

# Empathie und Affektive Störungen

Lassen sich soziale Schwierigkeiten unipolar depressiver oder bipolarer Patienten mit einem Defizit der Empathie erklären?

## Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät  
der Eberhard Karls Universität Tübingen  
zur Erlangung des Grades eines  
Doktors der Naturwissenschaften  
(Dr. rer. nat.)

vorgelegt von

**Dipl.-Psych. Fabienne Große Wentrup**  
aus Stuttgart

Tübingen  
2017

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät  
der Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation: 22.10.2018  
Dekan: Prof. Dr. Wolfgang Rosenstiel  
1. Berichterstatter: Prof. Dr. Martin Hautzinger  
2. Berichterstatter: Prof. Dr. Stefan Klingberg

Für meinen geliebten Sohn



# Danksagung

Mein besonderer Dank gilt all jenen Patienten und gesunden Probanden, die sich zur aufwändigen Teilnahme an der Erhebung bereit erklärt und damit die vorliegende Arbeit erst möglich gemacht haben. Prof. Dr. Martin Hautzinger möchte ich für die gute Betreuung und Unterstützung während meiner Promotionszeit danken. Dr. Larissa Wolkenstein danke ich herzlich für die Möglichkeit, im Rahmen eines spannenden Projektes und einer belebten Arbeitsgruppe mitzuarbeiten. Eben jener Arbeitsgruppe bin ich dankbar für die gute Zusammenarbeit und gegenseitige Ermutigung sowie Erheiterung. Besonders zu nennen sind hierbei Marjorie Kinney, die die Rückübersetzung des Fragebogens ermöglicht hat, aber auch Stefan Lüttke und Julia Feyerabend, die im selben Projekt tätig waren. Auch den Studentinnen und Studenten, die bei der Durchführung der beiden empirischen Arbeiten tatkräftig mitgearbeitet haben, sei herzlich gedankt. Großer Dank gebührt Sibylla Wolter für die wunderbare Zusammenarbeit zu Beginn meiner Promotionszeit und die nach der Zusammenarbeit weiter bestehende Freundschaft und mentale Unterstützung. Ich danke meinen Eltern, auf die ich mich immer verlassen kann, und die zur Fertigstellung der Arbeit viel Ermutigung sowie praktische Unterstützung beigetragen haben. Besonders meiner Mutter danke ich für ihren unermüdlichen Einsatz beim Korrekturlesen. Der Dank, der meinem Ehemann für seine Unterstützung auf so vielen Ebenen gebührt, ist nicht in Worte zu fassen.



# Inhaltsverzeichnis

<b>Abkürzungsverzeichnis</b>	<b>XI</b>
<b>Abbildungsverzeichnis</b>	<b>XIII</b>
<b>Tabellenverzeichnis</b>	<b>XVI</b>
<b>1 Einleitung</b>	<b>5</b>
1.1 Affektive Störungen . . . . .	6
1.1.1 Depressionen . . . . .	7
1.1.2 Bipolare Störungen . . . . .	7
1.1.3 Affektive Störungen als Krankheitsentität? . . . . .	12
1.2 Empathie . . . . .	13
1.3 Empathie bei Affektiven Störungen . . . . .	15
1.3.1 Emotionserkennung in Gesichtern . . . . .	17
1.3.2 Theory of Mind . . . . .	18
1.3.3 State vs. Trait . . . . .	19
1.3.4 Zusammenfassung und Ausblick . . . . .	21
1.4 Messung der Empathie . . . . .	22
1.4.1 Fragebögen vs. behaviorale Messung . . . . .	22
1.4.2 Selbsteinschätzung vs. Fremdbeurteilung . . . . .	25
1.4.3 Ausreichende Schwierigkeit der Instrumente? . . . . .	25
1.4.4 Zusammenfassung und Ausblick . . . . .	26
1.5 Ziele dieser Arbeit . . . . .	26
1.5.1 Ziele der Fragebogenvalidierung . . . . .	27
1.5.2 Ziele der Experimentalstudie . . . . .	27
<b>2 Fragebogenvalidierung: Methoden</b>	<b>29</b>
2.1 Übersetzung und Umformulierung des Fragebogens . . . . .	29
2.2 Stichprobenbeschreibung . . . . .	30
2.2.1 Ein- und Ausschlusskriterien . . . . .	31
2.2.2 Neuropsychologie deskriptiv . . . . .	32
2.3 Messinstrumente und Methoden . . . . .	32
2.3.1 Neuropsychologie . . . . .	32
2.3.2 Empathie und Social Functioning . . . . .	32
2.3.3 Weitere Messinstrumente . . . . .	33
2.4 Untersuchungsablauf . . . . .	33
2.5 Studiendesign . . . . .	34

2.6	Statistische Analysen . . . . .	34
2.6.1	Fragestellung V1 . . . . .	34
2.6.2	Fragestellung V2 . . . . .	35
2.6.3	Fragestellung V3 . . . . .	35
2.6.4	Fragestellung V4 . . . . .	35
2.6.5	Fragestellung V5 . . . . .	35
2.6.6	Fragestellung V6 . . . . .	35
<b>3</b>	<b>Fragebogenvalidierung: Ergebnisse</b>	<b>37</b>
3.1	<i>Fragestellung V1</i> : Deskriptive Beschreibung der Ergebnisse . . . . .	37
3.2	<i>Fragestellung V2</i> : Gütekriterien des EQ . . . . .	38
3.3	<i>Fragestellung V3</i> : Faktorenstruktur des EQ . . . . .	39
3.3.1	EQ-S . . . . .	39
3.3.2	EQ-F . . . . .	41
3.3.3	Auswertungsregeln nach PCA . . . . .	47
3.4	<i>Fragestellung V4</i> : Geschlechtsunterschiede . . . . .	47
3.5	<i>Fragestellung V5</i> : Konstruktvalidität des EQ . . . . .	49
3.5.1	Konvergente Validität . . . . .	49
3.5.2	Diskriminante Validität . . . . .	51
3.6	<i>Fragestellung V6</i> : Explorative Suche nach Einflussfaktoren auf die Varianz des EQ . . . . .	52
3.7	Zusammenfassung der Ergebnisse der Validierungsstudie . . . . .	53
<b>4</b>	<b>Experimentalstudie: Methoden</b>	<b>55</b>
4.1	Stichprobenbeschreibung . . . . .	55
4.1.1	Ein- und Ausschlusskriterien . . . . .	63
4.1.2	Diagnosen und Krankheitsgeschichte . . . . .	64
4.1.3	Neuropsychologie deskriptiv . . . . .	65
4.2	Messinstrumente und -methoden . . . . .	70
4.2.1	Allgemeine Diagnostik . . . . .	70
4.2.2	Neuropsychologie . . . . .	73
4.2.3	Empathie und Social Functioning . . . . .	74
4.2.4	Weitere Messinstrumente . . . . .	78
4.3	Experimentelle Manipulation . . . . .	79
4.4	Untersuchungsablauf . . . . .	79
4.4.1	Ablauf Telefoninterview . . . . .	79
4.4.2	Ablauf Diagnostik t0 . . . . .	80
4.4.3	Ablauf Fragebogenpaket . . . . .	80
4.4.4	Ablauf experimentelle Sitzung t1 und t2 . . . . .	81
4.4.5	Ablauf Nacherhebung t3 . . . . .	82
4.5	Studiendesign . . . . .	82
4.6	Statistische Analysen . . . . .	82
4.6.1	Fragestellung E1 . . . . .	83
4.6.2	Fragestellung E2 . . . . .	83
4.6.3	Fragestellung E3 . . . . .	84
4.6.4	Fragestellung E4 . . . . .	84
4.6.5	Fragestellung E5 . . . . .	84



4.6.6	Fragestellung E6 . . . . .	84
<b>5</b>	<b>Experimentalstudie: Ergebnisse</b>	<b>85</b>
5.1	<i>Fragestellung E1</i> : Fragebogenmaße des Social Functioning . . . . .	86
5.1.1	IIP . . . . .	86
5.1.2	F-SozU . . . . .	88
5.2	<i>Fragestellung E2</i> : Selbst- und Fremdeinschätzung der Empathie mit dem EQ . . . . .	89
5.2.1	Varianzanalyse zur Untersuchung von Gruppenunterschieden im EQ . . . . .	91
5.2.2	Unterschiede zwischen Teilnehmern mit und ohne Fremdbeurteilung . . . . .	92
5.2.3	Regressionsanalyse zur Erklärung der EQ-Varianz . . . . .	92
5.3	<i>Fragestellung E3</i> : Diskrepanz zwischen Selbst- und Fremdeinschätzung der Empathie . . . . .	95
5.3.1	Besteht eine Diskrepanz zwischen Selbst- und Fremdbeurteilung? . . . . .	95
5.3.2	Gruppenunterschiede bezüglich der Diskrepanz zwischen Selbst- und Fremdbeurteilung . . . . .	96
5.4	<i>Fragestellung E4</i> : Behaviorale Messung der Empathie . . . . .	97
5.4.1	CAM Gesamt . . . . .	97
5.4.2	MASC Gesamt . . . . .	103
5.4.3	Zusammenhang behavioraler Maße mit dem EQ . . . . .	104
5.5	<i>Fragestellung E5</i> : Empathie und Induzierbarkeit negativer Emotionen	106
5.5.1	Überprüfung der Emotionsinduktion . . . . .	106
5.5.2	Einfluss der Emotionsinduktion auf die behaviorale Messung der Empathie . . . . .	111
5.5.3	Interindividuelle Unterschiede der Induzierbarkeit negativer Emotionen . . . . .	119
5.6	<i>Fragestellung E6</i> : Zusammenhänge zwischen Empathie und Social Functioning . . . . .	122
5.6.1	Empathie, Social Functioning und Neuropsychologie . . . . .	122
5.6.2	EQ und Social Functioning . . . . .	124
5.6.3	EQ und behaviorale Messung der Empathie . . . . .	125
5.6.4	Behaviorale Messung der Empathie und Social Functioning . .	126
5.7	Zusammenfassung der Ergebnisse der Experimentalstudie . . . . .	127
<b>6</b>	<b>Diskussion</b>	<b>129</b>
6.1	Einordnung und Kritik der Fragebogenvalidierung . . . . .	131
6.2	Einordnung und Kritik der Experimentalstudie . . . . .	132
6.2.1	E1: Fragebogenmaße des Social Functioning . . . . .	132
6.2.2	E2: Selbst- und Fremdeinschätzung der Empathie mit dem EQ	133
6.2.3	E3: Diskrepanz zwischen Selbst- und Fremdeinschätzung der Empathie . . . . .	134
6.2.4	E4: Behaviorale Messung der Empathie . . . . .	134
6.2.5	E5: Empathie und Induzierbarkeit negativer Emotionen . . . .	136
6.2.6	E6: Zusammenhänge zwischen Empathie und Social Functioning	137

6.2.7	Stärken und Schwächen . . . . .	138
6.3	Offene Forschungsfragen . . . . .	141
6.4	Fazit . . . . .	142
<b>Literaturverzeichnis</b>		<b>XIV</b>
<b>A Empathy Quotient</b>		<b>XV</b>
A.1	Selbstbeurteilung . . . . .	XV
A.1.1	Instruktion . . . . .	XV
A.1.2	Items . . . . .	XV
A.1.3	Abschluss . . . . .	XVIII
A.2	Fremdbeurteilung . . . . .	XIX
A.2.1	Instruktion . . . . .	XIX
A.2.2	Items . . . . .	XIX
A.2.3	Abschluss . . . . .	XXIII
<b>B Diagnosen</b>		<b>XXV</b>
<b>C Medikation</b>		<b>XXIX</b>
<b>D Ablauf</b>		<b>XXXI</b>
<b>E PCA-Auswertung</b>		<b>XXXV</b>
E.1	Experimentelle Untersuchung: Fragestellung E2 . . . . .	XXXV
E.1.1	Varianzanalysen zur Untersuchung von Gruppenunterschieden im EQ nach PCA-Auswertung . . . . .	XXXIX
E.1.2	Regressionsanalyse zur Erklärung der EQ-Varianz . . . . .	XL
E.1.3	Unterschiede zwischen Personen mit und ohne Fremdbeurteilung . . . . .	XLII
E.2	Experimentelle Untersuchung: Fragestellung E3 . . . . .	XLII
E.2.1	EQ-Diff für PCA-Auswertung . . . . .	XLIII
E.3	Experimentelle Untersuchung: Fragestellung E5 . . . . .	XLIII
E.3.1	Zusammenhang zwischen EQ und dem Ausmaß der Induzierbarkeit negativer Stimmung . . . . .	XLIII
E.4	Experimentelle Untersuchung: Fragestellung E6 . . . . .	XLIV
E.4.1	IIP . . . . .	XLIV
E.4.2	F-SozU . . . . .	XLIV
<b>F Ergänzende Ergebnisse</b>		<b>XLVII</b>
F.1	Fragebogenvalidierung: Unterschiede zwischen Personen mit und ohne Fremdbeurteilung . . . . .	XLVII
F.2	Fragebogenvalidierung: Fragestellung V5 . . . . .	XLVII
F.2.1	Häufigkeit der Sozialkontakte . . . . .	XLVII
F.2.2	LSAS . . . . .	XLVIII
F.2.3	SIAS-SPS . . . . .	XLVIII
F.2.4	IMI . . . . .	XLVIII
F.2.5	IIP-Cc . . . . .	XLIX
F.3	Experimentelle Untersuchung: Fragestellung E1 . . . . .	L

F.3.1	Betrachtung der Unterskalen des IIP . . . . .	L
F.3.2	Betrachtung der Zusatzskalen und des Teil B des F-SozU . . . . .	LI
F.4	Experimentelle Untersuchung: Fragestellung E4 . . . . .	LIII
F.4.1	Negativity/Positivity Bias in der CAM gesamt . . . . .	LIII
F.5	Experimentelle Untersuchung: Fragestellung E5 . . . . .	LIV
F.5.1	Negativity/Positivity Bias in der CAM . . . . .	LIV



# Abkürzungsverzeichnis

$\rho$ .....	Korrelationskoeffizient Spearmans $\rho$ (schwach ab $\pm 1$ , mittel ab $\pm 3$ , stark ab $\pm 5$ )
APA .....	Amerikanische Psychiatrische Vereinigung
AV .....	Abhängige Variable
BDI .....	Beck-Depressions-Inventar
BR .....	Teilnehmergruppe mit remittierter bipolarer affektiver Störung
CALC .....	Kontrollaufgabe mit mathematischen Gleichungen
CAM .....	Face-Task der Cambridge Mindreading Face-Voice Battery
d .....	Effektstärkemaß Cohens d (kleiner Effekt ab 0.2, mittelgroßer Effekt ab 0.5, starker Effekt ab 0.8)
DSM-5 .....	Diagnostisches und Statistisches Manual Psychischer Störungen, Auflage 5
DSM-IV .....	Diagnostisches und Statistisches Manual Psychischer Störungen, Auflage 4
EQ .....	Empathy Quotient
EQ-Diff .....	Diskrepanz zwischen der Selbst- und Fremdeinschätzung der Empathie anhand des Empathy Quotient
EQ-F .....	Empathy Quotient Fremdbeurteilungsversion
EQ-S .....	Empathy Quotient Selbstbeurteilungsversion
EQ-SQ .....	Empathizing-Systemizing Quotient
F-SozU .....	Fragebogen zur Sozialen Unterstützung
GAF .....	Global Assessment of Functioning
HAMD-17 .....	Hamilton Depression Rating Scale, Version mit 17 Fragen
HAMD-IDS-Kombi	Kombiinterview aus HAMD-17 und IDS-C

ICD-10 .....	Internationale Klassifikation der Krankheiten, Version 10
IDS-C .....	Inventar Depressiver Symptome, Fremdbeurteilungsversion
IIP-C .....	Inventar zur Erfassung interpersonaler Probleme (Kurzversion)
IMI .....	Impact Message Inventory
IQ .....	Intelligenzquotient
IRI .....	Interpersonal Reactivity Index
KG-BR .....	Kontrollgruppe für die Patientengruppe mit remittierter bipolarer affektiver Störung
KG-UR .....	Kontrollgruppe für die Patientengruppe mit remittierter unipolarer affektiver Störung
LSAS .....	Liebowitz Social Anxiety Scale
MASC .....	Movie for the Assessment of Social Cognition
MI .....	Mood Induction, Emotionsinduktion
MSS .....	Manie-Selbstbeurteilungsskala
MWT-B .....	Mehrfachwahl-Wortschatz-Intelligenztest Version B
NMI .....	No Mood Induction, Sitzung ohne Emotionsinduktion
PANAS .....	Positive and Negative Affect Schedule
PCA-Auswertung ..	modifizierte Auswertungsregeln des EQ nach der Durchführung der Hauptkomponentenanalyse
r .....	Korrelationskoeffizient oder Effektstärkemaß Pearsons r (schwach ab $\pm 1$ , mittel ab $\pm 3$ , stark ab $\pm 5$ )
RME .....	Reading the Mind in the Eyes Test
SIAS-SPS .....	Social Interaction Anxiety Scale und Social Phobia Scale
SKID-I und -II ....	Strukturiertes Klinisches Interview für DSM-IV
t0 .....	Diagnostische Sitzung
t1 .....	Erste experimentelle Sitzung
t2 .....	Zweite experimentelle Sitzung
TMT .....	Trail Making Test
ToM .....	Theory of Mind

UA .....	Teilnehmergruppe mit akuter unipolarer affektiver Störung
UR .....	Teilnehmergruppe mit remittierter unipolarer affektiver Störung
UV .....	Unabhängige Variable
VLMT .....	Verbaler Lern- und Merkfähigkeitstest
WHO .....	Weltgesundheitsorganisation
YMRS .....	Young Mania Rating Scale





# Abbildungsverzeichnis

1.1	Komponenten der Empathie . . . . .	15
3.1	Screepplot EQ-S . . . . .	42
3.2	Screepplot EQ-F . . . . .	46
3.3	Geschlechtsunterschiede EQ-S und EQ-F . . . . .	49
4.1	Sozialkontakte . . . . .	57
4.2	Sozialkontakte . . . . .	77
5.1	Boxplot IIP gemittelte interpersonale Probleme . . . . .	87
5.2	Boxplot F-SozU Soziale Belastung . . . . .	89
5.3	Boxplot F-SozU Wahrgenommene soziale Unterstützung . . . . .	90
5.4	CAM mittlere Sicherheit gesamt . . . . .	99
5.5	CAM mittlere Schwierigkeit gesamt . . . . .	100
5.6	Deckeneffekt CALC . . . . .	102
5.7	Überprüfung Stimmungsinduktion negativ . . . . .	108
5.8	Überprüfung Stimmungsinduktion positiv . . . . .	109
5.9	Überprüfung Stimmungsinduktion Negativ . . . . .	111
5.10	Überprüfung Stimmungsinduktion Positiv . . . . .	112
5.11	Einfluss der Stimmungsinduktion auf die CAM . . . . .	113
5.12	Einfluss der Stimmungsinduktion auf den MASC . . . . .	115
5.13	Fehlerarten im MASC . . . . .	117
5.14	Arten der ToM im MASC . . . . .	118
5.15	Stimmungsinduktion . . . . .	120
D.1	Ablauf Hauptstudie . . . . .	XXXII
D.2	Ablauf Diagnostiksitzung . . . . .	XXXIII
D.3	Ablauf Experimentalsitzung . . . . .	XXXIV
E.1	Boxplot kognitive Empathie . . . . .	XL



# Tabellenverzeichnis

1.1	Vergleich der Symptome einer depressiven Episode in ICD-10, DSM-IV und DSM-5 . . . . .	8
1.2	Vergleich der Symptome einer manischen oder hypomanen Episode in ICD-10, DSM-IV und DSM-5 . . . . .	9
2.1	Soziodemographische Beschreibung der Teilnehmer . . . . .	31
3.1	Mittelwerte und Standardabweichungen der EQ-S und EQ-F Scores . . . . .	38
3.2	Ladungen der Hauptkomponentenanalyse des EQ-S . . . . .	40
3.3	EQ-S Items in den 3 Faktoren . . . . .	41
3.4	Ladungen der Hauptkomponentenanalyse des EQ-F . . . . .	44
3.5	EQ-F Items in den vier Faktoren . . . . .	45
3.6	Ergebnisse der multiplen Regression für den EQ-S sowie EQ-F . . . . .	52
4.1	Beschreibung der fünf Teilnehmergruppen: Geschlechterverteilung, Bildungsniveau, Erwerbstätigkeit und Regelmäßigkeit sozialer Kontakte . . . . .	56
4.2	Soziodemographische Beschreibung der Teilnehmergruppen . . . . .	58
4.7	Komorbiditäten, Familienanamnese und Krankheitsgeschichte . . . . .	65
4.8	Aktuelle Psychopharmakologische Medikation der Patientengruppen . . . . .	66
4.9	Aktuelle Symptombelastung . . . . .	67
4.10	Ergebnisse im MWT-B . . . . .	68
4.11	Ergebnisse im TMT (in Sekunden) . . . . .	68
4.12	Deskriptive Ergebnisse des VLMT: Mittelwert (Standardabweichung) . . . . .	69
5.1	Gemittelte interpersonale Probleme . . . . .	86
5.2	F-SozU soziale Belastung und soziale Unterstützung deskriptiv . . . . .	88
5.3	Regulär ausgewertete Summen des EQ-S und EQ-F für die verschiedenen Teilnehmergruppen . . . . .	91
5.4	Homogenitätstest für Probanden mit und ohne EQ-F . . . . .	92
5.5	Ergebnisse der multiplen Regression mit acht Prädiktoren für den regulär ausgewerteten EQ-S . . . . .	93
5.6	Ergebnisse der multiplen Regression mit vier Prädiktoren (post-hoc) für den normal ausgewerteten EQ-S . . . . .	94
5.7	Ergebnisse der multiplen Regression mit acht Prädiktoren für den normal ausgewerteten EQ-F . . . . .	94
5.8	Ergebnisse der multiplen Regression mit vier Prädiktoren (post-hoc) für den normal ausgewerteten EQ-F . . . . .	95

5.9	Deskriptive Beschreibung der Diskrepanz zwischen Selbst- und Fremdeinschätzung der Empathie . . . . .	96
5.10	Summe richtiger Antworten in der CAM Gesamt, deskriptiv . . . . .	97
5.11	Summe richtiger Antworten im MASC Gesamt deskriptiv . . . . .	103
5.12	Ergebnisse der multiplen Regression mit sechs Prädiktoren für den Anteil richtiger Antworten in der CAM . . . . .	105
5.13	Ergebnisse der multiplen Regression mit sechs Prädiktoren für die Summe richtiger Antworten im MASC . . . . .	106
5.14	Ergebnisse der multiplen Regression mit sechs Prädiktoren für die Differenz der positiven Stimmung durch Stimmungsinduktion . . . . .	121
5.15	Ergebnisse der multiplen Regression mit sechs Prädiktoren für die Differenz der negativen Stimmung durch Stimmungsinduktion . . . . .	122
5.16	Korrelationen zwischen EQ-S und EQ-F mit der MWT-B Summe . . . . .	123
5.17	Korrelationen zwischen CAM und MASC mit der MWT-B Summe . . . . .	123
5.18	Korrelationen zwischen F-SozU (wahrgenommene soziale Unterstützung = Unt; soziale Belastung = Bel) und IIP (gemittelte interpersonelle Probleme) mit der MWT-B Summe . . . . .	124
5.19	Korrelationen zwischen EQ-S und EQ-F mit den gemittelten interpersonellen Problemen im IIP . . . . .	124
5.20	Korrelationen zwischen EQ-S und EQ-F mit dem F-SozU . . . . .	125
5.21	Korrelationen zwischen EQ-S und EQ-F mit der CAM und dem MASC	126
5.22	Korrelationen zwischen CAM und MASC mit dem F-SozU . . . . .	127
5.23	Korrelationen zwischen CAM und MASC mit den gemittelten interpersonellen Problemen im IIP . . . . .	127
B.1	Genaue Aufschlüsselung der vergebenen Diagnosen in SKID-I und II . . . . .	XXV
C.1	Aktuelle Psychopharmakologische Medikation der Patientengruppen, detaillierte Aufschlüsselung . . . . .	XXIX
E.1	Summen des EQ-S und EQ-F nach PCA-Auswertung für die verschiedenen Teilnehmergruppen . . . . .	XXXVI
E.2	Unterskalen des EQ-S für die verschiedenen Teilnehmergruppen	XXXVII
E.3	Unterskalen des EQ-F für die verschiedenen Teilnehmergruppen	XXXVIII
E.4	Ergebnisse der multiplen Regression mit acht Prädiktoren für den nach Validierungsregeln ausgewerteten EQ-S . . . . .	XLI
E.5	Ergebnisse der multiplen Regression mit acht Prädiktoren für den nach Validierung ausgewerteten EQ-F . . . . .	XLII
E.6	Homogenitätstest für Probanden mit und ohne EQ-F . . . . .	XLII
E.7	Deskriptive Beschreibung der Diskrepanz zwischen Selbst- und Fremdeinschätzung der Empathie . . . . .	XLIII
F.1	Deskriptive Beschreibung des EQ-S für Probanden mit und ohne vorhandenem EQ-F sowie Ergebnisse des Homogenitätstests . . . . .	XLVII
F.2	F-Soz U weitere Skalen deskriptiv . . . . .	LII
F.3	F-Soz U weitere Skalen deskriptiv . . . . .	LIII

# Zusammenfassung

Zufriedenstellende zwischenmenschliche Beziehungen und Begegnungen sind den meisten Menschen ein wesentliches Bedürfnis. Patienten, die an affektiven Störungen leiden, erleben jedoch zusätzlich zu belastenden Symptomen wie niedergeschlagener Stimmung und Antriebslosigkeit auch Konflikte mit ihrer Familie, dem sozialen und beruflichen Umfeld, bis hin zur sozialen Isolation. Die Frage nach der Ursache der Beeinträchtigung des sozialen Funktionierens von Patienten mit affektiven Störungen bleibt bislang ungeklärt. Als mögliche Erklärung wird eine verminderte Empathiefähigkeit diskutiert.

Empathie ermöglicht es uns, uns in ein Gegenüber hineinzusetzen, dessen Stimmung und Emotionen nachzufühlen und seine Absichten, Ziele und Beweggründe zu verstehen. Im alltäglichen Miteinander ist Empathie eine unverzichtbare und allgegenwärtige Fähigkeit. Die Datenlage zur Empathiefähigkeit von Patienten mit affektiven Störungen ist uneinheitlich. Es gibt sowohl Befunde die für eine verringerte Empathie sprechen, als auch Studien die Hinweise darauf geben, dass Erkrankte besonders empathisch sind. Dies könnte an den sehr heterogenen Stichproben liegen. Patienten mit akuter oder remittierter affektiver Störung, mit unipolarer oder bipolarer Depression werden mitunter nicht getrennt untersucht. Zudem findet die aktuelle Stimmung zum Zeitpunkt der Erhebung zu wenig Beachtung. Im Rahmen der vorliegenden Arbeit sollen diese Punkte differenziert untersucht werden.

Wie lässt sich jedoch eine so grundlegende und wichtige Fähigkeit wie Empathie erfassen und messen? In der bisherigen Forschung wird hierbei häufig auf Fragebögen zur Selbsteinschätzung gesetzt. In der vorliegenden Arbeit wird dieser Ansatz um die Möglichkeit zur Fremdeinschätzung durch Menschen aus dem persönlichen Umfeld der Person ergänzt. Zudem erfolgen direkte Messungen der Erkennung von Emotionen sowie der Fähigkeit, über Emotionen, Absichten und Wünsche nachzudenken.

**Methoden:** Wie lässt sich ein mögliches Defizit der Empathiefähigkeiten messen? Können soziale Schwierigkeiten unipolar depressiver oder bipolarer Patienten mit einem Defizit der Empathiefähigkeiten erklärt werden? Zur Klärung dieser zentralen Fragen werden zwei empirische Untersuchungen durchgeführt: Im Rahmen der Fragebogenvalidierung wird an einer gesunden studentischen Stichprobe eine verbesserte deutsche Übersetzung des Empathy Quotient sowie eine Umformulierung desselben in ein Fremdbeurteilungsmaß validiert. Eine Experimentalstudie untersucht Patienten mit akuter oder remittierter unipolarer Depression, Patienten mit

bipolarer Störung in Vollremission sowie gesunde Kontrollen bezüglich ihrer Empathiefähigkeiten, gemessen mit den beiden Versionen des Empathy Quotient sowie behavioralen Tests zur Emotionserkennungsfähigkeit (Cambridge Mindreading Face-Voice Battery) und zur Fähigkeit, Schlussfolgerungen über mentale Zustände eines anderen Menschen zu treffen (Movie for the Assessment of Social Cognition). Zudem erfolgt eine experimentelle Manipulation der aktuellen Stimmung durch Musik.

**Ergebnisse:** Die Fragebogenvalidierung mit 140 gesunden studentischen Probanden zeigt, dass sowohl die Übersetzung des Empathy Quotient sowie die Fremdbeurteilungsversion als valide angesehen werden können. Beide Versionen verfügen über hohe interne Konsistenzen und weisen Faktorstrukturen auf, die der des englischen Originals weitgehend entsprechen. Sowohl Durchführungs-, Auswertungs- und Interpretationsobjektivität als auch Konstruktvalidität können als gegeben angenommen werden.

Im Rahmen der Experimentalstudie mit insgesamt 203 Teilnehmern finden sich Hinweise auf die erwarteten Defizite der Patientengruppen bezüglich des sozialen Funktionierens. Bezüglich der Selbst- und Fremdeinschätzung der Empathie und der behavioralen Messung der Empathie können jedoch – anders als erwartet – keine Unterschiede zwischen Patienten mit affektiven Störungen und gesunden Kontrollen gefunden werden. Obwohl die Induktion negativer Stimmung erfolgreich gewesen zu sein scheint, findet sich kein Hinweis auf einen Einfluss der aktuellen Stimmung auf die Emotionserkennung oder auf die Fähigkeit Schlussfolgerungen über mentale Zustände anderer Personen zu ziehen.

**Diskussion:** Die Begrenzungen der Studie – wie eventuelle Stichprobenverzerrungen, zur Erfassung möglicher subtiler Defizite möglicherweise nicht geeignete Messmethoden sowie unklare Definitionen der untersuchten Konzepte – sollten bei der Interpretation der Ergebnisse beachtet werden. Auf der Grundlage der vorliegenden empirischen Daten ist jedoch nicht davon auszugehen, dass soziale Schwierigkeiten von Patienten mit affektiven Störungen durch ein Defizit der Empathiefähigkeiten erklärt werden können. Die deutsche Übersetzung und Umformulierung des Empathy Quotient kann hingegen als geeigneter Fragebogen zur Selbst- oder Fremdeinschätzung der Empathie angesehen werden.

# Abstract

Satisfying interpersonal relationships and encounters are a fundamental need of most humans. Patients suffering from affective disorders experience not only burdening symptoms like subdued mood or lethargy. They are also facing interpersonal problems within their family, their social and professional environment and even social isolation. The reason for this decrease in social functioning remains unexplained. Empathy deficits are discussed as one possible explanation.

Empathy allows us to put oneself in the position of other people, to understand their mood and feelings as well as their goals and motivations. Empathy is an essential and ubiquitous skill in our everyday social interactions. Regarding empathy in patients with affective disorders the data are uncertain. Some experimental results suggest decreased empathy, whereas other studies indicate enhanced empathy in patients. This might be due to the fact that very heterogenous samples were examined. Patients with current episodes and in partial or complete remission, with unipolar and bipolar depression are investigated jointly. Furthermore the present mood at the time of the investigation is often ignored. In this work, a differentiated consideration of these issues is taken into account.

How can such a fundamental and important ability like empathy be captured and measured? Previous research has primarily been based on questionnaires and self-assessment. In the present work, this approach is expanded by external assessment by people from the personal environment of the participant. Further, direct measurement of recognition of emotions as well as reasoning about emotions, intentions and wishes in behavioural data are considered.

**Methods:** How can a possible deficit in empathy be captured? Can patients' social difficulties be explained by a deficit in empathy? To clarify these two central questions, two empirical studies are implemented: A validation study examines an enhanced German translation of the Empathy Quotient as well as a reformulated external assessment version in a student population. In an experimental study, patients with unipolar and bipolar depression in current episode or complete remission as well as healthy controls are tested with both versions of the Empathy Quotient as well as with behavioural tests for emotion recognition (Cambridge Mindreading Face-Voice Battery) and for reasoning about mental states of other people (Movie for the Assessment of Social Cognition). Furthermore the present mood is manipulated experimentally by music.

**Results:** The validation study with 140 healthy student participant shows that the German translation as well as the reworded external assessment version of the Empathy Quotient can be considered valid. Both versions seem to have high internal consistencies and factorial structures that comply with the English original. Implementation, evaluation and interpretation objectivity as well as construct validity are assumed as given.

The experimental study with a total of 203 participants provides evidence for the expected social difficulties in patients. Concerning self or external assessment of empathy and behavioural measurement of empathy – differently than expected – no discrepancies between patients and healthy controls were found. The induction of negative mood seems to be effective. Nevertheless the negative mood seems to have no effect on emotion recognition or reasoning about mental states of other people.

**Discussion:** The limitations of this study – for example potential sampling biases, possibly inappropriate measurement methods for the detection of deficits, that may be subtle, or ambiguous definitions of the investigated concepts – should be considered when interpreting the data. However, based on the empiric data at hand, it can not be assumed that patients' social difficulties can be explained by a deficit in empathy. On the other hand, the German translation and reformulation of the Empathy Quotient can be considered as an applicable questionnaire for self-report or external assessment of empathy.



# Kapitel 1

## Einleitung

Zu jedem Zeitpunkt leiden etwa 8.1% der Bevölkerung - also ungefähr sechseinhalb Millionen Menschen in Deutschland - an klinisch relevanter depressiver Symptomatik (Busch et al., 2013). Bei 6.0% der Deutschen, also fast fünf Millionen Menschen, wird innerhalb eines 12-Monatszeitraumes eine unipolare Depression diagnostiziert (Busch et al., 2013). Weitere schätzungsweise 1%, also 800 000 Menschen pro Jahr, erkranken an Bipolaren Störungen (Pini et al., 2005). Das bedeutet, es sind jedes Jahr etwa so viele Personen von affektiven Störungen betroffen, wie zusammengerechnet in Berlin, München und Stuttgart leben. Diese Betroffenen erleben nicht nur belastende Symptome wie niedergeschlagene Stimmung, Antriebslosigkeit, Schlafstörungen und Konzentrationsprobleme. Sie leiden zusätzlich unter sozialen Schwierigkeiten wie sozialem Rückzug, Konflikten mit Familie, sozialem und beruflichem Umfeld bis hin zur sozialen Isolation (Pompili et al., 2013). Wie Maslow (1943) beschreibt, streben die meisten Menschen nach erfüllenden Beziehungen im Rahmen einer Partnerschaft, Familie oder Freundschaft, sobald die grundlegendsten physiologischen und die Sicherheit betreffenden Bedürfnisse befriedigt sind. Affektive Störungen und die damit einhergehenden sozialen Schwierigkeiten und Konflikte beeinträchtigen die Betroffenen somit in der Erfüllung ihrer Grundbedürfnisse. Die Lebensqualität von Patienten mit unipolaren Depressionen ist mindestens genauso stark eingeschränkt, wie bei Patienten mit chronischen körperlichen Erkrankungen, und das soziale Funktionieren spielt hierfür eine bedeutende Rolle (Rapaport et al., 2005). Patienten mit bipolaren Störungen leiden selbst in Phasen der Remission unter einer deutlich eingeschränkten Lebensqualität, auch in Bezug auf soziales Funktionieren (Michalak et al., 2005). Eine Möglichkeit zur Erhöhung der Lebensqualität der enormen Anzahl Betroffener läge also in der Verbesserung des sozialen Funktionierens. Um das soziale Funktionieren zu verbessern, müssen zunächst diejenigen Mechanismen verstanden werden, die zu einer Beeinträchtigung desselben führen.

Verminderte Empathiefähigkeiten werden in Zusammenhang mit den sozialen Schwierigkeiten von Patienten mit affektiven Störungen diskutiert: Die Empathiefähigkeiten scheinen bei unipolaren sowie bipolaren affektiven Störungen auch während Phasen der Remission beeinträchtigt zu sein (Inoue et al., 2004). Wie Inoue et al. (2006) berichten, können Rückfälle unipolar depressiver Patienten anhand von Theory of Mind (ToM) Defiziten vorhergesagt werden: Je niedriger die Empa-

thiefähigkeit ausgeprägt ist, desto wahrscheinlicher sind eine erneute Episode sowie Einschränkungen des sozialen Funktionierens innerhalb eines Jahres. Empathie oder ToM ermöglicht es uns, uns oder anderen Personen mentale Zustände zuzuschreiben, das heißt Annahmen über die Erwartungen, Gefühle oder Wünsche zu treffen (Premack & Woodruff, 1978). Dass ein Defizit dieser Fähigkeiten zu Schwierigkeiten im sozialen Funktionieren führen könnte, erscheint plausibel. Unklar ist jedoch, wie diese Fähigkeiten mit affektiven Störungen zusammenhängen.

In der vorliegenden Arbeit wird daher der Zusammenhang zwischen Empathie und affektiven Störungen genauer untersucht. Im Folgenden werden die theoretischen Grundlagen sowie der bisherige Stand der Forschung ausführlich dargelegt. Zunächst werden die affektiven Störungen, unipolare Depressionen sowie bipolare Störungen genauer beschrieben. Anschließend wird der Begriff der Empathie dargestellt und es werden Befunde zur Empathie bei affektiven Störungen und Möglichkeiten der Messung der Empathie besprochen. Schließlich werden die offenen Fragen und Ziele der vorliegenden Arbeit aufgezeigt.

## 1.1 Affektive Störungen

Die in der vorliegenden Arbeit untersuchten psychischen Erkrankungen werden im Folgenden genauer beschrieben. Während der Durchführung der empirischen Arbeiten ergaben sich grundlegende Änderungen der Diagnosesysteme, weshalb die verschiedenen Versionen und Diagnosekriterien differenziert, sowie die Auswahl der hier verwendeten Variante dargestellt werden.

Zur Kategorisierung psychischer Erkrankungen existieren grundsätzlich zwei Systeme: Das Diagnostische und Statistische Manual Psychischer Störungen (DSM) der Amerikanischen Psychiatrischen Vereinigung (APA) sowie die Internationale Klassifikation der Krankheiten (ICD) der Weltgesundheitsorganisation (WHO). In Forschungsarbeiten wird hauptsächlich das DSM verwendet, während im Deutschen Gesundheitssystem die Kodierung nach ICD verpflichtend ist.

Im Jahr 1994 veröffentlichte die APA die vierte Auflage des Diagnostischen und Statistischen Manuals Psychischer Störungen (DSM-IV; American Psychiatric Association, 1994). Saß et al. (1996) gaben zwei Jahre später die deutsche Übersetzung heraus. Wittchen et al. (1997) veröffentlichte die deutsche Übersetzung des Strukturierten Klinischen Interviews für DSM-IV (SKID), ein Interviewverfahren zur Diagnosestellung psychischer Erkrankungen anhand der Kriterien des DMS-IV. Eine Textrevision des DMS-IV erschien 2000 zunächst auf Englisch und 2003 im Deutschen (American Psychiatric Association, 2000; Saß et al., 2003). Im Folgenden wird ausschließlich auf die Textrevision des DMS-IV Bezug genommen, wenn vom DSM-IV die Rede ist. Die fünfte Auflage des DSM erschien kurz vor Beginn der empirischen Arbeiten auf Englisch (DSM-5; American Psychiatric Association, 2013) und während der Durchführung der Untersuchung auf Deutsch (Falkai & Wittchen, 2015). Da jedoch zur Diagnosestellung im Rahmen der Untersuchung auf das SKID (Wittchen et al., 1997) zurückgegriffen wurde, erfolgte die Diagnostik basierend auf den Kriterien des DSM-IV. Eine an das DSM-5 angepasste Version des SKID erschien erst nach Abschluss der empirischen Arbeit in englischer Sprache (First et al., 2015). Im Deutschen ist sie bis zum aktuellen Zeitpunkt nicht verfügbar.

Das Kapitel V: Psychische und Verhaltensstörungen der ICD-10 (Dilling et al., 1991) befasst sich, ebenso wie das DSM, mit der Kategorisierung psychischer Erkrankungen. In den übrigen Kapiteln sind somatische Erkrankungen sowie zusätzliche Kodierungsmöglichkeiten enthalten. Das Erscheinen einer Revision (ICD-11) wird für 2018 erwartet (Weltgesundheitsorganisation, 2017).

Im Folgenden werden zunächst die einzelnen Krankheitsbilder beschrieben. Anschließend erfolgt eine Diskussion der gemeinsamen oder getrennten Kategorisierung dieser lange als affektive Störungen bekannten Erkrankungen.

### 1.1.1 Depressionen

In Tabelle 1.1 sind die Symptome einer einzelnen depressiven Episode, wie sie von den verschiedenen Diagnosesystemen beschrieben wird, dargestellt.

Sobald ein Patient mehrere depressive Episoden, aber nie manische oder hypomane, erlebt hat, sprechen sowohl ICD-10 als auch DSM-IV und DSM-5 von einer rezidivierenden depressiven Störung.

ICD-10 und DSM-IV beschreiben eine über mindestens zwei Jahre anhaltende depressive Verstimmung, die nie oder nur sehr selten so ausgeprägt ist, die Kriterien einer depressiven Episode zu erfüllen, als *Dysthymia (F34.1)* beziehungsweise *dysthyme Störung (300.4)*. Das DSM-5 bringt hier eine Neuerung mit sich: Die dort aufgeführte *persistierende depressive Störung (Dysthymie)* beschreibt ebenfalls eine anhaltende depressive Verstimmung an der Mehrzahl der Tage über einen mindestens zweijährigen Zeitraum, diese kann jedoch - anders als in ICD-10 und DSM-IV - auch das Ausmaß einer Major Depression annehmen.

### 1.1.2 Bipolare Störungen

In Tabelle 1.2 finden sich die in ICD-10, DSM-IV und DSM-5 beschriebenen Symptome hypomaner oder manischer Episoden. Hierbei unterscheidet die ICD-10 zwischen drei Schweregraden (Hypomanie, Manie ohne psychotische Symptome und Manie mit psychotischen Symptomen) während DSM-IV und DSM-5 ausschließlich zwischen Hypomanie und Manie unterscheiden. Das mögliche Auftreten psychotischer Symptome im Rahmen einer Manie ist jedoch auch in DSM-IV und DSM-5 beschrieben. Allen drei Systemen gemeinsam ist, dass eine Hypomanie ein kürzerer und weniger stark ausgeprägter Zustand gehobener Stimmung als eine Manie ist. Eine Hypomanie führt demnach zu beobachtbaren Auswirkungen auf die Lebensführung, jedoch nicht zu deutlichen Beeinträchtigungen der sozialen Aktivitäten oder der Arbeitsleistung oder gar zu einer Hospitalisierung.

Tabelle 1.1: Vergleich der Symptome einer depressiven Episode in ICD-10, DSM-IV und DSM-5

<b>ICD-10</b>	<b>DSM-IV</b>	<b>DSM-5</b>
Depressive Episode	Major Depression Episode	Major Depression
<i>F32</i>	<i>296.2</i>	
Für einen Zeitraum von mindestens 2 Wochen, mindestens zwei der <b>typischen Symptome</b> :	Mindestens 5 der folgenden Symptome über eine Dauer von mindestens 2 Wochen (darunter mindestens eines der fett gedruckten):	Mindestens 5 der folgenden Symptome über eine Dauer von mindestens 2 Wochen (darunter mindestens eines der fett gedruckten):
· gedrückte Stimmung	<b>1) depressive Verstimmung</b>	<b>1) depressive Verstimmung</b>
· Interessenverlust, Freudlosigkeit	<b>2) Verlust von Interesse oder Freude</b>	<b>2) Verlust von Interesse oder Freude</b>
· Verminderung des Antriebs, erhöhte Ermüdbarkeit	3) Gewichtsverlust/ Gewichtszunahme oder verminderter/gesteigerter Appetit	3) Gewichtsverlust/ Gewichtszunahme oder verminderter/gesteigerter Appetit
und zusätzlich <b>andere häufige Symptome</b> :	4) Schläflosigkeit oder vermehrter Schlaf	4) Insomnie oder Hypersomnie
1) verminderte Konzentration und Aufmerksamkeit	5) psychomotorische Unruhe oder Verlangsamung	5) psychomotorische Unruhe oder Verlangsamung
2) vermindertes Selbstwertgefühl und Selbstvertrauen	6) Erschöpfung oder Energieverlust	6) Müdigkeit oder Energieverlust
3) Schuldgefühle und Gefühle von Wertlosigkeit	7) Gefühl der Wertlosigkeit, Schuldgefühle (auch wahnhaft)	7) Gefühl der Wertlosigkeit, Schuldgefühle (auch wahnhaft)
4) negative und pessimistische Zukunftsperspektiven	8) verminderte Denk- und Konzentrationsfähigkeit oder Entscheidungsunfähigkeit	8) verminderte Fähigkeit zu denken, sich zu konzentrieren oder zu entscheiden
5) Suizidgedanken, erfolgte Selbstverletzung oder Suizidhandlungen	9) wiederkehrende Gedanken an den Tod, Selbstmordgedanken, ein Selbstmordversuch oder ein konkreter Selbstmordplan	9) wiederkehrende Gedanken an den Tod, Suizidvorstellungen, Suizidversuch oder Planung eines Suizids
6) Schlafstörungen		
7) verminderter Appetit		

Tabelle 1.2: Vergleich der Symptome einer manischen oder hypomanen Episode in ICD-10, DSM-IV und DSM-5

Teil 1 ICD-10	DSM-IV	DSM-5
Hypomanie <i>F30.0</i>	Hypomane Episode <i>296.89</i>	Hypomane Episode
<p>Für wenigstens <b>einige Tage anhaltend:</b></p> <ul style="list-style-type: none"> <li>· leicht gehobene Stimmung</li> <li>· gesteigerter Antrieb und Aktivität</li> <li>· auffallendes Gefühl von Wohlbefinden und körperlicher und seelischer Leistungsfähigkeit</li> </ul> <p><b>Zusätzlich möglich:</b></p> <ul style="list-style-type: none"> <li>· gesteigerte Geselligkeit, seltener Reizbarkeit, eingebildetes Auftreten und flegelhaftes Verhalten</li> <li>· Gesprächigkeit</li> <li>· übermäßige Vertraulichkeit</li> <li>· gesteigerte Libido</li> <li>· vermindertes Schlafbedürfnis</li> <li>· Beeinträchtigung der Konzentration und Aufmerksamkeit</li> <li>· etwas übertriebene Geldausgaben</li> </ul>	<p>Für <b>mindestens 4 Tage</b> anhaltende gehobene, expansive oder reizbare Stimmung.</p> <p><b>Zusätzlich</b> mindestens 3 (4 bei reizbarer Stimmung) weitere der Folgenden:</p> <ol style="list-style-type: none"> <li>1) übersteigertes Selbstwertgefühl oder Größenideen</li> <li>2) vermindertes Schlafbedürfnis</li> <li>3) Gesprächigkeit oder Rededrang</li> <li>4) Ideenflucht oder subjektive Erfahrung des Gedankenjagens</li> <li>5) Ablenkbarkeit</li> <li>6) Zunahme zielgerichteter Aktivitäten (sozial, Arbeit/Schule, sexuell) oder psychomotorische Unruhe</li> <li>7) übermäßige Beschäftigung mit angenehmen Aktivitäten</li> </ol>	<p>Anhaltend gehobene, expansive oder reizbare Stimmung <b>und</b> gesteigerte Aktivität oder Energie über <b>mindestens 4 Tage.</b></p> <p><b>Zusätzlich</b> mindestens 3 (4 bei reizbarer Stimmung) weitere der Folgenden:</p> <ol style="list-style-type: none"> <li>1) übersteigertes Selbstwertgefühl oder Größenideen</li> <li>2) vermindertes Schlafbedürfnis</li> <li>3) vermehrte Gesprächigkeit oder Rededrang</li> <li>4) Ideenflucht oder subjektives Gefühl des Gedankenrasens</li> <li>5) erhöhte Ablenkbarkeit</li> <li>6) Zunahme zielgerichteter Aktivitäten (sozial, Arbeit/Schule, sexuell) oder psychomotorische Unruhe</li> <li>7) übermäßige Beschäftigung mit Aktivitäten, die mit hoher Wahrscheinlichkeit unangenehme Konsequenzen nach sich ziehen</li> </ol>

Teil 2		
ICD-10	DSM-IV	DSM-5
Manie ohne psychotische Symptome <i>F30.1</i>	Manische Episode  <i>296.0</i>	Manische Episode
<p>Für <b>mindestens 1 Woche</b>:</p> <ul style="list-style-type: none"> <li>· situationsinadäquat gehobene Stimmung (sorglose Heiterkeit bis zu fast unkontrollierbarer Erregung, auch Reizbarkeit und Aggressivität möglich)</li> </ul> <p><b>Weitere Symptome:</b></p> <ul style="list-style-type: none"> <li>· Überaktivität</li> <li>· Rededrang</li> <li>· vermindertes Schlafbedürfnis</li> <li>· starke Ablenkbarkeit</li> <li>· Verlust üblicher sozialer Hemmungen</li> <li>· übertriebene Selbsteinschätzung, Größenideen oder maßloser Optimismus</li> <li>· Wahrnehmungsstörungen, z.B. Einschätzung von Farben als besonders lebhaft oder Hyperakusis</li> <li>· Beginn undurchführbarer Projekte, leichtsinnige Geldausgabe</li> </ul>	<p>Für <b>mindestens 1 Woche</b> anhaltende gehobene, expansive oder reizbare Stimmung.</p> <p><b>Zusätzlich</b> mindestens 3 (4 bei reizbarer Stimmung) weitere der Folgenden:</p> <ol style="list-style-type: none"> <li>1) übersteigertes Selbstwertgefühl oder Größenideen</li> <li>2) vermindertes Schlafbedürfnis</li> <li>3) Gesprächigkeit oder Rededrang</li> <li>4) Ideenflucht oder subjektive Erfahrung des Gedankenjagens</li> <li>5) Ablenkbarkeit</li> <li>6) Zunahme zielgerichteter Aktivitäten (sozial, Arbeit/Schule, sexuell) oder psychomotorische Unruhe</li> <li>7) übermäßige Beschäftigung mit angenehmen Aktivitäten</li> </ol>	<p>Anhaltend gehobene, expansive oder reizbare Stimmung <b>und</b> gesteigerte Aktivität oder Energie über <b>mindestens 1 Woche</b>.</p> <p><b>Zusätzlich</b> mindestens 3 (4 bei reizbarer Stimmung) weitere der Folgenden:</p> <ol style="list-style-type: none"> <li>1) übersteigertes Selbstwertgefühl oder Größenideen</li> <li>2) vermindertes Schlafbedürfnis</li> <li>3) vermehrte Gesprächigkeit oder Rededrang</li> <li>4) Ideenflucht oder subjektives Gefühl des Gedankenrasens</li> <li>5) erhöhte Ablenkbarkeit</li> <li>6) Zunahme zielgerichteter Aktivitäten (sozial, Arbeit/Schule, sexuell) oder psychomotorische Unruhe</li> <li>7) übermäßige Beschäftigung mit Aktivitäten, die mit hoher Wahrscheinlichkeit unangenehme Konsequenzen nach sich ziehen</li> </ol>

Teil 3 ICD-10	DSM-IV	DSM-5
Manie mit psychotischen Symptomen <i>F30.2</i>		
<hr/> <p><b>Zusätzlich</b> zu den oben genannten möglich:</p> <ul style="list-style-type: none"> <li>· Wahn, zum Beispiel religiöse Wahnvorstellungen welche die eigene Identität betreffen</li> <li>· Halluzinationen</li> <li>· Betroffene können aufgrund der Ideenflucht und Rededrang nicht mehr verstanden werden</li> <li>· Aggression und Gewalttätigkeit</li> <li>· Dehydratation und Verwahrlosung aufgrund einer Vernachlässigung der Nahrungsaufnahme und der persönlichen Hygiene</li> </ul> <hr/>		

Die ICD-10 ermöglicht die Diagnose einzelner manischer oder hypomaner Episoden, sofern zuvor oder später nie andere affektive Episoden auftreten. Sobald mehrere manische, hypomane, gemischte oder depressive Episoden auftreten, wird eine *bipolare affektive Störung (F31)* diagnostiziert.

Im DSM-IV können einzelne manische Episoden als *Bipolar I, einzelne manische Episode (296.0)* kodiert und nach Schwere- oder Remissionsgrad unterteilt werden. Sobald mehrere affektive Episoden in Form von manischen, depressiven Episoden auftreten, spricht das DSM-IV von *Bipolar I Störung* und erfordert die Angabe der jüngsten Episode, sowie des aktuellen Schwere- oder Remissionsgrades (z.B. *Bipolar I Störung, letzte Episode manisch, in Teilremission 296.45*). Von einer *Bipolar II Störung (296.89)* ist im DSM-IV die Rede, wenn bisher mehrere ausschließlich hypomane oder hypomane und depressive affektive Episoden aufgetreten sind.

Auch im DSM-5 wird zwischen Bipolar I und Bipolar II unterschieden: Sobald die Kriterien für mindestens eine manische Episode erfüllt sind, liegt die Diagnose einer *Bipolar-I-Störung* vor. Eine *Bipolar-II-Störung* besteht, wenn mindestens eine hypomane Episode, jedoch keine manischen, und mindestens eine Episode einer Major Depression aufgetreten sind.

Eine anhaltende und fluktuierende affektive Störung wird in der ICD-10 als *Zyklothymia (F34.0)* bezeichnet. Hierbei handelt es sich um eine andauernde Instabilität der Stimmung mit zahlreichen Perioden leichter depressiver Verstimmung und leicht gehobener Stimmung, die jedoch nicht ausgeprägt genug sind um die Diagnose einer Hypomanie oder leichten depressiven Episode zu rechtfertigen. Im DSM-IV

fällt die *Zyklothyme Störung* unter die *anderen bipolaren Störungen (296.80)* und bezeichnet einen mindestens zwei Jahre anhaltenden Zeitraum mit mehreren hypomanen Episoden und zahlreichen Episoden depressiver Verstimmung und Interessenverlust, die jedoch nicht die Kriterien für eine Major Depression erfüllen. Nach DSM-5 ist eine *Zyklothyme Störung* zu diagnostizieren, wenn über einen Zeitraum von zwei Jahren zahlreiche Perioden mit hypomanen oder depressiven Symptomen bestehen, die jeweils jedoch nie die Kriterien einer hypomanen Episode oder einer Episode einer Major Depression erfüllen. In diesem Zeitraum müssen die hypomanen oder depressiven Perioden für mindestens die Hälfte der Zeit vorhanden sein.

### 1.1.3 Affektive Störungen als Krankheitsentität?

Wie oben beschrieben unterscheiden sich die symptomatischen Beschreibungen der einzelnen Krankheitsbilder in den drei Diagnosesystemen nur wenig bis nicht. Die Einordnung in ein gemeinsames Kapitel der affektiven Störungen, wie sie in der ICD-10 und im DSM-IV stattgefunden hat, wird im DSM-5 nicht fortgesetzt. Das DSM-5 unterteilt die beschriebenen Krankheitsbilder in *bipolare und verwandte Störungen* und *depressive Störungen*, unterscheidet also grundsätzlich zwischen rein unipolaren affektiven Störungen und bipolaren Störungen (Ehret & Berking, 2013).

Durch die getrennte Betrachtung der affektiven Störungen soll die Stellung der *bipolaren und verwandten Störungen* zwischen *Schizophrenie-Spektrum und anderen psychotischen Störungen* sowie *depressiven Störungen* hinsichtlich Symptomatik, Familiengeschichte und Genetik verdeutlicht werden (Falkai & Wittchen, 2015). Aktuelle Forschungsergebnisse weisen beispielsweise auf eine gemeinsame genetische Grundlage von affektiven und psychotischen Störungen hin und gehen eher von einem Spektrum der affektiven und psychotischen Störungen aus als, von klar zu trennenden psychischen Erkrankungen (Green et al., 2010). Auch hinsichtlich der Erkrankungsraten von Kindern psychisch erkrankter Eltern zeigt sich nur eine teilweise Diagnose-spezifische Erhöhung des Risikos, sondern ebenfalls ein Zusammenhang der sogenannten schweren psychischen Erkrankungen, das heißt Schizophrenie, bipolare Störungen und Depressionen (Rasic et al., 2013). Eine Metaanalyse von Goldsmith et al. (2016) spricht für ähnliche Veränderungen des Immunsystems durch akute und chronische Erkrankungsphasen depressiver, bipolarer und psychotischer Störungen. Auch auf neuropsychologischer Ebene gibt es Hinweise auf kognitive Einschränkungen bei Patienten mit bipolaren Störungen oder psychotischen Erkrankungen die eher einem Kontinuum folgen (Hill et al., 2013).

Trotz der neuerlichen diagnostischen Trennung der affektiven Störungen in bipolare Störungen und depressive Erkrankungen sowie aufgrund der Hinweise auf eine Anordnung dieser Erkrankungen auf einem gemeinsamen Kontinuum werden in der vorliegenden Arbeit sowohl depressive Patienten als auch Patienten mit bipolarer Störung untersucht. Die bei beiden Krankheitsbildern auftretenden depressiven Episoden stehen im Fokus der Untersuchung. Zudem wird Bezug auf vorherige Forschung genommen, in der affektiven Störungen gemäß der Definition in ICD-10 und DSM-IV untersucht wurden. Obwohl in der vorliegenden Arbeit von affektiven Störungen die Rede ist, wird zwischen unipolaren affektiven Störungen und bipolaren Störungen differenziert, so dass die Untersuchung als zeitgemäß gelten kann.



## 1.2 Empathie

Für den Begriff der Empathie existieren eine Vielzahl unterschiedlicher Definitionen, die die Forschung und Arbeit in diesem Bereich erschweren (Cuff et al., 2016). Cuff et al. (2016) diskutieren und vergleichen 43 verschiedene Definitionen um zu der folgenden Eigenen zu kommen:

Empathy is an emotional response (affective), dependent upon the interaction between trait capacities and state influences. Empathic processes are automatically elicited but are also shaped by top-down control processes. The resulting emotion is similar to one's perception (directly experienced or imagined) and understanding (cognitive empathy) of the stimulus emotion, with recognition that the source of the emotion is not one's own.

Ein Blick zu den Anfängen der wissenschaftlichen Beschäftigung mit dem Konzept der Empathie führt zu den Arbeiten des deutschen Philosophen und Psychologen Theodor Lipps (1851 - 1914), der den Begriff der *Einfühlung* verwendete, welcher von Edward Titchener 1909 mit dem Neologismus *empathy* übersetzt wurde (Montag et al., 2008). Ausgehend von diesen Arbeiten entwickelte sich ein breites Forschungsfeld zur Empathie, das zunächst von Psychoanalytikern wie beispielsweise Sigmund Freud und Otto Fenichel bearbeitet wurde (Brothers, 1989). Entwicklungspsychologische, sozialpsychologische und biologische Arbeiten zur Empathie und damit eine Vielzahl an eher kognitiven sowie eher affektiven Definitionen des Begriffes folgten (Eisenberg & Miller, 1987). Davis (1983) versteht Empathie im weitesten Sinne als diejenigen Reaktionen eines Individuums, die entstehen wenn die Erfahrungen eines anderen Individuums beobachtet werden. Hierbei werden sowohl kognitive als auch affektive Komponenten von Davis als relevante Bestandteile eines multidimensionalen Konstruktes angesehen. Auch aktuelle Arbeiten vertreten zumeist die Ansicht, dass Empathie sowohl affektive als auch kognitive Komponenten enthält. Gleichzeitig hält eine Debatte über die Definition und die Messung der Empathie an (D. Jolliffe & Farrington, 2004). Die affektive Komponente wird als die Fähigkeit beschrieben, die Gefühle anderer zu teilen, während die kognitive Komponente erlaubt, Gefühle anderer zu verstehen und Schlussfolgerungen zu ziehen (Ang & Goh, 2010). Fan et al. (2011) beschreiben in ihrer Metaanalyse Hinweise auf verschiedene den beiden Komponenten unterliegende neuronale Prozesse aber auch auf gemeinsame der Empathie zugehörige Aktivierungsmuster im Gehirn.

Empathie wird als ein wesentlicher Bestandteil des normalen sozialen Funktionierens (*Social Functioning*) angesehen, da sie uns erlaubt, effektiv mit der sozialen Umwelt zu interagieren und zudem dafür sorgt, dass wir anderen Menschen helfen und uns davon abhält, andere zu verletzen (Baron-Cohen & Wheelwright, 2004). Auf Achse V des DSM-IV erfolgt eine globale Beurteilung der Leistungsfähigkeit, die sowohl psychisches als auch soziales und berufliches Funktionieren mit einschließt (Safé et al., 2003). Hierbei weisen soziale Integration und das Vorhandensein konfliktarmer wichtiger zwischenmenschlicher Beziehungen, zum Beispiel zu Freunden oder Familienangehörigen, auf ein gutes Social Functioning hin. Soziale Fähigkeiten wie beispielsweise die Fähigkeit, sich verbal auszudrücken und andere in soziale Interaktionen zu integrieren, die Fähigkeit verbale Äußerungen zu verstehen und ein

allgemeines Wissen über soziale Normen, werden als relevant für Social Functioning erachtet und in Verbindung mit empathischen Fähigkeiten gebracht (Riggio, 1986; Cliffordson, 2002). Die Ergebnisse von Zhou et al. (2002) sprechen für einen positiven Zusammenhang zwischen Empathie und Social Functioning. Je empathischer Grundschul Kinder sind, desto höher scheinen deren soziale Kompetenz und Social Functioning. Ein Mangel an Empathie scheint hingegen mit niedrigem sozialen Funktionieren zusammenzuhängen: Michaels et al. (2014) fanden Hinweise darauf, dass besonders Defizite der kognitiven Empathie erklärend für Einschränkungen des sozialen Funktionierens von Patienten mit Schizophrenie sein könnten. Personen, die andere beleidigen, weisen niedrige Werte besonders der kognitiven Empathie auf (D. Jolliffe & Farrington, 2004). Wie Ang & Goh (2010) zeigen, scheinen niedrige kognitive und affektive Empathie mit einer vermehrten Neigung zu Internet-Mobbing einherzugehen.

Premack & Woodruff (1978) sprachen erstmals von Theory of Mind (ToM) als der Fähigkeit, sich in ein Gegenüber hineinzusetzen. Ein Individuum verfügt laut Premack & Woodruff dann über eine ToM, wenn es in der Lage ist, sich und anderen mentale Zustände zuzuschreiben, also Rückschlüsse über Zustände wie Wollen, Denken oder Glauben zu ziehen, die nicht direkt beobachtbar sind. Sabbagh (2004) unterscheidet zwei aufeinander folgende Komponenten der ToM: Zunächst müssen die mentalen Zustände anderer auf der Basis direkt beobachtbarer Informationen erkannt und entschlüsselt werden (*Decoding*). Daran schließt sich das Schlussfolgern über diese mentalen Zustände an, um das Verhalten anderer erklären oder vorherzusagen zu können (*Reasoning*). Baron-Cohen & Wheelwright (2004) setzen ToM und die kognitive Komponente der Empathie gleich und unterscheiden sie hiermit von der affektiven Komponente. Eine Trennung zwischen der eher kognitiven Fähigkeit, mentale Zustände anderer Personen zu verstehen (kognitive Empathie, ToM oder *Mentalizing*) und der eher affektiven Fähigkeit, die Gefühle anderer Menschen zu teilen oder nachzufühlen (Empathie oder affektive Empathie) wird in der aktuellen Forschung häufig vollzogen (Singer, 2006; Völlm et al., 2006; Goldstein & Winner, 2012). Singer (2006) sieht die beiden Konzepte nicht ausschließlich als getrennt sondern als eng miteinander verflochten an, da diese sowohl neuronal als auch psychologisch verknüpft zu sein scheinen als auch ähnliche Entwicklungsgeschichten in Kindheit und Jugend aufweisen. Völlm et al. (2006) beschreiben ToM als die Fähigkeit, anderen mentale Zustände zuzuschreiben, und Empathie als die Fähigkeit, emotionale Zustände anderer nachzuvollziehen. Wie sie in einer Studie anhand von funktioneller Magnetresonanztomographie zeigen, werden bei Aufgaben die ToM oder Empathie erfordern, ähnliche neuronale Netzwerke aktiviert, die die Basis für das Schlussfolgern über mentale Zustände anderer zu sein scheinen. Aufgaben zur Empathie scheinen zusätzlich Aktivität in Gehirnregionen zu benötigen, die für die Verarbeitung emotionaler Inhalte bekannt sind (Völlm et al., 2006). Shamay-Tsoory (2011) sieht die ToM als einen Bestandteil der kognitiven Empathie und berichtet über unabhängige neuronale Grundlagen der ToM und der affektiven Empathie oder *emotionalen Ansteckung*, die jedoch je nach sozialer Situation zusammenwirken können. Goldstein & Winner (2012) unterscheiden ToM und Empathie, indem sie ToM als das Verstehen mentaler Zustände anderer definieren und Empathie als die Anpassung an den emotionalen Zustand anderer ansehen. In ihrer Studie zeigen sie

jedoch Hinweise auf eine Zunahme beider Fähigkeiten bei Kindern und Jugendlichen durch Schauspiel-Unterricht, was auf einen engen Zusammenhang der beiden Fähigkeiten deutet. Evidenz für eine enge Verknüpfung der beiden Konzepte gibt auch eine Studie von Ibanez et al. (2013), bei der (affektive) Empathie zur Vorhersage der ToM im Rahmen einer Pfadanalyse genutzt werden konnte. Insgesamt spricht der aktuelle Stand der Forschung also für das Vorhandensein einer affektiven sowie einer kognitiven Komponente der Empathie, die eng miteinander verknüpft zu sein scheinen, und für die eine Vielzahl an unterschiedlichen Bezeichnungen existiert.

Baron-Cohen & Wheelwright (2004) definieren Empathie wie folgt und decken dabei sowohl die affektive als auch die kognitive Komponente der Empathie ab:

It allows us to tune into how someone else is feeling, or what they might be thinking. Empathy allows us to understand the intentions of others, predict their behavior, and experience an emotion triggered by their emotion. In short, empathy allows us to interact effectively in the social world.

In der vorliegenden Arbeit soll diese aktuelle Definition der Empathie verwendet werden. In Anlehnung an Baron-Cohen & Wheelwright sind die beiden Komponenten der Empathie in Abbildung 1.1 grafisch dargestellt. Die affektive Komponente ist hierbei die Fähigkeit, eine Emotion zu erleben, ausgelöst durch die Beobachtung oder das Erlangen von Kenntnis über die Emotion eines anderen. Die kognitive Komponente wird gleichgesetzt mit dem Begriff der ToM und ist definiert als die Fähigkeit, zu verstehen oder vorherzusagen, was ein anderer denken, fühlen oder tun könnte (Baron-Cohen & Wheelwright, 2004).

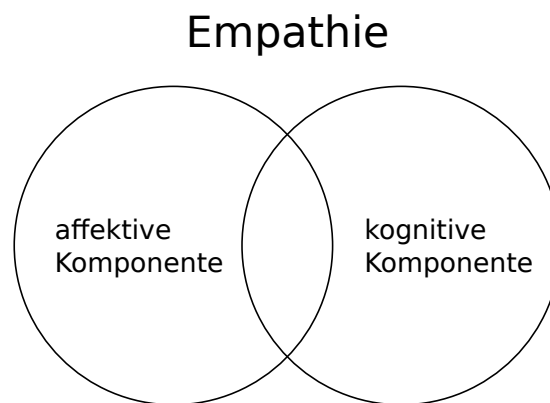


Abbildung 1.1: Die Komponenten der Empathie, in Anlehnung an Baron-Cohen & Wheelwright (2004)

### 1.3 Empathie bei Affektiven Störungen

Eine zu geringe Ausprägung einer so bedeutenden Fähigkeit wie der Empathie führt mit hoher Wahrscheinlichkeit zu Einschränkungen und Beeinträchtigungen im alltäglichen Leben und im zwischenmenschlichen Miteinander. Bekanntestes Beispiel hier-

für sind Personen, die unter einer Autismus-Spektrum-Störung leiden. Wie Baron-Cohen et al. (1985) zeigen konnten, scheint die ToM bei autistischen Kindern beeinträchtigt zu sein: Die autistischen Kinder konnten im Gegensatz zu Kindern mit Down-Syndrom oder zu gesunden Kindern im klassischen Sally-Anne-Test anderen Personen keine Überzeugungen zuschreiben und deren Verhalten nicht vorhersagen (*first-order belief attribution*). Weitere Untersuchungen an älteren Kindern und Jugendlichen mit Autismus gaben Hinweise darauf, dass es sich um eine spezifische Entwicklungsverzögerung handeln könnte. Individuen mit Autismus, die zwar anderen mentale Zustände zuschreiben können (*first-order belief attribution*), scheinen dennoch Schwierigkeiten mit der Lösung komplexerer und alltagsnäherer Aufgaben, bei denen einer anderen Person Wissen über mentale Zustände anderer Personen zugeschrieben werden muss (*second-order belief attribution*), zu haben (Baron-Cohen, 1989). Umfassende Forschung beschäftigte sich mit den Zusammenhängen zwischen Empathie und Autismus. Eine defizitäre ToM wurde jedoch nicht nur bei Autismus sondern auch bei einer Vielzahl an psychischen Erkrankungen wie beispielsweise der Schizophrenie, bipolaren Störungen, Demenzen und anderen, festgestellt (Brüne & Brüne-Cohrs, 2006). So weisen schon erkrankte Patienten mit Schizophrenie im Vergleich zu gesunden Kontrollen bedeutsame Defizite bezüglich der Emotionserkennung im Reading the Mind in the Eyes Test (RME) auf (Kettle et al., 2008). Vor allem aber die Zusammenhänge zwischen affektiven Störungen und Empathie sind laut Brüne & Brüne-Cohrs zu wenig erforscht und sollten in den Fokus zukünftiger Arbeiten genommen werden. Dies ist besonders insofern von großer Relevanz, da ein Defizit an Empathie als mögliche Erklärung der psychosozialen Beeinträchtigungen von Patienten mit affektiven Störungen in Frage kommt (Schreier et al., 2013). Patienten mit unipolaren Depressionen leiden nämlich zusätzlich zu den belastenden Symptomen unter deutlichen Beeinträchtigungen des globalen Funktionsniveaus wie auch des Social Functioning (Judd et al., 2000). Judd et al. finden Hinweise darauf, dass das Ausmaß der Beeinträchtigungen in einem Zusammenhang mit dem Ausmaß der Symptomschwere zu stehen scheint: Die Beeinträchtigungen können im Langzeitverlauf beherrschend und chronisch werden. Wenn jedoch Symptommfreiheit erreicht wird, scheinen die psychosozialen Beeinträchtigungen zu verschwinden. Ähnliche Beobachtungen wurden bei Patienten mit bipolaren Störungen gemacht (Judd et al., 2005): Die psychosozialen Beeinträchtigungen der Betroffenen scheinen von der Symptomschwere und dem Krankheitsverlauf abzuhängen. Eine ausgeprägtere depressive, manische oder hypomane Symptomatik scheint mit vermehrten psychosozialen Beeinträchtigungen einherzugehen. Selbst subklinische depressive Symptome scheinen das Social Functioning zu verschlechtern. Einzig subklinische hypomane Symptome bei Patienten mit Bipolar II führen scheinbar zu verbesserten sozialen Fähigkeiten (Judd et al., 2005).

Beruhend diese Einschränkungen des psychosozialen Funktionsniveaus auf einem Defizit der affektiven oder kognitiven Empathie, des Decodings oder Reasonings? Tritt ein mögliches Defizit der Empathie nur in Zusammenhang mit aktueller affektiver Symptomatik auf oder ist es eine Persönlichkeitseigenschaft Betroffener (state vs. trait)? Im Folgenden werden vorliegende Forschungsarbeiten zum Zusammenhang zwischen Empathie und affektiven Störungen dargestellt.

### 1.3.1 Emotionserkennung in Gesichtern

Viele Arbeiten beschäftigen sich mit der Emotionserkennung in Gesichtern, dem Decoding, also einem Teilbereich der Empathie. Schon Rubinow & Post (1992) fanden Hinweise auf Defizite akut depressiver Patienten bezüglich der Erkennung von Basisemotionen in Gesichtern, jedoch nicht in verbalen Äußerungen. In einer Studie von Surguladze et al. (2004) zeigten gegenwärtig depressive Patienten mit rezidivierender depressiver Störung im Vergleich zu Gesunden in der Emotionserkennung in Gesichtern leichte Defizite bezüglich der Anzahl richtiger Antworten und eine Tendenz wenig ausgeprägte glückliche Gesichter nicht als solche zu erkennen. Auch eine Studie von Joormann & Gotlib (2006) spricht dafür, dass akut depressive Probanden deutlich höhere Intensitäten zur Identifikation von Freude, jedoch signifikant niedrigere Intensitäten zur Erkennung von Traurigkeit benötigten als gesunde Kontrollen. Liu et al. (2012) fanden ebenfalls Evidenz für eine Tendenz depressiver Patienten zur Wahrnehmung negativer Emotionen und eine Überempfindlichkeit gegenüber ärgerlichen Signalen in emotionalen Gesichtern. Die Ergebnisse von Schaefer et al. (2010) sprechen hingegen dafür, dass zwar akut depressive Patienten mit bipolaren Störungen im Vergleich zu Gesunden signifikant höhere Intensitäten zur Emotionserkennung in Gesichtern benötigen, es konnten jedoch weder Unterschiede bezüglich der von unipolar depressiven Patienten benötigten Intensitäten, noch bezüglich der Anzahl richtiger Antworten in allen Gruppen gefunden werden. In diesen Studien wurden jedoch ausschließlich Basisemotionen in Form von statischen oder gemorphten Bildern verschiedener Intensitäten als Stimulusmaterial verwendet, was die ökologische Validität stark einschränkt.

Ökologisch valider ist die Verwendung komplexer Emotionen wie beispielsweise im RME. Lee et al. (2005) verwendeten den RME und fanden Hinweise auf ein defizitäres Decoding akut depressiver Patientinnen im Vergleich zu gesunden Kontrollen, das nicht durch generelle kognitive Beeinträchtigungen oder Wahrnehmungsdefizite erklärt werden konnte. Auch Nejati et al. (2012) konnten Defizite akut depressiver Patienten bezüglich der mit dem RME gemessenen Emotionserkennung replizieren. Der RME beruht jedoch auf statischem Bildmaterial in schwarz-weiß, das ausschließlich die Augenpartie von Schauspielern zeigt, die die gefragte Emotion sehr deutlich darstellen. Eine Übertragbarkeit der Ergebnisse auf das alltägliche zwischenmenschliche Miteinander scheint daher auch hier eingeschränkt.

