist dabei unter Berücksichtigung der Effektgröße sowie des Stichprobenumfangs auf dem 5%- Niveau signifikant ( $d = .05$ ,  $p < .05$ ) (vgl. Tabelle 23).



Anmerkungen. Als Außenkriterium dienen dichotomisierte MADR-Werte

Abbildung 19: Vergleich der ROC-Kurven von DIA-S und GDS



Anmerkungen. Frauen: n = 108 depressive, n = 135 nicht depressive, Männer : n = 43 depressive, n = 45 nicht depressiv.

Abbildung 20: Vergleich der ROC-Kurven von DIA-S und GDS15 nach Geschlecht

Ein jeweils ähnlicher Effekt zugunsten der neuen Skala ergibt sich auch, wenn die ROC-Analysen für die weibliche und männliche Teilstichprobe getrennt durchgeführt werden (siehe Abbildung 20). Bei entsprechend geringeren Stichprobengrößen überschreitet dieser jedoch nur für die weibliche Teilstichprobe die 5%-Irrtumswahrscheinlichkeit (Frauen:  $d = .05$ ,  $p < .05$ , Männer:  $d = .06$ , n.s.).

### 5.2.2.3 Testeffektivität in einzelnen Trennwerten

Tabelle 23 gibt einen Überblick über die jeweils besten Trennwerte der Skalen. Für die Gesamtstichprobe liegt der empirisch beste Trennwert der DIA-S zwischen 3 und 4 Punkten mit einer Sensitivität von 82% und einer Spezifität von 79%, während der beste Trennwert der GDS15 zwischen 5 und 6 Punkten liegt und eine Sensitivität von 76% sowie eine Spezifität von 71% erreicht.

*Tabelle 23: Vergleich von Sensitivität und Spezifität für die effektivsten Trennwerte der Skalen*

Depressive Patienten				McNemar-Test			
n	Skala	Trennwert	Sensitivität	$\chi^2$	df	p(zweiseitig)	p(einseitig)
Gesamt	DIA-S	3.5	.82				
(n= 151)	GDS15	5.5	.76	1.83	1	.176	.088
Frauen	DIA-S	3.5	.87				
(n=108)	GDS15	5.5	.77	4.35	1	.035*	.017*
Männer	DIA-S	2.5	.84				
(n=43)	GDS15	5.5	.72	-	-	-	-
Nicht depressive Patienten				McNemar-Test			
n	Skala	Trennwert	Spezifität	$\chi^2$	df	p(zweiseitig)	p(einseitig)
Gesamt	DIA-S	3.5	.79				
(n=180)	GDS15	5.5	.71	4.23	1	.040*	.020*
Frauen	DIA-S	3.5	.78				
(n=135)	GDS15	5.5	.69	3.78	1	.050	.025*
Männer	DIA-S	2.5	.80				
(n=45)	GDS15	5.5	.78	-	-	-	-

Anmerkungen. \*p<.05. Der McNemar-Test konnte für die männliche Teilstichprobe nicht durchgeführt werden, da die Zellenbesetzung der 4-Feldertafeln z.T. unter n=5 lag (vgl. Bortz & Liener, 1998).

Sensitivität und Spezifität in den jeweils besten Trennwerten der Skalen wurden mithilfe des Mc-Nemar-Tests auf signifikante Unterschiede überprüft (vgl. Trajman & Luiz, 2008). Dabei erzielt die neue Skala in der Spezifität hypothesenkonform signifikant bessere Ergebnisse ( $\chi^2$  (.05;1,N=180) = 4.23). In der Sensitivität kann unter Berücksichtigung der Stichprobengröße nicht von einer signifikanten Verbesserung ausgegangen werden, es zeichnet sich jedoch ein positiver Trend zugunsten des neuen Verfahrens ab ( $\chi^2$  (n.s.; 1, N=151) = 1.83).

Für beide Skalen ergeben sich Gendereffekte hinsichtlich der Klassifikationsraten. So liegt der beste Trennwert der GDS15 für beide Geschlechter zwischen 5 und 6 Punkten. Allerdings zeigt sich hier für die männliche Teilstichprobe eine etwas geringere Erkennens-Rate erkrankter Probanden, während die Spezifität des Verfahrens in der weiblichen Teilstichprobe um 10% geringer ist. Für die neue Skala konnten jeweils ähnlich gute Klassifikationsraten für die weiblichen (Sensitivität: 87%, Spezifität: 78%) und männlichen

Probanden (Sensitivität: 84%, Spezifität: 80%) erzielt werden, indem der Trennwert für die männliche Teilstichprobe um einen Punkt herabgesetzt wurde.

### *5.2.3 Einflussfaktoren auf die Testvalidität*

Da die Klassifikationsraten von Testverfahren nicht nur durch das Geschlecht, sondern auch durch weitere Eigenschaften des Probanden sowie des Erhebungssettings beeinflusst werden können, wurde der Einfluss möglicher Prädiktoren auf die Fehlerraten der Skalen jeweils durch hierarchisch logistische Regressionen überprüft. Dabei sollten Screeningverfahren idealiter auch unter heterogenen Erhebungsbedingungen konstant gültige Ergebnisse erzielen. Um diese Eigenschaft der Verfahren zu prüfen, wurden in die multivariaten Analysen alle erhobenen Prädiktoren einbezogen, die theoretisch für die Testperformanz relevant sein könnten; so war es nicht das Ziel der Analysen, auf der empirischen Datengrundlage statistisch besonders bedeutsame Modelle zu erstellen, sondern es sollte für jeden Prädiktor die Nullhypothese, dass kein Einfluss auf die Testqualität besteht, überprüft und gegebenenfalls bestätigt werden.

#### *5.2.3.1 Umgang mit fehlenden Werten*

Vollständige Assessmentdaten zu den Prädiktorvariablen, wie sie zur Berechnung der binär-logistischen Regressionsanalyse erforderlich sind, sind nur für einen Teildatensatz der untersuchten Stichprobe (n=195) verfügbar. Fälle mit fehlenden Werten wurden daher für die entsprechenden Analysen ausgeschlossen. Baltès-Götz (2008b) weist darauf hin, dass die Qualität der Schätzwerte bei Regressionsanalysen mit fallweisem Ausschluss fehlender Werte weniger von der Repräsentativität der Teilstichprobe vom Gesamtdatensatz als vielmehr von der Unabhängigkeit des Musters fehlender Werte von der Kriteriumsvariable abhängt. Entsprechend erscheint die Vorgehensweise des fallweisen Ausschlusses für die geplante Analyse zulässig. So weicht die vorliegende Teilstichprobe zwar hinsichtlich einiger Merkmale vom Hauptdatensatz ab (siehe Tabelle im Anhang), es lässt sich jedoch zeigen, dass die Zugehörigkeit zur Teilstichprobe keinen Einfluss auf die korrekte Zuordnung der

Probanden hat: Es ergibt sich für beide Messinstrumente kein signifikanter Zusammenhang mit dem Muster fehlender Werte in den Assessmentdaten (DIA-S:  $\chi^2(\text{n.s.};1, N=331)=.277$ , GDS15:  $\chi^2(\text{n.s.};1, N=331)=.156$ ).

#### *5.2.3.2 Erstellung der Regressionsmodelle*

Die Prädiktoren wurden den Modellen der logistischen Regression schrittweise beigefügt. Im ersten Schritt wurde lediglich die Klassifikationsschwierigkeit der Probanden in Abhängigkeit von Ihrem Erkrankungsschweregrad berücksichtigt. So ist davon auszugehen, dass besonders schwer depressiv erkrankte Probanden sowie Probanden, die eindeutig keine depressive Symptomatik aufweisen, durch ein Screeningverfahren häufiger korrekt klassifiziert werden, als Probanden, deren Symptomatik nahe an der „wahren“ Erkrankungsgrenze liegt. Entsprechend wurde die Klassifikationsschwierigkeit operationalisiert durch den Betrag des Abstands des MADR-Wertes eines Probanden vom Trennwert der MADR-S.

In den weiteren Schritten wurden den Modellen Alter (Schritt 2), Geschlecht (Schritt 3), kognitive und funktionelle Einschränkung des Probanden (Schritt 4), und schließlich der Anteil der tageslichtlosen Zeit zum Erhebungsdatum (Schritt 5) hinzugefügt. Als Kriterium diente hierbei jeweils die Übereinstimmung der jeweiligen Skala mit dem Außenkriterium („1“ = falsche Vorhersage, „0“ = korrekte Vorhersage), wobei der beste Kennwert für die Gesamtstichprobe zugrunde gelegt wurde.

#### *5.2.3.3 Überprüfung der Multikolarität*

Tabelle 24 gibt einen Überblick über die bivariaten Korrelationen der einzelnen Prädiktoren. In der Stichprobe ergeben sich erwartungsgemäß signifikante Zusammenhänge zwischen Alter und Geschlecht sowie zwischen kognitiver und funktionaler Einschränkung. Zudem zeigt sich ein saisonaler Einfluss auf die Klassifikationsschwierigkeit, bedingt durch einen Anstieg subdiagnostischer und leichter depressiver Symptomatik in der dunklen Jahreszeit. Bei insgesamt geringen Effektstärken ist auf der Grundlage der bivariaten Zusammenhänge nicht von einem problematischen Ausmaß an Multikolarität auszugehen. Eine lineare

Hilfsregression bestätigt diese Vermutung auf multivariater Ebene mit Toleranzkoeffizienten zwischen .90 und .96 (vgl. Albers & Skiera, 1999).

*Tabelle 24: Interkorrelationen der Prädiktoren für die hierarchische logistische Regression, n=195*

	Klassifizierungs- Schwierigkeit <i>R/R<sub>pb</sub></i>	Alter <i>R/R<sub>pb</sub></i>	Geschlecht <i>R<sub>pb</sub></i>	Kognitive Einschränkung <i>R/R<sub>pb</sub></i>	Funktionale Einschränkung <i>R/R<sub>pb</sub></i>
Klassifizierungs- Schwierigkeit	-				
Alter	.076	-			
Geschlecht	-.079	-.152*	-		
Kognitive Einschränkung	-.074	.118	.063	-	
Funktionale Einschränkung	-.025	.027	.051	.269**	-
Dunkelheit	.201**	-.051	-.076	.011	-.079

Anmerkungen. \*p<.05, \*\*p<.01, R= Korrelation nach Pearson, Rpb = punktbiseriale Korrelation

#### 5.2.3.4 Ergebnisse der Regressionsanalysen: Fehlklassifikation durch die GDS15

Für die multivariate Analyse zur Erklärung von Fehlklassifikationen durch die GDS15 standen insgesamt 195 Datensätze zur Verfügung, davon wurden n=141 durch den Test richtig und n=53 falsch klassifiziert. Eine globale Beurteilung der Modellgültigkeit unter Berücksichtigung der Hosmer-Lemeshow-Statistik war für alle Schritte der Regression akzeptabel  $\chi^2_{HL(n.s.;8;N=195)} = 2.586-12.614$ . Ein Vergleich der einzelnen Modellschritte unter Berücksichtigung der logarithmierten Likelihood-Quotienten ergibt nur für die ersten beiden Schritte der Regression eine signifikante Verbesserung des Modells (Schritt 1:  $\chi^2_{(.01;1; N=195)} = 9.643$ , Schritt 2  $\chi^2_{(.05;1; N=195)} = 5.898$ ). Entsprechend werden die Maße der Modellrelevanz jeweils nur für das Modell zu Schritt 2 unter Einbeziehung der Variablen Klassifikationsschwierigkeit und Alter berichtet. Ein Überblick zu den Kennwerten aller

Tabelle 25: Hierarchisch logistische Regression zur Erklärung von Fehlklassifizierungen durch die GDS15, n=195

Prädiktor	B	SE B	Wald	OR	95% KI		X <sup>2</sup>	df	Modell		
					UG	OG			Δ X <sup>2</sup>	R <sup>2</sup>	ΔR <sup>2</sup>
Schritt 1											
Klassifizierungsschwierigkeit	.153	.051	8.822**	1.165	1.053	1.288	9.64**	1	9.643**	.07	
Schritt 2											
Klassifizierungsschwierigkeit	.152	.053	8.311**	1.164	1.050	1.290					
Alter	.051	.024	4.510*	1.053	1.004	1.103	14.54**	2	4.898*	.104	.034
Schritt 3											
Klassifizierungsschwierigkeit	.150	.053	8.066**	1.161	1.047	1.288					
Alter	.049	.024	4.131*	1.051	1.002	1.102					
Geschlecht	.211	.374	.318	1.235	.594	2.586	14.86**	3	.322	.106	.002
Schritt 4											
Klassifizierungsschwierigkeit	.148	.054	7.640**	1.160	1.044	1.289					
Alter	.055	.025	4.848*	1.056	1.006	1.109					
Geschlecht	.187	.379	.243	1.206	.574	2.543					
Kognitive Einschränkung	-.097	.053	3.325	.908	.818	1.007					
Funktionelle Einschränkung	.005	.007	.418	1.005	.990	1.019	18.406**	5	3.543	.130	.024
Schritt 5											
Klassifizierungsschwierigkeit	.132	.055	5.881*	1.141	1.026	1.270					
Alter	.056	.025	5.193*	1.058	1.008	1.111					
Geschlecht	.171	.382	.200	1.187	.561	2.510					
Kognitive Einschränkung	-.102	.053	3.648	.903	.991	1.020					
Funktionelle Einschränkung	.006	.007	.597	1.006	.991	1.020					
Dunkelheit	.005	.004	1.806	1.005	.998	1.012	2.240**	6	1.834	.142	.012

Anmerkungen. \*p<.05, \*\*p>.01. R<sup>2</sup>= Varianzaufklärung nach Nagelkerke.

Tabelle 26: Hierarchisch logistische Regression zur Erklärung von Fehlklassifizierungen durch die DIA-S, n=195

Prädiktor	B	SE B	Wald	OR	95% KI		X <sup>2</sup>	df	Modell		
					UG	OG			Δ X <sup>2</sup>	R <sup>2</sup>	ΔR <sup>2</sup>
Schritt 1											
Klassifizierungsschwierigkeit	.224	.064	12.165**	1.251	1.1.03	1.419	14.348**	1	14.348**	11.4	
Schritt 2											
Klassifizierungsschwierigkeit	.224	.064	12.286**	1.252	1.1.4	1.419					
Alter	-.011	.025	.212	.989	.942	1.038	14.559**	2	.211	11.6	.2
Schritt 3											
Klassifizierungsschwierigkeit	.235	.065	13.061**	1.265	1.114	1.438					
Alter	-.005	.025	.047	.995	.946	1.045					
Geschlecht	-.645	.404	2.544	.525	.237	1.159	17.068**	3	2.508	13.5	1.9
Schritt 4											
Klassifizierungsschwierigkeit	.235	.065	13.082**	1.265	1.114	1.437					
Alter	-.006	.025	.060	.994	.945	1.045					
Geschlecht	-.627	.407	2.375	.534	.241	1.186					
Kognitive Einschränkung	.008	.056	.018	1.008	.902	1.126					
Funktionelle Einschränkung	.003	.008	.115	1.003	.987	1.019	17.238**	5	.170	13.6	.1
Schritt 5											
Klassifizierungsschwierigkeit	.233	.067	12.107**	1.262	1.107	1.439					
Alter	-.006	.025	.057	.994	.946	1.045					
Geschlecht	-.627	.407	2.375	.534	.241	1.186					
Kognitive Einschränkung	.007	.057	.017	1.007	.902	1.125					
Funktionelle Einschränkung	.003	.008	.124	1.003	.987	1.019					
Dunkelheit	.001	.004	.020	1.001	.992	1.009	17.258**	6.	.020	13.6	0

Anmerkungen. \*p<.05, \*\*p>.01. R<sup>2</sup>= Varianzaufklärung nach Nagelkerke.

Modelle mit den Statistiken der einzelnen Prädiktoren ist Tabelle 27 zu entnehmen. Bei einer akzeptablen Varianzaufklärung (Nagelkerke  $R^2 = .104$ ) ist die prädiktive Effizienz des Modells eher gering; so werden insgesamt 73.8% der Fälle korrekt vorhergesagt, wobei 9.3% der Fehlklassifikationen richtig zugeordnet werden im Gegensatz zu 98.6% der richtig diagnostizierten Patienten. Bei einer a priori – Wahrscheinlichkeit der richtigen Vorhersage des Kriteriums von 72.3% durch das Basismodell steigt somit die Vorhersagegenauigkeit unter Einbeziehung der Prädiktoren insgesamt nur marginal an.

