

# **Entwicklung und Überprüfung eines Fragebogens zur Erfassung relevanter Therapiebedingungen (FERT)**

## **Dissertation**

der Fakultät für Informations- und Kognitionswissenschaften  
der Eberhard-Karls-Universität Tübingen  
zur Erlangung des Grades eines  
Doktors der Naturwissenschaften  
(Dr. rer. nat.)

vorgelegt von  
Dipl.-Psych. Konstanze Vollmann  
aus Stuttgart

Tübingen  
2009

Tag der mündlichen Qualifikation:

10. Februar 2010

Dekan:

Prof. Dr. Oliver Kohlbacher

1. Berichterstatter:

Prof. Dr. Martin Hautzinger

2. Berichterstatterin:

PD Dr. Ute Strehl

Meinen Eltern

---

## Danksagung

An dieser Stelle möchte ich mich von Herzen bei all denjenigen bedanken, die dazu beigetragen haben, dass ich das Projekt „Promotion“ in Angriff genommen, durchgeführt und nun auch zu Ende gebracht habe.

Mein besonderer Dank gilt

... *PD Dr. Ute Strehl* für ihre umsichtigen und hilfreichen Rückmeldungen, die freundliche Beharrlichkeit, mit der sie mich immer motiviert hat, und ihre sonstige spezifische und unspezifische Unterstützung,

... *Prof. Dr. Martin Hautzinger* für seine wohlwollende Haltung und die stets vorhandene Bereitschaft, bei anstehenden Fragen zu helfen,

... *der Christoph-Dornier-Stiftung* für die Bereitstellung der strukturellen Rahmenbedingungen und die Förderung des Austausches zwischen den Mitarbeitern und Doktoranden der verschiedenen Institute,

... *Dr. Angelika Glöckner-Rist* und *Prof. Dr. Herbert Hoijtink* dafür, dass sie mich beraten, meine methodischen Kenntnisse erweitert und mir die Vorzüge des Programms m-Plus näher gebracht haben,

... meinem Lebensgefährten *Marcus Feichter* für seine Unterstützung, seine Geduld und seinen Einsatz, vor allem bei der Fertigstellung der Arbeit,

... allen Freunden und Kollegen, die mir in den verschiedenen Phasen der Arbeit mit Rat und Tat zur Seite gestanden sind,

... allen Therapeuten und Patienten, die durch ihre Bereitschaft zur Mitarbeit an diesem Projekt sein Gelingen überhaupt erst möglich gemacht haben,

... und nicht zuletzt meinen Eltern für all ihre Liebe, ihre unermüdliche Unterstützung und dafür, dass sie mir die Freude am wissenschaftlichen Arbeiten vermittelt haben.