Im Rahmen zweier Metaanalysen wird von einem Defizit depressiver Patienten bezüglich der Emotionserkennung in Gesichtern mit kleinen bis mittelgroßen Effektstärken ausgegangen (Demenescu et al., 2010; Dalili et al., 2015). Die Metaanalyse von Kohler et al. (2011) spricht für Defizite unipolar depressiver sowie bipolarer Patienten, Emotionen in Gesichtern zu erkennen und zu unterscheiden. Ein Review von Cusi et al. (2012) spricht für Defizite von Patienten mit unipolaren oder bipolaren Störungen bezüglich der Emotionserkennung in Gesichtern, das während akuter Phasen sowie in Remission zu bestehen und mit einer veränderten Gehirnaktivität einherzugehen scheint. Bourke et al. (2010) hingegen fassen die Datenlage in ihrem Review so zusammen, dass wenig Evidenz für ein generelles oder emotionsspezifisches Defizit in der Emotionserkennung unipolarer Patienten besteht, dass jedoch depressive Patienten dazu neigen, Gesichtsausdrücke eher als traurig beziehungsweise als weniger fröhlich einschätzen und eine erhöhte Vigilanz und selektive

Aufmerksamkeit auf traurige und weg von glücklichen Gesichtsausdrücken zeigen. Selbst verschiedene Überblicksarbeiten kommen also zu unterschiedlichen Schlussfolgerungen und die Frage, ob Patienten mit affektiven Störungen ein Defizit in der Emotionserkennung aufweisen, bleibt ungeklärt. Gründe hierfür könnten in der Vielzahl der verwendeten Paradigmen, der Heterogenität der untersuchten Stichproben sowie der zu kleinen Fallzahlen liegen (Bourke et al., 2010; Cusi et al., 2012).

### 1.3.2 Theory of Mind

Anstatt sich allein auf das Decoding zu konzentrieren, wie im vorherigen Unterkapitel beschrieben, gibt es einige Arbeiten, die verschiedene Aspekte der ToM oder Empathie abdecken. Wang et al. (2008) fanden Hinweise auf eine Beeinträchtigung akut depressiver Patienten im Vergleich zu Gesunden bezüglich verschiedener Aufgaben zur ToM (RME, Faux-Pas-Test), die stärker ausgeprägt zu sein scheint, wenn auch psychotische Symptome vorhanden waren. Auch bei der Einschätzung kurzer Videoclips in denen Schauspieler emotionale Ereignisse erzählen, scheinen akut Depressive im Vergleich zu gesunden Kontrollen Defizite bezüglich der Emotionserkennung und dem Mitfühlen aufzuweisen, sowie erhöhten Stress zu erleben (D. Schneider et al., 2012). Zwick & Wolkenstein (2017) konnten im Vergleich akut und remittiert depressiver sowie gesunder Kontrollen Hinweise auf Defizite akut depressiver Patienten bezüglich Decoding (Amsterdam Dynamic Stimuli) und Reasoning (Movie for the Assessment of Social Cognition, MASC) sowie milde Defizite remittiert depressiver Patienten bezüglich Decoding feststellen. Nach einer Metaanalyse von Bora & Berk (2016) kann von bedeutsamen Defiziten bezüglich der ToM depressiver Patienten mit mittelgroßen Effektstärken ausgegangen werden. Beeinträchtigungen konnten sowohl im Decoding als auch im Reasoning gefunden werden und scheinen von der Schwere der depressiven Symptomatik abhängig zu sein.

Selbst im euthymen Zustand scheinen Patienten mit Bipolar I im Vergleich zu gesunden Kontrollen unter leichten Defiziten der ToM zu leiden (Olley et al., 2005). Eine Studie zum Reasoning (MASC) gibt Hinweise darauf, dass bei euthymen Patienten mit bipolarer Störung vor allem die kognitive ToM, nicht aber die affektive ToM beeinträchtigt zu sein scheint (Montag et al., 2010). Hierbei scheint das Ausmaß der Defizite größer zu sein, je höher die Anzahl manischer oder hypomaner Episoden in der Anamnese ist. Auch im Vergleich akut depressiver, akut manischer und remittierter Patienten mit bipolaren Störungen sowie gesunden Kontrollen werden Hinweise auf ein spezifisches Defizit der ToM beobachtet, welches nicht vollständig durch eventuell gleichzeitig bestehende Defizite der Exekutivfunktionen erklärt werden kann (Wolf et al., 2010). Hierbei finden Wolf et al. keine Unterschiede bezüglich des Ausmaßes der ToM zwischen den verschiedenen Krankheitsstadien bipolarer Patienten. Auch aktuelle Ergebnisse von Yang et al. (2017) sprechen für behaviorale sowie in der Selbstbeurteilung auftauchende Defizite der Empathie gegenüber Schmerz von bipolaren Patienten in allen Krankheitsphasen im Vergleich zu gesunden Kontrollen. Bodnar & Rybakowski (2017) hingegen berichten von einem Defizit kognitiver Empathie sowohl während depressiver als auch manischer Phasen einer bipolaren Störung, jedoch von verbesserter affektiver Empathie während manischer Phasen.

Im Gegensatz zu den oben genannten Studien, finden die folgenden Arbeiten keine oder keine generellen Defizite bezüglich der Empathie von Patienten mit affektiven Störungen: In einer Studie von Thoma et al. (2011) schätzten sich akut depressive Patienten im Rahmen einer Selbstbeurteilung (IRI) zwar als empathischer ein als gesunde Kontrollen, in einer behavioralen Messung der Empathie zeigten sich hingegen keine Unterschiede zwischen Patienten und Kontrollen. Ebenfalls keine Hinweise auf behaviorale Defizite der ToM (MASC) fanden Wilbertz et al. (2010) beim Vergleich chronisch depressiver sowie gesunder Teilnehmer. Hier fiel die Selbsteinschätzung (IRI) sowie Fremdeinschätzung der depressiven Patienten jedoch geringer aus als die der Gesunden. Auch Wolkenstein et al. (2011) fanden beim Vergleich akut depressiver und gesunder Probanden keine Hinweise auf ein generelles Defizit bezüglich des Decoding (RME) und nur einen Trend auf Defizite bezüglich des Reasoning (MASC). Es zeigte sich jedoch eine Interaktion zwischen der Valenz der zu erkennenden Emotion und der Gruppenzugehörigkeit: Depressive scheinen negative Stimuli leichter zu erkennen als Gesunde. Die Ergebnisse von Hoffmann et al. (2016) sprechen für intakte empathische Fähigkeiten akut Depressiver, sofern nur einfache Projektionen erforderlich sind. Bei komplexeren Situationen mit inkongruenten emotionalen Zuständen zwischen Selbst und Gegenüber scheinen depressive Patienten laut Hoffmann et al. zu sogenannten *egocentric biases* sowie *altercentric biases* zu neigen. Das heißt, sie schreiben dem Gegenüber die eigene Stimmungslage zu (egocentric bias) oder neigen zu einer erhöhten Zuschreibung der Emotionen anderer als der eigenen (erhöhte emotionale Ansteckung, altercentric bias). Washburn et al. (2016) fanden nur dann Defizite depressiver Patienten im Decoding (RME) oder Reasoning, wenn zusätzlich zur Depression eine soziale Phobie diagnostiziert wurde.

Insgesamt bleiben also auch die Forschungsergebnisse zur Empathie oder ToM von Patienten mit affektiven Störungen uneinheitlich bis widersprüchlich, was an der Vielzahl verwendeter Methoden sowie heterogenen Stichproben liegen könnte.

### 1.3.3 State vs. Trait

Auch gesunde Probanden scheinen unter experimentell induzierter trauriger Stimmung Defizite bezüglich der Emotionserkennung zu entwickeln (Chepenik et al., 2007). Sind mögliche Defizite in der Empathie von Patienten mit affektiven Störungen also nur eine Folge der aktuellen gedrückten Stimmung? Bestehen Defizite über akute Krankheitsphasen hinaus oder gehen sie einer Erkrankung gar voraus? Besonders hinsichtlich der uneinheitlichen Forschungslage kann die Frage nach der Abhängigkeit vom aktuellen Zustand der Patienten möglicherweise Licht ins Dunkel bringen.

Leppänen (2006) berichtet in seinem Review zur Verarbeitung emotionaler Stimuli von einem Aufmerksamkeits- und Gedächtnis-Bias depressiver Patienten hin zu negativen und weg von positiven Reizen, der mit neuronalen Veränderungen einhergeht und auch bei gesunden Personen mit einem Risiko für die Entwicklung einer Major Depression zu beobachten ist. Dies kann also als Hinweis auf einen Persönlichkeitszug depressiver Patienten angesehen werden (*trait*). Auch die Ergebnisse von Milders et al. (2010) sprechen für Empathiedefizite als *trait* depressiver Pati-

enten: Milders et al. fanden bei Patienten mit akuter unipolarer Depression einen Bias hin zur Erkennung trauriger Gesichtsausdrücke, der auch bei der Abnahme der depressiven Symptomatik erhalten blieb.

Ladegaard et al. (2014) berichten Hinweise auf Defizite in allen Bereichen der übergeordneten sozialen Kognitionen, beispielsweise der ToM, schon bei der ersten Episode einer unipolaren depressiven Erkrankung. Die Ergebnisse von Cusi et al. (2011) sprechen für einen Zusammenhang des Ausmaßes der Empathie mit Krankheitsverlauf und Symptomschwere: Je größer die Defizite der Empathiefähigkeit sind, desto schwerere Symptomatik und schwerere Krankheitsverläufe werden beobachtet. Weitere Studien sprechen für einen Zusammenhang von Symptomschwere, Krankheitsverlauf und Empathie: Die Bewältigung einfacherer *first-order* ToM-Aufgaben scheint bei Patienten mit nur milder depressiver Symptomatik intakt, während sie in komplexen Aufgaben zu *second-order* ToM Defizite aufzuweisen scheinen (Cusi et al., 2013). Im Vergleich chronisch depressiver Patienten und Patienten mit episodisch wiederkehrenden Depressionen zeigen sich Hinweise auf eine niedrigere Empathie, niedrigere soziale Fähigkeiten und höhere Werte von Disstress in angespannten sozialen Situationen bei chronisch depressiven Patienten, also denjenigen Patienten mit schwereren Krankheitsverläufen (Domes et al., 2016). Diese Studien können also eher als Evidenz für ein im Auftreten und Ausmaß in Abhängigkeit von der depressiven Symptomatik schwankendes Defizit der Empathiefähigkeit Depressiver gewertet werden (*state*). Mit der state-Annahme vereinbaren lassen sich auch die Ergebnisse von Harkness et al. (2011): Die ToM erwachsener Frauen scheint zwar durch das Vorliegen einer positiven mütterlichen Anamnese für Depressionen verbessert, jedoch durch eine eigene Diagnose einer depressiven Erkrankung verschlechtert.

Evidenz für die state-Hypothese einer defizitären Empathiefähigkeit findet sich auch bei Patienten mit bipolaren Störungen: Kerr et al. (2003) fanden Hinweise auf eine beeinträchtigte ToM bipolarer Patienten im Vergleich zu gesunden Kontrollen, die sowohl in akut manischen als auch akut depressiven Phasen zu bestehen scheinen, jedoch außerhalb akuter Krankheitsepisoden verschwinden. In einer Studie von Purcell et al. (2013) konnten keine Unterschiede bezüglich der Anzahl richtiger Antworten bei einer Aufgabe zur Emotionserkennung (RME) zwischen remittierten Patienten mit bipolarer Störung, mit Major Depression und gesunden Kontrollen gefunden werden. Es gibt hingegen auch empirische Arbeiten, die für Beeinträchtigungen der Empathie als trait bipolarer Patienten sprechen: Ergebnisse von Bora et al. (2005) sprechen für beeinträchtigte Emotionserkennung sowie ToM (RME und Hinting Task) euthymer Patienten mit bipolarer Störung im Vergleich zu Gesunden. Lahera et al. (2012) fanden im Vergleich euthymer Patienten mit bipolarer Störung und gesunder Kontrollen Hinweise auf bedeutsame Defizite in der Emotionserkennung in Gesichtern, in der ToM und niedrigeres psychosoziales Funktionieren und Einschränkungen in der Patientengruppe. Auch Inoue et al. (2004) berichten von in Remission weiter bestehenden Schwierigkeiten unipolarer und bipolarer Patienten bei second-order ToM, jedoch nicht bei first-order ToM. Dies scheint vor allem deshalb sehr relevant, da die Defizite in der ToM, wie eine Katamnese mit derselben Stichprobe zeigte, mit einem erhöhten Rückfallrisiko sowie mit niedrigerem Social Functioning zusammen zu hängen scheinen. Die Studie von Martino et al. (2011) spricht für Einschränkungen der ToM (Faux-Pas-Test) sowie Emotionser-



kennung in Gesichtern bei euthymen Patienten mit Bipolar I oder II im Vergleich zu Gesunden, die jedoch zumindest teilweise von Defiziten in Aufmerksamkeit und Exekutivfunktionen sowie durch psychopharmakologische Behandlung mediert werden. Van Rheeën & Rossell (2014) beobachteten bei bipolaren Patienten in akuten Krankheitsphasen und in Remission im Vergleich zu gesunden Kontrollen ein über verschiedene Aufgaben konsistentes Defizit der Emotionserkennung in Gesichtern, das nicht von der aktuellen Stimmung abzuhängen scheint. Mercer & Becerra (2013) fassen in ihrem Review Ergebnisse zusammen, die für Defizite Bipolarer bezüglich Decoding und Reasoning sprechen, die auch außerhalb akuter Episoden anzuhalten scheinen. Die Ergebnisse der Metaanalyse von Samamé et al. (2012) sprechen ebenfalls für eine defizitäre Emotionserkennung in Gesichtern bei euthymen Patienten mit bipolarer Störung, jedoch mit kleinen Effektstärken und einer sehr heterogenen Verteilung der Effektstärken. Die Autorin kritisiert den Mangel an etablierten Instrumenten zur Erhebung sozialer Kognition sowie die zu seltene Beachtung potentieller Moderatoren (Samamé et al., 2012).

#### 1.3.4 Zusammenfassung und Ausblick

Zu den offenbar komplexen Zusammenhängen zwischen Empathie und affektiven Störungen bleiben also weiterhin viele offene Fragen. Es gibt, wie oben zusammengefasst, sowohl Studien deren Ergebnisse für als auch solche die gegen Defizite in einzelnen oder mehreren Bereichen der Empathie sprechen. Die genutzten Methoden sind jedoch uneinheitlich, wenig ökologisch valide und können schwer untereinander verglichen werden. Zudem werden heterogene Stichproben und kleine Fallzahlen verwendet. Viele Studien konzentrieren sich zudem auf nur eine spezifische Patientengruppe, also entweder auf unipolare oder bipolare Patienten, auf akut erkrankte oder remittierte Patienten.

Da dieser Bereich jedoch nicht zuletzt für die Therapie affektiver Störungen von hoher Relevanz ist (Dykieriek et al., 2012), und auch zur Klärung der Gemeinsamkeiten und Unterschiede der verschiedenen Erkrankungen auf dem Kontinuum von Schizophrenie, bipolaren Störungen und Major Depressionen beitragen könnte (Derntl et al., 2012), ist weitere Forschung unbedingt erforderlich. Diese sollte insbesondere bei der Auswahl der Maße der Empathie auf implizitere und ökologisch valide Methoden setzen (Schreiter et al., 2013), potentiellen Moderatoren ausreichende Beachtung schenken (Samamé et al., 2012) und klar zwischen Diagnosen sowie aktuellen Schwere- oder Remissionsgraden differenzieren. In der vorliegenden Arbeit wird daher ein breites Spektrum verschiedener, insbesondere auch ökologisch valider, Methoden zur Messung der Empathie eingesetzt. Diese werden nicht nur für sich gesehen betrachtet, sondern auch mit Maßen des Social Functioning in Verbindung gebracht. Weiterhin werde sowohl unipolare als auch bipolare Patienten in verschiedenen Krankheitsphasen rekrutiert und miteinander sowie mit gematchten gesunden Kontrollen verglichen. Neben dieser umfassenden Herangehensweise sowohl an die Messung der Empathie als auch an die verschiedenen Patientengruppen, hebt sich die vorliegende Arbeit durch eine experimentelle Manipulation der aktuellen Stimmung von bisherigen Forschungsarbeiten ab. Es konnte gezeigt werden, dass Musik eine effektive Möglichkeit zur Induktion verschiedener Stimmungen zu

sein scheint (Västfjäll, 2001), die in verschiedenen Arbeiten zu anderen Forschungsbereichen Anwendung findet (Sanchez et al., 2014; Fairclough et al., 2014). In der vorliegenden Arbeit erfolgt daher eine Induktion negativer Stimmung anhand von Musik.

## 1.4 Messung der Empathie

Die Verfahren zur Messung der Empathie sind so vielfältig und verschieden wie die große Anzahl existierender Definitionen (vgl. Kapitel 1.2 sowie Baldner & McGinley, 2014). Im Folgenden werden verschiedene Ansätze einander gegenüber gestellt und diskutiert werden.

### 1.4.1 Fragebögen vs. behaviorale Messung

Eine grundsätzliche Unterscheidung der Messinstrumente zur Empathie ist die Trennung zwischen Fragebögen und behavioralen Maßen. Im Rahmen von Fragebögen werden die zumeist introspektiven Einschätzungen der Probanden erfragt, während behaviorale Maße auf die direkte Beobachtung verschiedener Aspekte der Empathie abzielen.

Ein bekannter Fragebogen zur Empathie ist der Interpersonal Reactivity Index (IRI; Davis et al., 1980; Davis, 1983). Der IRI erfasst Empathie als multidimensionales Konstrukt auf den vier Subskalen *perspective taking*, *fantasy*, *empathic concern* und *personal distress*. Der Autor kritisiert die bis zur Entwicklung des IRI herrschende Konzentration auf nur einzelne Aspekte der Empathie oder die Zusammenfassung unterschiedlicher Aspekte zu einem Gesamtwert, denen durch die Einführung der Multidimensionalität begegnet werden soll. Eine deutsche Übersetzung und Überarbeitung mit guten Werten für Reliabilität, Validität und Itemtrennschärfen liegt vor (Paulus, 2009). Mischo (2003) stellt die Inhalts- sowie Konstruktvalidität vieler Selbstbeurteilungsinstrumente zur Empathie in Frage. Auch Baron-Cohen & Wheelwright (2004) beanstanden die zu weit gefassten Konstrukte älterer Fragebögen. Laut den Autoren messen die meisten bestehenden Selbstbeurteilungsinstrumente keine oder nicht nur Empathie sondern auch soziale Regeln, kulturelle Konventionen, strategisches Denken oder soziale Fähigkeiten insgesamt. Der IRI wird von Baron-Cohen & Wheelwright als anderen Fragebögen überlegen, aber dennoch zu weit gefasst angesehen. Daher wurde der Empathy Quotient (EQ; Baron-Cohen & Wheelwright, 2004) entwickelt, der sich als reliabel, valide und mit einer sinnvollen Faktorenstruktur zeigte und in gesunden wie auch Patientenstichproben einsetzbar scheint (Lawrence et al., 2004). Der EQ misst Empathie eindimensional. Diese setzt sich jedoch laut Lawrence et al. aus den unterliegenden Faktoren *cognitive empathy*, *emotional reactivity* und *social skills* zusammen. Eine deutsche Übersetzung wurde zwar erstellt (de Haen, o. J.) aber bisher nicht validiert. Samson & Huber (2010) konnten mit einer deutschen Kurzform aus EQ und Systemizing Quotient (SQ) einen Selbstbeurteilungsfragebogen mit zufriedenstellenden psychometrischen Eigenschaften validieren, der sich zur Messung der Empathiefähigkeiten sowie der Fähigkeiten Systeme zu verstehen, erstellen, vorhersagen und kontrollieren, eignet (EQ-SQ).

Baldner & McGinley (2014) hinterfragen in ihrer Arbeit die große Zahl verschiedener Selbstbeurteilungsmaße der Empathie, die auf der Vielfalt der Definitionen für das Konzept der Empathie beruht. In einer groß angelegten Studie an fast 500 Teilnehmern untersuchten die Autoren eine Auswahl von Fragebögen auf Korrelationen sowie unterliegende Faktoren (*EQ*, *IRI*, *How I feel in Different Situations Scale*, *Basic Empathy Scale*, *Toronto Empathy Questionnaire*, *Trait Sympathy Scale* und *Situational Sympathy Adjectives*). Dabei konnten sie lediglich niedrige bis moderate Korrelationen zwischen den verschiedenen Instrumenten beobachten und fanden zudem Items und Faktoren, die nicht zum Konzept der Empathie passen, jedoch große Teile der Varianz aufklären. Baldner & McGinley sehen außerdem die affektive sowie kognitive Komponente der Empathie aber auch die Fähigkeit zwischen Selbst und Gegenüber zu unterscheiden in aktuellen Fragebögen zur Empathie unterrepräsentiert.

Neben der Konstrukt- und Inhaltsvalidität existieren weitere zu beachtende Aspekte bei der Anwendung von Fragebögen zur Messung der Empathie: Das Problem der sozialen Erwünschtheit, also die Tendenz, in Selbstbeurteilungsinstrumenten so zu antworten, wie man gerne nach außen hin wirken möchte, ist schon seit langem bekannt (Crowne & Marlowe, 1960). Lawrence et al. (2004) empfehlen daher, bei der Messung von Empathie die Tendenz zu sozial erwünschten Antworten mit zu erheben. Baldner & McGinley (2014) hingegen fanden in ihrer Studie wenig Hinweise auf eine Beeinflussung der Ergebnisse von Selbstbeurteilungsinstrumenten zur Empathie durch soziale Erwünschtheit. Eisenberg & Fabes (1990) geben weiterhin zu bedenken, dass der erfolgreiche Einsatz von Fragebögen zur Selbstbeurteilung davon abhängt, ob die Probanden überhaupt in der Lage sind, die eigenen Fähigkeiten einschätzen zu können. Aber auch physiologische oder behaviorale Maße sind laut den Autoren nicht unproblematisch. So werden beispielsweise sichtbare emotionale Reaktionen der Gesichtsmuskeln mit zunehmendem Alter mehr und mehr unterdrückt und. Eisenberg & Fabes empfehlen daher die Kombination verschiedener Verfahren und Methoden, um einen umfassenden Blick auf die Empathiefähigkeiten der Probanden zu erhalten.

Die direkte Beobachtung der Empathie im Rahmen sogenannter behavioraler Maße hat - wie schon für die Fragebögen beschrieben - viele Gesichter: Da sich die behavioralen Maße oft auf einzelne Aspekte der Empathie konzentrieren, unterscheiden sie sich zum Teil enorm. Zudem besteht auch hier Uneinigkeit bezüglich der Definition sowie darüber, welches Instrument was misst (Thoma et al., 2013).

Ein großer Teil der Instrumente zur behavioralen Messung der Empathie erfasst diese über die Fähigkeit, Emotionen in Gesichtern zu erkennen (*Decoding*). Ekman & Friesen (1971) veröffentlichten einen über Kulturen hinweg einsetzbaren Satz von Schwarz-Weiß-Fotografien mit emotionalen Gesichtsausdrücken (die sogenannten Basisemotionen Freude, Wut, Ekel, Furcht, Verachtung, Traurigkeit und Überraschung). Diese Basisemotionen kommen seither in verschiedensten Studien zur Empathie oder Emotionserkennung zur Anwendung (z.B. Joormann & Gotlib, 2006; Bourke et al., 2010; Dalili et al., 2015). Baron-Cohen et al. (2001) geben jedoch zu bedenken, dass Basisemotionen in voller Ausprägung zur Erfassung subtiler Defizite in der Emotionserkennung nicht ausreichend sein könnten, da sie zu leicht erkennbar sind. Sie entwickelten daher ein Messinstrument zur Erkennung von

komplexen Emotionen in Fotografien von Gesichtern oder Augenregionen (Reading the Mind in the Eyes Test, RME Baron-Cohen et al., 1997, 2001). Da die statische Präsentation von emotionalen Gesichtsausdrücken in der Augenregion jedoch wenig ökologische Validität aufweist und auf eine enge Auswahl an Emotionen begrenzt bleibt, veröffentlichten (Golan et al., 2006) ein Maß zur Erkennung komplexer Emotionen in kurzen Videos emotionaler Gesichtsausdrücke (*Face-Task*) sowie kurzen Sprachaufnahmen (*Voice-Task*), die Cambridge Mindreading (CAM) Face-Voice Battery.

Behavioralen Maßen des Decoding gegenüber stehen Instrumente die *Reasoning* messen, oder auch Decoding und Reasoning vereinen. Diesen zuzuordnen ist beispielsweise der klassische Sally-Anne-Test von Baron-Cohen et al. (1985), der first-order false beliefs anhand eines Puppenspiels erfasst. Perner & Wimmer (1985) entwickelten einen Test zur Messung von second-order false beliefs, der höhere Empathiefähigkeiten erfordert als der Sally-Anne-Test. Da manche Personen mit Schwierigkeiten im alltäglichen sozialen Miteinander haben, aber dennoch Aufgaben lösen können, die sowohl first- als auch second-order false beliefs erfassen, entwickelte Happé (1994) die *Strange Stories*: Geschichten mit alltäglichen Situationen in denen die Protagonisten etwas sagen, das sie nicht wörtlich meinen (z.B. Lüge, Witz, Missverständnis). Zudem existieren Kontrollgeschichten zum Verständnis physikalischer Ereignisse. Die Autorin beschreibt den Test als realistisch und ökologisch valide, da der Kontext sozialer Situationen mit beachtet werden muss, um die Aufgabe erfolgreich lösen zu können. T. Jolliffe & Baron-Cohen (1999) und auch White et al. (2009) konnten die Ergebnisse von Happé replizieren und verstärken damit die Evidenz für die Eignung der Strange Stories zur Erfassung subtilerer Defizite der ToM. In eine ähnliche Richtung geht der Faux-Pas-Test Gregory et al. (2002) bei dem die Probanden Geschichten zu sozialen Interaktionen in denen jemand unabsichtlich etwas sagt, das er nicht hätte sagen sollen, von Kontrollgeschichten mit kleinen Konflikten aber ohne Faux Pas unterscheiden können müssen. Die Ergebnisse von Gregory et al. (2002) sprechen dafür, dass der Test in der Lage ist, Defizite der ToM zu erfassen, ohne dass die Exekutivfunktionen intakt sein müssen. Doch auch die ökologische Validität erzählter oder selbst gelesener Geschichten kann in Frage gestellt werden, weshalb Dziobek et al. (2006) ein Video-basiertes Verfahren zur behavioralen Messung sowohl des Decoding als auch des Reasoning vorstellten. Der Movie for the Assessment of Social Cognition (MASC) ist ein Filmtest, in dem vier Personen sich zu einem gemeinsamen Abendessen vereinbaren. Der Film stoppt an verschiedenen Stellen und die Probanden sollen Fragen zu den mentalen Zuständen der Protagonisten beantworten. Hierbei werden sowohl Gefühle (Decoding), als auch Gedanken und Absichten (Reasoning) der Charaktere erfragt. Der MASC deckt verschiedene Valenzen und klassische Konzepte sozialer Kognition (z.B. Faux Pas, Metaphern, Sarkasmus) ab und die Items weisen verschiedene Schwierigkeitsgrade auf. Die Daten von Dziobek et al. (2006) sprechen dafür, dass der MASC im Vergleich zu RME, Strange Stories und Erkennung von Basisemotionen die größte Sensitivität für subtile Defizite aufweist. Seit der Entwicklung wird der MASC an verschiedenen Stichproben erfolgreich eingesetzt (z.B. Wilbertz et al., 2010; Wolkenstein et al., 2011; Zwick & Wolkenstein, 2017; Fossati et al., 2017).

Die Durchführung behavioraler Maße ist oft mit einem höheren zeitlichen, personellen und instrumentellen Aufwand verbunden. Zudem konnten Melchers et al. (2015) beobachten, dass behaviorale Maße (RME, CAM und rotation task) wenig untereinander und auch wenig mit Fragebogenmaßen der Empathie korrelieren, was für den Einsatz zur Klärung spezifischer Fragen spricht. Melchers et al. sehen verschiedene Fragebögen zur Empathie (IRI und EQ) hingegen als valide und für den Einsatz in klinischer Forschung geeignet an. Thoma et al. (2013) empfehlen, die Leistungen der Probanden in verschiedenen behavioralen Maßen und Fragebögen miteinander in Verbindung zu bringen, um verlässliche Daten zu den Empathiefähigkeiten der Probanden zu erhalten.

### 1.4.2 Selbsteinschätzung vs. Fremdbeurteilung

Klassische Fragebögen zur Empathie Erwachsener sind häufig als Selbstbeurteilungsinstrumente konzipiert (z.B. Paulus, 2009; Baron-Cohen & Wheelwright, 2004; Spreng et al., 2009; D. Jolliffe & Farrington, 2006). Problematisch können hierbei jedoch soziale Erwünschtheit sowie mangelnde Fähigkeiten zur Einschätzung der eigenen Stärken und Schwächen sein (Lawrence et al., 2004; Eisenberg & Fabes, 1990). Es wird angenommen, dass das Ausmaß der Empathie die Qualität der sozialen Beziehungen maßgeblich mit beeinflusst, beziehungsweise dass ein Defizit der Empathiefähigkeit zu sozialen Schwierigkeiten führen kann (Zhou et al., 2002; Baron-Cohen & Wheelwright, 2004; D. Jolliffe & Farrington, 2004). Es liegt also nahe, die sozialen Interaktionspartner einer Person mit einzubeziehen, wenn das Ausmaß der Empathie dieser Person eingeschätzt werden soll. Schon Gagan (1983) erarbeiteten ein Maß zur Fremdbeurteilung der Empathie von Krankenschwestern durch Patienten oder Simulationspatienten. Auch zur Erhebung der Empathie von Kindern oder Jugendlichen existieren Fremdbeurteilungsinstrumente, die von Eltern oder anderen nahestehenden Personen ausgefüllt werden können (Auyeung et al., 2009, 2012). Mischo (2003) formulierte bestehende Selbstbeurteilungsinstrumente zur Fremdbeurteilung um, indem „ich“ durch „er/sie“ ausgetauscht wurde, und setzte diese an einer studentischen Stichprobe ein. Die Teilnehmer wurden gebeten, die Fremdbeurteilungsversion an Freunde, Bekannte oder Eltern auszuhändigen und in einem verschlossenen Umschlag wieder zurückgeben zu lassen. Die Ergebnisse von Mischo sprechen dafür, dass durch die Fremdbeurteilung relevante Aspekte ergänzt werden können. Mischo empfiehlt daher, die Selbstbeurteilung der Empathie durch Fremdbeurteilungsmaße sowie behaviorale Instrumente zu ergänzen.

### 1.4.3 Ausreichende Schwierigkeit der Instrumente?

Aufgrund der Vorreiterrolle der Forschung zu Autismus-Spektrum-Störungen und Empathie, sind viele Instrumente für den Einsatz an autistischen Stichproben entwickelt worden (vgl. Kapitel 1.3 sowie Brüne & Brüne-Cohrs, 2006). Beispiele für solche Instrumente sind der Sally-Anne-Test (Baron-Cohen et al., 1985), die Strange Stories (Happé, 1994) oder der RME (Baron-Cohen et al., 2001). Schon bei hochfunktionalen Erwachsene mit Autismus-Spektrum-Störungen können manche Verfahren nicht herausfordernd genug sein - trotz sozialer Schwierigkeiten im Alltag,

können diese Patienten die klassischen Aufgaben erfolgreich lösen (Dziobek et al., 2006). Auch bei Patienten mit affektiven Störungen könnte es zu einem Deckeneffekt bei Verwendung zu einfacher Maße der Empathie oder Emotionserkennung kommen (Zwick & Wolkenstein, 2017). Es scheint daher sinnvoll, auf ausreichende Schwierigkeit der Instrumente zu achten, um eine Detektion der eventuell subtilen Defizite der Probanden zu ermöglichen: Verfahren, die unter besonderer Berücksichtigung eines ausreichenden Schwierigkeitsgrades entwickelt wurden sind beispielsweise die CAM (Golan et al., 2006) sowie der MASC (Dziobek et al., 2006).

#### 1.4.4 Zusammenfassung und Ausblick

Zusammenfassend lässt sich feststellen, dass Empathie - ein so breites Konzept mit verschiedenen Komponenten - schwerlich durch den Einsatz eines einzigen Messinstrumentes erfasst werden kann. In der vorliegenden Arbeit werden daher mehrere Instrumente und Methoden angewendet und in Zusammenhang gebracht, um einen umfassenden Blick auf die empathischen Fähigkeiten der Probanden zu erhalten. Der Empfehlung, Fragebögen und behaviorale Maße zu kombinieren, wird gefolgt (Thoma et al., 2013). Zudem werden Selbstbeurteilungsinstrumente durch eine Fremdbeurteilung ergänzt (Mischo, 2003). Hierfür werden eine Umformulierung des EQ (Baron-Cohen & Wheelwright, 2004) in eine Fremdbeurteilungsversion sowie die deutsche Übersetzung der Selbstbeurteilungsversion validiert werden. Gerade bei der Auswahl der behavioralen Messinstrumente wird auf einen ausreichenden Schwierigkeitsgrad geachtet werden. Weiterhin werden die umfassend erhobenen Daten zur Empathie in Zusammenhang zu Maßen des Social Functioning gebracht, um mögliche Empathie-Defizite mit realen sozialen Interaktionen oder Schwierigkeiten in Verbindung bringen zu können.

### 1.5 Ziele dieser Arbeit

Wie dargelegt, scheinen die Zusammenhänge zwischen affektiven Störungen und Empathie bisher nicht ausreichend geklärt. Die vorliegende Arbeit soll anhand gründlicher Trennung der verschiedenen Diagnosen und Krankheitsphasen sowie durch Verwendung einer ausgewogenen und an die Stichprobe angepassten Auswahl an Messverfahren der Empathie einen differenzierten Blick auf die Zusammenhänge zwischen affektiven Störungen und Empathie ermöglichen. Hierfür werden zwei aufeinander folgende empirische Arbeiten durchgeführt. Im Rahmen einer Validierungsstudie werden eine deutsche Übersetzung des EQ als Selbstbeurteilungsversion sowie eine Umformulierung desselben als Fremdbeurteilungsversion an einer gesunden Stichprobe validiert. Anschließend werden diese Fragebögen sowie behaviorale Maße der Empathie und Daten zum Social Functioning an einer Stichprobe aus Patienten mit unipolaren und bipolaren Störungen in verschiedenen Krankheitsphasen sowie gematchten gesunden Kontrollen eingesetzt. In dieser Experimentalstudie wird zudem der Einfluss der aktuellen Stimmung durch eine Induktion negativer Stimmung untersucht werden. Im Folgenden werden die Ziele und Hypothesen der Fragebogenvalidierung sowie der Experimentalstudie genauer beschrieben.

### 1.5.1 Ziele der Fragebogenvalidierung

Wie lässt sich Empathie mit einem Fragebogen messen? Im Rahmen der ersten Studie wird die deutsche Übersetzung des EQ sowohl in der ursprünglichen Selbstbeurteilungsversion (EQ-S) als auch in einer neu erstellten Fremdbeurteilungsversion (EQ-F) an einer gesunden Stichprobe validiert. Dabei werden die interne Konsistenz, sowie die Konstruktvalidität und die Faktorstruktur untersucht. Zudem werden explorativ mögliche Einflussfaktoren auf die Ergebnisse im EQ identifiziert. Die folgenden Fragen stehen im Fokus der Validierungsstudie:

- V1) Wie lassen sich die Ergebnisse des EQ-S und -F deskriptiv beschreiben? Sind die Rohwerte vergleichbar mit den vom englischen Original bekannten Werten?
- V2) Wie fallen die klassischen psychometrischen Gütekriterien für die beiden Versionen des deutschen EQ aus?
- V3) Welche Faktorstruktur weisen EQ-S und -F auf? Sind die Faktorstrukturen des EQ-S und -F vergleichbar mit der des englischen Originals?
- V4) Sind die aus Untersuchungen mit dem englischen Original berichteten Geschlechtsunterschiede auch mit dem EQ-S und -F messbar?
- V5) Sind EQ-S und -F konstruktvalide, messen sie also tatsächlich Empathie?
  - a) Konvergente Validität: Messen EQ-S und -F dasselbe Konstrukt wie andere Maße der Empathie (RME, CAM und die Empathizing-Unterskala des EQ-SQ)?
  - b) Diskriminante Validität: Messen EQ-S und -F andere Konstrukte als Maße der Intelligenz (MWT-B) und des Systemizing (Unterskala des EQ-SQ)?
- V6) Welche Einflussfaktoren zeigen sich in einer explorativen regressionsanalytischen Betrachtung des EQ?

### 1.5.2 Ziele der Experimentalstudie

Nach Abschluss der Validierungsstudie wird der neu entwickelte und validierte EQ-S und EQ-F im Rahmen der Hauptstudie an einer Stichprobe aus Patienten mit affektiven Störungen sowie gematchten gesunden Kontrollen eingesetzt. Maße des Social Functioning werden eingesetzt, um erstens die Stichprobe auf die typischerweise bei Patienten vorhandenen sozialen Schwierigkeiten zu untersuchen und zweitens Zusammenhänge zwischen Empathie und Social Functioning zu untersuchen. Zur Messung der Empathie kommen nicht nur beide Versionen des EQ sondern auch behaviorale Maße der Empathie mit ausreichenden Schwierigkeitsgraden und hoher ökologischer Validität zur Anwendung. Zudem werden unipolar depressive und bipolare Patienten sowie verschiedene Krankheitsphasen differenziert betrachtet. Eine experimentelle Manipulation der aktuellen Stimmung beleuchtet die Zusammenhänge zwischen Empathie und affektiven Störungen näher. Folgende Hypothesen werden im Rahmen der Experimentalstudie untersucht:

- E1) Finden sich Unterschiede zwischen Patienten mit akuter oder remittierter, unipolarer oder bipolarer affektiver Störung und gesunden Kontrollen bezüglich des Social Functioning gemessen mit dem IIP oder dem F-SozU?
- $H_0 : IIP_{\text{KG-UR}} = IIP_{\text{UR}} = IIP_{\text{UA}} = IIP_{\text{BR}} = IIP_{\text{KG-BR}}$
  - $H_0 : F\text{-SozU}_{\text{F-SozU}} = F\text{-SozU}_{\text{UR}} = F\text{-SozU}_{\text{UA}} = F\text{-SozU}_{\text{BR}} = F\text{-SozU}_{\text{KG-BR}}$
- E2) Bestehen zwischen Patienten mit akuter oder remittierter unipolarer oder bipolarer affektiver Störung und gesunden Kontrollen Unterschiede im Ausmaß der Empathie gemessen mit dem EQ-S und EQ-F?
- $H_0 : EQ_{\text{KG-UR}} = EQ_{\text{UR}} = EQ_{\text{UA}} = EQ_{\text{BR}} = EQ_{\text{KG-BR}}$
- E3) Besteht eine Diskrepanz zwischen der Selbst- und Fremdeinschätzung der Empathie? Wenn ja, ist diese Diskrepanz abhängig von der Gruppenzugehörigkeit, also von dem Vorliegen oder Fehlen einer akuten oder remittierten affektiven Störung?
- $H_0 : EQ\text{-S} = EQ\text{-F}$
  - $H_0 : EQ\text{-Diff}_{\text{KG-UR}} = EQ\text{-Diff}_{\text{UR}} = EQ\text{-Diff}_{\text{UA}} = EQ\text{-Diff}_{\text{BR}} = EQ\text{-Diff}_{\text{KG-BR}}$
- E4) Gibt es Gruppenunterschiede bezüglich der behavioralen Messung der Empathie anhand der CAM (Emotionserkennung) sowie des MASC (Reasoning)? Besteht ein Zusammenhang zwischen der Messung der Empathie anhand des EQ sowie der behavioralen Messung?
- $H_0 : CAM_{\text{KG-UR}} = CAM_{\text{UR}} = CAM_{\text{UA}} = CAM_{\text{BR}} = CAM_{\text{KG-BR}}$
  - $H_0 : MASC_{\text{KG-UR}} = MASC_{\text{UR}} = MASC_{\text{UA}} = MASC_{\text{BR}} = MASC_{\text{KG-BR}}$
  - $H_0 : EQ \not\sim CAM$
  - $H_0 : EQ \not\sim MASC$
- E5) Sind negative Emotionen induzierbar? Beeinflusst die Emotionsinduktion die behaviorale Messung der Empathie? Gibt es Gruppenunterschiede bezüglich der Emotionsinduktion? Ist die Induzierbarkeit von Emotionen abhängig von der Selbst- oder Fremdbeurteilung der Empathie oder vom Ausmaß der Diskrepanz zwischen beidem?
- $H_0 : Stimmung_{\text{PräMI}} = Stimmung_{\text{PostMI}} \text{ und } Stimmung_{\text{MI}} = Stimmung_{\text{NMI}}$
  - $H_0 : CAM_{\text{MI}} = CAM_{\text{NMI}} \text{ und } MASC_{\text{MI}} = MASC_{\text{NMI}}$
  - $H_0 : Ind_{\text{KG-UR}} = Ind_{\text{UR}} = Ind_{\text{UA}} = Ind_{\text{BR}} = Ind_{\text{KG-BR}}$
  - $H_0 : Ind \not\sim EQ \text{ und } Ind \not\sim EQ\text{-Diff}$
- E6) Bestehen Zusammenhänge zwischen der Fragebogenmessung sowie der behavioralen Messung der Empathie, dem Social Functioning und der prämorbidem Intelligenz?
- $H_0 : \rho_{EQ\text{-IQ}} = 0 \text{ in allen Teilnehmergruppen}$
  - $H_0 : \rho_{EQ\text{-IIP}} = \rho_{EQ\text{-F-SozU}} = 0 \text{ in allen Teilnehmergruppen}$



# Kapitel 2

## Methoden der Fragebogenvalidierung

Folgend werden die Methoden der Fragebogenvalidierung beschrieben. Zunächst wird die Übersetzung sowie Umformulierung des Fragebogens dargestellt, im Anschluss sind Details zur Stichprobe, Ablauf und verwendeten Messverfahren zu finden.

### 2.1 Übersetzung und Umformulierung des Fragebogens

Zur Validierung einer deutschen Version des EQ wurde zunächst eine eigene Übersetzung der englischen Originalversion von Baron-Cohen & Wheelwright (2004) erstellt. Diese wurde anschließend mit der deutschen Version von Joern de Haen (Autism Research Center) verglichen und zeigte einen sehr hohen Übereinstimmungsgrad. Aufgrund der hohen Übereinstimmung konnten die beiden Versionen problemlos zusammengeführt werden. Die zusammengeführte Version wurde schließlich von einer englischen Muttersprachlerin zurück übersetzt. Die Rückübersetzung und das englische Original wurden auf die inhaltliche Übereinstimmung untersucht. Diese erwies sich als sehr hoch. Schließlich wurde die deutsche Fragebogenversion von fünf Mitarbeiterinnen der Arbeitsgruppe auf Verständlichkeit der Fragen und Durchführbarkeit getestet. Auf eine weitergehende Pilot-Studie konnte verzichtet werden, da das englische Original schon auf Boden- und Deckeneffekte untersucht wurde. Vor den eigentlichen 60 Items des EQ werden 4 Beispielitems präsentiert, um den Probanden das Ausfüllen erklären zu können. Die endgültige deutsche Version enthält, ebenso wie das englische Original, 40 Items zur Empathie sowie 20 Filler-Items. Die Fragen zur Empathie bestehen aus 21 positiv gepolten Items (1, 6, 19, 22, 25, 26, 35, 36, 37, 38, 41, 42, 43, 44, 52, 54, 55, 57, 58, 59, 60) und 19 negativ gepolten Items (4, 8, 10, 11, 12, 14, 15, 18, 21, 27, 28, 29, 32, 34, 39, 46, 48, 49, 50). Filler-Items sind Fragen ohne Zusammenhang zur eigentlichen Fragestellung, also ohne Bezug zu Empathie (2, 3, 5, 7, 9, 13, 16, 17, 20, 23, 24, 30, 31, 33, 40, 45, 47, 51, 53, 56). Alle Fragen wurden von den Teilnehmern auf einer vierstufigen Skala (3 = „stimme voll und ganz zu“, 2 = „stimme eher zu“, 1 = „stimme eher nicht zu“, 0 = „stimme überhaupt nicht zu“) beantwortet. Die vollständige Version inklusive der 40 Empathie-Fragen sowie 20 Filler-Items findet sich im Anhang (A.1).

Die Auswertung erfolgt analog zur Auswertung des englischen Originals: Die Scores der Filler-Items gehen nicht in die Gesamtsumme ein (= 0). Für die positiv gepolten Items werden die Antworten „stimme voll und ganz zu“ mit 2 und „stimme eher zu“ mit 1 berechnet, „stimme eher nicht zu“ und „stimme überhaupt nicht zu“ werden nicht berechnet (= 0). Die negativ gepolten Items müssen umgepolt werden so dass hier „stimme voll und ganz zu“ und „stimme eher zu“ nicht berechnet werden (= 0), während „stimme eher nicht zu“ mit 1 und „stimme überhaupt nicht zu“ mit 2 in die Gesamtsumme eingehen. So ergibt sich eine mögliche Gesamtsumme im Bereich von 0 bis 80 Punkten. Je höher hierbei die Gesamtsumme, desto höher schätzt der Ausfüllende die empathischen Fähigkeiten ein.

Ausgehend von dieser Selbstbeurteilungsvariante (EQ-S) wurden alle Items für die Fremdbeurteilung durch eine andere Person umformuliert (EQ-F). Aus dem Item „Ich erkenne leicht, ob jemand an einem Gespräch interessiert ist.“ wurde beispielsweise „Ich habe den Eindruck, Tim erkennt leicht, ob jemand an einem Gespräch interessiert ist.“. Das Item „Andere sagen häufig, ich sei unsensibel, auch wenn ich nicht immer verstehe warum.“ wurde in „Ich könnte mir vorstellen, dass Anne häufig gesagt bekommt, dass sie unsensibel ist, auch wenn sie nicht immer versteht warum“ umformuliert. Die Fremdbeurteilungsversion wurde von denselben fünf Mitarbeiterinnen auf Verständlichkeit der Fragen und Durchführbarkeit überprüft, wie die Übersetzung. Um die Beurteilung durch eine andere Person zu erleichtern, wurde eine Version erstellt in der vor dem Ausdrucken des Dokuments der Vorname der zu beurteilenden Person eingetragen werden kann. So wird die ausfüllende Person in jeder Frage daran erinnert, wen sie beurteilen soll. Zudem gibt es getrennte Versionen für männliche und weibliche Teilnehmer, damit alle Pronomina jeweils zur zu beurteilenden Person passen. Für die Anwendungsbereiche des Fragebogens in denen ein gezieltes Ausdrucken für jeden Probanden nicht möglich ist – zum Beispiel im Rahmen der Validierungsstudie – wurden ebenfalls je eine Version für männliche und weibliche Personen erstellt, jedoch ohne Felder zum Eintragen des Namens. Auch die vollständige Fremdbeurteilungsversion findet sich im Anhang (A.2).

## 2.2 Stichprobenbeschreibung

Im Rahmen einer größeren Untersuchung mit mehreren unabhängigen Fragestellungen wurden 140 Teilnehmer von studentischen Mitarbeitern rekrutiert. Die Rekrutierung fand hauptsächlich im persönlichen Umfeld der Mitarbeiterinnen statt. Unter den Probanden sind 106 Frauen (75.7%). Das Geschlechterverhältnis beträgt somit 3.12 : 1 und ist damit in einem ähnlichen Bereich wie im Artikel von Baron-Cohen & Wheelwright (2004) welches bei 2.6 : 1 liegt. Das durchschnittliche Alter beträgt 22.8 Jahre ( $SD = 4.17$ ,  $Min = 18$ ,  $Max = 42$ ). Die Stichprobe besteht zum Großteil aus Studenten ( $n = 125$ , 89.3%) mit regelmäßigen Sozialkontakten (mindestens ein Mal pro Woche,  $n = 137$ , 97.9%). Eine detaillierte soziodemographische Beschreibung der Teilnehmer findet sich in Tabelle 2.1. Alle Daten in der Stichprobenbeschreibung beziehen sich auf den vollständigen Datensatz mit allen 140 Teilnehmern.

Tabelle 2.1: Soziodemographische Beschreibung der Teilnehmer

<i>Teil 1</i>	Anzahl	Anteil
<b>höchster Schulabschluss</b>		
Sonderschulabschluss	1	0.7%
mittlere Reife	5	3.6%
Abitur / Fachhochschulreife	134	95.7%
<b>höchster erreichter Berufsabschluss</b>		
noch in Berufsausbildung	102	72.9%
Lehre / Fachschule	17	12.1%
Meister	2	1.4%
Fachhochschule / Universität	18	12.9%
ohne Abschluss	–	–
<b>aktuelle Lebenssituation</b>		
in einer Wohngemeinschaft lebend	79	56.4%
alleine lebend	20	14.3%
mit Ehe-/Lebenspartner lebend	23	16.4%
mit Eltern/Verwandten lebend	15	10.7%
Sonstige	3	2.1%
<b>Sozialkontakte</b>		
mehr als einmal die Woche	127	90.7%
höchstens einmal pro Woche	10	7.1%
einmal in zwei Wochen	2	1.4%
einmal im Monat	1	0.7%

### 2.2.1 Ein- und Ausschlusskriterien

Die folgenden Ausschlusskriterien gelten für alle Teilnehmer der Validierungsstudie:

- Unzureichende Kenntnis der deutschen Sprache
- Diagnose einer Erkrankung des ZNS (z.B. Morbus Parkinson)
- Alter unter 18 oder über 69 Jahren
- Diagnose einer psychischen Erkrankung aktuell oder anamnestisch vorliegend

## 2.2.2 Neuropsychologie deskriptiv

Die mit dem MWT-B erhobenen Daten ermöglichen eine Schätzung der allgemeinen Intelligenz. Von allen 140 Teilnehmern liegen hierfür gültige Werte vor. Für die Gesamtstichprobe beträgt der mittlere Rohwert 29.3 ( $SD = 2.97$ ,  $Min = 17$ ,  $Max = 36$ ). Daraus ergibt sich ein mittlerer geschätzter IQ von 107.7 ( $SD = 9.72$ ,  $Min = 85$ ,  $Max = 143$ ). 1.4% ( $n = 2$ ) der Teilnehmer weisen eine niedrige Intelligenz auf, 60% ( $n = 84$ ) eine mittlere, 35% ( $n = 49$ ) eine hohe und 3.6% ( $n = 5$ ) eine sehr hohe Intelligenz. Dies sind bei einer hauptsächlich studentischen Stichprobe zu erwartende Werte.

## 2.3 Messinstrumente und Methoden

### 2.3.1 Neuropsychologie

Die neuropsychologische Testung im Rahmen der Fragebogenvalidierung enthielt einzig den Mehrfachwahl-Wortschatz-Intelligenztest Version B (MWT-B; Lehrl, 1999). Details zu diesem Verfahren sind in Kapitel 4.2.2 zu finden.

### 2.3.2 Empathie und Social Functioning

Die folgenden Messinstrumente werden sowohl in der Validierungsstudie als auch in der experimentellen Untersuchung verwendet: Das Inventar zur Erfassung interpersonaler Probleme (IIP-Cc, Kurzversion mit 24 Items; Horowitz et al., 2000), die Face-Task der Cambridge Mindreading Face-Voice Battery (CAM; Golan et al., 2006) sowie der EQ-S und -F. Zur Beschreibung des IIP-C und der CAM sei auf Kapitel 4.2.3 verwiesen. Die Beschreibung des EQ findet sich in Kapitel 2.1.

#### **Empathizing-Systemizing Quotient (EQ-SQ)**

Der Empathizing-Systemizing Quotient (EQ-SQ; Samson & Huber, 2010) ist ein Fragebogen zur Erfassung der Empathiefähigkeit und der Systematisierungstendenz. Diese beiden Dimensionen werden mit jeweils 13 Fragen erfasst, so dass sich eine Gesamtsumme von 26 Items ergibt. „Ich erkenne leicht, ob jemand ein Gespräch anfangen möchte.“ ist ein Beispiel für ein Empathie-Item, während „Wenn es in meiner Wohnung ein Problem mit der Elektrik gibt, kann ich es selbst beheben.“ eines der Systematisierungs-Items ist. Für beide Dimensionen existieren sowohl positiv als auch negativ gepolte Items und alle Items werden von den Probanden auf einer vierstufigen Likert-Skala von „Stimme überhaupt nicht zu“ bis „Stimme voll zu“ eingeschätzt. Die Ergebnisse beider Dimensionen werden getrennt aufsummiert: Bei positiv gepolten Items werden für „Stimme eher zu“ 1 Punkt und für „Stimme voll zu“ 2 Punkte vergeben. Bei negativ gepolten Items erhalten Probanden 2 Punkte für „Stimme überhaupt nicht zu“ und einen Punkt für „Stimme eher nicht zu“. Alle übrigen Antworten gehen mit 0 Punkten in die jeweiligen Skalensummen ein. Es ergeben sich daher für beide Dimensionen Gesamtsummen von 0 bis 26. Die Daten von Samson & Huber (2010) sprechen dafür dass eine hohe interne Konsistenz der beiden Dimensionen sowie eine hohe Retest-Reliabilität gegeben sind.

### **Reading the Mind in the Eyes Test (RME)**

Der Reading the Mind in the Eyes Test (RME; Baron-Cohen et al., 2001) ist ein Computertest zur Messung der Emotionserkennungsfähigkeit. Die Probanden sehen 36 Schwarzweißfotografien mit Augenpaaren von Schauspielern die positive, negative sowie neutrale Emotionen zeigen. Es handelt sich um Schauspieler beiderlei Geschlechts und aller Altersklassen. Für jedes Augenpaar werden vier Auswahlmöglichkeiten dargeboten und die Probanden werden gebeten, sich für die am besten passende Emotion zu entscheiden. Es ergibt sich eine Gesamtsumme richtiger Antworten im Bereich von 0 bis 36.

### **Impact Message Inventory (IMI)**

Mit dem Fremdbeurteilungsfragebogen Impact Message Inventory (IMI; Kiesler & Schmidt, 1993) wird die zwischenmenschlichen Interaktion und Persönlichkeit des Probanden erfasst. Es ergeben sich Werte für die folgenden acht Faktoren: Freundlich, freundlich-submissiv, submissiv, feindselig-submissiv, feindselig, feindselig-dominant, dominant und freundlich-dominant. Damit kann die zwischenmenschliche Interaktion des Probanden im Rahmen eines Circumplexmodells eingeordnet werden.

## **2.3.3 Weitere Messinstrumente**

### **Liebowitz Social Anxiety Scale (LSAS)**

Die Liebowitz Social Anxiety Scale (LSAS; Liebowitz, 1987) ist ein Fragebogen mit 24 Items zur Einschätzung sozialer Ängstlichkeit. Verschiedene soziale Situationen, zum Beispiel „Vor Publikum auftreten“, werden von den Probanden bezüglich der Angstbesetztheit (keine, gering, mäßig, stark) und des Vermeidungsverhaltens (nie, selten, häufig, fast immer) auf einer vierstufigen Likert-Skala eingeschätzt. Daraus ergeben sich Gesamtwerte für die soziale Ängstlichkeit sowie für die Ausprägtheit des Vermeidungsverhaltens.

### **Social Interaction Anxiety Scale und Social Phobia Scale (SIAS-SPS)**

Die Social Interaction Anxiety Scale und Social Phobia Scale (SIAS-SPS; Stangier & Heidenreich, 1995) ist ein Fragebogen mit 40 Items, wie zum Beispiel „Ich fühle mich angespannt, wenn ich mit einer Person alleine zusammen bin.“, zur Erfassung sozialer Ängstlichkeit. Die Probanden werden gebeten, die verschiedenen positiv sowie negativ gepolten Aussagen dahingehend zu beurteilen, ob die Aussage auf sie zutrifft. Zur Beurteilung wird eine fünfstufige Likert-Skala von „überhaupt nicht zutreffend“ bis „sehr stark zutreffend“ genutzt. Es ergibt sich ein Gesamtwert für die Ausprägung der sozialen Angst des Probanden.

## **2.4 Untersuchungsablauf**

Sowohl die Selbst- als auch die Fremdbeurteilungsvariante des deutschen EQ wurden im Rahmen einer größeren Studie mit einer gesunden Stichprobe erstmalig einge-

setzt und validiert.

Die Erhebung fand in den Räumen des Fachbereich Psychologie der Eberhard Karls Universität Tübingen statt und dauerte ca. 120 Minuten. Die Probanden erhielten zunächst einige demographische Fragen. Darauf folgte der MWT-B zur Schätzung der allgemeinen Intelligenz. Anschließend füllten sie den EQ-SQ, das IIP-Cc und den EQ-S als Papierfragebögen aus. Daraufhin wurde der RME am PC durchgeführt. Es folgte der Aggressionsfragebogen nach Buss und Perry (Herzberg, 2003) in der Papierversion, der jedoch für die vorliegende Untersuchung nicht weiter beachtet wird. Anschließend bearbeiteten die Probanden die CAM am PC. Danach folgte ebenfalls am PC eine Aufgabe zur Arithmetic Inhibition aus dem Paced Auditory Serial Addition Test (Gronwall, 1977), der für die vorliegende Studie keine weitere Beachtung findet. Daraufhin wurden die LSAS und die SIAS-SPS zur Erfassung der sozialen Ängstlichkeit durchgeführt. Der dann folgende Cognitive Emotion Regulation Questionnaire (Loch et al., 2011) sowie eine weitere Aufgabe am PC (Babyfaces; Kinney, Torney & Wolkenstein) werden für die vorliegende Untersuchung nicht weiter beachtet.

Nach diesem Ablauf im Labor erhielten die Teilnehmer die Fremdbeurteilungsbögen EQ-F und das IMI sowie ein Anschreiben und einen frankierten Rückumschlag mit der Bitte, diese an eine Person weiterzugeben, die die Fremdbeurteilung übernehmen kann. Um eine unbeeinflusste Fremdbeurteilung zu ermöglichen, wurde auf die Rücksendung der Unterlagen per Post bestanden.

## 2.5 Studiendesign

Bei der vorliegenden Untersuchung handelt es sich um eine korrelative Studie im Querschnittsdesign zur Validierung eines Fragebogens.

## 2.6 Statistische Analysen

Zu den in der vorliegenden Arbeit verwendeten Auswertungssoftwares sowie grundsätzlichen Herangehensweisen in der Datenanalyse sei auf Kapitel 4.6 verwiesen. Die Vorgehensweise zur Klärung der einzelnen Fragestellungen ist im Folgenden dargestellt:

### 2.6.1 Fragestellung V1

Zur deskriptiven Beschreibung der Ergebnisse des EQ werden sowohl Mittelwerte und Standardabweichungen als auch Minima und Maxima der Summen beider Fragebogenvarianten für die gesamte Stichprobe und für männliche und weibliche Teilnehmer getrennt dargestellt. Um einen Vergleich mit dem englischen Original zu ermöglichen, werden die entsprechenden Ergebnisse denen von Baron-Cohen & Wheelwright (2004) und von Lawrence et al. (2004) gegenüber gestellt.

### 2.6.2 Fragestellung V2

Fragestellung V2 befasst sich mit den Gütekriterien des EQ. Es werden sowohl Durchführungs- als auch Auswertungsobjektivität und Interpretationsobjektivität diskutiert. Weiterhin wird die interne Konsistenz und Homogenität anhand des Reliabilitätskoeffizienten Cronbachs  $\alpha$  bestimmt. Zudem werden Trennschärfen und Itemschwierigkeiten beachtet.

### 2.6.3 Fragestellung V3

Die Untersuchung der Faktorenstruktur des EQ erfolgt anhand einer exploratorischen Faktorenanalyse mit Hauptkomponentenmethode, Varimax Rotation und Scree Plots. Hierbei werden Werte kleiner als .3 unterdrückt. Es wird weiterhin überprüft ob alle Items zumindest mit einem anderen ab .2 korrelieren und ob alle Items ausreichend hohe Gemeinsamkeitsgrade (Kommunalitäten) im finalen Modell haben. Items für die dies nicht als gegeben angenommen werden kann, werden ausgeschlossen. In der Validierung des englischen Originals wurden zudem Items ausgeschlossen, die signifikant mit einem Maß der sozialen Erwünschtheit korrelieren. In der vorliegenden Untersuchung fand keine Messung der sozialen Erwünschtheit statt, weshalb dieser Schritt in der Untersuchung der Faktorenstruktur nicht stattfinden kann.

### 2.6.4 Fragestellung V4

Mögliche Geschlechtsunterschiede bezüglich der Gesamtsummen des EQ in der regulären Auswertung werden mittels t-Tests beziehungsweise der nonparametrischen Alternative Mann-Whitney-U-Test untersucht. Zur Klärung der Frage ob Geschlechtsunterschiede bezüglich der einzelnen Unterskalen bestehen, wird eine MANOVA mit der unabhängigen Variable Geschlecht und den Unterskalen des EQ-S und -F als abhängigen Variablen berechnet.

### 2.6.5 Fragestellung V5

Die Konstruktvalidität setzt sich aus der konvergenten Validität und der diskriminanten Validität zusammen. Die konvergente Validität des EQ wird betrachtet, indem die Korrelationen des EQ mit anderen Maßen der Empathie berechnet werden. Die verwendeten Maße der Empathie sind die Empathizing-Unterskala des EQ-SQ, der RME sowie die CAM. Für die diskriminante Validität rücken Korrelationen des EQ mit Maßen anderer Konstrukte in den Vordergrund. Hierfür werden die Systemizing-Unterskala des EQ-SQ sowie die mit dem MWT-B geschätzte prämorbid Intelligenz herangezogen.

### 2.6.6 Fragestellung V6

Fragestellung V6 ermöglicht den explorativen Blick auf mögliche Einflussfaktoren auf den EQ im Rahmen jeweils einer linearen multiplen Regressionsanalyse für jede Fragebogenvariante. Anhand der regressionsanalytischen Betrachtung soll die Gesamtsumme des EQ durch Alter, Geschlecht, Sozialkontakte, Schulabschluss, Be-

rufsabschluss, Psychologiestudium, MWT-B-Summe und RME-Summe vorhergesagt werden.



# Kapitel 3

## Ergebnisse der Fragebogenvalidierung

Im Rahmen der Fragebogenvalidierung sollen die folgenden Fragen untersucht werden:

- V1) Wie lassen sich die Ergebnisse des EQ-S und -F deskriptiv beschreiben? Sind die Rohwerte vergleichbar mit den vom englischen Original bekannten Werten?
- V2) Wie fallen die klassischen psychometrischen Gütekriterien für die beiden Versionen des deutschen EQ aus?
- V3) Welche Faktorstruktur weisen EQ-S und -F auf? Sind die Faktorstrukturen des EQ-S und -F vergleichbar mit der des englischen Originals?
- V4) Sind die aus Untersuchungen mit dem englischen Original bekannten Geschlechtsunterschiede auch mit dem EQ-S und -F messbar?
- V5) Sind EQ-S und -F konstruktvalide, messen sie also tatsächlich Empathie?
  - a) Konvergente Validität: Messen EQ-S und -F dasselbe Konstrukt wie andere Maße der Empathie (RME, CAM und die Empathizing-Unterskala des EQ-SQ)?
  - b) Diskriminante Validität: Messen EQ-S und -F andere Konstrukte als Maße der Intelligenz (MWT-B) und des Systemizing (Unterskala des EQ-SQ)?
- V6) Welche Einflussfaktoren zeigen sich in einer explorativen regressionsanalytischen Betrachtung des EQ?

### 3.1 *Fragestellung V1: Deskriptive Beschreibung der Ergebnisse*

In Tabelle 3.1 sind die Mittelwerte und Standardabweichungen der EQ-S und EQ-F Gesamtscores sowohl für weibliche und männliche Teilnehmer getrennt als auch für

die Gesamtstichprobe dargestellt. Die Mittelwerte der EQ-S Scores ähneln denen der englischen Originalartikel (Baron-Cohen & Wheelwright, 2004; Lawrence et al., 2004): Die männlichen Teilnehmer der vorliegenden Studie erreichen einen Mittelwert von 39.0 ( $SD = 10.9$ ), diejenigen bei Baron-Cohen & Wheelwright von 41.8 ( $SD = 11.2$ ) und diejenigen bei Lawrence et al. von 41.3 ( $SD = 10.1$ ). Die weiblichen Teilnehmer der vorliegenden Arbeit haben einen mittleren EQ-S Score von 46.43 ( $SD = 10.6$ ), bei Baron-Cohen & Wheelwright von 47.2 ( $SD = 10.2$ ) und bei Lawrence et al. von 50.6 ( $SD = 9.2$ ).

Tabelle 3.1: Mittelwerte und Standardabweichungen der EQ-S und EQ-F Scores

	Anzahl	Mittelwert	$SD$	$Min$	$Max$
<b>EQ-S</b>					
männlich	34	39.0	10.9	19	68
weiblich	106	46.4	10.6	19	76
gesamt	140	44.6	11.1	19	76
<b>EQ-F</b>					
männlich	31	38.6	13.3	11	63
weiblich	97	46.7	12.2	13	69
gesamt	128	44.7	12.9	11	69

Insgesamt ähneln die vorliegenden Werte des EQ-S somit denen der englischen Originalartikel in der deskriptiven Beschreibung. Die Mittelwerte des EQ-F fallen in der deskriptiven Betrachtung ähnlich aus wie die des EQ-S, weisen jedoch leicht größere Standardabweichungen und tiefere Minima sowie Maxima auf.

## 3.2 Fragestellung V2: Gütekriterien des EQ

Ein Testverfahren verfügt über Durchführungsobjektivität wenn die Ergebnisse nicht vom jeweiligen Testleiter abhängig sind. Dies kann sowohl bei EQ-S als auch bei EQ-F als gegeben angesehen werden, da die Fragebögen von den Probanden selbst, beziehungsweise von einer vertrauten Person ausgefüllt werden. Auch sowohl Auswertungsobjektivität als auch Interpretationsobjektivität sind aufgrund fest vorgegebener Regeln zur Summenbildung und Interpretation als gegeben anzusehen.

Im Folgenden werden die 20 Filler-Items gemäß der Auswertungsregeln nicht weiter beachtet, es verbleiben also die 40 positiven und negativen Items zur Empathie. Das Umpolen der negativen Items sowie die Berechnung der Summen erfolgt wie in den Auswertungsregeln beschrieben (siehe Seite 30).

Um die interne Konsistenz oder Homogenität der Fragebogen-Versionen zu bestimmen, wird der Reliabilitätskoeffizient Cronbachs  $\alpha$  berechnet. Mit einem  $\alpha$  von .88 für den EQ-S und von 0.91 für den EQ-F weisen sowohl die Selbstbeurteilungsvariante als auch die Fremdbeurteilungsvariante eine gute interne Konsistenz auf.

Bei der Betrachtung der Trennschärfen zeigt sich, dass neun Items des EQ-S Trennschärfen unter 0.2 beziehungsweise einen negativen Wert aufweisen und daher

aus der Auswertung ausgeschlossen werden sollten. Es handelt sich um die folgenden Items (Trennschärfe in Klammern): EQ-S10 (.075), EQ-S11 (.073), EQ-S18 (.013), EQ-S21 (.110), EQ-S28 (.024), EQ-S37 (.192), EQ-S49 (.176), EQ-S57 (.172) und EQ-S60 (.095). Vier Items des EQ-F müssen aus dem selben Grund ausgeschlossen werden: EQ-F18 (.070), EQ-F28 (.120), EQ-F37 (.027) und EQ-F39 (-.003). Im nächsten Schritt werden Items aus der Auswertung ausgeschlossen deren Itemschwierigkeiten außerhalb des Bereichs von .2 und .8 liegen, deren Schwierigkeit also zu niedrig oder zu hoch ist. Im EQ-S findet sich nur ein Item mit einer zu hohen Itemschwierigkeit: EQ-S12 (.86). Alle Items des EQ-F liegen im gewünschten Bereich und es muss kein weiteres Item ausgeschlossen werden. Nach Ausschluss der genannten Items erhöht sich Cronbachs  $\alpha$  für den EQ-S auf .91 und bleibt für den EQ-F bei .91.

Zusammenfassend lässt sich feststellen, dass EQ-S und -F eine gute Durchführungs-, Auswertungs- und Interpretationsobjektivität aufweisen. Beide Fragebogenvarianten verfügen über eine als gut anzusehende interne Konsistenz.

### 3.3 *Fragestellung V3: Faktorenstruktur des EQ*

Die folgende Auswertung orientiert sich an der Validierung des englischen Originals des EQ (Lawrence et al., 2004). Zur Untersuchung der Faktorenstruktur der beiden deutschen Fragebogenversionen wird jeweils eine exploratorische Faktorenanalyse mit Hauptkomponentenmethode und Varimax Rotation sowie Screeplots durchgeführt (Werte kleiner als .3 werden unterdrückt). Es wird überprüft ob alle Items zumindest mit einem anderen ab .2 korrelieren und ob alle Items ausreichend hohe Gemeinsamkeitsgrade (Kommunalitäten) im finalen Modell haben. Dies ist für alle Items der Selbst- sowie der Fremdbeurteilungsversion gegeben, es müssen also keine weiteren Items ausgeschlossen werden.

#### 3.3.1 EQ-S

Für die deutsche Version des EQ-S liegen die Kommunalitäten zwischen 0.532 und 0.797. Es liegen Datensätze von 140 Probanden für 30 verbleibende Items vor, was annähernd der Regel entspricht, dass fünf Fälle pro Item vorhanden sein sollten. Das Kaiser-Meyer-Okin-Kriterium liegt bei 0.85 und der Bartlett-Test auf Sphärizität ist hoch signifikant, eine Faktorenanalyse kann also gerechnet werden. Screeplot 3.1 und Eigenwerte sprechen für drei Faktoren, die 44.2% der Gesamtvarianz erklären. Elf der Items laden auf keinen der Faktoren (EQ-S25, EQ-S58, EQ-S15, EQ-S8, EQ-S35, EQ-S29, EQ-S46, EQ-S34, EQ-S4, EQ-S14, EQ-S38), es bleiben 19 Items die mindestens mit .3 auf mindestens einen der drei Faktoren laden. Die finalen Ladungen der Items aus der Hauptkomponentenanalyse finden sich in Tabelle 3.2. Alle Items die im in der vorliegenden Untersuchung ermittelten Faktor 1 enthalten sind, sind auch im englischen Original in Faktor 1 enthalten. Die Benennung erfolgt daher analog zu Lawrence et al. in „kognitive Empathie“. Zwei der Items die in der vorliegenden Untersuchung auf Faktor 2 laden (EQ-S36, EQ-S52) sind im

englischen Original im Faktor 1 enthalten. Davon abgesehen entspricht der hier ermittelte Faktor 2 dem Faktor 2 von Lawrence et al. und wird analog zur Benennung des englischen Originals in „emotionale Reaktivität“ benannt. Der hier gefundene Faktor 3 zeigt keine Übereinstimmung mit dem Faktor 3 „soziale Fähigkeiten“ von Lawrence et al.. Anhand der inhaltlichen Betrachtung der enthaltenen Items wird Faktor 3 als „empathische Betroffenheit“ bezeichnet. In Tabelle 3.3 finden sich die Items des EQ-S in den drei Faktoren.