Auf Ebene der einzelnen Prädiktoren trägt neben der Klassifikationsschwierigkeit der Fälle, die erwartungsgemäß einen bedeutsamen Beitrag zur Vorhersage der Fehldiagnosen leistet, Schritt 2: Wald  $\chi^2_{(.01;1; N=195)}=8.311$ , das Alter der Probanden signifikant zur Erklärung falscher Testergebnisse bei, Wald  $\chi^2_{(.05;1; N=195)}=4.131$ . So steigt die Chance, durch die GDS15 falsch diagnostiziert zu werden, pro Lebensjahr des Probanden jeweils um 5% an.

#### *5.2.3.5 Ergebnisse der Regressionsanalysen: Fehlklassifikation durch die DIA-S*

Einen Überblick über die Ergebnisse der hierarchisch logistischen Regression zur Erklärung von Fehlklassifikationen durch die DIA-S gibt Tabelle 28. Grundlage der Analyse sind 195 Fälle, von denen jeweils  $n=37$  falsch und  $n=158$  durch den Test richtig klassifiziert werden. Das Regressionsmodell erreicht in allen Schritten eine hinreichende Anpassungsgüte  $\chi^2_{HL(n.s.;8;N=195)}= 5.925-13.419$ , jedoch trägt nur die Einbeziehung des Prädiktors Klassifizierungsschwierigkeit im ersten Schritt signifikant zu einer Modellverbesserung nach dem logarithmierten Likelihood-Quotienten bei,  $\chi^2_{(.01;1; N=195)} = 14.348$ , bei einer Varianzaufklärung nach Nagelkerke von .114. Entsprechend kann die Chance zur Fehlklassifikationen nur durch die Nähe der Symptomausprägung des Probanden zum wahren diagnostischen Kriterium erklärt werden. Eine Verbesserung der prädiktiven Vorhersagegenauigkeit kann dadurch im Vergleich zum Basismodell jedoch nicht erzielt werden (Gesamtklassifikationsrate jeweils 81% mit 100% korrekten Klassifikationen der richtigen und 0% der vorhergesagten Fehlklassifikationen).

### 5.2.3.6 Überprüfung der Modellrelevanz anhand bivariater Zusammenhänge

Da sich in den beschriebenen logistischen Regressionsanalysen möglicherweise Einschränkungen für die inhaltliche Interpretation ergeben, wird die inhaltliche Relevanz der Modelle im folgenden Abschnitt anhand der bivariaten Zusammenhänge der Klassifikationen überprüft. So ergibt sich für die Modelle einerseits das Problem, dass sich das dichotome Kriterium sehr ungleich auf die verschiedenen Gruppen verteilt, was dazu führt, dass jeweils nur eine verhältnismäßig geringe Anzahl an Fehlklassifikationen ( $n=54$  für die GDS15 und  $n=37$ ) für die Analysen zur Verfügung steht. Dabei wird in der Literatur (vgl. Baltes-Görtz, 2008c, Fromm, 2005) darauf hingewiesen, dass für die Erstellung von Regressionsmodellen das Verhältnis von Beobachtungen zu den erklärenden Prädiktoren berücksichtigt werden sollte, wobei insbesondere die Anzahl der Fälle in der schwächer besetzten Kategorie beachtet werden muss, da ansonsten die Teststärke des Verfahrens gering ist. Für die vorliegenden Analysen bedeutet dies, dass die Bestätigung der Nullhypothese für die später aufgenommenen Variablen eher auf die Größe der Standardfehler als auf zu geringe Effektstärken zurückzuführen ist. Die Berechnung der bivariaten Zusammenhänge zwischen den Fehlklassifikationen und den Prädiktoren (siehe Tabelle 28) unterstützt jedoch die Ergebnisse der multivariaten Analysen. So ergibt sich für die DIA-S auch hier nur hinsichtlich der Klassifikationsschwierigkeit ein signifikantes Ergebnis ( $R_{pb} = .260, p < .01$ ) während für die GDS15 darüber hinaus das Alter der Probanden von Bedeutung ist ( $R_{pb} = .217, p < .01$ ;  $R_{pb} = .163, p < .05$ ).

Eine weitere mögliche Einschränkung bei der Interpretation der Prädiktoren für Fehlklassifikationen ergibt sich aus der Zusammenfassung Falsch positiver und Falsch negativer Testergebnisse in eine Kategorie. So ist es denkbar, dass für die jeweilige Fehlerart unterschiedliche Faktoren relevant sind, bzw. dass der Prädiktor je nach Typ des Fehlers in verschiedene Richtungen wirkt. So zeigt sich beispielsweise in Abschnitt 5.2.2, dass bei weiblichen Probanden die Wahrscheinlichkeit falsch positiver Testergebnisse höher ist,

während männliche Probanden tendenziell häufiger falsch negative Testergebnisse aufweisen. Für die Anzahl falscher Testergebnisse allgemein spielt das Geschlecht dagegen kaum eine Rolle.

*Tabelle 27: Bivariate Zusammenhänge zwischen Fehlklassifikationen von DIA-S und GDS15 und den Prädiktoren*

Prädiktor	DIA-S			GDS15		
	Falsche Gesamt	Falsch Positive	Falsch Negative	Falsche Gesamt	Falsch Positive	Falsch Negative
	$R_{pb}/\chi^2$	$R_{pb}/\chi^2$	$R_{pb}/\chi^2$	$R_{pb}/\chi^2$	$R_{pb}/\chi^2$	$R_{pb}/\chi^2$
Klassifizierungs-Schwierigkeit	.260**	.398**	.137	.217**	.294**	.154
	n=195	n=94	n=101	n=195	n=94	n=101
	.280**	.379**	.181*	.259**	.319**	.204*
	N=331	n=180	n=151	N=331	n=180	n=151
Alter	-.020	-.032	-.009	.163*	.192	.130
	n=195	n=94	n=101	n=195	n=94	n=101
	.001	.023	-.031	.139	.196**	.058
	n=321	n=175	n=146	n=321	n=175	n=146
Geschlecht	1.611	.064	4.322*	1.186	.532	.736
	n=195	n=94	n=101	n=195	n=94	n=101
	1.923	.400	7.818**	.217	1.298	.376
	N=331	n=180	n=151	N=331	n=180	n=151
Kognitive Einschränkung	.010	.030	-.001	-.116	-.020	-.200*
	n=195	n=94	n=101	n=195	n=94	n=101
	.030	.059	.010	-.036	.097	-.177*
	N=311	n=172	n=139	n=311	n=169	n=139
Funktionale Einschränkung	.032	.046	.028	.013	.112	-.087
	n=195	n=94	n=101	n=195	n=94	n=101
	.003	.041	-.033	-.016	.108	-.135
	n=211	n=100	n=111	n=211	n=98	n=111
Dunkelheit	.068	.286**	-.069	.123	.134	.148
	n=195	n=94	n=101	n=195	n=94	n=101
	-.000	.035	-.028	.051	-.001	.131
	N=331	n=180	n=151	N=331	n=180	n=151

Anmerkungen. \* $p < .05$ , \*\* $p < .01$ ,  $R_{pb}$  = punktbiseriale Korrelation,  $\chi^2$  = Chi-Quadrat nach Pearson,  $df=1$ . Der obere Wert bezieht sich jeweils auf die Teilstichprobe der Teilnehmer mit allen Assessmentdaten, der für die multivariaten Analysen verwendet wurde, der untere Wert auf alle verfügbaren Daten.

Betrachtet man die in Tabelle 27 aufgeführten Zusammenhänge, so zeigen sich bei einer Aufteilung der Fehlklassifizierung nach Art des Fehlers hinsichtlich einiger Faktoren weitere Effekte. So ist die Schwierigkeit der Zuordnung erwartungsgemäß bei beiden Skalen für beide

Fehlertypen relevant, die Wahrscheinlichkeit für eine falsch positive Klassifizierung steigt jedoch deutlicher als die für eine falsch negative. Für die GDS15 steht das Alter des Probanden in positivem Zusammenhang mit beiden Fehlertypen, wobei der Zusammenhang für die falsch positiven Ergebnisse deutlicher ausfällt. Statistische Bedeutsamkeit wird allerdings aufgrund der geringeren Teststärke bei Aufteilung der Teilstichprobe nicht erreicht. Ein Vergleich mit der Korrelation unter Berücksichtigung aller verfügbaren Daten für den bivariaten Zusammenhang deutet zudem darauf hin, dass der Effekt für die falsch positiven Ergebnisse konstant zu sein scheint ( $n=94$ ,  $R_{pb}=.192$ , n.s.;  $n=175$ ,  $R_{pb}=.196$ ,  $p<.01$ ), während er für die falsch negativen Ergebnisse möglicherweise überschätzt wird ( $n=101$ ,  $R_{pb}=.130$ , n.s.,  $n = 146$   $R_{pb}=.058$ , n.s.). Ein weiterer Faktor für das Vorkommen falsch negativer Testergebnisse der GDS15 ist der kognitive Status der Probanden. So ist die Wahrscheinlichkeit, dass depressive Erkrankungen durch den Test nicht angezeigt werden bei Probanden mit geringerer kognitiver Einschränkung höher ( $R_{pb} = -.200$ ,  $p<.01$ ). Hinsichtlich der funktionellen Einschränkung ergibt sich bei der GDS15 eine Umkehrung des Effektes für den jeweiligen Fehlertypus. So ist eine funktionale Einschränkung erwartungsgemäß eher mit falsch positiven Testergebnissen assoziiert während funktionelle Selbstständigkeit eher in Zusammenhang mit falsch negativen Ergebnissen steht. Bei insgesamt geringen Effektstärken sowie mangelnder statistischer Bedeutsamkeit sind entsprechende Ergebnisse jedoch unter Vorbehalt zu sehen.

Für die DIA-S ergibt sich bei einer Differenzierung der Fehlklassifizierung ein signifikanter Geschlechtseffekt für das Vorkommen falsch negativer Testergebnisse  $\chi^2_{(.05;1; N=101)} = 4.322$ . Ein entsprechender Zusammenhang ergibt sich für die vorliegenden Daten allerdings dadurch, dass für die Analysen für beide Geschlechter der gleiche Trennwert zugrunde gelegt wurde. Mit einer Absenkung des Trennwertes für die männlichen Probanden, wie er in Abschnitt 5.2.2 diskutiert wird, fällt die Sensitivität für beide Geschlechter ähnlich hoch aus (vgl. Tabelle x). Zudem sind falsch-positive Testergebnisse der DIA-S unter

Berücksichtigung der bivariaten Analysen mit saisonalen Einflüssen assoziiert ( $n= 94$ ,  $R_{pb}=.286$ ,  $p < .01$ ), allerdings ergeben sich, wenn man für den Prädiktor Dunkelheit jeweils die Ergebnisse der Teilstichprobe mit denen der Gesamtstichprobe vergleicht, wesentlich stärkere Schwankungen in den Effekten als für die anderen Variablen, sodass der verhältnismäßig starke Zusammenhang möglicherweise auf ein statistisches Artefakt zurückzuführen ist ( $n= 180$ ,  $R_{pb}=.035$ , n.s.).

Da bivariate Analysen aufgrund der Vernachlässigung möglicher Interaktionseffekte sowie aufgrund der Problematik des wiederholten Testens unter Vorbehalt zu sehen sind, wurden für die unterschiedlichen Fehlertypen der Skalen zusätzlich die hierarchisch logistischen Regressionsmodelle berechnet, wie sie auch für die Fehlklassifikationen allgemein verwendet wurden. Dabei weisen die Gesamtmodelle aufgrund der geringen Anzahl an Fällen jeweils nur geringe statistische und praktische Bedeutsamkeit auf und werden daher an dieser Stelle nicht diskutiert. Auf Ebene der einzelnen Prädiktoren bestätigen sich jedoch jeweils die in den bivariaten Analysen gefundenen Zusammenhänge (vgl. Tabellen 28-31).

Tabelle 28: Hierarchisch logistische Regression zur Erklärung von falsch- positiv- Klassifizierungen durch die GDS15

Prädiktor	95% KI						Modell				
	<i>B</i>	<i>SE B</i>	<i>Wald</i>	<i>OR</i>	<i>UG</i>	<i>OG</i>	<i>X<sup>2</sup></i>	<i>df</i>	$\Delta X^2$	<i>R<sup>2</sup></i>	$\Delta R^2$
Schritt 1											
Klassifizierungsschwierigkeit	.202	.073	7.583	1.224	1.060	1.413	8.292**	1	8.292**	.120	.120
Schritt 2											
Klassifizierungsschwierigkeit	.210	.076	7.644	1.234	1.063	1.432	12.058**	2	3.766	.171	.159
Alter	.063	.034	3.355	1.065	.996	1.139					
Schritt 3											
Klassifizierungsschwierigkeit	.214	.077	7.630	1.238	1.064	1.441	12.133**	3	.075	.172	.001
Alter	.065	.036	3.363	1.068	.996	1.145					
Geschlecht	-.151	.551	.075	.860	.292	2.531					
Schritt 4											
Klassifizierungsschwierigkeit	.212	.079	7.301	1.237	1.060	1.443	13.495*	5	1.362	.190	.018
Alter	.069	.036	3.694	1.071	.999	1.148					
Geschlecht	-.190	.556	.116	.827	.278	2.460					
Kognitive Einschränkung	-.047	.073	.419	.954	.827	1.100					
Funktionelle Einschränkung	.011	.010	1.242	1.011	.992	1.031					
Schritt 5											
Klassifizierungsschwierigkeit	.192	.081	5.624*	1.211	1.034	1.419	14.719*	6	1.224	.206	.016
Alter	.074	.036	4.093*	1.076	1.002	1.156					
Geschlecht	-.142	.568	.062	.868	.285	2.643					
Kognitive Einschränkung	-.055	.073	.562	.947	.820	1.093					
Funktionelle Einschränkung	.013	.010	1.680	1.013	.993	1.034					
Dunkelheit	.007	.007	1.270	1.007	.994	1.021					

Anmerkungen. \* $p < .05$ , \*\* $p > .01$ .  $R^2$ = Varianzaufklärung nach Nagelkerke.

Tabelle 29: Hierarchisch logistische Regression zur Erklärung von falsch- positiv- Klassifizierungen durch die DIA-S

Prädiktor	95% KI						Modell				
	<i>B</i>	<i>SE B</i>	<i>Wald</i>	<i>OR</i>	<i>UG</i>	<i>OG</i>	<i>X<sup>2</sup></i>	<i>df</i>	$\Delta X^2$	<i>R<sup>2</sup></i>	$\Delta R^2$
Schritt 1											
Klassifizierungsschwierigkeit	.345	.089	12.305**	1.412	1.164	1.731	15.933**	1	15.933**	.246	.246
Schritt 2											
Klassifizierungsschwierigkeit	.345	.089	12.311**	1.411	1.164	1.711	16.034**	2	.101	.247	.001
Alter	-.011	.035	.101	.989	.924	1.059					
Schritt 3											
Klassifizierungsschwierigkeit	.348	.099	12.263**	1.416	1.166	1.721	16.095**	3	.061	.248	.001
Alter	-.008	.037	.048	.992	.923	1.066					
Geschlecht	-.162	.652	.061	.851	.237	3.055					
Schritt 4											
Klassifizierungsschwierigkeit	.346	.100	12.097**	1.414	1.163	1.719	16.121**	5	0.26	.248	.000
Alter	-.008	.037	.051	.992	.922	1.066					
Geschlecht	-.151	-.656	.051	.860	.238	3.109					
Kognitive Einschränkung	.012	.084	.019	1.012	.858	1.192					
Funktionelle Einschränkung	.000	.011	.001	1.00	.979	1.022					
Schritt 5											
Klassifizierungsschwierigkeit	.310	.103	9.082**	1.3631	1.114	1.667	20.492**	6	4.372*	.309	.061
Alter	-.002	.039	.003	.998	.925	1.077					
Geschlecht	.111	.703	.025	1.117	.282	4.429					
Kognitive Einschränkung	-.006	.087	.004	.994	.839	1.179					
Funktionelle Einschränkung	.005	.012	.214	.994	.839	1.029					
Dunkelheit	.018	.009	3.933*	1.018	1.000	1.036					

Anmerkungen. \*p<.05, \*\*p>.01. R<sup>2</sup>= Varianzaufklärung nach Nagelkerke.