---

# Inhaltsverzeichnis

I. Einleitung.....	7
II. Theorieteil.....	8
1. Psychotherapieforschung.....	8
1.1. Grundüberlegungen.....	8
1.2. Kontrollgruppen in der Psychotherapieforschung.....	9
2. Das Placebo-Konzept.....	11
2.1. Das Placebo-Konzept in der Medizin.....	11
2.2. Das Placebo-Konzept in der Psychotherapie.....	13
2.3. Placebo-Faktoren versus unspezifische/allgemeine Wirkfaktoren.....	13
3. Unspezifische/allgemeine Wirkfaktoren in der Psychotherapie.....	16
3.1. Theoretische Konzeptionen allgemeiner Wirkfaktoren.....	16
3.1.1. Das Modell von Frank (1961).....	16
3.1.2. Das Modell von Garfield (1982).....	17
3.1.3. Das Modell von Orlinsky & Howard (1987).....	17
3.1.4. Das Modell von Grawe et al. (1994; Grawe, 2000, 2005).....	18
3.1.5. Vergleich der Modelle .....	19
3.1.6. Das Modell von Schulte (1996).....	21
3.1.7. Der Ansatz von Wampold (2005).....	22
3.1.8. Fazit.....	23
3.2. Das Konzept der therapeutischen Beziehung.....	24
3.3. Empirische Untersuchungen zu allgemeinen Wirkfaktoren.....	30
3.3.1. Zur Effektivität allgemeiner Wirkfaktoren: Metaanalysen .....	30
3.3.2. Zur Effektivität allgemeiner Wirkfaktoren: Studien zu einzelnen Variablen.....	37
3.3.3. Zusammenfassung.....	44
3.4. Messinstrumente zur Erfassung allgemeiner Wirkfaktoren.....	45
3.4.1. Fazit.....	48
III. Empirischer Teil .....	49
1. Konzeption des Fragebogens zur Erfassung allgemeiner Wirkfaktoren.....	49
1.1. Vorgehen.....	49
1.2. Formale Aspekte.....	49
1.2.1. Wahl des Verfahrens.....	49
1.2.2. Vorteile der Fragebogenmethode.....	50
1.2.3. Schwächen der Fragebogenmethode und Lösungsansätze.....	50
1.2.4. Wahl der Perspektive.....	52
1.3. Inhaltliche Aspekte.....	53
1.3.1. Festlegung der Fragebogenskalen.....	53
1.3.2. Itemsammlung.....	55
2. Expertenbefragung.....	56
2.1. Ziel und Vorgehensweise.....	56
2.2. Stichprobenbeschreibung.....	57
2.3. Auswertung: Itemvorauswahl und Skalen-Neubestimmung.....	58
3. Überarbeitungen und Erstellung der ersten Version des Fragebogens.....	67

---

3.1. Überarbeitungen der Formulierungen.....	67
3.2. Erstellen der 1. Version des Fragebogens.....	67
3.2.1. Antwortformat.....	67
3.2.2. Festlegung des Bezugszeitraumes.....	69
3.2.3. Itemreihung und Ergänzung.....	70
4. Fragebogenuntersuchung.....	70
4.1. Stichprobe.....	70
4.1.1. Kriterien für die Stichprobengewinnung.....	70
4.1.2. Vorgehen.....	73
4.1.3. Beschreibung der Stichprobe.....	75
5. Auswertung und Ergebnisse.....	86
5.1. Deskriptive Auswertung und Berechnung der Itemschwierigkeiten.....	86
5.1.1. Vorgehen.....	86
5.1.2. Ergebnisse .....	87
5.2. Faktorenanalysen.....	92
5.2.1. Vorgehen.....	92
5.2.2. Explorative Faktorenanalyse.....	97
5.2.3. Konfirmatorische Faktorenanalyse.....	102
5.2.4. Ergebnisse .....	112
6. Darstellung der revidierten Form des Fragebogens.....	140
6.1. Darstellung der Skalen und Items .....	140
6.2. Beurteilung des endgültigen Fragebogens.....	143
6.2.1. Gütekriterien.....	143
6.2.2. Inhaltliche Darstellung.....	149
6.2.3. Vergleich mit den A-priori-Skalen.....	151
6.3. Bezeichnung des Fragebogens.....	152
IV. Diskussion und Ausblick.....	154
1. Diskussion der Vorgehensweise.....	154
1.1. Konzeption.....	154
1.2. Expertenbefragung.....	156
1.3. Fragebogenuntersuchung.....	157
1.3.1. Durchführung.....	157
1.3.2. Auswertung.....	159
2. Diskussion der Ergebnisse und Ausblick.....	161
3. Zusammenfassung.....	165
V. Literaturangaben.....	167
VI. Anhang.....	177
Anhang A: Expertenbefragung.....	177
Anhang B: Fragebogenuntersuchung.....	184

---

# I. Einleitung

In einer Zeit knapper werdender Ressourcen und einer ansteigenden gesellschaftlichen und staatlichen Verschuldung verändern sich die Rahmenbedingungen der Kostenerstattung für alle Leistungen im Gesundheitsbereich und so auch für den Bereich der Psychotherapie. In der Bundesrepublik Deutschland sind ebenso wie in anderen Industrieländern die Forderungen nach Effektivitäts-Nachweisen von Behandlungsprogrammen und Interventionen, welche als Legitimation und Grundvoraussetzung für die Kostenerstattung von Seiten der Kostenträger gelten sollen, lauter geworden (z.B. Richter & Harst, 2006, Kendall et al., 2004, Burlingame et al., 1995)