Tabelle 3.2: Ladungen der Hauptkomponentenanalyse des EQ-S

	1	2	3
EQ-S41	0.783		
EQ-S44	0.751		
EQ-S26	0.737		
EQ-S1	0.674		
EQ-S19	0.668		
EQ-S55	0.559		
EQ-S54	0.464	0.409	
EQ-S43		0.753	
EQ-S6		0.658	
EQ-S59		0.578	0.458
EQ-S36		0.478	
EQ-S52	0.356	0.473	
EQ-S22		0.376	
EQ-S48		0.317	
EQ-S32			0.784
EQ-S27			0.708
EQ-S42			0.577
EQ-S39			0.538
EQ-S50		0.407	0.457

Tabelle 3.3: EQ-S Items in den 3 Faktoren

---

<b>kognitive Empathie</b>	
EQ-S41	Ich erkenne leicht, ob jemand sich für das, was ich sage, interessiert, oder davon gelangweilt ist.
EQ-S44	Ich bemerke, wenn ich störe, auch wenn die andere Person es nicht sagt.
EQ-S26	Ich bemerke es schnell, wenn sich jemand in einer Gruppe unwohl oder unbehaglich fühlt.
EQ-S1	Ich erkenne leicht, ob jemand an einem Gespräch interessiert ist.
EQ-S19	Ich merke leicht, wenn jemand etwas anderes sagt, als er meint
EQ-S55	Ich bemerke, ob jemand seine wahren Gefühle verbirgt
EQ-S54	Ich finde leicht heraus, worüber jemand anderes gerne sprechen würde
<hr/>	
<b>emotionale Reaktivität</b>	
EQ-S43	Freunde erzählen mir von ihren Problemen, weil sie sagen, ich sei verständnisvoll
EQ-S6	Es macht mir große Freude, mich um andere Leute zu kümmern
EQ-S59	Ich neige dazu, von Problemen meiner Freunde emotional betroffen zu sein
EQ-S36	Andere sagen, ich verstehe gut was sie fühlen und denken
EQ-S52	Ich kann mich schnell und intuitiv darauf einstellen, wie es jemandem geht
EQ-S22	Ich finde es leicht, mich in jemanden hineinzusetzen
EQ-S48	Andere sagen häufig, ich sei unsensibel, auch wenn ich nicht verstehe warum
<hr/>	
<b>empathische Betroffenheit</b>	
EQ-S32	Leute weinen zu sehen macht mir nichts aus
EQ-S27	Wenn ich etwas sage, dass jemanden verletzt, dann ist das dessen Problem
EQ-S42	Es erschüttert mich, wenn ich in Nachrichtensendungen Leute leiden sehe
EQ-S39	Ich kann Entscheidungen treffen, ohne mich von den Gefühlen anderer beeinflussen zu lassen
EQ-S50	Wenn ich einen Film sehe, bleibe ich in der Regel emotional unberührt

---

### 3.3.2 EQ-F

Die Kommunalitäten der Items des EQ-F liegen zwischen 0.487 und 0.829. Für 36 Items sind 128 Fälle vorhanden, die fünf-Fälle-pro-Item-Regel wird damit nicht erfüllt. Mit einem Kaiser-Meyer-Olkin-Kriterium von 0.84 und einem hoch signifikanten Bartlett-Test sind die Voraussetzungen für eine Hauptkomponentenanalyse gegeben. Anhand Screeplot (siehe Abbildung 3.2) und Eigenwerten werden vier Fak-

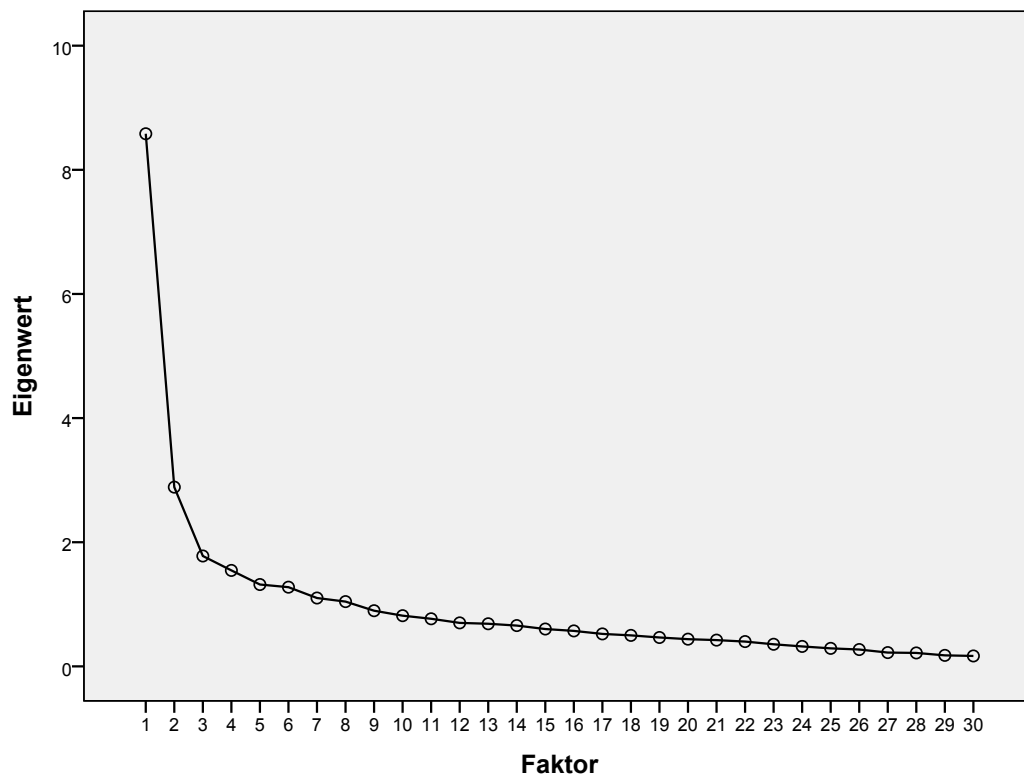


Abbildung 3.1: Screeplot für den EQ-S

toren ermittelt, die 44.2% der Gesamtvarianz erklären. Neun Items laden auf keinen der Faktoren und werden ausgeschlossen (EQ-F49, EQ-F15, EQ-F22, EQ-F21, EQ-F50, EQ-F38, F60, F4, F57), es verbleiben also 27 Items. Die Ladungen der einzelnen Items auf die vier Faktoren finden sich in Tabelle 3.4. Der hier ermittelte Faktor 1 weist weder mit einem Faktor bei Lawrence et al. noch bei der Faktorenanalyse des EQ-S eine Übereinstimmung auf. Inhaltlich passend zu den enthaltenen Items wird er „soziales Geschick“ genannt. Alle im Faktor 2 enthaltenen Items sind bei Lawrence et al. sowie im EQ-S in Faktor 1 enthalten. Als Benennung wird entsprechend „kognitive Empathie“ gewählt. Der Faktor 3 des EQ-F weist beachtliche Übereinstimmungen mit Faktor 2 des englischen Originals und des EQ-S auf und wird „emotionale Reaktivität“ genannt. Geringe Übereinstimmungen liegen für den hier gefundenen Faktor 4 mit Faktor 3 des EQ-S vor. Entsprechend der inhaltlichen Betrachtung der Items wird er jedoch ebenfalls als „empathische Betroffenheit“ bezeichnet. Der genaue Wortlaut der Items der vier Faktoren findet sich in Tabelle 3.5.

Tabelle 3.4: Ladungen der Hauptkomponentenanalyse des EQ-F

	1	2	3	4
EQ-F34	0.681			
EQ-F11	0.616			
EQ-F10	0.605			
EQ-F27	0.594			
EQ-F14	0.586			
EQ-F29	0.551			
EQ-F48	0.514			
EQ-F46	0.412			
EQ-F32	0.400		0.315	
EQ-F41		0.809		
EQ-F1		0.730		
EQ-F26		0.591	0.300	
EQ-F19		0.496	0.397	
EQ-F44	0.321	0.473	0.305	
EQ-F25		0.456	0.387	0.405
EQ-F43			0.741	
EQ-F36			0.695	
EQ-F55			0.678	
EQ-F54		0.447	0.492	
EQ-F52		0.339	0.441	
EQ-F42				0.643
EQ-F59				0.606
EQ-F6				0.567
EQ-F12				0.532
EQ-F58		0.386	0.314	0.447
EQ-F35				0.411
EQ-F8				0.380



Tabelle 3.5: EQ-F Items in den vier Faktoren

---

<b>soziales Geschick</b>	
EQ-F34	Sehr direkt, was manche Leute für unhöflich halten
EQ-F11	Stört ihn nicht besonders, wenn er zu Treffen mit Freunden zu spät kommt
EQ-F10	Geht oft zu weit, um sich in Diskussionen durchzusetzen
EQ-F27	Ist der Meinung, wenn er etwas Verletzendes sagt, ist es das Problem des anderen
EQ-F14	Schwierig zu beurteilen, ob etwas höflich oder unhöflich ist
EQ-F29	Kann nicht immer verstehen, warum sich jemand durch eine Bemerkung verletzt gefühlt haben sollte
EQ-F48	Bekommt häufig gesagt, dass er unsensibel ist
EQ-F46	Geht manchmal beim Necken zu weit
EQ-F32	Leute weinen zu sehen, macht ihm nicht wirklich was aus

---

<b>kognitive Empathie</b>	
EQ-F41	Erkennt leicht, ob jemand sich für das was er sagt interessiert
EQ-F1	Erkennt leicht, ob jemand an Gespräch interessiert ist
EQ-F26	Bemerkt schnell, wenn sich jemand in einer Gruppe unwohl fühlt
EQ-F19	Bemerkt es leicht, wenn jemand etwas anderes sagt, als er meint
EQ-F44	Bemerkt, wenn er stört, auch wenn andere Person es nicht sagt
EQ-F25	Kann gut vorhersehen, wie sich jemand fühlen wird

---

<b>emotionale Reaktivität</b>	
EQ-F43	Ich erzähle ihm gewöhnlich von meinen Problemen, weil er verständnisvoll ist
EQ-F36	Ist gut darin zu verstehen, was ich fühle und denke
EQ-F55	Bemerkt, ob jemand seine wahren Gefühle verbirgt
EQ-F54	Kann leicht herausfinden, worüber jemand anderes gern sprechen würde
EQ-F52	Kann sich schnell und intuitiv darauf einstellen, wie es jemandem geht

---

<b>empathische Betroffenheit</b>	
EQ-F42	Ist erschüttert, wenn er in Nachrichtensendungen Leute leiden sieht
EQ-F59	Neigt dazu, von Problemen seiner Freunde emotional betroffen zu sein
EQ-F6	Große Freude, sich um andere zu kümmern
EQ-F12	Freundschaften und Beziehungen zu kompliziert
EQ-F58	Kann gut vorhersehen, was jemand tun wird
EQ-F35	Findet soziale Situationen nicht verwirrend
EQ-F8	Schwierigkeit zu erkennen, was er in einer sozialen Situation tun soll

---

Alles in allem scheinen die Faktorenstrukturen des EQ-S wie auch des EQ-F der des englischen Originals vor allem auf inhaltlicher Ebene zu ähneln.

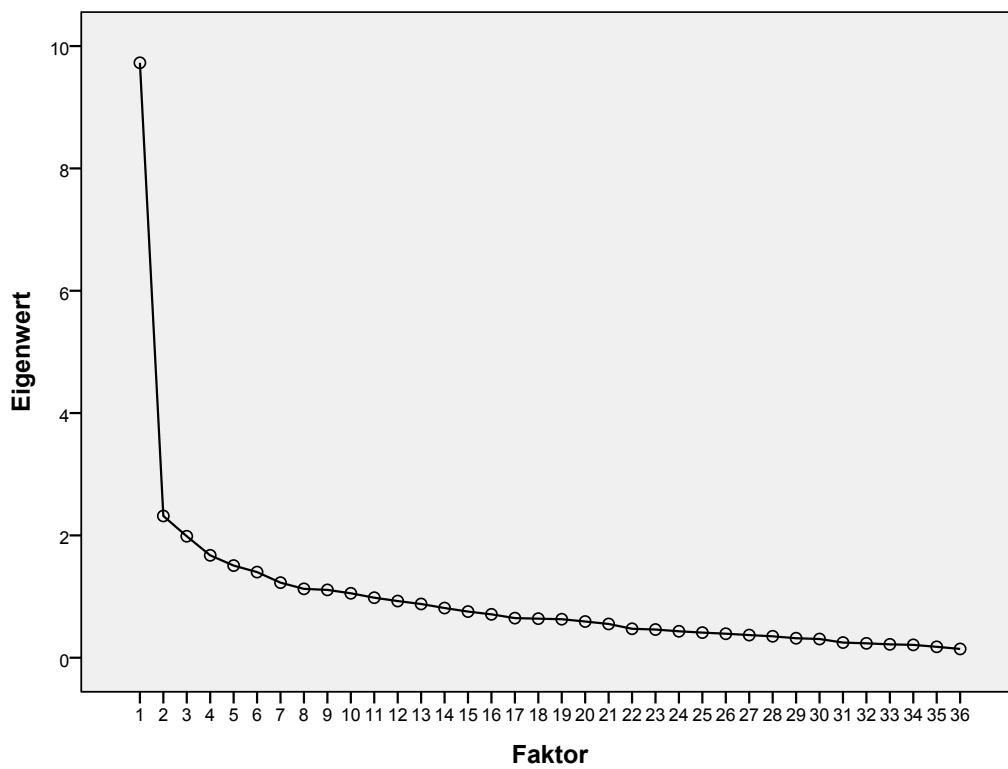


Abbildung 3.2: Screeplot für den EQ-F

### 3.3.3 Auswertungsregeln nach PCA

Nach der Faktorenanalyse der Selbst- und der Fremdbeurteilungsversion des deutschsprachigen EQ ergibt sich eine modifizierte Auswertung. Im Anhang A.1 und A.2 sind die in der Auswertung verbleibenden Items fett gedruckt. Der EQ-S enthält in der Auswertung nach der Hauptkomponentenanalyse (PCA-Auswertung) 19 Items, davon 14 positive und 5 negative. Der Gesamtscore kann zwischen 0 und 38 liegen. Zudem können Werte für die drei Faktoren ermittelt werden: Faktor 1 kognitive Empathie mit 7 Items, Faktor 2 emotionale Reaktivität mit 7 Items und Faktor 3 empathische Betroffenheit mit 5 Items. Die PCA-Auswertung des EQ-F bezieht 27 Items mit ein, davon 16 positive und 11 negative. Der Gesamtscore kann daher zwischen 0 und 54 liegen. Außerdem können Werte für die vier Faktoren ermittelt werden: Faktor 1 soziales Geschick mit 9 Items, Faktor 2 kognitive Empathie mit 6 Items, Faktor 3 emotionale Reaktivität mit 5 Items und Faktor 4 empathische Betroffenheit mit 7 Items.

Zur Untersuchung der Verteilungsform der Gesamtscores in der PCA-Auswertung sowie der einzelnen Faktoren des EQ-S und EQ-F wird der Shapiro-Wilk-Test auf Normalverteilung durchgeführt. Für die EQ-S Summe in der PCA-Auswertung kann die Normalverteilungsannahme aufrecht erhalten werden ( $W = 0.99$ ,  $df = 140$ ,  $p = .33$ ). Die Normalverteilungsannahme ist jedoch abzulehnen für die einzelnen Unterfaktoren des EQ-S (Faktor 1 kognitive Empathie  $W = 0.96$ ,  $df = 140$ ,  $p < .001$ ; Faktor 2 emotionale Reaktivität  $W = 0.97$ ,  $df = 140$ ,  $p = .006$ ; Faktor 3 empathische Betroffenheit  $W = 0.95$ ,  $df = 140$ ,  $p < .001$ ), für die EQ-F Summe in der PCA-Auswertung ( $W = 0.98$ ,  $df = 128$ ,  $p = .026$ ) sowie für die einzelnen Unterfaktoren des EQ-F (Faktor 1 soziales Geschick  $W = 0.95$ ,  $df = 128$ ,  $p < .001$ ; Faktor 2 kognitive Empathie  $W = 0.97$ ,  $df = 128$ ,  $p = .006$ ; Faktor 3 emotionale Reaktivität  $W = 0.95$ ,  $df = 128$ ,  $p < .001$ ; Faktor 4 empathische Betroffenheit  $W = 0.96$ ,  $df = 128$ ,  $p = .001$ ).

Die Summen des EQ-S und EQ-F in der PCA-Auswertung korrelieren mittelstark miteinander ( $\rho = .35$ ,  $n = 128$ ,  $p < .001$ ). Weiterhin werden die Korrelationen zwischen den sich entsprechenden Unterfaktoren in der Selbst- und Fremdbeurteilungsvariante berichtet. Sowohl für die kognitive Empathie (EQ-S Faktor 1 und EQ-F Faktor 2:  $\rho = .22$ ,  $n = 128$ ,  $p = .013$ ), als auch für die emotionale Reaktivität (EQ-S Faktor 2 und EQ-F Faktor 3:  $\rho = .28$ ,  $n = 128$ ,  $p = .001$ ) sowie für die empathische Betroffenheit (EQ-S Faktor 3 und EQ-F Faktor 4:  $\rho = .26$ ,  $n = 128$ ,  $p = .003$ ) finden sich signifikante Korrelationen zwischen Selbst- und Fremdbeurteilung.

## 3.4 Fragestellung V4: Geschlechtsunterschiede

Wie in den englischen Originalartikeln finden sich auch für die vorliegende deutsche Übersetzung Geschlechtsunterschiede sowohl für den EQ-S in der regulären Auswertung ( $t = -3.56$ ,  $df = 138$ ,  $p = .001$ ,  $d = 0.70$  (mittlerer Effekt), bei gegebener Varianzgleichheit) als auch für EQ-F in der regulären Auswertung ( $t = -3.16$ ,  $df = 126$ ,  $p = .002$ ,  $d = 0.64$  (mittlerer Effekt), bei gegebener Varianzgleichheit). Im Shapiro-Wilk-Test zeigt sich, dass die Normalverteilungsannahme für den Gesamtscore des EQ-S aufrecht erhalten werden kann ( $W = 0.99$ ,  $df = 128$ ,  $p = .88$ ). Für den

Gesamtscore des EQ-F hingegen ist sie abzulehnen ( $W = 0.98$ ,  $df = 128$ ,  $p = .020$ ). Der EQ-S weist eine Schiefe von 0.10 und Kurtosis von 0.030 auf, der EQ-F eine Schiefe von -0.45 und Kurtosis von -0.30. Zur Untersuchung der Geschlechtsunterschiede bezüglich des EQ-F wird daher der nonparametrische Mann-Whitney-U-Test berechnet. Auch hier zeigen sich signifikante Unterschiede zwischen Männern ( $Md = 38$ ) und Frauen ( $Md = 48$ ) für den EQ-F ( $U = 964$ ,  $p = .003$ ,  $d = 0.64$  (mittlerer Effekt)).

Um zu untersuchen, ob Geschlechtsunterschiede bezüglich der einzelnen Unterskalen bestehen, die in der Faktorenanalyse identifiziert wurden, wird eine MANOVA mit der unabhängigen Variable Geschlecht und den sieben Unterskalen des EQ-S und des EQ-F als abhängige Variablen berechnet. Da der Box-Test nicht signifikant wird ( $p = .33$ ) kann die Homogenität der Kovarianzmatrizen als gegeben angenommen werden. Die einzelnen Variablen scheinen nach den Ergebnissen des Shapiro-Wilk-Tests nicht normalverteilt zu sein. Damit kann die multivariate Normalverteilung nicht als gegeben angenommen werden. Der Levene-Test auf Gleichheit der Fehlervarianzen ist nicht signifikant ( $p > 0.05$ ) für alle abhängigen Variablen, das heißt, die Annahme der Varianzhomogenität kann aufrecht erhalten werden. Die Verwendung der Pillai-Spur (ebenso wie Wilks-Lambda, Hotelling-Spur und größte charakteristische Wurzel nach Roy) zeigt einen signifikanten Effekt des Geschlechts auf die sieben verschiedenen Unterskalen der Selbst- und Fremdbeurteilung der Empathie ( $V = 0.22$ ,  $F(7, 120) = 4.75$ ,  $p < .001$ ,  $f^2 = .22$  (mittlerer Effekt)). Auch einzelne ANOVAs zeigen signifikante Haupteffekte des Geschlechts auf fünf der sieben Skalen: EQ-S Faktor 2 emotionale Reaktivität ( $F(1, 126) = 20.0$ ,  $p < .001$ ,  $f = 0.37$  (mittlerer Effekt)), EQ-S Faktor 3 empathische Betroffenheit ( $F(1, 126) = 23.3$ ,  $p < .001$ ,  $f = 0.39$  (mittlerer Effekt)), EQ-F Faktor 1 soziales Geschick ( $F(1, 126) = 4.13$ ,  $p = .044$ ,  $f = 0.18$  (kleiner Effekt)), EQ-F Faktor 3 emotionale Reaktivität ( $F(1, 126) = 4.65$ ,  $p = .033$ ,  $f = 0.19$  (kleiner Effekt)) und EQ-F Faktor 4 empathische Betroffenheit ( $F(1, 126) = 8.54$ ,  $p = .004$ ,  $f = 0.25$  (mittlerer Effekt)). Das Geschlecht zeigt jedoch weder für EQ-S Faktor 1 kognitive Empathie ( $F(1, 126) = 2.27$ ,  $p = 0.13$ ) noch für EQ-F Faktor 2 kognitive Empathie ( $F(1, 126) = 1.12$ ,  $p = 0.29$ ) einen signifikanten Haupteffekt. Zur Untersuchung der Beziehung zwischen den verschiedenen abhängigen Variablen, wird eine Diskriminanzanalyse an die MANOVA angeschlossen: Es zeigt sich eine Diskriminanzfunktion die 100% der Varianz mit einer kanonischen Korrelation von  $R^2 = 0.47$  erklärt. Diese Diskriminanzfunktion unterscheidet signifikant zwischen Männern und Frauen ( $\Lambda = 0.78$ ,  $\chi^2(7) = 30.0$ ,  $p < .001$ ). Die Korrelationen zwischen den Ergebnissen und der Diskriminanzfunktion zeigen, dass EQ-S Faktor 3 empathische Betroffenheit ( $r = .82$ ), EQ-S Faktor 2 emotionale Reaktivität ( $r = .76$ ) und EQ-F Faktor 4 empathische Betroffenheit ( $r = .50$ ) relativ hoch auf die Diskriminanzfunktion laden. Auch EQ-F Faktor 3 emotionale Reaktivität ( $r = .37$ ) und EQ-F Faktor 1 soziales Geschick ( $r = .34$ ) laden mittelstark auf die Diskriminanzfunktion während EQ-S Faktor 1 kognitive Empathie ( $r = .26$ ) und EQ-S Faktor 2 kognitive Empathie ( $r = .18$ ) recht niedrig auf die Diskriminanzfunktion laden.

Zusammenfassend lässt sich annehmen, dass die aus dem englischen Original berichteten Geschlechtsunterschiede sowohl in der deutschen Übersetzung des EQ, als

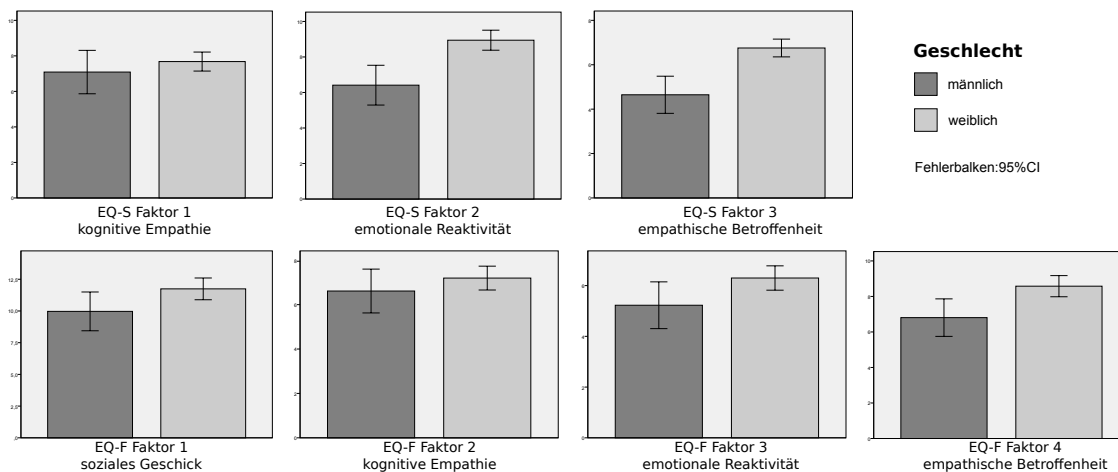


Abbildung 3.3: Geschlechtsunterschiede in den Unterfaktoren des EQ-S und EQ-F

auch in der neu entwickelten Fremdbeurteilungsversion beobachtet werden können. Bei weiblichen Teilnehmern scheint die Empathiefähigkeit höher ausgeprägt zu sein, als bei männlichen.

### 3.5 Fragestellung V5: Konstruktvalidität des EQ

Um die Konstruktvalidität genau zu untersuchen, müssen sowohl die konvergente Validität, als auch die diskriminante Validität Beachtung finden. Zur Überprüfung der konvergenten Validität kommen Testverfahren zum Einsatz, die dasselbe oder ähnliche Konstrukte messen. In der vorliegenden Untersuchung sind das der Empathizing-Score des EQ-SQ, der RME und die CAM, die alle zur Messung der Theory of Mind (ToM) beziehungsweise Empathie entwickelt wurden. Die Überprüfung der diskriminanten Validität erfolgt mit Testverfahren die Konstrukte abbilden, die nicht mit dem im Fokus stehenden überlappen. Im vorliegenden Fall kommen dafür der Systemizing-Score des EQ-SQ und die mit dem MWT-B geschätzte prä-morbide Intelligenz zum Einsatz.

Die Ergebnisse der Selbstbeurteilung im EQ-S und und der Fremdbeurteilung im EQ-F korrelieren mittelstark ( $\rho = .38$ ,  $n = 128$ ,  $p < .001$ ) miteinander.

#### 3.5.1 Konvergente Validität

##### EQ-SQ

Der mit dem EQ-SQ ermittelte Gesamtwert für Empathizing liegt im Mittel bei 14.50 (SD = 4.30) und korreliert erwartungsgemäß stark mit der Gesamtsumme des EQ-S in der regulären Auswertung ( $\rho = .73$ ,  $n = 139$ ,  $p < .001$ ). Auch mit dem EQ-F in der regulären Auswertung findet sich eine schwach positive Korrelation ( $\rho = .26$ ,  $n = 127$ ,  $p = .003$ ).

Wie erwartet korreliert die mit dem EQ-SQ gemessene Empathie ebenfalls positiv mit der EQ-S Summe nach PCA-Auswertung ( $\rho = .76$ ,  $n = 139$ ,  $p < .001$ ) und mit der EQ-F Summe nach PCA-Auswertung ( $\rho = .25$ ,  $n = 127$ ,  $p = .005$ ). Bei der

Betrachtung der einzelnen Unterfaktoren des EQ-S und EQ-F finden sich folgende signifikante Zusammenhänge: Die mit dem EQ-SQ gemessene Empathie korreliert positiv mit dem EQ-S Faktor 1 kognitive Empathie ( $\rho = .78$ ,  $n = 139$ ,  $p < .001$ ), mit dem EQ-S Faktor 2 emotionale Reaktivität ( $\rho = .68$ ,  $n = 139$ ,  $p < .001$ ), mit dem EQ-S Faktor 3 empathische Betroffenheit ( $\rho = .29$ ,  $n = 139$ ,  $p < .001$ ), mit dem EQ-F Faktor 2 kognitive Empathie ( $\rho = .19$ ,  $n = 127$ ,  $p = .037$ ), mit dem EQ-F Faktor 3 emotionale Reaktivität ( $\rho = .24$ ,  $n = 127$ ,  $p = .008$ ) sowie mit dem EQ-F Faktor 4 empathische Betroffenheit ( $\rho = .26$ ,  $n = 127$ ,  $p = .003$ ).

## RME

Die Anzahl richtiger Antworten im RME liegt im Mittel bei 26.0 ( $SD = 2.84$ ,  $Min = 19$ ,  $Max = 32$ ), die mittlere Reaktionszeit beträgt 7680 ms ( $SD = 2791$ ,  $Min = 3475$ ,  $Max = 27278$ ). Somit ähneln die Ergebnisse den Daten der Normierung des RME (Baron-Cohen et al., 2001) mit 26.2 für Normalbevölkerung beziehungsweise 28 für studentische Stichproben und den Ergebnissen von Lawrence et al. (2004) mit 27.6. Da die Normalverteilungsannahme für die Anzahl richtiger Antworten im RME nicht aufrecht erhalten werden kann wird der Spearmansche Korrelationskoeffizient berechnet: Der RME korreliert weder signifikant mit dem EQ-S in der regulären Auswertung ( $\rho = .12$ ,  $n = 140$ ,  $p = .17$ ), noch mit dem EQ-F in der regulären Auswertung ( $\rho = .12$ ,  $n = 128$ ,  $p = .17$ ).

Weiterhin finden sich keine Korrelationen des RME mit EQ-Faktoren die soziale Fähigkeiten messen (z.B. EQ-F Faktor 1 soziales Geschick). Es findet sich aber eine leichte Korrelation zwischen der Gesamtsumme im RME und dem EQ-S Faktor 2 emotionale Reaktivität ( $\rho = .19$ ,  $p = .023$ ). Keiner der anderen Faktoren der Selbst- oder Fremdbeurteilungsversion des EQ korreliert signifikant ( $p < .05$ ) mit der Anzahl richtiger Antworten im RME.

## CAM

Die Summe richtiger Antworten der CAM korreliert nicht signifikant mit dem EQ-S in der regulären Auswertung, es zeigt sich jedoch eine Tendenz ( $\rho = .17$ ,  $n = 139$ ,  $p = .052$ ). Mit dem EQ-F in der regulären Auswertung findet sich eine signifikante schwache Korrelation ( $\rho = .20$ ,  $n = 127$ ,  $p = .023$ ). Wird die CAM aufgeteilt in positive, negative und neutrale Items, zeigen weder positive noch neutrale Items signifikante Korrelationen mit EQ-S oder EQ-F in der regulären Auswertung. Die negativen Items hingegen korrelieren schwach mit dem EQ-S ( $\rho = .23$ ,  $n = 139$ ,  $p = .008$ ) und dem EQ-F ( $\rho = .18$ ,  $n = 127$ ,  $p = .047$ ).

Die Summe richtiger Antworten der CAM insgesamt korreliert signifikant mit der EQ-F Summe nach PCA-Auswertung ( $\rho = .22$ ,  $n = 127$ ,  $p = .017$ ), mit dem EQ-S Faktor 2 emotionale Reaktivität ( $\rho = .20$ ,  $n = 139$ ,  $p = .018$ ), mit dem EQ-F Faktor 1 soziales Geschick ( $\rho = .21$ ,  $n = 127$ ,  $p = .017$ ), mit dem EQ-F Faktor 2 kognitive Empathie ( $\rho = .18$ ,  $n = 127$ ,  $p = .047$ ) sowie mit dem EQ-F Faktor 3 emotionale Reaktivität ( $\rho = 0.19$ ,  $n = 127$ ,  $p = .031$ ). Werden die Ergebnisse der CAM für negative, positive und neutrale Items getrennt untersucht, so finden sich signifikante Korrelationen der Summe richtiger Antworten bei negativen Items mit

der EQ-S Summe ( $\rho = .20$ ,  $n = 139$ ,  $p = .002$ ) sowie dem EQ-S Faktor 2 emotionale Reaktivität ( $\rho = .26$ ,  $n = 139$ ,  $p = .002$ ).

### 3.5.2 Diskriminante Validität

#### EQ-SQ

Der Systemizing Gesamtwert liegt im Mittel bei 10.64 ( $SD = 4.59$ ) und korreliert nicht mit dem EQ-S in der regulären Auswertung ( $\rho = -.15$ ,  $n = 137$ ,  $p = .075$ ), aber schwach negativ mit dem EQ-F in der regulären Auswertung ( $\rho = -.23$ ,  $n = 125$ ,  $p = .011$ ).

In der PCA-Auswertung korreliert das Systemizing schwach negativ mit der EQ-S Summe ( $\rho = -.24$ ,  $n = 137$ ,  $p = .005$ ) und der EQ-F Summe ( $\rho = -.20$ ,  $n = 125$ ,  $p = .028$ ). Bei der Betrachtung der einzelnen Unterfaktoren des EQ-S und EQ-F finden sich folgende signifikante Zusammenhänge: Das Systemizing korreliert negativ mit dem EQ-S Faktor 1 kognitive Empathie ( $\rho = -.21$ ,  $n = 137$ ,  $p = .015$ ), mit dem EQ-S Faktor 2 emotionale Reaktivität ( $\rho = -.22$ ,  $n = 137$ ,  $p = .010$ ), mit dem EQ-S Faktor 3 empathische Betroffenheit ( $\rho = -.22$ ,  $n = 137$ ,  $p = .011$ ), mit dem EQ-F Faktor 1 soziales Geschick ( $\rho = -.19$ ,  $n = 125$ ,  $p = .035$ ) und dem EQ-F Faktor 4 empathische Betroffenheit ( $\rho = -.18$ ,  $n = 125$ ,  $p = .040$ ).

#### Intelligenz

Lawrence et al. (2004) maßen die verbale Intelligenz mit dem National Adult Reading Test und fanden einen durchschnittlichen geschätzten Intelligenzquotienten (IQ) von 120.48 ( $SD = 4.7$ ). In der vorliegenden Studie wurde die allgemeine Intelligenz anhand der Summe richtiger Antworten im MWT-B erhoben: der geschätzte IQ liegt im Mittel bei 107.65 ( $SD = 9.72$ ). Das Ergebnis des MWT-B korreliert wie bei Lawrence et al. nicht mit dem EQ-S in der regulären Auswertung ( $\rho = .018$ ,  $n = 140$ ,  $p = .83$ ) und ebenfalls nicht mit dem EQ-F in der regulären Auswertung ( $\rho = -.022$ ,  $n = 128$ ,  $p = .81$ ). Anders als bei Lawrence et al. findet sich keine signifikante Korrelation der Intelligenz mit der Anzahl richtiger Antworten im RME ( $\rho = -.036$ ,  $n = 140$ ,  $p = .67$ ) und auch nicht mit der mittleren Reaktionszeit im RME ( $\rho = .063$ ,  $n = 140$ ,  $p = .46$ ).

Auch in der PCA-Auswertung finden sich keinerlei signifikante Korrelationen zwischen dem anhand des MWT-B geschätzten IQ und den Gesamtscores oder einzelnen Unterfaktoren des EQ-S und EQ-F.

Insgesamt kann davon ausgegangen werden, dass EQ-S und -F tatsächlich Empathie messen, dass sie also konstruktvalide sind. Die vorliegenden Daten geben Hinweise auf eine gute konvergente sowie diskriminante Validität, vor allem im Vergleich mit anderen Fragebogenmaßen. Behaviorale Maße der Empathie weisen uneindeutige Ergebnisse auf, was Hinweise auf eine Differenz der zugrundeliegenden Fähigkeiten hinweisen könnte.

### 3.6 Fragestellung V6: Explorative Suche nach Einflussfaktoren auf die Varianz des EQ

Eine lineare multiple Regressionsanalyse zur Vorhersage der Gesamtsumme im EQ-S anhand der acht Prädiktoren Alter, Geschlecht, Sozialkontakte, Schulabschluss, Berufsabschluss, Psychologiestudium, MWT-B-Summe und RME-Summe, ergibt eine Varianzaufklärung von 17.5% ( $R^2 = .18$ ,  $p = .001$ ,  $f^2 = .21$  (mittlerer Effekt)). Eine analoge Regressionsanalyse für die Fremdbeurteilungsversion zeigt: Dieselben Prädiktoren erklären 18.7% der Varianz des EQ-F ( $R^2 = .19$ ,  $p = .001$ ,  $f^2 = .23$  (mittlerer Effekt)). Die Ergebnisse für die einzelnen Prädiktoren finden sich in Tabelle 3.6.

Tabelle 3.6: Ergebnisse der multiplen Regression für den EQ-S sowie EQ-F

	B	SE B	$\beta$
<b>EQ-S</b>			
(Konstante)	-1.59	23.7	
Alter	0.13	0.25	.049
Geschlecht	4.96	2.31	.19 *
Sozialkontakte	2.77	2.23	.11
RME	0.60	0.32	.15 +
MWT-B	-0.013	0.31	-.004
höchster Schulabschluss	0.12	2.90	.003
höchster Berufsabschluss	-0.51	0.93	-.049
Psychologiestudium	5.90	1.95	.27 **
<b>EQ-F</b>			
(Konstante)	-30.1	34.3	
Alter	0.085	0.31	.028
Geschlecht	4.50	2.81	.15
Sozialkontakte	6.28	2.61	.22 *
RME	0.68	0.38	.15 +
MWT-B	-0.40	0.37	-.093
höchster Schulabschluss	4.92	5.79	.074
höchster Berufsabschluss	-0.60	1.20	-.049
Psychologiestudium	5.83	2.43	.23 *

Signifikanzniveaus: +  $p < .1$ , \*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$

Als Einflussfaktoren auf die Ergebnisse im EQ-S und -F kommen somit das Geschlecht, die Leistung im RME, die Anzahl der Sozialkontakte sowie ein aktuell oder in der Vergangenheit absolviertes Psychologiestudium in Frage.



## 3.7 Zusammenfassung der Ergebnisse der Validierungsstudie

Zusammenfassend lassen sich die anfangs gestellten Fragen auf Basis der vorliegenden Daten wie folgt beantworten:

- V1) Die vorliegenden Werte des EQ-S ähneln denen der englischen Originalartikel in der deskriptiven Beschreibung. Deskriptiv betrachtet fallen die Mittelwerte des EQ-F ähnlich aus wie die des EQ-S, weisen jedoch leicht größere Standardabweichungen und tiefere Minima sowie Maxima auf.
- V2) Sowohl EQ-S als auch -F weisen eine gute Durchführungs-, Auswertungs- und Interpretationsobjektivität auf. Die interne Konsistenz ist in beiden Fragebogenvarianten als gut anzusehen.
- V3) Für den EQ-S ergeben sich drei Faktoren, während für den EQ-F vier gefunden werden können. Inhaltlich scheinen die Strukturen daher vergleichbar mit denen des englischen Originals zu sein.
- V4) Die aus dem englischen Original berichteten Geschlechtsunterschiede lassen sich sowohl in der deutschen Übersetzung des EQ, als auch der neu entwickelten Fremdbeurteilungsversion beobachten: Frauen scheinen über eine höhere Empathie zu verfügen als Männer.
- V5) Es kann davon ausgegangen werden, dass sowohl EQ-S als auch -F konstruktvalide sind, dass sie also Empathie messen. Die vorliegenden Daten geben Hinweise auf eine gute konvergente sowie diskriminante Validität.
- V6) In einer explorativen regressionsanalytischen Betrachtung finden sich Hinweise für einen Einfluss des Geschlechts, der Ergebnisse im RME, der Anzahl der Sozialkontakte sowie eines aktuell oder in der Vergangenheit absolvierten Psychologiestudiums auf die Ergebnisse im EQ-S und -F.



# Kapitel 4

## Methoden der experimentellen Untersuchung

Im vorliegenden Kapitel werden die Methoden der mit Patienten und gesunden Kontrollen durchgeführten experimentellen Untersuchung detailliert besprochen. Es erfolgt eine Beschreibung der Rekrutierung und der Stichprobe mit demographischen Daten sowie der Ergebnisse der neuropsychologischen Testung. Die Ein- und Ausschlusskriterien, die verwendeten Fragebögen und Testverfahren ebenso wie die Abläufe und statistischen Verfahren kommen zur Sprache.

### 4.1 Stichprobenbeschreibung

Alle in der vorliegenden Arbeit beschriebenen Daten beziehen sich auf die von Juli 2013 bis 18. August 2015 untersuchte Stichprobe. Zur Rekrutierung der Teilnehmer wurde auf den Probandenpool der Abteilung für Klinische Psychologie und Psychotherapie am Fachbereich Psychologie der Eberhard Karls Universität Tübingen zugegriffen. In diesem Probandenpool sind sowohl gesunde Personen als auch Personen mit psychischer Erkrankung eingetragen, die Interesse an Studienteilnahmen haben. Weiterhin wurde auf der Internetseite der Abteilung für Klinische Psychologie und Psychotherapie, sowie in regionalen Zeitungen, in Internetforen für Betroffene mit affektiven Erkrankungen, in Arztpraxen und Kliniken, über Rundmails innerhalb der Eberhard Karls Universität und in Selbsthilfegruppen anhand von Flyern, Vorträgen und Artikeln Werbung für die Teilnahme an der Studie gemacht. Für die Teilnahme an der Studie wurde eine Aufwandsentschädigung von acht Euro pro Stunde beziehungsweise ein Maximalbetrag von 100 Euro angeboten. Die interessierten Personen meldeten sich telefonisch oder per Email und wurden im Rahmen eines ausführlichen Telefoninterviews auf mögliche Passung der Ein- und Ausschlusskriterien gescreent. Im genannten Zeitraum wurden nach erfolgreichen Telefoninterviews insgesamt 280 Personen zur ausführlichen Diagnostik eingeladen. Im Rahmen der Diagnostik mussten insgesamt 67 Personen (20 unipolar akute, 17 unipolar remittierte, 2 bipolar akute und 14 bipolar remittierte Patienten sowie 14 gesunde Kontrollen) aufgrund Nichterfüllung der Einschlusskriterien beziehungsweise Vorliegen von Ausschlusskriterien oder aufgrund von Problemen bei der Eyetracking-Messung von der Teilnahme ausgeschlossen werden. Daher ergab sich zum 18. August 2015 ein Da-

tensatz mit insgesamt 213 Personen, die beide experimentellen Sitzungen absolviert hatten. Im Nachhinein mussten bei der Durchsicht des Datensatzes acht weitere Personen aufgrund nicht vorliegender Einschlusskriterien beziehungsweise Vorliegen von Ausschlusskriterien ausgeschlossen werden. Weiterhin konnten für die Gruppe der Patienten mit akuter Bipolarer Störung lediglich zwei Probanden gewonnen werden, so dass diese wegen der geringen Gruppengröße ebenfalls ausgeschlossen wurden. Es ergibt sich ein Datensatz mit 203 Personen, der für die folgenden Beschreibungen der Stichprobe sowie für die statistischen Analysen herangezogen wird.

Die 130 Teilnehmer mit psychischer Erkrankung wurden anhand der gestellten Diagnose in drei Gruppen aufgeteilt: 45 Patienten mit remittierter unipolarer affektiver Störung (UR), 43 Patienten mit akuter unipolarer affektiver Störung (UA) und 42 Patienten mit remittierter bipolarer affektiver Störung (BR). Die 73 gesunden Kontrollen wurden bezüglich Alter, Geschlecht und Bildungsniveau gematcht in zwei Kontrollgruppen unterteilt: 38 gesunde Kontrollen für die unipolar remittierten Patienten (KG-UR) und 35 gesunde Kontrollen für die bipolar remittierten Patienten (KG-BR). In Tabelle 4.1 finden sich die Daten zur Geschlechterverteilung, Bildungsniveau, Erwerbstätigkeit und Regelmäßigkeit sozialer Kontakte für die fünf genannten Teilnehmergruppen.

Tabelle 4.1: Beschreibung der fünf Teilnehmergruppen: Geschlechterverteilung, Bildungsniveau, Erwerbstätigkeit und Regelmäßigkeit sozialer Kontakte

	BR	KG-BR	UA	UR	KG-UR
<b>Geschlecht</b>					
weiblich	64.3%	62.9%	74.4%	82.2%	81.6%
<b>Alter</b>					
<i>M</i>	47.9	48.0	41.0	46.3	42.3
<i>SD</i>	12.4	12.2	14.8	14.5	13.9
<i>Min</i>	20	20	19	20	18
<i>Max</i>	66	68	65	68	66
<b>Höchster Schulabschluss</b>					
Abitur	64.3%	74.3%	60.5%	68.9%	78.9%
<b>Aktuelle Beschäftigungssituation</b>					
erwerbstätig	58.5%	77.1%	27.9%	53.3%	57.9%
<b>Regelmäßigkeit der Sozialkontakte</b>					
mind. 1 pro Woche	78.6%	82.9%	65.1%	75.6%	92.1%

Anhand von  $\chi^2$ -Tests wird die Stichprobe auf ihre Homogenität untersucht. Mit  $\chi^2 = 6.90$  ( $p = .14$ ) sprechen die Daten nicht dafür, dass sich die Gruppen bezüglich der Geschlechterverteilung unterscheiden. Auch bezüglich des Bildungsniveaus ergeben sich mit  $\chi^2 = 4.11$  ( $p = .39$ ) keine Hinweise auf bedeutsame Gruppenunterschiede. Ebenso scheint das Alter mit  $\chi^2 = 647$  ( $p = .50$ ) in den Teilnehmergruppen nicht unterschiedlich zu sein. Das Matching der Patienten- und Kontrollgruppen hinsichtlich Geschlecht, Alter und Bildungsniveau kann also als gegeben angesehen

werden. Mit  $\chi^2 = 9.30$  ( $p = .054$ ) wird der  $\chi^2$ -Test für Gruppenzugehörigkeit und die Regelmäßigkeit der Sozialkontakte marginal signifikant. Dies kann als schwacher Hinweis auf eine höhere Regelmäßigkeit der Sozialkontakte in den Kontrollgruppen verstanden werden. Zur Illustration sind in Abbildung 4.1 die Anzahl von Personen mit oder ohne regelmäßige, das heißt mindestens einmal wöchentliche, Sozialkontakte dargestellt. Für die Gruppenzugehörigkeit und die Erwerbstätigkeit ergeben sich mit  $\chi^2 = 19.91$  ( $p = .001$ ) starke Hinweise auf Unterschiede bezüglich der Erwerbstätigkeit zwischen den Teilnehmergruppen. Insbesondere die akut erkrankten Patienten weisen eine geringere Erwerbstätigkeit auf, wie der Blick in die deskriptiven Daten zeigt.

Eine ausführlichere soziodemographische Beschreibung der Teilnehmergruppen ist in den Tabellen 4.2 bis 4.6 zu finden.

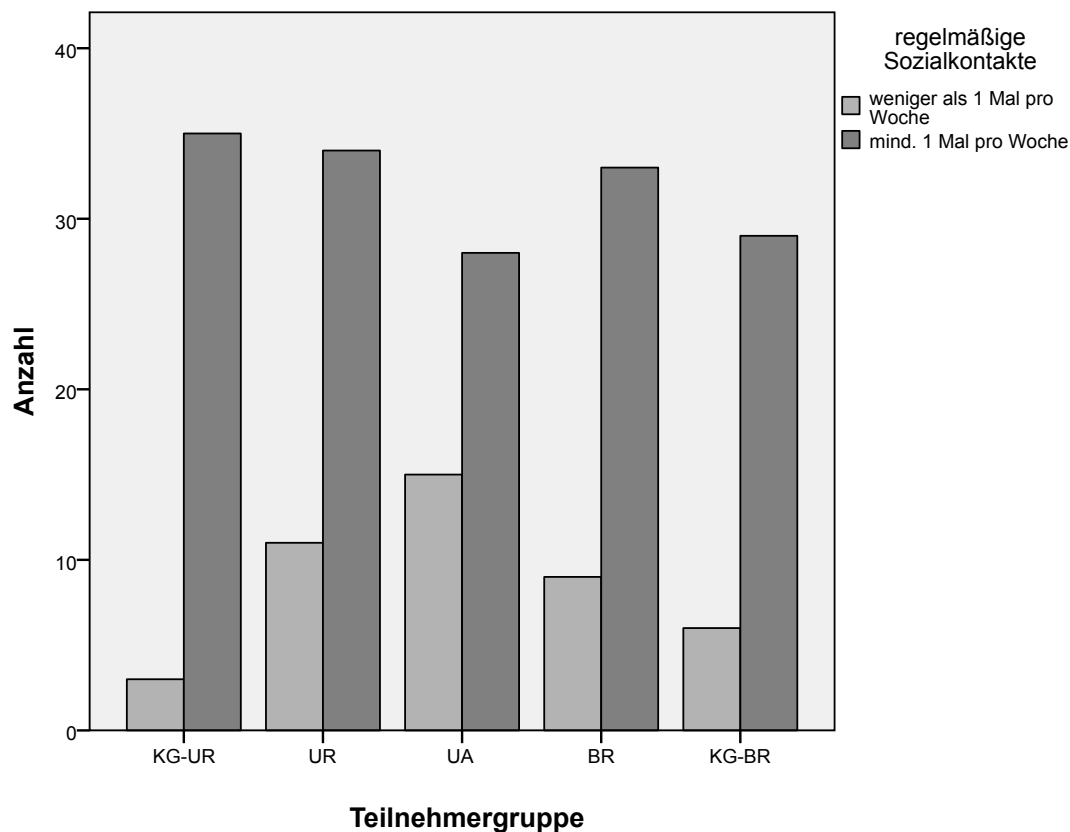


Abbildung 4.1: Sozialkontakte in den verschiedenen Teilnehmergruppen

Tabelle 4.2: Soziodemographische Beschreibung der Teilnehmergruppen

<i>Teil 1</i>	BR	KG-BR	UA	UR	KG-UR
<b>höchster Schulabschluss</b>					
noch Schüler	2.4%	—	—	—	—
Hauptschulabschluss	4.8%	2.9%	11.6%	6.7%	2.6%
mittlere Reife	28.6%	22.9%	27.9%	24.4%	18.4%
Abitur / Fachhochschulreife	64.3%	74.3%	60.5%	68.9%	78.9%
<b>höchster erreichter Berufsabschluss</b>					
noch in Berufsausbildung	4.8%	5.7%	20.9%	17.8%	13.2%
Lehre / Fachschule	42.9%	31.4%	32.6%	40.0%	28.9%
Meister	4.8%	2.9%	11.6%	2.2%	2.6%
Fachhochschule / Universität	45.2%	57.1%	27.9%	37.8%	42.1%
ohne Abschluss	2.4%	2.9%	4.7%	2.2%	13.2%

Tabelle 4.3: Soziodemographische Beschreibung der Teilnehmergruppen

<i>Teil 2</i>	BR	KG-BR	UA	UR	KG-UR
<b>Finanzierung des Lebensunterhalts</b>					
Eltern	4.8%	2.9%	25.6%	2.2%	15.8%
Partner	19.0%	11.4%	7.0%	17.8%	18.4%
sonstige Angehörige	—	2.9%	—	—	—
tariflich bezahlte Arbeit	21.4%	74.3%	23.3%	37.8%	44.7%
Arbeitslosenversicherung	2.4%	5.7%	2.3%	4.4%	5.3%
Krankengeld	11.9%	—	18.6%	6.7%	—
Sozialleistungen (z.B. ALG II)	2.4%	—	—	—	—
Rente	26.2%	2.9%	16.3%	20.0%	2.6%
Sonstiges	11.9%	—	7.0%	11.1%	13.2%

Tabelle 4.4: Soziodemographische Beschreibung der Teilnehmergruppen

<i>Teil 3</i>	BR	KG-BR	UA	UR	KG-UR
<b>derzeitige Arbeitssituation</b>					
berufstätig Vollzeit	21.4%	34.3%	9.3%	24.4%	26.3%
berufstätig Teilzeit	23.8%	40.8%	16.3%	26.7%	31.6%
berufstätig gelegentlich	11.9%	2.9%	2.3%	2.2%	—
Hausfrau/-mann, nicht berufstätig	2.4%	2.9%	—	2.2%	5.3%
Ausbildung, Umschulung	4.8%	11.4%	25.6%	15.6%	21.1%
arbeitslos gemeldet	4.8%	5.7%	2.3%	4.4%	7.9%
Rentenverfahren, Erwerbsunfähigkeitsrente	23.8%	2.9%	11.6%	15.6%	—
Altersrente, Pension	7.1%	—	2.3%	4.4%	2.6%
Witwen(r)rente	—	—	2.3%	—	—
anderweitig ohne berufliche Beschäftigung	—	—	9.3%	2.2%	2.6%
Sonstiges	—	—	18.6%	2.2%	2.6%



Tabelle 4.5: Soziodemographische Beschreibung der Teilnehmergruppen

<i>Teil 4</i>	BR	KG-BR	UA	UR	KG-UR
<b>derzeitig ausgeübter Beruf</b>					
nicht erwerbstätig	42.5%	17.1%	53.5%	42.2%	26.3%
un-/angelernter Arbeiter, Jobs, Hiwi	7.1%	11.4%	18.6%	2.2%	13.2%
Facharbeiter, unselbstständiger Handwerker	—	—	2.3%	2.2%	5.3%
einfacher Angestellter, Beamter im unteren Dienst	7.1%	14.3%	4.7%	11.1%	5.3%
mittlerer Angestellter, Beamter im mittleren Dienst	9.5%	20.0%	9.3%	17.8%	18.4%
hoch qualifizierter Angestellter, Beamter im gehobenen Dienst	11.9%	22.9%	7.0%	11.1%	21.1%
Angestellter, Beamter im höheren Dienst	9.5%	8.6%	2.3%	8.9%	2.6%
Freiberufler, größerer Unternehmer	7.1%	5.7%	2.3%	4.4%	5.3%
Sonstiges	2.4%	—	—	—	2.6%

Tabelle 4.6: Soziodemographische Beschreibung der Teilnehmergruppen

<i>Teil 5</i>	BR	KG-BR	UA	UR	KG-UR
<b>aktuelle Lebenssituation</b>					
alleine lebend	33.3%	28.6%	46.5%	42.2%	26.3%
mit Ehe-/Lebenspartner lebend	52.4%	51.4%	30.2%	51.1%	60.5%
mit Eltern/Verwandten lebend	9.5%	14.3%	7.0%	2.2%	2.6%
Sonstige	4.8%	5.7%	16.3%	4.4%	10.5%
<b>Sozialkontakte</b>					
mehr als einmal die Woche	57.1%	71.4%	48.8%	53.3%	76.3%
höchstens einmal pro Woche	21.4%	11.4%	16.3%	22.2%	15.8%
einmal in zwei Wochen	11.9%	8.6%	11.6%	6.7%	5.3%
einmal im Monat	7.1%	5.7%	9.3%	13.3%	2.6%
keine Treffen außer im Hausflur / bei der Arbeit	2.4%	2.9%	9.3%	4.4%	—
trifft Bekannte/Freunde unter keinen Umständen	—	—	4.7%	—	—

### 4.1.1 Ein- und Ausschlusskriterien

Im Rahmen der Telefoninterviews sowie der ausführlichen Diagnostiksitung wurde eine Reihe von Ein- und Ausschlusskriterien überprüft. Die folgenden Ausschlusskriterien gelten ausnahmslos für alle Teilnehmer der Untersuchung:

- Unzureichende Kenntnis der deutschen Sprache
- Diagnose einer Erkrankung des ZNS (z.B. Morbus Parkinson)
- Alter unter 18 oder über 69 Jahren
- Abhängigkeit von Alkohol oder illegalen Drogen (aktuell oder anamnestisch falls weniger als drei Jahre abstinent)
- Missbrauch von Alkohol oder illegalen Drogen (aktuell)
- Psychotische Störung (aktuell und anamnestisch): Schizophrenie, schizophreniforme Störung, schizoaffektive Störung, wahnhaftige Störung, psychotische Merkmale während einer affektiven Störung (mit Ausnahme von stimmungskongruentem Wahn)
- Anorektische Essstörung (aktuell, BMI  $\leq 18$  kg/m<sup>2</sup>)
- Folgende Achse-II Störungen: Cluster A (schizoid, schizotyp, paranoid), Antisozial, Borderline

Für gesunde Kontrollprobanden war zusätzlich jegliche aktuell oder anamnestisch vorliegende psychische Erkrankung ein Ausschlusskriterium. Hierbei wurden sowohl die Untersuchung anhand SKID-I und II aber auch bisherige psychotherapeutische oder psychopharmakologische Behandlungen beachtet. Die störungsspezifischen Einschlusskriterien für die drei beziehungsweise vier Patientengruppen werden im Folgenden genannt.

#### **Patienten mit akuter unipolarer Depression:**

- Folgende Diagnosen: Major Depression einzelne Episode, Rezidivierende Major Depression, Dysthyme Störung
- Keine oder seit mindestens vier Wochen stabile psychotrope Medikation, keine Änderung während der Untersuchung
- HAMD-17  $> 13$  (entspricht ca. IDS-C30  $> 23$ )

#### **Patienten mit remittierter unipolarer Depression:**

- Folgende Diagnosen: Major Depression Episode, Rezidivierende Major Depression, Dysthyme Störung; gegenwärtig remittiert seit mindestens 8 Wochen
- Keine oder seit mindestens vier Wochen stabile psychotrope Medikation, keine Änderung während der Untersuchung
- HAMD-17  $< 8$  (entspricht ca. IDS-C30  $< 12$ )

**Patienten mit akuter Bipolarer Störung:**

- Folgende Diagnosen: Bipolar I, gegenwärtig depressiv; Bipolar II, gegenwärtig depressiv; Zykllothyme Störung, gegenwärtig depressiv
- Keine oder seit mindestens vier Wochen stabile psychotrope Medikation, keine Änderung während der Untersuchung
- HAMD-17 > 13 (entspricht ca. IDS-C30 > 23)
- YMRS < 12

**Patienten mit remittierter Bipolarer Störung:**

- Folgende Diagnosen: Bipolar I, Bipolar II, Zykllothyme Störung; gegenwärtig remittiert seit mindestens 8 Wochen
- Keine oder seit mindestens vier Wochen stabile psychotrope Medikation, keine Änderung während der Untersuchung
- HAMD 17 < 8 (entspricht ca. IDS-C30 < 12)
- YMRS < 12

**4.1.2 Diagnosen und Krankheitsgeschichte**

Anhand des SKID I und II wurden alle Teilnehmer auf psychische Erkrankungen untersucht. In Tabelle 4.7 finden sich deskriptive Daten zum Vorliegen komorbider Diagnosen, zu psychisch erkrankten Familienangehörigen sowie zu Suizidversuchen und Behandlungsdaten. Wie durch die Ein- und Ausschlusskriterien vorgegeben, sind in den Kontrollgruppen weder Diagnosen auf Achse I oder II noch Suizidversuche oder eine Behandlung aufgrund einer psychischen Erkrankung aufgetreten.

$\chi^2$ -Homogenitätstests weisen darauf hin, dass sich die Teilnehmergruppen bezüglich der Komorbiditäten auf Achse I ( $\chi^2 = 65.2$ ,  $p < .001$ ), auf Achse II ( $\chi^2 = 14.2$ ,  $p = .007$ ) und bezüglich der Familienanamnese psychischer Erkrankungen ( $\chi^2 = 52.4$ ,  $p < .001$ ) signifikant unterscheiden.

In Tabelle 4.8 ist die aktuelle psychopharmakologische Medikation in den Patientengruppen anhand der Substanzklassen und Anzahl der Wirkstoffe beschrieben. In Tabelle C.1 im Anhang findet sich eine detaillierte Aufschlüsselung der eingesetzten Medikamente in den Teilnehmergruppen.

**Aktuelle Symptombelastung**

In Tabelle 4.9 findet sich eine Beschreibung der aktuellen Symptombelastung der Teilnehmer anhand der deskriptiven Daten von BDI und MSS für alle Teilnehmergruppen und zusätzlich HAMD, IDS-C und YMRS für die Patientengruppen. Da das Erreichen bestimmter Werte in den Fremdbeurteilungsmaßen als Einschlusskriterium galt, wird auf den statistischen Vergleich der Teilnehmergruppen bezüglich der Symptombelastung verzichtet.

Tabelle 4.7: Komorbiditäten, Familienanamnese und Krankheitsgeschichte

	BR	KG-BR	UA	UR	KG-UR
<b>Komorbidie Diagnose vorhanden</b>					
Achse I	40.5%	—	62.8%	57.8%	—
Achse II	4.80%	—	16.3%	4.40%	—
<b>Familienanamnese</b>					
mind. 1 psychisch erkrankter Verwandter	92.9%	34.3%	67.4%	75.6%	26.3%
<b>Krankheitsgeschichte</b>					
Suizidversuch in der Vorgeschichte	38.1%	—	30.2%	22.2%	—
Behandlung einer psychischen Erkrankung	93.2%	—	95.5%	93.3%	—
Behandlungsdauer in Jahren [ $M$ ( $SD$ )]	17.6 (12.5)	—	12.1 (8.89)	16.7 (14.1)	—
aktuelle psychopharmakologische Medikation	95.2%	—	83.7%	48.9%	—

### 4.1.3 Neuropsychologie deskriptiv

Um etwaige Gruppenunterschiede bezüglich der zu untersuchenden Fragestellungen besser interpretieren zu können, wurde eine neuropsychologische Testung durchgeführt. Mit dem MWT-B kann die prämorbid beziehungsweise allgemeine Intelligenz anhand verbaler Fähigkeiten geschätzt werden. Der TMT-A und B misst die einfache und geteilte Aufmerksamkeit. Die Merkfähigkeit wird mit dem VLMT untersucht.

#### MWT-B

Die Ergebnisse des MWT-B sind in Tabelle 4.10 zu finden. Es sind sowohl die Rohwerte als auch geschätzte IQ-Werte und die Intelligenzstufen berichtet. Anhand von  $\chi^2$ -Homogenitätstests werden die Daten auf mögliche Gruppenunterschiede untersucht. Weder für Gruppenzugehörigkeit und den Rohwert ( $\chi^2 = 78.2$ ,  $p = .29$ ) noch für Gruppenzugehörigkeit und geschätzten IQ ( $\chi^2 = 78.2$ ,  $p = .29$ ) oder Gruppenzugehörigkeit und Intelligenzstufen ( $\chi^2 = 11.3$ ,  $p = .51$ ) können signifikante Unterschiede beobachtet werden. Es gibt also keine Hinweise auf unterschiedliche Intelligenz in den verschiedenen Teilnehmergruppen.

Tabelle 4.8: Aktuelle Psychopharmakologische Medikation der Patientengruppen

	BR	UA	UR
<b>Substanzklassen</b>			
Antidepressiva	42.5%	81.4%	44.4%
Mood Stabilizer	78.6%	9.3%	4.4%
Neuroleptika	42.5%	20.9%	4.4%
Sedativa	—	7.0%	4.4%
<b>Anzahl Wirkstoffe</b>			
Antidepressiva			
1 Wirkstoff	35.7%	55.6%	40%
2 Wirkstoffe	9.5%	20.9%	4.4%
3 Wirkstoffe	—	4.7%	—
Mood Stabilizer			
1 Wirkstoff	61.9%	9.3%	2.2%
2 Wirkstoffe	16.7%	—	2.2%
Neuroleptika			
1 Wirkstoff	40.5%	20.9%	4.4%
2 Wirkstoffe	2.4%	—	—
3 Wirkstoffe	2.4%	—	—
Sedativa			
1 Wirkstoff	—	4.7%	4.4%
2 Wirkstoffe	—	2.3%	—

Tabelle 4.9: Aktuelle Symptombelastung

	BR	KG-BR	UA	UR	KG-UR
<b>BDI</b>					
<i>M</i>	5.33	2.37	26.6	4.82	1.63
<i>SD</i>	5.11	3.14	8.67	5.67	2.25
<i>Min</i>	0	0	5	0	0
<i>Max</i>	17	14	47	28	9
<b>HAMD</b>					
<i>M</i>	2.29	—	17.88	2.02	—
<i>SD</i>	2.30	—	3.28	2.37	—
<i>Min</i>	0	—	14	0	—
<i>Max</i>	8	—	28	7	—
<b>IDS-C</b>					
<i>M</i>	4.83	—	34.07	4.02	—
<i>SD</i>	5.18	—	7.40	5.03	—
<i>Min</i>	0	—	21	0	—
<i>Max</i>	19	—	52	21	—
<b>MSS</b>					
<i>M</i>	2.79	1.46	3.38	2.56	1.39
<i>SD</i>	4.23	2.09	3.23	2.97	2.01
<i>Min</i>	0	0	0	0	0
<i>Max</i>	20	8	13	10	8
<b>YMRS</b>					
<i>M</i>	1.02	—	—	—	—
<i>SD</i>	1.91	—	—	—	—
<i>Min</i>	0	—	—	—	—
<i>Max</i>	7	—	—	—	—

Tabelle 4.10: Ergebnisse im MWT-B

	BR	KG-BR	UA	UR	KG-UR
<b>Deskriptive Daten:</b>					
<i>M (SD)</i>					
Rohwert der Summe	31.1 (3.48)	32.3 (3.48)	30.3 (4.27)	30.2 (4.42)	31.8 (3.40)
geschätzter IQ	117 (17.2)	122 (15.4)	114 (16.8)	114 (16.0)	119 (16.1)
<b>Intelligenzstufen</b>					
niedrige Intelligenz	2.4%	—	2.3%	4.4%	—
durchschnittliche Intelligenz	38.1%	22.9%	48.8%	40.0%	34.2%
hohe Intelligenz	23.8%	40.0%	20.9%	31.1%	28.9%
sehr hohe Intelligenz	35.7%	37.1%	27.9%	24.4%	36.8%

**TMT-A und B**

In Tabelle 4.11 sind die Ergebnisse für den TMT-A und -B dargestellt. Mittels  $\chi^2$ -Homogenitätstests werden sowohl die Rohwerte in beiden Testteilen als auch die Verteilung auf die Kategorien perfectly normal, normal, mild/moderately impaired und moderately/severely impaired in den Teilnehmergruppen untersucht. Mit  $\chi^2 = 169$  ( $p = .54$ ) für Teil A und  $\chi^2 = 317$  ( $p = .41$ ) für Teil B ergeben sich keine Hinweise auf Unterschiede bezüglich der Rohwerte zwischen den Teilnehmergruppen. Auch bei Einteilung der Ergebnisse in die genannten Kategorien scheinen sich die Teilnehmergruppen weder in Teil A ( $\chi^2 = 12.1$ ,  $p = .44$ ) noch in Teil B ( $\chi^2 = 14.3$ ,  $p = .28$ ) bedeutsam zu unterscheiden. Insgesamt scheinen also keine Gruppenunterschiede bezüglich der einfachen oder geteilten Aufmerksamkeit zu bestehen.

Tabelle 4.11: Ergebnisse im TMT (in Sekunden)

	BR	KG-BR	UA	UR	KG-UR
<b>Teil A</b>					
<i>M</i>	31.7	27.8	29.4	30.5	26.8
<i>SD</i>	11.9	8.95	11.7	15.1	8.48
<b>Teil B</b>					
<i>M</i>	76.2	63.7	68.4	72.9	57.3
<i>SD</i>	38.4	17.9	29.1	31.3	17.5

**VLMT**

Die ausführlichen Ergebnisse des VLMT sind in Tabelle 4.12 zu finden. Wie dort zu sehen ist, wurden alle Kennwerte des VLMT erfasst. Besonders relevant sind die Gesamtlernleistung, die Abrufleistung der Lernliste nach Interferenz, der Verlust nach Interferenz, der Verlust nach zeitlicher Verzögerung sowie die korrigierte



Tabelle 4.12: Deskriptive Ergebnisse des VLMT: Mittelwert (Standardabweichung)

	BR	KG-BR	UA	UR	KG-UR
Supraspanne	7.98 (2.16)	8.20 (2.10)	8.28 (2.20)	8.44 (2.20)	8.97 (2.12)
Lernleistung	12.5 (3.05)	13.8 (1.32)	13.5 (1.59)	13.6 (1.60)	13.8 (1.18)
Gesamtlernleistung	55.8 (9.33)	59.5 (7.38)	58.4 (9.43)	58.4 (8.96)	61.0 (7.36)
Abruffleistung Lernliste nach Interferenz	11.2 (2.89)	12.4 (2.66)	11.6 (3.39)	12.4 (2.51)	12.7 (2.09)
Abruffleistung Lernliste nach Verzögerung	11.3 (2.76)	12.6 (2.57)	12.1 (2.73)	12.5 (2.44)	13.0 (1.91)
Verlust nach Interferenz (Dg5-Dg6)	1.26 (2.54)	1.37 (1.85)	1.83 (2.66)	1.20 (1.90)	1.13 (1.40)
Verlust nach Verzögerung (Dg5-Dg6)	1.17 (2.28)	1.17 (1.86)	1.40 (1.79)	1.04 (1.43)	0.79 (1.36)
Wiedererkennensleistung	14.4 (0.80)	14.5 (1.09)	14.3 (1.21)	14.5 (1.10)	14.5 (0.76)
korrigierte Wiedererken- nensleistung	13.4 (2.82)	13.9 (1.96)	14.0 (1.86)	14.2 (1.86)	14.2 (1.26)
Perseverationen (Dg1 bis Dg5)	3.31 (2.90)	3.03 (3.88)	2.64 (3.44)	2.98 (4.06)	4.18 (7.52)
Inferenzen (Dg1 bis Dg5)	0.88 (1.89)	0.34 (0.80)	0.44 (1.08)	0.64 (1.46)	0.50 (0.95)

Wiedererkennensleistung. Anhand von  $\chi^2$ -Homogenitätstests werden die Teilnehmergruppen auf Unterschiede bezüglich der Lern- und Merkfähigkeit untersucht. Weder für die Gesamtlernleistung ( $\chi^2 = 120$ ,  $p = .93$ ), die Abruffleistung Lernliste nach Interferenz ( $\chi^2 = 59.5$ ,  $p = .22$ ), den Verlust nach Interferenz ( $\chi^2 = 44.0$ ,  $p = .78$ ), den Verlust nach zeitlicher Verzögerung ( $\chi^2 = 49.0$ ,  $p = .28$ ) noch für die korrigierte Wiedererkennensleistung ( $\chi^2 = 31.4$ ,  $p = .83$ ) sprechen die Ergebnisse für Unterschiede zwischen den Teilnehmergruppen. Auch bezüglich der zusätzlichen Kennwerte werden keine Hinweise auf Unterschiede zwischen den Teilnehmergruppen gefunden: Die Gruppen scheinen sich weder bezüglich der Supraspanne ( $\chi^2 = 51.2$ ,  $p = .21$ ), der Lernleistung ( $\chi^2 = 25.5$ ,  $p = .90$ ), der Abruffleistung Lernliste nach Verzögerung ( $\chi^2 = 40.1$ ,  $p = .46$ ), der Wiedererkennensleistung ( $\chi^2 = 16.4$ ,  $p = .69$ ), der Perseverationen ( $\chi^2 = 57.4$ ,  $p = .82$ ) noch der Inferenzen ( $\chi^2 = 20.8$ ,  $p = .84$ ) zu unterscheiden. Es kann also davon ausgegangen werden, dass sich die Patienten- und Kontrollgruppen bezüglich ihrer verbalen Lern- und Merkfähigkeit nicht unterscheiden.

## 4.2 Messinstrumente und -methoden

### 4.2.1 Allgemeine Diagnostik

#### **Strukturiertes Klinisches Interview für DSM-IV (SKID-I und -II)**

Zur allgemeinen Diagnostik psychischer Erkrankungen wurden alle potentiellen Teilnehmer von geschulten Interviewern mit dem ausführlichen und weit verbreiteten SKID-I und -II (Strukturiertes Klinisches Interview für DSM-IV; Wittchen et al., 1997) untersucht. Dieses Interviewverfahren ist an der vierten Auflage des Diagnostischen und Statistischen Manual Psychischer Störungen (DSM-IV; Saß et al., 2003) orientiert. Entsprechend den Achsen I und II des DSM-IV gibt es beim SKID einen Teil I der aktuellen und vergangene psychischen Erkrankungen erfasst und einen Teil II der zur Diagnose von Persönlichkeitsstörungen dient.

Das SKID-I enthält Kapitel zur Diagnostik affektiver Störungen, psychotischer Störungen, Substanzmissbrauch und Substanzabhängigkeiten, Angststörungen, Somatoforme Störungen, Essstörungen sowie Anpassungsstörungen und weitere optionale DSM-IV Störungen. In diesen Kapiteln erfolgt eine Erfassung der aktuell oder ehemals vorhandenen Symptome mittels Fragen in vorgegebener Reihenfolge und Wortlaut. Die freien Antworten des Probanden werden vom Interviewer beurteilt (? = keine Information; 1 = keine Hinweise für das Vorliegen des Kriteriums; 2 = eindeutige Hinweise aber nicht kriteriumsgemäßes Vorliegen des Merkmals; 3 = das Kriterium ist voll erfüllt) und so das Vorliegen bestimmter Kriterien und Diagnosen eingeschätzt. Auch die Chronologie und der Schweregrad der diagnostizierten Erkrankungen werden erfragt. Aufgrund von Sprungregeln ist eine zeitlich ökonomische und gleichzeitig umfassende Diagnosestellung möglich. Die Durchführungsdauer des SKID-I variiert sehr stark, abhängig von den vorliegenden Diagnosen. Mit gesunden Kontrollen werden etwa 20-30 Minuten benötigt, mit Patienten etwa 120 Minuten. Zusätzlich zu den genannten Kapiteln verschiedener Störungen ist mit dem SKID-I eine Erfassung der psychosozialen Beeinträchtigungen entsprechend der Achse IV des DSM-IV und eine globale Beurteilung der Leistungsfähigkeit entsprechend der Achse V des DSM-IV möglich. Letzteres wird als Global Assessment of Functioning (GAF) bezeichnet und bedeutet eine Einschätzung des aktuellen allgemeinen Funktionsniveaus auf einer Skala von 1 (Ständige Gefahr) bis 100 (Gute Leistungsfähigkeit auf allen Gebieten). Hierbei sind Zwischenwerte wie beispielsweise 68 möglich. Im Rahmen der vorliegenden Studie wurden alle Kapitel des SKID-I sowie die Erfassung des GAF durchgeführt.

Das SKID-II ist ein zweistufiges Verfahren, bestehend aus einem sogenannten Screening-Fragebogen und einem Interview zur genaueren Untersuchung der Kriterien der einzelnen Persönlichkeitsstörungen. Es werden die folgenden Persönlichkeitsstörungen entsprechend der Systematik des DSM-IV erfasst: Selbstunsichere, Dependente, Zwanghafte, Negativistische, Depressive, Paranoide, Schizotypische, Schizoide, Histrionische, Narzisstische, Borderline sowie Antisoziale Persönlichkeitsstörung. Im Rahmen des Screening-Fragebogens erhalten die Probanden alle Leitfragen des späteren Interviews zur Beantwortung. Die Probanden sollen beantworten, ob die genannten Empfindungen, Gefühle, Einstellungen und Verhaltensweisen in deren bisherigen Leben, das heißt in den letzten fünf bis zehn Jahren, be-

schreiben, wie die Probanden sich im Allgemeinen sehen. Es stehen nur die Antwortmöglichkeiten Ja oder Nein zur Auswahl. Im folgenden Interview-Teil werden all diejenigen Kapitel durchgearbeitet, bei denen die Anzahl der Ja-Antworten im Screening-Fragebogen eine Diagnose möglich macht. Anhand der Screening-Fragen sowie vertiefender Fragen prüft der Interviewer das Vorliegen der einzelnen Kriterien und damit der jeweiligen Persönlichkeitsstörung. Nach vorheriger Beantwortung des Screening-Fragebogens dauert der Interview-Teil etwa 30 Minuten. Die Probanden der vorliegenden Studie wurden mit beiden Stufen des SKID-II auf Persönlichkeitsstörungen untersucht.

Obwohl im Mai 2013, also kurz vor Beginn der Datenerhebung, die englischsprachige Neuauflage des Diagnostischen und Statistischen Manuals Psychischer Störungen das DSM-V (?) und im Dezember 2014 auch die deutsche Übersetzung der Neuauflage (?) publiziert wurden, wird in der vorliegenden Arbeit mit den Diagnosen nach DSM-IV gearbeitet. Ein strukturiertes Interview für DSM-V ist erst seit Ende 2014 und nur auf Englisch verfügbar und kam damit für die Mitte 2013 beginnende Datenerhebung nicht in Frage.

### **Anamnesebogen**

Mittels eines ausführlichen Anamnesebogens erfragt der Interviewer für jeden Probanden soziodemographische Daten und Daten zur Krankheitsgeschichte. Der Anamnesebogen erfasst Informationen zu Geburtsdatum, Geschlecht, Schul- und Berufsabschluss, Bestreitung des Lebensunterhalts, aktueller Arbeitssituation und Beruf sowie aktueller Lebenssituation und sozialen Kontakten. Außerdem werden krankheitsgeschichtliche Daten wie der Beginn der Behandlung einer psychischen Erkrankung, Anzahl und Zeitpunkte von durchgeführten Suizidversuchen, aktuelle psychopharmakologische und andere Medikation, Nikotin-Konsum, Zyklus sowie eine Familienanamnese erhoben. Die Beantwortung des Anamnesebogens dauert etwa 15 Minuten.