Tabelle 30: Hierarchisch logistische Regression zur Erklärung von falsch- negativ- Klassifizierungen durch die GDS15

Prädiktor	95% KI						Modell				
	<i>B</i>	<i>SE B</i>	<i>Wald</i>	<i>OR</i>	<i>UG</i>	<i>OG</i>	<i>X<sup>2</sup></i>	<i>df</i>	$\Delta X^2$	<i>R<sup>2</sup></i>	$\Delta R^2$
Schritt 1											
Klassifizierungsschwierigkeit	.109	.071	2.320	1.115	.969	1.282	2.520	1	2.520	.036	.036
Schritt 2											
Klassifizierungsschwierigkeit	.102	.073	1.973	1.108	.960	1.277	3.930	2	1.510	.056	.020
Alter	.040	.035	1.325	1.041	.972	1.114					
Schritt 3											
Klassifizierungsschwierigkeit	.103	.073	1.977	1.108	.960	1.278	4.680	3	.750	.067	.011
Alter	.040	.035	1.334	1.041	.972	1.115					
Geschlecht	.456	.537	.720	1.578	.550	4.522					
Schritt 4											
Klassifizierungsschwierigkeit	.075	.078	.924	1.077	.925	1.254	1.062	5	5.382	.139	.072
Alter	.068	.040	2.859	1.071	.989	1.159					
Geschlecht	.414	.564	.539	1.512	.501	4.566					
Kognitive Einschränkung	-.172	.083	4.347*	.842	.716	.990					
Funktionelle Einschränkung	-.006	.012	.275	.994	.971	1.017					
Schritt 5											
Klassifizierungsschwierigkeit	.059	.078	.569	1.061	.910	1.236	11.406	6	1.344	.157	.018
Alter	.067	.040	2.847	1.070	.989	1.157					
Geschlecht	.360	.568	.402	1.434	.471	4.366					
Kognitive Einschränkung	-.177	.084	4.445*	.838	.711	.988					
Funktionelle Einschränkung	-.006	.012	.215	.994	.971	1.018					
Dunkelheit	.006	.005	1.286	1.006	.996	1.016					

Anmerkungen. \*p<.05, R<sup>2</sup>= Varianzaufklärung nach Nagelkerke.

Tabelle 31: Hierarchisch logistische Regression zur Erklärung von falsch- negativ- Klassifizierungen durch die DIA-S

Prädiktor	95% KI						Modell					
	<i>B</i>	<i>SE B</i>	<i>Wald</i>	<i>OR</i>	<i>UG</i>	<i>OG</i>	<i>X<sup>2</sup></i>	<i>df</i>	$\Delta X^2$	<i>R<sup>2</sup></i>	$\Delta R^2$	
Schritt 1												
Klassifizierungsschwierigkeit	.113	.083	1.840	1.119	.951	1.318	1.886	1	1.886	.033	.033	
Schritt 2												
Klassifizierungsschwierigkeit	.115	.083	1.896	1.122	.953	1.320	2.088	2	.065	.034	.001	
Alter	-.009	.036	.066	.991	.923	1.063						
Schritt 3												
Klassifizierungsschwierigkeit	.123	.086	2.063	1.131	.956	1.338	6.303	3	4.215*	.099	.065	
Alter	-.009	.037	.066	.991	.922	1.064						
Geschlecht	-1.122	.544	4.260*	.326	.112	.945						
Schritt 4												
Klassifizierungsschwierigkeit	.127	.087	2.101	1.135	.956	1.347	6.349	5	.045	.100	.001	
Alter	-.011	.037	.084	.989	.920	1.064						
Geschlecht	-1.099	.554	3.937*	.333	.112	.987						
Kognitive Einschränkung	.003	.078	.002	1.003	.862	1.168						
Funktionelle Einschränkung	.003	.013	.040	1.003	.977	1.029						
Schritt 5												
Klassifizierungsschwierigkeit	.140	.091	2.376	1.150	.963	1.374	6.842	6	.493	.108	.008	
Alter	-.011	.038	.077	.990	.919	1.066						
Geschlecht	-1.067	.559	3.646	.344	.115	1.029						
Kognitive Einschränkung	-.001	.078	.000	.999	.857	1.164						
Funktionelle Einschränkung	.001	.013	.007	1.001	.975	1.027						
Dunkelheit	-.004	.005	.494	.996	.986	1.007						

Anmerkungen. \*p<.05, R<sup>2</sup>= Varianzaufklärung nach Nagelkerke.

#### 5.2.4 Zusammenfassung der Ergebnisse und Hypothesenprüfung

Vergleicht man die neu entwickelte DIA-S mit der GDS15 auf Itemebene, so ergeben sich hinsichtlich der durchschnittlichen Übereinstimmung mit dem Außenkriterium sowie der durchschnittlichen Trennschärfe jeweils Kennwerte, die für eine bessere empirische Qualität der neu entwickelten Items sprechen. Dabei enthalten beide Skalen vergleichbar viele Items in einem mittleren und hohem Schwierigkeitsbereich. Die drei Items der GDS15 mit geringen Schwierigkeiten (Items 02, 09 und 13) weisen bei Übereinstimmungen mit dem Außenkriterium von unter 30% jeweils den vermuteten Zusammenhang zwischen inhaltlichen Ungenauigkeiten, geringer Schwierigkeit und niedriger Validität auf (Unterstützung von Hypothese 01). Eine Ausnahme bildet mit zufriedenstellenden Kennwerten lediglich Item 04. Bei einem Vergleich der Verfahren auf Skalenebene zeigen sich für die DIA-S jeweils höhere Werte in der internen Konsistenz, in der Korrelation mit der MADRS sowie in der Skaleneffektivität, wobei sich neben einer signifikant höheren Spezifität auch eine tendenziell höhere Sensitivität ergibt (Unterstützung von Hypothese 2, Hypothese 3a und Hypothese 3b). Während die Wahrscheinlichkeit von Fehlklassifikationen bei beiden Testverfahren erwartungsgemäß mit zunehmender Klassifizierungsschwierigkeit ansteigt, ergibt sich für die GDS15 eine höhere Wahrscheinlichkeit von Fehldiagnosen mit höherem Alter der Probanden. Nach Fehlertypen getrennte Analysen von Einflussfaktoren ergeben für die GDS15 einen Zusammenhang von höherem Alter und falsch positiver Klassifizierung sowie einem hohen Mentalstatus und falsch negativer Klassifizierung. Ein Zusammenhang zwischen dem Geschlecht des Probanden und falsch negativer Klassifizierung durch die DIA-S lässt sich durch die Absenkung des Trennwertes der Skala um einen Punkt für die männlichen Probanden dahingehend ausgleichen, dass Sensitivität und Spezifität des neuen Tests für beide Geschlechter besser ausfallen als für die GDS15. Für die GDS15 lässt sich der Gendereffekt nicht durch eine Absenkung des Trennwertes der Skala ausgleichen. Ein Einfluss geringerer Tageslichtdauer auf die Wahrscheinlichkeit falsch

positiver Ergebnisse durch die DIA-S ergibt sich nur für eine Teilstichprobe des Datensatzes.  
Auch Hypothese 3c wird somit durch den Datensatz weitgehend unterstützt.

## 6. Diskussion

### 6.1 Interpretation der Ergebnisse

In der vorliegenden Studie werden die neu entwickelte DIA-S und die Geriatrische Depressionsskala (GDS15) im Hinblick auf ihre Testperformanz in der stationären Geriatrie miteinander verglichen. Die Stichprobe entspricht dabei in allen relevanten Merkmalen der Population, für die der neue Test intendiert ist und es wurde ein Untersuchungsdesign gewählt, das eine strenge empirische Prüfung der Skalen ermöglicht (vgl. Pepe, 2004). So wurden zur empirischen Prüfung der Skalen nicht bereits vordiagnostizierte Patienten mit besonders eindeutiger Symptomatik ausgewählt, um die Ergebnisse nicht im Sinne eines so genannten Spektrum-Bias positiv zu verzerren. Zudem wurden kognitiv eingeschränkte Patienten in die Analysen einbezogen, deren Depressionsstatus durch Screeningverfahren schwieriger zu ermitteln ist. Dadurch erklärt sich einerseits die Tatsache, dass die Ergebnisse für Sensitivität und Spezifität der GDS15 etwas niedriger ausfallen, als es in einer aktuellen Metastudie berichtet wird (Wancata et al., 2006). Andererseits ist durch diese Art der Prüfung davon auszugehen, dass die ermittelten Qualitätskriterien mit hoher Wahrscheinlichkeit auch der Performanz der Skalen in der Praxis entsprechen. Die Ergebnisse sprechen für bessere diagnostische Qualitäten der DIA-S, wobei sich der Unterschied sowohl auf der Ebene der einzelnen Items als auch im Gesamtvergleich der Skalen ergibt.

#### 6.1.1 Itemanalysen

Die theoretischen Überlegungen dieser Arbeit werden besonders durch die Ergebnisse der Itemanalysen unterstützt. So sind besonders diejenigen Items der GDS15 von geringem diagnostischen Wert, die inhaltlich schlecht zwischen Depressivität, körperlichen Bedingungen und den gegenwärtigen Lebensumständen der Probanden differenzieren (z.B. „Fühlen Sie sich voller Energie?“). Andererseits sind die Qualitätsunterschiede in den Skalen nicht lediglich auf einzelne, für den Kontext unpassende Fragen zurückzuführen, sondern es zeigt sich ein verhältnismäßig konstanter Effekt je nach Skalenzugehörigkeit. Zudem ergibt

sich für die GDS15 eine relativ hohe Anzahl an Items, die eine bedeutsame Anzahl fehlender Werte aufweisen. Da die entsprechenden Items auch bereits in einer Voruntersuchung zu dieser Arbeit unter anderen Erhebungsbedingungen von vielen Probanden nicht beantwortet wurden (vgl. Heidenblut & Zank, 2005), erscheint ein Zusammenhang mit dem Inhalt naheliegend. So handelt es sich überwiegend um Fragen, die zwar eindeutig die Stimmung der Probanden erfassen (z.B. „Sind Sie die meiste Zeit guter Laune?“, „Fühlen Sie sich die meiste Zeit über glücklich?“) und entsprechend befriedigende Werte in den Itemvaliditäten aufweisen, die jedoch so niederschwellig formuliert sind, dass Sie von Patienten, die sich gerade in einer schwierigen Lebenssituation befinden, leicht als unpassend empfunden werden können.

Wenn man sich über die rein psychometrischen Eigenschaften der Items hinaus die diagnostische Situation einer Fragebogenerhebung vorstellt, so ist es relativ wahrscheinlich, dass die Konfrontation mit Fragen, von denen sich Probanden möglicherweise nicht ernstgenommen fühlen, zu einer Art negativem Halo-Effekt führen kann, der die Auskunftsbereitschaft des Befragten insgesamt reduziert. Unter Berücksichtigung dieser Überlegung erscheint das Ergebnis eines durchgängigen Qualitätseffektes der Items je nach Skalenzugehörigkeit weniger überraschend.

Es ist also davon auszugehen, dass die Validität eines Verfahrens in der Praxis stark von der Akzeptanz abhängt, die es sowohl bei Mitarbeitern als auch bei der intendierten Zielgruppe der Patienten erfährt. Die inhaltliche Angemessenheit eines Items ist dabei nicht nur eine Frage der fairen Testung, sondern sie ist ein notwendiges Kriterium der psychometrischen Qualität. Entsprechend sollte die inhaltliche Passung von einzelnen Items und der Gesamtsituation der diagnostischen Erhebung ein Kriterium sein, das während der Testkonstruktion stärker berücksichtigt werden sollte. Denn die Aussagekraft der meisten Validierungsstudien für die tatsächliche Performanz eines Verfahrens unter Praxisbedingungen ist in der bisherigen Forschung eine relativ unbekanntes Größe. Stattdessen

werden viele Verfahren unter optimalen personellen und situativen Bedingungen getestet, woraus dann Werte ermittelt werden, die als Qualitätsnachweise für die praktische Anwendbarkeit gelten. Es bedarf keines besonders hohen Maßes an Phantasie, um sich vorzustellen, dass diese Verfahrensweise häufig zu falschen Annahmen führt. Praktische Anwendbarkeit, Testfairness, Transparenz und Akzeptanz von Testverfahren sind also Gütekriterien, die bei vielen Testentwicklungen eine wesentlich zentralere Rolle spielen sollten.

### *6.1.2 Skalenreliabilität*

Ein Vergleich der Testreliabilitäten ergibt eine höhere Zuverlässigkeit der neuen Skala, wobei der Unterschied zwischen den Skalen augenscheinlich nicht sehr groß wirkt (GDS15:  $\alpha = .75$ , DIA-S:  $\alpha = .84$ ). Dabei ist allerdings zu beachten, dass Skalen von unterschiedlicher Länge hinsichtlich ihrer Messgenauigkeit nur schwer zu vergleichen sind. So lässt sich die Reliabilität eines Tests nach den Axiomen der Klassischen Testtheorie als Anteil der wahren Varianz an der durch das Messverfahren ermittelten Gesamtvarianz beschreiben (Amelang & Schmidt-Atzert, 2006). Dabei wird vorausgesetzt, dass eine perfekte Korrelation zwischen den wahren Werten eines Tests besteht, während die Messfehler jeweils unkorreliert sind. Inhaltlich bedeutet dies, dass ohne Messfehler eine vollkommene Übereinstimmung zwischen den Testwerten bestehen würde, während Fehler sich bei wiederholter Messung gegenseitig ausgleichen. Aus dieser Annahme lässt sich ein Zusammenhang zwischen der Länge eines Tests und dessen Zuverlässigkeit ableiten, der mathematisch durch die sogenannte Spearman-Brown-Formel ausgedrückt wird.

Mithilfe dieser Formel kann man schätzen, wie sich die Reliabilität eines Tests verringern würde, wenn man ihn um eine bestimmte Anzahl an Items kürzen würde:

$$\text{corr } r_{tt} = \frac{n \times r_{tt}}{1 + (n - 1)r_{tt}} \quad (1)$$

mit

$n$  = Anzahl der Items vor der Testkürzung /Anzahl der Items nach der Testkürzung.

$\text{corr } r_{tt}$  = Reliabilität nach der Testkürzung

$r_{tt}$  = Reliabilität vor der Testkürzung

(vgl. Amelang und Schmidt-Atzert, 2006, S.47)

*Formel (1): die Spearman-Brown-Formel*

Dabei ist zu beachten, dass die Spearman-Brown-Formel zwar eine rein theoretische Ableitung aus den Axiomen der KTT darstellt und ihre Anwendung daher nicht bewiesen werden kann. Durch Simulationsstudien (Cortina, 1993) lässt sich der Effekt jedoch gut empirisch im Sinne eines erheblichen Einflusses der Testlänge auf den Alpha-Koeffizienten nachweisen.

Eine Anwendung von Gleichung 1 auf den Vergleich der Skalenreliabilitäten der vorliegenden Untersuchung würde bedeuten, dass sich die Zuverlässigkeit der GDS15 bei einer Kürzung auf die Länge der DIA-S, also auf 10 Items, von  $\alpha = .75$  auf  $\alpha = .67$  vermindern würde, was bereits einer relativ ungenauen Messung entspricht (Fisseni, 2004). Unter Berücksichtigung dieser Überlegung ergibt sich also bereits ein wesentlich bedeutsamerer Unterschied zwischen den Skalen. Die Annahme, dass eine Kürzung der GDS zu einer deutlichen Reduktion der Internen Konsistenz führen würde, wird auch durch empirische Daten bestärkt. So werden in der Literatur für kürzere Versionen der Skala mit vier bzw. fünf Items Alpha-Koeffizienten von .53 bzw. .49 berichtet (Goring et al., 2004; Chattat, Ellena, Cucinotta, Savorani & Mucciarelli, 2001). Die hohe Interne Konsistenz der

DIA-S deutet dagegen darauf hin, dass eventuell eine weitere Kürzung der Skala für bestimmte Anlässe möglich ist, ohne dass die Messgenauigkeit des Instrumentes dadurch grundsätzlich infrage gestellt wäre. So würde sich beispielsweise die Interne Konsistenz nach Gleichung 1 bei einer Kürzung der DIA-S auf fünf Items auf  $\alpha = .72$  reduzieren, was einer zufriedenstellenden Zuverlässigkeit entspricht.