Der Kampf um die knappen Ressourcen scheint seit einigen Jahren härter geworden zu sein und langfristig werden sich nur diejenigen Therapieprogramme halten können, denen es gelingt, ihren Erfolg zu belegen und ihre Überlegenheit hinsichtlich Wirksamkeit bzw. Kostenersparnis gegenüber anderen Programmen nachzuweisen. In diesem Zusammenhang sind Wirksamkeitsstudien und Kosten-Nutzen-Analysen im Rahmen der Psychotherapieforschung zunehmend in den Fokus der Aufmerksamkeit gerückt (z.B. Margraf, 2008) und ihre Ergebnisse haben an Bedeutsamkeit und gesundheitsökonomischer Relevanz, aber auch an Brisanz gewonnen.

Umso wichtiger erscheint es nun in Hinblick auf diese Entwicklung, dass Studien und Untersuchungen, auf deren Grundlage möglicherweise gesundheitspolitische Entscheidungen mit großer Tragweite getroffen werden, eine hohe methodische Qualität aufweisen und damit über eine zuverlässige Aussagekraft verfügen (vgl. auch Grawe, 2005).

Ziel der vorliegenden Arbeit ist es, ein Instrument zu entwickeln, das zur Verbesserung der Aussagekraft von psychotherapeutischen Wirksamkeitsstudien beitragen kann. Konkret wird die Konstruktion und Überprüfung eines Fragebogens dargestellt, mit dessen Hilfe allgemeine Wirkfaktoren in der Psychotherapie erfasst werden sollen. Dieser Fragebogen soll helfen, die Äquivalenz von Kontroll- und Therapiegruppen in Psychotherapiestudien in Hinblick auf die Ausprägung dieser allgemeinen Faktoren zu kontrollieren, um so die Aussagekraft dieser Studien hinsichtlich der Wirksamkeit der spezifischen Interventionen zu erhöhen.

Um die Konzeption und Entwicklung des im weiteren Verlauf dieser Arbeit darzustellenden Projektes plausibel darlegen zu können, erscheint zunächst eine kurze Einführung in das weite Themengebiet der Psychotherapieforschung angebracht, um eine theoretische Einbettung zu ermöglichen.

## II. Theorieteil

### 1. Psychotherapieforschung

#### 1.1. Grundüberlegungen

Die Aufgabe der Psychotherapieforschung besteht darin, den Wissensstand zu verschiedenen Aspekten der Psychotherapie zu erweitern und Bedingungen und Faktoren herauszufinden und zu überprüfen, die für den therapeutischen Prozess und das Therapieergebnis von Bedeutung sind. Der Fokus von Therapie-Studien kann auf ganz unterschiedliche Aspekte der Psychotherapie ausgerichtet sein, wie z.B. auf die Übereinstimmung zwischen Therapeut und Patient, auf die Aktivitätsrate des Therapeuten, auf Faktoren, die einen frühen Abbruch voraussagen können, u.v.a., aber ein Großteil der Studien setzt sich mit der Wirksamkeit von Psychotherapie auseinander und speziell um diese Studien soll es im Weiteren gehen.

Um bedeutsame Aussagen hinsichtlich der Wirksamkeit von Psychotherapien treffen zu können, müssen in den entsprechenden Untersuchungen allgemeine Grundprinzipien, die für Forschungsdesigns gelten, zur Anwendung kommen (Garfield, 1982). Generell gilt für das Design von experimentellen Studien, dass Rückschlüsse über den Einfluss einer einzelnen spezifischen Variable auf eine oder mehrere andere Kriteriums-Variablen nur dann zulässig sind, wenn alle anderen Variablen, die in der jeweiligen Situation eine Wirkung auf die abhängige Variable bzw. die abhängigen Variablen ausüben können, unter Kontrolle gehalten werden.