### **Kombiinterview aus Hamilton Depression Rating Scale und Inventar Depressiver Symptome (HAMD-IDS-Kombi)**

Zur Erfassung der aktuellen Stärke der depressiven Symptomatik in den vergangenen sieben Tagen wurde bei allen Teilnehmern aus den Patientengruppen eine Fremdbeurteilung durch geschulte Interviewer anhand eines Kombiinterviews aus der Hamilton Depression Rating Scale (HAMD-17 Hamilton, 1960) und dem Inventar Depressiver Symptome (IDS-C; Hautzinger & Bailer, 1994) durchgeführt. Die HAMD ist ein Fremdbeurteilungsinstrument mit 17 Items zur Einteilung der Schwere depressiver Symptome wie zum Beispiel depressive Verstimmung, Schuldgefühle oder psychomotorische Verlangsamung auf einer Skala von 0 (nicht vorhanden) bis 2 oder 4 (schwere Ausprägung). Es ergibt sich eine mögliche Gesamtsumme von 0 bis 52 Punkten. Werte von 0 bis 7 sprechen gegen das Vorliegen einer Depression. Werte von 8 bis 13 für eine leichte, von 14 bis 19 für eine mittelschwere, von 20 bis 25 für eine schwere und von 26 bis 52 für eine sehr schwere Depression. Das IDS-C ermöglicht ebenfalls eine Fremdbeurteilung der Schwere depressiver Symptome anhand einer Skala von 0 (nicht vorhanden) bis 3 (schwere Ausprägung) und enthält

30 verschiedene Items die zu einem Gesamtwert zwischen 0 und 84 addiert werden. Werte von 0 bis 13 geben keine Hinweise auf das Vorliegen einer Depression. Bei Werten von 14 bis 25 kann von einer leichten Depression ausgegangen werden, von 26 bis 38 von einer mittelschweren, von 39 bis 48 von einer schweren und von 49 bis 84 von einer sehr schweren. Rush et al. (1996) berichten eine ausgezeichnete interne Konsistenz mit Cronbachs  $\alpha = 0.94$  für das IDS-C eine gute interne Konsistenz mit Cronbachs  $\alpha = 0.89$  für die HAMD-17. Die Items des IDS-C decken sich zum Teil mit denen der HAMD-17, beide enthalten jedoch auch zusätzliche Items wie die Krankheitseinsicht in der HAMD-17 oder die zwischenmenschliche Empfindlichkeit im IDS-C. Um beide Skalen gleichzeitig erheben zu können, wurde ein strukturiertes Interview zur Kombination der beiden Verfahren von Kobak et al. (2007) entwickelt, welches sich an den strukturierten Interview-Richtlinien für die Hamilton Depression Rating Scale von Williams (1988) orientiert. Die Dauer des Kombiinterviews liegt bei ungefähr 25 Minuten.

### **Beck-Depressions-Inventar (BDI)**

Als Selbstbeurteilungsmaß der Schwere der depressiven Symptomatik wurde das Beck-Depressions-Inventar (BDI; Hautzinger et al., 1994) von allen Teilnehmern ausgefüllt. Es enthält 21 Items zu den verschiedenen Symptomen einer Depression. Für jedes Symptom sind vier verschiedene Auswahlmöglichkeiten zur Beschreibung des Zustands in der letzten Woche gegeben die eine Einschätzung des Schweregrads eines jeden Symptoms von 0 bis 3 ermöglichen. Für die depressive Verstimmung lauten die Möglichkeiten beispielsweise wie folgt: 0 = Ich bin nicht traurig; 1 = Ich bin traurig; 2 = Ich bin die ganze Zeit traurig und komme nicht davon los; 3 = Ich bin so traurig oder unglücklich, dass ich es kaum ertrage. So ergibt sich eine mögliche Gesamtsumme im Bereich von 0 bis 63 Punkten. Bei Werten von 0 bis 8 kann davon ausgegangen werden, dass keine Depression vorliegt. Werte von 9 bis 13 sprechen für eine minimale, von 14 bis 19 für eine leichte, von 20 bis 28 für eine mittelschwere und von 29 bis 63 für eine schwere Depression. Hautzinger (1991) berichtet eine gute interne Konsistenz, gute Korrelationen zu anderen Selbstbeurteilungsmaßen für Depression sowie eine hohe Änderungssensitivität. Die Bearbeitungsdauer liegt bei etwa zehn Minuten.

### **Young Mania Rating Scale (YMRS)**

Bei allen Teilnehmern in den Patientengruppen wurde eine Fremdbeurteilung der manischen Symptome in der letzten Woche anhand der Young Mania Rating Scale (YMRS; Young et al., 1978) durchgeführt. Es wird die Schwere elf verschiedener Symptome wie beispielsweise gehobene Stimmung oder gesteigerte motorische Aktivität auf einer Skala von 0 (nicht vorhanden) bis 4 oder 8 (schwere Ausprägung) erfasst. Die Gesamtsumme kann zwischen 0 und 60 Punkten liegen. Werte ab 13 weisen auf eine Hypomanie oder Manie hin. Die Durchführungsdauer beträgt etwa 15 Minuten. Mit einem Cronbachs Alpha von 0.74 weist die deutsche Version des YMRS eine akzeptable interne Konsistenz auf und kann insgesamt als valides, reliables und änderungssensitives Verfahren zur Erfassung manischer Symptome angesehen werden (Muehlbacher et al., 2011).

### **Manie-Selbstbeurteilungsskala (MSS)**

Die Manie-Selbstbeurteilungsskala (MSS; Krüger et al., 1997) wurde zur Einschätzung der manischen Symptome aller Teilnehmer der vorliegenden Studie genutzt. Es handelt sich um einen Fragebogen mit 48 Aussagen, für die angekreuzt werden soll, ob sie im Verlauf der vergangenen sieben Tage auf die Probanden zutrafen. Die Items decken manische und hypomanische Symptome ab, wie zum Beispiel „Ich konnte nicht ruhig sitzen bleiben.“ oder „Ich habe Musik lauter gehört als sonst.“, Antwortmöglichkeiten sind „Ja“ oder „Nein“. Zur Erfassung der Gesamtsumme werden alle Ja-Antworten addiert, es ergibt sich eine Gesamtsumme im Bereich von 0 bis 48. Werte ab 14 weisen auf eine aktuelle manische Episode hin. Bräunig et al. (1996) berichten eine gute interne Konsistenz ( $\alpha = 0.94$ ), Änderungssensitivität, Reliabilität und Validität. Die Bearbeitung der MSS dauert etwa zehn Minuten.

### **4.2.2 Neuropsychologie**

Im Rahmen der Diagnostiksituation wurde mit allen Teilnehmern der vorliegenden Studie eine neuropsychologische Untersuchung durchgeführt.

#### **Mehrfachwahl-Wortschatz-Intelligenztest (MWT-B)**

Beim Mehrfachwahl-Wortschatz-Intelligenztest (MWT-B; Lehrl, 1999) handelt es sich um einen Multiple Choice Test in Paper-Pencil-Form zur Schätzung der verbalen Intelligenz. Dieser ermöglicht eine valide Schätzung der prämorbidem allgemeinen Intelligenz und sollte in jeder Studie mit Patientenstichproben erfolgen (Lehrl et al., 1995). Jedes der 37 Items besteht aus einer Reihe von fünf Wörtern, von denen es sich bei nur einem um ein tatsächliches Wort handelt. Die Probanden werden aufgefordert, dasjenige Wort durchzustreichen, das ihnen bekannt vorkommt. Im Beispiel „Umma – Pamme – Nelle – Ampe – Amme“ wäre die richtige Antwort „Amme“. Die Schwierigkeit der Items soll sich kontinuierlich steigern. Die Durchführung des MWT-B dauert etwa fünf Minuten. Antretter et al. (2013) geben zu bedenken, dass der Test mit den wahrscheinlich veralteten Normen für die heutigen Bildungsstandards zu leicht sein könnte, was für die Interpretation der Ergebnisse beachtet werden sollte.

#### **Trail Making Test (TMT-A und -B)**

Der Trail Making Test (TMT Reitan, 1992) als Paper-Pencil-Version wurde in der vorliegenden Arbeit zur Erfassung der einfachen (TMT-A) und der geteilten (TMT-B) Aufmerksamkeit verwendet. Im Teil A müssen die Probanden die auf dem Blatt verstreuten Zahlen von 1 bis 25 in aufsteigender Reihenfolge mit einer Linie verbinden. Teil B enthält sowohl Zahlen von 1 bis 13 als auch Buchstaben von A bis L die abwechselnd in aufsteigender Reihenfolge mit einer Linie verbunden werden sollen. Für beide Testteile werden vorab zur Verdeutlichung der Instruktionen Beispiele zur Übung gegeben. Das Beispiel zu Teil B enthält die Zahlen-Buchstabenfolge „1 – A – 2 – B – 3 – C – 4 – D“. Bei der Testdurchführung wird die Zeit gestoppt, Fehler werden währenddessen korrigiert und führen so zu einer Verlängerung der gemessenen

Zeit und damit zu einem schlechteren Ergebnis. Anhand der Rohwerte in Sekunden werden die Ergebnisse für beide Teile in jeweils vier Kategorien (perfectly normal, normal, mild/moderately impaired und moderately/severely impaired) eingeteilt. Für die Bearbeitung beider Testteile werden etwa fünf Minuten benötigt.

### **Verbaler Lern- und Merkfähigkeitstest (VLMT)**

Der Verbale Lern- und Merkfähigkeitstest (VLMT) von Helmstädter et al. (2001) kam in der vorliegenden Arbeit in der Version A zur Anwendung. Eine Wortliste mit 15 Wörtern (Lernliste) wird dem Probanden fünf mal hintereinander vorgelesen und jeweils anschließend die vom Probanden richtig erinnerten Wörter, falsch positive Wörter sowie Perseverationen notiert. Daraufhin wird eine ebenfalls 15 Wörter lange Interferenzliste einmal vorgelesen und die richtig erinnerten Wörter erfasst. Anschließend wird der Proband gebeten ohne weitere Präsentation die Wörter der Lernliste zu nennen. Hierbei werden die Anzahl der richtigen Wörter aber auch die Anzahl der Interferenzen, also der Wörter aus der Interferenzliste, erfasst. Bis hierhin dauert die Durchführung etwa 15 Minuten. Nach einer Pause von mindestens 20 und maximal 30 Minuten werden der verzögerte Abruf sowie die Wiedererkennung geprüft: Ohne eine nochmalige Präsentation der Wörter, soll der Proband die Wörter der ersten Liste erinnern, es werden sowohl richtige Antworten als auch Interferenzen erfasst. Anschließend wird die Wiedererkennung geprüft indem eine Liste von Wörtern, bestehend aus denen der Lernliste, denen der Interferenzliste und völlig neuen Wörtern, vorgelesen wird. Der Proband aufgefordert wird, mit „Ja“ zu antworten, falls es sich um ein Wort der Lernliste handelt. Hierbei werden die Anzahl der richtig wiedererkannten Wörter, die Anzahl der Fehler aus der Interferenzliste und die Anzahl der Fehler mit neuen Wörtern notiert. Der Teil nach der Pause dauert etwa 10 Minuten. Es ergibt sich eine Vielzahl von Kennwerten, relevant sind vor allem die folgenden: Die Gesamtlernleistung bezeichnet die Summe richtiger Antworten in den Durchgängen 1 bis 5. Die Abrufleistung der Lernliste nach Interferenz ergibt sich aus der Anzahl richtiger Antworten bei Erinnerung der Lernliste nach der Präsentation der Interferenzliste. Der Verlust nach Interferenz berechnet sich aus der Differenz zwischen der Anzahl richtiger Antworten in Durchgang fünf (also direkt vor der Interferenz) und der Abrufleistung der Lernliste nach Interferenz. Der Verlust nach zeitlicher Verzögerung bezeichnet die Differenz zwischen der Anzahl richtiger Antworten in Durchgang fünf und der Anzahl richtiger Antworten beim freien Erinnern nach der Pause. Die korrigierte Wiedererkennensleistung ergibt sich aus der Anzahl korrekt wiedererkannter Wörter abzüglich der Anzahl der falsch positiv wiedererkannten Wörter.

### **4.2.3 Empathie und Social Functioning**

#### **Empathy Quotient (EQ-S und -F)**

Der von Baron-Cohen & Wheelwright (2004) entwickelte und validierte (Lawrence et al., 2004) Selbstbeurteilungsfragebogen Empathy Quotient (EQ) wurde wie in Kapitel 2 beschrieben ins Deutsche übersetzt und validiert. Es handelt sich um einen Fragebogen mit 60 Aussagen zur Empathie sowie Filler-Items, die der Pro-

band auf einer vierstufigen Likert-Skala von „stimme voll und ganz zu“ bis „stimme überhaupt nicht zu“ einschätzen soll. Weiterhin wurde im Rahmen der vorliegenden Arbeit eine Fremdbeurteilungsversion (EQ-F) analog zur Selbstbeurteilung (EQ-S) entwickelt, in der Items wie „Leute sagen mir oft, ich sei sehr unberechenbar.“ zu „Leute sagen ihm/ihr/[Name] oft, er/sie sei sehr unberechenbar.“ umformuliert wurden. Es kommen sowohl die ursprüngliche Auswertungsvariante bei der eine Summe aller Empathie-Items berechnet wird sowie eine Auswertungsvariante bei der sowohl eine Summe als auch Unterskalen berechnet werden können zum Einsatz. Für den EQ-S werden die Unterskalen Faktor 1 kognitive Empathie, Faktor 2 emotionale Reaktivität und Faktor 3 empathische Betroffenheit berechnet. Der EQ-F enthält die folgenden Unterskalen: Faktor 1 soziales Geschick, Faktor 2 kognitive Empathie, Faktor 3 emotionale Reaktivität und Faktor 4 empathische Betroffenheit. In der Auswertung der Experimentalstudie werden zumeist die Summenwerte nach ursprünglicher Auswertungsvariante verwendet, da auch das englische Original vorwiegend in der vollständigen Variante mit allen 60 Items zur Anwendung kommt. Für manche Fragestellungen werden zusätzlich die ermittelten Unterskalen in Betracht gezogen. Zudem finden sich im Anhang E einzelne Ergebnisse für den Gesamtwert nach PCA-Auswertung. Sowohl EQ-S ( $\alpha = 0.88$  für die ursprüngliche Auswertung inklusive aller Items und  $\alpha = 0.91$  nach Ausschluss von Items mit zu geringen Trennschärfen oder zu geringen oder zu hohen Schwierigkeiten) als auch EQ-F ( $\alpha = 0.91$  ursprüngliche Auswertung,  $\alpha = 0.91$  nach Ausschluss) weisen eine sehr gute interne Konsistenz auf. Darüber hinaus ähnelt die Faktorstruktur der deutschen Übersetzung der des englischen Originals sehr.

### **Fragebogen zur Sozialen Unterstützung (F-SozU)**

In Teil A des Fragebogens zur Sozialen Unterstützung (F-SozU Fydrich et al., 2007) werden Aussagen zu sozialer Unterstützung auf einer fünfstufigen Likert-Skala von „trifft nicht zu“ bis „trifft genau zu“ eingeschätzt. In der vorliegenden Arbeit wurde die Standardform mit 54 Items verwendet. „Bei Bedarf kann ich mir Werkzeug oder Lebensmittel ausleihen“ ist ein Beispielitem für die Unterskala praktische Unterstützung. Des Weiteren werden die Unterskalen emotionale Unterstützung, soziale Integration, wahrgenommene soziale Unterstützung (Gesamtwert der drei Unterstützungsskalen), Soziale Belastung, sowie die Zusatzskalen Reziprozität sozialer Unterstützung, Verfügbarkeit einer Vertrauensperson und die Zufriedenheit mit sozialer Unterstützung gemessen. Teil B des F-SozU ermöglicht die Erfassung genauer Anzahlen unterstützender und belastender Personen aus verschiedenen Bereichen (z.B. Familie). Die somit mögliche eher qualitative Auswertung steht für die vorliegende Arbeit eher im Hintergrund. Einzig die durchschnittliche Anzahl unterstützender und belastender Personen wird ausgewertet. Fydrich et al. (1999) berichten gute bis sehr gute innere Konsistenzen für die Hauptskalen ( $\alpha = 0.81$  bis  $\alpha = 0.93$ ). Die Beantwortung beider Teile dauert etwa 20 Minuten.

### **Inventar zur Erfassung interpersonaler Probleme (IIP-C)**

Das Inventar zur Erfassung interpersonaler Probleme (IIP-C Horowitz et al., 2000) kam in der vorliegenden Arbeit in der aktuell gebräuchlichen Kurzversion mit 64

Items zur Anwendung. Es handelt sich um einen aus zwei Teilen bestehenden Fragebogen, dessen Items auf einer fünfstufigen Likert-Skala von „nicht“ bis „sehr“ dahingehend beurteilt werden sollen, ob der Proband darunter leidet. Teil 1 behandelt Aspekte die im Umgang mit anderen schwierig sein können, wie zum Beispiel „Es fällt mir schwer anderen Menschen zu vertrauen“ oder „Es fällt mir schwer mich Gruppen anzuschließen“. Teil 2 enthält Aspekte, die man im Übermaß tun kann, wie zum Beispiel „Ich bin gegenüber anderen zu aggressiv“ oder „Ich vertraue anderen zu leicht“. Es ergibt sich ein Gesamtwert für die gemittelten interpersonalen Probleme sowie Werte für die folgenden acht Unterskalen: zu autokratisch/dominant, zu streitsüchtig/konkurrierend, zu abweisend/kalt, zu introvertiert/sozial vermeidend, zu selbstunsicher/unterwürfig, zu ausnutzbar/nachgiebig, zu fürsorglich/freundlich, zu expressiv/aufdringlich. Locke (2005) berichtet für die Unterskalen interne Konsistenzen zwischen  $\alpha = 0.66$  und  $\alpha = 0.87$ .

### Face-Task der Cambridge Mindreading Face-Voice Battery (CAM)

Die Cambridge Mindreading (CAM) Face-Voice Battery von Golan et al. (2006) erfasst die Erkennung komplexer Emotionen über kurze Sprachaufnahmen oder Videos am PC. In der vorliegenden Arbeit kam ausschließlich die sogenannte Face-Task, das heißt die Messung der Emotionserkennung anhand kurzer Videos, zur Anwendung. Hierbei sehen die Probanden Filmaufnahmen in denen Personen beider Geschlechts, verschiedener Altersklassen und Ethnien komplexe Emotionen zeigen. Im Anschluss werden vier Antwortmöglichkeiten präsentiert, von denen die Probanden eine auswählen sollen. Die Antwort erfolgt mittels eines von SR Research modifizierten Microsoft USB Gamepads, welches eine genaue Erfassung der Reaktionszeiten ermöglicht. Die gezeigten Emotionen decken verschiedene Valenzen (negativ vs. neutral vs. positiv) und verschiedene Emotionsgruppen ab (zum Beispiel traurig, interessiert oder glücklich). Beispiele für solche komplexen Emotionen sind ernst, einfühlsam oder beruhigt. Zudem werden zum Abschluss jeden Trials die empfundene Sicherheit und Schwierigkeit der Emotionserkennung erfasst. In Abbildung 4.2 ist ein Trial der CAM grafisch dargestellt. Die Face-Task der CAM wurde in einer noch nicht publizierten Untersuchung von Wolkenstein et al. (ca. 2012) ins Deutsche übersetzt und validiert. Von den 50 ursprünglichen Items wurden fünf als Beispiel-Items ausgewählt, deren Itemschwierigkeiten und Trennschärfen sich nicht als optimal erwiesen hatten. Es verblieben also 45 Items die für die vorliegende Arbeit in zwei gematchte Gruppen von jeweils 22 (Teil A) beziehungsweise 23 (Teil B) Videos aufgeteilt wurden. Um eine gleichzeitige Messung der Augenbewegungen und Muskelaktivität der Probanden zu ermöglichen, wurde die CAM mit dem Experimental Builder von SR Research neu programmiert. Hierbei wurden fünf Blöcke mit jeweils randomisierter Präsentation der jeweiligen Trials eingeführt. In der vorliegenden Arbeit wird nur der erste der fünf Blöcke in die Auswertung einbezogen, da keine Analyse der Blickbewegungen und Muskelaktivitäten erfolgt. Die Durchführung eines Teils der CAM inklusive Messung der Blickbewegungen und Muskelaktivität sowie CALC (siehe 4.2.4) dauert etwa 90 Minuten. Ein Block, das heißt ein Durchgang aller Items eines Teils, dauert etwa 10 bis 15 Minuten. Die Durchführung der kompletten CAM ohne Messung der Blickbewegungen und Muskelaktivitäten benötigt etwa 25 Minuten.



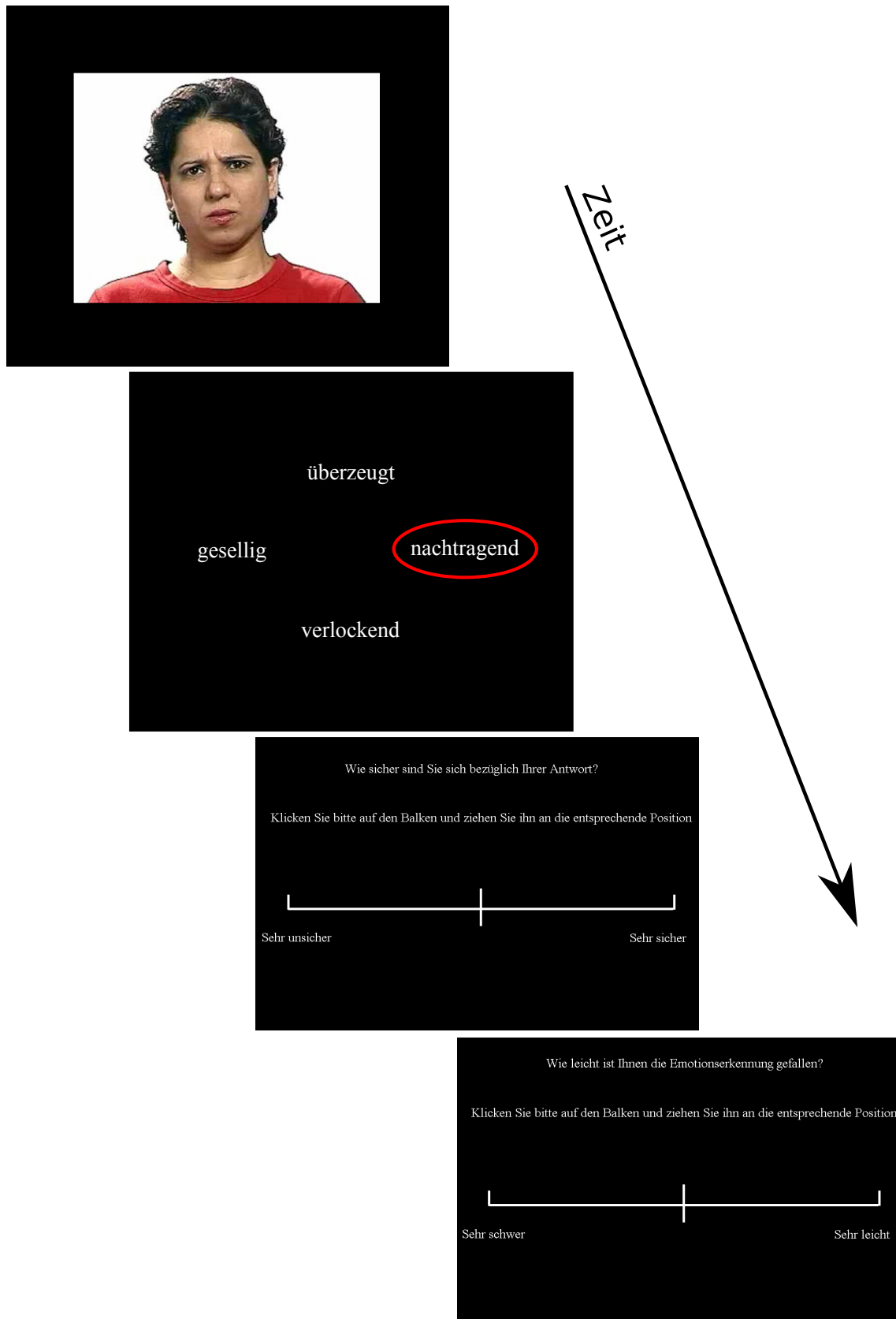


Abbildung 4.2: Ablauf eines Trials der CAM Face-Task

### **Movie for the Assessment of Social Cognition (MASC)**

Der Filmtest Movie for the Assessment of Social Cognition (MASC; Dziobek et al., 2006) ist ein Verfahren zur Messung emotionaler und kognitiver Empathie mit sehr hoher ökologischer Validität. Hierbei sehen die Probanden einen kurzen Film über eine Gruppe von Personen die einen Abend gemeinsam verbringt. Der Film wird immer wieder angehalten und die Probanden werden gebeten Fragen über die mentalen Zustände der Filmcharaktere zu beantworten. Hierbei werden jeweils vier Antwortmöglichkeiten dargeboten: eine richtige Antwort, eine Antwort ganz ohne ToM, eine mit zu wenig ToM und eine mit zu viel ToM. Ein Beispiel für eine Frage ist „Was empfindet Klaus?“, die Antwortmöglichkeiten sind hier „a) Die Wohnung ist schön“ (ohne ToM), „b) Er hat Angst vor dem Hund“ (richtig), „c) Er fühlt sich unwohl mit Sandra alleine zu sein“ (zu viel ToM) und „d) Er ist überrascht, dass sie einen Hund hat“ (zu wenig ToM). Der MASC enthält 45 Items und wird mit dem Versuchsleiter gemeinsam am PC bearbeitet. Der Versuchsleiter startet die jeweiligen Videos, liest die eingeblendeten Fragen vor und notiert die Antworten der Probanden. Die Durchführung des MASC in der Gesamtversion dauert etwa 45 Minuten. Um eine Messung sowohl in der Sitzung mit als auch ohne Stimmungsinduktion zu ermöglichen, wurde der MASC für die vorliegende Arbeit in zwei Teile aufgeteilt, für die jeweils etwa 30 Minuten benötigt werden.

### **4.2.4 Weitere Messinstrumente**

#### **Kontrollaufgabe mit mathematischen Gleichungen (CALC)**

Um emotionsspezifische Effekte bei der CAM zu identifizieren, wurde eine Kontrollaufgabe mit mathematischen Gleichungen entwickelt und gewissermaßen als zusätzlicher Block direkt im Anschluss an die CAM dargeboten. Zur Programmierung wurde entsprechend der Experimental Builder von SR Research genutzt. Die Trials der CALC laufen ab wie folgt: Zunächst sehen die Probanden zwei Rechenaufgaben, eine auf der linken Seite des Bildschirms und eine auf der rechten Seite des Bildschirms. Die Aufgabe der Teilnehmer ist es, die Rechenaufgabe mit dem größeren Ergebnis auszuwählen. Im Anschluss werden die Probanden gebeten die Schwierigkeit und die Sicherheit ihrer Entscheidung einzuschätzen. Damit ist der Ablauf analog zum Ablauf der Trials der CAM. Ein Beispiel für zwei Rechenaufgaben ist „ $11 + 5$  |  $21 - 3$ “, hierbei ist das Ergebnis der rechten Aufgabe größer. Wie bei der CAM antworten die Probanden anhand des von SR Research modifizierten Microsoft USB Gamepads um die Reaktionszeiten messen zu können. Sowohl für Teil A als auch B der CAM existiert jeweils ein Set von 20 Trials (CALC Teil A und CALC Teil B). Ein Teil der CALC dauert etwa zehn Minuten.

#### **Positive and Negative Affect Schedule (PANAS)**

Zur Messung der aktuellen positiven und negativen Stimmung wurde eine Kurzform der Positive and Negative Affect Schedule (PANAS; Watson et al., 1988) entwickelt. Die Originalform der PANAS enthält 20 Adjektive, zur Hälfte negative Emotionen, zur Hälfte positive, für die die Probanden einschätzen sollen, wie intensiv sie diejenige Emotion im Moment erleben. Die Einschätzung erfolgt auf einer fünfstufigen

Likert-Skala von „gar nicht“ bis „äußerst“. In der vorliegenden Arbeit werden lediglich fünf der Gefühle erfragt, es handelt sich um die Folgenden: bekümmert, freudig erregt, schlecht, zufrieden, traurig. Damit sind in der Kurzform drei negative und zwei positive Adjektive enthalten und die Bearbeitungsdauer verkürzt sich auf nur ein bis zwei Minuten.

### 4.3 Experimentelle Manipulation

Als experimentelle Manipulation wird eine Induktion negativer Stimmung durch Musik durchgeführt. Die Probanden hören das Stück *Åses Tod* aus der *Peer-Gynt-Suite* Nr. 1, Op. 46 von Edvard Grieg in verlangsamtem Tempo. Anstatt der üblichen vier bis sechs Minuten dauert die hier verwendete Version zehn Minuten. Vor Beginn der Musik erhalten die Probanden die folgende Instruktion vom Versuchsleiter:

Sie werden jetzt ein zehnmütiges Musikstück hören. Währenddessen sollen Sie sich auf ein trauriges Ereignis in Ihrem Leben konzentrieren, das mindestens zwei Jahre zurück liegt. Bitte überlegen Sie sich jetzt – bevor ich das Musikstück starte – an welches Ereignis Sie sich erinnern möchten. Lassen Sie sich dann von der Musik tragen. Sagen Sie mir bitte Bescheid, wenn Sie bereit sind!

Auch auf dem Bildschirm erhalten die Probanden folgende schriftliche Instruktion direkt vor Beginn des Musikstücks:

Hören Sie bitte in den nächsten 10 Minuten das folgende Musikstück, konzentrieren Sie sich dabei auf ein trauriges Ereignis in Ihrem Leben das mindestens zwei Jahre zurück liegt. Lassen Sie sich von der Musik tragen.

Während die Musik läuft, ist der Bildschirm schwarz. Um die experimentelle Manipulation der Stimmung im Verlauf des Experimentes auffrischen zu können, existiert eine gekürzte Version des Musikstücks die nur fünf Minuten dauert (MI 2). Das Tempo in dem *Åses Tod* gespielt wird, ist hierbei ebenso verlangsamt wie bei der ersten Stimmungsinduktion (MI 1).

### 4.4 Untersuchungsablauf

Folgend wird der Ablauf der Untersuchung detailliert beschrieben. Im Anhang D finden sich grafische Darstellungen sowohl des Ablaufs der Untersuchung insgesamt als auch des Ablaufs der diagnostischen Sitzung und der experimentellen Sitzungen im Detail.

#### 4.4.1 Ablauf Telefoninterview

Als erster Kontakt mit den potentiellen Probanden erfolgt ein etwa fünfzehnminütiges Telefoninterview. Im Rahmen des Telefoninterviews werden demographische

Daten wie Alter, Geschlecht und Muttersprache erfasst. Außerdem erfolgt eine erste grobe Überprüfung der Ein- und Ausschlusskriterien anhand des Screening-Teils des SKID-I und einiger Fragen zu affektiven Störungen. Weiterhin werden relevante Informationen zur aktuellen Medikation und zur Durchführbarkeit des Eyetrackings sowie der Elektromyographie gesammelt. Für den Probanden besteht die Möglichkeit zur Klärung offener Fragen, zum Beispiel auch organisatorischer Art. Wenn sowohl Versuchsleiter als auch Proband im Rahmen des Telefoninterviews die Entscheidung zu einer weiteren Teilnahme treffen, erfolgt die Terminvereinbarung sowie der Versand einiger Unterlagen per Post oder per Email an den Probanden. Diese Unterlagen beinhalten ein Informationsblatt zur Untersuchung, die Einverständniserklärung, die Erklärung zum Datenschutz, eine Wegbeschreibung sowie den SKID-II-Screening-Fragebogen. Die Bearbeitung durch den Probanden dauert etwa 30 Minuten.

#### 4.4.2 Ablauf Diagnostik t<sub>0</sub>

Zu Beginn der diagnostischen Sitzung (t<sub>0</sub>) werden gegebenenfalls offene Fragen des Probanden bezüglich der Untersuchung geklärt und die Einverständniserklärung sowie die Erklärung zum Datenschutz eingeholt. Anschließend erfolgt die allgemeine Diagnostik psychischer Erkrankungen anhand des Anamnesefragebogens und des SKID-I. Probanden aus den Patientengruppen werden mit dem HAMD-IDS-Kombiinterview (alle Patientengruppen) und dem YMRS-Interview (bipolare Patienten) auf ihre aktuelle Symptombelastung untersucht. Alle Teilnehmer füllen die Selbstbeurteilungsbögen BDI und MSS aus. Währenddessen wertet der Versuchsleiter den vorab ausgefüllten Screening-Fragebogen für SKID-II aus, um auffällige Sektionen im Anschluss mit dem Probanden durchzuführen. Sobald an irgendeiner Stelle der diagnostischen Sitzung klar wird, dass die Ein- oder Ausschlusskriterien nicht erfüllt sind, wird die Untersuchung des Probanden beendet. Bei vorzeitiger Beendigung durch den Versuchsleiter oder den Probanden erhält dieser die Aufwandsentschädigung für die bis dahin benötigte Zeit. Nach Prüfung aller Ein- und Ausschlusskriterien wird die Gruppenzuordnung festgelegt und es erfolgt die neuropsychologische Testung des Probanden (siehe unten). Insgesamt dauert t<sub>0</sub> etwa vier Stunden, wobei die benötigte Zeit stark variiert, je nachdem wie viele psychische Erkrankungen beim Probanden vorliegen.

#### Ablauf Neuropsychologie

Die neuropsychologische Testung beginnt mit Teil 1 des VLMT, daraufhin werden MWT-B und TMT-A sowie -B durchgeführt. Wenn zu diesem Zeitpunkt die erforderliche Wartezeit erfüllt ist, erfolgt Teil 2 des VLMT. Sollte die Wartezeit noch nicht erfüllt sein, werden Fragebögen aus dem Fragebogenpaket (siehe 4.4.3) vorgezogen, bis die Wartezeit erfüllt ist.

#### 4.4.3 Ablauf Fragebogenpaket

Nach der diagnostischen Sitzung erhalten die Probanden ein Fragebogenpaket mit verschiedenen Fragebögen zu Empathie und Social Functioning mit nach Hause. Für

die Bearbeitung des Fragebogenpaketes benötigen die Probanden etwa 30 Minuten. Darin enthalten sind der IIP-C, der F-SozU und der EQ-S. Weiterhin sind der EQ-F, ein frankierter Rückumschlag sowie ein Anschreiben für die Fremdbeurteilung beigelegt, die die Probanden an eine vertraute Person weiter geben sollen. Für andere Fragestellungen die im Rahmen der vorliegenden Arbeit nicht betrachtet werden, bearbeiten die Probanden auch den Cognitive Emotion Regulation Questionnaire sowie den Emotion Regulation Questionnaire.

#### 4.4.4 Ablauf experimentelle Sitzung t1 und t2

Die erste experimentelle Sitzung (t1) findet im Abstand von maximal einer Woche nach t0 statt, die zweite experimentelle Sitzung (t2) im Abstand von maximal einer Woche nach t1. Der Ablauf von t1 und t2 ist identisch bis auf die experimentelle Manipulation: Die Stimmungsinduktion erfolgt randomisiert entweder in t1 oder t2, so dass jeder Proband eine experimentelle Sitzung mit (MI) und eine ohne (NMI) Stimmungsinduktion durchläuft. Bei t1 werden zunächst die Fragebögen des Fragebogenpaketes eingesammelt und auf Vollständigkeit kontrolliert. Zu Beginn der experimentellen Sitzungen werden dann die Elektroden für die – im Rahmen der vorliegenden Arbeit nicht beachtete – Elektromyographie aufgebracht. Anschließend wird der Proband in der abgeschirmten Kabine vor dem PC positioniert und die aktuelle Stimmung mit dem PANAS erhoben (Zeitpunkt PräMI1). In der MI-Sitzung erfolgt nun die erste Stimmungsinduktion (MI 1) und anschließend eine erneute Messung der aktuellen Stimmung (Zeitpunkt PostMI1). Darauf folgt die CAM inklusive Elektromyographie, Eyetracking und Videoüberwachung. Teil A und B der CAM werden randomisiert in t1 und t2 dargeboten, daraus ergeben sich vier mögliche Reihenfolgen der experimentellen Sitzungen:

- 1) t1: MI und Teil A der CAM + t2: NMI und Teil B der CAM
- 2) t1: MI und Teil B der CAM + t2: NMI und Teil A der CAM
- 3) t1: NMI und Teil A der CAM + t2: MI und Teil B der CAM
- 4) t1: NMI und Teil B der CAM + t2: MI und Teil A der CAM

Zum Abschluss der CAM bearbeiten die Probanden die Kontrollaufgabe CALC. Daraufhin wird die aktuelle Stimmung (Zeitpunkt PräMI2) erhoben. Nach der Entfernung der Elektroden wird in der experimentellen Sitzung mit Stimmungsinduktion eine Auffrischung derselben (MI 2) sowie eine erneute Messung der aktuellen Stimmung (Zeitpunkt PostMI2) durchgeführt. Anschließend bearbeiten die Probanden den MASC. Aufgrund des durchgehenden Handlungsablaufs kann hier nicht randomisiert werden, Teil 1 erfolgt in t1 und Teil 2 in t2. Nach dem MASC wird die aktuelle Stimmung abschließend gemessen (Zeitpunkt Abschluss). In der zweiten experimentellen Sitzung werden weiterhin Vermutungen zum Nutzen der Elektroden erfragt und organisatorische Details zur Nacherhebung t3 geklärt. Die experimentellen Sitzungen dauern jeweils etwa 2:40 Stunden.

### 4.4.5 Ablauf Nacherhebung t3

Neun Monate nach t2 werden alle Probanden der Patientengruppen zu einer Nacherhebung (t3) eingeladen, in deren Rahmen der Anamnesefragebogen aktualisiert, die Maße der aktuellen Symptombelastung (HAMD-IDS, YMRS, BDI und MSS) erneut durchgeführt und ein Interview über den Verlauf der psychischen Erkrankung seit t2 erhoben werden. Die Nacherhebung dauert etwa anderthalb Stunden. Die Ergebnisse werden im Rahmen der vorliegenden Arbeit nicht ausgewertet.

## 4.5 Studiendesign

Bei der vorliegenden Studie handelt es sich um ein Experiment im 5 x 2 Design mit Messwiederholung. Es werden die Einflüsse des Between-Subject-Faktors Gruppe (Patienten mit akuter unipolarer Depression vs. Patienten mit remittierter unipolarer Depression vs. gesunde Kontrollprobanden für die unipolaren Patienten vs. Patienten mit remittierter bipolarer Störung vs. gesunde Kontrollprobanden für die bipolaren Patienten) und des Within-Subject-Faktors Messwiederholung (Induktion negativer Stimmung vs. keine Stimmungsinduktion) auf verschiedene Aspekte der Empathie erhoben. Als unabhängige Variablen neben der Gruppenzugehörigkeit werden zudem die vorab gemessene Selbst- und Fremdbeurteilung der Empathie sowie die aktuelle Symptombelastung in die Auswertung mit einbezogen. Abhängige Variablen sind die mit der CAM gemessene Emotionserkennung sowie die im Rahmen des MASC erhobene emotionale oder kognitive Empathie.

## 4.6 Statistische Analysen

Zur Beantwortung der Fragestellungen wird generell die folgende Vorgehensweise angewandt: Zunächst erfolgt eine deskriptive Beschreibung der erhobenen Daten. Nach Prüfung der Vorannahmen werden die statistischen Tests zur Klärung der Hypothesen durchgeführt. Im Anschluss werden die Ergebnisse der statistischen Tests anhand von Post-Hoc-Analysen genauer analysiert sowie auf mögliche Konfundierungen untersucht.

Die statistischen Analysen der vorliegenden Arbeit werden zum größten Teil mit IBM SPSS Statistics in der Version 23.0.0.0 durchgeführt. Die Berechnung einiger Effektstärken erfolgt zusätzlich mit G\*Power 3.1.9.2 (Faul et al., 2007). Sofern nicht explizit anders beschrieben, wird das Signifikanzniveau auf 5% (zweiseitig) festgelegt. Wenn von einer Normalverteilung der Daten ausgegangen werden kann, wird der Korrelationskoeffizient Pearsons  $r$  (schwach ab  $\pm 0.1$ , mittel ab  $\pm 0.3$ , stark ab  $\pm 0.5$ ) berechnet. Falls die Normalverteilung anzuzweifeln ist, wird hingegen Spearmans  $\rho$  (schwach ab  $\pm 0.1$ , mittel ab  $\pm 0.3$ , stark ab  $\pm 0.5$ ) berichtet. Als Effektstärkemaß kommen je nach Anwendungsfeld Pearsons  $r$ , Cohens  $d$  (kleiner Effekt ab 0.2, mittelgroßer Effekt ab 0.5, starker Effekt ab 0.8) oder Cohens  $f^2$  (kleiner Effekt ab .02, mittlerer Effekt ab .15, großer Effekt ab .35) zum Einsatz.

Deskriptive Beschreibungen der vorliegenden Daten erfolgen anhand des arithmetischen Mittels, der Standardabweichung sowie falls passend der Minima und Maxima. Die Prüfung der Normalverteilung erfolgt anhand des Shapiro-Wilk-Tests. Un-

tersuchungen verschiedener Gruppen auf Unterschiede bezüglich bestimmter Merkmale (z.B. der Intelligenz oder soziodemographischer Daten) werden mit dem  $\chi^2$ -Homogenitätstest gemacht. Dieser Test untersucht, ob die verschiedenen Gruppen bezüglich des Merkmals aus der gleichen Grundgesamtheit gezogen wurden. Zur Untersuchung des Einflusses unabhängiger Variablen auf abhängige Variablen wird zumeist auf die klassische parametrische Varianzanalyse zurückgegriffen. Nonparametrische Alternativen fehlen für komplexe Fragestellungen wie einige der vorliegenden. Bei weniger komplexen Fragestellungen wird die nonparametrische Alternative verwendet. Weiterhin kann davon ausgegangen werden, dass die ANOVA ein Verfahren ist, welches gegenüber Verletzungen der Normalverteilung bei ausreichender Größe der Gruppen robust ist (Schmider et al., 2010). Bei ähnlichen Gruppengrößen wird der Levene-Test durchgeführt, Verletzungen der Varianzhomogenität jedoch als unproblematisch angesehen. Weist der Mauchly-Test auf eine Verletzung der Sphärizität hin, wird die Greenhouse-Geisser-Korrektur angewandt. Sofern im Rahmen von varianzanalytischen oder entsprechenden nonparametrischen Verfahren signifikante Haupteffekte gefunden werden, werden Post-Hoc-Tests angewandt um die Haupteffekte genauer differenzieren zu können. Als Post-Hoc-Tests kommen je nach Situation Bonferroni, Gabriel oder Games-Howell zur Verwendung. Zur Erklärung der Varianz bestimmter Merkmale wird eine multiple lineare Regression durchgeführt. Dabei wird ein Modell aufgestellt um die Varianz einer abhängigen Variable aus mehreren unabhängigen Variablen (Prädiktoren) erklären zu können. Die Anzahl der möglichen unabhängigen Variablen ist hierbei abhängig von der Größe des Datensatzes. Häufig wird angenommen, dass pro Prädiktor 10 bis 15 Teilnehmer vorhanden sein sollen. Danach ist der vorliegende Datensatz für Regressionsanalysen mit 12 bis 18 Prädiktoren geeignet. Restriktiver ist die Regel, für  $n$  Prädiktoren  $3 \times n^2$  Teilnehmer zu rekrutieren, damit sind hier Regressionsanalysen mit 8 Prädiktoren möglich.

#### 4.6.1 Fragestellung E1

Mögliche Gruppenunterschiede bezüglich des Gesamtwerts interpersoneller Probleme werden anhand des Kruskal-Wallis-Tests analysiert. Weiterhin wird im Rahmen einer multivariaten Varianzanalyse der Einfluss der UV Gruppenzugehörigkeit auf die beiden AVs soziale Belastung und wahrgenommene soziale Unterstützung, gemessen mit dem F-SozU, untersucht.

#### 4.6.2 Fragestellung E2

Zur Klärung der Frage, ob zwischen den verschiedenen Gruppen Unterschiede im Ausmaß der mit dem EQ gemessenen Empathie bestehen, wird der Einfluss der unabhängigen Variable (UV) Gruppenzugehörigkeit auf die beiden abhängigen Variablen (AV) Gesamtsumme des EQ-S und Gesamtsumme des EQ-F im Rahmen einer multivariaten Varianzanalyse untersucht. Post-Hoc werden mögliche Unterschiede zwischen Personen mit und ohne vorhandener Fremdbeurteilung anhand von  $\chi^2$ -Homogenitätstests analysiert und eine Konfundierung der Ergebnisse durch

diese Unterschiede diskutiert. Im Anschluss erfolgt eine regressionsanalytische Betrachtung der EQ-Varianz zur explorativen Suche nach möglichen Einflussfaktoren.

### 4.6.3 Fragestellung E3

Zur Klärung der Frage, ob eine Diskrepanz zwischen Selbst- und Fremdeinschätzung der Empathie besteht, wird zunächst die Korrelation zwischen EQ-S und EQ-F betrachtet. Anschließend wird anhand eines gepaarten t-Tests für die Summen des EQ-S und EQ-F analysiert, ob sich Selbst- und Fremdeinschätzung systematisch unterscheiden. Die Differenz zwischen EQ-S und EQ-F (EQ-Diff) kann als Maß für die Größe der Diskrepanz zwischen EQ-S und EQ-F herangezogen werden. Anhand einer ANOVA wird untersucht ob Gruppenunterschiede bezüglich dieser Diskrepanz bestehen.

### 4.6.4 Fragestellung E4

Die Frage nach möglichen Gruppenunterschieden bezüglich der behavioralen Messung der Empathie wird anhand von Varianzanalysen für die CAM als auch den MASC untersucht. Post hoc werden nicht nur die richtigen Antworten in beiden Verfahren sondern auch Reaktionszeit, Sicherheits- und Schwierigkeitseinschätzung in der CAM sowie verschiedene Fehlerarten und verschiedene Arten der Empathie getrennt varianzanalytisch betrachtet. Zur näheren Beleuchtung des Zusammenhangs zwischen Ergebnissen in der Fragebogenmessung und in der behavioralen Messung der Empathie werden Regressionsanalysen durchgeführt.

### 4.6.5 Fragestellung E5

Der mögliche Einfluss einer experimentell induzierten negativen Stimmung auf die behaviorale Messung der Empathie soll anhand von Varianzanalysen untersucht werden. Vorab wird ebenfalls mit varianzanalytischen Methoden überprüft, ob eine Induktion negativer Stimmung erfolgreich war. Im Rahmen einer multivariaten Varianzanalyse sollen Gruppenunterschiede bezüglich der Induzierbarkeit negativer Emotionen genauer beleuchtet werden. Mögliche Einflussfaktoren auf das Ausmaß der Induzierbarkeit werden durch eine Regressionsanalyse untersucht.

### 4.6.6 Fragestellung E6

Mögliche Zusammenhänge zwischen EQ, CAM und MASC und der mit dem MWT-B geschätzten prämorbidem Intelligenz sowie den mit dem IIP erfassten interpersonellen Problemen und den Werten sozialer Belastung und sozialer Unterstützung aus dem F-SozU werden anhand von Korrelationen genauer beleuchtet. Hierfür interessieren sowohl die Korrelationen in der Gesamtgruppe als auch in den einzelnen Teilnehmergruppen um eventuelle spezifische Effekte in den Untergruppen beobachten zu können.



# Kapitel 5

## Ergebnisse der experimentellen Untersuchung

Die folgenden Fragen sollen im Rahmen der experimentellen Hauptstudie untersucht werden:

- E1) Finden sich Unterschiede zwischen Patienten mit akuter und remittierter affektiver Störung und gesunden Kontrollen bezüglich anderer Maße der Empathie beziehungsweise des Social Functioning?
- E2) Bestehen zwischen Patienten mit akuter und remittierter affektiver Störung und gesunden Kontrollen Unterschiede bezüglich der Selbst- und Fremdeinschätzung der Empathie durch den EQ?
- E3) Ist die Diskrepanz zwischen Selbst- und Fremdeinschätzung abhängig von der Gruppenzugehörigkeit, also von dem Vorliegen oder Fehlen einer akuten oder remittierten affektiven Störung?
- E4) Besteht ein Zusammenhang zwischen der Messung der Empathie anhand des EQ sowie der behavioralen Messung anhand der CAM (Emotionserkennung) sowie des MASCS (Reasoning)?
- E5) Gibt es einen Einfluss der Empathie auf die Induzierbarkeit von Emotionen? Bestehen Gruppenunterschiede bezüglich der Emotionsinduktion? Ist es bei Personen mit hoher Empathie leichter, negative Emotionen zu induzieren? Spielt das Ausmaß der Diskrepanz zwischen Selbst- und Fremdeinschätzung eine Rolle für die Induzierbarkeit negativer Emotionen?
- E6) Bestehen Zusammenhänge zwischen der Fragebogenmessung sowie der behavioralen Messung der Empathie, dem Social Functioning und der prä-morbiden Intelligenz?

## 5.1 Fragestellung E1: Fragebogenmaße des Social Functioning

Im Folgenden werden die Ergebnisse zu weiteren Fragebogenmaßen des Social Functioning berichtet. Es handelt sich hierbei um den IIP in der Variante IIP-C mit 64 Items sowie den F-SozU Version S mit 54 Items und Teil B.

### 5.1.1 IIP

Es wird der Hauptwert des IIP, die gemittelten interpersonalen Probleme, genauer betrachtet. In Tabelle 5.1 sind die deskriptiven Ergebnisse für die einzelnen Teilnehmergruppen dargestellt.

Tabelle 5.1: Gemittelte interpersonale Probleme

	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>
KG-UR	0.86	0.34	0.17	1.58
UR	1.38	0.52	0.23	2.38
UA	1.79	0.31	1.14	2.38
BR	1.43	0.52	0.47	2.28
KG-BR	0.78	0.46	0.06	1.58
Gesamt	1.27	0.57	0.06	2.38

Anhand des Shapiro-Wilk-Tests muss davon ausgegangen werden, dass für BR ( $W = 0.95$ ,  $df = 39$ ,  $p = .065$ ) sowie KG-BR ( $W = 0.94$ ,  $df = 35$ ,  $p = .038$ ) keine Normalverteilung vorliegt. In den anderen Teilnehmergruppen kann die Normalverteilungsannahme aufrecht erhalten werden ( $p > .1$ ). Die Größen der einzelnen Teilnehmergruppen können als vergleichbar groß (KG-UR  $n = 36$ , UR  $n = 40$ , UA  $n = 41$ , BR  $n = 39$  und KG-BR  $n = 35$ ) angesehen werden. Der Levene-Test auf Varianzhomogenität wird signifikant ( $p = .001$ ), es scheint also keine homogenen Varianzen zu geben. Damit sind die Voraussetzungen für eine ANOVA nicht erfüllt. Die non-parametrische Alternative, der Kruskal-Wallis-Test, wird signifikant ( $H(4) = 83.8$ ,  $p < .001$ ), was für einen signifikanten Effekt der Gruppenzugehörigkeit spricht. Die aufgrund der Robustheit des Verfahrens zusätzlich durchgeführte ANOVA zeigt ebenfalls einen signifikanten Haupteffekt der Gruppenzugehörigkeit (Welchs  $F(4, 91.1) = 51.3$ ,  $p < .001$ ,  $f = 0.66$ ). Post-Hoc werden mit Games-Howell Unterschiede zwischen den einzelnen Gruppen untersucht: Es zeigen sich signifikante Unterschiede zwischen KG-UR und UR ( $p < .001$ ,  $d = 1.19$ ), KG-UR und UA ( $p < .001$ ,  $d = 2.84$ ), KG-BR und BR ( $p < .001$ ,  $d = 1.32$ ), UR und UA ( $p < .001$ ,  $d = 0.97$ ) sowie BR und UA ( $p = .003$ ,  $d = 0.86$ ). Keine signifikanten Unterschiede sind zwischen UR und BR ( $p = .99$ ) sowie KG-UR und KG-BR ( $p = .93$ ) zu finden. Die Ergebnisse sind in Abbildung 5.1 als Boxplot dargestellt.

Insgesamt sprechen die Ergebnisse also für vermehrte interpersonelle Probleme in den Patientengruppen, die während einer akuten Krankheitsepisode noch verstärkt werden.

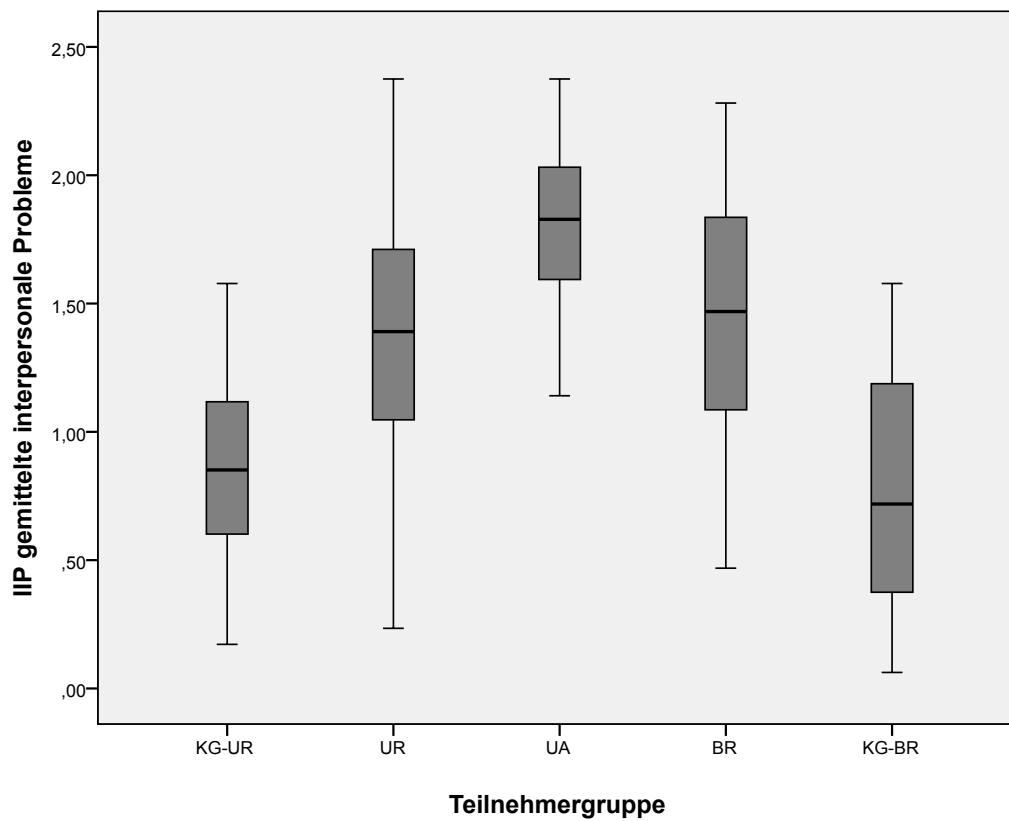


Abbildung 5.1: Boxplot IIP gemittelte interpersonale Probleme in den verschiedenen Teilnehmergruppen

### 5.1.2 F-SozU

Die vier mit dem F-SozU ermittelten Hauptskalen (soziale Belastung und die drei Unterstützungsskalen emotionale Unterstützung, praktische Unterstützung und soziale Integration) stehen im Fokus der Betrachtung. Der Gesamtwert wahrgenommene soziale Unterstützung vereint hierbei die Ergebnisse der drei Unterstützungsskalen. Shapiro-Wilk-Tests sprechen gegen eine Normalverteilung ( $p < 0.05$ ) der sozialen Belastung in einer Teilnehmergruppe (drei Gruppen mit Tendenz zur Ablehnung,  $p < 0.1$ ) und der wahrgenommenen sozialen Unterstützung (drei Gruppen mit Tendenz zur Ablehnung). In Tabelle 5.2 sind die Ergebnisse deskriptiv dargestellt.

Tabelle 5.2: F-SozU soziale Belastung und soziale Unterstützung deskriptiv

	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>
<b>Soziale Belastung</b>				
KG-UR	1.65	0.45	1.00	2.67
UR	2.08	0.80	1.00	4.33
UA	2.30	0.78	1.00	4.00
BR	2.30	0.83	1.08	3.92
KG-BR	1.72	0.52	1.00	3.08
<b>Wahrgenommene soziale Unterstützung</b>				
KG-UR	4.32	0.38	3.41	4.97
UR	3.85	0.63	2.08	4.90
UA	3.42	0.76	1.93	4.85
BR	3.59	0.66	2.21	4.63
KG-BR	4.16	0.57	2.75	4.87

Zur Untersuchung des Einflusses der Gruppenzugehörigkeit auf die Ergebnisse im F-SozU wird eine multivariate Varianzanalyse mit Gruppenzugehörigkeit als unabhängiger Variable und sozialer Belastung sowie wahrgenommener sozialer Unterstützung als abhängige Variablen durchgeführt. Der Levene-Test wird für beide abhängige Variablen signifikant ( $p < 0.01$ ), Varianzhomogenität kann daher nicht angenommen werden. Des Weiteren scheint die Gleichheit der Kovarianzmatrizen bei signifikantem Box-Test ( $p < 0.01$ ) nicht gegeben zu sein. Die Verwendung der Pillai-Spur (ebenso wie Wilks-Lambda, Hotelling-Spur und Größte charakteristische Wurzel nach Roy) weist auf einen signifikanten Effekt der Gruppenzugehörigkeit auf die soziale Belastung und die wahrgenommene soziale Unterstützung hin ( $V = 0.25$ ,  $F(8, 370) = 6.61$ ,  $p < .001$ ,  $f^2 = 0.14$ ). In einzelnen ANOVAs zeigen sich post hoc signifikante Unterschiede für die soziale Belastung ( $F(4, 185) = 7.15$ ,  $p < .001$ ,  $f = 0.37$ ) und für die wahrgenommene soziale Unterstützung ( $F(4, 185) = 14.0$ ,  $p < .001$ ,  $f = 0.48$ ). Diese Haupteffekte der Gruppenzugehörigkeit werden aufgrund fehlender Varianzhomogenität mit Games-Howell als Post-Hoc-Test genauer untersucht. Für die soziale Belastung finden sich signifikante Unterschiede ( $p < 0.05$ ) zwischen KG-UR und UR ( $d = 0.65$ ), zwischen KG-UR und UA ( $d = 1.03$ ) sowie zwischen KG-BR und BR ( $d = 0.83$ ). Die wahrgenommene soziale Unterstützung weist signifikante Unterschiede zwischen KG-UR und UR ( $d = 0.91$ ), zwischen KG-UR und

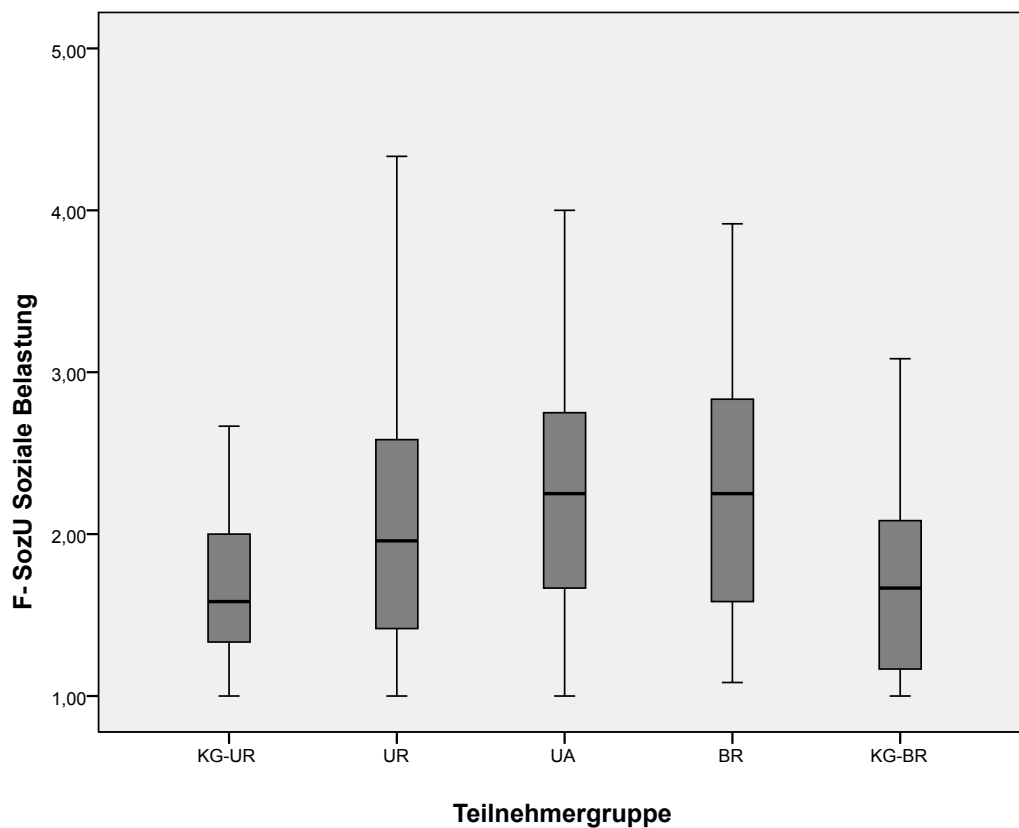


Abbildung 5.2: Boxplot F-SozU Soziale Belastung in den verschiedenen Teilnehmergruppen

UA ( $d = 1.50$ ), zwischen UR und UA ( $d = 0.62$ ) sowie zwischen KG-BR und BR ( $d = 0.93$ ) auf. In Abbildung 5.2 und 5.3 sind die Ergebnisse für die soziale Belastung respektive die wahrgenommene soziale Unterstützung als Boxplots dargestellt.

Insgesamt geben die Ergebnisse des F-SozU Hinweise darauf, dass Patienten mit affektiver Störung eine erhöhte soziale Belastung sowie geringere soziale Unterstützung erleben.

## 5.2 Fragestellung E2: Selbst- und Fremdeinschätzung der Empathie mit dem EQ

Die deskriptive Beschreibung der Ergebnisse des EQ für die verschiedenen Teilnehmergruppen sowie für die gesamte Stichprobe der Hauptstudie ist tabellarisch dargestellt: In Tabelle 5.3 befinden sich die Summen für EQ-S und EQ-F nach regulärer Auswertung.

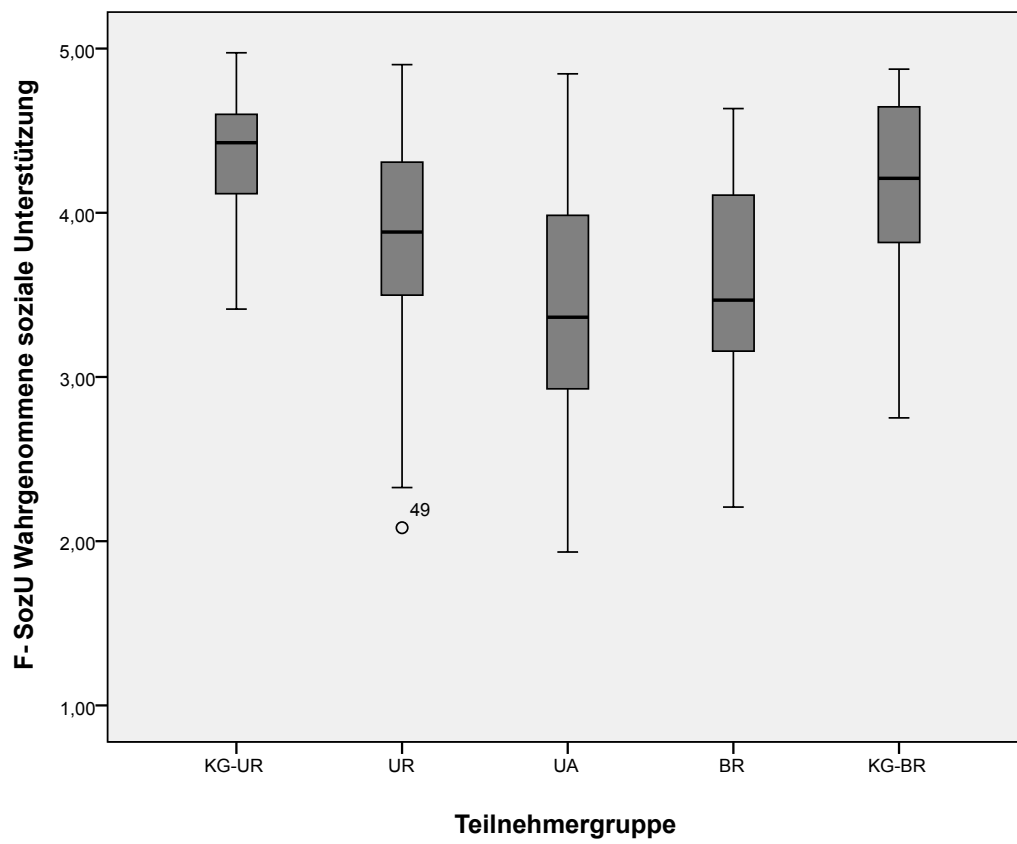


Abbildung 5.3: Boxplot F-SozU Wahrgenommene soziale Unterstützung in den verschiedenen Teilnehmergruppen

Tabelle 5.3: Regulär ausgewertete Summen des EQ-S und EQ-F für die verschiedenen Teilnehmergruppen

	M	SD	Min	Max
<b>KG-UR:</b>				
EQ-S	47.5	11.5	27	68
EQ-F	43.5	12.8	15	63
<b>UR:</b>				
EQ-S	41.2	9.46	22	61
EQ-F	38.0	14.7	8	69
<b>UA:</b>				
EQ-S	42.1	10.9	17	63
EQ-F	42.3	11.4	23	68
<b>BR:</b>				
EQ-S	42.6	11.3	17	64
EQ-F	38.3	13.5	7	70
<b>KG-BR:</b>				
EQ-S	46.6	9.35	32	68
EQ-F	42.0	11.0	23	70
<b>Gesamt:</b>				
EQ-S	43.7	10.8	17	68
EQ-F	40.6	12.9	7	70

Mit dem Shapiro-Wilk-Test werden die EQ-Summen auf Abweichungen von der Normalverteilung untersucht. Sowohl für die Summe des EQ-S als auch des EQ-F in der regulären Auswertung findet sich in keiner der Teilnehmergruppen ein Hinweis auf eine Verletzung der Normalverteilungsannahme.

### 5.2.1 Varianzanalyse zur Untersuchung von Gruppenunterschieden im EQ

Im Rahmen einer multivariaten Varianzanalyse wird der Einfluss der fünfstufigen unabhängigen Variable Gruppenzugehörigkeit (KG-UR vs. UR vs. UA vs. BR vs. KG-BR) auf die abhängigen Variablen EQ-S Summe und EQ-F Summe untersucht. Die Homogenität der Kovarianzmatrizen kann als gegeben angenommen werden, da der Box-Test auf Gleichheit der Kovarianzmatrizen nicht signifikant wird ( $p = .79$ ). Die multivariate Normalverteilung ist in SPSS nicht prüfbar, die Normalverteilung der einzelnen Variablen in den verschiedenen Gruppen kann jedoch als gegeben angenommen werden (siehe oben). Der Levene-Test auf Gleichheit der Fehlervarianzen ist weder für den EQ-S ( $p = .64$ ) noch für den EQ-F ( $p = .48$ ) signifikant, das heißt, die Annahme der Varianzgleichzeit kann aufrecht erhalten werden. Bei Verwendung der Pillai-Spur (ebenso wie Wilks-Lambda und Hotelling-Spur) zeigt sich kein signifikanter Effekt der Gruppenzugehörigkeit auf die Selbst- und Fremdbeurteilung der Empathie ( $V = 0.074$ ,  $F(8, 346) = 1.66$ ,  $p = .11$ ,  $f^2 = 0.038$ ).

Dieses Ergebnis weist also nicht auf einen Einfluss der Gruppenzugehörigkeit auf die Selbst- oder Fremdeinschätzung der Empathie hin. Das heißt, Patienten mit affektiven Störungen scheinen sich bezüglich des Ausmaßes an Empathie nicht von gesunden Kontrollen zu unterscheiden.

### 5.2.2 Unterschiede zwischen Teilnehmern mit und ohne Fremdbeurteilung

Von den insgesamt 203 Teilnehmern liegen für 178 Personen sowohl die Selbst- als auch die Fremdeinschätzung der Empathie anhand des EQ vor. Bei 20 Teilnehmern ist der EQ-S vorhanden während der EQ-F fehlt. Zwei Teilnehmer haben zwar eine Fremdeinschätzung erhalten, der EQ-S fehlt jedoch. Und bei drei weiteren Teilnehmern fehlen sowohl EQ-S als auch EQ-F. Es stellt sich die Frage, ob Unterschiede zwischen den Teilnehmern mit und ohne Fremdbeurteilung bestehen.

Tabelle 5.4: Homogenitätstest für Probanden mit und ohne EQ-F

	EQ-F vorhanden		EQ-F fehlt		$\chi^2$ -Test	
	M	SD	M	SD	$\chi^2$	<i>p</i>
EQ-S Summe	43.7	10.8	42.0	9.03	39.8	.69
Erwerbstätigkeit (derzeitige Arbeitssituation)	0.54	0.50	0.52	0.51	9.20	.51
Anzahl der Sozialkontakte	0.78	0.42	0.83	0.39	1.55	.91
BDI	8.41	10.8	8.61	12.5	32.7	.63
MSS	2.32	3.16	2.74	2.86	10.6	.65
MWT-B	31.1	4.07	30.9	3.99	15.8	.61

Die Ergebnisse der  $\chi^2$ -Homogenitätstests (siehe Tabelle 5.4) geben keinen Hinweis auf Unterschiede bezüglich der Ergebnisse des EQ-S, Depressivität, Manie, Intelligenz, Erwerbstätigkeit oder Sozialkontakten zwischen Personen mit und ohne Fremdbeurteilungs-Rating. Es wird daher angenommen, dass das Fehlen der Fremdbeurteilung nicht systematisch ist und somit Personen mit und ohne Fremdbeurteilung gemeinsam ausgewertet werden können, bzw. im Rahmen der multivariaten Varianzanalyse auf die Personen ohne Fremdbeurteilung verzichtet werden kann.

### 5.2.3 Regressionsanalyse zur Erklärung der EQ-Varianz

Zur explorativen Erklärung der Varianz des EQ wird eine lineare Regression durchgeführt. Hypothesengeleitet bieten sich die folgenden Variablen als Prädiktoren an:

1. Patient (ja vs. nein)
2. aktuelle Depressivität (BDI)
3. Geschlecht



4. regelmäßige Sozialkontakte (vorhanden vs. nicht vorhanden)
5. Intelligenz (MWT-B)
6. gemittelte interpersonale Probleme (IIP)
7. wahrgenommene Unterstützung (F-SozU)
8. soziale Belastung (F-SozU)

Zunächst wird eine Regressionsanalyse mit allen acht Prädiktoren durchgeführt, anschließend werden diejenigen Prädiktoren ausgeschlossen, die nicht signifikant zur Varianzaufklärung beitragen. Es folgt eine Post-Hoc durchgeführte Regressionsanalyse mit verminderter Prädiktorenanzahl.

### EQ-S Summe

Keine der Prädiktoren korrelieren zu hoch, das heißt über .90, miteinander, es besteht also keine Multikollinearität. Die Durbin-Watson-Statistik spricht mit 1.96 dafür, dass die Annahme der unabhängigen Fehler aufrechterhaltbar ist. Das Ergebnis der Regressionsanalyse ist mit  $R^2 = 0.299$  hoch signifikant ( $p < .001$ ,  $f^2 = 0.43$ ), die gewählten Prädiktoren erklären 29.9% der Varianz der EQ-S Summe. In Tabelle 5.5 finden sich die Ergebnisse für die einzelnen Prädiktoren.

Tabelle 5.5: Ergebnisse der multiplen Regression mit acht Prädiktoren für den regulär ausgewerteten EQ-S

	<i>B</i>	<i>SE B</i>	$\beta$
(Konstante)	17.4	9.62	
Patient	-0.95	1.83	-.043
BDI	0.15	0.079	.16 <sup>+</sup>
Geschlecht	7.68	1.61	.32 <sup>***</sup>
Regelmäßige Sozialkontakte	-0.10	1.79	-.004
Intelligenz	-0.019	0.043	-.029
Gemittelte interpersonale Probleme	-4.83	1.85	-.26 <sup>*</sup>
Wahrgenommene soziale Unterstützung	4.47	1.40	.29 <sup>**</sup>
Soziale Belastung	1.75	1.23	.12

Signifikanzniveaus: <sup>+</sup>  $p < .1$ , <sup>\*</sup>  $p < .05$ , <sup>\*\*</sup>  $p < .01$ , <sup>\*\*\*</sup>  $p < .001$

Drei der acht Prädiktoren tragen signifikant zur Varianzaufklärung bei (Geschlecht, gemittelte interpersonale Probleme und wahrgenommene soziale Unterstützung), ein weiterer Prädiktor (BDI) ist marginal signifikant. Diese vier Prädiktoren werden für eine erneute Regressionsanalyse herangezogen. In der Korrelationsmatrix zeigt sich kein Hinweis auf Multikollinearität und die Annahme der unabhängigen Fehler kann bei einer Durbin-Watson-Statistik von 1.96 aufrecht erhalten werden. Die aufgestellte Regressionsgleichung wird mit  $R^2 = 0.29$  hoch signifikant ( $p$

(<.001,  $f^2 = 0.40$ ), die vier Prädiktoren erklären somit 28.8% der Varianz der EQ-S Summe. Das Regressionsmodell mit nur vier Prädiktoren bietet also eine vergleichbar hohe Varianzaufklärung wie das ursprüngliche Modell mit acht Prädiktoren. Die detaillierten Ergebnisse der einzelnen Prädiktoren finden sich in Tabelle 5.6.

Tabelle 5.6: Ergebnisse der multiplen Regression mit vier Prädiktoren (post-hoc) für den normal ausgewerteten EQ-S

	<i>B</i>	<i>SE B</i>	$\beta$
(Konstante)	19.3	5.98	
BDI	0.14	0.077	.14 <sup>+</sup>
Geschlecht	8.05	1.58	.34 ***
Gemittelte interpersonale Probleme	-4.19	1.60	-.22 *
Wahrgenommene soziale Unterstützung	3.80	1.23	.25 **

Signifikanzniveaus: <sup>+</sup>  $p < .1$ , \*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$

### EQ-F Summe

Auch zur Erklärung der Varianz des EQ-F werden die oben genannten acht Prädiktoren herangezogen. Multikollinearität liegt nicht vor und mit Durbin-Watson-Statistik = 2.26 ist die Annahme unabhängiger Fehler aufrechterhaltbar. Mit  $R^2 = 0.16$  ( $p = .001$ ,  $f^2 = 0.19$ ) erklärt das aufgestellte Regressionsmodell 15.8% der Varianz und ist hoch signifikant. In Tabelle 5.7 finden sich die Ergebnisse für die einzelnen Prädiktoren. Einzig das Geschlecht trägt im aufgestellten Modell einen signifikanten Beitrag zur Varianzerklärung bei, die wahrgenommene soziale Unterstützung wird marginal signifikant.