Neben der numerischen Beurteilung der Internen Konsistenz anhand von Richtwerten ist die Übereinstimmung der Items auch inhaltlich von wichtiger Bedeutung und stellt eine Voraussetzung für die Testvalidität dar. So stellt sich bei Skalen, die zu Summenscores zusammengefasst werden, jeweils die Frage, ob die Aufsummierung der einzelnen Itembeantwortungen zu einem gemeinsamen Wert inhaltlich gerechtfertigt werden kann. Eine hinreichende Iteminterkorrelation ist dabei eine notwendige Bedingung dafür, dass durch die Items dasselbe Konstrukt gemessen wird. Ein hoher Alpha-Koeffizient bedeutet in diesem Zusammenhang, dass ein Test wenig Item-spezifische Varianz aufweist (Cortina, 1993), er ist jedoch keine hinreichende Bedingung für die Eindimensionalität einer Skala (Cortina, 1993, Schmitt, 1996). Im Hinblick auf die DIA-S erscheint die Zusammenfassung der einzelnen Itembeantwortungen zu einem Summenscore unter Berücksichtigung der verhältnismäßig homogenen Interkorrelationsmatrix der Items dabei eher zulässig als bei der Geriatrischen Depressionsskala, bei der einige Items trotz thematisch gleichgesetzter Polung negativ miteinander korrelieren. Dennoch wäre bei einer weiteren empirischen Bewährung des neuen Instrumentes eine Analyse der Faktorstruktur der DIA-S empfehlenswert.

### *6.1.3 Testeffektivität*

In der Testeffektivität erzielt die DIA-S bessere Ergebnisse als die GDS15, wobei sich die Überlegenheit der neuen Skala erwartungsgemäß vor allem in Spezifität zeigt. Dabei kann die Bedeutung von Unterschieden in der diagnostischen Treffsicherheit von Testverfahren für die klinische Praxis anhand der Kennwerte Sensitivität und Spezifität nur bedingt abgelesen werden. So stellen Sensitivität und Spezifität jeweils die bedingte Wahrscheinlichkeit eines

bestimmten Testergebnisses bei Vorliegen bzw. Nicht-Vorliegen der wahren Erkrankung dar (Wirtz & Nachtigall, 1998). Im klinischen Alltag interessiert dagegen eher die positive bzw. negative Vorhersagegenauigkeit eines Tests (PPV = positive predictive value, NPV = negative predictive value), d.h. die Wahrscheinlichkeiten mit denen bei Vorliegen eines Testergebnisses auch auf den zugrundeliegenden Erkrankungszustand geschlossen werden kann. Diese sind, wie aus Gleichung 2 und Gleichung 3 ersichtlich wird, nicht nur von den Testeigenschaften sondern auch von der Prävalenz des Merkmals in der untersuchten Population abhängig:

$$PPV = \frac{\text{Sensitivität} \times \text{Prävalenz}}{\text{Sensitivität} \times \text{Prävalenz} + (1 - \text{Spezifität}) \times (1 - \text{Prävalenz})} \quad (2)$$

$$NPV = \frac{\text{Spezifität} \times (1 - \text{Prävalenz})}{(1 - \text{Sensitivität}) \times \text{Prävalenz} + \text{Spezifität} \times (1 - \text{Prävalenz})} \quad (3)$$

(vgl. Altman & Bland, 1994, S. 309)

*Formel (2), (3): Die Anwendung des Base-Theorems auf negative und positive predictive values*

D.h., mit steigender Prävalenz des Merkmals steigt die positive Vorhersagegenauigkeit bei gleichbleibenden Eigenschaften des Testverfahrens an, während die negative Vorhersagegenauigkeit absinkt. Da die Prävalenz depressiver Erkrankungen in der vorliegenden Stichprobe aufgrund der Schichtung der Stichprobe künstlich angehoben wurde, wäre eine Schätzung der jeweiligen Vorhersagegenauigkeit der Testverfahren auf der Grundlage der vorliegenden Daten entsprechend wenig aussagekräftig. Würde man eine Punktprävalenz depressiver Erkrankungen von 30% zugrunde legen, wie Sie durch eine 2005 veröffentlichte große Prävalenzstudie für stationäre geriatrischen Patienten ermittelt wurde (Mc Cusker et al., 2005), so läge die positive Vorhersagegenauigkeit für die GDS15 bei 53%

und für die DIA-S bei 63%, während die negative Vorhersagegenauigkeit für die GDS15 87% und für die DIA-S 91,% betragen würde. Anhand dieser Ergebnisse wird deutlich, dass die, nennen wir sie gefühlte Qualität eines Testverfahrens, stark von den jeweils berichteten relativen Häufigkeiten des vier-Felder-Schemas, wie es in Abbildung 17 gezeigt wird, abhängt. Wenn die Wahrscheinlichkeit eines Testergebnisses bei gegebener Erkrankung erfragt wird, so hat dies den Vorteil, dass damit ein weitgehend stichprobenunabhängiges Urteil über den Test gefällt werden kann. Es hat allerdings den Nachteil, dass dann im Fall einer einzelnen Testung daraus nur nach Anwendung von Formel 2 eine Aussage über die Wahrscheinlichkeit einer manifesten Erkrankung gemacht werden kann. Dabei gibt es zahlreiche Untersuchungen darüber, dass die Kennwerte Sensitivität und Spezifität von praktischen Ärzten immer wieder mit der positiven bzw. negativen Vorhersagegenauigkeit verwechselt werden, was in der Praxis verheerende Folgen haben kann. Gigeranzer (1999) weist darauf hin, dass diesem Problem z.T. dadurch begegnet werden kann, dass zur Darstellung von Ergebnissen eher auf relative Häufigkeiten statt auf Prozentwerte verwiesen werden sollte.

Wirft man einen Blick auf die Vorhersagegenauigkeit der beiden Skalen, so wirken die 53% der GDS15 aber auch die 63% der neuen Skala zunächst verhältnismäßig gering, wenn man bedenkt, dass die korrekte Erfassung erkrankter Subjekte das wichtigste Testmerkmal darstellt. Wird das Zahlenbeispiel jedoch in relativen Häufigkeiten ausgedrückt, so ergibt sich bei einer Gesamtanzahl von 100 Probanden eine Anzahl von 30 erkrankten und 70 gesunden Probanden. Die DIA-S würde davon 25 der erkrankten Patienten erfassen, die GDS15 23. Bei der DIA-S hätten darüber hinaus 15 der gesunden Patienten ein auffälliges Testergebnis bei der GDS15 wären es mit 20 Patienten ein Viertel mehr. Insgesamt würde sich also durch den Einsatz des neuen Verfahrens der Anteil fehlerhaft eingeordneter Patienten von 27% auf 20% reduzieren. Das wären bei 100 neu diagnostizierten Patienten pro Monat aufs Jahr gerechnet 94 Patienten, die durch das neue Verfahren profitieren würden.

Dabei handelt es sich bei einer solchen Rechnung natürlich zunächst um Spekulation; es lässt sich jedoch zeigen, dass auch verhältnismäßig gering wirkende Unterschiede schon eine hohe Praxisrelevanz aufweisen können.

#### *6.1.3.1 Geschlechtsunterschiede in der Testeffektivität*

Obwohl das Geschlecht der Patienten keinen direkten Einfluss auf die Klassifikationsraten in der logistischen Regression für die Gesamtergebnisse aufweist, zeigen sich Einflüsse auf Sensitivität und Spezifität der Verfahren. So ergibt sich für männliche Probanden generell eine stärkere Tendenz zu falsch negativen Ergebnissen während die Selbsteinschätzung weiblicher Probanden häufiger zu falsch positiven Testergebnissen führt. Diese Tendenz ist intuitiv gut nachvollziehbar, wenn man berücksichtigt, dass das freimütige Äußern aversiver Gefühle Frauen in unserer Gesellschaft eher zugestanden wird als Männern, da sie mit der weiblichen Geschlechtsrolle wesentlich leichter vereinbar sind. Methodisch lässt sich diesem Trend, der in der Literatur bereits mehrfach beschrieben wurde (Camozzato, Hidalgo, Souza, & Chaves, 2007; Sigmon et al., 2005; Semmler & Klumb, 2004) für die DIA-S dahingehend entgegenwirken, als eine Absenkung des Trennwertes für die männliche Teilstichprobe für beide Geschlechter zu hohen Kennwerten in Sensitivität und Spezifität führt. Bei der GDS15 war eine entsprechende Maßnahme nicht möglich.

Die meisten Screeningverfahren zur Erfassung depressiver Symptome haben geschlechtsunabhängige Normwerte, was bei einem bekannten deutlichen Gendereffekt im Hinblick auf das Artikulieren von Emotionen eigentlich eine überraschende Tatsache ist. Eine systematische Untersuchung des Antwortverhaltens von Männern und Frauen im Hinblick auf verschiedene Selbstbeurteilungsverfahren, die Depressionen bzw. andere affektive Erkrankungen betreffen, wäre also ein Forschungsfeld, was sowohl die akademische als auch die klinische Diagnostik voran bringen könnte.

#### *6.1.4 Konvergente Validität*

Vergleicht man die Korrelationskoeffizienten in der Gesamtheit miteinander, so fällt einerseits auf, dass der Zusammenhang der neuen Skala mit der GDS15 unter Berücksichtigung der formalen Ähnlichkeit der Skalen sowie der Übereinstimmung des diagnostischen Aufgabengebietes verhältnismäßig gering ausfällt. Andererseits erscheint der Zusammenhang zwischen DIA-S und MADR-Skala trotz der Unterschiede in den Verfahren relativ hoch. Diese Ergebnisse sprechen dafür, dass die intendierte inhaltliche Abweichung der neuen Skala von der GDS15 realisiert werden konnte. Darüber hinaus ist die gute Übereinstimmung von DIA-S und MADR-S ein Hinweis dafür, dass die neu entwickelte Skala neben einer Grobeinschätzung des Krankheitsstatus eines Patienten auch Hinweise auf den Schweregrad der Symptomatik liefern könnte.

### *6.2 Überlegungen zur Anwendung der Skala*

#### *6.2.1 Klinische Anwendbarkeit der Skala.*

Die DIA-S wurde als Screeningverfahren konzipiert, das leicht und praktikabel einsetzbar sein soll. Durch den geringen Umfang der Skala, deren Beantwortung etwa zehn Minuten benötigt und durch das einfache Auswertungsschema erwies sich die DIA-S in den Untersuchungen als leicht handhabbar. Bei der Erhebung der Skala in der Praxis sollte allerdings berücksichtigt werden, dass mündlich erhobene Selbstaussagen zur Stimmung für die Patienten ein schwieriges Gesprächsthema darstellen, das mit der nötigen Sensibilität gehandhabt werden und ggfs. in einen breiteren diagnostischen Kontext, etwa ein Anamnesegespräch, gestellt werden sollte. So ist eine tragfähige Beziehung zwischen Diagnostiker und Patient Voraussetzung für eine faire Testung und somit auch für die Adhärenz eines Klienten zu einer ehrlichen Beantwortung der Items. Die Skalen wurden bisher nur an einem der drei Erhebungsorte durch Klinikmitarbeiter erhoben, die die Handhabbarkeit des Instrumentes grundsätzlich als gut bewerteten, und die die Skala nach Beendigung des Forschungsprojektes auch bereits in ihren klinischen Alltag implementiert

haben. Bisher liegen jedoch keine systematischen Angaben zur Einschätzung der klinischen Anwendbarkeit der Skala durch stationäre Einrichtungen vor. Im Rahmen einer Implementierung des Instrumentes in die Praxis wäre eine entsprechende Erhebung empfehlenswert. Darüber hinaus erscheint die Erstellung eines Praxisratgebers zur Erfassung depressiver Störungen in der Geriatrie eine lohnenswerte Aufgabe zu sein.

### *6.2.2 Stellenwert des Instrumentes im diagnostischen Prozess.*

Die DIA-S stellt ein klassisches Screeningverfahren zur Ermittlung depressiver Störungen dar. D.h., das Instrument soll breit und ökonomisch anwendbar sein und möglichst auch geringer ausgeprägte Erkrankungen anzeigen. Das bedeutet gleichzeitig, dass darüber hinausgehende diagnostische Fragestellungen durch die Skala nicht beantwortet werden können, da hierdurch die Screeningqualität des Instrumentes beeinträchtigt würde. So würde ein systematisches Abfragen von Suizidalität die Trennschärfe des Instruments mindern und Items zur differentialdiagnostischen Abklärung dementieller Syndrome oder von Angsterkrankungen würden sich negativ auf die Zuverlässigkeit und Ökonomie der Skala auswirken. Bei der Anwendung in der Praxis ist daher zu berücksichtigen, dass im Falle eines auffälligen Testergebnisses in der Regel weitere Diagnostik erforderlich ist. So sollten Fragestellungen wie die Einschätzung von Suizidalität affektiv auffälliger Patienten oder eine differentialdiagnostische Unterscheidung zwischen einer depressiven Erkrankung und einer Angsterkrankung im Anschluss an die Erhebung der Skala klinisch erfolgen.

## *6.3 Grenzen der vorliegenden Arbeit und Empfehlungen für künftige Untersuchungen*

### *6.3.1 Entwicklung des Instrumentes*

Bei der Entwicklung der DIA-S mussten neben methodischen Überlegungen immer auch forschungspraktische Einschränkungen berücksichtigt werden. Diese bestanden zum einen in der zeitlichen Beschränkung der Erhebungsphase, in einer begrenzten Anzahl kooperierender Kliniken und in einer begrenzten Anzahl an vorhandenem Forschungspersonal. So erfolgte die Hauptvalidierungsstudie bereits mit einer auf 10 Items reduzierten Skala, obwohl ein Pool

mit 31 inhaltsvaliden Items zur Verfügung stand. Die Itemanalysen, auf deren Grundlage die Gesamtskala zusammengestellt wurde, beruhte insgesamt auf einer Stichprobe von N=44 Probanden, also auf einem verhältnismäßig kleinen Datensatz. Dabei wurde versucht, dem geringen Umfang der Stichprobe durch eine Schichtung der depressiven Krankheitsbilder zumindest teilweise entgegenzuwirken.

Eine Haupterhebung, die sich lediglich auf 10 Items bezieht, kann als riskantes Unternehmen angesehen werden. Eine Alternative wäre die Erhebung des Pools von 31 Items gewesen. Im Anschluss an die Haupterhebung hätte man daraufhin eine kurze Skala aus den besten Items zusammenstellen können. Das hätte allerdings den Nachteil gehabt, dass es für diese Skala nach Abschluss des Forschungsprojektes keine echten Normwerte gegeben hätte und somit eine weitere Validierungsstudie notwendig gewesen wäre, um die Skala in der Praxis einzusetzen. So hat nämlich der Erhebungskontext eines Items, wie bereits bei der Interpretation der Ergebnisse diskutiert wurde, entgegen den Axiomen der KTT durchaus einen Einfluss auf dessen Beantwortung. Wäre dem nicht so, so wäre auch die Reihenfolge, in der einzelne Items präsentiert werden, irrelevant. Da jedoch zahlreiche Lehrbücher (z.B. Fisseni, 2004; Porst, 2000) beispielsweise auf ein sinnvolles Verhältnis von positiv und negativ gepolten Items hinweisen, bzw. auch sogenannte Icebraker-Fragen in der Praxis üblich sind, muss von einem Einfluss des Gesamtfragebogens auf die Beantwortung von Einzelitems ausgegangen werden. Die Entwicklungsstudie der GDS (Yesavage et al., 1983) hat diese Tatsache nicht berücksichtigt. Somit ist wahrscheinlich schon in der ersten Validierungsstudie der GDS, die aus der Validierung eines Gesamitempools von insgesamt 100 Fragen bestand, der Grundstein für eine systematische methodische Überschätzung dieses Verfahrens gelegt worden.