Für die Effektivitätsstudien in der Psychotherapieforschung bedeutet dies in erster Linie, dass in jede Untersuchung mindestens eine Kontrollgruppe als Vergleichsbedingung aufgenommen werden muss und dass die Gruppenzuweisung randomisiert erfolgen sollte. Erst dann können Unterschiede in der Wirksamkeit auf Differenzen in der Therapie, und nicht auf Stichprobenunterschiede oder andere äußere Faktoren zurückgeführt werden (Kendall et al., 2004). Diese Überlegungen wurden bei der Entwicklung von Kriterien bzw. Standards berücksichtigt, die bei der Begutachtung der mittlerweile umfangreichen Forschungsliteratur helfen sollen zu entscheiden, ob eine bestimmte Behandlung als „empirically supported“ angesehen werden kann: „Empirically supported treatments may be defined as treatments shown to be efficacious in randomized clini-



cal research trials with given populations“ (American Psychological Association Task Force on Promotion and Dissemination of Psychological Procedures, 1995, zit. nach Kendall et al., 2004). Um dem Anspruch einer kontrollierten Untersuchung Genüge zu leisten, werden verschiedene Arten von Kontrollgruppen als Vergleichsbedingung eingesetzt. Diese sollen im Folgenden kurz dargestellt werden.

### **1.2. Kontrollgruppen in der Psychotherapieforschung**

Die einfachste Form der Kontrollbedingung stellt eine *nicht-behandelte Kontrollgruppe* dar. Über den gesamten Zeitverlauf der in Frage stehenden Therapie hinweg werden dabei die Probanden der Kontrollgruppe den gleichen Untersuchungen unterzogen wie die Teilnehmer der Behandlungsbedingung. Bedeutsame Veränderungen, die sich im Vergleich zwischen den Teilnehmern der Therapiebedingung und denen der unbehandelten Kontrollgruppe zeigen, werden der Wirkung der Behandlung zugeschrieben. Mit Hilfe dieses Vorgehens können manche alternative Erklärungsmöglichkeiten für die beobachtbaren Veränderungen ausgeschlossen werden (wie z.B. Spontanremission; Reifung). Allerdings können bei Verwendung dieses Designs einige potentiell bedeutsame Faktoren nicht kontrolliert werden.

Dazu zählen z.B. die Antizipation der Behandlung, die Erwartung von Veränderung durch eine Therapie und der Kontakt mit einem Therapeuten (Kendall et al., 2004). Das größte Manko dieses Kontrollgruppendesigns stellt aber die ethische Fragwürdigkeit dar: um eine Vergleichbarkeit der Gruppen erreichen zu können, müssen die Probanden den Untersuchungsbedingungen randomisiert zugewiesen werden. Bei „echten“ Psychotherapiestudien würde dies bedeuten, dass etwa der Hälfte der behandlungsbedürftigen Probanden eine Therapie vorenthalten wird, die möglicherweise zu einer Linderung des psychischen Leides führen würde. Dass dieses Vorgehen bei Patienten mit voll ausgebildeten Störungen nicht zu vertreten ist, liegt auf der Hand.

Eine Erweiterung dieses Konzeptes stellt die *Wartelisten-Kontrollgruppe* dar. Auch in diesem Fall bleiben die Probanden über die Zeit der Untersuchung hinweg, d.h. während der Behandlungsdauer der Therapiegruppe, ohne Behandlung. Die Teilnehmer der Wartelisten-Kontrollgruppe wissen aber, dass sie nach Abschluss der Untersuchung einen Behandlungsplatz bekommen werden. Der Vorteil dieses Vorgehens gegenüber der unbehandelten Kontrollgruppe ist in zweifacher Hinsicht gegeben: zum einen wird das Problem der Nicht-Behandlung einer Gruppe von Probanden umgangen, zum anderen können zusätzliche Einflussvariablen kontrolliert werden. In