Tabelle 5.7: Ergebnisse der multiplen Regression mit acht Prädiktoren für den normal ausgewerteten EQ-F

	<i>B</i>	<i>SE B</i>	$\beta$
(Konstante)	32.9	13.3	
Patient	-0.15	2.52	-.006
BDI	0.16	0.11	.14
Geschlecht	4.98	2.23	.18 *
Regelmäßige Sozialkontakte	0.36	2.39	.012
Intelligenz	-0.063	0.059	-.083
Gemittelte interpersonale Probleme	-3.70	2.51	-.17
Wahrgenommene soziale Unterstützung	3.31	1.93	.19 <sup>+</sup>
Soziale Belastung	-1.25	1.72	-.075

Signifikanzniveaus: <sup>+</sup>  $p < .1$ , \*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$

Analog zur Post-Hoc Regression für den EQ-S, wird auch für den EQ-F eine Regressionsanalyse mit den vier Prädiktoren BDI, Geschlecht, gemittelte interpersonale Probleme und wahrgenommene soziale Unterstützung durchgeführt. Multikollinearität liegt nicht vor, die Annahme der unabhängigen Fehler ist aufrecht zu erhalten (Durbin-Watson-Statistik = 2.23). Die aufgestellte Regressionsgleichung erklärt mit  $R^2 = 0.15$  ( $p < .001$ ,  $f^2 = 0.18$ ) 15.0% der Varianz der EQ-F Summe. Die detaillierten Ergebnisse für die vier Prädiktoren sind Tabelle 5.8 zu entnehmen. Sowohl Geschlecht als auch wahrgenommene soziale Unterstützung leisten einen signifikanten Beitrag, während die gemittelten interpersonalen Probleme marginal signifikant sind.

Tabelle 5.8: Ergebnisse der multiplen Regression mit vier Prädiktoren (post-hoc) für den normal ausgewerteten EQ-F

	<i>B</i>	<i>SE B</i>	$\beta$
(Konstante)	20.6	8.10	
BDI	0.17	0.10	.15
Geschlecht	4.75	2.17	.17 *
Gemittelte interpersonale Probleme	-4.02	2.16	-.19 +
Wahrgenommene soziale Unterstützung	4.15	1.66	.24 *

Signifikanzniveaus: +  $p < .1$ , \*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$

### 5.3 *Fragestellung E3: Diskrepanz zwischen Selbst- und Fremdeinschätzung der Empathie*

Im Folgenden soll zunächst untersucht werden, ob ein Zusammenhang zwischen Selbst- und Fremdbeurteilung besteht und ob systematische Unterschiede zwischen beiden zu beobachten sind. Anschließend sollen mögliche Gruppenunterschiede bezüglich des Ausmaßes der möglichen Diskrepanz betrachtet werden.

#### 5.3.1 Besteht eine Diskrepanz zwischen Selbst- und Fremdbeurteilung?

Die Summen des EQ-S ( $M = 43.7$ ,  $SD = 10.8$ ) und des EQ-F ( $M = 40.6$ ,  $SD = 12.9$ ) korrelieren mittelstark miteinander ( $r = .46$ ,  $n = 178$ ,  $p < .001$ ). Zur Klärung der Frage, ob die Mittelwerte der beiden Fragebogenvarianten systematisch unterschiedlich sind, wird ein gepaarter t-Test durchgeführt. Es ergibt sich ein signifikanter Unterschied zwischen Selbst- und Fremdbeurteilung ( $t = 3.33$ ,  $p < .001$ ,  $d = .25$ ). Das heißt, die Teilnehmer schätzen sich selbst als signifikant empathischer ein, als sie von ihren Fremdbeurteilern eingeschätzt werden.

Tabelle 5.9: Deskriptive Beschreibung der Diskrepanz zwischen Selbst- und Fremdeinschätzung der Empathie

	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>
KG-UR	4.00	12.8	-25	29
UR	3.21	12.5	-22	38
UA	-0.18	12.4	-24	25
BR	4.31	13.3	-22	27
KG-BR	4.66	10.7	-25	32
Gesamt	3.11	12.4	-25	38

### 5.3.2 Gruppenunterschiede bezüglich der Diskrepanz zwischen Selbst- und Fremdbeurteilung

Zur Untersuchung des Ausmaßes der Diskrepanz zwischen Selbst- und Fremdeinschätzung der Empathie wird für jeden Probanden die Gesamtsumme des EQ-F von der Gesamtsumme des EQ-S subtrahiert. Ein positiver Wert bedeutet in beiden Auswertungs-Szenarien dass die Selbstbeurteilung höher ausfällt als die Fremdbeurteilung. Das heißt, die Person schreibt sich selbst ein größeres Ausmaß an Empathie zu, als sie von ihrem Fremdbeurteiler zugeschrieben bekommt. Umgekehrt bedeutet ein negativer Wert, dass die Selbstbeurteilung niedriger ist als die Fremdbeurteilung, das heißt die Person sieht sich als weniger empathisch, als ihr Fremdbeurteiler sie sieht. Dieser Wert für die Diskrepanz zwischen Selbst- und Fremdeinschätzung der Empathie wird im folgenden als EQ-Diff bezeichnet. In Tabelle 5.9 sind die deskriptiven Werte der EQ-Diff zu sehen.

Da der Shapiro-Wilk-Test auf Normalverteilung in keiner der Gruppen signifikant wird ( $p > .05$  in allen Gruppen), wird davon ausgegangen, dass EQ-Diff in allen Teilnehmergruppen normalverteilt ist. Der Levene-Test spricht für Varianzgleichheit ( $F(4, 173) = 0.85, p = .49$ ). Eine einfaktorielle Varianzanalyse mit der Gruppenzugehörigkeit (fünfstufig) als unabhängiger Variable und der Diskrepanz zwischen Selbst- und Fremdeinschätzung als abhängiger Variable weist nicht auf einen signifikanten Einfluss der Gruppenzugehörigkeit auf EQ-Diff hin ( $F(4, 173) = 0.91, p = .46$ ). Es scheint also keine bedeutsamen Gruppenunterschiede bezüglich der Diskrepanz zwischen Selbst- und Fremdeinschätzung zu geben. Die sehr großen Standardabweichungen sowie weiten Spannen zwischen Minimum und Maximum, die in der deskriptiven Beschreibung auffallen, sprechen für große interindividuelle Unterschiede bezüglich EQ-Diff in allen Teilnehmergruppen. Wie an den Rohwerten zu sehen ist, fällt der Mittelwert der unipolar akuten Patienten geringer aus, als die Mittelwerte der anderen Gruppen. Ein signifikanter Gruppenunterschied bei größeren Gruppengrößen erscheint möglich.

## 5.4 Fragestellung E4: Behaviorale Messung der Empathie

Als behaviorale Maße der Empathie wurden in der vorliegenden Studie die CAM für die Messung des Decoding (Emotionserkennung), sowie der MASC zur Messung des Reasoning erhoben. Da beide zur experimentellen Manipulation auf zwei Sitzungen aufgeteilt wurden und diese Manipulation für das vorliegende Unterkapitel nicht von Relevanz ist, werden die beiden Teile sowohl der CAM als auch des MASC zu einem Gesamtwert zusammengeführt. Dies wird im Folgenden als CAM Gesamt beziehungsweise MASC Gesamt bezeichnet. Durch die Pseudo-Randomisierung der Testteile und der experimentellen Manipulation spielen etwaige Einflüsse der Reihenfolge und der experimentellen Sitzung keine Rolle für den Vergleich der Gesamtergebnisse der beiden Testverfahren mit dem EQ.

### 5.4.1 CAM Gesamt

Die deskriptive Beschreibung der CAM Gesamt ist in Tabelle 5.10 zu finden. Aufgrund des Shapiro-Wilk-Tests auf Normalverteilung muss die Normalverteilungsannahme für KG-UR abgelehnt werden ( $W = 0.92$ ,  $df = 38$ ,  $p = .011$ ), zudem besteht eine Tendenz zur Ablehnung für UA ( $W = 0.95$ ,  $df = 43$ ,  $p = .077$ ). Für die verbleibenden Teilnehmergruppen kann Normalverteilung angenommen werden ( $p > 0.05$ ). Der Levene-Test auf Varianzhomogenität wird nicht signifikant ( $p = .378$ ), die Annahme der Varianzhomogenität kann daher aufrecht erhalten werden. Trotz der Verletzung der Normalverteilungsannahme wird die ANOVA als robustes Verfahren (siehe Kapitel 4.6) mit der Gruppenzugehörigkeit (fünfstufig) als unabhängiger Variable und der Summe richtiger Antworten in der CAM Gesamt als abhängiger Variable durchgeführt. Es kann kein signifikanter Haupteffekt der Gruppenzugehörigkeit gefunden werden ( $F(4, 198) = 0.28$ ,  $p = .89$ ).

Tabelle 5.10: Summe richtiger Antworten in der CAM Gesamt, deskriptiv

	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>
<b>Anteil richtiger Antworten in Prozent</b>				
KG-UR	69.3	10.9	44.7	84.3
UR	68.4	9.65	44.6	84.4
UA	68.9	10.2	37.9	86.5
BR	67.2	8.4	48.6	88.9
KG-BR	68.0	11.0	42.5	86.8

Es stellt sich die Frage, ob der erwartete Einfluss der Gruppenzugehörigkeit auf die Emotionserkennung sich anstatt in der Summe richtiger Antworten in Reaktionszeit, Sicherheits- oder Schwierigkeitseinschätzung niederschlägt, beziehungsweise ob das offenbare Fehlen des Einflusses der Gruppenzugehörigkeit durch diese Maße erklärt werden kann. Daher werden einfaktorielle Varianzanalysen mit der Gruppenzugehörigkeit als unabhängiger Variable für die Reaktionszeit, die mittlere Sicherheit sowie die mittlere Schwierigkeit durchgeführt:

### Reaktionszeit

Der Shapiro-Wilk-Test spricht gegen eine Normalverteilung der Reaktionszeit in den Teilnehmergruppen KG-UR ( $W = 0.90$ ,  $df = 38$ ,  $p = .003$ ), UA ( $W = 0.93$ ,  $df = 43$ ,  $p = .013$ ) und KG-BR ( $W = 0.92$ ,  $df = 35$ ,  $p = .010$ ). Für UR und BR kann die Normalverteilungsannahme aufrecht erhalten werden ( $p > 0.1$ ). Mit einem nicht signifikanten Levene-Test ( $p = .46$ ) kann die Annahme der Varianzhomogenität aufrecht erhalten werden. Im Rahmen der Varianzanalyse wird kein signifikanter Haupteffekt der Gruppenzugehörigkeit auf die Reaktionszeit in der CAM Gesamt beobachtet ( $F(4, 198) = 0.44$ ,  $p = .78$ ).

### Mittlere Sicherheit

In den Teilnehmergruppen KG-UR ( $W = 0.94$ ,  $df = 38$ ,  $p = .041$ ) sowie UA ( $W = 0.932$ ,  $df = 43$ ,  $p = .013$ ) spricht der Shapiro-Wilk-Test gegen eine Normalverteilung der Sicherheitseinschätzung. Für UR, BR und KG-BR kann die Normalverteilungsannahme aufrecht erhalten werden ( $p > .10$ ). Der Levene-Test spricht für Varianzhomogenität ( $p = .061$ ). Mittels der Varianzanalyse zeigt sich ein signifikanter Haupteffekt der Gruppenzugehörigkeit ( $F(4, 198) = 3.98$ ,  $p = .004$ ,  $f = 0.27$ ). Post-Hoc-Einzelvergleiche nach Gabriel zeigen einen signifikanten Unterschied ( $p < .05$ ) zwischen KG-UR und UA ( $d = 0.97$ ). In Abbildung 5.4 sind die Ergebnisse der Sicherheitseinschätzung als Boxplot dargestellt. Wie dort zu sehen ist, sind sich akut depressive Teilnehmer signifikant unsicherer als die gesunde Kontrollgruppe.

### Mittlere Schwierigkeit

Die Normalverteilungsannahme kann für UA ( $W = 0.91$ ,  $df = 43$ ,  $p = .003$ ) sowie KG-BR ( $W = 0.94$ ,  $df = 35$ ,  $p = .044$ ) nicht aufrecht erhalten werden. Der Shapiro-Wilk-Test spricht hingegen für Normalverteilung der mittleren Schwierigkeit in den Teilnehmergruppen UR, BR und KG-UR ( $p > .10$ ). Die Annahme der Varianzhomogenität kann auf Basis des Levene-Tests ( $p = .075$ ) aufrecht erhalten werden. Die einfaktorielle Varianzanalyse zeigt einen signifikanten Haupteffekt der Gruppenzugehörigkeit ( $F(4, 198) = 3.46$ ,  $p = .009$ ,  $f = 0.26$ ). In Post-Hoc-Einzelvergleichen nach Gabriel findet sich ein signifikanter Unterschied ( $p < .05$ ) zwischen KG-UR und UA ( $d = 0.47$ ). Wie in Abbildung 5.5 zu sehen ist, empfinden akut depressive Patienten die Emotionserkennung als signifikant schwieriger als gesunde Kontrollen.

Insgesamt lässt sich also sagen, dass akut depressive Patienten sich sowohl signifikant unsicherer bezüglich der Emotionserkennung fühlen als auch die Schwierigkeit der Aufgaben zur Emotionserkennung signifikant schwieriger einstufen als gesunde Kontrollen. Gleichzeitig scheinen sie weder langsamer zu sein, noch eine geringere Anzahl richtiger Antworten zu geben. Zwischen Patienten mit remittierter unipolarer und bipolarer Störung und gesunden Kontrollen sind in keinem der Bereiche Unterschiede zu beobachten. Es scheint also kein Zusammenhang zwischen affektiven Störungen und Beeinträchtigungen oder anderen Veränderungen der Emotionserkennung zu bestehen.

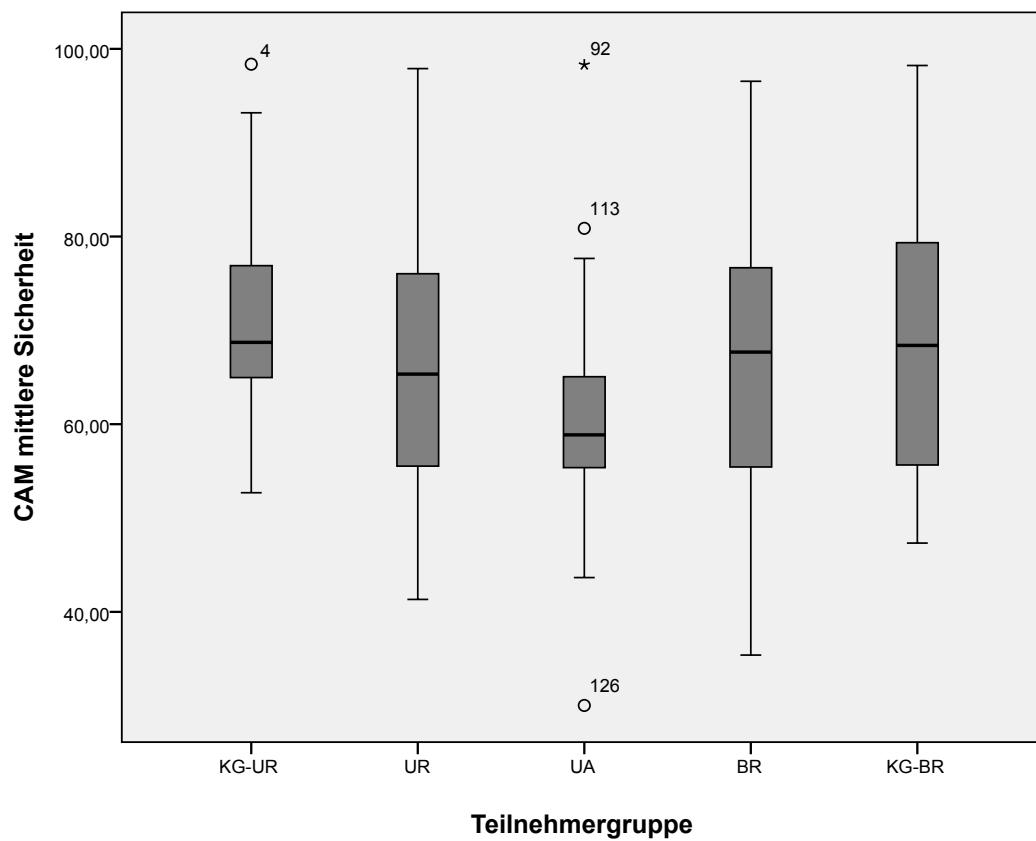


Abbildung 5.4: Boxplot mittlere Sicherheit CAM Gesamt in den verschiedenen Teilnehmergruppen

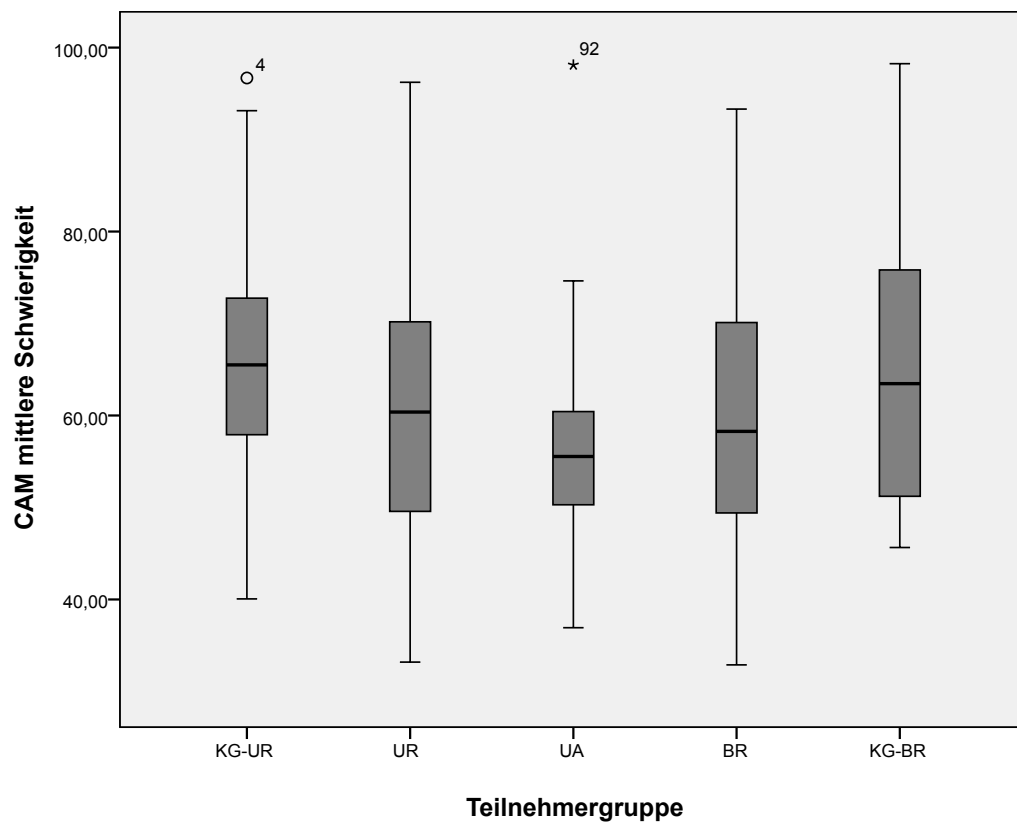


Abbildung 5.5: Boxplot mittlere Schwierigkeit CAM Gesamt in den verschiedenen Teilnehmergruppen



Bezüglich der geringeren Sicherheits- und höheren Schwierigkeitseinschätzung akut depressiver Patienten, stellt sich die Frage ob dies ein Effekt ist, der spezifisch für die Emotionserkennung ist. Möglicherweise nehmen sich akut Depressive hinsichtlich jeglicher Fragestellungen als unsicherer wahr und empfinden jegliche Fragestellungen als schwieriger als gesunde Probanden. Zur Untersuchung dieser Möglichkeit werden im Folgenden die Ergebnisse der Kontrollaufgabe CALC herangezogen.

### CALC

Anhand des Shapiro-Wilk-Tests muss angenommen werden, dass die Summe richtiger Antworten in keiner der Teilnehmergruppen normalverteilt ist ( $p < .01$  in allen Gruppen). Die Annahme der Varianzhomogenität kann auf Basis des Levene-Tests ( $p = .49$ ) aufrecht erhalten werden. In einer einfaktoriellen Varianzanalyse zeigt sich kein signifikanter Haupteffekt der Gruppenzugehörigkeit auf die Summe richtiger Antworten ( $F(4, 196) = 1.04, p = .39$ ). Auch mittels der nonparametrischen Alternative Kruskal-Wallis lässt sich kein Einfluss der Gruppenzugehörigkeit auf die Summe richtiger Antworten feststellen ( $H(4) = 7.95, p = .093$ ). Wie in Abbildung 5.6 zu sehen ist, handelt es sich um einen Deckeneffekt. Die Aufgabe wird von allen Teilnehmergruppen äußerst erfolgreich gelöst.

Die Reaktionszeit scheint in keiner der Gruppen normalverteilt zu sein (Shapiro-Wilk:  $p < .05$  in allen Gruppen). Bei einem signifikanten Levene-Test ( $p = .044$ ) muss die Annahme der Varianzhomogenität abgelehnt werden. Eine einfaktorielle Varianzanalyse kann keinen signifikanten Haupteffekt der Gruppenzugehörigkeit auf die Reaktionszeit finden (Welchs  $F(4, 97.423) = 2.45, p = .051$ ). Es besteht jedoch eine Tendenz zur Signifikanz, daher werden Post-Hoc-Einzelvergleiche nach Gabriel betrachtet: Hier finden sich keine signifikanten Unterschiede zwischen den einzelnen Teilnehmergruppen. Auch die nonparametrische Alternative zur Varianzanalyse, der Kruskal-Wallis-Test, findet keinen signifikanten Effekt der Gruppenzugehörigkeit ( $H(4) = 7.51, p = .11$ ).

Auch für die Sicherheitseinschätzung muss die Normalverteilungsannahme in allen Gruppen abgelehnt werden (Shapiro-Wilk:  $p < .05$  in allen Gruppen). Die Annahme der Varianzhomogenität kann anhand des Levene-Tests aufrecht erhalten werden ( $p = .29$ ). Im Rahmen einer einfaktoriellen Varianzanalyse zeigt sich kein signifikanter Einfluss der Gruppenzugehörigkeit auf die Sicherheitseinschätzung ( $F(4, 196) = 0.44, p = .78$ ). Ebenso zeigt sich auch im nonparametrischen Kruskal-Wallis-Test kein Haupteffekt der Gruppenzugehörigkeit ( $H(4) = 3.63, p = .46$ ).

Wie der Shapiro-Wilk-Test zeigt ( $p < 0.05$ ), scheint die Schwierigkeitseinschätzung in vier Gruppen nicht normal verteilt zu sein. Für KG-UR besteht eine Tendenz zur Ablehnung der Normalverteilungsannahme ( $p < .10$ ). Der Levene-Test wird nicht signifikant ( $p = .71$ ), Varianzhomogenität kann also angenommen werden. Weder im Rahmen der einfaktoriellen Varianzanalyse ( $F(4, 196) = 1.21, p = .31$ ) noch in der nonparametrischen Alternative Kruskal-Wallis ( $H(4) = 4.88, p = .30$ ) kann ein signifikanter Einfluss der Gruppenzugehörigkeit auf die mittlere Schwierigkeit beobachtet werden.

Anders als bei der CAM finden sich also im CALC keine Hinweise auf eine erhöhte Unsicherheit und erhöhte Schwierigkeitswahrnehmung akut depressiver Patienten.

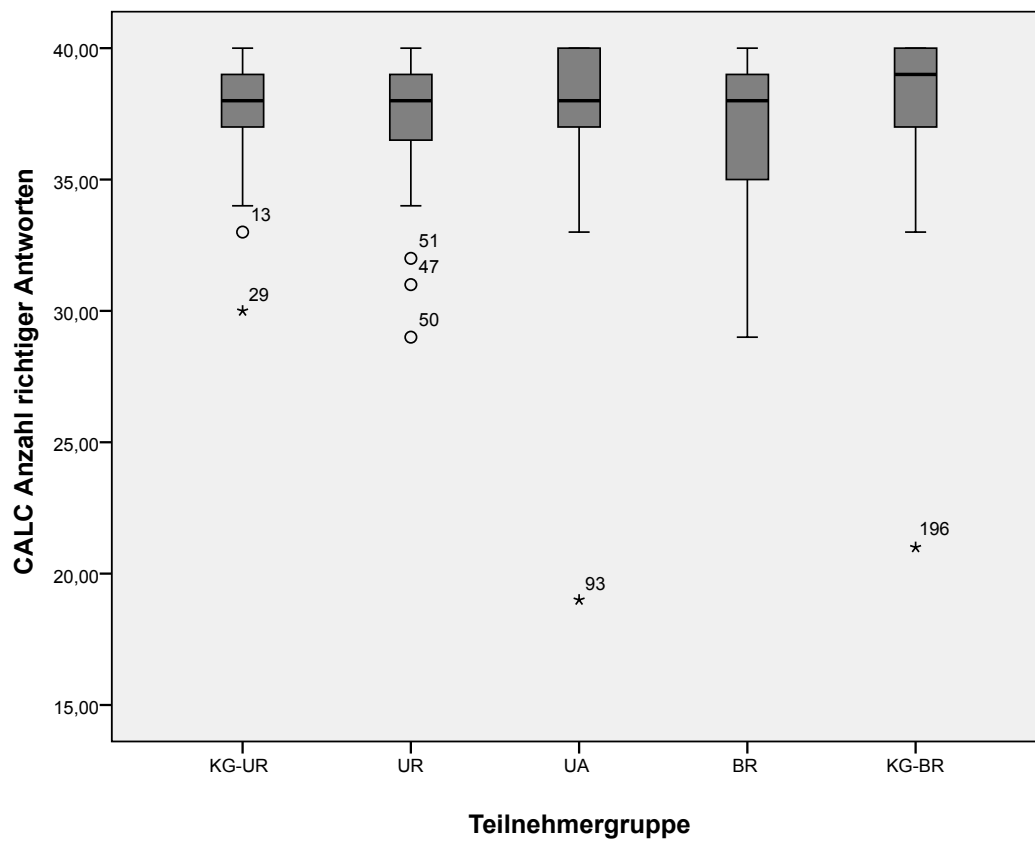


Abbildung 5.6: Boxplots der Anzahl richtiger Antworten im CALC für die verschiedenen Teilnehmergruppen

Dies spricht für einen für die Emotionserkennung spezifischen Effekt der verminderten Sicherheitseinschätzung und erhöhten Schwierigkeitseinschätzung akut depressiver Patienten.

### 5.4.2 MASC Gesamt

In Tabelle 5.11 ist die Anzahl richtiger Antworten im MASC Gesamt deskriptiv dargestellt.

Tabelle 5.11: Summe richtiger Antworten im MASC Gesamt deskriptiv

	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>
<b>Anzahl richtiger Antworten</b>				
KG-UR	36.7	3.98	23	43
UR	33.6	4.11	22	43
UA	34.3	5.16	18	44
BR	34.6	4.99	20	43
KG-BR	34.4	4.59	26	43

Um den Einfluss der Gruppenzugehörigkeit auf die Anzahl richtiger Antworten zu untersuchen, wird eine univariate Varianzanalyse mit der Gruppenzugehörigkeit als unabhängiger Variable (fünfstufig) und der Anzahl richtiger Antworten als abhängiger Variable durchgeführt. Shapiro-Wilk-Tests auf Normalverteilung sprechen dafür, die Normalverteilungsannahme für die Gruppen KG-UR ( $W = 0.92$ ,  $df = 38$ ,  $p = .012$ ), UA ( $W = 0.92$ ,  $df = 43$ ,  $p = .005$ ) und BR ( $W = 0.93$ ,  $df = 42$ ,  $p = .016$ ) abzulehnen. Für UR und KG-BR scheint die Anzahl richtiger Antworten im MASC Gesamt hingegen normalverteilt zu sein ( $p > .1$ ). Aufgrund der Robustheit des Verfahrens (siehe Kapitel 4.6), wird die ANOVA trotz der Verletzung der Normalverteilungsannahme berechnet. Mit einem nicht signifikanten Levene-Test ( $p = .40$ ) kann die Annahme der Varianzhomogenität aufrecht erhalten werden. Es zeigt sich ein signifikanter Haupteffekt der Gruppenzugehörigkeit ( $F(4, 198) = 2.47$ ,  $p = .046$ ,  $f = 0.22$ ). In Post-Hoc Einzelvergleichen nach Gabriel werden signifikante ( $p < .05$ ) Gruppenunterschiede ausschließlich zwischen KG-UR und UR ( $d = 0.75$ ) beobachtet.

Neben der Anzahl richtiger Antworten lässt sich der MASC hinsichtlich der Art der gemachten Fehler auswerten. Für jede Frage stehen vier Antwortmöglichkeiten zur Auswahl, neben der korrekten Antwort, drei verschiedene falsche Antworten. Eine deren Inhalt keine ToM erfordert, eine mit wenig ToM sowie eine mit zu viel ToM. Eine multivariate Varianzanalyse soll den Einfluss der Gruppenzugehörigkeit (fünfstufig) auf die unterschiedlichen Fehlerarten (keine ToM vs. wenig ToM vs. zu viel ToM) genauer beleuchten. Anhand von Shapiro-Wilk-Tests wird die Normalverteilungsannahme der einzelnen Fehlerarten in den Teilnehmergruppen überprüft: Die Anzahl der Fehler keine ToM scheint in keiner Gruppe normalverteilt zu sein ( $p < .01$  in allen Gruppen). Auch die Anzahl der Fehler scheint in vier von fünf Gruppen nicht normalverteilt zu sein ( $p < .05$ ), einzig für KG-BR kann die Normalverteilungsannahme aufrecht erhalten werden ( $p = .14$ ). Für die Anzahl der Fehler

zu viel ToM muss die Normalverteilungsannahme in allen Gruppen abgelehnt werden ( $p < .05$ ). Die multivariate Normalverteilung scheint also wegen fehlender Normalverteilung der einzelnen Variablen nicht gegeben zu sein. An der Durchführung der MANOVA wird aufgrund der Robustheit des Verfahrens festgehalten (siehe Kapitel 4.6). Der Box-Test auf Gleichheit der Kovarianzmatrizen wird signifikant ( $p = .001$ ), die Homogenität der Kovarianzmatrizen kann daher nicht als gegeben angenommen werden. Da der Box-Test als sehr sensibler Test kritisiert und seine Anwendung in Frage gestellt wird, wird die MANOVA dennoch durchgeführt. Der Levene-Test auf Gleichheit der Fehlervarianzen wird für alle Fehlerarten nicht signifikant ( $p > .05$ ), dies spricht dafür, die Annahme der Varianzhomogenität aufrecht zu erhalten. Bei Verwendung der Pillai-Spur (ebenso wie Wilks-Lambda und Hotelling-Spur, aber nicht Größte charakteristische Wurzel nach Roy) zeigt sich kein signifikanter Effekt der Gruppenzugehörigkeit auf Fehler im MASC ( $V = 0.078$ ,  $F(12, 594) = 1.32$ ,  $p = .21$ ). Die verschiedenen Teilnehmergruppen scheinen sich also nicht bezüglich der Art der gemachten Fehler zu unterscheiden.

Die Items des MASC lassen sich weiterhin in zwei Bereiche aufteilen: Solche für die emotionale ToM (Decoding) benötigt wird und solche die eher kognitive ToM (Reasoning) erfordern. Es wird untersucht, ob es spezifische Effekte der Gruppenzugehörigkeit je nach Art der ToM gibt. Dazu wird eine multivariate Varianzanalyse mit der Gruppenzugehörigkeit als unabhängiger Variable (fünfstufig) und jeweils sowohl der Anzahl richtiger Antworten als auch den Anzahl der verschiedenen Fehlerarten (keine ToM vs. wenig ToM vs. zu viel ToM) für die emotionale ToM und die kognitive ToM. Es ergeben sich also acht abhängige Variablen. Der Box-Test kann nicht berechnet werden, da weniger als zwei nichtsinguläre Zellen-Kovarianzmatrizen vorhanden sind. Die multivariate Normalverteilung scheint nicht gegeben zu sein, da die Normalverteilungsannahme der einzelnen Variablen jeweils in zwei bis fünf der Teilnehmergruppen abzulehnen ist (Shapiro-Wilk  $p < .05$ ). Der Levene-Test auf Gleichheit der Fehlervarianzen ist für alle abhängigen Variablen nicht signifikant ( $p > .05$ ), das heißt die Annahme der Varianzhomogenität kann aufrecht erhalten werden. Bei Verwendung der Pillai-Spur (ebenso wie Wilks-Lambda und Hotelling-Spur, aber nicht Größte charakteristische Wurzel nach Roy) zeigt sich kein signifikanter Effekt der Gruppenzugehörigkeit auf MASC-Ergebnisse ( $V = 0.18$ ,  $F(32, 776) = 1.14$ ,  $p = .27$ ). Die Gruppenzugehörigkeit scheint also keinen Einfluss auf die mit dem MASC gemessene emotionale oder kognitive ToM zu haben.

### 5.4.3 Zusammenhang behavioraler Maße mit dem EQ

Im Folgenden soll der Zusammenhang der Messung der Empathie anhand des EQ mit der behavioralen Messung derselben im Rahmen der CAM und des MASC näher beleuchtet werden. Hierzu werden zunächst signifikante Korrelationen der EQ-S oder EQ-F Summen mit den Ergebnissen der CAM und des MASC berichtet: Der Anteil richtiger Antworten in der CAM korreliert leicht positiv mit dem EQ-S ( $\rho = .16$ ,  $n = 198$ ,  $p = .028$ ) und mittelstark positiv mit dem EQ-F ( $\rho = .21$ ,  $n = 180$ ,  $p = .004$ ). Die mittlere Reaktionszeit in der CAM korreliert schwach negativ mit dem EQ-S ( $\rho = -.16$ ,  $n = 198$ ,  $p = .026$ ). Die mittlere Sicherheit korreliert mittelstark mit dem EQ-S ( $\rho = .21$ ,  $n = 198$ ,  $p = .003$ ) und dem EQ-F ( $\rho = .21$ ,  $n = 180$ ,  $p$

= .006). Auch die mittlere Schwierigkeit korreliert mittelstark mit EQ-S ( $\rho = .20$ ,  $n = 198$ ,  $p = .004$ ) sowie EQ-F ( $\rho = .26$ ,  $n = 180$ ,  $p = .001$ ). Die Positivity in der CAM korreliert mittelstark negativ mit dem EQ-S ( $\rho = -.19$ ,  $n = 198$ ,  $p = .007$ ). Die Anzahl richtiger Antworten im MASC korreliert schwach positiv mit dem EQ-S ( $\rho = .18$ ,  $n = 198$ ,  $p = .012$ ) und mittelstark positiv mit dem EQ-F ( $\rho = .22$ ,  $n = 180$ ,  $p = .004$ ). Die Anzahl der Fehler keine ToM korreliert schwach negativ mit dem EQ-F ( $\rho = -.157$ ,  $n = 180$ ,  $p = .035$ ). Die Anzahl der Fehler wenig ToM korreliert schwach negativ mit dem EQ-S ( $\rho = -.19$ ,  $n = 198$ ,  $p = .007$ ) und dem EQ-F ( $\rho = -.16$ ,  $n = 180$ ,  $p = .028$ ). Insgesamt scheinen also leichte bis mittelstarke Zusammenhänge zwischen der Messung der Empathie anhand eines Fragebogens und anhand behavioraler Aufgaben zu bestehen.

Anhand von Regressionsanalysen sollen die Zusammenhänge zwischen behavioraler Messung und Fragebogenmessung genauer untersucht werden. Es wird eine Regression für den Anteil richtiger Antworten in der CAM mit sechs Prädiktoren durchgeführt: Als Prädiktoren werden die Anzahl richtiger Antworten im MASC, die EQ-S Gesamtsumme, die EQ-F Gesamtsumme, der BDI, das Vorliegen einer psychischen Erkrankung (Patient vs. gesunde Kontrolle) sowie das Geschlecht herangezogen. Die Korrelationsmatrix zeigt, dass keine der Prädiktoren zu hoch miteinander korrelieren, es liegt also keine Multikollinearität vor. Anhand einer Durbin-Watson-Statistik von 1.89 ist die Annahme der unabhängigen Fehler aufrechterhaltbar. Mit  $R^2 = 0.16$  ( $p < .001$ ,  $f^2 = 0.19$ ) erklären die Prädiktoren 15.8% der Varianz des Anteils richtiger Antworten in der CAM. In Tabelle 5.12 sind die detaillierten Ergebnisse für die einzelnen Prädiktoren zu finden. Es ist zu sehen, dass die Anzahl richtiger Antworten im MASC der einzige Prädiktor zu sein scheint, der signifikant zur Varianzaufklärung beiträgt.

Tabelle 5.12: Ergebnisse der multiplen Regression mit sechs Prädiktoren für den Anteil richtiger Antworten in der CAM

	<i>B</i>	<i>SE B</i>	$\beta$
(Konstante)	37.9	6.26	
MASC Summe	0.75	0.15	.36***
EQ-S Summe	0.030	0.081	.033
EQ-F Summe	0.10	0.062	.13
BDI Summe	0.022	0.072	.024
Patient vs. Kontrolle	0.90	1.72	.043
Geschlecht	-1.00	1.78	-.044

Signifikanzniveaus: <sup>+</sup>  $p < .1$ , \*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$

Auch für die Anzahl richtiger Antworten im MASC wird eine Regressionsanalyse mit sechs Prädiktoren durchgeführt: Der Anteil richtiger Antworten in der CAM, die Gesamtsumme des EQ-S, die Gesamtsumme des EQ-F, der BDI, das Vorliegen einer psychischen Erkrankung (Patient vs. gesunde Kontrolle) sowie das Geschlecht werden als Prädiktoren herangezogen. In der Korrelationsmatrix ist zu sehen, dass keine der Prädiktoren zu hoch miteinander korrelieren, es liegt also keine Multikollinearität vor. Mit Durbin-Watson-Statistik = 1.84 ist die Annahme der unabhängigen

Fehler aufrechterhaltbar. Es ergibt sich eine Regressionsgleichung mit  $R^2 = 0.16$  ( $p < .001$ ,  $f^2 = 0.19$ ), das heißt, die Prädiktoren erklären 15.9% der Varianz der Anzahl richtiger Antworten im MASC. Die detaillierten Ergebnisse für die einzelnen Prädiktoren finden sich in Tabelle 5.13. Wie dort zu sehen ist, scheinen nur der Anteil richtiger Antworten in der CAM und das Vorliegen einer psychischen Erkrankung einen signifikanten Anteil zur Varianzaufklärung beizutragen. Das Geschlecht weist eine Tendenz zur Signifikanz auf.

Tabelle 5.13: Ergebnisse der multiplen Regression mit sechs Prädiktoren für die Summe richtiger Antworten im MASC

	B	SE B	$\beta$
(Konstante)	22.0	2.78	
CAM Anteil korrekt	0.17	0.034	.35***
EQ-S Summe	-0.031	0.038	-.069
EQ-F Summe	0.035	0.029	.094
BDI Summe	0.014	0.034	.031
Patient vs. Kontrolle	-1.88	0.80	-.19*
Geschlecht	1.420	0.829	.130 <sup>+</sup>

Signifikanzniveaus: <sup>+</sup>  $p < .1$ , \*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$

## 5.5 Fragestellung E5: Empathie und Induzierbarkeit negativer Emotionen

Diese Fragestellung befasst sich mit Empathie und der Induzierbarkeit negativer Emotionen. Hierzu wird zunächst überprüft, ob die Emotionsinduktion erfolgreich und anhaltend ist. Im Anschluss wird untersucht, ob die experimentelle Manipulation der Stimmung einen Einfluss auf die Ergebnisse in CAM und MASC hat. Im nächsten Schritt werden Gruppenunterschiede bezüglich der Induzierbarkeit negativer Emotionen betrachtet. Weiterhin werden Zusammenhänge zwischen dem Ausmaß der mit dem EQ gemessenen Empathie und der Induzierbarkeit negativer Emotionen analysiert. In Fragestellung 6 wird eine explorative Herangehensweise zur Eröffnung neuer Forschungsfragen gewählt.

### 5.5.1 Überprüfung der Emotionsinduktion

Jeweils vor Beginn der CAM und dem MASC wurde eine zehn- respektive fünfminütige Stimmungsinduktion durchgeführt. Die initiale Induktion negativer Emotionen vor der CAM wird im Folgenden als Emotionsinduktion 1 (MI1) und die Wiederholung der Induktion vor dem MASC als Emotionsinduktion 2 (MI2) bezeichnet. Für beide Zeitpunkte stellt sich die Frage, ob die Emotionsinduktion tatsächlich eine Erhöhung des negativen Affektes und eine Senkung des positiven Affektes bewirkt hat. Hierzu wird ein Vergleich der Messung der Stimmung vor (Prä MI) und direkt nach (Post MI) der Emotionsinduktion herangezogen. Zudem wird die Messung der Stimmung nach der Emotionsinduktion (Post MI) mit der Messung der Stimmung in der

Sitzung ohne Emotionsinduktion (NMI) jeweils direkt vor Durchführung der CAM (Prä CAM) beziehungsweise MASC (Prä MASC) verglichen. Zur Untersuchung der Dauerhaftigkeit der Emotionsinduktion wird die Messung der Stimmung direkt vor Beginn der Aufgabe (Post MI 1 bzw. 2) mit der Messung nach der Aufgabe (Prä MI2 bzw. Abschluss) verglichen.

Zur Untersuchung dieser Fragen werden zwei- bzw. mehrfaktorielle Varianzanalysen im gemischten Design (Mixed-Design ANOVA) mit den unabhängigen Variablen Gruppenzugehörigkeit (fünfstufig; Zwischensubjektfaktor), dem Zeitpunkt (fünfstufig innerhalb der MI-Sitzung, dreistufig beim Vergleich beider Sitzungen; Innersubjektfaktor) und der Sitzung (fällt weg innerhalb der MI-Sitzung, zweistufig beim Vergleich beider Sitzungen; Innersubjektfaktor) durchgeführt.

### MI-Sitzung

Im Rahmen einer solchen Mixed-Design-ANOVA soll der Einfluss der Gruppenzugehörigkeit und des Zeitpunkts (fünfstufig: Prä MI1 vs. Post MI1 vs. Prä MI2 vs. Post MI2 vs. Abschluss) auf den negativen Affekt in der MI-Sitzung untersucht werden. Der Box-Test auf Gleichheit der Kovarianzmatrizen wird signifikant ( $p < .001$ ). Da der Levene-Test zu allen Zeitpunkten signifikant ist ( $p < 0.01$ ), muss die Annahme der Varianzhomogenität verworfen werden und die Ergebnisse werden entsprechend angepasst. Aufgrund eines signifikanten Mauchly-Tests ( $p < .001$ ) muss davon ausgegangen werden, dass die Annahme der Sphärizität verletzt ist und die Greenhouse-Geisser-Korrektur kommt zur Anwendung. Es zeigen sich ein signifikanter Haupteffekt des Zeitpunkts ( $F(2.34, 459) = 151, p < .001, f = 0.88$ ), ein signifikanter Haupteffekt der Gruppenzugehörigkeit ( $F(4, 196) = 28.1, p < .001, f = 0.76$ ) und eine signifikante Interaktion zwischen Zeitpunkt und Gruppenzugehörigkeit ( $F(9.38, 459) = 3.40, p < .001, f = 0.26$ ). Bonferroni-korrigierte paarweise Vergleiche weisen auf signifikante Unterschiede ( $p < .05$ ) zwischen Prä MI1 und Post MI1 hin. Post MI1 sowie Post MI2 scheinen sich von allen anderen Zeitpunkten signifikant zu unterscheiden. Es finden sich keine signifikanten Unterschiede zwischen Prä MI1 und Prä MI2, sowie zwischen Prä MI1 und Abschluss. Paarweise Vergleiche der Teilnehmergruppen (Bonferroni-korrigiert) zeigen signifikante Unterschiede zwischen KG-R und UA, zwischen UR und UA sowie zwischen UA und BR. In Abbildung 5.7 sind diese Daten grafisch dargestellt. Insgesamt sprechen die Ergebnisse also für eine Erhöhung des negativen Affektes durch die Emotionsinduktionen. Die Erhöhung des negativen Affektes scheint in den Patientengruppen größer auszufallen als bei gesunden Kontrollen. Ein dauerhafter Effekt der Emotionsinduktion scheint nicht gegeben zu sein, zu den Zeitpunkten nach der Durchführung der CAM beziehungsweise des MASC ist keine vermehrt negative Stimmung mehr messbar.

Eine analoge Mixed-Design ANOVA wird für den Einfluss der Gruppenzugehörigkeit (fünfstufig) und des Zeitpunkts (fünfstufig) auf den positiven Affekt in der MI-Sitzung durchgeführt. Der Box-Test auf Gleichheit der Kovarianzmatrizen ist signifikant ( $p = .030$ ). Bei signifikantem Mauchly-Test auf Sphärizität ( $p < .001$ ) wird die Greenhouse-Geisser Korrektur angewendet. Da der Levene-Test zu keinem Zeitpunkt signifikant wird ( $p > .05$ ) kann Varianzhomogenität angenommen werden. Sowohl der Haupteffekt des Zeitpunkts ( $F(3.32, 650) = 109, p < .001, f = 0.75$ ), als auch der Haupteffekt der Gruppenzugehörigkeit ( $F(4, 196) = 16.6, p < .001, f =$

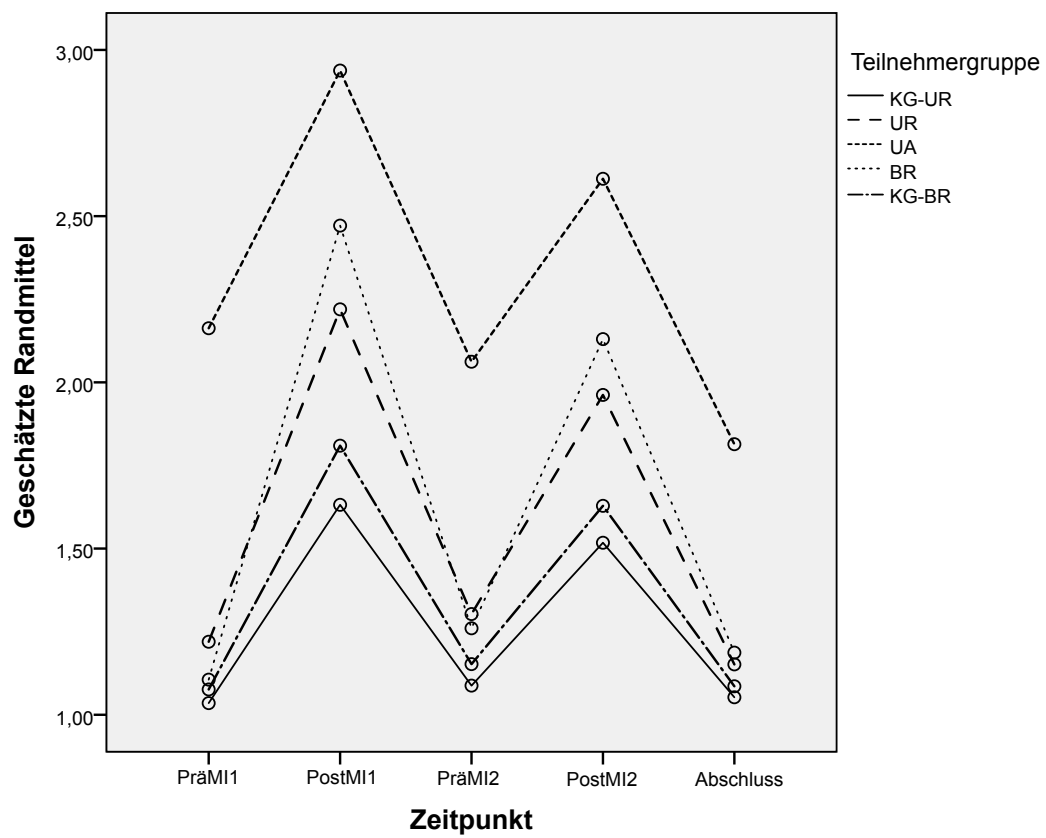


Abbildung 5.7: Negative Stimmung im Sitzungsverlauf mit Induktion in den verschiedenen Teilnehmergruppen



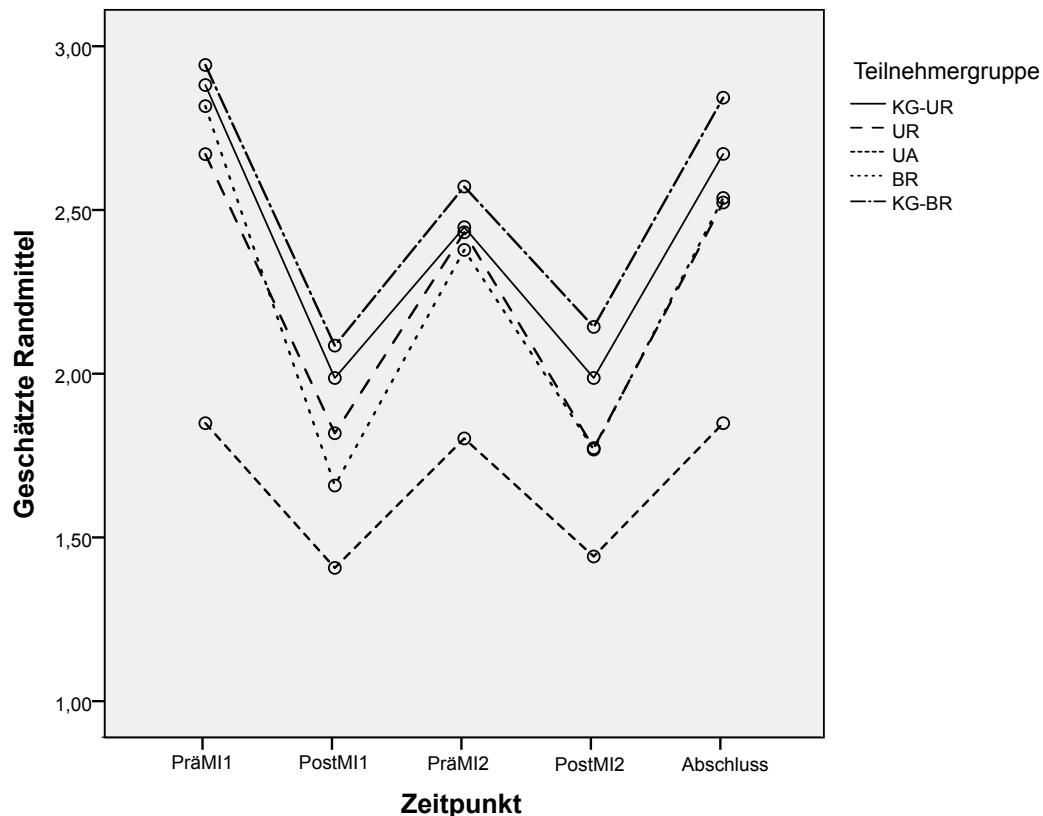


Abbildung 5.8: Positive Stimmung im Sitzungsverlauf mit Induktion in den verschiedenen Teilnehmergruppen

0.58) und Interaktion zwischen Zeitpunkt und Gruppenzugehörigkeit ( $F(13.3, 650) = 1.97, p = .020, f = 0.20$ ) werden signifikant. Paarweise Vergleiche (Bonferroni-korrigiert) weisen auf signifikante Unterschiede ( $p < 0.05$ ) zwischen Prä MI1, Prä MI2 und Abschluss jeweils zu allen anderen Zeitpunkten hin. Bonferroni-korrigierte paarweise Vergleiche sprechen weiterhin für signifikante Unterschiede zwischen KG-UR und UA, UR und UA sowie UA und BR. Eine grafische Darstellung der Ergebnisse ist in Abbildung 5.8 zu sehen. Die Ergebnisse sprechen für eine Verminderung des positiven Affekts durch die Emotionsinduktionen. Diese Verminderung des positiven Affekts scheint in den Patientengruppen größer auszufallen. Eine dauerhafte Veränderung des positiven Affekts scheint ähnlich wie beim negativen Affekt eher unwahrscheinlich. Nach der Durchführung der CAM beziehungsweise des MASC steigt die positive Stimmung im Vergleich zum Zeitpunkt direkt nach der Emotionsinduktion signifikant an.

### MI-Sitzung vs. NMI-Sitzung

Im Rahmen einer dreifaktoriellen Mixed-Design ANOVA wird der Einfluss der Gruppenzugehörigkeit (fünfstufig), des Zeitpunkts (dreistufig: Prä CAM vs. Prä MASC vs. Abschluss) und der Sitzung (zweistufig: MI vs. NMI) auf die negative Stimmung

untersucht. Der Box-Test auf Gleichheit der Kovarianzmatrizen ist signifikant ( $p < .001$ ). Aufgrund eines signifikanten Mauchly-Tests auf Sphärizität ( $p = .003$ ) wird wo nötig die Greenhouse-Geisser Korrektur angewendet. Die Annahme der Varianzhomogenität muss aufgrund signifikanter Levene-Tests zu allen Zeitpunkten ( $p < 0.01$ ) verworfen werden. Eine entsprechende Anpassung der Ergebnisse erfolgt. Es finden sich signifikante Haupteffekte des Zeitpunkts ( $F(1.89, 373) = 168, p < .001, f = 0.92$ ), der Sitzung ( $F(1,197) = 154, p < .001, f = 0.88$ ) sowie der Gruppenzugehörigkeit ( $F(4, 197) = 31.8, p < .001, f = 0.80$ ). Auch die Interaktionen zwischen Zeitpunkt und Gruppenzugehörigkeit ( $F(7.57, 373) = 4.76, p < .001, f = 0.31$ ), zwischen Sitzung und Gruppenzugehörigkeit ( $F(4, 197) = 2.95, p = .021, f = 0.25$ ) und zwischen Zeitpunkt und Sitzung ( $F(1.81, 357) = 112, p < .001, f = 0.75$ ) werden signifikant. Es kann keine signifikante Dreifachinteraktion zwischen Zeitpunkt, Sitzung und Gruppenzugehörigkeit festgestellt werden ( $F(7.24, 357) = 1.85, p = .074$ ). Es werden Bonferroni-korrigierte paarweise Vergleiche für alle unabhängigen Variablen durchgeführt: Signifikante Unterschiede ( $p < .05$ ) bestehen zwischen allen drei Zeitpunkten. Weiterhin zeigen sich signifikante Unterschiede zwischen KG-UR und UA, zwischen UR und UA, zwischen UA und BR sowie zwischen KG-BR und BR. Eine grafische Darstellung der Ergebnisse ist in Abbildung 5.9 zu finden. Die Ergebnisse sprechen dafür, dass die negative Stimmung in der MI-Sitzung im Vergleich zur Sitzung ohne Emotionsinduktion deutlich erhöht ist. Direkt nach der Emotionsinduktion, das heißt zum Zeitpunkt Prä CAM und Prä MASC in der MI-Sitzung ist die Stimmung signifikant negativer als zum Abschluss. Die Abschlusszeitpunkte beider Sitzungen scheinen sich nicht zu unterscheiden. Außerdem sprechen die Daten dafür, dass die negative Stimmung in den Patientengruppen ausgeprägter ist und dass die Patientengruppen stärker auf die Emotionsinduktion reagieren.

Auch für die positive Stimmung wird eine dreifaktorielle Mixed-Design ANOVA mit der Gruppenzugehörigkeit (fünfstufig), dem Zeitpunkt (dreistufig) und der Sitzung (zweistufig) als unabhängige Variablen durchgeführt. Der Box-Test auf Gleichheit der Kovarianzmatrizen ist signifikant ( $p = .002$ ). Da der Mauchly-Test auf Sphärizität signifikant wird ( $p = .006$ ), wird die Greenhouse-Geisser Korrektur an den entsprechenden Stellen angewendet. Bei nicht signifikanten Levene-Tests ( $p > .05$ ) kann die Annahme der Varianzhomogenität aufrecht erhalten werden. Die Haupteffekte des Zeitpunkts ( $F(1.90, 371) = 88.2, p < .001, f = 0.67$ ), der Sitzung ( $F(1,195) = 114, p < .001, f = 0.76$ ) sowie der Gruppenzugehörigkeit werden signifikant ( $F(4, 195) = 19.4, p < .001, f = 0.63$ ). Außerdem finden sich signifikante Interaktionen zwischen Sitzung und Gruppenzugehörigkeit ( $F(4, 195) = 3.27, p = .013, f = 0.26$ ) sowie zwischen Zeitpunkt und Sitzung ( $F(2, 390) = 41.8, p < .001, f = 0.36$ ). Die Interaktion zwischen Zeitpunkt und Gruppenzugehörigkeit weist eine Tendenz zur Signifikanz auf ( $F(7.61, 371) = 1.89, p = .063$ ). Die Dreifachinteraktion wird nicht signifikant ( $F(8, 389) = 1.47, p = .17$ ). Zur genaueren Betrachtung der Effekte werden Bonferroni-korrigierte paarweise Vergleiche durchgeführt. Es zeigen sich signifikante Unterschiede zwischen den Zeitpunkten Prä CAM und Abschluss sowie Prä MASC und Abschluss. Weiterhin können signifikante Unterschiede zwischen KG-UR und UR, zwischen UR und UA und zwischen UA und BR gefunden werden. In Abbildung 5.10 sind die Ergebnisse grafisch dargestellt. Die Daten sprechen für eine Verminderung der positiven Stimmung in der Sitzung mit Emotionsinduktion zu den

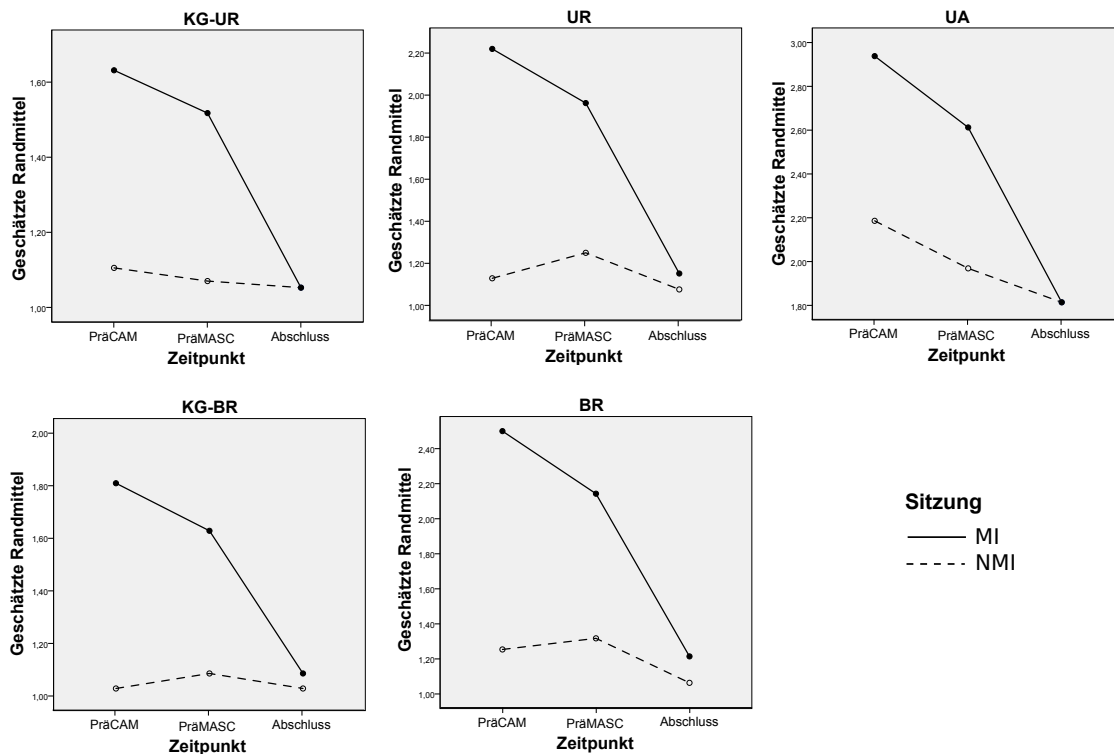


Abbildung 5.9: Negative Stimmung im Vergleich zwischen den Sitzungen in den verschiedenen Teilnehmergruppen

Zeitpunkten direkt nach der Emotionsinduktion. Zu Abschluss der Sitzungen werden keine signifikanten Unterschiede gefunden. Dies spricht dafür, dass die positive Stimmung in der MI-Sitzung auf das selbe Niveau ansteigt wie in der NMI-Sitzung. Die Verminderung der positiven Stimmung scheint in den Patientengruppen stärker auszufallen als bei gesunden Kontrollprobanden.

Zusammenfassend lässt sich also sagen, dass der negative Affekt sowohl durch die Emotionsinduktion 1 als auch 2 verstärkt zu werden scheint. Gleichzeitig sprechen die Ergebnisse für eine Verminderung des positiven Affekts durch die Emotionsinduktion 1 und 2. Die Emotionsinduktion scheint also erfolgreich zu sein. Die Daten sprechen dafür, dass die Emotionsinduktion in allen Teilnehmergruppen erfolgreich ist. Die Dauerhaftigkeit dieses Effekts hingegen ist fraglich, da zwar der positive Affekt auch nach der Durchführung der CAM und des MASC vermindert bleibt, die Verstärkung der negativen Stimmung scheint sich jedoch bis zum Ende der Aufgabe zurück zu bilden. Des Weiteren scheinen sich die Sitzungen mit und ohne Emotionsinduktion zum Abschlusszeitpunkt weder bezüglich der positiven noch der negativen Stimmung zu unterscheiden.

### 5.5.2 Einfluss der Emotionsinduktion auf die behaviorale Messung der Empathie

Im Folgenden soll der Einfluss der als erfolgreich jedoch fraglich dauerhaft anzusehenden Emotionsinduktion auf die behavioralen Maße der Empathie betrachtet

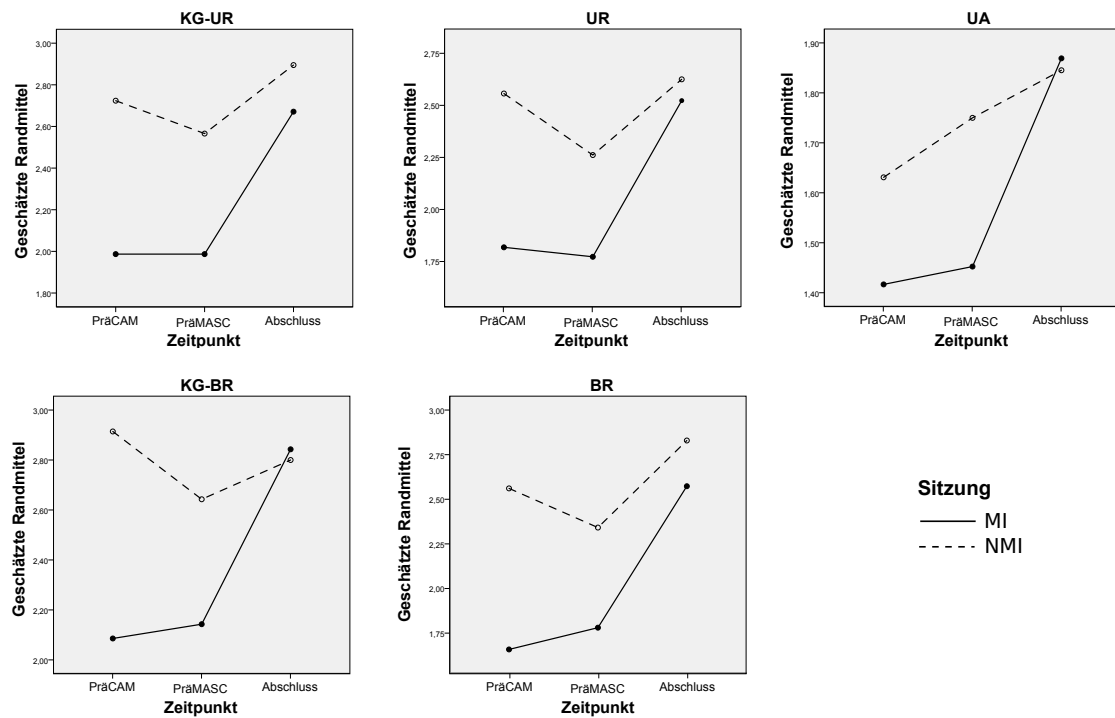


Abbildung 5.10: Positive Stimmung im Vergleich zwischen den Sitzungen in den verschiedenen Teilnehmergruppen

werden. Zunächst werden die Ergebnisse dieser experimentellen Manipulation für die CAM und anschließend für den MASC berichtet.

### Einfluss der induzierten Stimmung auf die CAM

Im Rahmen einer Mixed-Design ANOVA wird der Einfluss des Messwiederholungsfaktors Stimmungsinduktion (zweistufig) und des Zwischensubjektfaktors Gruppenzugehörigkeit (fünfstufig) auf den Anteil richtiger Antworten in der CAM untersucht. Der Box-Test auf Gleichheit der Kovarianzmatrizen wird nicht signifikant ( $p = .62$ ), es kann also von Homogenität der Kovarianzmatrizen ausgegangen werden. Aufgrund des nur zweistufigen Messwiederholungsfaktors muss der Mauchly-Test nicht berechnet werden. Der Levene-Test wird zu keinem Zeitpunkt signifikant ( $p > .05$ ), das heißt die Varianzhomogenität scheint gegeben. Es können kein signifikanter Haupteffekt der Stimmungsinduktion ( $F(1, 198) = 0.35, p = .56$ ), kein signifikanter Haupteffekt der Gruppenzugehörigkeit ( $F(4, 198) = 0.081, p = .99$ ) und keine signifikante Interaktion zwischen Stimmungsinduktion und Gruppenzugehörigkeit gefunden werden ( $F(4, 198) = 0.28, p = .89$ ). Die Ergebnisse sprechen also gegen einen bedeutsamen Einfluss der experimentellen Manipulation oder einer psychischen Erkrankung auf die behaviorale Messung der Emotionserkennung. In Abbildung 5.11 sind die Ergebnisse als Box-Plots dargestellt.

Nachdem sich der erwartete Einfluss der Stimmungsinduktion sowie der Gruppenzugehörigkeit auf den Anteil richtiger Antworten in der CAM nicht beobachten lässt, stellt sich die Frage, ob diese beiden unabhängigen Variablen einen Einfluss auf

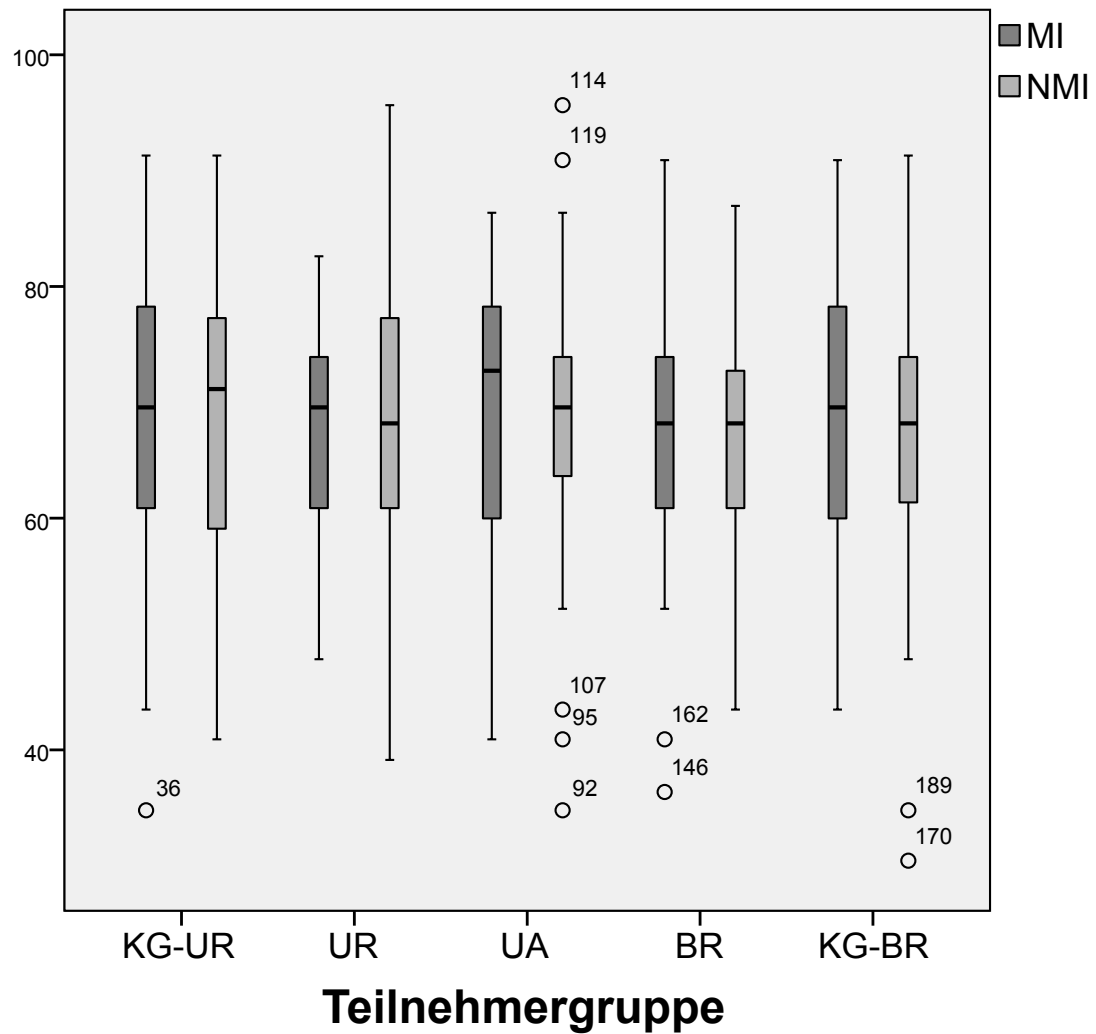


Abbildung 5.11: Einfluss der Stimmungsinduktion auf die Anzahl richtiger Antworten in der CAM in den verschiedenen Teilnehmergruppen

die Reaktionszeit, Sicherheits- oder Schwierigkeitseinschätzung haben. Diese Frage soll anhand von weiteren Varianzanalysen untersucht werden.

In einer Mixed-Design ANOVA mit der Stimmungsinduktion und der Gruppenzugehörigkeit als unabhängige Variablen wird der Einfluss auf die Reaktionszeit in der CAM betrachtet. Der Box-Test auf Gleichheit der Kovarianzmatrizen wird nicht signifikant ( $p = .39$ ), der Mauchly-Test muss nicht berechnet werden und der Levene-Test wird ebenfalls zu keinem Zeitpunkt signifikant ( $p > .1$ ). Es zeigt sich weder ein signifikanter Haupteffekt der Stimmungsinduktion ( $F(1, 198) = 1.01, p = .32$ ), noch der Gruppenzugehörigkeit ( $F(4, 198) = 0.44, p = .78$ ), noch eine signifikante Interaktion zwischen beiden unabhängigen Variablen ( $F(4, 198) = 0.21, p = .93$ ).

Eine weitere Mixed-Design ANOVA für die Sicherheitseinschätzung (Box-Test nicht signifikant:  $p = .41$ ; Mauchly-Test aufgrund zweistufigen Messwiederholungsfaktors nicht berechnet; Levene-Test signifikant für MI:  $p = .046$ , nicht signifikant für NMI:  $p > .05$ ) weist weder auf einen signifikanten Haupteffekt der Stimmungsinduktion ( $F(1, 198) = 0.018, p = .89$ ) noch auf eine signifikante Interaktion zwischen Stimmungsinduktion und Gruppenzugehörigkeit ( $F(4, 198) = 0.729, p = .573$ ) hin. Es zeigt sich jedoch ein signifikanter Haupteffekt der Gruppenzugehörigkeit auf die Sicherheitseinschätzung ( $F(4, 198) = 3.98, p = .004, f = 0.28$ ). Post-Hoc-Tests nach Gabriel sprechen für signifikante Unterschiede zwischen KG-UR und UA ( $p = .002$ ).

Die Varianzanalyse zur Untersuchung des Einflusses der Stimmungsinduktion und der Gruppenzugehörigkeit auf die Schwierigkeitseinschätzung (Box-Test signifikant:  $p = .044$ ; Mauchly-Test aufgrund zweistufigen Messwiederholungsfaktors nicht berechnet; Levene-Test zu keinem Zeitpunkt signifikant  $p > .05$ ) zeigt weder einen signifikanten Haupteffekt der Stimmungsinduktion ( $F(1, 198) = 0.006, p = .94$ ) noch eine Interaktion zwischen Stimmungsinduktion und Gruppenzugehörigkeit ( $F(4, 198) = 0.52, p = .72$ ). Es kann jedoch ein signifikanter Haupteffekt der Gruppenzugehörigkeit auf die Schwierigkeitseinschätzung beobachtet werden ( $F(4, 198) = 3.457, p = .009, f = 0.26$ ). In Post-Hoc-Tests nach Games-Howell zeigen sich signifikante Unterschiede zwischen KG-UR und UA.

Zusammenfassend lässt sich also sagen, dass die Ergebnisse keine Hinweise auf einen bedeutsamen Einfluss der Stimmungsinduktion auf die CAM geben. Die Stimmungsinduktion scheint weder die Anzahl richtiger Antworten, noch Reaktionszeit, Sicherheits- oder Schwierigkeitseinschätzung zu verändern. Die Gruppenzugehörigkeit hingegen scheint einen kleinen Einfluss auf die Schwierigkeits- und Sicherheitseinschätzung zu haben, wie schon für die CAM Gesamt beschrieben.