Dennoch wäre die Erprobung eines größeren Pools von Items für die neue Skala an einer größeren Stichprobe sinnvoll gewesen. So ist das Konzept, inhaltlich stärker auf eine depressive Symptomatik abzielende und somit schwerer zu lösende Items zu konzipieren,

zwar aufgegangen, aber der Preis der hohen Spezifität der Skala hätte leicht in einer zu geringen Sensitivität liegen können. Dem war nicht so. Et hät noch immer jot jejang.

Die Tatsache allerdings, dass bereits die erste Validitätsstudie mit entsprechenden Vergleichsdaten (Heidenblut & Zank, 2009) veröffentlicht werden konnte, kommt der Implementierung des Instrumentes in Forschung und Praxis zugute (vgl. Interessenten der DIA-S, im Anhang).

### *6.3.2 Einflussfaktoren auf die Validität der Testverfahren*

Der Fokus der vorliegenden Untersuchung lag in erster Linie auf dem Vergleich der neu entwickelten DIA-S und der GDS15 hinsichtlich ihrer Übereinstimmung mit dem Außenkriterium. Bei den darüber hinaus erhobenen Daten zur kognitiven und funktionalen Leistungsfähigkeit der Probanden handelt es sich um routinemäßig erhobenes Datenmaterial, das der Studie von den kooperierenden Kliniken zur Verfügung gestellt wurde. Die Erhebung entsprechender Daten wird dabei in erster Linie durch die Anforderungen des klinischen Alltags bestimmt. So konnte beispielsweise nur bedingt Einfluss darauf genommen werden, dass die von den Kliniken erhobenen Daten zum kognitiven Status der Probanden sowie zur funktionellen Selbstständigkeit zeitnah zu den Terminen der Depressionsdiagnostik zur Verfügung standen. Dieses Problem stellt sich für das Datenmaterial zum Barthel-Index stärker als für die MMSE-Messungen, da die funktionelle Selbstständigkeit der Patienten im Gegensatz zum Mentalstatus ein Merkmal ist, dass sich bei vielen Patienten innerhalb kurzer Zeitabstände erheblich verändern kann. Entsprechend ist fraglich, inwiefern die vorliegenden Daten zum Barthel-Index und z.T. auch diejenigen zum MMSE reliable und valide Messungen für die intendierte Untersuchung darstellen.

Darüber hinaus stehen vollständige Daten der Probanden nur für einen Teildatensatz zur Verfügung, sodass die Teststärke der multivariaten Analysen reduziert ist. Entsprechend wird der Einfluss der kognitiven und funktionellen Selbstständigkeit auf die Testvalidität der Skalen möglicherweise unterschätzt. Zudem wären theoretisch weitere mögliche

Einflussfaktoren auf das Antwortverhalten der Probanden denkbar, wie z.B. die Länge des gegenwärtigen Krankenhausaufenthaltes zum Zeitpunkt der Befragung oder die Anzahl chronischer Erkrankungen etc. Dabei stellt die systematische Untersuchung von Faktoren, die die Validität von Testverfahren im Bereich der Depressionsdiagnostik beeinträchtigen können, einen Forschungsbereich dar, der zwar einige praktische Relevanz hat, bisher jedoch nur durch wenige Studien angegangen wurde. So ließen sich entsprechende Analysen relativ leicht in Validierungsstudien implementieren, wenn die Erhebung der interessierenden Zielvariablen bei der Versuchsplanung berücksichtigt werden könnte.

#### *6.3.4 Neutralität der Untersuchungssituation*

Die Erhebung der DIA-S sowie der GDS15 erfolgte durch Personal des Forschungsprojektes bzw. in Ausnahmefällen durch Klinikpersonal. Dabei wurden die Untersucher zwar einerseits hinsichtlich der Bedeutung eines möglichst neutralen Auftretens gegenüber den Probanden geschult, andererseits war es aus forschungspraktischen Gründen nicht möglich, sie über die Hypothesen der Arbeit im Unklaren zu lassen. Entsprechend wären Versuchsleitereffekte zugunsten der neu entwickelten DIA-Skala durch die Art der Datenerhebung nicht auszuschließen. Da sich die Hypothesen der Untersuchung jedoch nicht auf ein bestimmtes Antwortverhalten der Teilnehmer an sich beziehen, sondern auf ein Antwortverhalten in Abhängigkeit zum wahren Depressionsstatus der Probanden, ist jedoch nicht von eindeutigen Wirkungen zugunsten der neuen Skala auszugehen. So wusste der entsprechende Untersucher zum Zeitpunkt der Befragung nicht, ob bei dem Teilnehmer laut diagnostischem Kriterium eine depressive Verstimmung vorlag und konnte das Antwortverhalten des Teilnehmers weder bewusst noch unbewusst in die „richtige“ Richtung lenken.

Dennoch wäre es wünschenswert, weitere Validierungsstudien zur DIA-S beziehungsweise Vergleichsstudien der Skala mit anderen Messinstrumenten jeweils durch Personal durchführen zu lassen, dass keine Vorerwartung bezüglich der diagnostischen Qualität der einzelnen Skalen hat.

#### 6.3.4 *Setting der Untersuchungssituation*

Die Ergebnisse der vorliegenden Arbeit beziehen sich auf stationäre geriatrische Patienten. Dabei wurde die DIA-S so konstruiert, dass es theoretisch keine Einschränkung zur Verwendung der Skala in anderen klinischen Bereichen, etwa in tagesklinischen Einrichtungen oder in ambulanten Bereichen bei Haus- oder Fachärzten gibt. Eine Erprobung der Skala an entsprechenden Stichproben wäre wünschenswert.

#### 6.4 *Ergänzende Überlegungen zur Depressionsdiagnostik*

Die Ergebnisse dieser Arbeit sind im Kontext des aktuellen wissenschaftlichen Konsenses zur Klassifikation psychischer Störungen zu sehen. So erfolgte die Erhebung der klinischen Diagnosen als Außenkriterium nach den Richtlinien der ICD-10. In diesem Klassifikationssystem werden Erkrankungen im höheren Lebensalter nicht gesondert klassifiziert bzw. beschrieben. Die Setzung einer ICD-10- Diagnose als „echtes“ Außenkriterium und somit als „wahres“ Urteil über das Ausmaß einer Störung ist also nicht unproblematisch, denn in der Forschungsliteratur kontrovers diskutierte Besonderheiten des Krankheitsbildes depressiver Störungen im Alter finden dabei keine Berücksichtigung.

So beschreibt beispielsweise Ernst (1997) den epidemiologischen Widerspruch zwischen im Alter vermehrt auftretenden depressiven Symptomen und sinkenden Prävalenzraten für krankheitswertige Störungen und führt diesen unter anderem auf die unzureichende Beschreibung im Alter häufig auftretender Depressionssyndrome durch die ICD-10 zurück. Auch Davidson und Neale (2002) weisen darauf hin, dass die gängigen Standardkriterien der gegenwärtigen Klassifikationssysteme unter Umständen zur Erfassung der bei geriatrischen Patienten verbreiteten Krankheitsbilder nicht ausreichen. So kann beispielsweise die in der Forschungsliteratur vielfach beschriebene subdiagnostische depressive Symptomatik (Helmchen et al., 1996, Heuft, Kruse & Radebold, 2000) durch eine ICD-10-Diagnose nicht abgebildet werden. Zudem werden in der Literatur die diagnostischen Kriterien für depressive Störungen bei gleichzeitigem Vorliegen von koronaren

Herzerkrankungen kritisiert (Hasnain, Vieweg, Lesnefsky & Pedurangi, 2011), bzw. es werden Subtypen depressiver Erscheinungsbilder bei körperlich kranken Patienten beschrieben (Guidi et al., 2011; Thorman, Himmerich & Steinberg, 2011).

Die durch die Festlegung des Außenkriteriums getroffenen Entscheidungen für die diagnostischen Kriterien der Störung sind jedoch von erheblicher praktischer Relevanz für die Beurteilung eines Testverfahrens. Denn der empfohlene Cutoff-Score eines Messinstrumentes wird nach empirischen Analysen auf denjenigen Wert festgesetzt, durch den die meisten Übereinstimmungen mit dem Außenkriterium erreicht werden. Steht die diagnostische Validität dieses Kriteriums infrage, so muss auch die Gültigkeit des Trennwertes bzw. des Testverfahrens unter Vorbehalt gesehen werden. Die Entscheidung über den Krankheitswert einer depressiven Verstimmung stellt somit ein erkenntnistheoretisches Problem dar, dass nicht mit vollkommener Sicherheit gelöst werden kann. Hautzinger (1988) beschreibt dieses Problem wie folgt: Die Frage, „wann und wodurch die Grenze zwischen . . . normalen Reaktionen und den als klinisch auffällig betrachteten Symptomen überschritten wird, gehört unverändert zu den ungelösten Fragen im Zusammenhang mit depressiven Störungen“ (S.3). Das Ergebnis eines Testverfahrens kann dem klinischen Diagnostiker diese Frage nicht zweifelsfrei beantworten; ein zuverlässiges Messinstrument kann jedoch einen wichtigen Beitrag zur Erstellung gültiger Diagnosen und somit zu einer besseren Behandlung der Erkrankten leisten. Die Gültigkeit von Testverfahren und der ihnen übergeordneten Konstrukte sollte aus diesem Grund immer wieder überprüft werden.

#### *6.4.1 Das Recht eines Patienten, die Testung zu verweigern*

Die Multidisziplinarität in Behandlung und Diagnostik gilt heutzutage als ein unumstößliches Paradigma in der geriatrischen Praxis. Was bedeutet es aber beispielsweise für einen ansonsten gesunden Patienten, der aufgrund eines Unfalls in die stationäre Geriatrie eingeliefert wird? Dieser Patient durchläuft dann eine Reihe von Untersuchungen, mit denen überprüft wird, ob er von häufigen geriatrischen Erkrankungen wie etwa Demenz oder

Depression etc. betroffen ist. Dabei hätte der Betroffene vielleicht im Fall einer psychischen Komorbidität für einige Erkrankungen, die bei ihm im Rahmen der Gesamtdiagnostik festgestellt werden, gar keinem Behandlungsauftrag zugestimmt.

Die ganzheitliche Erfassung der Erkrankungs- und Lebenssituation eines Patienten ist natürlich grundsätzlich eine sinnvolle Einrichtung des geriatrischen Versorgungssystems, trägt sie doch der Tatsache der Multimorbidität vieler ihrer Klienten in adäquater Weise Rechnung. Eine mögliche Nebenwirkung dieser effektiven Verfahrensweise liegt jedoch in der potentiellen Entmündigung des Patienten, für den ohne Einbeziehung seiner eigenen Sichtweise entschieden wird, was für ihn das Beste sei. Wie ist dieses Dilemma in der Praxis zu lösen? In der diagnostischen Praxis ergibt sich immer wieder die Erkenntnis, dass ein möglichst transparenter Umgang mit dem Klienten dazu führt, dass die meisten Untersuchungen problemlos durchgeführt werden können. So kann ein Hinweis auf die strukturellen Bedingungen der Klinik und die sich daran anschließende Tatsache, dass routinemäßig nach diversen Erkrankungen gescreened wird, die Verwunderung der meisten Patienten darüber, dass sie nicht nur zu ihrer Stimmung befragt werden, sondern auch noch einen „Idiotentest“ (MMSE) machen müssen, erheblich abgemildert werden.

Dennoch sollte dem Recht des Patienten, eine solche Routineuntersuchung zu verweigern, trotz dem Druck von Krankenkassen und ihren entsprechenden Abrechnungssystemen bei Bedarf umgehend stattgegeben werden. Denn ein beispielsweise kognitiv und affektiv unbeeinträchtigter hochaltriger Patient, der ahnt, dass seine weitere Lebenszeit begrenzt ist, möchte wohlmöglich seine verbleibende Zeit nicht mit der Beantwortung von Assessmentverfahren verbringen.

Entsprechend sei für alle Bereiche der Testdiagnostik darauf hingewiesen, dass kein Verfahren den gesunden Menschenverstand bzw. ein menschliches Einfühlungsvermögen ersetzen kann und sollte.

#### *6.4.2 Das Verhältnis von Diagnostik und Behandlung*

Obwohl es zahlreiche Bemühungen gibt, die psychotherapeutische Behandlung geriatrischer Patienten zu verbessern, ist es in der Praxis gerade in ländlichen Bereichen noch unmöglich, älteren depressiven Probanden Psychotherapie zu vermitteln. Darüber hinaus gibt es in den meisten geriatrischen Kliniken, in denen Depressionsdiagnostik erfolgt, keine Kapazität für weitere stützende Gespräche mit dem Klienten. Entsprechend erfährt ein depressiver Patient, der in einer Klinik nach seiner Stimmung befragt wird, meist eine deutliche Diskrepanz zwischen diagnostischen Maßnahmen und der sich daran anschließenden Behandlung. Darüber hinaus werden die meisten Patienten auch nicht explizit über die Gabe von antidepressiver Medikation aufgeklärt. Der Patient muss also den Eindruck gewinnen, er wäre grundlos nach seinem gegenwärtigen Gefühlserleben gefragt worden. Das ist nicht schön, lässt sich aber nur teilweise leicht ändern. Dabei stellt die Aufklärung und ggfs. Einbeziehung der Patienten in die medikamentöse Behandlungsplanung das Minimum eines adäquaten Umgangs zwischen Behandelndem und Patient dar. Darüber hinaus wären natürlich auch eine weitere therapeutische Einbettung der Patienten während des stationären Klinikaufenthaltes und ggfs. auch eine entsprechende therapeutische Entlassungsplanung i.S. von Case Management wünschenswert. Beispielsweise ließe sich überlegen, ob Kapazitäten, die bisher für das immer noch weitverbreitete und ebenso weitgehend sinnlose Gedächtnistraining genutzt werden, nicht besser depressiven Patienten in Einzel- oder Gruppensettings zugutekommen sollten. Dabei gilt, wie durch die WHO (Perel-Levin, 2008) postuliert wurde, der Grundsatz, dass allein eine adäquate Behandlung die Rechtfertigung für ein flächendeckendes Screening sein kann.

#### *6.5 Ausblick für Forschung und Praxis*

Eine erste Veröffentlichung zu den Ergebnissen der vorliegenden Studie (Heidenblut & Zank, 2009) wurde bereits vereinzelt als Grundlage für weitere Erhebungen genutzt. So wird die Skala in einer Studie von Haas et al. (in Arbeit) an einer Stichprobe von ambulanten Klienten

einer Gedächtnisambulanz validiert und weitere Erhebungen in stationären sowie ambulanten und teilstationären Einrichtungen sind in Planung (vgl. die Auflistung im Anhang).

Hinsichtlich der praktischen Anwendung im Klinikalltag wurde das Instrument bereits von einigen stationären Geriatrien in ganz Deutschland bei den Autoren angefordert (siehe Anhang), wobei die mündlichen Rückmeldungen der Anwender vielversprechend sind. Erste Erfahrungsberichte aus dem Herz-Jesu-Krankenhaus in Lindlar, in dem die Skala bereits routinemäßig eingesetzt wird, sprechen für eine hohe Akzeptanz und Praktikabilität des Verfahrens.

Als weitere Schritte sind eine Übersetzung der Skala ins englische (Heidenblut & Zank, in Arbeit) sowie eine weitere Akquise von Klinikeinrichtungen, die die Skala in der Praxis anwenden werden, geplant. Darüber hinaus wird die Erstellung eines Ratgebers zur Erfassung von depressiven Störungen in der Geriatrie (Heidenblut & Feis, in Planung) erwogen.

Insgesamt erscheint die Schnittstelle zwischen akademischer und klinischer Diagnostik gerade für das Feld der Gerontologie ein besonders fruchtbares Untersuchungsfeld zu sein, bei dem Gegebenheiten wie Multimorbidität, die Endlichkeit des Lebens, die Subjektivität von Lebensqualität etc., die Forschung immer wieder vor neue interessante und lohnenswerte Herausforderungen stellen.