erster Linie geht es dabei um das Wissen der Probanden, dass sie eine Therapie erhalten werden und die damit verbundenen Veränderungserwartungen. Trotz der Vorzüge dieser Vorgehensweise gegenüber eines Designs mit einer unbehandelten Kontrollgruppe bleiben einige Schwierigkeiten. So ist es möglich, dass sich der Zustand von Probanden während der Wartezeit so sehr verschlechtert, dass ein unmittelbares Eingreifen erforderlich wird. Die Zumutbarkeit der Wartezeit hängt unmittelbar mit der Studiendauer zusammen. Bei Kurzinterventionen erscheint diese noch akzeptabel, bei einer längeren Therapiedauer wird das kritisch. Zumal dann auch die Wahrscheinlichkeit für einen Schwund der Probanden in der Kontrollgruppe steigt, was zu einer Veränderung der Gruppenzusammensetzung und damit zu einem Verlust der Vergleichbarkeit bzw. Repräsentativität führen kann.

Der Einsatz von Wartelistenkontrollgruppen erscheint daher als sinnvolles und wertvolles Vorgehen, um die generelle Wirksamkeit einer Behandlung zu belegen. Sobald dies aber erfolgt ist, sollte aufgrund der bestehenden Schwierigkeiten auf andere Methoden zurückgegriffen werden (vgl. Kendall et al., 2004). Vorteile gegenüber den beiden oben dargestellten Vorgehensweisen bieten zwei andere Kontrollgruppendesigns: die Verwendung einer *Standard-Behandlung* („treatment as usual“) als Vergleichsbedingung und der Einsatz einer *Placebo- bzw. unspezifischen Behandlungsbedingung*. In beiden Fällen haben die Teilnehmer der Kontrollgruppe parallel zu den Probanden der Therapiegruppe Kontakt zu Therapeuten und werden behandelt. Somit wird keinem Patienten eine Wartezeit zugemutet bzw. eine therapeutische Betreuung vorenthalten. Gleichzeitig wird die Vergleichbarkeit der beiden Gruppenbedingungen erhöht und das Risiko von drop-outs gegenüber einer Wartelisten-Kontrollgruppe reduziert. Im Fall der „treatment-as-usual“-Bedingung werden die Ergebnisse einer neuen Behandlungsform mit denen der augenblicklich als Standardvorgehen eingesetzten Therapie für die in Frage stehende Problematik verglichen. Bei dieser Vorgehensweise muss allerdings sichergestellt werden, dass in beiden Therapiegruppen, d.h. sowohl in der Behandlungsgruppe mit der zu überprüfenden Therapieform, als auch in der Kontrollgruppe, in der die Standardtherapie durchgeführt wird, diejenigen Wirkfaktoren und Bedingungen gleichermaßen ausgeprägt sind, die nicht für die eigentliche Therapie spezifisch sind (Kendall et al., 2004). Dann kann die relative Wirksamkeit der „neuen“ Therapie gegenüber dem bewährten Verfahren festgestellt werden.

In der Placebo- bzw. unspezifischen Behandlungsbedingung haben die Probanden regelmäßigen therapeutischen Kontakt und erfahren Zuwendung und Aufmerksamkeit. Ihnen wird ein Therapieerfolg vermittelt, das die Erwartung positiver Veränderungen wecken soll. Das Ziel einer solchen Kontrollgruppenbedingung besteht darin, ein Therapieangebot zur Verfügung zu stellen, das demjenigen der Behandlungsbedingung gleicht, mit Ausnahme der theoriebedingt als wirksam erachteten therapeutischen Interventionen. Unter der Voraussetzung, dass es gelingt, die unspezifischen Wirkfaktoren in der Kontrollgruppe gleichermaßen zu realisieren bzw. zu aktivieren wie in der Behandlungsbedingung, sollten Unterschiede im Therapieerfolg der beiden Gruppen auf die durchgeführten Interventionen zurückzuführen sein.