### **Einfluss der Stimmungsinduktion auf den MASC**

Eine Mixed-Design ANOVA mit dem Messwiederholungsfaktor Stimmungsinduktion (zweistufig) und der Gruppenzugehörigkeit (fünfstufig) soll den Einfluss dieser beiden unabhängigen Variablen auf den Anteil richtiger Antworten im MASC genauer beleuchten. Da der Box-Test auf Gleichheit der Kovarianzmatrizen nicht signifikant wird ( $p = .48$ ), kann von Homogenität der Kovarianzmatrizen ausgegangen werden. Aufgrund des nur zweistufigen Messwiederholungsfaktors muss der Mauchly-Test nicht berechnet werden. Der Levene-Test wird zu keinem Zeitpunkt signifikant ( $p > .05$ ), die Varianzhomogenität scheint also gegeben. Es findet sich weder ein signifi-

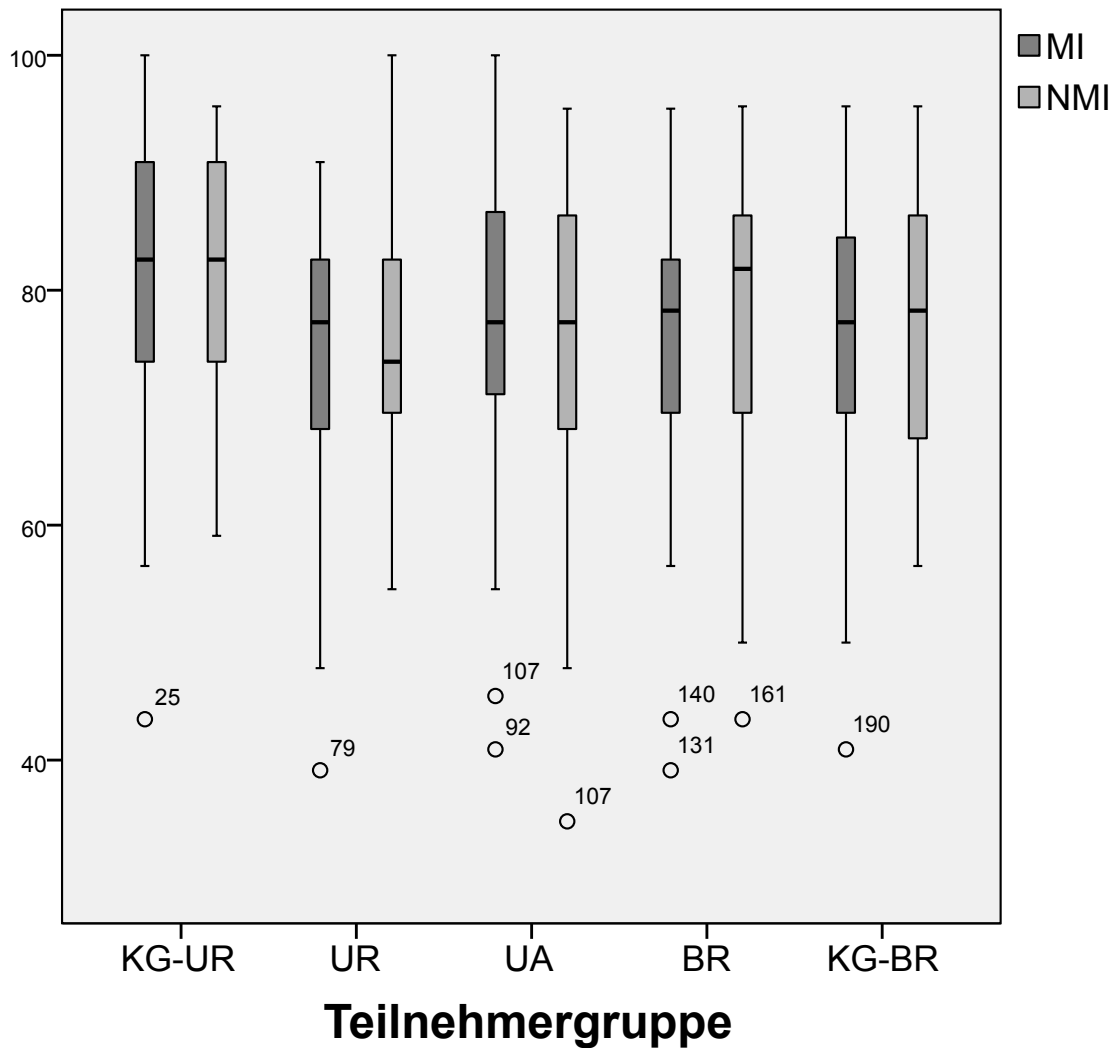


Abbildung 5.12: Einfluss der Stimmungsinduktion auf den MASC in den verschiedenen Teilnehmergruppen

kanter Haupteffekt der Stimmungsinduktion ( $F(1, 198) = 0.084, p = .77$ ) noch eine signifikante Interaktion zwischen Stimmungsinduktion und Gruppenzugehörigkeit ( $F(4, 198) = 1.02, p = .40$ ). Es zeigt sich jedoch ein signifikanter Haupteffekt der Gruppenzugehörigkeit ( $F(4, 198) = 2.47, p = .046, f = 0.23$ ). Post-Hoc Vergleiche nach Gabriel weisen daraufhin, dass der einzige signifikante Unterschied zwischen KG-UR und UR ( $p = .03$ ) besteht. In Abbildung 5.12 sind die Ergebnisse grafisch dargestellt. Anders als erwartet, scheint die Stimmungsinduktion also keinen oder kaum Einfluss auf die richtigen Antworten im MASC zu haben. Es stellt sich die Frage, ob stattdessen ein für bestimmte Fehlerarten spezifischer Einfluss festzustellen ist oder ob die verschiedenen Arten der im MASC geprüften Theory of Mind eine unterschiedliche Beeinflussung durch die Stimmungsinduktion aufweisen.

Die im MASC vorkommenden falschen Antwortmöglichkeiten lassen sich in solche ohne ToM, mit zu wenig ToM oder mit zu viel ToM aufteilen. Im Rahmen einer Mixed-Design Varianzanalyse soll der Einfluss der Stimmungsinduktion (zweistufig),

der Gruppenzugehörigkeit (fünfstufig) und der Fehlerart (dreistufig) auf den Anteil der gemachten Fehler untersucht werden. Der Box-Test auf Gleichheit der Kovarianzmatrizen wird signifikant ( $p = .046$ ). Der Mauchly-Test muss für den zweistufigen Innersubjektfaktor Stimmungsinduktion nicht berechnet werden, sowohl für die Fehlerart als auch für die Interaktion zwischen Fehlerart und Sitzung wird er signifikant ( $p < .001$ ). Der Levene-Test wird signifikant für die Fehlerart zu wenig ToM in der Sitzung ohne Stimmungsinduktion ( $p = .038$ ), ansonsten ist er nicht signifikant ( $p > .1$ ). Es zeigen sich keine signifikanten Interaktionen zwischen Stimmungsinduktion und Gruppenzugehörigkeit ( $F(4, 198) = 1.10, p = .36$ ), zwischen Sitzung und Fehlerart ( $F(1.77, 351) = 0.70, p = .48$ ), zwischen Fehlerart und Gruppenzugehörigkeit ( $F(6.76, 334) = 1.00, p = .43$ ) oder zwischen Sitzung, Fehlerart und Gruppenzugehörigkeit ( $F(7.10, 351) = 1.15, p = .33$ ). Es können aber signifikante Haupteffekte der Fehlerart ( $F(1.69, 334) = 69.5, p < .001, f = 0.59$ ) und der Gruppenzugehörigkeit ( $F(4, 198) = 2.43, p = .049, f = 0.22$ ) gefunden werden. Ein Haupteffekt der Stimmungsinduktion ist hingegen nicht feststellbar ( $F(1, 198) = 0.12, p = .73$ ). Post-Hoc werden Bonferroni-korrigierte paarweise Vergleiche durchgeführt. Dabei sprechen die Daten für signifikante Unterschiede zwischen KG-UR und UR, außerdem zwischen den Fehlerarten keine ToM und zu wenig ToM und zwischen keine ToM und zu viel ToM. Die Fehlerarten zu wenig und zu viel ToM scheinen sich nicht signifikant zu unterscheiden. In Abbildung 5.13 sind die Ergebnisse für die drei Fehlerarten dargestellt. Wie dort zu erkennen ist, werden Fehler ganz ohne ToM seltener gemacht als solche mit zu wenig oder zu viel ToM.

Im Rahmen des MASC werden Items erhoben, die eher der emotionalen ToM – also dem Decoding – zuzuordnen sind und solche die eher kognitive ToM – das heißt Reasoning-Fähigkeiten – erfordern. Im Rahmen einer Mixed-Design ANOVA sollen der Einfluss der Stimmungsinduktion (zweistufig), der Gruppenzugehörigkeit (fünfstufig) und der Art der ToM (zweistufig) auf den Anteil richtiger Antworten untersucht werden. Der Box-Test auf Gleichheit der Kovarianzmatrizen wird nicht signifikant ( $p = .24$ ). Der Mauchly-Test muss nicht berechnet werden, da sowohl Sitzung als auch Art der ToM nur zweistufig sind. Der Levene-Test wird ausschließlich für die Anzahl richtiger Antworten bei Items mit kognitiver ToM in der Sitzung ohne Stimmungsinduktion signifikant ( $p = .022$ ). Es werden keine signifikanten Interaktionen zwischen Stimmungsinduktion und Gruppenzugehörigkeit ( $F(4, 198) = 0.53, p = .72$ ), zwischen Art der ToM und Gruppenzugehörigkeit ( $F(4, 198) = 1.23, p = .30$ ) oder zwischen Stimmungsinduktion und Art der ToM ( $F(1, 198) = 0.20, p = .17$ ) noch eine Dreifach-Interaktion zwischen allen Faktoren ( $F(4, 198) = 0.57, p = .69$ ) festgestellt. Es zeigt sich weder ein signifikanter Haupteffekt der Gruppenzugehörigkeit ( $F(4, 198) = 1.64, p = .17$ ) noch der Stimmungsinduktion ( $F(1, 198) = 1.43, p = .23$ ). Einzig ein signifikanter Haupteffekt der Art der ToM ( $F(1, 198) = 6251, p < .001, f = 5.59$ ) scheint zu bestehen. Wie in Abbildung 5.14 zu sehen ist, ist der Anteil der richtigen Antworten für die emotionale höher als für die kognitive ToM.

Insgesamt scheint die Stimmungsinduktion also keinen Einfluss auf die Ergebnisse im MASC zu haben. Auch bei näherer Betrachtung der unterschiedlichen Fehlerarten oder Arten der ToM sprechen die Daten nicht für eine Veränderung der mit dem MASC gemessenen Empathie durch die Induktion negativer Stimmung.



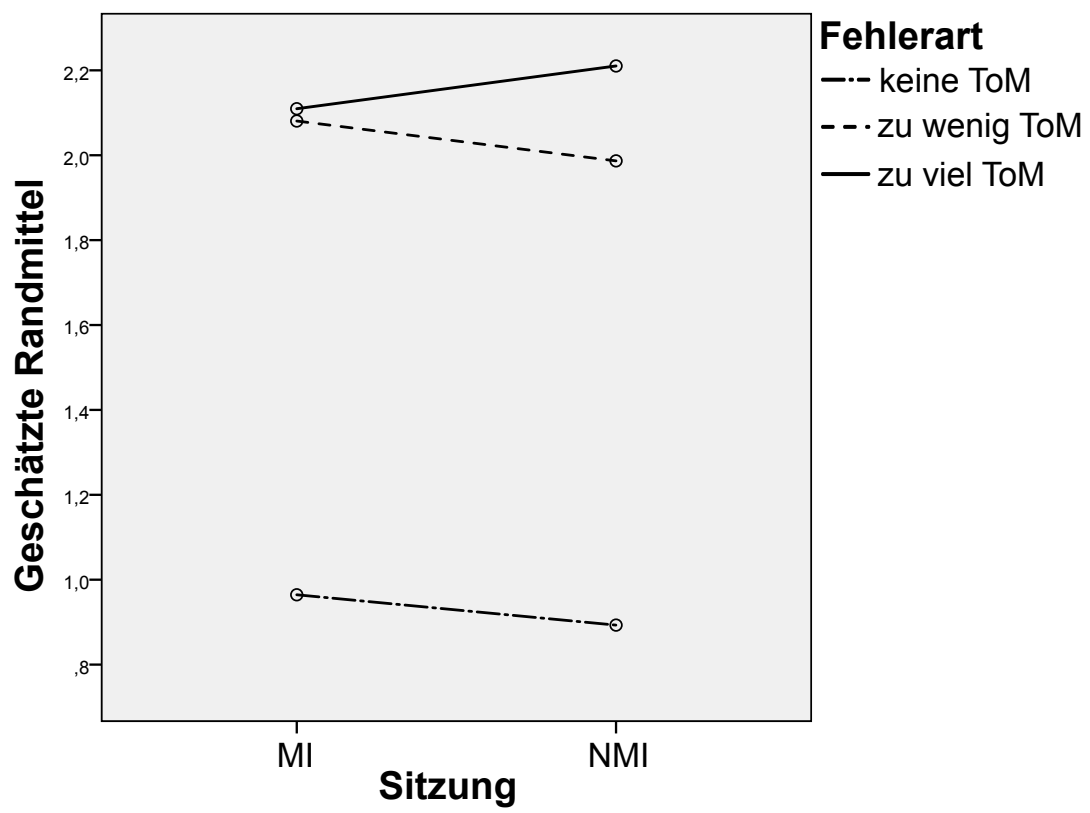


Abbildung 5.13: Einfluss der Stimmungsinduktion auf die Fehlerarten des MASC

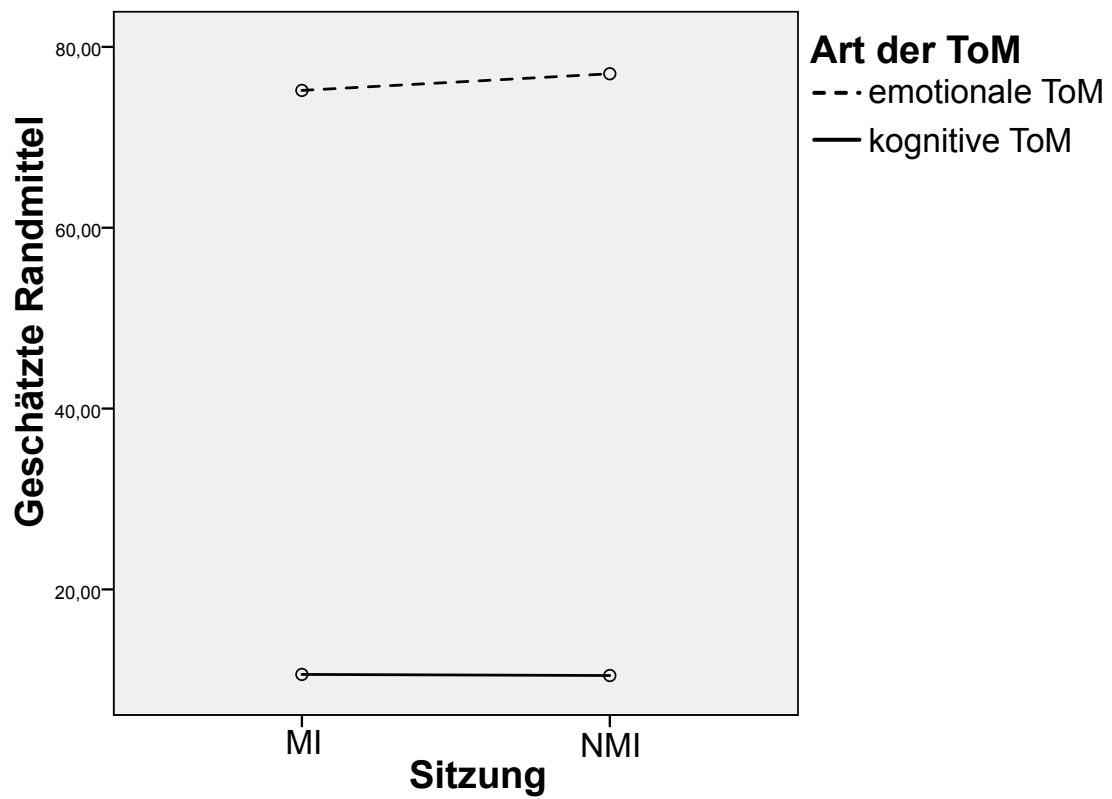


Abbildung 5.14: Einfluss der Stimmungsinduktion auf die vom MASC gemessenen Arten der ToM

### 5.5.3 Interindividuelle Unterschiede der Induzierbarkeit negativer Emotionen

Im Folgenden soll der Frage nachgegangen werden, ob es interindividuelle Unterschiede bezüglich der Induzierbarkeit negativer Emotionen gibt. Unterscheiden sich beispielsweise Personen mit oder ohne psychische Erkrankungen in der Reaktion auf die Induktion negativer Stimmung? Oder hängt das Ausmaß der Induzierbarkeit negativer Emotionen von der mit dem EQ oder anderen Fragebögen gemessenen Empathie ab? Spielt die Differenz zwischen Eigen- und Selbstwahrnehmung bei der Einschätzung der Empathie eine Rolle für die Induzierbarkeit negativer Emotionen?

Zur Untersuchung dieser Fragen wird zunächst das Ausmaß der Induzierbarkeit von Emotionen berechnet. Dafür werden sowohl für die negative als auch für die positive Stimmung die Differenzen, die durch die Stimmungsinduktion entstehen, betrachtet. Der Wert nach der Stimmungsinduktion (post) wird vom Wert vor der Stimmungsinduktion subtrahiert. Für die negative Stimmung bedeuten somit negative Werte einen Anstieg der negativen Stimmung, das heißt einen Einfluss der Emotionsinduktion wie er zu erwarten ist. Bei der positiven Stimmung weisen positive Differenzwerte auf einen Abfall der positiven Stimmung hin, was ebenfalls dem entspricht, was durch die Emotionsinduktion zu erwarten ist. Je negativer der Differenzwert der negativen Stimmung und je positiver der Differenzwert der positiven Stimmung, desto höher ist das Ausmaß der Induzierbarkeit negativer Emotionen.

#### **Induzierbarkeit negativer Emotionen bei Patienten mit psychischen Erkrankungen und bei gesunden Kontrollen.**

Mit einer multivariaten Varianzanalyse soll der Einfluss der Gruppenzugehörigkeit (fünfstufig) auf die beiden abhängigen Variablen, nämlich die Differenzwerte der positiven und negativen Stimmung, untersucht werden. Anhand des Shapiro-Wilk-Tests ist die Normalverteilungsannahme abzulehnen ( $p < .05$ ) für die positive Stimmung in den Gruppen KG-UR, UR, UA und KG-BR sowie für die negative Stimmung in den Gruppen KG-UR, UR, BR und KG-BR. Der Box-Test auf Gleichheit der Kovarianzmatrizen wird signifikant ( $p < .001$ ), die Homogenität der Kovarianzmatrizen kann also nicht als gegeben angenommen werden. Der Levene-Test auf Gleichheit der Fehlervarianzen wird nicht signifikant für positive Stimmung ( $p = .11$ ), jedoch für negative Stimmung ( $p = .003$ ). Bei der Verwendung der Pillai-Spur (ebenso wie Wilks-Lambda, Hotelling-Spur und Größte charakteristische Wurzel nach Roy) zeigt sich ein signifikanter Effekt der Gruppenzugehörigkeit auf die Differenz der positiven und negativen Stimmung ( $V = 0.17$ ,  $F(8, 396) = 4.70$ ,  $p < .001$ ,  $f^2 = 0.095$ ). Post-Hoc werden in univariaten Varianzanalysen (Tests der Zwischensubjekteffekte) sowohl für die Differenz positiver Stimmung ( $F(4, 198) = 4.75$ ,  $p = .001$ ,  $f = 0.30$ ) als auch für die Differenz negativer Stimmung ( $F(4, 198) = 5.73$ ,  $p < .001$ ,  $f = 0.35$ ) signifikante Haupteffekte sichtbar. Post-Hoc Tests mit Games-Howell weisen für die Differenz positiver Stimmung auf signifikante Unterschiede zwischen KG-UR und UA ( $d = 0.64$ ) und zwischen UA und BR ( $d = 1.01$ ) sowie für die Differenz negativer Stimmung zwischen UA und BR ( $d = 0.69$ ) und zwischen BR und KG-BR ( $d = 0.80$ ) hin. In Abbildung 5.15 sind die Ergebnisse als Boxplots dargestellt.

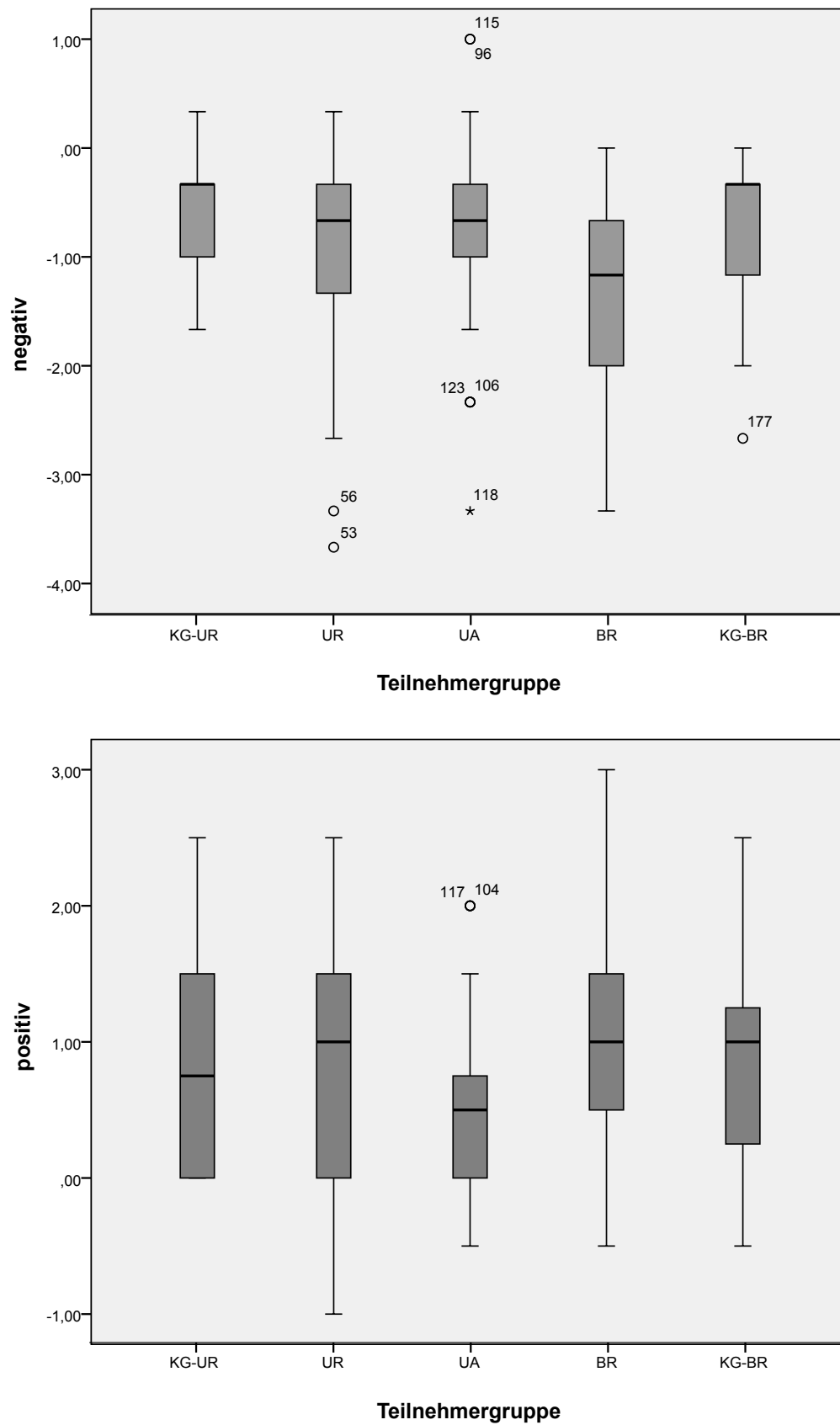


Abbildung 5.15: Einfluss der Stimmungsinduktion in den unterschiedlichen Teilnehmergruppen: Ausmaß der Differenz positiver sowie negativer Stimmung

### Induzierbarkeit negativer Emotionen und das Ausmaß der in Fragebögen gemessenen Empathie

Zur Untersuchung des Zusammenhangs der Induzierbarkeit negativer Emotionen mit den Ergebnissen im EQ und anderen Fragebögen zur Empathie sowie mit der Differenz zwischen Selbst- und Fremdbild werden zunächst signifikante Korrelationen ( $p < .05$ ) berichtet:

Die Differenzwerte der positiven und negativen Stimmung durch die Emotionsinduktion korrelieren mittelstark negativ untereinander ( $\rho = -.46$ ,  $n = 203$ ,  $p < .001$ ), was dafür spricht, dass es durch die Stimmungsinduktion sowohl zu einem Abfall positiver als auch einem Anstieg negativer Stimmung kommt. Die EQ-F Summe korreliert schwach positiv mit dem Differenzwert negativer Stimmung ( $\rho = .148$ ,  $n = 180$ ,  $p = .047$ ). Die Differenz der Selbst- und Fremdeinschätzung anhand des EQ korreliert schwach negativ mit der Differenz negativer Stimmung durch die Emotionsinduktion ( $\rho = -.17$ ,  $n = 178$ ,  $p = .026$ ).

Zur Erklärung der Varianz der Differenz positiver Stimmung wird eine lineare Regressionsanalyse durchgeführt. Als Prädiktoren werden die Summen des EQ-S und des EQ-F, die Differenz zwischen Selbst- und Fremdeinschätzung im EQ, das Geschlecht, das Vorliegen einer psychischen Erkrankung und der aktuelle BDI herangezogen. Die Korrelationsmatrix zeigt, dass keine Prädiktoren zu hoch miteinander korrelieren, es liegt also keine Multikollinearität vor. Mit einer Durbin-Watson-Statistik von 2.38 ist die Annahme der unabhängigen Fehler aufrecht zu erhalten. Es ergibt sich ein  $R^2 = 0.098$  ( $p = .003$ ,  $f^2 = 0.022$ ), das heißt die Prädiktoren erklären 9.8% der Varianz der Differenz positiver Stimmung durch die Emotionsinduktion. Die Differenz zwischen Selbst- und Fremdeinschätzung im EQ wird dabei aus dem Modell ausgeschlossen. In Tabelle 5.14 sind die Ergebnisse für die einzelnen Prädiktoren dargestellt. Wie dort zu sehen ist, trägt einzig der aktuelle BDI signifikant zur Varianzaufklärung bei.

Tabelle 5.14: Ergebnisse der multiplen Regression mit sechs Prädiktoren für die Differenz der positiven Stimmung durch Stimmungsinduktion

	<i>B</i>	<i>SE B</i>	$\beta$
(Konstante)	0.84	0.32	
EQ-S Summe	0.002	0.007	0.030
EQ-F Summe	-0.008	0.005	-0.13
EQ Selbst-Fremd			
Geschlecht	0.18	0.14	0.10
Patient	0.17	0.14	0.10
BDI	-0.022	0.006	-0.31 ***

Signifikanzniveaus: +  $p < .1$ , \*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$

Auch für die Differenz negativer Stimmung durch die Emotionsinduktion wird eine lineare Regression mit den Prädiktoren EQ-S und EQ-F Summen, Differenz zwischen Selbst- und Fremdeinschätzung anhand des EQ, Geschlecht, Vorliegen einer psychischen Erkrankung und aktueller BDI durchgeführt. Keine der Prädiktoren korrelieren zu hoch miteinander, es liegt also keine Multikollinearität vor. Bei einer

Durbin-Watson-Statistik = 2.03 ist die Annahme der unabhängigen Fehler aufrecht-erhaltbar. Mit  $R^2 = 0.12$  ( $p < .001$ ,  $f^2 = 0.14$ ), erklären die Prädiktoren 12.4% der Varianz der Differenz negativer Stimmung durch die Emotionsinduktion. Erneut wird die Differenz zwischen Selbst- und Fremdeinschätzung aus dem Modell ausgeschlossen. In Tabelle 5.15 sind die Ergebnisse der einzelnen Prädiktoren zu finden. Der aktuelle BDI, aber auch das Vorliegen einer psychischen Erkrankung und die EQ-F Summe erbringen einen signifikanten Beitrag zur Varianzaufklärung.

Tabelle 5.15: Ergebnisse der multiplen Regression mit sechs Prädiktoren für die Differenz der negativen Stimmung durch Stimmungsinduktion

	<i>B</i>	<i>SE B</i>	$\beta$
(Konstante)	-0.26	0.35	
EQ-S Summe	-0.011	0.007	-0.13
EQ-F Summe	0.013	0.005	0.19 *
EQ Selbst-Fremd			
Geschlecht	-0.25	0.16	-0.12
Patient	-0.57	0.15	-0.31 ***
BDI	0.016	0.006	0.19 *

Signifikanzniveaus: +  $p < .1$ , \*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$

Zusammenfassend lässt sich feststellen, dass die Induzierbarkeit negativer Emotionen in Zusammenhang mit der aktuellen Stimmung und dem Vorliegen einer aktuellen oder remittierten psychischen Erkrankung stehen könnte. Es gibt wenige bis keine Hinweise auf eine Beeinflussung der Induzierbarkeit negativer Emotionen durch die Empathie.

## 5.6 *Fragestellung E6: Zusammenhänge zwischen Empathie und Social Functioning*

Im Folgenden sollen Zusammenhänge zwischen Fragebogenmessungen der Empathie, Fragebogenmaßen des Social Functioning und der behavioralen Messung der Empathie betrachtet werden. Aber auch Zusammenhänge mit neuropsychologischen Tests werden beschrieben.

### 5.6.1 Empathie, Social Functioning und Neuropsychologie

Ein zentrales Maß der neuropsychologischen Testung ist der MWT-B zur Schätzung der prämorbidem Intelligenz. Im Folgenden soll genauer betrachtet werden ob es Zusammenhänge zwischen Intelligenz und Empathie sowie Social Functioning gibt.

Die MWT-B Summe, anhand derer der prämorbidem Intelligenzquotient geschätzt wird, korreliert weder mit der Summe des EQ-S noch des EQ-F. Es scheint also kein Zusammenhang zwischen der Selbst- oder Fremdeinschätzung der Empathie und der Intelligenz zu bestehen. In Tabelle 5.16 sind die Korrelationen der EQ-Summen mit

der MWT-B-Summe für die einzelnen Teilnehmergruppen sowie für die Gesamtgruppe dargestellt. Wie dort zu sehen ist, bestehen einzig zwischen der Fremdbeurteilung und der MWT-B-Summe in den Gruppen KG-UR und BR signifikante Korrelationen.

Tabelle 5.16: Korrelationen zwischen EQ-S und EQ-F mit der MWT-B Summe

	$\rho_{EQ-S}$	$\rho_{EQ-F}$
KG-UR	-.054	.35*
UR	-.032	-.19
UA	.074	-.009
BR	-.15	-.46**
KG-BR	-.13	.16
Gesamt	-.011	-.043

Signifikanzniveaus: +  $p < .1$ , \*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$

Auch die behaviorale Messung der Empathie durch CAM und MASC zeigt in der Gesamtgruppe keinerlei signifikante Korrelationen mit der durch den MWT-B geschätzten prämorbidem Intelligenz. In Tabelle 5.17 sind die Korrelationsquotienten für die einzelnen Teilnehmergruppen aufgeführt. Einzig in KG-BR gibt es Hinweise auf eine signifikant positive Korrelation zwischen der Intelligenz und der mit dem MASC gemessenen Empathie.

Tabelle 5.17: Korrelationen zwischen CAM und MASC mit der MWT-B Summe

	$\rho_{CAM}$	$\rho_{MASC}$
KG-UR	.035	-.088
UR	.049	.22
UA	.14	.14
BR	.14	-.12
KG-BR	-.007	.35*
Gesamt	.051	.11

Signifikanzniveaus: +  $p < .1$ , \*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$

In Tabelle 5.18 sind die Korrelationskoeffizienten für die Hauptmaße des Social Functioning mit der durch den MWT-B gemessenen prämorbidem Intelligenz aufgeführt. Die mit dem IIP gemessenen gemittelten interpersonellen Probleme korrelieren weder in der Gesamtgruppe noch in einer der einzelnen Teilnehmergruppen signifikant mit der Intelligenz. Auch für die wahrgenommene soziale Unterstützung gemessen mit dem F-SozU zeigen sich keine signifikanten Korrelationen. Die soziale Belastung gemessen mit dem F-SozU hingegen korreliert in der Gesamtgruppe leicht negativ mit der Intelligenz, in der Gruppe der Bipolar Remittierten sogar mittelstark negativ.

Tabelle 5.18: Korrelationen zwischen F-SozU (wahrgenommene soziale Unterstützung = Unt; soziale Belastung = Bel) und IIP (gemittelte interpersonelle Probleme) mit der MWT-B Summe

	$\rho_{\text{Unt}}$	$\rho_{\text{Bel}}$	$\rho_{\text{IIP}}$
KG-UR	-.20	-.28 <sup>+</sup>	-.28 <sup>+</sup>
UR	-.12	.054	.21
UA	-.065	.051	-.045
BR	-.039	-.36 <sup>*</sup>	-.14
KG-BR	.057	.13	.12
Gesamt	.019	-.14 <sup>*</sup>	-.14

Signifikanzniveaus: <sup>+</sup>  $p < .1$ , \*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$

Insgesamt sprechen die Zusammenhänge des MWT-B mit CAM und MASC oder F-SozU und IIP nicht für einen bedeutsamen Einfluss der Intelligenz auf die Empathie oder das Social Functioning. Lediglich einzelne Untergruppen weisen signifikante Korrelationen auf.

## 5.6.2 EQ und Social Functioning

### IIP

Die mit dem IIP gemessenen interpersonalen Probleme korrelieren in der Gesamtgruppe leicht negativ sowohl mit der Summe des EQ-S als auch des EQ-F. In Tabelle 5.19 sind die Ergebnisse für die einzelnen Teilnehmergruppen aufgetragen. Wie dort zu sehen ist, bestehen leichte bis mittelstarke negative Zusammenhänge zwischen den Ergebnissen der Messung der Empathie mit dem EQ und der Messung interpersonaler Probleme mit dem IIP. Signifikante Korrelationen finden sich nur zwischen der Fremdbeurteilung des EQ und den gemittelten interpersonalen Problemen in den Gruppen KG-UR, UR und UA.

Tabelle 5.19: Korrelationen zwischen EQ-S und EQ-F mit den gemittelten interpersonalen Problemen im IIP

	$\rho_{\text{EQ-S}}$	$\rho_{\text{EQ-F}}$
KG-UR	-.21	-.38 <sup>*</sup>
UR	-.27 <sup>+</sup>	-.40 <sup>*</sup>
UA	-.30 <sup>+</sup>	-.44 <sup>**</sup>
BR	-.10	-.023
KG-BR	-.30 <sup>+</sup>	.25
Gesamt	-.26 <sup>**</sup>	-.24 <sup>**</sup>

Signifikanzniveaus: <sup>+</sup>  $p < .1$ , \*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$



**F-SozU**

In Tabelle 5.20 sind die Korrelationen zwischen EQ-S und EQ-F mit der mit dem F-SozU gemessenen sozialen Unterstützung (Summe der drei Unterstützungsskalen) und der sozialen Belastung zu lesen. Für die Gesamtgruppe finden sich mittelstarke positive Korrelationen zwischen sozialer Unterstützung und EQ. Aufgeteilt nach Teilnehmergruppen schwankt die Stärke der Korrelation erheblich. Für die soziale Belastung ergibt sich ein leicht negativer Zusammenhang in der Gesamtgruppe. In den einzelnen Teilnehmergruppen sind keine bis leicht negative Korrelationen zwischen EQ und sozialer Belastung zu beobachten.

Tabelle 5.20: Korrelationen zwischen EQ-S und EQ-F mit dem F-SozU

<b>Soziale Unterstützung</b>	$\rho_{EQ-S}$	$\rho_{EQ-F}$
KG-UR	.40*	.12
UR	.54**	.44**
UA	.16	.40*
BR	.16	.23
KG-BR	.51**	.31
Gesamt	.39**	.33**
<b>Soziale Belastung</b>	$\rho_{EQ-S}$	$\rho_{EQ-F}$
KG-UR	-.15	-.21
UR	-.19	-.24
UA	.05	-.28 <sup>+</sup>
BR	-.027	-.016
KG-BR	-.13	-.19
Gesamt	-.13 <sup>+</sup>	-.18*

Signifikanzniveaus: <sup>+</sup>  $p < .1$ , \*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$

**5.6.3 EQ und behaviorale Messung der Empathie**

In der Gesamtgruppe aller Teilnehmer zeigen sich leichte bis mittelstarke Korrelationen zwischen EQ und behavioraler Messung der Empathie durch CAM und MASC. In den einzelnen Teilnehmergruppen sind jedoch nur vereinzelt signifikante Korrelationen zu beobachten. Die Ergebnisse sind in Tabelle 5.21 detailliert dargestellt.

Tabelle 5.21: Korrelationen zwischen EQ-S und EQ-F mit der CAM und dem MASC

<b>CAM</b>	$\rho_{\text{EQ-S}}$	$\rho_{\text{EQ-F}}$
KG-UR	.30 <sup>+</sup>	.10
UR	.29 <sup>+</sup>	.33*
UA	.19	.31 <sup>+</sup>
BR	-.17	.002
KG-BR	.15	.083
Gesamt	.16*	.21**
<b>MASC</b>	$\rho_{\text{EQ-S}}$	$\rho_{\text{EQ-F}}$
KG-UR	.25	.13
UR	.18	.18
UA	.19	.26
BR	-.090	.034
KG-BR	.16	.42*
Gesamt	.18*	.22**

Signifikanzniveaus: <sup>+</sup>  $p < .1$ , \*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$

#### 5.6.4 Behaviorale Messung der Empathie und Social Functioning

In Tabelle 5.22 finden sich die Korrelationen zwischen CAM, MASC und der mit dem F-SozU gemessenen wahrgenommenen sozialen Unterstützung sowie sozialen Belastung. In der Gesamtgruppe scheinen leicht positive Korrelationen zwischen Unterstützung und CAM sowie MASC zu bestehen, während für die Belastung keine signifikanten Korrelationen beobachtet werden können.

Tabelle 5.22: Korrelationen zwischen CAM und MASC mit dem F-SozU

<b>Soziale Unterstützung</b>	$\rho_{CAM}$	$\rho_{MASC}$
KG-UR	.044	.10
UR	.40**	.050
UA	.21	.22
BR	.11	.007
KG-BR	-.092	.43*
Gesamt	.14*	.19**
<b>Soziale Belastung</b>	$\rho_{CAM}$	$\rho_{MASC}$
KG-UR	-.067	-.212
UR	-.15	.063
UA	-.16	-.092
BR	-.043	.048
KG-BR	-.032	-.27
Gesamt	-.081	-.11

Signifikanzniveaus: +  $p < .1$ , \*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$

Die mit dem IIP erhobenen gemittelten interpersonalen Probleme scheinen in der Gesamtgruppe nicht mit der CAM und nur leicht negativ mit dem MASC zu korrelieren. In der Gruppe der Unipolar Akuten zeigen sich jedoch starke negative Korrelationen zwischen interpersonalen Problemen und der mit CAM und MASC gemessenen Empathiefähigkeiten, wie in Tabelle 5.23 zu sehen ist.

Tabelle 5.23: Korrelationen zwischen CAM und MASC mit den gemittelten interpersonalen Problemen im IIP

	$\rho_{CAM}$	$\rho_{MASC}$
KG-UR	-.21	.084
UR	-.12	-.26
UA	-.50**	-.54***
BR	.14	-.062
KG-BR	.042	.009
Gesamt	-.080	-.19*

Signifikanzniveaus: +  $p < .1$ , \*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$

## 5.7 Zusammenfassung der Ergebnisse der Experimentalstudie

Zum Abschluss des vorliegenden Kapitels sollen die anfangs gestellten Fragen wieder aufgegriffen und soweit möglich beantwortet werden:

- E1) Die Patienten mit akuter und remittierter affektiver Störung scheinen im Vergleich zu gesunden Kontrollen Defizite im Social Functioning aufzuweisen: Sowohl im IIP als auch im F-SozU unterscheiden sich die verschiedenen Teilnehmergruppen signifikant.
- E2) Es finden sich keine Hinweise auf Unterschiede bezüglich der Selbst- und Fremdeinschätzung der Empathie zwischen Patienten mit akuter und remittierter affektiver Störung und gesunden Kontrollen.
- E3) Die Diskrepanz der mit dem EQ gemessenen Selbst- und Fremdeinschätzung der Empathie scheint unabhängig von dem Vorliegen oder Fehlen einer akuten oder remittierten affektiven Störung zu sein. Die Teilnehmergruppen weisen keine signifikanten Unterschiede bezüglich der Differenz zwischen Selbst- und Fremdeinschätzung anhand des EQ auf.
- E4) Bei der behavioralen Messung der Empathie anhand der CAM und des MASC können wenig bis keine Unterschiede zwischen Patienten und Kontrollen beobachtet werden. Es finden sich leichte bis mittlere Korrelationen zwischen der Fragebogenmessung anhand des EQ und der behavioralen Messung der Empathie. Im Rahmen einer Regressionsanalyse tragen die Ergebnisse der Fragebogenmessung nicht signifikant zur Varianzaufklärung der behavioralen Messung bei.
- E5) Die Emotionsinduktion scheint zu einer bedeutsamen Verschlechterung der Stimmung zu führen, die Dauerhaftigkeit dieses Effektes ist jedoch fraglich. Es ist kein Einfluss der Stimmungsinduktion auf die Ergebnisse der behavioralen Messung der Empathie beobachtbar. Weiterhin scheint die Fragebogenmessung der Empathie nicht in einem Zusammenhang mit der Induzierbarkeit negativer Emotionen zu stehen.
- E6) Die vorliegenden Daten sprechen nicht für einen Einfluss der prämorbidem Intelligenz auf Empathie oder Social Functioning. Zwischen den verschiedenen Maßen der Empathie und des Social Functioning bestehen teilweise leichte bis mittelstarke Korrelationen, die geringer ausfallen als zu erwarten sein könnte.

# Kapitel 6

## Diskussion

Im Zentrum der vorliegenden Arbeit steht die Frage nach den Zusammenhängen zwischen affektiven Störungen und Empathie: Patienten mit unipolaren Depressionen oder bipolarer Störung leiden häufig unter sozialen Schwierigkeiten. Sind verringerte Empathiefähigkeiten eine mögliche Ursache dieser Schwierigkeiten? Gerade angesichts der großen Anzahl Betroffener (Busch et al., 2013; Pini et al., 2005) und der deutlich verminderten Lebensqualität der Patienten (Rapaport et al., 2005; Michalak et al., 2005), sollte dem Zusammenhang zwischen affektiven Störungen und Empathie in der Forschung Beachtung geschenkt werden. Um diese Thematik genauer zu beleuchten, wurden zwei empirische Studien durchgeführt. Im Rahmen der Fragebogenvalidierung wurde ein neues Messverfahren zur Erfassung der Empathie entwickelt und erprobt. Dieses kam in der Experimentalstudie gemeinsam mit anderen Maßen der Empathie sowohl bei Betroffenen mit affektiven Störungen als auch bei gesunden Kontrollen zur Anwendung.

Im Folgenden werden die in beiden Teilen untersuchten Fragestellungen und die Beantwortung derselben durch die Ergebnisse zusammengefasst. Anschließend werden beide Studien kritisch beleuchtet und alternative Erklärungsmöglichkeiten sowie offene Forschungsfragen diskutiert.

### **Zusammenfassung der Ergebnisse der Fragebogenvalidierung**

Im Rahmen der ersten Studie wurde der Empathy Quotient aus dem Englischen ins Deutsche übersetzt und an einer hauptsächlich studentischen Stichprobe validiert. Neben der ursprünglichen Selbstbeurteilungsversion EQ-S kam eine neu entwickelte Fremdbeurteilungsvariante EQ-F zum Einsatz.

Die Rohwerte der beiden deutschen Fragebogenversionen sind vergleichbar mit den vom englischen Original bekannten Werten. Deskriptiv betrachtet weisen die deutschen Versionen ähnliche Mittelwerte auf, haben aber leicht größere Standardabweichungen sowie ausgeprägtere Minima und Maxima. EQ-S und -F scheinen über gute klassische psychometrische Gütekriterien zu verfügen. Sowohl Durchführungs-, Auswertungs- und Interpretationsobjektivität als auch die interne Konsistenz sind als gut anzusehen. Exploratorische Faktorenanalysen geben Hinweise dafür, dass die deutsche Übersetzung und Umformulierung dem englischen Original inhaltlich ähneln. Weiterhin können die aus der Untersuchung mit dem englischen Original bekannten Geschlechtsunterschiede auch mit dem deutschen EQ-S und -F beobach-

tet werden: Frauen scheinen über eine sowohl in Selbst- als auch in Fremdratings höher eingeschätzte Empathie zu verfügen. Im Vergleich der Ergebnisse im EQ sowie anderen Maßen der Empathie zeigen sich gute Übereinstimmungen vor allem mit anderen Fragebogenmessungen. Eine gute konvergente Validität kann also als gegeben angenommen werden. Die diskriminante Validität scheint, wie Vergleiche mit Maßen der Intelligenz und des Systemizing zeigen, ebenfalls gegeben zu sein. Es kann also davon ausgegangen werden, dass die deutsche Übersetzung sowie die Fremdbeurteilungsversion des EQ konstruktvalid sind, das heißt, dass sie tatsächlich Empathie messen. Explorativ konnten in einer Regressionsanalyse verschiedene mögliche Einflussfaktoren auf die Ergebnisse im EQ-S und -F ermittelt werden: Das Geschlecht, die Ergebnisse in einer behavioralen Messung der Emotionserkennung (RME), die Anzahl der Sozialkontakte sowie ein aktuell oder in der Vergangenheit absolviertes Psychologiestudium erklären einen bedeutsamen Anteil der Varianz des EQ.

### **Zusammenfassung der Ergebnisse der Experimentalstudie**

Die zweite Studie beinhaltete die Untersuchung von Patienten mit affektiven Störungen und gesunden Kontrollen hinsichtlich ihrer Empathiefähigkeiten sowie eine experimentelle Manipulation der aktuellen Stimmung. Im Rahmen der Experimentalstudie wurden sowohl der im ersten Teil validierte EQ als auch behaviorale Maße der Empathie eingesetzt.

In der untersuchten Stichprobe konnten wie erwartet Defizite im Social Functioning bei den Patientengruppen beobachtet werden: Die Patienten scheinen mehr interpersonale Probleme, erhöhte soziale Belastungen und weniger soziale Unterstützung zu erleben als die gematchten gesunden Kontrollen. Im EQ jedoch finden sich keine Hinweise auf Unterschiede bezüglich der Selbst- und Fremdeinschätzung des Ausmaßes der Empathie von Patienten und gesunden Kontrollen. Die Ergebnisse weisen auf eine in allen Teilnehmergruppen im Vergleich zur Fremdbeurteilung höhere Selbsteinschätzung der Empathie hin. Die Diskrepanz der Selbst- und Fremdeinschätzung im EQ scheint jedoch unabhängig von dem Vorliegen oder Fehlen einer akuten oder remittierten affektiven Störung zu sein. Auch in der behavioralen Messung der Empathie können kaum bis keine Unterschiede zwischen Betroffenen und gesunden Kontrollen beobachtet werden: Die Fähigkeiten zur Emotionserkennung und zum Schlussfolgern über mentale Zustände anderer Personen scheinen bei Patienten mit affektiven Störungen nicht niedriger oder höher ausgeprägt zu sein als bei Gesunden. Die Zusammenhänge zwischen der Fragebogenmessung und den Verhaltensdaten scheinen nur leicht bis mittelstark ausgeprägt zu sein, was dafür spricht, dass die Verfahren unterschiedliche Aspekte des weitreichenden Konstruktes der Empathie abdecken. Es erfolgte eine experimentelle Manipulation der aktuellen Stimmung vor der Durchführung der behavioralen Messung der Empathie. Die experimentelle Manipulation kann als erfolgreich angesehen werden, da sie zu einer bedeutsamen Verschlechterung der Stimmung geführt zu haben scheint. Die Dauerhaftigkeit der Manipulation muss jedoch als fraglich eingeschätzt werden, da die Stimmung nach Abschluss der Tests jeweils wieder auf dem Ausgangsniveau angekommen war. Ein Einfluss der Stimmungsinduktion auf die Ergebnisse bezüglich der Emotionserkennung oder dem Schlussfolgern über mentale Zustände anderer

Personen konnte nicht beobachtet werden. Die Induzierbarkeit negativer Emotionen scheint in den Patientengruppen ausgeprägter zu sein. Patienten mit affektiven Störungen scheinen leichter in negative Stimmung zu bringen zu sein. Es finden sich keine Hinweise auf einen Zusammenhang der Fragebogenmessung der Empathie mit dem Ausmaß der Induzierbarkeit von Emotionen durch Musik. Die vorliegenden Ergebnisse weisen nicht auf einen Einfluss der prämorbidem Intelligenz auf Empathie oder Social Functioning hin. Es können leichte bis mittelstarke Zusammenhänge zwischen den verschiedenen Maßen der Empathie und des Social Functioning beobachtet werden. Diese fallen jedoch geringer aus, als zu erwarten gewesen wäre.

## 6.1 Einordnung und Kritik der Fragebogenvalidierung

Die Ergebnisse zur Validierung der deutschen Übersetzung und Umformulierung des EQ entsprechen bezüglich der Rohwerte sowie der psychometrischen Gütekriterien dem englischen Original von Baron-Cohen & Wheelwright (2004) und Lawrence et al. (2004). Bezüglich der Faktorstruktur sowie der Detektion von Geschlechtsunterschieden mithilfe des Fragebogens zeigen sich ebenfalls große Überlappungen zwischen der vorgelegten Übersetzung sowie Umformulierung und der englischen Originalversion. Auch zu japanischen (Wakabayashi et al., 2007), französischen (Bertozzi et al., 2008) und italienischen (Preti et al., 2011) Übersetzungen des EQ oder seiner Kurzfassung werden ähnliche Ergebnisse berichtet. Die Anwendung einer chinesischen Übersetzung der Kurzform des EQ hingegen wies zwar ebenfalls zufriedenstellende psychometrische Gütekriterien auf, konnte jedoch die zu erwartenden Geschlechtsunterschiede nicht erfassen (Guan et al., 2012). Insgesamt scheint der EQ ein über Kulturen und Sprachen hinweg einsetzbares Maß der Empathie zu sein, was auch durch die vorliegenden Daten bekräftigt wird. Allison et al. (2011) haben die englische Version des EQ mittels verschiedener statistischer Herangehensweisen untersucht und sprechen sich für eine Nutzung des EQ zur Messung der Empathie auf einer einzelnen Dimension aus. Eine Nutzung des Summenscores des EQ wird also empfohlen. Die vorgelegte Arbeit folgt dieser Empfehlung. Die leichten Abweichungen der Faktorstruktur zwischen englischem Original und deutscher Übersetzung sowie die Abweichungen zwischen englischem Original und deutscher Fremdbeurteilungsversion können daher toleriert werden. Sollen jedoch einzelne Unterfaktoren des EQ betrachtet werden, kann die dem englischen Original nicht ganz entsprechende Faktorstruktur der deutschen Übersetzung und Umformulierung als Kritikpunkt genannt werden. Auch die Unterschiede der Faktorstruktur zwischen den deutschen Selbst- und Fremdbeurteilungsversionen sollten bei der Interpretation von Ergebnissen beachtet werden. Wie Mischke (2003) beschreibt, werden durch eine Fremdeinschätzung der Empathie relevante Aspekte ergänzt, die alleine in der Selbstbeurteilung nicht erfasst werden können. Dies könnte ein Erklärungsansatz für die unterschiedliche Faktorstruktur der Selbst- und Fremdbeurteilungsversion des deutschen EQ sein. In englischer Sprache existiert keine Fremdbeurteilungsversion des EQ für Erwachsene und die bestehenden Versionen für Kinder und Jugendliche (Auyeung et al., 2009, 2012) wurden nicht hinsichtlich ihrer Faktorstruktur unter-

sucht. Zukünftige Forschung sollte hier weitere empirische Daten zur Faktorstruktur des EQ in Selbst- und Fremdbeurteilungsversionen verschiedener Sprachen erheben um eine bessere Interpretation der Daten zu ermöglichen.

Die vorliegende Validierungsstudie kann aufgrund ihrer hauptsächlich studentischen Stichprobe kritisiert werden. Die Ergebnisse von Kettle et al. (2008) sprechen dafür, dass universitäre Stichproben über besonders gut ausgeprägte empathische Fähigkeiten verfügen. Zudem sind die meisten untersuchten Studenten für Psychologie eingeschrieben. Psychotherapeuten scheinen sich von gematchten Kontrollen bezüglich der Fähigkeit verbale Äußerungen zu interpretieren und durch geringere Werte persönlichen Disstresses zu unterscheiden, scheinen also zumindest in Teilbereichen eine ausgeprägtere Empathie zu haben (Hassenstab et al., 2007). Unter Psychologiestudenten könnten also ebenfalls erhöhte Empathiefähigkeiten vorhanden sein. Die Generalisierbarkeit der erhobenen Daten auf Stichproben aus der Normalbevölkerung oder auf klinische Stichproben ist daher in Frage zu stellen. Auch die hohe Anzahl an Sozialkontakten und Wohngemeinschaften sprechen für eine ausgeprägte soziale Kompetenz der vorliegenden Stichprobe, mit möglicherweise besser ausgeprägten Empathiefähigkeiten als dies in einer Stichprobe der Normalbevölkerung zu erwarten wäre. In zukünftigen Arbeiten sollte daher auf repräsentative und größere Stichproben gesetzt werden.

Anders als von Lawrence et al. (2004) empfohlen, wurde in der vorliegenden Studie keine Messung der Tendenz zu sozial erwünschten Antworten durchgeführt. Eine Messung der Tendenz zu sozial erwünschten Antworten hätte wertvolle Zusatzinformationen bieten können, da die Fremdbeurteilung durch den Probanden nahe stehende Personen möglicherweise sehr anfällig für sozial erwünschte Antworten ist. Auch die soziale Nähe zwischen Probanden und Versuchsleitern, die sich zumeist aus dem Studium kannten, könnte zu vermehrten sozial erwünschten Antworten führen. Zukünftige Arbeiten sollten daher ein Maß der sozialen Erwünschtheit mit erheben. Zur Vermeidung sozialer Erwünschtheit in der vorliegenden Arbeit könnte jedoch das Ausfüllen der Fragebögen außerhalb des Labors, das heißt ohne Anwesenheit des Versuchsleiters beziehungsweise des Probanden, beigetragen haben. Zudem gibt es Hinweise darauf, dass Selbstbeurteilungsmaße der Empathie wenig durch soziale Erwünschtheit beeinflusst werden (Baldner & McGinley, 2014).

Ein weiterer Kritikpunkt ist die Auswahl des Fremdbeurteilers durch den jeweiligen Probanden selbst. Die Auswahl vertrauter Personen, die mit dem Probanden in einer positiven sozialen Beziehung stehen, könnte dazu führen, dass die empathischen Fähigkeiten im Rahmen der Fremdeinschätzung überschätzt werden.

## **6.2 Einordnung und Kritik der Experimentalstudie**

### **6.2.1 E1: Fragebogenmaße des Social Functioning**

Die Patientienstichproben der vorliegenden Arbeit scheinen im Vergleich zu gesunden Kontrollen Defizite im Social Functioning aufzuweisen. Hierbei scheinen sowohl unipolare als auch bipolare Patienten in Remission im Vergleich zu gesunden Kontrollen unter erhöhten interpersonalen Problemen, erhöhten sozialen Belastungen und verminderter sozialer Unterstützung zu leiden. Die akut depressiven Patien-



ten scheinen sich bezüglich dieser Werte sowohl von den gesunden Kontrollen als auch von remittierten Patienten zu unterscheiden und somit das niedrigste Social Functioning aufzuweisen. Diese Ergebnisse entsprechen weitgehend den Beobachtungen von Judd et al. (2000) zu Patienten mit unipolarer Depression und von Judd et al. (2005) zu Patienten mit bipolarer Störung: Auch Judd et al. konnten Hinweise auf Einschränkungen des psychosozialen Funktionsniveaus bei Patienten mit affektiven Störungen feststellen, die von Symptomschwere und Krankheitsverlauf abhängen. Aktuelle Überblicksarbeiten gehen ebenfalls von Defiziten unipolar oder bipolar erkrankter Personen in verschiedenen Bereichen des Social Functioning aus (Kupferberg et al., 2016; Samamé et al., 2012). Es kann also davon ausgegangen werden, dass die vorliegende Stichprobe die typischen für die Erkrankungen bekannten Einschränkungen des Social Functioning aufweist und sich somit zur Untersuchung der Zusammenhänge zwischen Empathie und affektiven Störungen eignet. Zudem erbringt die vorliegende Studie weitere Evidenz für einen Zusammenhang der Social Functioning Defizite mit aktuellen Krankheitsphasen beziehungsweise der Symptomschwere von Patienten mit unipolaren Depressionen.

### 6.2.2 E2: Selbst- und Fremdeinschätzung der Empathie mit dem EQ

In der vorliegenden Arbeit können keine Hinweise auf Unterschiede bezüglich der per Fragebogen gemessenen Empathie von Patienten oder gesunden Kontrollen gefunden werden. Weder in der Selbsteinschätzung mit dem EQ-S noch in der Fremdeinschätzung anhand des EQ-F scheinen sich die verschiedenen Teilnehmergruppen voneinander zu unterscheiden. Widersprechende Ergebnisse erbrachte die Studie von Yang et al. (2017), die in der Selbstbeurteilung bipolarer Patienten mit dem IRI niedrigere Werte für empathic concern und perspective taking im Vergleich zu Gesunden beobachteten. Yang et al. untersuchten jedoch Patienten mit bipolarer Störung in verschiedenen Krankheitsphasen, während in der vorliegenden Arbeit ausschließlich euthyme Patienten mit bipolarer Störung rekrutiert werden konnten. Wilbertz et al. (2010) konnte bei chronisch depressiven Patienten Hinweise auf eine verminderte Selbsteinschätzung der Empathie finden. Thoma et al. (2011) beobachteten Hinweise auf erhöhten Distress in der Selbsteinschätzung der Empathie depressiver Patienten anhand des IRI. Berthoz et al. (2008) berichten in ihrer Arbeit hingegen eine signifikant negative Korrelation zwischen der mit dem EQ gemessenen Empathie und dem BDI. Im Rahmen der vorliegenden Arbeit jedoch erbringt die mit dem BDI gemessene Schwere der aktuellen depressiven Symptomatik keinen signifikanten Beitrag zur Varianzaufklärung des EQ-S. Die Übersichtsarbeit von Schreier et al. (2013) zu den Zusammenhängen zwischen Empathie und klinischer oder subklinischer depressiver Symptomatik zeigt, dass eine Vielzahl von Studien mit verschiedensten Messmethoden durchgeführt wurden. Bezüglich der Selbstbeurteilung der Empathie werden sowohl niedrigere als auch höhere Empathiefähigkeiten Depressiver berichtet. In einigen Studien werden jedoch auch nicht signifikante Ergebnisse beschrieben. Die Datenlage scheint also uneinheitlich. Insgesamt gehen die Autoren davon aus, dass depressive Symptome mit erhöhten Werten für empathischen Stress zusammenhängen, der durch eine erhöhte Selbstaufmerksamkeit ausgelöst sein könnte

(Schreiter et al., 2013). Der in der vorliegenden Arbeit eingesetzte EQ wird bisher wenig verwendet. Die neu formulierte Fremdbeurteilungsvariante wurde bisher noch nicht eingesetzt. Insgesamt finden sich in der bestehenden Literatur zu den Empathiefähigkeiten von Patienten mit affektiven Störungen keine Arbeiten, die auf Fremdbeurteilungsmaße setzen.

### **6.2.3 E3: Diskrepanz zwischen Selbst- und Fremdeinschätzung der Empathie**

Die Teilnehmer der Experimentalstudie scheinen sich selbst signifikant empathischer einzuschätzen als dies ihre Fremdbeurteiler tun. Diese Diskrepanz zwischen Selbst- und Fremdeinschätzung scheint unabhängig von dem Vorliegen einer affektiven Störung und von der aktuellen Symptomschwere. Aufgrund des Mangels an Arbeiten, die Fremdeinschätzungen der Empathie an Patienten mit affektiven Störungen einsetzen, können für die Einordnung dieser Ergebnisse keine bestehenden Studien herangezogen werden. Häufig kommt eine Fremdeinschätzung der Empathie nur dann zum Einsatz, wenn eine Selbsteinschätzung nicht möglich scheint – zum Beispiel bei Kindern oder Jugendlichen mit Autismus-Spektrum-Störungen (Auyeung et al., 2009, 2012). Daten zur Diskrepanz zwischen Selbst- und Fremdeinschätzung werden daher in diesen Einsatzgebieten nicht erhoben. Auch Studien zur Empathie von Krankenschwestern scheinen sich klassischerweise ausschließlich auf Fremdbeurteilungen zu stützen (z.B. Gagan, 1983). Ein direkter Vergleich der Selbst- und Fremdeinschätzung der Empathie in einer studentischen Stichprobe von Mischo (2003) gibt Hinweise darauf, dass die Fremdeinschätzung sich nicht mit der Selbsteinschätzung deckt, sondern zusätzliche Informationen bietet. Mischo empfiehlt daher die Kombination beider Ansätze sowie objektiver Maße der Empathie, wie in der vorgelegten Arbeit realisiert.

### **6.2.4 E4: Behaviorale Messung der Empathie**

Die behaviorale Messung der Empathie erfolgte in der vorliegenden Studie mit der Face-Task der CAM und dem MASC. Die verschiedenen Teilnehmergruppen scheinen sich in beiden Verfahren nicht bis kaum bezüglich der Anzahl richtiger Antworten zu unterscheiden. Für die Reaktionszeit in der CAM gibt es ebenfalls keine Hinweise auf Gruppenunterschiede. Bezüglich der Sicherheitseinschätzung sowie Schwierigkeitseinschätzung jedoch scheinen akut Depressive sich signifikant von gesunden Kontrollen zu unterscheiden: Akut depressive Patienten scheinen sich bezüglich ihrer Entscheidungen unsicherer zu sein und die Aufgabe als schwerer zu empfinden als gesunde Kontrollen. Wie eine Kontrollaufgabe mit Rechenaufgaben zeigte, scheint dieser Effekt nicht auf andere Aufgaben generalisiert werden zu können. Für den MASC gibt es keine Hinweise auf Gruppenunterschiede bezüglich der Fehlerarten oder bezüglich der Anzahl richtiger Antworten und Fehlerarten bei Unterscheidung zwischen Items mit emotionaler oder kognitiver ToM. Einzig zwischen unipolar Remittierten und gesunden Kontrollen scheint ein signifikanter Unterschied (mittelgroßer Effekt) zwischen der Anzahl richtiger Antworten zu bestehen. In einer Metaanalyse von Bora & Berk (2016) wurden bestehende Forschungsarbeiten

zusammengefasst und Hinweise auf eine Beeinträchtigung von Patienten mit Major Depression im Vergleich zu Gesunden bezüglich des Decoding sowie Reasoning berichtet. Die vorliegenden Ergebnisse aus dem MASC entsprechen nur teilweise diesen Befunden, da Hinweise auf eine Beeinträchtigung remittierter Patienten mit unipolarer Depression beobachtet wurden. Eine Beeinträchtigung akut depressiver Patienten scheint in der vorliegenden Studie jedoch nicht vorhanden zu sein. Auch die berichteten Ergebnisse zum mit der CAM gemessenen Decoding entsprechen nicht denen von Bora & Berk (2016): Keine der Patientengruppen in der vorliegenden Stichprobe scheint bezüglich des Decoding beeinträchtigt zu sein. Wolkenstein et al. (2011) konnten ebenfalls keine Unterschiede zwischen Depressiven und Gesunden bezüglich des Decoding (gemessen mit dem RME) und nur einen Trend auf Defizite bezüglich des Reasoning (gemessen mit dem MASC) beobachten. Auch die Ergebnisse von Washburn et al. (2016) sprechen nicht für ein defizitäres Decoding bei depressiven Patienten ohne komorbide Angsterkrankung. Keine Hinweise auf Unterschiede bezüglich des mit dem MASC gemessenen Reasoning zwischen chronisch Depressiven und gesunden Kontrollen werden von Wilbertz et al. (2010) berichtet. Zwick & Wolkenstein (2017) hingegen konnten Hinweise auf Defizite akut Depressiver bezüglich Decoding sowie Reasoning und auf Defizite remittiert Depressiver bezüglich Decoding finden.

Die in der vorliegenden Stichprobe untersuchten bipolaren Patienten scheinen sich bezüglich des Decoding oder Reasoning nicht von gesunden Kontrollen zu unterscheiden. Damit widersprechen die Ergebnisse denen von Olley et al. (2005) sowie (Montag et al., 2010). In beiden Studien wurden Defizite remittiert Bipolarer im Vergleich zu Gesunden beobachtet. Auch die Ergebnisse von Wolf et al. (2010) und Yang et al. (2017) sprechen für ToM-Defizite von Patienten mit bipolarer Störung. Diese beiden Arbeiten untersuchten jedoch akut erkrankte Patienten, während in der vorliegenden nur remittierte Patienten mit bipolarer Störung rekrutiert werden konnten. Eine Erklärung für die fehlende Übereinstimmung mit der schon vorhandenen Literatur könnte im Publikationsbias zu finden sein: Studien, die wie die vorliegende keine Unterschiede zwischen Patienten und Gesunden beobachten können, haben eventuell eine geringere Wahrscheinlichkeit publiziert zu werden.

Die Zusammenhänge zwischen Fragebogenmessung und behavioraler Messung der Empathie scheinen in der vorliegenden Arbeit nur leicht bis mittelstark zu sein. Im Rahmen einer Regressionsanalyse gibt es keine Hinweise darauf, dass die Varianz der CAM durch die Fragebogenmessung oder die Zugehörigkeit zu einer der Patienten-Gruppen erklärt werden könnte. Das Vorliegen einer psychischen Erkrankung scheint jedoch signifikant zur Varianzaufklärung des MASC beizutragen. Außerdem scheinen die Leistung in MASC und CAM gegenseitig einen signifikanten Beitrag zur Varianzaufklärung zu erbringen. Die verschiedenen Verfahren zur Messung der Empathie scheinen also nicht deckungsgleich zu sein (vgl. Thoma et al., 2013). Im Fragebogen und in den verschiedenen behavioralen Messmethoden werden wahrscheinlich unterschiedliche Aspekte des sehr breiten Konstrukts der Empathie erhoben. Die Vielzahl der existierenden Definitionen (Cuff et al., 2016) erschweren eine Zusammenführung verschiedener Messmethoden und deren Ergebnisse zudem.

### 6.2.5 E5: Empathie und Induzierbarkeit negativer Emotionen

Trotz scheinbar erfolgreicher Induktion negativer Stimmung, lassen sich in den vorliegenden Daten keine Hinweise auf einen Einfluss aktuell negativer Stimmung auf die behaviorale Messung der Empathie finden. Die vorliegenden Ergebnisse entsprechen daher nicht denen von Chepenik et al. (2007), die Hinweise auf eine beeinträchtigte Emotionserkennung Gesunder nach Induktion trauriger Stimmung geben. Remittierte Patienten mit rezidivierenden Depressionen scheinen zudem unter Induktion negativer Stimmung vor allem glückliche Gesichter weniger gut erkennen zu können als gesunde Kontrollen (LeMoult et al., 2009). Die durchgeführte Induktion negativer Stimmung entspricht der Vorgehensweise von Rojas et al. (2014) und scheint erfolgreich zu sein, wofür die Messung der aktuellen Stimmung direkt nach dem Anhören der traurigen Musik spricht. Wie die Messungen der aktuellen Stimmung nach Abschluss der behavioralen Aufgaben jedoch zeigen, scheint die Dauerhaftigkeit der Stimmungsinduktion fraglich zu sein. Besonders bezüglich des hier ausgewerteten ersten Blocks der CAM lässt sich nichts darüber aussagen, ob die induzierte negative Stimmung während der Durchführung noch vorherrschend war, da erst nach Abschluss aller fünf Blöcke sowie der Kontrollaufgabe CALC eine erneute Messung der Stimmung erfolgte. Die Kontrollmessung nach Abschluss des MASC hingegen erfolgte ohne Verzögerungen, was die Annahme nahelegt, dass die induzierte Stimmung während des MASC nicht dauerhaft gegeben war. Zur scheinbar nicht ausreichenden Dauer der Stimmungsinduktion könnte die Konzentration der Teilnehmer auf die mitunter positiven Emotionen während der beiden Aufgaben beigetragen haben: Wie Sanchez et al. (2014) berichten, kann die Konzentration auf glückliche Gesichtsausdrücke zu einer ausgeprägteren Erholung von einer Induktion negativer Stimmung führen.

Bei der Analyse interindividueller Unterschiede bezüglich der Induzierbarkeit negativer Emotionen ergeben sich Hinweise auf eine stärkere Reaktion der Patientenstichproben als der gesunden Kontrollen. Sowohl unipolar depressive als auch bipolare Patienten scheinen mit einer stärkeren Verschlechterung der Stimmung zu reagieren als gesunde Kontrollen. Für Patienten mit Schizophrenie ist ein gegenteiliger Effekt bekannt (F. Schneider et al., 1995). Euthyme Patienten mit bipolarer Störung scheinen hingegen stärker als Gesunde auf eine Induktion positiver Stimmung zu reagieren (Roiser et al., 2009). In einer Studie mit akut depressiven sowie remittierten Patienten mit Major Depression und gesunden Kontrollen konnten zwar keine Hinweise auf eine stärkere Reaktion der Patienten auf die Emotionsinduktion gefunden werden, die Patientengruppen scheinen sich jedoch im Gegensatz zu den Gesunden durch aktive Erinnerung positiver Ereignisse nicht von der induzierten Stimmung zu erholen (Joormann et al., 2007). Studien die, wie die vorliegende, eine stärkere Verschlechterung der Stimmung für Patienten mit affektiven Störungen im Vergleich zu Gesunden berichten, sind nicht bekannt. Wie eine Studie von Segal et al. (2006) zeigt, scheint das Thema der Induzierbarkeit negativer Emotionen und besonders der kognitiven Reaktion auf induzierte Stimmungen jedoch ein relevantes zu sein: Je höher das Ausmaß dysfunktionaler Gedanken in Reaktion auf eine Induktion trauriger Stimmung ist, umso höher scheint die Rückfallrate remittierter Patienten mit Major Depression zu sein (Segal et al., 2006). Zukünftige Forschung sollte daher

die vorliegenden Befunde zur erhöhten Induzierbarkeit negativer Emotionen weiter vertiefen.

### 6.2.6 E6: Zusammenhänge zwischen Empathie und Social Functioning

In den vorliegenden Daten finden sich kaum Hinweise auf einen Zusammenhang der Fragebogenmessung der Empathie mit der prämorbidem Intelligenz. Die Selbsteinschätzung durch EQ-S scheint weder in einzelnen Untergruppen noch in der Gesamtgruppe mit der prämorbidem Intelligenz zu korrelieren. Auch die Fremdbeurteilung korreliert in der Gesamtgruppe nicht signifikant. Einzig in einer Kontrollgruppe und bei bipolar Remittierten zeigen sich signifikante Korrelationen. Weiterhin scheinen keine Zusammenhänge zwischen der Leistung in der CAM und der prämorbidem Intelligenz zu bestehen. Nur in einer Kontrollgruppe lässt sich eine signifikante Korrelation zwischen Ergebnissen im MASC und der prämorbidem Intelligenz beobachten. Auch die interpersonellen Probleme und die wahrgenommene soziale Unterstützung scheinen nicht signifikant mit der prämorbidem Intelligenz zu korrelieren. Für die wahrgenommene soziale Belastung hingegen finden sich negative Korrelationen mit der prämorbidem in der Gesamtgruppe (leicht) und bei bipolar remittierten Patienten (mittelstark). Insgesamt sprechen die vorliegenden Daten also nicht bis kaum für einen Zusammenhang zwischen Empathie und prämorbidem Intelligenz. Damit stimmen sie mit den Ergebnissen von Schwenck et al. (2014) überein, die nur sehr kleine Zusammenhänge zwischen allgemeiner Intelligenz und kognitiver Empathie und keine Zusammenhänge zwischen allgemeiner Intelligenz und emotionaler Empathie bei sich normal entwickelnden Kindern und Jugendlichen beobachten konnten. Ibanez et al. (2013) hingegen konnten im Rahmen einer Pfadanalyse Hinweise auf einen kausalen Zusammenhang zwischen fluider Intelligenz und ToM finden.

Zwischen den mit dem IIP gemessenen interpersonalen Problemen und der mit dem EQ gemessenen Empathie lassen sich mittelstarke negative Korrelationen für die Gesamtstichprobe beobachten. In den einzelnen Untergruppen finden sich signifikante Korrelationen jedoch nur zwischen EQ-F und IIP für unipolare Patienten und die entsprechende Stichprobe. Zwischen IIP und CAM lässt sich ausschließlich für akut depressive Patienten eine signifikante Korrelation beobachten. IIP und MASC korrelieren ebenfalls signifikant negativ in der Gruppe der Unipolar Akuten und leicht negativ in der Gesamtgruppe. Auch zwischen der Fragebogenmessung der Empathie und den, mit dem F-SozU gemessenen, Werten für wahrgenommene soziale Unterstützung sowie Belastung scheinen Zusammenhänge zu bestehen: In der Gesamtgruppe korrelieren EQ-S sowie EQ-F, CAM und MASC mittelstark positiv mit der wahrgenommenen sozialen Unterstützung. Zudem korreliert der EQ-F leicht negativ mit der wahrgenommenen sozialen Belastung. In den einzelnen Untergruppen schwankt das Ausmaß der Korrelationen jedoch sehr. Nur vereinzelt können leichte bis mittelstarke Korrelationen beobachtet werden. Die Zusammenhänge zwischen Fragebogenmessung, behavioraler Messung der Empathie und Social Functioning scheinen also nicht in allen Gruppen so deutlich zu sein, wie zu erwarten wäre (Riggio, 1986; Cliffordson, 2002). Dies widerspricht den in anderen Studien beobachteten Ergebnissen, die für einen Zusammenhang zwischen Empathie und Social

Functioning in verschiedenen Stichproben sprechen (Zhou et al., 2002; Michaels et al., 2014; D. Jolliffe & Farrington, 2004; Ang & Goh, 2010; Jeung & Herpertz, 2014; Preller et al., 2014; Bailey et al., 2008).

Auch die Zusammenhänge zwischen Fragebogenmessung und behavioraler Messung der Empathie sind nicht so stark wie zu erwarten sein könnte: In der Gesamtgruppe lassen sich leichte bis mittelstarke Korrelationen finden. In den einzelnen Untergruppen sind jedoch nur vereinzelt signifikante Korrelationen zu verzeichnen. Dies ähnelt den Beobachtungen von Melchers et al. (2015), die ebenfalls kaum Korrelationen zwischen verschiedenen behavioralen Maßen sowie mit Selbstbeurteilungs-Fragebögen zur Empathie finden konnten. Dies könnte an der Uneinigkeit bezüglich der Definition der Empathie aber auch bezüglich dessen, welche Konzepte durch welche Verfahren erfasst werden können, liegen (Thoma et al., 2013). Auch andere Bereiche der Psychologie und Neurowissenschaften berichten von einer Vielzahl widersprüchlicher Begriffe und Konzepte, die den Erkenntnisgewinn behindern (Poldrack & Yarkoni, 2016; Mischel, 2008). Poldrack & Yarkoni schlagen einen Ansatz zur Identifikation einzelner kognitiver Funktionen mithilfe Big Data vor. Auf der Basis großer Datensätze aus der Bildgebung sollen Zusammenhänge zwischen neurologischen Vorgängen und kognitiven Funktionen erschlossen werden. So sollen einheitliche Ontologien in den kognitiven Neurowissenschaften angestrebt werden. Eine solche Neuordnung der Begriffe anhand von aktuellen statistischen Herangehensweisen könnte auch bezüglich der Empathie, des Social Functioning oder der untersuchten Krankheitsentitäten eine Chance für ein tieferes Verständnis und eine zukunftssträchtige Ausrichtung des Forschungsbereichs sein.