## LITERATUR

- Adams, K., Matto, H. & Sanders, S. (2004). Confirmatory Factor Analysis of the Geriatric Depression Scale. *The Gerontologist*, 44, 818-826.
- AGAST (1997). *Geriatrisches Basisassessment. Handlungsanleitungen für die Praxis*. München: Medizin Verlag.
- Agrell, B. & Dehlin, O. (1989). Comparison of Six Depression Rating Scales in Geriatric Stroke Patients. *Stroke*, 20, 1190-1194.
- Albers, S. & Skiera, B. (1999). Regressionsanalyse. In A. Herrmann, & C. Homburg (Hrsg.), *Marktforschung. Grundlagen - Methoden – Anwendungen* ( S. 205-236). Wiesbaden: Gabler.
- Altman, D.G. & Bland, J.M. (1994). Diagnostic tests 2: predictive values. *British Medical Journal*, 9, 309.
- Amelang, M. & Schmidt-Atzert, L. (2006). *Psychologische Diagnostik und Intervention*. Heidelberg: Springer.
- Amelang, M. & Zielinski, W. (1997). *Psychologische Diagnostik und Intervention* (2. Aufl.).Berlin: Springer.
- Amelang, M. & Zielinski, W. (2002). Gütekriterien psychometrischer Tests und ihre Bestimmung. In *Psychologische Diagnostik und Intervention* (3. Aufl.) (S.144-195). Berlin: Springer.
- Anderson, I.M. (2000). Selective serotonin reuptake inhibitors versus tricyclic antidepressants: A metaanalysis of efficacy and tolerability. *J Affective Disord*, 58, 19-36.
- Arolt, V. & Driessen, M. (1997). Depressive Störungen bei älteren Patienten im Allgemeinkrankenhaus. In R.D. Hirsch et al. (Hrsg.), *Altern und Depressivität* (S.83-86). Bern: Hans Huber.
- Bach, M., Nikolaus, T., Oster, P. & Schlierf, G. (1995). Depressionsdiagnostik im Alter. Die „Geriatric Depression Scale“. *Zeitschrift für Gerontologie und Geriatrie*, 28, 42-46.
- Baltes, P. B., & Baltes, M. M. (1989). Optimierung durch Selektion und Kompensation: Ein psychologisches Modell erfolgreichen Alterns. *Zeitschrift für Pädagogik*, 35, 85–105.
- Baltes-Götz, B. (2008b). *Behandlung fehlender Werte in SPSS und Amos*. Zugriff am 23.08.2011 unter <http://www.uni-trier.de/fileadmin/urt/doku/bfw/bfw.pdf>
- Baltes-Götz, B. (2006c). *Logistische Regressionsanalyse mit SPSS*. Zugriff am 23.08.2011 unter <http://www.uni-trier.de/fileadmin/urt/doku/logist/logist.pdf>
- Bamberg, G. & Baur, F.(2002). *Statistik*. München: Oldenbourg.

- Bayes, T. (1908). Versuch zur Lösung eines Problems eines Problems der Wahrscheinlichkeitsrechnung (H.E. von Timmerling, Trans.). Leipzig: Wilhelm Engelmann. (Originalarbeit erschienen 1763)
- Beck, A.T. (1967). *Depression: Clinical, experimental and theoretical aspects*. New York: Harper & Row.
- Beck, A.T. (1986). Hopelessness as a predictor of eventual suicide. In J.J. Mann & M. Stanley (eds.), *Psychobiology of suicidal Behavior*. New York: New York Academy of Sciences.
- Beck, J.R. & Shultz, E.K. (1986). The Use of Relative Operating Characteristic (ROC) Curves in Test Performance Evaluation. *Archives of Pathology and Laboratory Medicine*, 110, 13-20.
- Berrios, E. (1991). Affektive Disorders in old age: a conceptual history. *International Journal of Geriatric Psychiatry*, 6, 337-346.
- Bickel, H. (1997). Epidemiologie psychischer Erkrankungen im Alter. In H. Förstl (Hrsg.), *Lehrbuch der Gerontopsychiatrie* (S.1-13). Stuttgart: Enke.
- Blazer, D. (1982). The Epidemiology of Late Life Depression. *Journal of the American Geriatrics Society*, 30, 587-592.
- Blazer, D. (2004): The epidemiology of depressive disorders in late life. In P. Roose & H.a. Sackeim (eds.), *Late-Life Depression* (3-11). Oxford: University Press.
- Blöschl, L. (1998). Depressive Störungen: Ätiologie/ Bedingungsanalyse. In U. Baumann & M. Perrez (Hrsg.), *Lehrbuch Klinische Psychologie – Psychotherapie* (S.853-879). Bern: Hans Huber.
- Birbaumer, N., Schmidt, R.F. (2010). *Biologische Psychologie*. Heidelberg: Springer.
- Bouman, T.K. (1993). Einschätzung von Stimmungsstörungen. In T. A. Albersnagel, P.M. Emmelkamp & R. van de Hoofdakker (Hrsg.), *Depression* (V. Brezinka, Übers.)
- Bortz, J. & Döring, N. (2003). *Forschungsmethoden und Evaluation*. Berlin: Springer
- Bortz, J. & Lienert, G. (1998). *Kurzgefasste Statistik für die klinische Forschung*. Berlin: Springer.
- Bortz, J., Lienert, G. & Boehnke, K. (2000). *Verteilungsfreie Methoden in der Biostatistik*. Heidelberg: Springer.
- Brandstädter, J. & Renner, G. (1992). Coping with discrepancies between aspirations and achievements in adult development: A dual-process model. In L. Montada, S.-H. Philipp & M.R. Lerner (Eds.), *Life crises and experiences of loss in adulthood* (pp. 301-320). Hillsdale, NJ: Erlbaum.

- Brink, T., Yesavage, J., Lum, O., Heersuma, P., Adey, M., Rose, T. (1982). Screening tests for geriatric depression. *Clinical Gerontologist*, 1, 37- 45.
- Brink, T., Curran, P., Dorr, M.L., Janson, E., McNulty, U., Messina, M. (1985). Geriatric Depression Scale : Order, examiner and reminiscence effects. *Clinical Gerontologist*, 3, 57-60.
- Broekman, B.F.P., Niti, M., Nyunt, M., Ko, S., Kumar, R., & Pin, T.(2011). Validation of a Brief Seven-Item Response Bias-Free Geriatric Depression Scale. *American Journal of Geriatric Psychiatry*, 19, 589-596.
- Burke, W., Houston M.J., Boust S.J. & Roccaforte, W.H. (1989). Use of the Geriatric Depression Scale in dementia of the Alzheimer Type. *Journal of the American Geriatrics Society*, 37, 856-60.
- Burke, W., Roccaforte, W. & Wengel, S. (1991). The Short Form of the Geriatric Depression Scale: A Comparison with the 30-Item Form. *Journal of Geriatric Psychiatry and Neurology*, 4, 173-177.
- Burn, W.K., Davies, K., McKenzie, F.R. & Brothwell, J.A. (1993). The Prevalence of Psychiatric Illness in acute geriatric Admissions. *International Journal of Geriatric Psychiatry*, 8, 171-174.
- Camozzato, A., Hidalgo, M. Souza, S. & Chaves, M.(2007). Association among items from the self-report version of the Hamilton Depression Scale and respondents sex. *Psychological Reports*, 101, 291-301.
- Chattat, R., Ellena, L., Cucinotta, D., Savorani, G. & Mucciarelli, G. (2001). A Study on the validity of different short versions of the geriatric depression scale. *Archives of Gerontology and Geriatrics*, 7, 81-86.
- Cole, H., McCusker, J., Ciampi, A., Windholz, S., Latimer, E., Belzile, E. (2006). The prognosis of major and minor depression in older medical inpatients. *The American Journal of Geriatric Psychiatry*, 14, 966-975.
- Cole, H., McCusker, J., Ciampi, A., Belzile, E. (2008). Risk Factors of major depression in older medical inpatients. *The American Journal of Geriatric Psychiatry*, 16, 175-178.
- Cole, H., McCusker, J., Dufouil, C., Ciampi, A., Belzile, E. (2007). Short-term stability of diagnoses of major and minor depression in older medical inpatients. *Psychosomatics*, 48, 38-45.
- Cole, H., McCusker, J., Elie, M., Dendukuri, N., Latimer, E. & Belzile, E. (2005). Systematic detection and multidisciplinary care of depression in older medical inpatients: a randomized trial. *The Japan Medical Association Journal*, 174, 38-44.

- Comer, R.J. (1995). Affektive Störungen. In G. Sortory & J. Metsch (Hrsg.), *Klinische Psychologie*. (S.281-289). Heidelberg: Spektrum.
- Comer, R.J. (2008). *Klinische Psychologie*. Heidelberg: Spektrum.
- Copeland, J.R.M., Dewey, M.E. & Griffiths-Jones, H.M. (1986). A computerized psychiatric diagnostic system and case nomenclature for elderly subjects: GMS und AGE CAT. *Psychological Medicine*, 16, 89-99.
- Copeland, J.R.M., Dewey, M.E. & Griffiths-Jones, H.M. (1990). Dementia und depression in elderly persons: AGE CAT compared with DSM III and pervasive illness. *International Journal of Geriatric Psychiatry*, 5, 47-51.
- Cortina, J.M. (1993). What is Coefficient Alpha? An Examination of Theory and Applications. *Journal of Applied Psychology*, 78, 98-104.
- Craen, A., Heeren, T. & Gussekloo, J. (2003). Accuracy of the 15- Item Geriatric Depression Scale (GDS15) in a community sample of the oldest old. *International Journal of Geriatric Psychiatry*, 18, 63-66.
- D' Ath, P., Katona, P., Mullan, E., Evans, S. & Katona, C. (1994). Screening, detection and Management of depression in elderly primary care attenders: The acceptability and performance of the 15 Item Geriatric Depression Scale (GDS15) and the development of short versions. *Family Practice*, 11 260-266.
- Davison, G. & Neale, J. (2002). Affektive Störungen. In M. Hautzinger (Hrsg.), *Klinische Psychologie* (S.302-355) Weinheim: Beltz.
- Davison, G. & Neale, J. (2002). Psychische Störungen im Alter. In M. Hautzinger (Hrsg.), *Klinische Psychologie* (S.583-632). Weinheim: Beltz.
- Davidson, G., Neale, J. & Hautzinger, M. (2007). *Klinische Psychologie*. Weinheim: Beltz.
- De Bruyn, E. (2003). Assessment Process. In R. Fernandez-Ballesteroz (Hrsg.), *Encyclopedia of psychological assessment* (S. 93-97). London: Sage.
- De Craen, A.J., Gussekloo, J., van der Mast, R.C., le Cessie, S., Lemkes, J.W. & Westendorp, R.G. (2005). Seasonal mood variation in the elderly : the Leiden 85-plus study. *International Journal of Geriatric Psychiatry*, 20, 269-273.
- DeLong, E.R., DeLong, D.M. & Clarke-Pearson, D.L. (1988). Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*, 44, 837-845.
- DeRubeis, R.J., Gelfand, L.A., Tang, T.Z. & Simons, A.D. (1999). Medications versus cognitive behavior therapy for severely depressed outpatients: mega-analysis of four randomized comparisons. *American Journal of Psychiatry*, 156, 1007-1013.

- Disner, S.G., Beevers, C.G., Haigh, E.A.P. & Beck, A.T. (2011). Neural mechanisms of the cognitive model of depression. *Nature Reviews*, 12, 467-477.
- Djernes, J.K. (2006). Prevalence and predictors of depression in populations of elderly: a review. *Acta Psychiatrica Scandinavia*, 113, 372-387.
- Ebel, H., Beichert, K. (2002). Depressive Störungen bei Patienten der Allgemeinmedizin: Früherkennung und therapeutische Ansätze. *Deutsches Ärzteblatt*, 99, 99-124.
- Elsen, A., Müller-Thomsen, T.(2003). Depression nach Schlaganfall bei älteren Patienten in der neurologischen Rehabilitation – Häufigkeit, Prädiktoren und Folgen für den kurzfristigen Rehabilitationserfolg. *Neurologische Rehabilitation*, 9 (3-4), 129-133.
- ESEMeD/MEDEA 2000 Investigators (2004). Prevalence of mental disorders in Europe: Results from the European Study of the Epidemiology of Mental Disorders (ESEMeD) project. *Acta Psychiatrica Scandinavica*, 109, 21-27.
- Evans, M.E. (1993). Depression in elderly physically ill Inpatients: A 12-Month Prospective Study. *International Journal of Geriatric Psychiatry*, 8, 587-592.
- Fenton, F.R., Cole, M.G., Engelsmann, F. & Monsouri, I. (1994). Depression in older medical inpatients. *International Journal of Geriatric Psychiatry*, 9, 279-284.
- Fiest, K.M., currie, S.R., Williams, J.V.A. & Wang, J.L. (2011). Chronic conditions and major depression in community-dwelling older adults. *Journal of Affective Disorders*, 131, 172-178.
- Fisseni, H.-J.(1997). *Lehrbuch der psychologischen Diagnostik* (2.Aufl.). Göttingen: Hogrefe.
- Fisseni, H.-J.(2004). *Lehrbuch der psychologischen Diagnostik* (3.Aufl.). Göttingen: Hogrefe.
- Folstein, M., Folstein, S. & Mc Hugh (1975). Mini Mental State: A practical method for grading the cognitive state of patients for the clinicians. *Journal of Psychiatric Research*, 12, 189-198.
- Folstein, M., Folstein, S. & Mc Hugh (1975). Mini Mental State: A practical method for grading the cognitive state für the clinicians. *Journal of Psychiatric Research*, 12,189-198.
- Fydrich, T. (1997). Diagnostik und Intervention in der Klinischen Psychologie. In M Amelang & W. Zielinski (Hrsg.), *Psychologische Diagnostik und Intervention* (S.443-487). Berlin: Springer.
- Freud, S. (1917). Mourning and Melancholia. In *Collected papers* (vol.4). London: Hogarth and the Institute of Psychoanalysis, 1950.

- Fromm, S. (2005). *Binäre logistische Regressionsanalyse. Eine Einführung für Sozialwissenschaftler mit SPSS für Windows*. Zugriff am 23.08.2011 Unter [http://uni-bamberg.de/fileadmin/uni/fakultaeten/sowi\\_lehrstuehle/empirische\\_sozialforschung/pdf/bambergerbeitraege/Log-Reg-BBES.pdf](http://uni-bamberg.de/fileadmin/uni/fakultaeten/sowi_lehrstuehle/empirische_sozialforschung/pdf/bambergerbeitraege/Log-Reg-BBES.pdf)
- Gallagher-Thomson, D., Wolford-Thomson, L.(1995). Probleme des Alterns. In G. Sortory & J. Metsch (Hrsg.), *Klinische Psychologie*. (S.719-730). Heidelberg: Spektrum.
- Gantner, a., Schubert, D., Wolf, S.& Creps, P. (2003). Screening for Depression in a Geriatric Rehabilitation Sample. *International Journal of Psychiatry in Medicine*, 33(4),S.333-341.
- Gauggel , S., Birkner, B. (1998). Diagnostik depressiver Störungen bei älteren Menschen. Eine Übersicht über die Entwicklung und Evaluation der „Geriatric Depression Scale“ (GDS). *Zeitschrift für Gerontopsychologie und Psychiatrie*, 11, 159-171.
- Gauggel, S. & Birkner, B. (1999). Validität und Reliabilität einer deutschen Version der Geriatrischen Depressionsskala. *Zeitschrift für klinische Psychologie*, 28, 18-27.
- Gauggel, S. & Schmidt, A. (1995). Was leistet die deutsche Version der Geriatrischen Depressionsskala (GDS)?. *Geriatric Praxis*, 6, 33-36.
- Geiger-Kabisch, C. & Weyerer, S. (1991). Zur Reliabilität und Validität von Screening-Instrumenten zur Erfassung von Depressionen bei älteren Menschen: Ein Überblick. *Zeitschrift für Gerontopsychologie und –psychiatrie*, 4, 75-89.
- Gensichen, J. Beyer, M. Schwäbe, N. & Gerlach, F.M. (2004). Hausärztliche Begleitung von Patienten durch Case Menagement – Ein BMBF-Projekt. *Zeitschrift für Allgemeine Medizin*, 80, 507-511.
- Gesellschaft für Versicherungswirtschaft und –gestaltung e.V.(2006). *6. Nationales Gesundheitsziel: Depressive Erkrankungen: verhindern, früh erkennen, nachhaltig behandeln*. Berlin: Bundesministerium für Gesundheit.
- Gigeranzer, G., Swijtnk, Z., Porter, T., Daston, L., Beatty, J., Krüger, L.(1999). *Das Reich des Zufalls: Wissen zwischen Wahrscheinlichkeiten, Häufigkeiten und Unschärfen*. Heidelberg: Spektrum.
- Goring, H., Baldwin, R., Marriott, A., Pratt, H. & Roberts, C. (2004). Validation of short screening tests for depression and cognitive impairment in older medically ill inpatients. *International Journal of Geriatric Psychiatry*, 19, 465-471.
- Gremaud-Heitz, D.J., Steart, J.W., Dammann, G. (2010). Konzept der atypischen Depression und deutsche Pbersetzung der “Atypical Deprssion Diagnostic Scale (ADDs)””. *Schweizer Archiv für Neurologie und Psychiatrie*, 162, 148-154.