Sowohl das Placebo-Konzept, als auch das Konzept der unspezifischen Wirkfaktoren sind im Kontext der Psychotherapieforschung nicht unumstritten und bedürfen einer näheren Erläuterung. In den folgenden Kapiteln soll daher auf beide Konzepte näher eingegangen werden.

## 2. Das Placebo-Konzept

### 2.1. Das Placebo-Konzept in der Medizin

Der Begriff „Placebo“ stammt aus dem Lateinischen und bedeutet „ich werde angenehm sein“. Er geht zurück auf Vers 9 des 114. Psalms („Placebo domino in regione vivorum“- „Ich werde dem Herrn angenehm sein im Lande des Lebens.“), der als Kehr- und Rahmenvers seit dem frühen Mittelalter bei der Totenvesper gesungen wurde. Der Ausdruck „ein Placebo singen“ erhielt später die Bedeutung „einer hochgestellten Person schmeicheln“ und der Begriff „Placebo“ bezeichnete dann auch den Schmeichler selbst (Stuedel, 1956<sup>1</sup>). Nach Brown (1998, zit. nach Herbert & Gaudiano, 2005) und Walach (2003, zit. nach Wampold et al., 2005) wurden im 13. Jahrhundert „professionelle Trauernde“ engagiert, um die Totenvesper zu singen. Diese wurden dann als „Placebos“ betitelt, womit der Begriff erstmals in Verbindung mit der Bedeutung „anstelle von....“ gebraucht wurde.

Im 19. Jahrhundert wurde diese Bezeichnung in den Bereich der Medizin übernommen und bezeichnete dann zunächst Substanzen, die den Patienten verabreicht wurden, um ihnen den Anschein einer Behandlung zu vermitteln, obwohl die Mittel für unwirksam erachtet wurden. Mit

---

1 Dank an Prof. Dr. Gerhard Fichtner, Institut für Ethik und Geschichte der Medizin in Tübingen, für diesen Literaturhinweis.

der Einführung von Doppel-Blind-Studien im frühen 20. Jahrhundert bekam der Begriff noch eine zusätzliche Bedeutung. Seit dieser Zeit werden pharmakologisch unwirksame Substanzen Placebos genannt, die in Studien eingesetzt werden, um die „echten“ biochemischen Effekte eines Medikamentes von den psychologischen Mechanismen, d.h. in erster Linie den Wirkungserwartungen der Probanden, zu unterscheiden (Leber, 2000). Somit bezeichnet der Begriff Placebo zum einen eine Behandlung bzw. Behandlungskomponente, die „nur“ aufgrund der durch sie angestoßenen psychischen Prozesse Veränderung bewirkt, und zum anderen eine Kontrollbedingung in experimentellen Studien im Bereich der Medizin bzw. Pharmakologie (vgl. Herbert & Gaudiano, 2005).

Die Wirksamkeit medizinischer Placebo-Behandlungen gilt für den Bereich der Pharmakotherapie mittlerweile als gut belegt und sogar im Feld chirurgischer Interventionen konnten Erfolge von Placebo-Behandlungen nachgewiesen werden (z.B. Benedetti et al., 2003; Mayberg et al., 2002; zit. nach Herbert & Gaudiano, 2005), doch gelten diese Effekte gleichwohl als nebensächlich im Vergleich zur Wirkung der eigentlichen Behandlung (vgl. Wampold et al., 2005). Daher bleibt die Grundidee der klinischen Untersuchungen, dass die Placebo-Wirkung zu unterscheiden ist von derjenigen der eigentlich wirksamen Behandlung, sei es in Form eines biochemischen Prozesses oder eines chirurgischen Eingriffes.