### 6.2.7 Stärken und Schwächen

Zur Diskussion der Stärken und Schwächen der vorliegenden Arbeit muss zunächst die verwendete Stichprobe genannt werden: Trotz der vergleichsweise hohen Zahl von über 200 Teilnehmern sind die einzelnen Untergruppen mit 35 bis 45 Teilnehmern nicht allzu groß gewählt. Bei einer Varianzanalyse mit fünf Gruppen zu je ca. 40 Teilnehmern und einem  $\alpha$ -Niveau von .05 lassen sich mittelgroße Effekte finden. Da es sich jedoch möglicherweise um kleine Effekte handelt, könnten Gruppengrößen von je 240 Teilnehmern pro Gruppe nötig sein (Cohen, 1992). Zudem bestehen leichte Unterschiede zwischen den Gruppengrößen der verschiedenen Untergruppen. Diese kleinen Unterschiede sollten die Power der Tests aber nicht beeinträchtigen (Lachin, 1988; Schulz & Grimes, 2002). Die Rekrutierung von akut depressiven Patienten mit bipolarer Störung hat sich leider als nicht möglich erwiesen, da hier die Ein- und Ausschlusskriterien aufgrund der Medikation häufig nicht eingehalten werden konnten und zudem die Teilnahme an der sehr zeitaufwändigen Studie ein zu großes Hindernis zu sein schien. Es war daher nicht möglich, wie eigentlich geplant, einen umfassenden Blick zu ermöglichen, indem sowohl unipolar depressive als auch bipolare Patienten jeweils sowohl in akut depressiver als auch in remittierter Krankheitsphase verglichen werden. Weiterhin gibt es nur für die unipolar Remittierten sowie für die bipolar Remittierten gematchte gesunde Kontrollen. Eine Kontrollgruppe für unipolar akute Patienten fehlt in der vorliegenden Arbeit, weshalb Vergleiche hier zwischen den remittierten und akuten Patienten mit Major Depression sowie der

für die remittierten Patienten gematchten Kontrollgruppe gezogen werden mussten. Möglicherweise wären also bei einer noch größeren und ausgewogeneren Stichprobe Unterschiede bezüglich der Empathie messbar geworden. Nicht aus den Augen verloren werden sollte zudem die Diagnostik anhand des SKID I und II, die zwar dem Goldstandard entspricht, die jedoch in der vorliegenden Arbeit mit Nutzung der Sprungregeln durchgeführt wurde und daher keine solide Erfassung der Interrater-Reliabilität ermöglicht (vgl. Joormann & Gotlib, 2006). Eine Stärke der vorliegenden Arbeit ist die gründliche Untersuchung aller Teilnehmer – auch der gesunden Kontrollprobanden – mit dem SKID I und II, was in der bestehenden Literatur nicht immer umgesetzt wurde (z.B. Wilbertz et al., 2010).

Weiterhin muss die Möglichkeit einer Verzerrung der gezogenen Stichprobe (Heckman, 1977) in zwei Richtungen erwähnt werden: Erstens ist es denkbar, dass diejenigen Patienten, die an Forschung zur eigenen Erkrankung interessiert sind und sich als Studienteilnehmer bewerben, über besonders gute soziale und empathische Fähigkeiten verfügen. Zweitens ist es vorstellbar, dass diejenigen Gesunden, besonders jenseits des studentischen Alters, die sich als Versuchspersonen melden, sozial und beruflich weniger eingebunden sind, woraus sich die zeitliche Möglichkeit und der finanzielle Anreiz ergeben, an bezahlten Studien teilzunehmen. Diese geringere soziale und berufliche Einbindung könnte in verminderten sozialen und empathischen Fähigkeiten begründet sein. Aus diesen beiden Verzerrungen könnte sich daher eine Verschleierung eines tatsächlichen Unterschiedes bezüglich der Empathie von Patienten mit affektiven Störungen im Vergleich zu gesunden Kontrollen ergeben.

Für die vorliegende Arbeit wurde explizit darauf geachtet, möglichst ökologisch valide und ausreichend schwere behaviorale Maße der Empathie auszuwählen (Dziobek et al., 2006; Golan et al., 2006). Dennoch ist es denkbar, dass vorhandene – eventuell subtile – Unterschiede bezüglich der empathischen Fähigkeiten von Patienten mit affektiven Störungen mit CAM und MASC nicht detektiert werden konnten, da die ökologische Validität sowie die Schwierigkeitsgrade dieser Verfahren – obgleich schon recht hoch – noch nicht ausreichend waren. Aber auch in der Fragebogenmessung der Empathie konnten die erwarteten Unterschiede zwischen Patienten und Gesunden nicht erfasst werden, was ein konsistentes Bild mit der behavioralen Messung der Empathie ergibt. Die beobachteten hypothesenkonformen Einschränkungen des Social Functioning in den Patientengruppen können daher in der vorliegenden Arbeit nicht durch eine defizitäre Empathie erklärt werden. Alternative Erklärungsmöglichkeiten dieser Social Functioning Defizite sollten also in den Blick zukünftiger Forschungsarbeiten genommen werden.

Zu hinterfragen ist hierbei die bisher ausschließliche Selbsteinschätzung des Social Functionings durch die Probanden selbst: Sind die scheinbaren Einschränkungen des Social Functionings von Patienten mit affektiven Störungen keine tatsächlich erhöhten Konflikte, verminderte soziale Unterstützung und erhöhte soziale Belastung, sondern sind sich Patienten in sozialen Kontakten vermehrt unsicher und empfinden diese als schwieriger als gesunde Kontrollen? Erste Hinweise hierfür könnten die in der vorliegenden Arbeit beobachteten Gruppenunterschiede bezüglich der Sicherheits- und Schwierigkeitseinschätzung in der CAM sein, die spezifisch für die Emotionserkennung zu sein scheinen. Zukünftige Studien sollten daher zur Erfassung des Social Functioning Fremdbeurteilungen sowie objektive Maße mit einschließen.

So könnten beispielsweise beobachtende Methoden wie Rollenspiele zur Anwendung kommen (Norton & Hope, 2001).

Die Durchführung der behavioralen Messung der Empathie am Computer könnte besonders für diejenigen Teilnehmer ein Hindernis dargestellt haben, die wenig Erfahrung mit Computern im Allgemeinen und Spielcontrollern im Speziellen haben (Parshall & Kromrey, 1993). Besonders relevant ist dies für die CAM, da der Proband diese alleine in der Versuchskabine durchführte. Schwierigkeiten bei der Bedienung wurden mit Erklärungen und Übungstrials vor Beginn der eigentlichen Messung sowie mit einer auf dem Platz des Probanden sichtbar ausgelegten Bedienungsanleitung des Controllers begegnet. Dennoch sind Probleme bei der Bedienung nicht auszuschließen. Für den MASC dürfte dies eine deutlich geringere Rolle spielen, da hier der Versuchsleiter die Bedienung des Computers sowie die Protokollierung der Antwortauswahl übernahm.

Die Dauerhaftigkeit der Stimmungsinduktion muss – wie oben angesprochen – infrage gestellt werden. Direkt nach der Manipulation schien eine deutliche Verschlechterung der Stimmung zu beobachten. Nach Abschluss der CAM beziehungsweise des MASC scheint sich die Stimmung jedoch wieder auf Ausgangsniveau verbessert zu haben. Besonders bei der CAM ist hierbei aber zu beachten, dass für die vorliegende Arbeit nur der erste von fünf Blöcken ausgewertet wurde, während die Stimmung nach Abschluss der fünf Blöcke sowie der Kontrollaufgabe CALC erhoben wurde. Ob und wie lange die Induktion negativer Stimmung angehalten hat, ist also rein spekulativ. Für die Durchführung des MASC hingegen sprechen die vorliegenden Daten recht eindeutig dafür, dass die induzierte Stimmung zumindest nicht bis zum Abschluss der Aufgabe angehalten hat. Während beider Aufgaben scheint die Verminderung positiver Stimmung anhaltend. Da dies jedoch auch in der Sitzung ohne Stimmungsinduktion zu beobachten scheint, könnte dies ein Effekt des recht langen und anstrengenden Versuchsablaufs sein, der unabhängig von der vorgenommenen experimentellen Manipulation zu sein scheint. Von anderen Verfahren zur Stimmungsinduktion ist eine recht kurze Dauerhaftigkeit bekannt: So scheint sich induzierte niedergeschlagene Stimmung schon innerhalb eines zehnminütigen Wartezeitraumes signifikant zu verbessern und durch das Lesen positiver Sätze ganz zu erholen (Frost & Green, 1982).

Ein weiterer Kritikpunkt bezüglich der experimentellen Manipulation betrifft die Art der Stimmungsinduktion: In der vorliegenden Studie wurde negative Stimmung anhand trauriger Musik und der Erinnerung an eigene traurige Erlebnisse induziert. Obgleich diese Methode sich als erfolgreich erwiesen hat (Rojas et al., 2014), wären auch andere Arten der Stimmungsinduktion denkbar, die engere Zusammenhänge mit empathischen Fähigkeiten aufweisen könnten. Beispielsweise kann negative Stimmung auch anhand einer Kombination aus trauriger Musik und Bildern oder anhand trauriger Filmszenen induziert werden, in denen Probanden mit den Protagonisten mitfühlen (Zhang et al., 2014; Martin, 1990).

Bezüglich der Fragebogenmessung der Empathie sowie des Social Functioning gibt es verschiedene Schwächen und Stärken, die schon bezüglich der Fragebogenvalidierung diskutiert wurden. Das Fehlen eines Maßes der sozialen Erwünschtheit trifft auf die Experimentalstudie ebenso zu wie auf die Fragebogenvalidierung. Zudem wurden auch in der Experimentalstudie die Fragebögen von den Probanden



zu Hause ausgefüllt. Und auch hier wurden die Fremdbeurteiler durch die Probanden ausgewählt, was möglicherweise eine positive Beziehung und damit eine positive Sicht auf die empathischen Fähigkeiten des Probanden begünstigt.

## 6.3 Offene Forschungsfragen

Für die zukünftige Forschung ergeben sich auf Grundlage der vorliegenden Arbeit offene Fragen, die im Folgenden zusammengefasst werden.

1. Die in der vorliegenden Arbeit vielversprechende Validierung der deutschen Übersetzung sowie Umformulierung des EQ sollte an weiteren klinischen sowie gesunden Stichproben repliziert werden. Eine Mitbeachtung der sozialen Erwünschtheit wäre hierbei empfehlenswert.
2. Die neu entwickelte Fremdbeurteilungsversion des EQ sollte zudem mit Fremdbeurteilern ohne eine positive Beziehung zum Probanden beziehungsweise mit nicht vom Probanden selbst ausgewählten Fremdbeurteilern eingesetzt werden. Die Diskrepanz zwischen Selbst- und Fremdeinschätzung sollte an verschiedenen Stichproben ermittelt werden. Das Ausmaß dieser Diskrepanz und Zusammenhänge mit verschiedenen Erkrankungen sollten näher beleuchtet werden.
3. Die im Rahmen von Selbstbeurteilungen beobachteten Social Functioning Defizite von Patienten mit affektiven Störungen sollten anhand von Fremdbeurteilungsmaßen sowie objektiven Messverfahren, wie beispielsweise Rollenspielen, weiter untermauert werden. Durch diese Erweiterung der Messung des Social Functioning soll einer Verzerrung der Ergebnisse durch eine möglicherweise negative Selbstsicht der Patienten mit affektiven Störungen entgegengewirkt werden.
4. Zukünftige Arbeiten sollten bei einer Manipulation der aktuellen Stimmung sicherstellen, dass die induzierte Stimmung ausreichend lange anhält, um die Dauer der zu bearbeitenden Aufgaben abzudecken. Hierfür erscheinen effektive Verfahren der Stimmungsinduktion, eine gründliche und zeitnahe Messung der aktuellen Stimmung sowie eher kurze Aufgaben von Vorteil. Eine Untersuchung der Zusammenhänge zwischen Selbst- oder Fremdeinschätzung der Empathie mit der Induzierbarkeit negativer oder positiver Stimmung anhand von Verfahren, die Empathie erfordern – wie beispielsweise Filmszenen – wäre ein zu verfolgender Ansatz. Zudem sollten interindividuelle Unterschiede bezüglich der Induzierbarkeit negativer Stimmung näher erforscht werden: Die vorliegende Studie gibt Hinweise auf eine erhöhte Induzierbarkeit negativer Stimmung bei Patienten mit affektiven Störungen, die in der bisherigen Forschung nicht diskutiert wurde. Eine Replikation und weitere Klärung der Zusammenhänge könnte einen Beitrag zum Verständnis der Entstehung oder Aufrechterhaltung affektiver Störungen darstellen.
5. Trotz der breiten und umfassenden Herangehensweise an die Empathiefähigkeiten von Patienten mit affektiven Störungen in der vorliegenden Arbeit, blei-

ben Fragen offen. In der vorliegenden Arbeit konnten – anders als auf Grundlage der bestehenden Literatur zu erwarten – keine Hinweise auf Empathie-Defizite von akut depressiven oder remittierten Patienten mit Major Depression oder bipolarer Störung gefunden werden. Diese Ergebnisse sollten mit größeren und möglichst repräsentativen Stichproben repliziert werden. Eine Erfassung der empathischen Fähigkeiten akut depressiver Bipolarer sowie von Patienten aus stationären Settings sollte angestrebt werden, um dem Selection Bias entgegen zu wirken. Die Nutzung von Messverfahren mit noch höherer ökologischer Validität sowie Schwierigkeit könnte weiterhin dazu beitragen subtile Defizite zu erfassen.

6. Zur Untersuchung der Zusammenhänge zwischen verschiedenen Maßen der Empathie sowie zwischen Empathie, Social Functioning und affektiven Störungen, erscheint eine klare Definition sowie Operationalisierung der zu untersuchenden Konstrukte unerlässlich. Die Einigung auf eine gemeinsame Definition und Verwendung einheitlicher Untersuchungsmethoden sollte zukünftigen Forschungsarbeiten zugrunde gelegt werden, um eine bessere Vergleichbarkeit der Ergebnisse und damit einen Fortschritt des Forschungsbereichs zu ermöglichen. Eine Chance hierfür könnten Methoden des Big Data bieten.

## 6.4 Fazit

Die Ergebnisse der vorliegenden Arbeit sprechen dafür, dass der EQ ein geeignetes Maß der Empathie für die Selbst- sowie Fremdbeurteilung im Rahmen eines Fragebogens ist. Beide Versionen des EQ können in der deutschen Übersetzung als valide und reliabel angesehen werden und scheinen über gute psychometrische Eigenschaften zu verfügen.

Insgesamt lässt sich jedoch feststellen, dass die sozialen Schwierigkeiten von Patienten mit affektiven Störungen nicht durch ein Defizit in der Empathie erklärt werden können. In der hier untersuchten Stichprobe berichten die Patienten von den erwarteten sozialen Schwierigkeiten. Sie weisen jedoch in den verschiedensten Maßen eine ebenso ausgeprägte Empathie auf wie gesunde Kontrollen. Die bestehenden Social-Functioning-Defizite scheinen somit eine andere Ursache als Empathie-Defizite zu haben.

# Literaturverzeichnis

- Allison, C., Baron-Cohen, S., Wheelwright, S., Stone, M. H. & Muncer, S. J. (2011). Psychometric analysis of the empathy quotient (eq). *Personality and Individual Differences*, 51 (7), 829–835.
- American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders - DSM-IV*. American Psychological Association, Washington, DC.
- American Psychiatric Association. (2000). *DSM-IV-TR: diagnostic and statistical manual of mental disorders, text revision*. American Psychiatric Association, Washington, DC.
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5®)*. American Psychiatric Publishing, Arlington VA.
- Ang, R. P. & Goh, D. H. (2010). Cyberbullying among adolescents: The role of affective and cognitive empathy, and gender. *Child Psychiatry & Human Development*, 41 (4), 387–397.
- Antretter, E., Dunkel, D. & Haring, C. (2013). Wie zeitgemäß sind die in der deutschsprachigen Erwachsenenpsychiatrie verwendeten psychologischen Leistungstests? *Psychiatrische Praxis*, 40 (03), 120–129.
- Auyeung, B., Allison, C., Wheelwright, S. & Baron-Cohen, S. (2012). Brief report: development of the adolescent empathy and systemizing quotients. *Journal of autism and developmental disorders*, 42 (10), 2225–2235.
- Auyeung, B., Wheelwright, S., Allison, C., Atkinson, M., Samarawickrema, N. & Baron-Cohen, S. (2009). The children's empathy quotient and systemizing quotient: Sex differences in typical development and in autism spectrum conditions. *Journal of autism and developmental disorders*, 39 (11), 1509.
- Bailey, P. E., Henry, J. D. & Von Hippel, W. (2008). Empathy and social functioning in late adulthood. *Aging and Mental Health*, 12 (4), 499–503.
- Baldner, C. & McGinley, J. J. (2014). Correlational and exploratory factor analyses (EFA) of commonly used empathy questionnaires: New insights. *Motivation and Emotion*, 38 (5), 727–744.

- Baron-Cohen, S. (1989). The autistic child's theory of mind: A case of specific developmental delay. *Journal of Child Psychology and Psychiatry*, 30 (2), 285–297.
- Baron-Cohen, S., Leslie, A. M. & Frith, U. (1985). Does the autistic child have a "theory of mind"? *Cognition*, 21 (1), 37–46.
- Baron-Cohen, S. & Wheelwright, S. (2004). The Empathy Quotient: an investigation of adults with asperger syndrome or high functioning autism, and normal sex differences. *Journal of autism and developmental disorders*, 34 (2), 163–175.
- Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y. & Plumb, I. (2001). The "Reading the Mind in the Eyes" test revised version: A study with normal adults, and adults with asperger syndrome or high-functioning autism. *Journal of child psychology and psychiatry*, 42 (2), 241–251.
- Baron-Cohen, S., Wheelwright, S., Jolliffe & Therese. (1997). Is there a language of the eyes"? evidence from normal adults, and adults with autism or asperger syndrome. *Visual Cognition*, 4 (3), 311–331.
- Berthoz, S., Wessa, M., Kedia, G., Wicker, B. & Grèzes, J. (2008). Cross-cultural validation of the empathy quotient in a french-speaking sample. *The Canadian Journal of Psychiatry*, 53 (7), 469–477.
- Bodnar, A. & Rybakowski, J. K. (2017). Increased affective empathy in bipolar patients during a manic episode. *Revista Brasileira de Psiquiatria (AHEAD)*, 0–0.
- Bora, E. & Berk, M. (2016). Theory of mind in major depressive disorder: A meta-analysis. *Journal of affective disorders*, 191, 49–55.
- Bora, E., Vahip, S., Gonul, A., Akdeniz, F., Alkan, M., Ogut, M. & Eryavuz, A. (2005). Evidence for theory of mind deficits in euthymic patients with bipolar disorder. *Acta Psychiatrica Scandinavica*, 112 (2), 110–116.
- Bourke, C., Douglas, K. & Porter, R. (2010). Processing of facial emotion expression in major depression: a review. *Australian and New Zealand Journal of Psychiatry*, 44 (8), 681–696.
- Bräunig, P., Shugar, G. & Krüger, S. (1996). An investigation of the Self-Report Manic Inventory as a diagnostic and severity scale for mania. *Comprehensive psychiatry*, 37 (1), 52–55.
- Brothers, L. (1989). A biological perspective on empathy. *American Journal of Psychiatry*, 146 (1), 10–19.
- Brüne, M. & Brüne-Cohrs, U. (2006). Theory of mind—evolution, ontogeny, brain mechanisms and psychopathology. *Neuroscience & Biobehavioral Reviews*, 30 (4), 437–455.

- Busch, M., Maske, U., Ryl, L., Schlack, R. & Hapke, U. (2013). Prävalenz von depressiver Symptomatik und diagnostizierter Depression bei Erwachsenen in Deutschland. *Bundesgesundheitsblatt-Gesundheitsforschung-Gesundheitsschutz*, 56 (5-6), 733–739.
- Chepenik, L. G., Cornew, L. A. & Farah, M. J. (2007). The influence of sad mood on cognition. *Emotion*, 7 (4), 802.
- Cliffordson, C. (2002). The hierarchical structure of empathy: Dimensional organization and relations to social functioning. *Scandinavian Journal of Psychology*, 43 (1), 49–59.
- Cohen, J. (1992). A power primer. *Psychological bulletin*, 112 (1), 155.
- Crowne, D. P. & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of consulting psychology*, 24 (4), 349.
- Cuff, B. M., Brown, S. J., Taylor, L. & Howat, D. J. (2016). Empathy: a review of the concept. *Emotion Review*, 8 (2), 144–153.
- Cusi, A. M., MacQueen, G. M., Spreng, R. N. & McKinnon, M. C. (2011). Altered empathic responding in major depressive disorder: relation to symptom severity, illness burden, and psychosocial outcome. *Psychiatry research*, 188 (2), 231–236.
- Cusi, A. M., Nazarov, A., Holshausen, K., MacQueen, G. M. & McKinnon, M. C. (2012). Systematic review of the neural basis of social cognition in patients with mood disorders. *Journal of psychiatry & neuroscience: JPN*, 37 (3), 154.
- Cusi, A. M., Nazarov, A., MacQueen, G. M. & McKinnon, M. C. (2013). Theory of mind deficits in patients with mild symptoms of major depressive disorder. *Psychiatry research*, 210 (2), 672–674.
- Dalili, M., Penton-Voak, I., Harmer, C. & Munafò, M. (2015). Meta-analysis of emotion recognition deficits in major depressive disorder. *Psychological medicine*, 45 (06), 1135–1144.
- Davis, M. H. (1983). Measuring individual differences in empathy: evidence for a multidimensional approach. *Journal of personality and social psychology*, 44 (1), 113.
- Davis, M. H., Association, A. P. et al. (1980). A multidimensional approach to individual differences in empathy.
- de Haen, J. (o.J.). *Empathy Quotient – Deutsch*. [https://www.autismresearchcentre.com/arc\\_tests](https://www.autismresearchcentre.com/arc_tests). (Abgerufen am 20.06.2017)
- Demescu, L. R., Kortekaas, R., den Boer, J. A. & Aleman, A. (2010). Impaired attribution of emotion to facial expressions in anxiety and major depression. *PLoS One*, 5 (12), e15058.

- Derntl, B., Seidel, E.-M., Schneider, F. & Habel, U. (2012). How specific are emotional deficits? a comparison of empathic abilities in schizophrenia, bipolar and depressed patients. *Schizophrenia research*, 142 (1), 58–64.
- Dilling, H., Mombour, W., Schmidt, M. H., Organization, W. H. et al. (1991). Internationale Klassifikation psychischer Störungen: ICD-10, Kapitel V F, klinisch-diagnostische Leitlinien.
- Domes, G., Spenthof, I., Radtke, M., Isaksson, A., Normann, C. & Heinrichs, M. (2016). Autistic traits and empathy in chronic vs. episodic depression. *Journal of affective disorders*, 195, 144–147.
- Dykieriek, P., Schramm, E. & Berger, M. (2012). Bedeutung der Theory of Mind für die Psychotherapie der Depression. In *Theory of mind* (S. 319–335). Springer.
- Dziobek, I., Fleck, S., Kalbe, E., Rogers, K., Hassenstab, J., Brand, M., . . . Convit, A. (2006). Introducing MASC: a movie for the assessment of social cognition. *Journal of autism and developmental disorders*, 36 (5), 623–636.
- Ehret, A. M. & Berking, M. (2013). DSM-IV und DSM-5: Was hat sich tatsächlich verändert? *Verhaltenstherapie*, 23 (4), 258–266.
- Eisenberg, N. & Fabes, R. A. (1990). Empathy: Conceptualization, measurement, and relation to prosocial behavior. *Motivation and Emotion*, 14 (2), 131–149.
- Eisenberg, N. & Miller, P. A. (1987). The relation of empathy to prosocial and related behaviors. *Psychological bulletin*, 101 (1), 91.
- Ekman, P. & Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of personality and social psychology*, 17 (2), 124.
- Faireclough, S. H., van der Zwaag, M., Spiridon, E. & Westerink, J. (2014). Effects of mood induction via music on cardiovascular measures of negative emotion during simulated driving. *Physiology & behavior*, 129, 173–180.
- Falkai, P. & Wittchen, H.-U. (2015). *Diagnostisches und statistisches Manual psychischer Störungen - DSM-5*. Hogrefe, Göttingen.
- Fan, Y., Duncan, N. W., de Greck, M. & Northoff, G. (2011). Is there a core neural network in empathy? an fmri based quantitative meta-analysis. *Neuroscience & Biobehavioral Reviews*, 35 (3), 903–911.
- Faul, F., Erdfelder, E., AG, L. & Buchner, A. (2007). G\*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39, 175-191.
- First, J., MB an Williams, Karg, R. & Spitzer, R. (2015). *Structured clinical interview for DSM-V – research version (SCID-5 for DSM-V, research version SCID-5-RV)*. American Psychiatric Publishing, Arlington VA.

- Fossati, A., Borroni, S., Dziobek, I., Fonagy, P. & Somma, A. (2017). Thinking about assessment: Further evidence of the validity of the movie for the assessment of social cognition as a measure of mentalistic abilities.
- Frost, R. O. & Green, M. L. (1982). Velten mood induction procedure effects: Duration and postexperimental removal. *Personality and Social Psychology Bulletin*, 8 (2), 341–347.
- Fydrich, T., Geyer, M., Hessel, A., Sommer, G. & Brähler, E. (1999). Fragebogen zur Sozialen Unterstützung (F-SozU): Normierung an einer repräsentativen Stichprobe. *Diagnostica*, 45, 212–216.
- Fydrich, T., Sommer, G. & Brähler, E. (2007). Fragebogen zur Sozialen Unterstützung. *Hogrefe, Göttingen*.
- Gagan, J. M. (1983). Methodological notes on empathy. *Advances in Nursing Science*, 5 (2), 65–72.
- Golan, O., Baron-Cohen, S. & Hill, J. (2006). The cambridge mindreading (CAM) face-voice battery: Testing complex emotion recognition in adults with and without asperger syndrome. *Journal of autism and developmental disorders*, 36 (2), 169–183.
- Goldsmith, D., Rapaport, M. & Miller, B. (2016). A meta-analysis of blood cytokine network alterations in psychiatric patients: comparisons between schizophrenia, bipolar disorder and depression. *Molecular psychiatry*, 21 (12), 1696–1709.
- Goldstein, T. R. & Winner, E. (2012). Enhancing empathy and theory of mind. *Journal of Cognition and Development*, 13 (1), 19–37.
- Green, E. K., Grozeva, D., Jones, I., Jones, L., Kirov, G., Caesar, S., ... others (2010). The bipolar disorder risk allele at CACNA1C also confers risk of recurrent major depression and of schizophrenia. *Molecular psychiatry*, 15 (10), 1016–1022.
- Gregory, C., Lough, S., Stone, V., Erzinclioglu, S., Martin, L., Baron-Cohen, S. & Hodges, J. R. (2002). Theory of mind in patients with frontal variant frontotemporal dementia and alzheimer's disease: theoretical and practical implications. *Brain*, 125 (4), 752–764.
- Gronwall, D. (1977). Paced auditory serial-addition task: a measure of recovery from concussion. *Perceptual and motor skills*, 44 (2), 367–373.
- Guan, R., Jin, L. & Qian, M. (2012). Validation of the empathy quotient–short form among chinese healthcare professionals. *Social Behavior and Personality: an international journal*, 40 (1), 75–84.
- Hamilton, M. (1960). A rating scale for depression. *Journal of Neurology, Neurosurgery & Psychiatry*, 23 (1), 56–62.

- Happé, F. (1994). An advanced test of theory of mind: Understanding of story characters' thoughts and feelings by able autistic, mentally handicapped, and normal children and adults. *Journal of autism and Developmental disorders*, 24 (2), 129–154.
- Harkness, K. L., Washburn, D., Theriault, J. E., Lee, L. & Sabbagh, M. A. (2011). Maternal history of depression is associated with enhanced theory of mind in depressed and nondepressed adult women. *Psychiatry research*, 189 (1), 91–96.
- Hassenstab, J., Dziobek, I., Rogers, K., Wolf, O. T. & Convit, A. (2007). Knowing what others know, feeling what others feel: a controlled study of empathy in psychotherapists. *The Journal of nervous and mental disease*, 195 (4), 277–281.
- Hautzinger, M. (1991). Das Beck-Depressioninventar (BDI) in der Klinik. *Der Nervenarzt*.
- Hautzinger, M. & Bailer, M. (1994). Das Inventar Depressiver Symptome (IDS). *Beltz Test, Weinheim*.
- Hautzinger, M., Bailer, M., Worall, H. & Keller, F. (1994). Beck-Depressions-Inventar (BDI). Bearbeitung der deutschen Ausgabe. Testhandbuch. *Bern: Huber*.
- Heckman, J. J. (1977). *Sample selection bias as a specification error (with an application to the estimation of labor supply functions)*. National Bureau of Economic Research Cambridge, Mass., USA.
- Helmstädter, C., Lendt, M. & Lux, S. (2001). Verbaler Lern- und Merkfähigkeitstest. *Hogrefe, Göttingen*.
- Herzberg, P. Y. (2003). Faktorstruktur, Gütekriterien und Konstruktvalidität der deutschen Übersetzung des Aggressionsfragebogens von Buss und Perry. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 24 (4), 311–323.
- Hill, S. K., Reilly, J. L., Keefe, R. S., Gold, J. M., Bishop, J. R., Gershon, E. S., ... Sweeney, J. A. (2013). Neuropsychological impairments in schizophrenia and psychotic bipolar disorder. *American Journal of Psychiatry*, 170 (11), 1275–1284.
- Hoffmann, F., Banzhaf, C., Kanske, P., Gärtner, M., Bermpohl, F. & Singer, T. (2016). Empathy in depression: Egocentric and altercentric biases and the role of alexithymia. *Journal of affective disorders*, 199, 23–29.
- Horowitz, L. M., Strauß, B. & Kordy, H. (2000). *Inventar zur Erfassung interpersonaler Probleme: IIP-D; deutsche Version*. Beltz.
- Ibanez, A., Huepe, D., Gempp, R., Gutiérrez, V., Rivera-Rei, A. & Toledo, M. I. (2013). Empathy, sex and fluid intelligence as predictors of theory of mind. *Personality and Individual Differences*, 54 (5), 616–621.



- Inoue, Y., Tonooka, Y., Yamada, K. & Kanba, S. (2004). Deficiency of theory of mind in patients with remitted mood disorder. *Journal of affective disorders*, *82* (3), 403–409.
- Inoue, Y., Yamada, K. & Kanba, S. (2006). Deficit in theory of mind is a risk for relapse of major depression. *Journal of affective disorders*, *95* (1), 125–127.
- Jeung, H. & Herpertz, S. C. (2014). Impairments of interpersonal functioning: empathy and intimacy in borderline personality disorder. *Psychopathology*, *47* (4), 220–234.
- Jolliffe, D. & Farrington, D. P. (2004). Empathy and offending: A systematic review and meta-analysis. *Aggression and violent behavior*, *9* (5), 441–476.
- Jolliffe, D. & Farrington, D. P. (2006). Development and validation of the basic empathy scale. *Journal of adolescence*, *29* (4), 589–611.
- Jolliffe, T. & Baron-Cohen, S. (1999). The strange stories test: A replication with high-functioning adults with autism or asperger syndrome. *Journal of autism and developmental disorders*, *29* (5), 395–406.
- Joormann, J. & Gotlib, I. H. (2006). Is this happiness i see? biases in the identification of emotional facial expressions in depression and social phobia. *Journal of abnormal psychology*, *115* (4), 705.
- Joormann, J., Siemer, M. & Gotlib, I. H. (2007). Mood regulation in depression: Differential effects of distraction and recall of happy memories on sad mood. *Journal of abnormal psychology*, *116* (3), 484.
- Judd, L. L., Akiskal, H. S., Schettler, P. J., Endicott, J., Leon, A. C., Solomon, D. A., ... Keller, M. B. (2005). Psychosocial disability in the course of bipolar i and ii disorders: a prospective, comparative, longitudinal study. *Archives of general psychiatry*, *62* (12), 1322–1330.
- Judd, L. L., Akiskal, H. S., Zeller, P. J., Paulus, M., Leon, A. C., Maser, J. D., ... others (2000). Psychosocial disability during the long-term course of unipolar major depressive disorder. *Archives of general psychiatry*, *57* (4), 375–380.
- Kerr, N., Dunbar, R. I. & Bentall, R. P. (2003). Theory of mind deficits in bipolar affective disorder. *Journal of affective disorders*, *73* (3), 253–259.
- Kettle, J. W., O'Brien-Simpson, L. & Allen, N. B. (2008). Impaired theory of mind in first-episode schizophrenia: comparison with community, university and depressed controls. *Schizophrenia research*, *99* (1), 96–102.
- Kiesler, D. & Schmidt, J. (1993). The impact message inventory: Form IIA octant scale version. *Richmond: Virginia Commonwealth University*.
- Kobak, K. A., Williams, J. B. & Rush, A. J. (2007). Strukturierte interview-richtlinien für die hamilton depression scale und das inventar depressiver symptome (sighd-ids).

- Kohler, C. G., Hoffman, L. J., Eastman, L. B., Healey, K. & Moberg, P. J. (2011). Facial emotion perception in depression and bipolar disorder: a quantitative review. *Psychiatry research*, *188* (3), 303–309.
- Krüger, S., Bräunig, P. & Shugar, G. (1997). Manie-Selbstbeurteilungsskala. *Beltz Test, Göttingen*.
- Kupferberg, A., Bicks, L. & Hasler, G. (2016). Social functioning in major depressive disorder. *Neuroscience & Biobehavioral Reviews*, *69*, 313–332.
- Lachin, J. M. (1988). Statistical properties of randomization in clinical trials. *Controlled clinical trials*, *9* (4), 289–311.
- Ladegaard, N., Larsen, E. R., Videbech, P. & Lysaker, P. H. (2014). Higher-order social cognition in first-episode major depression. *Psychiatry research*, *216* (1), 37–43.
- Lahera, G., Ruiz-Murugarren, S., Iglesias, P., Ruiz-Bennasar, C., Herrería, E., Montes, J. M. & Fernández-Liria, A. (2012). Social cognition and global functioning in bipolar disorder. *The Journal of nervous and mental disease*, *200* (2), 135–141.
- Lawrence, E., Shaw, P., Baker, D., Baron-Cohen, S. & David, A. (2004). Measuring empathy: reliability and validity of the Empathy Quotient. *Psychological medicine*, *34* (05), 911–920.
- Lee, L., Harkness, K. L., Sabbagh, M. A. & Jacobson, J. A. (2005). Mental state decoding abilities in clinical depression. *Journal of affective disorders*, *86* (2), 247–258.
- Lehrl, S. (1999). *Mehrfachwahl-Wortschatz-Intelligenztest: MWT-B*. Spitta.
- Lehrl, S., Triebig, G. & Fischer, B. (1995). Multiple choice vocabulary test MWT as a valid and short test to estimate premorbid intelligence. *Acta Neurologica Scandinavica*, *91* (5), 335–345.
- LeMoult, J., Joormann, J., Sherdell, L., Wright, Y. & Gotlib, I. H. (2009). Identification of emotional facial expressions following recovery from depression. *Journal of abnormal psychology*, *118* (4), 828.
- Leppänen, J. M. (2006). Emotional information processing in mood disorders: a review of behavioral and neuroimaging findings. *Current opinion in psychiatry*, *19* (1), 34–39.
- Liebowitz, M. R. (1987). *Social phobia*. Karger Publishers.
- Liu, W.-h., Huang, J., Wang, L.-z., Gong, Q.-y. & Chan, R. C. (2012). Facial perception bias in patients with major depression. *Psychiatry research*, *197* (3), 217–220.
- Loch, N., Hiller, W. & Witthöft, M. (2011). Der cognitive emotion regulation questionnaire (CERQ). *Zeitschrift für Klinische Psychologie und Psychotherapie*.

- Locke, K. D. (2005). Interpersonal problems and interpersonal expectations in everyday life. *Journal of social and clinical psychology, 24* (7), 915–931.
- Martin, M. (1990). On the induction of mood. *Clinical Psychology Review, 10* (6), 669–697.
- Martino, D. J., Strejilevich, S. A., Fassi, G., Marengo, E. & Igoa, A. (2011). Theory of mind and facial emotion recognition in euthymic bipolar i and bipolar ii disorders. *Psychiatry research, 189* (3), 379–384.
- Maslow, A. H. (1943). A theory of human motivation. *Psychological review, 50* (4), 370.
- Melchers, M. C., Montag, C., Markett, S. & Reuter, M. (2015). Assessment of empathy via self-report and behavioural paradigms: data on convergent and discriminant validity. *Cognitive neuropsychiatry, 20* (2), 157–171.
- Mercer, L. & Becerra, R. (2013). A unique emotional processing profile of euthymic bipolar disorder? a critical review. *Journal of affective disorders, 146* (3), 295–309.
- Michaels, T. M., Horan, W. P., Ginger, E. J., Martinovich, Z., Pinkham, A. E. & Smith, M. J. (2014). Cognitive empathy contributes to poor social functioning in schizophrenia: evidence from a new self-report measure of cognitive and affective empathy. *Psychiatry research, 220* (3), 803–810.
- Michalak, E. E., Yatham, L. N. & Lam, R. W. (2005). Quality of life in bipolar disorder: a review of the literature. *Health and Quality of Life Outcomes, 3* (1), 72.
- Milders, M., Bell, S., Platt, J., Serrano, R. & Runcie, O. (2010). Stable expression recognition abnormalities in unipolar depression. *Psychiatry research, 179* (1), 38–42.
- Mischel, W. (2008). The toothbrush problem. *APS Observer, 21* (11).
- Mischo, C. (2003). Wie valide sind Selbsteinschätzungen der Empathie? *Gruppendynamik und Organisationsberatung, 34* (2), 187–202.
- Montag, C., Ehrlich, A., Neuhaus, K., Dziobek, I., Heekeren, H. R., Heinz, A. & Gallinat, J. (2010). Theory of mind impairments in euthymic bipolar patients. *Journal of affective disorders, 123* (1), 264–269.
- Montag, C., Gallinat, J. & Heinz, A. (2008). Theodor lipps and the concept of empathy: 1851–1914. *American Journal of Psychiatry, 165* (10), 1261–1261.
- Muehlbacher, M., Egger, C., Kaplan, P., Simhandl, C., Grunze, H., Geretsegger, C., ... Stuppaeck, C. (2011). Reliabilität und Übereinstimmungsvalidität der deutschen Version der Young Mania Rating Scale (YMRS-D). *Neuropsychiatrie, 25* (1), 1–10.

- Nejati, V., Zabihzadeh, A., Maleki, G. & Tehrani, A. (2012). Mind reading and mindfulness deficits in patients with major depression disorder. *Procedia-Social and Behavioral Sciences*, *32*, 431–437.
- Norton, P. J. & Hope, D. A. (2001). Analogue observational methods in the assessment of social functioning in adults. *Psychological Assessment*, *13* (1), 59.
- Olley, A. L., Malhi, G. S., Bachelor, J., Cahill, C. M., Mitchell, P. B. & Berk, M. (2005). Executive functioning and theory of mind in euthymic bipolar disorder. *Bipolar Disorders*, *7* (s5), 43–52.
- Parshall, C. G. & Kromrey, J. D. (1993). Computer testing versus paper-and-pencil testing: An analysis of examinee characteristics associated with mode effect.
- Paulus, C. (2009). Der Saarbrücker Persönlichkeitsfragebogen SPF (IRI) zur Messung von Empathie: Psychometrische Evaluation der deutschen Version des Interpersonal Reactivity Index.
- Perner, J. & Wimmer, H. (1985). “john thinks that mary thinks that...” attribution of second-order beliefs by 5-to 10-year-old children. *Journal of experimental child psychology*, *39* (3), 437–471.
- Pini, S., de Queiroz, V., Pagnin, D., Pezawas, L., Angst, J., Cassano, G. B. & Wittchen, H.-U. (2005). Prevalence and burden of bipolar disorders in european countries. *European Neuropsychopharmacology*, *15* (4), 425–434.
- Poldrack, R. A. & Yarkoni, T. (2016). From brain maps to cognitive ontologies: informatics and the search for mental structure. *Annual review of psychology*, *67*, 587–612.
- Pompili, M., Innamorati, M., Gonda, X., Serafini, G., Sarno, S., Erbuto, D., ... others (2013). Affective temperaments and hopelessness as predictors of health and social functioning in mood disorder patients: a prospective follow-up study. *Journal of affective disorders*, *150* (2), 216–222.
- Preller, K. H., Hulka, L. M., Vonmoos, M., Jenni, D., Baumgartner, M. R., Seifritz, E., ... Quednow, B. B. (2014). Impaired emotional empathy and related social network deficits in cocaine users. *Addiction biology*, *19* (3), 452–466.
- Premack, D. & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and brain sciences*, *1* (04), 515–526.
- Preti, A., Vellante, M., Baron-Cohen, S., Zucca, G., Petretto, D. R. & Masala, C. (2011). The empathy quotient: A cross-cultural comparison of the italian version. *Cognitive Neuropsychiatry*, *16* (1), 50–70.
- Purcell, A. L., Phillips, M. & Gruber, J. (2013). In your eyes: Does theory of mind predict impaired life functioning in bipolar disorder? *Journal of affective disorders*, *151* (3), 1113–1119.

- Rapaport, M. H., Clary, C., Fayyad, R. & Endicott, J. (2005). Quality-of-life impairment in depressive and anxiety disorders. *American Journal of Psychiatry*, *162* (6), 1171–1178.
- Rasic, D., Hajek, T., Alda, M. & Uher, R. (2013). Risk of mental illness in offspring of parents with schizophrenia, bipolar disorder, and major depressive disorder: a meta-analysis of family high-risk studies. *Schizophrenia Bulletin* (2014), *40*.
- Reitan, R. (1992). Trail making test: Manual for administration and scoring. *Reitan Neuropsychology Laboratory, South Tuscon, AZ*.
- Riggio, R. E. (1986). Assessment of basic social skills. *Journal of Personality and social Psychology*, *51* (3), 649.
- Roiser, J., Farmer, A., Lam, D., Burke, A., O’neill, N., Keating, S., ... McGuffin, P. (2009). The effect of positive mood induction on emotional processing in euthymic individuals with bipolar disorder and controls. *Psychological medicine*, *39* (5), 785–791.
- Rojas, R., Geissner, E. & Hautzinger, M. (2014). Kognitive Reaktivität und Stressbelastung als Prädiktoren eines Rezidivs bei remittiert depressiven Personen. *Zeitschrift für Klinische Psychologie und Psychotherapie*.
- Rubinow, D. R. & Post, R. M. (1992). Impaired recognition of affect in facial expression in depressed patients. *Biological psychiatry*, *31* (9), 947–953.
- Rush, A. J., Gullion, C. M., Basco, M. R., Jarrett, R. B. & Trivedi, M. H. (1996). The inventory of depressive symptomatology (ids): psychometric properties. *Psychological medicine*, *26* (03), 477–486.
- Saß, H., Wittchen, H.-U., Zaudig, M. & Houben, I. (2003). *Diagnostisches und Statistisches Manual Psychischer Störungen - Textrevision - DSM-IV-TR*. Hogrefe, Göttingen.
- Sabbagh, M. A. (2004). Understanding orbitofrontal contributions to theory-of-mind reasoning: Implications for autism. *Brain and cognition*, *55* (1), 209–219.
- Samamé, C., Martino, D. & Strejilevich, S. (2012). Social cognition in euthymic bipolar disorder: systematic review and meta-analytic approach. *Acta Psychiatrica Scandinavica*, *125* (4), 266–280.
- Samson, A. C. & Huber, O. W. (2010). Short german versions of empathizing and systemizing self-assessment scales. *Swiss Journal of Psychology*.
- Sanchez, A., Vazquez, C., Gomez, D. & Joormann, J. (2014). Gaze-fixation to happy faces predicts mood repair after a negative mood induction. *Emotion*, *14* (1), 85.
- Saß, H., Wittchen, H.-U. & Zaudig, M. (1996). *Diagnostisches und statistisches Manual psychischer Störungen - DSM-IV. Deutsche Bearbeitung u. Einleitung*. Hogrefe, Göttingen.

- Schaefer, K. L., Baumann, J., Rich, B. A., Luckenbaugh, D. A. & Zarate, C. A. (2010). Perception of facial emotion in adults with bipolar or unipolar depression and controls. *Journal of psychiatric research*, *44* (16), 1229–1235.
- Schmider, E., Ziegler, M., Danay, E., Beyer, L. & Bühner, M. (2010). Is it really robust? *Methodology*.
- Schneider, D., Regenbogen, C., Kellermann, T., Finkelmeyer, A., Kohn, N., Derntl, B., ... Habel, U. (2012). Empathic behavioral and physiological responses to dynamic stimuli in depression. *Psychiatry research*, *200* (2), 294–305.
- Schneider, F., Gur, R. C., Gur, R. E. & Shtasel, D. L. (1995). Emotional processing in schizophrenia: neurobehavioral probes in relation to psychopathology. *Schizophrenia research*, *17* (1), 67–75.
- Schreiter, S., Pijnenborg, G. & Aan Het Rot, M. (2013). Empathy in adults with clinical or subclinical depressive symptoms. *Journal of Affective Disorders*, *150* (1), 1–16.
- Schulz, K. F. & Grimes, D. A. (2002). Unequal group sizes in randomised trials: guarding against guessing. *The Lancet*, *359* (9310), 966–970.
- Schwenck, C., Göhle, B., Hauf, J., Warnke, A., Freitag, C. M. & Schneider, W. (2014). Cognitive and emotional empathy in typically developing children: The influence of age, gender, and intelligence. *European Journal of Developmental Psychology*, *11* (1), 63–76.
- Segal, Z. V., Kennedy, S., Gemar, M., Hood, K., Pedersen, R. & Buis, T. (2006). Cognitive reactivity to sad mood provocation and the prediction of depressive relapse. *Archives of General Psychiatry*, *63* (7), 749–755.
- Shamay-Tsoory, S. G. (2011). The neural bases for empathy. *The Neuroscientist*, *17* (1), 18–24.
- Singer, T. (2006). The neuronal basis and ontogeny of empathy and mind reading: review of literature and implications for future research. *Neuroscience & Biobehavioral Reviews*, *30* (6), 855–863.
- Spreng, R. N., McKinnon, M. C., Mar, R. A. & Levine, B. (2009). The toronto empathy questionnaire: Scale development and initial validation of a factor-analytic solution to multiple empathy measures. *Journal of personality assessment*, *91* (1), 62–71.
- Stangier, U. & Heidenreich, T. (1995). Die deutsche Bearbeitung der Social Phobia Scale (SPS) und Social Interaction Anxiety Scale (SIAS) von Mattick und Clarke. In *Vortrag auf dem Kongreß der Deutschen Gesellschaft für Psychologie in Hamburg*.
- Surguladze, S. A., Young, A. W., Senior, C., Brébion, G., Travis, M. J. & Phillips, M. L. (2004). Recognition accuracy and response bias to happy and sad facial expressions in patients with major depression. *Neuropsychology*, *18* (2), 212.

- Thoma, P., Friedmann, C. & Suchan, B. (2013). Empathy and social problem solving in alcohol dependence, mood disorders and selected personality disorders. *Neuroscience & Biobehavioral Reviews*, *37* (3), 448–470.
- Thoma, P., Zalewski, I., von Reventlow, H. G., Norra, C., Juckel, G. & Daum, I. (2011). Cognitive and affective empathy in depression linked to executive control. *Psychiatry research*, *189* (3), 373–378.
- Van Rheeën, T. E. & Rossell, S. L. (2014). Let's face it: facial emotion processing is impaired in bipolar disorder. *Journal of the International Neuropsychological Society*, *20* (02), 200–208.
- Västfjäll, D. (2001). Emotion induction through music: A review of the musical mood induction procedure. *Musicae Scientiae*, *5* (1\_suppl), 173–211.
- Völlm, B. A., Taylor, A. N., Richardson, P., Corcoran, R., Stirling, J., McKie, S., ... Elliott, R. (2006). Neuronal correlates of theory of mind and empathy: a functional magnetic resonance imaging study in a nonverbal task. *Neuroimage*, *29* (1), 90–98.
- Wakabayashi, A., Baron-Cohen, S., Uchiyama, T., Yoshida, Y., Kuroda, M. & Wheelwright, S. (2007). Empathizing and systemizing in adults with and without autism spectrum conditions: Cross-cultural stability. *Journal of autism and developmental disorders*, *37* (10), 1823–1832.
- Wang, Y.-g., Wang, Y.-q., Chen, S.-l., Zhu, C.-y. & Wang, K. (2008). Theory of mind disability in major depression with or without psychotic symptoms: a componential view. *Psychiatry research*, *161* (2), 153–161.
- Washburn, D., Wilson, G., Roes, M., Rnic, K. & Harkness, K. L. (2016). Theory of mind in social anxiety disorder, depression, and comorbid conditions. *Journal of anxiety disorders*, *37*, 71–77.
- Watson, D., Clark, L. A. & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the panas scales. *Journal of personality and social psychology*, *54* (6), 1063.
- Weltgesundheitsorganisation. (2017). *The 11th Revision of the International Classification of Diseases (ICD-11) is due by 2018!* <http://www.who.int/classifications/icd/revision/en/>. (Abgerufen am 24.04.2017)
- White, S., Hill, E., Happé, F. & Frith, U. (2009). Revisiting the strange stories: revealing mentalizing impairments in autism. *Child development*, *80* (4), 1097–1117.
- Wilbertz, G., Brakemeier, E.-L., Zobel, I., Härter, M. & Schramm, E. (2010). Exploring preoperational features in chronic depression. *Journal of affective disorders*, *124* (3), 262–269.
- Williams, J. B. (1988). A structured interview guide for the Hamilton Depression Rating Scale. *Archives of general psychiatry*, *45* (8), 742–747.

- Wittchen, H.-U., Zaudig, M. & Fydrich, T. (1997). SKID. Strukturiertes klinisches Interview für DSM-IV. Achse I und II. Handanweisung.
- Wolf, F., Brüne, M. & Assion, H.-J. (2010). Theory of mind and neurocognitive functioning in patients with bipolar disorder. *Bipolar disorders*, 12 (6), 657–666.
- Wolkenstein, L., Schönenberg, M., Schirm, E. & Hautzinger, M. (2011). I can see what you feel, but i can't deal with it: Impaired theory of mind in depression. *Journal of affective disorders*, 132 (1), 104–111.
- Yang, J., Hu, X., Li, X., Zhang, L., Dong, Y., Li, X., ... others (2017). Decreased empathy response to other people's pain in bipolar disorder: evidence from an event-related potential study. *Scientific Reports*, 7.
- Young, R., Biggs, J., Ziegler, V. & Meyer, D. (1978). A rating scale for mania: reliability, validity and sensitivity. *The British Journal of Psychiatry*, 133 (5), 429–435.
- Zhang, X., Hui, W. Y. & Barrett, L. F. (2014). How does this make you feel? a comparison of four affect induction procedures. *Frontiers in psychology*, 5.
- Zhou, Q., Eisenberg, N., Losoya, S. H., Fabes, R. A., Reiser, M., Guthrie, I. K., ... Shepard, S. A. (2002). The relations of parental warmth and positive expressiveness to children's empathy-related responding and social functioning: A longitudinal study. *Child development*, 73 (3), 893–915.
- Zwick, J. C. & Wolkenstein, L. (2017). Facial emotion recognition, theory of mind and the role of facial mimicry in depression. *Journal of Affective Disorders*, 210, 90–99.



# Anhang A

## Empathy Quotient

### A.1 Selbstbeurteilung

#### A.1.1 Instruktion

Im weiteren Verlauf finden Sie eine Liste von Aussagen. Bitte lesen Sie jede Aussage sehr gründlich und bewerten Sie, wie sehr Sie dieser Aussage zustimmen oder sie ablehnen.

Es gibt keine richtigen oder falschen Antworten und auch keine Fangfragen.

Es ist wichtig, dass jede Frage beantwortet wird, denn nur dann ist eine gültige Auswertung möglich.

#### A.1.2 Items

Im Folgenden sind alle Items der Selbstbeurteilungsvariante aufgezählt. Beispiel-Items sind mit [B], Filler-Items mit [F], positiv gepolte Items mit [+] und negativ gepolte Items mit [-] markiert. Alle Items werden auf einer vierstufigen Skala mit 3 = „stimme voll und ganz zu“, 2 = „stimme eher zu“, 1 = „stimme eher nicht zu“ 0 = „stimme überhaupt nicht zu“ beantwortet.

- a) Ich würde mich sehr aufregen, wenn ich nicht jeden Tag Musik hören könnte. [B]
- b) Ich rede mit meinen Freunden lieber am Telefon, statt ihnen Briefe zu schreiben. [B]
- c) Ich habe kein Bedürfnis, andere Teile der Welt zu bereisen. [B]
- d) Ich lese lieber, als zu tanzen. [B]
- 1. Ich erkenne leicht, ob jemand an einem Gespräch interessiert ist. [+]
- 2. Ich mag Tiere lieber als Menschen. [F]
- 3. Ich bemühe mich, mit den aktuellen Trends und Moden mitzugehen. [F]

4. Ich finde es schwierig, anderen Dinge zu erklären, die mir leicht fallen, wenn sie es nicht auf Anhieb verstehen. [-]
5. Nachts träume ich meistens. [F]
6. **Es macht mir große Freude, mich um andere Leute zu kümmern.** [+]
7. Ich versuche lieber, meine Probleme alleine zu lösen, anstatt sie mit anderen zu besprechen. [F]
8. Ich finde es schwierig zu erkennen, was ich in einer sozialen Situation tun soll. [-]
9. Morgens bin ich Bestform.[F]
10. Ich bekomme oft gesagt, dass ich zu weit gegangen bin, um mich in einer Diskussion durchzusetzen. [-]
11. Es stört mich nicht besonders, wenn ich zu einem Treffen mit Freunden zu spät komme. [-]
12. Freundschaften und Beziehungen sind einfach zu kompliziert sind, deshalb belaste ich mich lieber nicht damit. [-]
13. Ich würde nie gegen ein Gesetz verstoßen, egal wie geringfügig. [F]
14. Ich finde es oft schwierig zu beurteilen, ob etwas höflich oder unhöflich ist. [-]
15. In einer Unterhaltung neige ich dazu, mich eher auf meine eigenen Gedanken zu konzentrieren, als darauf, was mein Zuhörer denken könnte. [-]
16. Ich ziehe praktische Scherze erzählten Witzen vor. [F]
17. Ich lebe eher für das Heute als für die Zukunft. [F]
18. Als ich ein Kind war, zerteilte ich gerne Würmer, um zu sehen, was passieren würde. [-]
19. **Ich bemerke leicht, wenn jemand etwas anderes sagt, als er meint.** [+]
20. Ich habe strenge Vorstellungen von Moral.[F]
21. Es fällt mir schwer zu verstehen, warum manche Dinge Leute so sehr aufregen. [-]
22. **Ich finde es leicht, mich in jemand anderen hineinzusetzen.** [+]
23. Ich glaube, dass gute Manieren das Wichtigste sind, was Eltern ihrem Kind beibringen können. [F]
24. Ich tue Dinge gerne spontan. [F]

25. Ich kann gut vorhersehen, wie sich jemand fühlen wird. [+]
26. **Ich bemerke es schnell, wenn sich jemand in einer Gruppe unwohl oder unbehaglich fühlt.** [+]
27. **Wenn ich etwas sage, das jemand anderen verletzt, dann glaube ich, dass es dessen Problem ist, und nicht meins.** [-]
28. Wenn mich jemand fragt, was ich von seiner Frisur halte, antworte ich wahrheitsgemäß, auch wenn sie mir nicht gefällt. [-]
29. Ich kann nicht immer verstehen, warum sich jemand durch eine Bemerkung verletzt gefühlt haben sollte. [-]
30. Leute sagen mir oft, ich sei sehr unberechenbar. [*F*]
31. Bei sozialen Anlässen stehe ich allgemein gern im Mittelpunkt. [*F*]
32. **Leute weinen zu sehen, macht mir nicht wirklich etwas aus.** [-]
33. Ich unterhalte mich gern über Politik. [*F*]
34. Ich bin sehr direkt, was manche Leute für unhöflich halten, auch wenn es keine böse Absicht ist. [-]
35. Im Allgemeinen finde ich soziale Situationen nicht verwirrend. [+]
36. **Andere sagen mir, ich sei gut darin zu verstehen, was sie fühlen und was sie denken.** [+]
37. Wenn ich mich mit anderen unterhalte, rede ich eher über deren Erfahrungen als über meine. [+]
38. Es erschüttert mich zu sehen, wenn ein Tier Schmerzen hat. [+]
39. **Ich kann Entscheidungen treffen, ohne mich von den Gefühlen anderer beeinflussen zu lassen.** [-]
40. Ich kann mich nicht entspannen, solange ich nicht alles geschafft habe, was ich für den Tag geplant habe. [*F*]
41. **Ich erkenne leicht, ob jemand sich für das, was ich sage, interessiert, oder davon gelangweilt ist.** [+]
42. **Es erschüttert mich, wenn ich in Nachrichtensendungen Leute leiden sehe.** [+]
43. **Freunde erzählen mir gewöhnlich von ihren Problemen, weil sie sagen, ich sei sehr verständnisvoll.** [+]
44. **Ich bemerke, wenn ich störe, auch wenn die andere Person es nicht sagt.** [+]

45. Ich beginne oft neue Hobbies, bin aber schnell davon gelangweilt und suche mir etwas anderes. [*F*]
46. Manchmal sagen mir Leute, ich sei beim Necken zu weit gegangen. [-]
47. Ich wäre zu ängstlich, um auf einer großen Achterbahn zu fahren. [*F*]
48. **Andere sagen häufig, ich sei unsensibel, auch wenn ich nicht immer verstehe warum.** [-]
49. Ich finde, wer neu in eine Gruppe kommt, muss sich selbst um Kontakt bemühen. [-]
50. **Wenn ich einen Film ansehe, bleibe ich in der Regel davon emotional unberührt.** [-]
51. Im Alltag ziehe ich es vor, gut organisiert zu sein und mache mir oft Pläne, was zu tun ist.[*F*]
52. **Ich kann mich schnell und intuitiv darauf einstellen, wie es jemandem geht.** [+]
53. Ich gehe nicht gern Risiken ein. [*F*]
54. **Ich kann leicht herausfinden, worüber jemand anderes vielleicht gern sprechen würde.** [+]
55. **Ich bemerke, ob jemand seine wahren Gefühle verbirgt.** [+]
56. Bevor ich eine Entscheidung treffe, wäge ich immer die Vor- und Nachteile ab.[*F*]
57. Ich erarbeite mir nicht bewusst die Regeln des sozialen Zusammenlebens. [+]
58. Ich kann gut vorhersehen, was jemand tun wird. [+]
59. **Ich neige dazu, von Problemen meiner Freunde emotional betroffen zu sein.** [+]
60. Ich kann im Allgemeinen die Meinung einer anderen Person anerkennen, auch wenn ich anderer Meinung bin. [+]

### A.1.3 Abschluss

Herzlichen Dank für das Ausfüllen des Fragebogens!

## A.2 Fremdbeurteilung

### A.2.1 Instruktion

Bitte geben Sie hier an, in welcher Beziehung, Sie zu [Name]/ derjenigen Person stehen, die Ihnen diesen Fragebogen gegeben hat.

[Name]/ Die Person von der ich diesen Fragebogen bekommen habe, ist mein/e:

Partner/in  Vater/ Mutter  Sohn/ Tochter  Freund/in  Arbeitskollege/in

Im weiteren Verlauf finden Sie eine Liste von Aussagen zu [Name]/ dieser Person. Bitte lesen Sie jede Aussage sehr gründlich und bewerten Sie, wie sehr Sie dieser Aussage zustimmen oder sie ablehnen.

Es gibt keine richtigen oder falschen Antworten und auch keine Fangfragen.

Es ist wichtig, dass jede Frage beantwortet wird, denn nur dann ist eine gültige Auswertung möglich.

### A.2.2 Items

Im Folgenden sind alle Items der Fremdbeurteilungsvariante aufgezählt. Beispiel-Items sind mit [B], Filler-Items mit [F], positiv gepolte Items mit [+] und negativ gepolte Items mit [-] markiert. Alle Items werden auf einer vierstufigen Skala mit 3 = „stimme voll und ganz zu“, 2 = „stimme eher zu“, 1 = „stimme eher nicht zu“ 0 = „stimme überhaupt nicht zu“ beantwortet.

- a) Ich könnte mir vorstellen, dass es [Name]/ihn/sie sehr aufregen würde, wenn er/sie nicht jeden Tag Musik hören könnte. [B]
  - b) [Name]/Er/Sie redet mit seinen/ihren Freunden lieber am Telefon, statt ihnen Briefe zu schreiben. [B]
  - c) Ich habe den Eindruck, [Name]/er/sie hat kein Bedürfnis, andere Teile der Welt zu bereisen. [B]
  - d) [Name]/Er/Sie liest lieber, als zu tanzen. [B]
1. **Ich habe den Eindruck, [Name]/er/sie erkennt leicht, ob jemand an einem Gespräch interessiert ist.** [+]
  2. Ich könnte mir vorstellen, dass [Name]/er/sie Tiere lieber mag als Menschen. [F]
  3. [Name]/Er/Sie geht mit den aktuellen Trends und Moden mit. [F]
  4. Ich habe den Eindruck, dass [Name]/er/sie es schwierig findet, anderen Dinge zu erklären, die ihm/ihr leicht fallen, wenn die anderen es nicht auf Anhieb verstehen. [-]
  5. Ich könnte mir vorstellen, dass [Name]/er/sie nachts meistens träumt. [F]

6. **Ich habe den Eindruck, dass es [Name]/ihm/ihr große Freude macht, sich um andere Leute zu kümmern.** [+]
7. Ich habe den Eindruck, [Name]/er/sie versucht lieber, seine/ihre Probleme alleine zu lösen, anstatt sie mit anderen zu besprechen. [F]
8. **Ich habe den Eindruck, [Name]/er/sie findet es schwierig zu erkennen, was er/sie in einer sozialen Situation tun soll.** [-]
9. Ich habe den Eindruck, [Name]/er/sie ist morgens in Bestform. [F]
10. **Ich finde, [Name]/er/sie geht oft zu weit, um sich in Diskussionen durchzusetzen.** [-]
11. **Ich habe den Eindruck, [Name]/ihn/sie stört es nicht besonders, wenn er/sie zu einem Treffen mit Freunden zu spät kommt.** [-]
12. **Ich habe den Eindruck, dass [Name]/ihm/ihr Freundschaften und Beziehungen einfach zu kompliziert sind und er/sie sich deshalb nicht damit belastet.** [-]
13. [Name]/Er/Sie würde nie gegen ein Gesetz verstoßen, egal wie geringfügig. [F]
14. **Ich habe den Eindruck, dass [Name]/er/sie es oft schwierig findet zu beurteilen, ob etwas höflich oder unhöflich ist.** [-]
15. Ich habe den Eindruck, in einer Unterhaltung neigt [Name]/er/sie dazu, sich eher auf seine/ihre eigenen Gedanken zu konzentrieren, als darauf, was sein/ihr Zuhörer denken könnte. [-]
16. Ich könnte mir vorstellen, dass [Name]/er/sie praktische Scherze erzählten Witzen vorzieht. [F]
17. Ich habe den Eindruck, [Name]/er/sie lebt eher für das Heute als für die Zukunft. [F]
18. Ich könnte mir vorstellen, dass [Name]/er/sie als er/sie ein Kind war gerne Würmer zerteilte, um zu sehen, was passieren würde. [-]
19. **Ich habe den Eindruck, [Name]/er/sie bemerkt es leicht, wenn jemand etwas anderes sagt, als er meint.** [+]
20. Ich habe den Eindruck, [Name]/er/sie hat strenge Vorstellungen von Moral. [F]
21. Ich habe den Eindruck, [Name]/ihm/ihr fällt es schwer zu verstehen, warum manche Dinge Leute so sehr aufregen. [-]
22. Ich habe den Eindruck, dass es [Name]/ihm/ihr leicht fällt, sich in jemand anderen hineinzuversetzen. [+]

23. Ich könnte mir vorstellen, dass [*Name*]/er/sie glaubt, dass gute Manieren das Wichtigste sind, was Eltern ihrem Kind beibringen können. [*F*]
24. Ich habe den Eindruck, [*Name*]/er/sie tut Dinge gerne spontan [*F*]
25. **Ich habe den Eindruck, dass [*Name*]/er/sie gut vorhersehen kann, wie sich jemand fühlen wird.** [+]
26. **Ich habe den Eindruck, dass [*Name*]/er/sie es schnell bemerkt, wenn sich jemand in einer Gruppe unwohl oder unbehaglich fühlt.** [+]
27. **Ich habe den Eindruck, dass [*Name*]/er/sie der Meinung ist, wenn er/sie etwas sagt, was jemand anderen verletzt, ist das dessen Problem und nicht seines/ihr.** [-]
28. Ich habe den Eindruck, wenn jemand [*Name*]/ihn/sie fragt, was er/sie von seiner Frisur hält, antwortet er/sie wahrheitsgemäß, auch wenn sie ihm/ihr nicht gefällt. [-]
29. **Ich habe den Eindruck, [*Name*]/er/sie kann nicht immer verstehen, warum sich jemand durch eine Bemerkung verletzt gefühlt haben sollte.** [-]
30. Leute sagen [*Name*]/ihm/ihr oft, er/sie sei sehr unberechenbar. [*F*]
31. Ich habe den Eindruck, dass [*Name*]/er/sie bei sozialen Anlässen allgemein gern im Mittelpunkt steht. [*F*]
32. **Ich habe den Eindruck, Leute weinen zu sehen, macht [*Name*]/ihm/ihr nicht wirklich etwas aus.** [-]
33. [*Name*]/Er/Sie unterhält sich gern über Politik. [*F*]
34. [*Name*]/Er/Sie ist sehr direkt, was manche Leute für unhöflich halten, auch wenn es keine böse Absicht ist. [-]
35. **Ich habe den Eindruck, dass [*Name*]/er/sie soziale Situationen im Allgemeinen nicht verwirrend findet.** [+]
36. **Ich finde, [*Name*]/er/sie ist gut darin zu verstehen, was ich fühle und denke.** [+]
37. Wenn [*Name*]/er/sie sich mit anderen unterhält, redet er/sie eher über deren Erfahrungen als über seine/ihre. [+]
38. Ich habe den Eindruck, [*Name*]/ihn/sie erschüttert es zu sehen, wenn ein Tier Schmerzen hat. [+]
39. Ich habe den Eindruck, [*Name*]/er/sie kann Entscheidungen treffen, ohne sich von den Gefühlen anderer beeinflussen zu lassen. [-]

40. Ich habe den Eindruck, [Name]/er/sie kann sich nicht entspannen, solange er/sie nicht alles geschafft hat, was er/sie für den Tag geplant hat. [F]
41. **Ich habe den Eindruck, [Name]/er/sie erkennt leicht, ob jemand sich für das, was er/sie sagt, interessiert, oder davon gelangweilt ist. [+]**
42. **Ich habe den Eindruck, dass es [Name]/ihn/sie erschüttert, wenn er/sie in Nachrichtensendungen Leute leiden sieht. [+]**
43. **Ich erzähle [Name]/ihm/sie gewöhnlich von meinen Problemen, weil er/sie sehr verständnisvoll ist. [+]**
44. **Ich habe den Eindruck, [Name]/er/sie bemerkt, wenn er/sie stört, auch wenn die andere Person es ihm/ihr nicht sagt. [+]**
45. [Name]/Er/Sie beginnt oft neue Hobbies, ist aber schnell davon gelangweilt und sucht sich etwas anderes. [F]
46. **Manchmal geht [Name]/er/sie beim Necken zu weit. [-]**
47. Ich könnte mir vorstellen, dass [Name]/er/sie zu ängstlich wäre, um auf einer großen Achterbahn zu fahren. [F]
48. **Ich könnte mir vorstellen, dass [Name]/er/sie häufig gesagt bekommt, dass er/sie unsensibel ist, auch wenn er/sie nicht immer versteht warum. [-]**
49. Ich habe den Eindruck, dass [Name]/er/sie findet, wer neu in eine Gruppe kommt, muss sich selbst um Kontakt bemühen. [-]
50. Ich habe den Eindruck, [Name]/er/sie bleibt in der Regel davon emotional unberührt, wenn er/sie einen Film ansieht. [-]
51. Ich habe den Eindruck, [Name]/er/sie ist im Alltag gut organisiert und macht sich oft Pläne, was zu tun ist.[F]
52. **Ich habe den Eindruck, [Name]/er/sie kann sich schnell und intuitiv darauf einstellen, wie es jemandem geht. [+]**
53. Ich habe den Eindruck, [Name]/er/sie geht nicht gern Risiken ein. [F]
54. **Ich habe den Eindruck, [Name]/er/sie kann leicht herausfinden, worüber jemand anderes vielleicht gern sprechen würde. [+]**
55. **Ich habe den Eindruck, [Name]/er/sie bemerkt, ob jemand seine wahren Gefühle verbirgt. [+]**
56. Ich könnte mir vorstellen, dass [Name]/er/sie immer die Vor- und Nachteile abwägt, bevor er eine Entscheidung trifft.[F]
57. Ich habe den Eindruck, [Name]/er/sie erarbeitet sich nicht bewusst die Regeln des sozialen Zusammenlebens. [+]



58. Ich habe den Eindruck, dass [*Name*]/er/sie gut vorhersehen kann, was jemand tun wird. [+]
59. Ich habe den Eindruck, [*Name*]/er/sie neigt dazu, von Problemen seiner/ihrer Freunde emotional betroffen zu sein. [+]
60. Ich habe den Eindruck, dass [*Name*]/er/sie im Allgemeinen die Meinung einer anderen Person anerkennen kann, auch wenn er/sie anderer Meinung ist. [+]

### A.2.3 Abschluss

Ihre Antworten werden streng vertraulich behandelt und weder an [*Name*]/ ihn/ sie noch an Dritte weitergegeben. Bitte geben Sie den Fragebogen im verschlossenen Briefumschlag zurück.

Herzlichen Dank für das Ausfüllen des Fragebogens!