- Gullikson, H. (1950). *Theorie of mental tests*. New York: Wiley.
- Guidi, J., Fava, G.A., Picardi, A., Porcelli, P., Bellomo, a., Grandi, S., Grassi, L., Pasquini, P., Quartesan, R., Rafanelli, C., Rigatelli, M. & Sonino, N. (2011). Sybtyping depression in the medically ill by cluster analysis. *Journal of Affective Disorders*, 132, 383-388.
- Haas, et al. (in Arbeit). Depression in geriatric outpatients in a German memory clinic: Assessing the reliability and validity of the Depression- In-old-Age-Skale (DIA-S), the Geriatric Depression scale (GDS15) and the association of depression with cognitive functioning.
- Hammen, C. & Watkins, E. (2008). *Depression*. Hove: Psychology Press.
- Hammond, M.F. (2004). Doctors`and nurses`observations on the Geriatric Depression Rating Scale. *Age and Aging*, 33, 189-192.
- Han, L., McCusker, J., Cole, M., Abrahamowicz, M. & Capek, R. (2006). The temporal relationship between Depression symptoms and cognitive functioning in older medical patients – perspective or concurrent?. *The Journals of Gerontology, A*, 61a. 1319-1323.
- Han, L., McCusker, J., Cole, M., Abrahamowicz, M, & Capek, R. (2008). 12-month cognitive outcomes of major and minor depression in older medical patients. *The American Journal of Geriatric Psychiatry*, 16, 742-751.
- Hanley, J.H. & Mc Neil, B.J. (1982). The Meaning and Use of the Area under a Receiver Operating Characteristic (ROC) Curve. *Radiology*, 143, 29-36.
- Harper, R.G., Kotik Harper, D. & Kirby, H. (1990). Psychometric assessment of depression in an elderly general medical population: Over –or underassessment ? *Journal of Nervous and Mental Disease*, 178, 113-119.
- Hasnain, M., Vieweg, W.V.R., Lesnefsky, E.J., & Pandurangi, A.K. (2011). Depression screening in patients with coronary heart disease: A critical evaluation of the AHA guidelines. *Journal of Psychosomatic Research*, 71, 6-12.
- Hautzinger, M.(1998). *Depression, Band 4* (S.1-22). Göttingen: Hogrefe.
- Hautzinger, M. (2002). Depressive Störungen. In A. Maercker (Hrsg.), *Alterspsychiatrie und klinische Gerontopsychologie* (S. 141-163). Berlin: Springer.
- Hautzinger, M. & Meyer, T.D. (2002). *Diagnostik affektiver Störungen*. Göttingen: Hogrefe.
- Hautzinger, M. & de Jong-Meyer, R. (2003). Depressionen. In H. Reinecker (Hrsg.), *Lehrbuch der Klinischen Psychologie und Psychotherapie* (S.215-259).Göttingen: Hogrefe.

- Hautzinger, M. & Radebold, H. (1997). Wirk- und Einflussfaktoren auf Depressionen im Alter. In R. D. Hirsch et al. (Hrsg.), *Altern und Depressivität* (S.112-114).Bern: Hans Huber.
- Hautzinger, M. & Welz, S. (2004). Kognitive Verhaltenstherapie bei Depressionen im Alter. Ergebnisse einer kontrollierten Vergleichsstudie unter ambulanten Bedingungen an Depressionen mittleren Schweregrads. *Zeitschrift für Gerontologie und Geriatrie*, 37, 427-435.
- Heidenblut, S. & Zank, S. (2005). Depressionsdiagnostik bei geriatrischen Patienten. Validierung und Modifikation der geriatrischen Depressionsskala (GDS):*Vortrag, gehalten auf der gemeinsamen Jahrestagung der Sektionen III und IV der Deutschen Gesellschaft für Gerontologie und Geriatrie in Mannheim, 29.-30-9.*
- Heidenblut, S. & Zank, S. (2009). Versorgung von Depressionen im Alter. In G. Adler, H.Gutzmann, M. Haupt, R. Kortus & D.K.Wolter (Hrsg.), *Seelische Gesundheit und Lebensqualität im Alter* (S. 15-20). Stuttgart: Kohlhammer.
- Heidenblut, S. & Zank, S. (2010). Entwicklung eines neuen Depressionsscreenings für den Einsatz in der Geriatrie.[Development of a new screening instrument for geriatric depression]. *Zeitschrift für Gerontologie und Geriatrie*, 43 (3),.170-176.
- Heidenblut, S. & Zank, S. ( in Arbeit) Screening for depression among geriatric inpatients: Diagnostic accuracy of the Depression in old Age-Scale (DIA-S) and the GDS15 considering patient – and setting-characteristics
- Helmchen H., Baltes M.M., Geiselman B., Kanowski S., Linden M., Reischies, F.M., Wagner M., et al. (1996). Psychische Erkrankungen im Alter. In K.U. Mayer & P.B. Baltes (Hrsg.), *Die Berliner Altersstudie* (S.185-216). Berlin: Akademie Verlag.
- Helmchen, H., Baltes, M. M., Geiselman, B., Kanowski, S., Linden, M., Reischies, Wagner, F.M. & Wilms, H.-U. (2010). Psychische Erkrankungen im Alter. In Lindenberger, U., Smith, J., Mayer, K. U., & Baltes, P. B. (Hrsg.). (2010). *Die Berliner Altersstudie*, S. 209-245, Berlin: Akademie Verlag.
- Henkel, V., Mergl, R., Kohnen, R., Möller, H-J., Hegerl, U. & Allgaier, A.-K.(2004). Use of brief depression screening tools in primary care: consideration of heterogeneity in performance in different patient groups. *General Hospital Psychiatry*, 23, 190-198.
- Herrman, N., Mittman, N., Silver, I.L., Shulman, K.I., Busto, U.A., Shear, N.H. & Naranjo,C.A.(1996). A Validation Study of the Geriatric Depression Scale short form. *International Journal of Geriatric Psychiatry*, 11, 457-460.

- Heuft, G. (1992). Formen und Erscheinungsbilder der depressiven Syndrome. In R.D. Hirsch (Hrsg.), *Altern und Depressivität* (S.29-43). Bern: Hans Huber.
- Heuft, G., Kruse, A., Radebold, H. (2000). Die depressiven Erkrankungen. In *Lehrbuch der Gerontopsychosomatik und Alterspsychotherapie* (S. 123-132). München:Ernst Reinhardt Verlag.
- Heuft, G., Kruse, A. & Radebold, H. (2006). *Lehrbuch der Gerontopsychosomatik und Alterspsychotherapie*. Weinheim: Beltz.
- Hörning (2004). Grundlagen der Sozialmedizin – Gesundheit, Krankheit, Behinderung, Public Health. In W.Schwarzer (Hrsg.), *Lehrbuch der Sozialmedizin*. Dortmund: Borgmann.
- Jackson, R. & Baldwin, B. (1993). Detecting depression en elderly medically ill patients: The use of the Geriatric Depression Scale compared with medical and nursing Observations. *Age and Aging*, 22, 349-353.
- Jacobi, F., Höfler, M. & Wittchen, H.U. (2002). Prävalenz, Erkennens- und Verschreibungsverhalten bei depressiven Syndromen – eine bundesdeutsche Hausarztstudie. *Der Nervenarzt*, 7, 651-658.
- Jong-Meyer, R. de, Hautzinger, M., Kühner, C. & Schramm, E. (im Druck). Leitlinien: Psychotherapie Affektiver Störungen. Fachgruppe Klinische Psychologie und Psychotherapie der Deutschen Gesellschaft für Psychologie. [www.klinischepsychologiepsychotherapie.de/dateien/lleitlinien2004ptllversionmrz05.pdf](http://www.klinischepsychologiepsychotherapie.de/dateien/lleitlinien2004ptllversionmrz05.pdf) (Letzter Zugriff 20.12.2005).
- Katona, C. (1994). *Depression in old age*. Chichester: John Wiley & Sons.
- Karakus, M.C.,& Patton, L.C. (2011). Depression and the Onset of Chronic Illness in Older Adults: A 12-Year Prospective Study. *Journal of Behavioral Health Services & Research*, 38, 373-382.
- Kenny Weeks, S., McGann, P., King Michaels, T. & Penninx, B. (2003). Comparing various short-form Geriatric depression scales leads to the GDS-5/15. *Journal of Nursing Scholarship*, 35, 133-137.
- Kessler, R. C. (2003). Epidemiology of women and depression. *Journal of Affective Disorders*, 74, 5-13.
- Kessler, R.C., Berglund, P., Demler, O., Jin, R. Koretz, D., Merikangas, K.R. et al. (2003). The epidemiology of major depressive disorder. *Journal of the American Medical Association*, 289, 3095-3105.

- Koenig, H. (1997). Differences in psychosocial and health correlates of major and minor depression in medically ill older adults. *Journal of the American Geriatrics Society*, 45(12), 1487-1495.
- Koenig, H. G., Meador, K. G. Cohen, H.J. & Blazer, D.G. (1988c). Self rating depression Scales and screening for major depression in the older hospitalized patient with medical illness. *Journal of the American Geriatric Society*, 36, 699-706.
- Koenig, H.G., Meador, K.G., Shelp, F., Goli, V., Cohen, H.J. &Blazer, D.G. (1991). Major Depressive Disorder in Hospitalized Medically Ill Patients: An Examination of Young and Elderly Male Veterans. *Journal of the American Geriatrics Society*, 39, 881-890.
- Kok, R.M., Heeren, T.J., Hooijer, C., Dinkgreve, M. & Rooijmans, H. (1995). The Prevalence of Depression in Elderly medical Inpatients, *Journal of Affective Disorders*, 33, 77-82.
- Kørner, A., Lauritzen, L., Abelskov, K., Gulmann, N., Brodersen, A.M., Wedervang-Jensen, T., Kjeldgaard, K.M, (2005). The Geriatric Depression Scale and the Cornell Scale for Depression in Dementia. A validity study. *Nordic Journal of Psychiatry*, 50, 360-364.
- Krauth, J. (1995). *Testkonstruktion und Testtheorie*. Weinheim: Beltz.
- Kruse, A. (1992). Depressionen bei Pflegebedürftigkeit. In R. D. Hirsch et al. (Hrsg.), *Altern und Depressivität* (S.181-194). Bern: Hans Huber.
- Kurz, A. (1997). Depressionen im Alter: Klassifikation, Differentialdiagnose und Psychopathologie. In H. Radebold et al. (Hrsg.), *Depressionen im Alter*. (S.18-33). Darmstadt: Steinkopf.
- Laidlaw, K., Thomson, L.W., Dick-Siskin, L. & Gallagher-Thomson, D. (2003). *Cognitive Behaviour Therapy with older People*. West Sussex: Wiley.
- Lämmle, G. (09.07.2005). Vorstellung des EGZB im Rahmen eines Praktikumsinserats auf der Homepage der Deutschen Gesellschaft für Neuropsychologie. File: [gnp.de/html/stellenbörse/praktikumsplätze/index.php](http://gnp.de/html/stellenbörse/praktikumsplätze/index.php).
- Lasser, R., Siegel, E., Dukoff, R., & Sunderland, T. (1998). Diagnosis and treatment of geriatric depression. *CNS Drugs*, 1, 17-30.
- Leentjens, A., Verhey, F., Lousberg, R., Spitsbergen, H. & Wilmsink, F. (2000). The validity of the Hamilton and Montgomery-Asberg Depression Rating Scales as screening and diagnostic tools for depression in Parkinson Disease. *International Journal of Geriatric Psychiatry*, 15, 644-649.

- Leshner, E.L. & Berryhill, J.S. (1994). Validation of the Geriatric Depression Scale- short form among inpatients. *Journal of Clinical Psychology, 50*, 256-260.
- Lewinsohn, P.M. (1974). A behavioural approach to depression. In M.M. Friedman, R.J. Katz (eds.), *The psychology of depression*. New York: Wiley.
- Lewinsohn, P.M. (1981). Depression related cognitions: Antecedent or consequences? *Journal of Abnormal Psychology, 90*, 213-219.
- Lieberman et al. (1999). Geriatric Depression Screening Scale (GDS) in Patients Hospitalized for physical rehabilitation. *International Journal of Geriatric Psychiatry, 14*, 549-555.
- Lienert, G.A. & Raatz, U. (1998). *Testaufbau und Testanalyse*. Weinheim: Beltz.
- Linka, E., Bartko, G., Agardi, T. & Kemeny, K. (2000). Dementia and Depression in elderly medical Inpatients. *International Psychogeriatrics, 12(1)*, 67-75.
- Lopez, A. Mathers, C., Ezzati, M. Jamison, D. & Murray, C. (2006). *Global Burden Of Disease and Risk Factors*. Oxford: University Press.
- Mahoney, F.I. & Barthel. D (1965). Funktional Evaluation: The Barthel Index. *Maryland State Medical Journal, 14*, 56-61.
- Margraf, J. (2003). *Lehrbuch der Verhaltenstherapie* (Band 1). Berlin: Springer.
- Mausback, B.T., Chattillion, E.A., Moore, R.C., Roepke, S., Depp, C.A. & Roesch, S. (2011). Activity restriction and depression in medical patients and their caregivers: A meta-analysis. *Clinical Psychology Review, 31*, 900-908.
- McCusker, J., Cole, M., Dufouil, C., Dendukuri, N., Latimer, E., Windholz, S. & Elie, M. (2005). The prevalence and correlates of major and minor depression in older medical inpatients. *Journal of the American Geriatrics Society, 53*, 1344-1353.
- McCusker, J., Cole, M., Ciampi, A., Windholz, S., & Belzile, E. (2007). Major depression in older medical inpatients predicts poor physical and mental health status over 12 months. *General Hospital Psychiatry, 29*, 340-348.
- Meller, M., Fichter, M. & Schröppel, H. (1997). Psychosoziale Risikofaktoren für Depressionen bei Hochbetagten – Ergebnisse aus der Münchener Hochbetagten-Studie. In H. Radebold et al. (Hrsg.), *Depressionen im Alter* (S.119-122). Darmstadt: Steinkopf.
- Mertens, W. (2002). Grundlagen psychoanalytischer Psychotherapie. In W. Senf & M. Broda (Hrsg.), *Praxis der Psychotherapie* (S.130-168). Stuttgart: Thieme.
- Möller, H. (1997). Medikamentöse Behandlung der Depression im höheren Lebensalter. In R.D. Hirsch (Hrsg.), *Altern und Depressivität* (S.41-50). Bern: Hans Huber.

- Möller, H., Laux, G. & Deister, A.-J. (2005). *Psychiatrie und Psychotherapie*. Stuttgart: Thieme.
- Möller, H.-J., Laux, G. & Kapfhammer, H.P. (2008). *Psychiatrie und Psychotherapie Band 2*. Heidelberg: Springer.
- Montgomery, S.A., Asberg, M., Perris, C., Schalling, D., Sedvall, G.(1978). A comprehensive psychopathological rating-scale. *Acta psychiatrica scandinavica*, 271, 5-25.
- Montgomery, S. & Asberg, M. (1979). A new depression scale designed to be sensitive to change. *British Journal of Psychiatry*, 134, 382-389.
- Morishita, L., Boulton, C., Ebbitt, B., Rambel, M., Fallstrom, K. & Gooden, T.(1995). Concurrent Validity of Administering the Geriatric Depression Scale and the Physical Of the SIP by Telephone. *Journal of the American Geriatric Society*, 43(6), 674-679.
- Müller-Spahn, F. (2002). Depressionen im höheren Lebensalter. In W. Gaebel & F. Müller-Spahn (Hrsg.), *Diagnostik und Therapie psychischer Störungen* (S. 449-463). Stuttgart: Kohlhammer.
- Müller-Spahn, F. & Hock, C. (1997). Neurobiologische Faktoren und somatische Aspekte der Depressionen im höheren Lebensalter. In H. Radebold et al. (Hrsg.), *Depressionen im Alter* (S.119-122). Darmstadt: Steinkopf.
- Müller-Thomsen, T., Meins, W. & Manecke, S. (1999). Psychiatrische Störungen im alter und deren psychosoziale Hintergründe. *Psychiatrische Praxis*, 267-272.
- Müller-Thomsen, T., Tabrizian, S., Mittermeier, O. (2003). Depression bei geriatrischen Patienten mit hüftgelenksnahen Frakturen und deren Auswirkung auf den Rehabilitationsverlauf. *Zeitschrift für Gerontologie und Geriatrie*, 36, 138-142.
- Mottram, P., Wilson, K. & Copeland, J. (2000). Validation of the Hamilton Depression Rating Scale and Montgommery and Asperg Rating Scales in terms of AGE-CAT Depression cases. *International Journal of Geriatric Psychiatry*, 12, 1113-1119.
- Murray; c. & Lopez, A. (1997). Alternative projections of mortality and disability by cause 1990-2020. Global Burden of Disease Study. *The Lancets* 349, 1498-1504.
- National Institute of Clinical Excellence (NICE) (Hrsg.) (2004), NICE guideline: quick reference guide – depression. London.  
<http://www.nice.org.uk/pdf/CG023quickrefguide.pdf> (Letzter Zugriff 20.12.2005).
- Neumann, N.-U. & Schulte, R.-M. (1989). Montgomery and Asberg Depression Rating Scale. Deutsche Fassung. Erlangen: Perimed Fachbuch Verlagsgesellschaft.
- Nordhus, I.H. (2008). Manifestations of derpression and anxiety in older adults. *Handbook of the Clinical Psychology of Ageing*, (99-110). Chichester: John Wiley & Sons.

- Ostermann, K. (1997). Depressivität bei stationär behandelten geriatrischen PatientInnen. In R.D. Hirsch (Hrsg.), *Altern und Depressivität* (S.86-88). Bern: Hans Huber.
- Parmelee, P., Lawton, M. & Katz (1989). Psychometric Properties of the Geriatric Depression Scale Among the Institutionalized Aged. *Psychological Assessment*, 1, 331-338.
- Pawlik, K. (2003). Theoretical Perspektive: Psychometrics. In R. Fernandez-Ballesteroz (Hrsg.), *Encyclopedia of psychological assessment* (S. 1019-1023). London: Sage.
- Paykel, E.S., Brugha, T., & Fryers, T. (2005). Size and burden of depressive disorders in Europe. *European Neuropsychopharmacology*, 15, 411-423.
- Peter, M.(2004). *Klinische Entwicklungspsychologie des Alters – Grundlagen für psychosoziale Beratung und Psychotherapie*. Göttingen: Vandenhoeck & Ruprecht.
- Pinel, J. (1997). Die biopsychologischen Grundlagen von Emotionen, Streß und psychischen Störungen. In W. Boucsein (Hrsg.), *Biopsychologie* (S. 458-490). Heidelberg: Spektrum.
- Parmelee, P., Lawton, M. & Katz (1989). Psychometric Properties of the Geriatric Depression Scale Among the Institutionalized Aged. *Psychological Assessment*, 1, 331-338.
- Pepe, M.(2004). *The Statistical Evaluation of Medical Tests for Classification and Prediction*. Oxford: University Press.
- Perel-Levin, S. (2008). *Discussing screening for elder abuse at primary health care level*. World Health Organization.
- Pomeroy, I., Clark, C. & Philp, I.(2001). The effectiveness of very short scales for depression screening in elderly medical patients. *International Journal of Geriatric Psychiatry*, 16, 321-326.
- Porst, R. (2000). *Question Wording – Zur Formulierung von Fragebogen-Fragen*. ZUMA How-to-Reihe, 2.
- Pössel, P., & Knopf, K. (2010). Bridging the Gaps: An Attempt to Integrate Three Major Cognitive Depression Models. *Cognitive Theory Research*, 35, 342-358
- Radebold, H.(1997). Psychoanalytische Psychotherapie von Depressionen über 60jähriger. In R.D. Hirsch (Hrsg.), *Altern und Depressivität* (S.51-59). Bern: Hans Huber.
- Ramsay, R., Wright, P., Katz, A. & Bielawska, C. (1991). The detection of psychiatric Morbidity and its effects on outcome in acute elderly medical admissions. *International Journal of Geriatric Psychiatry*, 6, 861-866.

- Renteln-Kruse, W. & Ebert, D. (2003), Merkmale hospitalisierter geriatrischer Patienten – Zwei Kohorten im Vergleich unter Verwendung des Screenings der Raasch, C. (1996). Depressionsmessung in der geriatrischen Rehabilitation. *Geriatric Forschung*, 6, 117-122.
- Reischies, F.M. (2003). Depression. In E. Steinhagen-Thiessen & B. Hanke (Hrsg.), *Neurogeriatrie* (S. 87-128). Berlin-Wien: Blackwell Verlag GmbH.
- Renteln-Kruse, W. & Ebert, D. (2003) Merkmale hospitalisierter geriatrischer Patienten – Zwei Kohorten im Vergleich unter Verwendung des Screenings der Arbeitsgemeinschaft Geriatrisches Basisassessment (AGAST). *Zeitschrift für Gerontologie und Geriatrie*, 36, 223-232.
- Richter, T. (2001). Psychische Erkrankungen im Alter: Vorbeugen und Therapieren. Verfügbar unter : [http://www.who-tag.de/pdf/2006richter\\_r\\_abstract.pdf](http://www.who-tag.de/pdf/2006richter_r_abstract.pdf). Zugriff am 12.11.2007.
- Rinaldi, P., Mecocci, P., Benedetti, C., Ercolani, S., Bregnocchi, M., Menculini, G., Catani, M., Senin, U., & Cherubini, A. (2003). Validation of the five-item Geriatric Depression Scale in elderly subjects in three different settings. *Journal of the American Geriatrics Society*, 51, 694-698.
- Rovner, B.W., Shmueli-Dulitzki, Y. (1997). Screening for depression in low-vision elderly. *International Journal of Geriatric Psychiatry*, 12, 955-959.
- Ryan, D.H. (1995). Depression and Dementia in Geriatric Inpatients: Diagnostic Comparisons between Psychiatrists, Geriatricians and Test Scores. *International Journal of Geriatric Psychiatry*, 10, 447-456.
- Schaller, S. & Göbel, L. (1988). Zur Messgenauigkeit der Erfassung von Beschwerden im Alter. In K. Böhme & E. Lungershausen (Hrsg.), *Beiträge zur Erforschung selbstdestruktiven Verhaltens, Band 8, Suizid und Depression im Alter* (S.66-77). Regensburg: S. Roderer Verlag.
- Schlager, Schwartz & Bromet (1993). Seasonal variations of current symptoms in a healthy population. *The British Journal of Psychiatry*, 163, 322-326.
- Schmeling-Kludas, C., Jäger, K. & Niemann, B.M. (2000). Diagnostik und Bedeutung Psychischer Störungen bei körperlich kranken geriatrischen Patienten. *Zeitschrift für Gerontologie und Geriatrie*, 33, 36-43.
- Smith, D., Dempster, C., Glanville, J., Freemantle, N. & Anderson, I. (2002). Efficacy and tolerability of venlafaxine with selective serotonin reuptake inhibitors and other antidepressants: a metaanalysis. *British Journal of Psychiatry* 180, 396-494.

- Schmitz, N., Kruse, J., Heckrath, C., Alberti, L., Tress, W. (1999). Diagnosing mental disorders in primary care: the General Health Questionnaire (GHQ) and the Symptom Checklist (SCL-90-R) as screening instruments. *Soc Psychiatry Psychiatr Epidemiol*, 34, 360-366.
- Schneider, G., Kruse, A., Nehen, H.-G., Senf, W. & Heuft, G. (2000). The Prevalence and Differential Diagnosis of Subclinical Depressive Syndromes in Inpatients 60 Years and Older. *Psychotherapy and Psychosomatics*, 69, 251-260.
- Schott, H. & Tölle, R. (2006). *Geschichte der Psychiatrie. Krankheitslehren, Irrwege, Behandlungsformen*. München: C. H. Beck.
- Shah, A., Phongasthorn, V., Bielawska, C. & Katona, C. (1996). Screening for depression among geriatric inpatients with short versions of the geriatric depression scale. *International Journal of Geriatric Psychiatry*, 11, 915-918.
- Schwärzler, F. (2003). Depression im Alter (S. 45-59). In H. Wormstall & H.-J. Wilhelm (Hrsg.), *Alterspsychiatrie im Wandel* (S. 45-59). Oberhausen: Athena
- Seidel, U., Pantel, J., Re, S., Schröder, J. (2004) Depressive Störung und Spätdepression. In A. Kruse und M. Martin (Hrsg.), *Enzyklopädie der Gerontopsychiatrie*, S. 240-254. Göttingen: Hans Huber.
- Seligman, M.E.P. (1974). Depression and learned helplessness. In R.J. Friedman & M.M. Katz (eds.), *The psychology of depression. Contemporary theory and research*. Washington: Winston-Wiley.
- Semmler, G. & Klumb, P.L. (2004). Geschlechtsunterschiede in der Prävalenz depressiver Symptomatik: Ein Resultat differentieller Validität der Erhebungsinstrumente? [Gender differences in the prevalence of depressive symptoms: A consequence of differentially valid measurement instruments?]. (*Journal of Public Health*, 12, 139-146.
- Shiekh, J., Yesavage, J. (1986). Geriatric Depression Scale; recent findings and development Of a short version. In T. Brink (Hrsg.), *Clinical gerontology: a guide to assessment and intervention*. New York: Howarth Press.
- Sigmon, S., Pells, J., Boulard, N., Whitcomb-Smith, S., Edenfield, T., Hermann, B., LaMattina, S., Schartel, J. & Kubik, E. (2005). Gender differences in self-reports of depression: the response bias hypotheses revisited. *Sex Roles*, 53, 401-411.
- Snowdon, J. (2001). Is depression more prevalent in old age? *Australian and New Zealand Journal of Psychiatry*, 35, 782-787.

- Stieglitz, R.-D. (1994). Selbstbeurteilungsverfahren. In R.-D. Stieglitz & U. Baumann (Hrsg.), *Psychodiagnostik psychischer Störungen*. Stuttgart: Enke.
- Stieglitz, R.-D. (2000). *Diagnostik und Klassifikation psychischer Störungen*. Göttingen: Hogrefe.
- Storosum, J.G., Elferink, A.J.A., van Zwieten, B.J., van den Brink, W., Gersons, B.P.R., van Strik, R. & Broekmans, A.W. (2001). Short-term efficacy of tricyclic antidepressants revisited: a metaanalytic study. *European Neuropsychopharmacology* 11, 173-180.
- Tabachnick, B.G. & Fidell, L.S. (2007). *Using Multivariate Statistics*. Boston: Pearson.
- Thormann, J., Himmerich, H. & Steinberg, H. (2011). Das Konzept der „vegetativen Depression“ (1949) von Rudolf Lemke – Ein obsoletes Konzept? *Psychiatrie Praxis*, 38, 91-96.
- Trajman, A. & Luiz, R.R. (2008). McNemar  $\chi^2$  test revisited: comparing sensitivity and specificity of diagnostic examinations. *The Scandinavian Journal of Clinical & Laboratory Investigation*, 68, 77-80.
- Unnützer, J. Katon, W. Callahan, C. Williams, J., Hunkeler, E., Harpole, L. hoffing, M., Penna, R., Noel, P., Lin, E., Arean, P. Hegel, M., Tang, L., Belin, T. Oishi, S. & Lanston, c. (2001). Collaborative Care Management of Late-Life Depression in the Primary Care Setting. *Journal of the American Medical Association*, 288, 2836-2845.
- Van de Velde, S., Bracke, P., Levecque, K. & Meuleman, B. (2010). Gender differences in depression in 25 European Countries after eliminating measurement bias in the CES-D 8. *Social Science Research*, 39, 396-404.
- van Gunten, A., Herrmann, F.R., Elliott, R. & Duc, R. (2011). Abnormal sensitivity to negative feedback in late-life depression. *Psychiatry and clinical Neurosciences*, 65, 333-340.
- Waller (2002). Krankheit und Behinderung: medizinische und sozialwissenschaftliche Konzepte. In: *Sozialmedizin – Grundlagen und Praxis*, S. 9-41. Stuttgart: Kohlhammer.
- Wancata, J., Alexandrowicz, R., Marquart, B., Weiss, M. & Friedrich, F. (2006). The criterion validity of the Geriatric Depression Scale: a systematic review. *Acta Psychiatrica Scandinavica*, 114, 398-410.
- Wancata, J., Meise, U. & Sachs, G. (2000). Prävalenz seelischer Erkrankungen unter älteren Patienten an internen und chirurgischen Stationen. *Psychiatrische Praxis*, 27, 170-175.

- Wernicke, T. & Linden, M. (1997). Psychopharmakotherapie bei Depressionen im Alter – Die Berliner Altersstudie (BASE). In H. Radebold, R. D. Hirsch, J. Kopp, R. Kortus, G. Stoppe, B. Struwe et al. (Hrsg.) *Depressionen im Alter* (S. 152-153). Darmstadt: Stinkopff.
- Wetterling, T. (1999). Besonderheiten affektiver Störungen im Alter. In P. Hartwich, S. Haas, K. Maurer & B. Pflug (Hrsg.), *Affektive Erkrankungen und Lebensalter* (S. 29-37). Berlin: Wissenschaft & Praxis.
- WHO (2000). *Die vielen Gesichter des psychischen Leids – Das offizielle Fallbuch der WHO zur ICD-10 Kapitel V (F) – Falldarstellungen von Erwachsenen*. Bern: Hans Huber.
- WHO (2005). Affektive Störungen. In H. Dilling, W. Mombour & M. H. Schmidt (Hrsg.), *Internationale Klassifikation psychischer Störungen ICD-10 Kapitel V (F), Klinisch-diagnostische Leitlinien* (5. Aufl.) (S. 129-152). Bern: Hans Huber.
- WHO (2011). *Internationale Klassifikation psychischer Störungen – ICD-10 Kapitel V (F) – Diagnostische Kriterien für Forschung und Praxis*. Bern: Hans Huber.
- Weyerer, S. & Bickel, H. (2007). *Epidemiologie psychischer Erkrankungen im höheren Lebensalter*. Grundriss Gerontologie Band 14. Stuttgart: Kohlhammer.
- Yesavage, J., Brink, T., Rose, T., Lum, O., Huang, V., Adey, m. & Leirer, O. (1983). Development and validation of a geriatric depression scale: a Preliminary report. *Journal of Psychiatric Research*, 17, 37-49.
- Zank, S. (2002). Einstellungen alter Menschen zur Psychotherapie und Prädiktoren der Behandlungsbereitschaft bei Psychotherapeuten. *Verhaltenstherapie und Verhaltensmedizin*, 23, 181-193.
- Zank, S. & Schacke, C. (2001). Evaluation von Effekten gerontopsychiatrischer und geriatrischer Tagesstätten auf ihre Besucher(innen) und deren Angehörigen. *Schriftreihe des Bundesministeriums für Familie, Senioren, Frauen und Jugend*. Band 210. Stuttgart: Kohlhammer.
- Zank, S., Schacke, C., Heidenblut, S. & Opterbeck, I. (2008). Zwischenbericht im Rahmen der Längsschnittstudie zur Belastung pflegender Angehöriger von demenziell Erkrankten (LEANDER-Phase III). *Bundesministeriums für Familie, Senioren, Frauen und Jugend*.
- Zank, S. Peters, M. & Wilz, G. (2010). *Klinische Psychologie und Psychotherapie des Alters*. Stuttgart: Kohlhammer.