# Anhang B

## Diagnosen

Tabelle B.1: Genaue Aufschlüsselung der vergebenen Diagnosen in SKID-I und II

<i>Teil 1</i>	BR	UA	UR
<b>Hauptdiagnose (Achse I)</b>			
Major Depression, einzelne Episode (296.2)	—	9.3%	2.2%
Major Depression rezidivierend (296.3)	—	86.0%	2.2%
Double Depression (Dysthymie 300.4 + MDE)	—	4.7%	—
remittierte Major Depression, einzelne Episode (296.26)	—	—	35.6%
remittierte Major Depression rezidivierend (296.36)	—	—	60%
Bipolar I, remittiert (DSM 296.x6)	54.8%	—	—
Bipolar II, remittiert (DSM 296.89)	45.2%	—	—

Tabelle B.2: Genaue Aufschlüsselung der vergebenen Diagnosen in SKID-I und II

<i>Teil 2</i>	BR	UA	UR
<b>Komorbide Diagnosen (Achse I)</b>			
kein Befund	59.5%	37.2%	42.2%
Soziale Phobie lifetime (300.23)	4.8%	9.3%	6.7%
Zwangsstörung aktuell (300.3)	2.4%	2.3%	—
Alkoholmissbrauch lifetime (305.00)	4.8%	2.3%	13.3%
Spezifische Phobie aktuell (300.29)	7.1%	11.6%	11.1%
Soziale Phobie aktuell (300.23)	2.4%	14.0%	6.7%
Spezifische Phobie lifetime (300.29)	16.7%	2.3%	—
Panikstörung mit Agoraphobie aktuell (300.21)	—	7.0%	—
Agoraphobie ohne Panikstörung aktuell (300.22)	2.4%	2.3%	4.4%
Generalisierte Angststörung aktuell (300.02)	—	4.7%	—
Schmerzstörung aktuell (307.8)	—	7.0%	8.9%
Panikstörung ohne Agoraphobie lifetime (300.01)	—	7.0%	6.7%
Bulimie lifetime (307.51)	—	7.0%	8.9%
Sedativaabhängigkeit lifetime (304.10)	—	—	2.2%
PTBS aktuell (309.81)	—	2.3%	2.2%

Tabelle B.3: Genaue Aufschlüsselung der vergebenen Diagnosen in SKID-I und II

<i>Teil 3</i>	BR	UA	UR
<b>Komorbide Diagnosen (Achse I)</b>			
Anorexia Nervosa lifetime (307.1)	2.4%	4.7%	2.2%
Bulimie aktuell (307.51)	—	4.7%	—
Panikstörung mit Agoraphobie lifetime (300.21)	2.4%	2.3%	6.7%
Alkoholabhängigkeit lifetime (303.90)	4.8%	4.7%	13.3%
Dysthymie lifetime (300.4)	—	—	2.2%
unspezifische somatoforme Störung aktuell (300.82)	—	4.7%	2.2%
Cannabismissbrauch aktuell (305.20)	2.4%	—	—
PTBS lifetime (309.81)	7.1%	4.7%	2.2%
Körperwahrnehmungsstörung (300.7)	—	4.7%	—
ADHS? (314.0)	—	2.3%	—
Generalisierte Angststörung lifetime (300.02)	2.4%	—	2.2%
Panikstörung ohne Agoraphobie aktuell (300.01)	—	2.3%	—
Zwangsstörung lifetime (300.3)	2.4%	—	—
Cannabisabhängigkeit lifetime (304.30)	2.4%	—	—

Tabelle B.4: Genaue Aufschlüsselung der vergebenen Diagnosen in SKID-I und II

<i>Teil 4</i>	BR	UA	UR
<b>Komorbide Diagnosen (Achse II)</b>			
kein Befund	95.2%	86.0%	95.6%
selbstunsicher vermeidend (301.82)	—	9.3%	4.4%
zwanghaft (301.4)	2.4%	2.3%	—
passiv aggressiv negativistisch (–)	2.4%	—	—
depressiv (–)	—	2.3%	2.2%



# Anhang C

## Medikation

Tabelle C.1: Aktuelle Psychopharmakologische Medikation der Patientengruppen, detaillierte Aufschlüsselung

<i>Teil 1</i>	BR	UA	UR
<b>Antidepressiva</b>			
Tetrazyklische Antidepressiva	2.4%	20.9%	4.4%
Trizyklische Antidepressiva	4.8%	14.0%	11.1%
Selektive Serotoninwiederaufnahmehemmer	21.4%	32.6%	15.6%
Selektive Noradrenalin- und Dopaminwiederaufnahmehemmer	2.4%	11.6%	2.2%
Selektive Noradrenalinwiederaufnahmehemmer	2.4%	2.3%	—
Serotonin-Noradrenalinwiederaufnahmehemmer (SSNRI)	16.7%	23.3%	13.3%
Melatoninantagonisten	4.8%	2.3%	2.2%
<b>Mood Stabilizer</b>			
Antiepileptika	35.7%	7.0%	2.2%
Lithium	40.5%	2.3%	—
<b>Neuroleptika</b>			
Atypische Neuroleptika	33.3%	20.9%	4.4%
Niederpotente Neuroleptika	14.3%	—	—
<b>Sedativa</b>			
Benzodiazepine	—	2.3%	—
Sonstige Sedativa	—	7.0%	4.4%







# Anhang D

## Ablauf

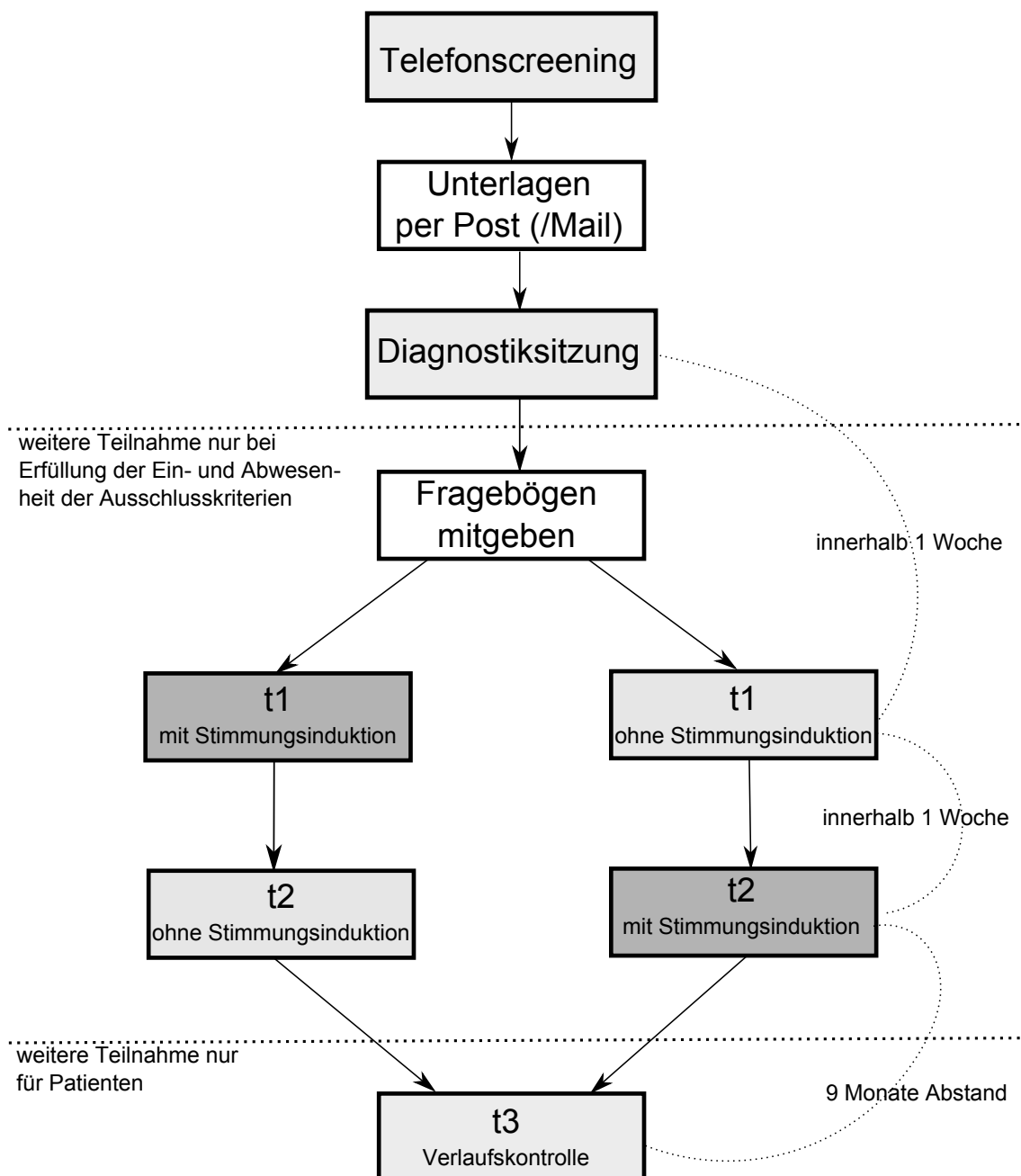


Abbildung D.1: Überblick über den Ablauf der Hauptstudie

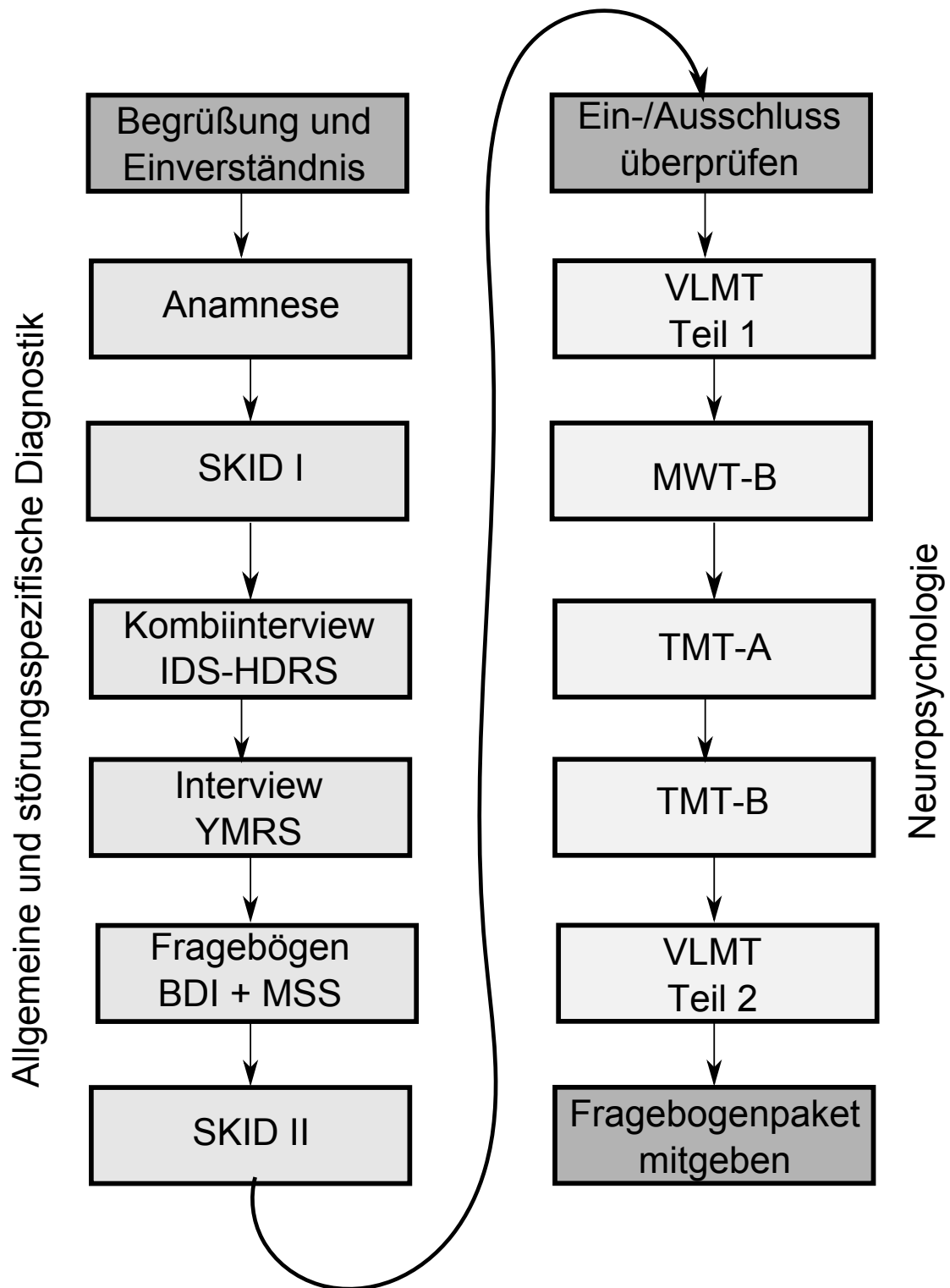


Abbildung D.2: Überblick über den Ablauf der Diagnostiksituation

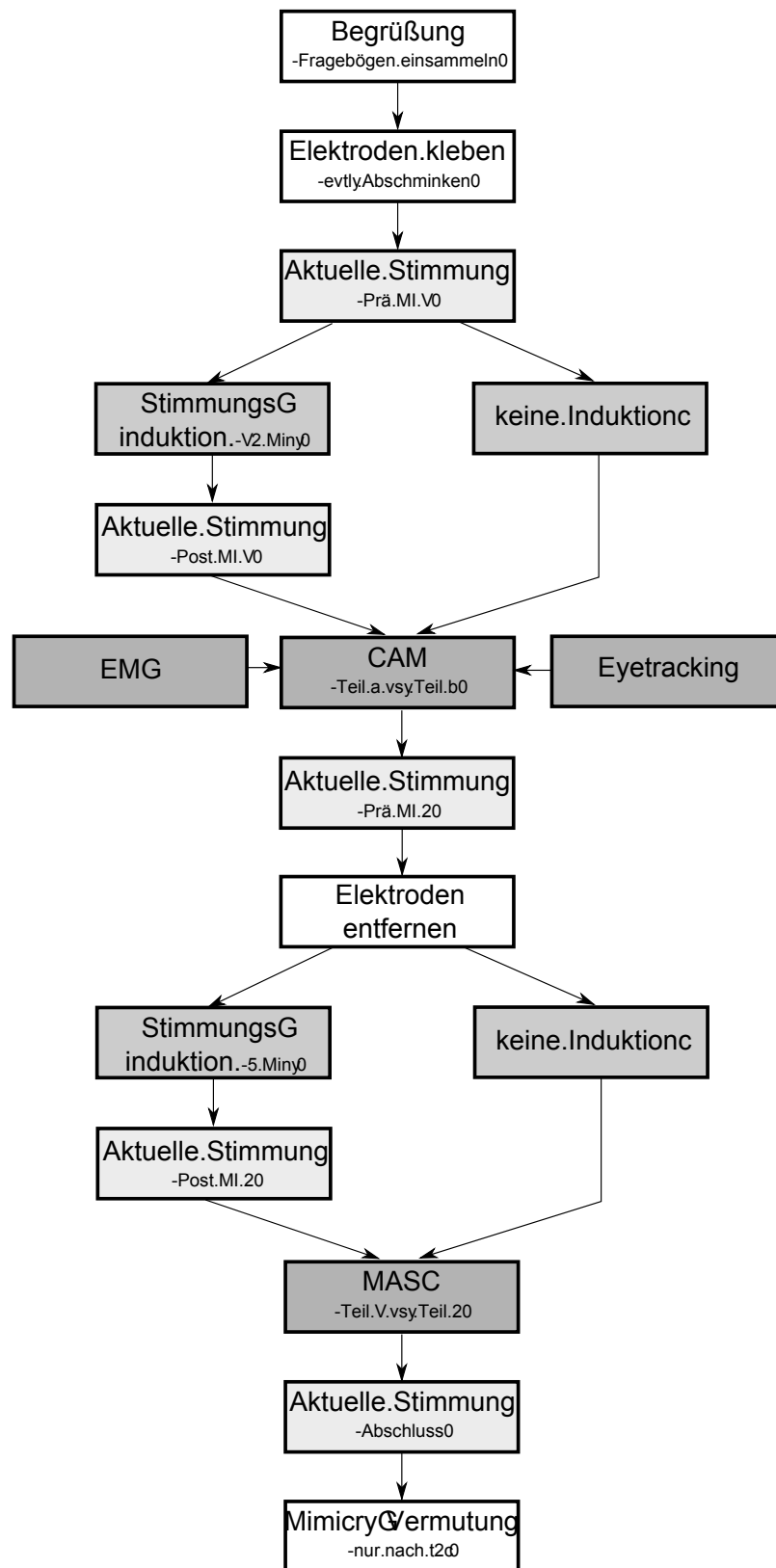


Abbildung D.3: Überblick über den Ablauf der Experimentalsitzung

# Anhang E

## Ergebnisse bei Verwendung der PCA-Auswertung

### E.1 Experimentelle Untersuchung: Fragestellung E2

In Tabelle E.1 sind die Summen nach PCA-Auswertung sowie zur besseren Vergleichbarkeit die Summe des EQ-F, wenn dieser nach den PCA-Regeln für den EQ-S ausgewertet wird, zu finden. In Tabelle E.2 und E.3 sind die Ergebnisse für die jeweiligen Unterskalen des EQ-S und EQ-F zu lesen.

Tabelle E.1: Summen des EQ-S und EQ-F nach PCA-Auswertung für die verschiedenen Teilnehmergruppen

	M	SD	Min	Max
<b>Kontrollgruppe UR:</b>				
EQ-S	22.8	6.80	8	35
EQ-F	32.0	10.1	9	46
EQ-F (analog zu EQ-S)	22.2	8.00	4	34
<b>Unipolar Remittiert:</b>				
EQ-S	20.2	5.77	8	30
EQ-F	26.6	11.7	3	53
EQ-F (analog zu EQ-S)	19.1	8.05	2	36
<b>Unipolar Akut:</b>				
EQ-S	20.5	6.84	6	34
EQ-F	30.1	8.79	14	49
EQ-F (analog zu EQ-S)	21.8	6.74	11	37
<b>Bipolar Remittiert:</b>				
EQ-S	22.1	7.20	5	36
EQ-F	27.0	9.93	3	48
EQ-F (analog zu EQ-S)	19.8	7.19	4	35
<b>Kontrollgruppe BR:</b>				
EQ-S	22.7	5.54	12	34
EQ-F	30.5	7.74	15	51
EQ-F (analog zu EQ-S)	20.7	5.95	9	33
<b>Gesamt:</b>				
EQ-S	21.6	6.52	5	36
EQ-F	29.1	9.94	3	53
EQ-F (analog zu EQ-S)	20.7	7.29	2	37

Tabelle E.2: Unterskalen des EQ-S für die verschiedenen Teilnehmergruppen

	M	SD	Min	Max
<b>Kontrollgruppe UR:</b>				
Faktor 1 kognitive Empathie	8.21	3.05	3	14
Faktor 2 emotionale Reaktivität	8.52	2.88	1	13
Faktor 3 empathische Betroffenheit	6.06	2.00	1	9
<b>Unipolar Remittiert:</b>				
Faktor 1 kognitive Empathie	7.15	2.57	1	14
Faktor 2 emotionale Reaktivität	7.38	3.25	1	14
Faktor 3 empathische Betroffenheit	5.69	1.99	0	10
<b>Unipolar Akut:</b>				
Faktor 1 kognitive Empathie	6.50	3.28	2	14
Faktor 2 emotionale Reaktivität	7.34	3.24	1	13
Faktor 3 empathische Betroffenheit	6.61	2.22	2	10
<b>Bipolar Remittiert:</b>				
Faktor 1 kognitive Empathie	7.92	3.22	0	14
Faktor 2 emotionale Reaktivität	8.10	3.39	1	14
Faktor 3 empathische Betroffenheit	6.08	1.98	2	10
<b>Kontrollgruppe BR:</b>				
Faktor 1 kognitive Empathie	8.90	2.88	4	14
Faktor 2 emotionale Reaktivität	8.34	2.60	3	14
Faktor 3 empathische Betroffenheit	5.48	1.98	2	10
<b>Gesamt:</b>				
Faktor 1 kognitive Empathie	7.66	3.09	0	14
Faktor 2 emotionale Reaktivität	7.90	3.12	1	14
Faktor 3 empathische Betroffenheit	6.01	2.05	0	10

Tabelle E.3: Unterskalen des EQ-F für die verschiedenen Teilnehmergruppen

	M	SD	Min	Max
<b>Kontrollgruppe UR:</b>				
Faktor 1 soziales Geschick	10.91	3.91	3	17
Faktor 2 kognitive Empathie	7.12	2.77	1	11
Faktor 3 emotionale Reaktivität	5.64	2.70	0	10
Faktor 4 empathische Betroffenheit	8.30	2.82	1	14
<b>Unipolar Remittiert:</b>				
Faktor 1 soziales Geschick	8.85	4.98	1	18
Faktor 2 kognitive Empathie	6.28	3.20	0	12
Faktor 3 emotionale Reaktivität	4.74	2.56	0	10
Faktor 4 empathische Betroffenheit	6.77	2.84	2	14
<b>Unipolar Akut:</b>				
Faktor 1 soziales Geschick	10.58	3.74	5	17
Faktor 2 kognitive Empathie	7.21	3.10	0	12
Faktor 3 emotionale Reaktivität	5.47	2.53	0	10
Faktor 4 empathische Betroffenheit	6.84	2.72	1	12
<b>Bipolar Remittiert:</b>				
Faktor 1 soziales Geschick	9.10	3.84	1	15
Faktor 2 kognitive Empathie	5.95	2.45	0	11
Faktor 3 emotionale Reaktivität	4.79	2.45	0	10
Faktor 4 empathische Betroffenheit	7.10	3.13	1	13
<b>Kontrollgruppe BR:</b>				
Faktor 1 soziales Geschick	10.17	3.91	2	18
Faktor 2 kognitive Empathie	7.17	2.51	1	12
Faktor 3 emotionale Reaktivität	4.97	2.34	0	10
Faktor 4 empathische Betroffenheit	8.14	2.13	3	11
<b>Gesamt:</b>				
Faktor 1 soziales Geschick	9.87	4.16	1	18
Faktor 2 kognitive Empathie	6.71	2.86	0	12
Faktor 3 emotionale Reaktivität	5.11	2.52	0	10
Faktor 4 empathische Betroffenheit	7.37	2.82	1	14

Mit dem Shapiro-Wilk-Test werden die EQ-Summen und Unterskalen auf Abweichungen von der Normalverteilung untersucht. Sowohl für die EQ-S Summe in der PCA-Auswertung als auch für die EQ-F Summe wenn diese analog zur PCA-Auswertung des EQ-S ausgewertet wird, findet sich in keiner der Teilnehmergruppen eine Verletzung der Normalverteilungsannahme. Die EQ-F Summe in der PCA-Auswertung hingegen scheint für KG-UR nicht normalverteilt zu sein ( $W = .94$ ,  $df = 33$ ,  $p = .049$ ). Bei den folgenden Unterskalen kann die Normalverteilung in allen Teilnehmergruppen als gegeben angenommen werden: EQ-S Faktor 2 emotionale Reaktivität, EQ-S Faktor 3 empathische Betroffenheit und EQ-F Faktor 3 emotionale Reaktivität. Eine Verletzung der Normalverteilungsannahme hingegen muss für



die folgenden Unterskalen angenommen werden: EQ-S Faktor 1 kognitive Empathie für UA ( $W = .94$ ,  $df = 38$ ,  $p = .047$ ) und für BR ( $W = .93$ ,  $df = 39$ ,  $p = .013$ ), EQ-F Faktor 1 soziales Geschick für UA ( $W = .93$ ,  $df = 38$ ,  $p = .020$ ), EQ-F Faktor 2 für KG-UR ( $W = .93$ ,  $df = 33$ ,  $p = .034$ ) sowie EQ-F Faktor 3 empathische Betroffenheit für KG-BR ( $W = .92$ ,  $df = 29$ ,  $p = .036$ ).

### E.1.1 Varianzanalysen zur Untersuchung von Gruppenunterschieden im EQ nach PCA-Auswertung

Analog zur Analyse der regulär ausgewerteten EQ-Summen wird für die Summen in der PCA-Auswertung ebenfalls eine multivariate Varianzanalyse mit der unabhängigen Variable Gruppenzugehörigkeit (fünfstufig) und den beiden abhängigen Variablen EQ-S Summe und EQ-F Summe (beide PCA-Auswertung) durchgeführt. Da der Box-Test auf Gleichheit der Kovarianzmatrizen nicht signifikant wird ( $p = .62$ ) kann die Homogenität der Kovarianzmatrizen als gegeben angenommen werden. Auch die Annahme der Varianzhomogenität kann bei nicht signifikantem Levene-Test als gegeben angenommen werden ( $p = .64$  für EQ-S und  $p = .17$  für EQ-F). Da die Normalverteilung der Summen in der PCA-Auswertung nicht für alle Gruppen als gegeben angenommen werden kann (siehe oben), wird die Voraussetzung der multivariaten Normalverteilung nicht erfüllt. Wie in Kapitel 4.6 dargelegt, wird dennoch an der parametrischen Auswertung festgehalten. Bei Verwendung der Pillai-Spur (ebenso wie Wilks-Lambda, Hotelling-Spur und Größter charakteristischer Wurzel nach Roy) zeigt sich kein signifikanter Effekt der Gruppenzugehörigkeit auf die Selbst- und Fremdbeurteilung der Empathie ( $V = 0.070$ ,  $F(8, 346) = 1.58$ ,  $p = .131$ ,  $f^2 = 0.036$ ).

Der Einfluss der Gruppenzugehörigkeit auf die Unterskalen des EQ-S und EQ-F wird im Rahmen einer multivariaten Varianzanalyse mit der unabhängigen Variable Gruppenzugehörigkeit (fünfstufig) und den sieben abhängigen Variablen EQ-S Unterskalen (3) und EQ-F Unterskalen (4) untersucht. Der Box-Test auf Gleichheit der Kovarianzmatrizen wird nicht signifikant ( $p = .62$ ), die Homogenität der Kovarianzmatrizen kann also als gegeben angenommen werden. Auch die Annahme der Varianzhomogenität kann bei nicht-signifikantem Levene-Test auf Gleichheit der Fehlervarianzen aufrecht erhalten werden ( $p > .05$  für alle abhängigen Variablen). Die multivariate Normalverteilung hingegen wird nicht erfüllt, da die Normalverteilung der einzelnen Variablen (siehe oben) nicht als gegeben angenommen werden kann. Wie oben beschrieben fällt dennoch die Entscheidung zur parametrischen Auswertung. Bei Verwendung der Pillai-Spur (ebenso wie Wilks-Lambda, Hotelling-Spur und Größter charakteristischer Wurzel nach Roy) zeigt sich ein signifikanter Effekt der Gruppenzugehörigkeit auf die Selbst- und Fremdbeurteilung der Empathie ( $V = 0.25$ ,  $F(28, 680) = 1.62$ ,  $p = .024$ ,  $f^2 = 0.067$ ). Dieser Einfluss der Gruppenzugehörigkeit wird im Rahmen von Post-Hoc-Tests weiter differenziert: In einzelnen Tests der Zwischensubjekteffekte zeigt sich einzig für den EQ-S Faktor 1 kognitive Empathie ein signifikanter Effekt der Gruppenzugehörigkeit ( $F(4, 173) = 3.25$ ,  $p = .013$ ,  $f = 0.26$ ). Für EQ-F Faktor 4 empathische Betroffenheit ist lediglich eine Tendenz erkennbar ( $F(4, 173) = 2.38$ ,  $p = .054$ ,  $f = 0.23$ ). Für alle anderen Unterskalen können keine signifikanten Effekte der Gruppenzugehörigkeit

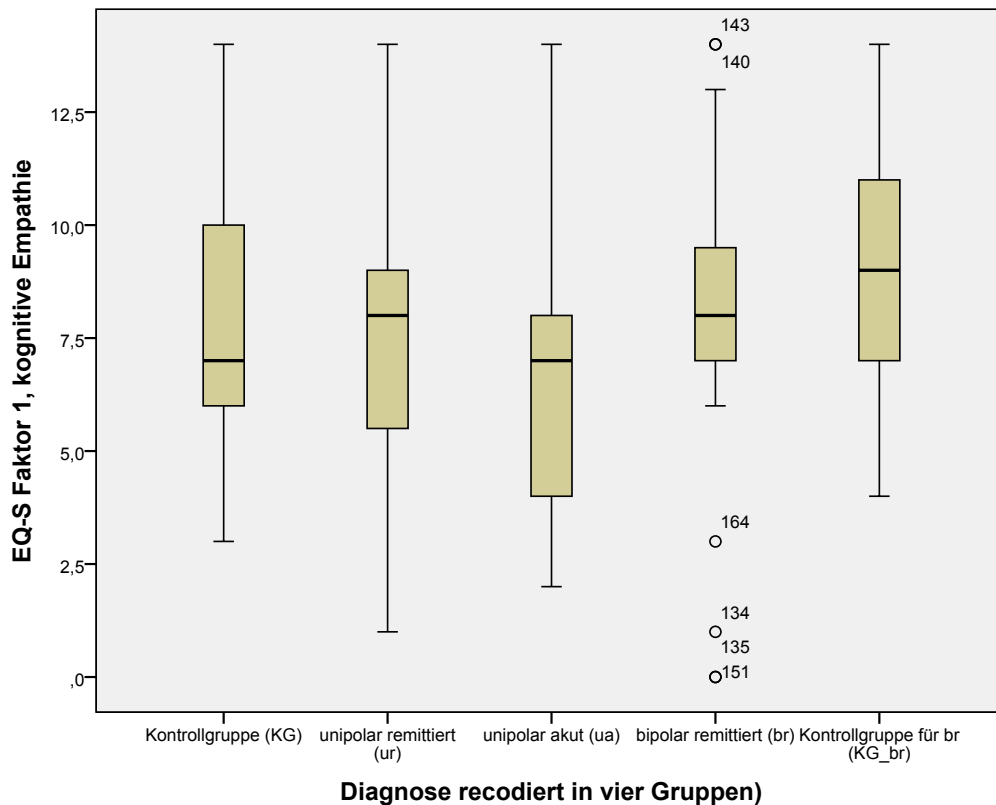


Abbildung E.1: Boxplot EQ-S Faktor 1 kognitive Empathie in den verschiedenen Teilnehmergruppen

gefunden werden (EQ-S Faktor 2 emotionale Reaktivität  $F(4, 173) = 1.08$ ,  $p = .37$ , EQ-S Faktor 3 empathische Betroffenheit  $F(4, 173) = 1.55$ ,  $p = .19$ , EQ-F Faktor 1 soziales Geschick  $F(4, 173) = 1.79$ ,  $p = .13$ , EQ-F Faktor 2 kognitive Empathie  $F(4, 173) = 1.58$ ,  $p = .18$ , EQ-F Faktor 3 emotionale Reaktivität  $F(4, 173) = 0.94$ ,  $p = .44$ ). Einzelvergleiche sprechen dafür dass sich nur KG-BR und UA bezüglich des EQ-S Faktor 1 kognitive Empathie signifikant unterscheiden ( $p = .015$ ,  $d = 0.78$ ). In Abbildung E.1 sind die Ergebnisse des EQ-S Faktor 1 als Boxplot dargestellt. Da es sich bei dem signifikanten Ergebnis um eine nicht zur Patientengruppe passende Kontrollgruppe handelt, und sich weder KG-BR von BR noch UA von KG-UR zu unterscheiden scheinen, ist dieser Befund wenig aussagekräftig. Es wird daher davon ausgegangen, dass es keinen signifikanten Einfluss der Gruppenzugehörigkeit auf die Unterskalen des EQ gibt.

## E.1.2 Regressionsanalyse zur Erklärung der EQ-Varianz

### EQ-S Summe nach PCA-Auswertung

Die oben genannten acht Prädiktoren werden zur Erklärung der EQ-S Summe in der PCA-Auswertung herangezogen, um mögliche Unterschiede zwischen der regu-

lären und der Auswertung nach den in der Validierungsstudie aufgestellten Regeln zu untersuchen. Multikollinearität liegt nicht vor und die Annahme der unabhängigen Fehler kann aufrecht erhalten werden (Durbin-Watson-Statistik = 2.05). Mit  $R^2 = 0.26$  ( $p < .001$ ,  $f^2 = 0.35$ ) erklären die Prädiktoren 25.9% der Varianz der EQ-S Summe nach PCA-Auswertung. Die Ergebnisse für die Prädiktoren sind in Tabelle E.4 zu sehen. Geschlecht, wahrgenommene soziale Unterstützung sowie soziale Belastung tragen signifikant zur Varianzaufklärung bei.

Tabelle E.4: Ergebnisse der multiplen Regression mit acht Prädiktoren für den nach Validierungsregeln ausgewerteten EQ-S

	<i>B</i>	<i>SE B</i>	$\beta$
(Konstante)	6.27	5.93	
Patient	-0.28	1.13	-.021
BDI	0.060	0.049	.10
Geschlecht	4.33	0.99	.30 ***
Regelmäßige Sozialkontakte	0.096	1.10	.006
IQ	-0.041	0.027	-.11
Gemittelte interpersonale Probleme	-1.77	1.14	-.16
Wahrgenommene soziale Unterstützung	2.93	0.86	.32 *
Soziale Belastung	1.58	0.76	.18 *

Signifikanzniveaus: <sup>+</sup>  $p < .1$ , \*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$

### EQ-F Summe nach PCA-Auswertung

Auch die EQ-F Summe nach PCA-Auswertung wird im Rahmen einer Regressionsanalyse mit den oben genannten acht Prädiktoren untersucht. Es liegt keine Multikollinearität vor und unabhängige Fehler können angenommen werden (Durbin-Watson-Statistik = 2.20). Die Regressionsgleichung erklärt mit  $R^2 = 0.17$  ( $p < .001$ ,  $f^2 = 0.20$ ) 16.8% der Varianz der EQ-F Summe nach Validierungsauswertungsregeln. Wie in Tabelle E.5 zu sehen ist, trägt das Geschlecht einen signifikanten Beitrag zur Varianzaufklärung bei, die gemittelten interpersonalen Probleme werden marginal signifikant.

Tabelle E.5: Ergebnisse der multiplen Regression mit acht Prädiktoren für den nach Validierung ausgewerteten EQ-F

	<i>B</i>	<i>SE B</i>	$\beta$
(Konstante)	27.3	10.2	
Gruppenzugehörigkeit	-0.74	1.93	-.037
BDI	0.11	0.082	.13
Geschlecht	3.99	1.70	.18 *
Regelmäßige Sozialkontakte	0.003	1.83	.001
IQ	-0.050	0.045	-.086
Gemittelte interpersonale Probleme	-3.50	1.93	-.21 +
Wahrgenommene soziale Unterstützung	1.92	1.48	.14
Soziale Belastung	-1.11	1.32	-.086

Signifikanzniveaus: +  $p < .1$ , \*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$ 

### E.1.3 Unterschiede zwischen Personen mit und ohne Fremdbeurteilung

Tabelle E.6: Homogenitätstest für Probanden mit und ohne EQ-F

		EQ-F vorhanden		EQ-F fehlt		$\chi^2$ -Test	
		M	SD	M	SD	$\chi^2$	<i>p</i>
EQ-S	Summe	21.6	6.52	20.6	4.84	11.9	.99
(PCA)							
EQ-S	Faktor 1	7.66	3.09	7.70	2.76	7.61	.91
EQ-S	Faktor 2	7.90	3.12	8.05	2.46	12.4	.50
EQ-S	Faktor 3	6.01	2.05	4.80	1.40	15.7	.11

## E.2 Experimentelle Untersuchung: Fragestellung E3

Es besteht eine mittelstarke Korrelation der EQ-Summen in der PCA-Auswertung ( $\rho = .34$ ,  $n = 178$ ,  $p < .001$ ). Des Weiteren korreliert die kognitive Empathie, also der EQ-S Faktor 1 mit dem EQ-F Faktor 2 mittelstark ( $\rho = .23$ ,  $n = 178$ ,  $p = .002$ ). Auch die emotionale Reaktivität, also EQ-S Faktor 2 und EQ-F Faktor 3 korreliert mittelstark ( $\rho = .38$ ,  $n = 178$ ,  $p < .001$ ). Einzig die empathische Betroffenheit gemessen durch EQ-S Faktor 3 und EQ-F Faktor 4 korreliert nicht signifikant zwischen Selbst- und Fremdbeurteilung.

Tabelle E.7: Deskriptive Beschreibung der Diskrepanz zwischen Selbst- und Fremdeinschätzung der Empathie

	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>
<b>PCA-Auswertung</b>				
Kontrollgruppe für UR	0.58	8.17	-18	14
Unipolar Remittiert	1.31	7.26	-12	21
Unipolar Akut	-1.39	7.39	-16	13
Bipolar Remittiert	2.31	7.91	-16	20
Kontrollgruppe für BR	2.03	6.45	-18	19
Gesamt	0.93	7.52	-18	21

### E.2.1 EQ-Diff für PCA-Auswertung

Die Auswertungsregeln nach der Validierungsstudie unterscheiden sich für EQ-S und EQ-F, daher wird hier der EQ-S in der PCA-Auswertung mit dem EQ-F in der PCA-Auswertung des EQ-S und nicht der eigenen PCA-Auswertung verglichen.

Da der Shapiro-Wilk-Test auf Normalverteilung in keiner der Gruppen signifikant wird ( $p > .05$  in allen Gruppen), wird davon ausgegangen, dass EQ-Diff in allen Teilnehmergruppen normalverteilt ist. Der Levene-Test spricht für Varianzgleichheit ( $F(4, 173) = 0.86, p = .49$ ). Eine einfaktorielle Varianzanalyse mit der Gruppenzugehörigkeit (fünfstufig) als unabhängiger Variable und der Diskrepanz zwischen Selbst- und Fremdeinschätzung als abhängiger Variable weist für EQ-Diff in der PCA-Auswertung nicht auf einen signifikanten Einfluss der Gruppenzugehörigkeit hin ( $F(4, 173) = 1.45, p = .22$ ).

## E.3 Experimentelle Untersuchung: Fragestellung E5

### E.3.1 Zusammenhang zwischen EQ und dem Ausmaß der Induzierbarkeit negativer Stimmung

Zwischen EQ-F Summe in der PCA-Auswertung und Induzierbarkeit von Emotionen findet sich eine mittelstark positive Korrelation ( $\rho = .201, n = 180, p = .007$ ). Der EQ-S Faktor 3 empathische Betroffenheit korreliert mittelstark negativ mit der Differenz negativer Stimmung ( $\rho = -.20, n = 198, p = .005$ ). Es findet sich eine mittelstark positive Korrelation zwischen der Differenz negativer Stimmung und dem EQ-S Faktor 2 kognitive Empathie ( $\rho = .21, n = 180, p = .004$ ). Der EQ-F Faktor 4 empathische Betroffenheit korreliert schwach positiv mit der Differenz negativer Stimmung ( $\rho = .16, n = 180, p = .035$ ). In der PCA-Auswertung korreliert die Differenz zwischen Selbst- und Fremdeinschätzung mittelstark negativ mit der Differenz negativer Stimmung ( $\rho = -.22, n = 178, p = .003$ ).

## E.4 Experimentelle Untersuchung: Fragestellung E6

### E.4.1 IIP

Die mit dem IIP gemessenen gemittelten interpersonalen Probleme korrelieren mit dem EQ-S schwach negativ in der PCA-Auswertung ( $\rho = .22$ ,  $n = 128$ ,  $p = .013$ ) und mittelstark negativ mit dem EQ-F in der PCA-Auswertung ( $\rho = -.28$ ,  $n = 171$ ,  $p < .001$ ). Signifikante Korrelationen der einzelnen Unterbereichen des IIP mit EQ-S oder EQ-F werden im Folgenden berichtet:

Die Unterskala zu streitsüchtig/konkurrieren korreliert leicht negativ mit dem EQ-S ( $\rho = -.17$ ,  $n = 189$ ,  $p = .018$ ) in der regulären Auswertung sowie mittelstark mit dem EQ-F in beiden Auswertungsvarianten (regulär:  $\rho = -.25$ ,  $n = 171$ ,  $p = .001$ ; PCA:  $\rho = -.22$ ,  $n = 171$ ,  $p = .004$ ). Zu abweisend korreliert mittelstark negativ mit dem EQ-S (regulär:  $\rho = -.31$ ,  $n = 189$ ,  $p < .001$ ; PCA:  $\rho = -.33$ ,  $n = 189$ ,  $p < .001$ ). Zu introvertiert/sozial vermeidend korreliert mittelstark negativ mit dem EQ-S (regulär:  $\rho = -.27$ ,  $n = 189$ ,  $p < .001$ ; PCA:  $\rho = -.30$ ,  $n = 189$ ,  $p < .001$ ) sowie schwach negativ mit dem EQ-F in der regulären Auswertung ( $\rho = -.18$ ,  $n = 171$ ,  $p = .021$ ). Es finden sich mittelstarke positive Korrelationen zwischen der Unterskala zu ausnutzbar/nachgiebig und dem EQ-S (regulär:  $\rho = .27$ ,  $n = 189$ ,  $p < .001$ ; PCA:  $\rho = .20$ ,  $n = 189$ ,  $p = .006$ ) sowie dem EQ-F (regulär:  $\rho = .28$ ,  $n = 171$ ,  $p < .001$ ; PCA:  $\rho = .26$ ,  $n = 171$ ,  $p = .001$ ). Zu fürsorglich/freundlich korreliert mittelstark positiv mit dem EQ-S (regulär:  $\rho = .23$ ,  $n = 189$ ,  $p = .001$ ; PCA:  $\rho = .25$ ,  $n = 189$ ,  $p = .001$ ) und schwach positiv mit dem EQ-F in der regulären Auswertung ( $\rho = .15$ ,  $n = 189$ ,  $p = .048$ ). Für die Unterskala zu expressiv/aufdringlich finden sich mit dem EQ-S in der regulären Auswertung leicht positive ( $\rho = .17$ ,  $n = 189$ ,  $p = .016$ ) und in der PCA-Auswertung mittelstark positive ( $\rho = .30$ ,  $n = 189$ ,  $p < .001$ ) Korrelationen. Die Unterskalen zu autokratisch/dominant und zu selbstunsicher/unterwürfig korrelieren nicht mit den EQ-Summen.

### E.4.2 F-SozU

Aus den drei Unterstützungsskalen des F-SozU lässt sich als Gesamtwert die wahrgenommene soziale Unterstützung berechnen. Diese korreliert mittelstark positiv sowohl mit EQ-S (regulär:  $\rho = .39$ ,  $n = 190$ ,  $p < .001$ ; PCA:  $\rho = .32$ ,  $n = 190$ ,  $p < .001$ ) als auch mit EQ-F (regulär:  $\rho = .33$ ,  $n = 171$ ,  $p < .001$ ; PCA:  $\rho = .33$ ,  $n = 171$ ,  $p < .001$ ). Einzeln betrachtet finden sich ebenfalls Korrelationen der Unterstützungsskalen (Hauptskalen) mit den EQ-Summen: Die emotionale Unterstützung korreliert mittelstark positiv mit dem EQ-S (regulär:  $\rho = .35$ ,  $n = 195$ ,  $p < .001$ ; PCA:  $\rho = .32$ ,  $n = 195$ ,  $p < .001$ ) und dem EQ-F (regulär:  $\rho = .26$ ,  $n = 176$ ,  $p = .001$ ; PCA:  $\rho = .27$ ,  $n = 176$ ,  $p < .001$ ). Auch die praktische Unterstützung korreliert mittelstark positiv mit dem EQ-S (regulär:  $\rho = .40$ ,  $n = 195$ ,  $p < .001$ ; PCA:  $\rho = .33$ ,  $n = 195$ ,  $p < .001$ ) sowie dem EQ-F (regulär:  $\rho = .29$ ,  $n = 175$ ,  $p < .001$ ; PCA:  $\rho = .29$ ,  $n = 175$ ,  $p < .001$ ). Ebenso die soziale Integration mit dem EQ-S (regulär:  $\rho = .35$ ,  $n = 195$ ,  $p < .001$ ; PCA:  $\rho = .25$ ,  $n = 195$ ,  $p = .001$ ) und dem EQ-F (regulär:  $\rho = .28$ ,  $n = 175$ ,  $p < .001$ ; PCA:  $\rho = .29$ ,  $n = 175$ ,  $p < .001$ ). Die vierte Hauptskala, die soziale Belastung korreliert mit dem EQ-F in der regulären

Auswertung leicht negativ ( $\rho = -.18$ ,  $n = 178$ ,  $p = .017$ ) und mittelstark negativ in der PCA-Auswertung ( $\rho = -.202$ ,  $n = 178$ ,  $p = .007$ )

Auch unter den Zusatzskalen und den Ergebnissen aus Teil B finden sich Korrelationen mit den EQ-Summen: Die Reziprozität korreliert mittelstark positiv mit EQ-S (regulär:  $\rho = .45$ ,  $n = 196$ ,  $p < .001$ ; PCA:  $\rho = .48$ ,  $n = 196$ ,  $p < .001$ ) sowie EQ-F (regulär:  $\rho = .23$ ,  $n = 176$ ,  $p = .002$ ; PCA:  $\rho = .23$ ,  $n = 176$ ,  $p = .002$ ). Auch die Zufriedenheit mit sozialer Unterstützung korreliert mittelstark positiv mit dem EQ-S in der regulären Auswertung ( $\rho = .19$ ,  $n = 197$ ,  $p = .006$ ) sowie dem EQ-F in beiden Auswertungsvarianten (regulär:  $\rho = .19$ ,  $n = 177$ ,  $p = .010$ ; PCA:  $\rho = .20$ ,  $n = 177$ ,  $p = .008$ ). Die Zusatzskala Vertrauensperson korreliert mittelstark positiv mit dem EQ-S (regulär:  $\rho = .34$ ,  $n = 198$ ,  $p < .001$ ; PCA:  $\rho = .32$ ,  $n = 198$ ,  $p < .001$ ) und dem EQ-F (regulär:  $\rho = .21$ ,  $n = 178$ ,  $p = .005$ ; PCA:  $\rho = .22$ ,  $n = 178$ ,  $p = .003$ ). Die durchschnittliche Anzahl unterstützender Personen insgesamt korreliert mit dem EQ-S in der regulären Auswertung mittelstark positiv ( $\rho = .22$ ,  $n = 172$ ,  $p = .004$ ) sowie schwach positiv in der PCA-Auswertung ( $\rho = .15$ ,  $n = 172$ ,  $p = .043$ ) und mit dem EQ-F (regulär:  $\rho = .16$ ,  $n = 156$ ,  $p = .049$ ; PCA:  $\rho = .16$ ,  $n = 156$ ,  $p = .041$ ). Die durchschnittliche Anzahl belastender Personen insgesamt korreliert leicht negativ mit dem EQ-F in der regulären Auswertung ( $\rho = -.18$ ,  $n = 165$ ,  $p = .024$ ) und mittelstark negativ mit dem EQ-F in der PCA-Auswertung ( $\rho = -.21$ ,  $n = 165$ ,  $p = .008$ ).





# Anhang F

## Ergebnisse die über die Klärung der Fragestellungen hinausgehen

### F.1 Fragebogenvalidierung: Unterschiede zwischen Personen mit und ohne Fremdbeurteilung

Im Rahmen von  $\chi^2$ -Test Homogenitätstests für die EQ-S-Gesamtsumme sowie die einzelnen Faktoren der Selbstbeurteilung können keine signifikanten Unterschiede zwischen Probanden mit und ohne vorhandenem Fremdbeurteilungs-Rating festgestellt werden. In Tabelle F.1 sind die deskriptiven Daten sowie die Teststatistiken dargestellt.

Tabelle F.1: Deskriptive Beschreibung des EQ-S für Probanden mit und ohne vorhandenem EQ-F sowie Ergebnisse des Homogenitätstests

	EQ-F vorhanden		EQ-F fehlt		$\chi^2$ -Test	
	M	SD	M	SD	$\chi^2$	<i>p</i>
EQ-S Summe	22.3	6.65	20.3	8.03	39.0	.15
EQ-S Faktor 1	7.65	2.90	6.33	3.53	16.6	.22
EQ-S Faktor 2	8.38	3.19	7.83	3.10	11.0	.69
EQ-S Faktor 3	6.25	2.31	6.17	2.73	6.51	.77

### F.2 Fragebogenvalidierung: Fragestellung V5

#### F.2.1 Häufigkeit der Sozialkontakte

Die Häufigkeit der Sozialkontakte weist keine signifikanten Korrelationen mit der Selbstbeurteilungsversion des EQ auf. Sie korreliert jedoch signifikant mit dem EQ-F Faktor 2 kognitive Empathie ( $\rho = .24$ ,  $n = 128$ ,  $p = .007$ ), mit dem EQ-F Faktor 3 emotionale Reaktivität ( $\rho = .24$ ,  $n = 128$ ,  $p = .008$ ) und mit dem EQ-F Faktor 4 empathische Betroffenheit ( $\rho = .18$ ,  $n = 128$ ,  $p = .041$ ).

### F.2.2 LSAS

Mit der LSAS werden soziale Angst sowie Vermeidung erfasst. Die soziale Angst an sich korreliert weder mit EQ-S ( $\rho = -.50$ ,  $n = 135$ ,  $p = .57$ ) noch mit dem EQ-F ( $\rho = .039$ ,  $n = 123$ ,  $p = .67$ ). Die Vermeidung jedoch korreliert schwach negativ mit dem EQ-S ( $\rho = -.25$ ,  $n = 135$ ,  $p = .004$ ), aber nicht mit dem EQ-F ( $\rho = -.064$ ,  $n = 124$ ,  $p = .48$ ). Die Gesamtsumme des EQ-S in der PCA-Auswertung und Vermeidung korrelieren schwach negativ ( $\rho = -.21$ ,  $n = 135$ ,  $p = .017$ ). Der Vermeidungs-Score korreliert ebenfalls schwach negativ mit dem EQ-S Faktor 1 kognitive Empathie ( $\rho = -.24$ ,  $n = 135$ ,  $p = .005$ ) sowie dem EQ-S Faktor 2 emotionale Reaktivität ( $\rho = -.26$ ,  $n = 135$ ,  $p = .002$ ). Die soziale Angst hingegen korreliert schwach positiv mit dem EQ-S Faktor 3 empathische Betroffenheit ( $\rho = .24$ ,  $n = 135$ ,  $p = .006$ ). Zwischen der Fremdbeurteilungsversion des EQ und der LSAS finden sich keine signifikanten Korrelationen.

### F.2.3 SIAS-SPS

Das Ergebnis des SIAS-SPS, der ebenfalls soziale Angst misst, korreliert schwach negativ mit dem EQ-S ( $\rho = -.19$ ,  $n = 129$ ,  $p = .030$ ), aber nicht mit dem EQ-F ( $\rho = -.151$ ,  $n = 118$ ,  $p = .10$ ). Ab einem Gesamtwert von 43 kann von klinisch relevanter sozialer Angst ausgegangen werden. Das Vorhandensein klinisch relevanter sozialer Angst korreliert mit dem EQ-S Faktor 1 kognitive Empathie ( $\rho = -.19$ ,  $n = 129$ ,  $p = .031$ ), mit dem EQ-S Faktor 2 emotionale Reaktivität ( $\rho = -.22$ ,  $n = 129$ ,  $p = .014$ ), mit der EQ-F Gesamtsumme in der PCA-Auswertung ( $\rho = -.21$ ,  $n = 118$ ,  $p = .022$ ), mit dem EQ-F Faktor 3 emotionale Reaktivität ( $\rho = -.19$ ,  $n = 118$ ,  $p = 0.045$ ) sowie mit dem EQ-F Faktor 4 empathische Betroffenheit ( $\rho = -0.260$ ,  $n = 118$ ,  $p = 0.004$ ). Signifikante Korrelationen mit der SIAS-Summe finden sich für den EQ-S Faktor 2 emotionale Reaktivität ( $\rho = -.21$ ,  $n = 129$ ,  $p = .019$ ), für den EQ-F Faktor 3 emotionale Reaktivität ( $\rho = -.21$ ,  $n = 118$ ,  $p = .023$ ) und den EQ-F Faktor 4 empathische Betroffenheit ( $\rho = -.24$ ,  $n = 118$ ,  $p = .010$ ).

### F.2.4 IMI

Mit dem IMI werden acht Faktoren des Circumplexmodells gemessen: Freundlich, Freundlich-Submissiv, Submissiv, Submissiv-Feindselig, Feindselig, Feindselig-Dominant, Dominant, Dominant-Freundlich. Im folgenden werden alle signifikanten Korrelationen mit dem EQ-S oder EQ-F berichtet. Die Subskala Freundlich korreliert mittelstark mit dem EQ-F ( $\rho = .53$ ,  $n = 125$ ,  $p < .001$ ). Die Subskala Freundlich-Submissiv korreliert schwach mit dem EQ-S ( $\rho = .21$ ,  $n = 127$ ,  $p = .020$ ) und dem EQ-F ( $\rho = .19$ ,  $n = 127$ ,  $p = .035$ ). Die Subskala Feindselig-Submissiv korreliert leicht negativ mit dem EQ-F ( $\rho = -.18$ ,  $n = 127$ ,  $p = .041$ ). Die Subskala Feindselig korreliert mittelstark negativ mit dem EQ-F ( $\rho = -.50$ ,  $n = 126$ ,  $p < .001$ ). Die Subskala Feindselig-Dominant korreliert mittelstark negativ mit dem EQ-F ( $\rho = -.29$ ,  $n = 124$ ,  $p = .001$ ). Die Subskala Freundlich-Dominant korreliert mittelstark mit dem EQ-F ( $\rho = .39$ ,  $n = 128$ ,  $p < .001$ ). Die Skala Freundlich korreliert signifikant mit der Gesamtsumme des EQ-S in der PCA-Auswertung ( $\rho = .19$ ,  $n = 125$ ,  $p = .036$ ), mit dem EQ-S Faktor 3 empathische Betroffenheit ( $\rho = .25$ ,  $n = 125$ ,  $p = .005$ ),

mit der Gesamtsumme des EQ-F in der PCA-Auswertung ( $\rho = .53$ ,  $n = 125$ ,  $p < .001$ ), mit dem EQ-F Faktor 1 soziales Geschick ( $\rho = .37$ ,  $n = 125$ ,  $p < .001$ ), mit dem EQ-F Faktor 2 kognitive Empathie ( $\rho = .37$ ,  $n = 125$ ,  $p < .001$ ), mit dem EQ-F Faktor 3 emotionale Reaktivität ( $\rho = .54$ ,  $n = 125$ ,  $p < .001$ ) und mit dem EQ-F Faktor 4 empathische Betroffenheit ( $\rho = .49$ ,  $n = 125$ ,  $p < .001$ ). Die Skala Freundlich-Submissiv korreliert signifikant mit der Gesamtsumme des EQ-S in der PCA-Auswertung ( $\rho = .20$ ,  $n = 127$ ,  $p = .027$ ), mit dem EQ-S Faktor 2 emotionale Reaktivität ( $\rho = .20$ ,  $n = 127$ ,  $p = .024$ ) sowie mit dem EQ-F Faktor 3 emotionale Reaktivität ( $\rho = .25$ ,  $n = 127$ ,  $p = .005$ ). Die Skala Submissiv korreliert signifikant mit dem EQ-F Faktor 4 empathische Betroffenheit ( $\rho = -.34$ ,  $n = 127$ ,  $p < .001$ ). Die Skala Feindselig-Submissiv korreliert signifikant mit der Gesamtsumme des EQ-F in der PCA-Auswertung ( $\rho = -.23$ ,  $n = 126$ ,  $p = .009$ ) und mit dem EQ-F Faktor 4 empathische Betroffenheit ( $\rho = -.33$ ,  $n = 127$ ,  $p < .001$ ). Die Skala Feindselig korreliert signifikant mit der Gesamtsumme des EQ-F in der PCA-Auswertung ( $\rho = -.53$ ,  $n = 126$ ,  $p < .001$ ), mit dem EQ-F Faktor 1 soziales Geschick ( $\rho = -.39$ ,  $n = 126$ ,  $p < .001$ ), mit dem EQ-F Faktor 2 kognitive Empathie ( $\rho = -.35$ ,  $n = 126$ ,  $p < .001$ ), mit dem EQ-F Faktor 3 emotionale Reaktivität ( $\rho = -.49$ ,  $n = 126$ ,  $p < .001$ ) und mit dem EQ-F Faktor 4 empathische Betroffenheit ( $\rho = -.48$ ,  $n = 126$ ,  $p < .001$ ). Die Skala Feindselig-Dominant korreliert signifikant mit der Gesamtsumme des EQ-F in der PCA-Auswertung ( $\rho = -.29$ ,  $n = 124$ ,  $p = .001$ ), mit dem EQ-F Faktor 1 soziales Geschick ( $\rho = -.44$ ,  $n = 124$ ,  $p < .001$ ) sowie mit dem EQ-F Faktor 2 kognitive Empathie ( $\rho = -.21$ ,  $n = 124$ ,  $p = .020$ ). Die Skala Dominant korreliert signifikant mit dem EQ-F Faktor 1 soziales Geschick ( $\rho = -.23$ ,  $n = 123$ ,  $p = .012$ ) und mit dem EQ-F Faktor 4 empathische Betroffenheit ( $\rho = .317$ ,  $n = 123$ ,  $p < .001$ ). Die Skala Freundlich-Dominant korreliert signifikant mit der Gesamtsumme des EQ-F in der PCA-Auswertung ( $\rho = .45$ ,  $n = 128$ ,  $p < .001$ ), mit dem EQ-F Faktor 2 kognitive Empathie ( $\rho = .35$ ,  $n = 128$ ,  $p < .001$ ), mit dem EQ-F Faktor 3 emotionale Reaktivität ( $\rho = .38$ ,  $n = 128$ ,  $p < .001$ ) sowie mit dem EQ-F Faktor 4 empathische Betroffenheit ( $\rho = .51$ ,  $n = 128$ ,  $p < .001$ ).

### F.2.5 IIP-Cc

Der IIP-Cc erfasst acht Unterskalen zu verschiedenen Bereichen und Problemen interpersonalem Verhaltens: Zu autokratisch/dominant, zu streitsüchtig/konkurrierend, zu abweisend/kalt, zu introvertiert/sozial vermeidend, zu selbstunsicher/unterwürfig, zu ausnutzbar/nachgiebig, zur fürsorglich/freundlich, zu expressiv/aufdringlich. Auch hier werden im Folgenden nur die signifikanten Korrelationen mit dem EQ-S oder dem EQ-F genannt: Zu kalt/abweisend korreliert mittelstark negativ mit dem EQ-S ( $\rho = -.50$ ,  $n = 140$ ,  $p < .001$ ) und schwach negativ mit dem EQ-F ( $\rho = -.19$ ,  $n = 128$ ,  $p = .028$ ). Zu fürsorglich/freundlich korreliert mittelstark mit dem EQ-S ( $\rho = .29$ ,  $n = 140$ ,  $p = 0.001$ ) und dem EQ-F ( $\rho = .29$ ,  $n = 128$ ,  $p = .001$ ). Zu introvertiert/sozial vermeidend korreliert schwach negativ mit dem EQ-S ( $\rho = -.19$ ,  $n = 140$ ,  $p = .025$ ). Die Unterskala kalt-abweisend korreliert signifikant mit der EQ-S Summe in der PCA-Auswertung ( $\rho = -.45$ ,  $n = 140$ ,  $p < .001$ ), mit dem EQ-S Faktor 1 kognitive Empathie ( $\rho = -.32$ ,  $n = 140$ ,  $p < .001$ ), mit dem EQ-S Faktor 2 emotionale Reaktivität ( $\rho = -.43$ ,  $n = 140$ ,  $p < .001$ ), mit dem EQ-S Faktor

empathische Betroffenheit ( $\rho = -.33$ ,  $n = 140$ ,  $p < .001$ ), mit der EQ-F Summe in der PCA-Auswertung ( $\rho = .26$ ,  $n = 128$ ,  $p = .003$ ) und mit dem EQ-F Faktor 4 empathische Betroffenheit ( $\rho = -.21$ ,  $n = 128$ ,  $p = .017$ ). Die Unterskala fürsorglich-freundlich korreliert signifikant mit der EQ-S Summe in der PCA Auswertung ( $\rho = .36$ ,  $n = 140$ ,  $p < .001$ ), mit dem EQ-S Faktor 2 emotionale Reaktivität ( $\rho = .36$ ,  $n = 140$ ,  $p < .001$ ), mit dem EQ-S Faktor empathische Betroffenheit ( $\rho = .42$ ,  $n = 140$ ,  $p < .001$ ) und mit dem EQ-F Faktor 1 soziales Geschick ( $\rho = .30$ ,  $n = 128$ ,  $p = .001$ ). Die Unterskala introvertiert-sozial vermeidend korreliert signifikant mit der EQ-S Summe in der PCA Auswertung ( $\rho = -.23$ ,  $n = 140$ ,  $p = .006$ ), mit dem EQ-S Faktor 1 kognitive Empathie ( $\rho = -.27$ ,  $n = 140$ ,  $p = .001$ ), mit dem EQ-S Faktor 2 emotionale Reaktivität ( $\rho = -.21$ ,  $n = 140$ ,  $p = .013$ ) und mit dem EQ-F Faktor 4 empathische Betroffenheit ( $\rho = -.28$ ,  $n = 128$ ,  $p = .001$ ).

## F.3 Experimentelle Untersuchung: Fragestellung E1

### F.3.1 Betrachtung der Unterskalen des IIP

Mit dem IIP lassen sich acht Unterskalen interpersonaler Probleme ermitteln. Mit einer multivariaten Varianzanalyse sollen Gruppenunterschiede bezüglich der ipsatierten Unterskalen untersucht werden. Die Ergebnisse des Shapiro-Wilk-Tests sprechen für eine Verletzung der Normalverteilungsannahme bei zwei Skalen für eine TN-Gruppe, bei einer Skala für zwei TN-Gruppen. Die übrigen fünf Unterskalen scheinen normalverteilt zu sein. Ein signifikanter Levene-Test ( $p < .05$ ) weist auf fehlende Varianzhomogenität in den Unterskalen zu streitsüchtig/konkurrierend, zu abweisend/kalt und zu introvertiert/sozial vermeidend hin. Für die übrigen fünf Unterskalen kann Varianzhomogenität angenommen werden. Die MANOVA mit Gruppenzugehörigkeit (fünfstufig) als unabhängiger Variable und den acht ipsatierten IIP-Unterskalen als abhängigen Variablen zeigt bei Verwendung der Pillai-Spur (ebenso wie Wilks-Lambda, Hotelling-Spur und Größte charakteristische Wurzel nach Roy) einen signifikanten Effekt der Gruppenzugehörigkeit auf die verschiedenen IIP-Unterskalen ( $V = 0.49$ ,  $F(28, 732) = 3.67$ ,  $p < .001$ ,  $f^2 = 0.14$ ). Post-Hoc berechnete univariate Varianzanalysen weisen auf signifikante Haupteffekte der Gruppenzugehörigkeit für die folgenden Unterskalen hin: zu autokratisch/dominant ( $F(4, 189) = 19.9$ ,  $p < .001$ ,  $f = 0.55$ ), zu streitsüchtig/konkurrierend ( $F(4, 189) = 3.79$ ,  $p = .005$ ,  $f = 0.29$ ), zu introvertiert/sozial vermeidend ( $F(4, 189) = 3.15$ ,  $p = .015$ ,  $f = 0.251$ ), zu selbstunsicher/unterwürfig ( $F(4, 189) = 7.86$ ,  $p < .001$ ,  $f = 0.39$ ) und zu fürsorglich/freundlich ( $F(4, 189) = 3.67$ ,  $p = .007$ ,  $f = 0.27$ ). Eine Tendenz zur Signifikanz lässt sich für zu ausnutzbar/nachgiebig ( $F(4, 189) = 2.01$ ,  $p = .095$ ,  $f = 0.20$ ) und für zu expressiv/aufdringlich ( $F(4, 189) = 2.13$ ,  $p = .079$ ,  $f = 0.22$ ) erkennen. Zu abweisend/kalt weist keine signifikanten Gruppenunterschiede auf ( $F(4, 189) = 0.67$ ,  $p = .61$ ,  $f = 0.12$ ). Die gefundenen Haupteffekte der Gruppenzugehörigkeit werden aufgrund teilweise fehlender Varianzhomogenität in Post-Hoc-Tests nach Games-Howell genauer betrachtet. Gruppenunterschiede mit  $p < .05$  werden als signifikant angesehen,  $p < .1$  weist auf eine Tendenz zu signifikanten Unterschieden hin. Nicht-signifikante Ergebnisse sowie irrelevante Vergleiche (z.B. UR und KG-BR) werden nicht berichtet. Für die Unterskala zu autokratisch/dominant finden

sich signifikante Unterschiede zwischen KG-UR und UA ( $d = 1.64$ ), UR und UA ( $d = 1.12$ ), BR und UA ( $d = 2.46$ ). Für zu streitsüchtig/konkurrierend kann eine Tendenz zu signifikanten Unterschieden zwischen UR und UA ( $d = 0.61$ ) beobachtet werden. Die Unterskala zu introvertiert/sozial vermeidend zeigt signifikante Unterschiede zwischen KG-UR und UA ( $d = 0.67$ ). Für zu selbstunsicher/unterwürfig zeigen sich signifikante Unterschiede zwischen KG-UR und UA ( $d = 1.02$ ) und zwischen BR und UA ( $d = 1.01$ ). Die Unterskala zu fürsorglich/freundlich weist signifikante Unterschiede zwischen UR und UA ( $d = 0.63$ ) auf.

### F.3.2 Betrachtung der Zusatzskalen und des Teil B des F-SozU

Folgend werden die Zusatzskalen Reziprozität, Zufriedenheit mit sozialer Unterstützung und Vertrauensperson sowie aus dem Teil B die durchschnittliche Anzahl unterstützender Personen insgesamt und die durchschnittliche Anzahl belastender Personen insgesamt betrachtet.

Die deskriptive Beschreibung der einzelnen Unterstützungsskalen, der Zusatzskalen und der Ergebnisse aus Teil B sind in Tabelle F.3 zu finden.

Im Rahmen einer multivariaten Varianzanalyse soll der Einfluss der Gruppenzugehörigkeit auf die drei verbleibenden Hauptskalen, die drei Zusatzskalen und die zwei Kennwerte aus Teil B untersucht werden. Aufgrund eines signifikanten Levene-Tests ( $p < 0.05$ ) muss für emotionale Unterstützung, soziale Integration, Zufriedenheit mit sozialer Unterstützung und Vertrauensperson eine Verletzung der Varianzhomogenität angenommen werden. Für Reziprozität besteht des Weiteren eine Tendenz ( $p < 0.1$ ) Varianzhomogenität abzulehnen. Der Box-Test auf Gleichheit der Kovarianzmatrizen wird signifikant ( $p < 0.01$ ), das heißt die Gleichheit der Kovarianzmatrizen kann also nicht als gegeben angenommen werden. Die Verwendung der Pillai-Spur (ebenso wie Wilks-Lambda, Hotelling-Spur und Größte charakteristische Wurzel nach Roy) zeigt einen signifikanten Effekt der Gruppenzugehörigkeit auf die abhängigen Variablen ( $V = 0.41$ ,  $F(32, 616) = 2.19$ ,  $p < .001$ ,  $f^2 = 0.11$ ). Post-Hoc werden einzelne Varianzanalysen zur Differenzierung des Effektes durchgeführt. Es zeigen sich signifikante Gruppenunterschiede für die emotionale Unterstützung ( $F(4, 158) = 6.59$ ,  $p < .001$ ,  $f = 0.38$ ), für die praktische Unterstützung ( $F(4, 158) = 5.67$ ,  $p < .001$ ,  $f = 0.36$ ), für die soziale Integration ( $F(4, 158) = 18.3$ ,  $p < .001$ ,  $f = 0.56$ ), für die Reziprozität ( $F(4, 158) = 4.24$ ,  $p = .003$ ,  $f = 0.31$ ), für die Zufriedenheit mit sozialer Unterstützung ( $F(4, 158) = 12.0$ ,  $p < .001$ ,  $f = 0.48$ ), für die Vertrauensperson ( $F(4, 158) = 4.13$ ,  $p = .003$ ,  $f = 0.31$ ) und für die durchschnittliche Anzahl belastender Personen insgesamt ( $F(4, 158) = 2.78$ ,  $p = .029$ ,  $f = 0.26$ ). Einzig für die Anzahl unterstützender Personen finden sich keine Hinweise auf signifikante Gruppenunterschiede. Aufgrund fehlender Varianzhomogenität werden die Gruppenunterschiede anhand von Post-Hoc-Tests nach Games-Howell genauer untersucht. Unterschiede werden bei  $p < 0.05$  als signifikant und bei  $p < 0.1$  als tendenziell signifikant bezeichnet. Für die emotionale Unterstützung zeigt sich ein signifikanter Unterschied zwischen KG-UR und UA ( $d = 1.13$ ) sowie eine Tendenz zu einem signifikanten Unterschied zwischen KG-UR und UR ( $d = 0.65$ ). Bezüglich der praktischen Unterstützung unterscheiden sich KG-UR und UA ( $d = 0.89$ ) signifikant

Tabelle F.2: F-Soz U weitere Skalen deskriptiv

<i>Teil 1</i>	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>
<b>Emotionale Unterstützung (Hauptskala)</b>				
Kontrollgruppe für UR	4.47	0.42	3.50	5.00
Unipolar Remittiert	4.11	0.66	2.13	4.94
Unipolar Akut	3.74	0.81	2.13	5.00
Bipolar Remittiert	3.81	0.73	2.19	4.88
Kontrollgruppe für BR	4.23	0.77	2.13	5.00
<b>Praktische Unterstützung (Hauptskala)</b>				
Kontrollgruppe für UR	4.22	0.55	2.89	5.00
Unipolar Remittiert	3.89	0.67	2.22	5.00
Unipolar Akut	3.60	0.83	1.67	5.00
Bipolar Remittiert	3.55	0.74	2.11	4.78
Kontrollgruppe für BR	4.07	0.69	2.56	5.00
<b>Soziale Integration (Hauptskala)</b>				
Kontrollgruppe für UR	4.23	0.35	3.62	4.92
Unipolar Remittiert	3.66	0.72	1.46	4.77
Unipolar Akut	3.07	0.88	1.38	4.69
Bipolar Remittiert	3.27	0.68	1.77	4.38
Kontrollgruppe für BR	4.11	0.66	2.38	4.92
<b>Reziprozität (Zusatzskala)</b>				
Kontrollgruppe für UR	3.95	0.57	2.25	5.00
Unipolar Remittiert	3.51	0.63	2.25	5.00
Unipolar Akut	3.40	0.78	1.75	4.50
Bipolar Remittiert	3.57	0.72	2.50	5.00
Kontrollgruppe für BR	3.91	0.79	2.50	5.00
<b>Zufriedenheit mit sozialer Unterstützung (Zusatzskala)</b>				
Kontrollgruppe für UR	4.16	0.56	3.00	5.00
Unipolar Remittiert	3.67	0.81	1.80	4.80
Unipolar Akut	3.02	0.99	1.20	5.00
Bipolar Remittiert	3.17	0.95	1.20	5.00
Kontrollgruppe für BR	3.99	0.81	1.60	5.00
<b>Vertrauensperson (Zusatzskala)</b>				
Kontrollgruppe für UR	4.63	0.54	3.00	5.00
Unipolar Remittiert	4.21	0.82	1.50	5.00
Unipolar Akut	3.90	0.99	1.25	5.00
Bipolar Remittiert	3.98	0.97	1.25	5.00
Kontrollgruppe für BR	4.40	0.92	1.00	5.00

Tabelle F.3: F-Soz U weitere Skalen deskriptiv

<i>Teil 2</i>	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>
<b>Durchschnittliche Anzahl unterstützender Personen insgesamt (Teil B)</b>				
Kontrollgruppe für UR	4.64	2.73	0.83	14.00
Unipolar Remittiert	4.07	2.57	0.50	12.00
Unipolar Akut	3.36	1.95	0.67	9.00
Bipolar Remittiert	3.68	1.83	1.00	8.50
Kontrollgruppe für BR	4.81	3.72	0.33	19.33
<b>Durchschnittliche Anzahl belastender Personen insgesamt (Teil B)</b>				
Kontrollgruppe für UR	0.42	0.48	0.00	1.75
Unipolar Remittiert	0.89	1.37	0.00	8.25
Unipolar Akut	1.10	1.62	0.00	9.75
Bipolar Remittiert	1.25	0.97	0.00	4.00
Kontrollgruppe für BR	0.61	0.86	0.00	3.00

sowie KG-BR und BR ( $d = 0.73$ ) tendenziell signifikant. Bei der sozialen Integration finden sich signifikante Unterschiede zwischen KG-UR und UR ( $d = 1.01$ ), zwischen KG-UR und UA ( $d = 1.74$ ), zwischen UA und UR ( $d = 0.74$ ) sowie zwischen KG-BR und BR ( $d = 1.26$ ). Für die Reziprozität zeigen sich signifikante Unterschiede zwischen KG-UR und UR ( $d = 0.74$ ) und zwischen KG-UR und UA ( $d = 0.81$ ). Die Zufriedenheit mit sozialer Unterstützung weist signifikante Unterschiede zwischen KG-UR und UR ( $d = 0.71$ ), zwischen KG-UR und UA ( $d = 1.43$ ), zwischen UR und UA ( $d = 0.72$ ) sowie zwischen KG-BR und BR ( $d = 0.93$ ) auf. Bezüglich der Vertrauensperson unterscheiden sich KG-UR und UA ( $d = 0.91$ ) signifikant. Der einzige signifikante Gruppenunterschied für die durchschnittliche Anzahl belastender Personen besteht zwischen KG-UR und BR, also einem irrelevanten Vergleich, da für BR eine eigene Kontrollgruppe vorhanden ist. Eine Tendenz zu signifikanten Unterschieden besteht zwischen BR und ebendieser KG-BR ( $d = 0.70$ ).

## F.4 Experimentelle Untersuchung: Fragestellung E4

### F.4.1 Negativity/Positivity Bias in der CAM gesamt

Der Negativity Bias bezeichnet die Verzerrung der Antworten in der Emotionserkennungsaufgabe in Richtung vermehrt negativer Emotionen. Errechnet wird die Negativity aus dem Anteil richtiger Antworten der negativen Items. Entsprechend bezeichnet der Positivity Bias eine Verzerrung in Richtung positiver Emotionen und wird aus dem Anteil richtiger Antworten positiver Items errechnet.

Die Negativity scheint anhand des Shapiro-Wilk-Tests in vier Gruppen nicht normalverteilt zu sein ( $p < .05$ ), UR weist eine Tendenz zur Ablehnung der Normalverteilungsannahme auf ( $p < .10$ ). Für die Positivity kann die Normalverteilungsannahme in vier Gruppen aufrecht erhalten werden ( $p > .10$ ), nur für BR muss sie abgelehnt werden ( $p < .05$ ). Damit scheint die multivariate Normalverteilung, de-

ren Voraussetzung einzelne Normalverteilung ist, nicht gegeben. Der Box-Test auf Gleichheit der Kovarianzmatrizen wird signifikant ( $p = .019$ ), die Homogenität der Kovarianzmatrizen kann daher nicht als gegeben angenommen werden. Der Levene-Test auf Gleichheit der Fehlervarianzen ist für alle Gruppen für Negativity sowie Positivity nicht signifikant ( $p > .05$ ) das heißt, die Annahme der Varianzhomogenität kann aufrecht erhalten werden. Bei Verwendung der Pillai-Spur (ebenso wie Wilks-Lambda, Hotelling-Spur und Größte charakteristische Wurzel nach Roy) zeigt sich ein signifikanter Effekt der Gruppenzugehörigkeit auf Positivity oder Negativity ( $V = 0.077$ ,  $F(8, 396) = 1.98$ ,  $p = .048$ ,  $f^2 = 0.040$ ). Post-Hoc durchgeführte einzelne ANOVAS (Tests der Zwischensubjekteffekte) zeigen keinen signifikanten Gruppenunterschied bezüglich der Negativity ( $F(4, 198) = 0.93$ ,  $p = .45$ ), aber einen signifikanten Unterschied für die Positivity ( $F(4, 198) = 2.59$ ,  $p = .038$ ,  $f = 0.22$ ). In Post-Hoc-Tests nach Games-Howell lassen sich jedoch keine signifikanten Unterschiede zwischen einzelnen Teilnehmergruppen zeigen.

## F.5 Experimentelle Untersuchung: Fragestellung E5

### F.5.1 Negativity/Positivity Bias in der CAM

Eine Mixed-Design ANOVA zur Prüfung des Einflusses der Stimmungsinduktion und der Gruppenzugehörigkeit auf die Negativity (Box-Test nicht signifikant:  $p = .15$ ; Mauchly-Test aufgrund zweistufigen Messwiederholungsfaktors nicht berechnet; Levene-Test zu keinem Zeitpunkt signifikant  $p > .1$ ) zeigt keinen signifikanten Haupteffekt der Stimmungsinduktion ( $F(1, 198) = 0.13$ ,  $p = .72$ ), keinen signifikanten Haupteffekt der Gruppenzugehörigkeit ( $F(4, 198) = 0.93$ ,  $p = .45$ ) und keine signifikante Interaktion zwischen Stimmungsinduktion und Gruppenzugehörigkeit ( $F(4, 198) = 0.15$ ,  $p = .96$ ).

Die Varianzanalyse für die Positivity mit Stimmungsinduktion und Gruppenzugehörigkeit als unabhängige Variablen (Box-Test nicht signifikant:  $p = .33$ ; Mauchly-Test aufgrund zweistufigen Messwiederholungsfaktors nicht berechnet; Levene-Test zu keinem Zeitpunkt signifikant  $p > .1$ ) zeigt weder einen signifikanten Haupteffekt der Stimmungsinduktion ( $F(1, 198) = 0.98$ ,  $p = .32$ ) noch eine signifikante Interaktion zwischen Stimmungsinduktion und Gruppenzugehörigkeit ( $F(4, 198) = 0.21$ ,  $p = .93$ ). Es gibt jedoch Hinweise für einen signifikanten Haupteffekt der Gruppenzugehörigkeit ( $F(4, 198) = 2.59$ ,  $p = .038$ ,  $f = 0.20$ ). Post-Hoc zeigen sich jedoch weder in Bonferroni-korrigierten paarweisen Vergleichen noch nach Gabriel signifikante Unterschiede zwischen einzelnen Gruppen.