Daher muss eine neue Therapie ihre Wirksamkeit dadurch belegen, dass die erzielten Effekte diejenigen der Placebo-Bedingung übertreffen. In doppel-blinden placebo-kontrollierten Untersuchungen wird der biologische, also spezifische Effekt aus der Differenz zwischen den Effekten der Behandlungs- und Placebobedingung errechnet. Es wird also angenommen, dass sich die Placebo-Effekte und die Medikamenteneffekte in ihrer Wirkung addieren (vgl. Kirsch, 2005). Die psychischen Prozesse, die der Placebo-Wirkung zugrunde liegen, sind für die medizinisch/pharmakologische Forschung über lange Jahre nur von sehr untergeordnetem Interesse gewesen. Mittlerweile zeichnet sich diesbezüglich allerdings eine Veränderung ab und mit dem steigenden Interesse an der Neurobiologie des Placeboprozesses rücken zunehmend auch die vermittelnden psychischen Prozesse in den Fokus des Forschungsinteresses (vgl. Oken, 2008; Enck et al., 2008).

## **2.2. Das Placebo-Konzept in der Psychotherapie**

In den letzten vierzig Jahren hielt das aus der Medizin stammende Placebo-Modell Einzug in die Psychotherapieforschung und es wurden zunehmend Placebo-Kontrollgruppen bei der Evaluation von Therapieverfahren eingesetzt. Dies lag in erster Linie an dem Bemühen, die interne Validität und damit die Aussagekraft der Untersuchungen zu erhöhen. Weiterhin bestand vermutlich auch der Wunsch, die „gold standards“ der medizinischen Forschung auf die Psychotherapie zu übertragen, um deren Status als legitime Behandlungsform zu stärken (vgl. Herbert & Gaudiano, 2005). Die Grundidee beim Einsatz psychotherapeutischer Placebo-Kontrollgruppen besteht darin, eine Vergleichsbedingung zu schaffen, mit deren Hilfe möglichst alle Faktoren kontrolliert werden, mit Ausnahme der spezifischen Techniken oder Prozesse, die theoriegemäß wirksam sind und zu positiven Therapieeffekten führen.

Die Übertragung dieses Konzepts von der Medizin in die Psychotherapie ist allerdings problematisch. Die Logik der Placebo-Bedingung in der Medizin ist recht einleuchtend: Placebos werden eingesetzt, um durch psychische Faktoren vermittelte Prozesse zu trennen von Vorgängen, die durch biochemische Faktoren angestoßen wurden bzw. Folgen invasiver Eingriffe darstellen. Somit werden zwei qualitativ verschiedene Klassen unabhängiger Variablen unterschieden.

In der Psychotherapie hingegen sind notwendigerweise alle auslösenden Faktoren bzw. alle vermittelnden Prozesse psychischer Natur, so dass mit der Trennung zwischen Placebo-Faktoren und spezifischen Therapiefaktoren innerhalb derselben Klasse, nämlich zwischen verschiedenen Formen psychischer Faktoren unterschieden werden muss (z.B. Herbert & Gaudiano, 2005; Kirsch, 2005). Damit stellt sich die schwierige Frage, bei welchen psychologischen Faktoren es sich um „echte“ Faktoren handelt und bei welchen um Placebos (Kirsch, 2005). Um diese konzeptionelle Schwierigkeit zu umgehen, wurden alternativ die Konzepte der unspezifischen (z.B. Lohr et al., 2003) bzw. allgemeinen Wirkfaktoren (z.B. Critelli & Neumann, 1984) eingeführt.

## **2.3. Placebo-Faktoren versus unspezifische/allgemeine Wirkfaktoren**

Die Begriffe „unspezifische“ bzw. „allgemeine“ Wirkfaktoren bezeichnen diejenigen Faktoren im Rahmen einer Psychotherapie, die nicht spezifisch sind für eine besondere Therapietechnik

oder Intervention, sondern in den meisten Therapieformen zur Anwendung kommen<sup>2</sup> (vgl. Herbert & Gaudiano, 2005). Eine Schwierigkeit dieses Konzeptes stellt die Trennung zwischen spezifischen und unspezifischen Faktoren dar: Wie charakteristisch für eine Therapieform muss ein Faktor sein, um ihn als „spezifisch“ zu bezeichnen?

Und umgekehrt: Ist ein Faktor bereits dann den allgemeinen Faktoren zuzurechnen, wenn er in mehr als einer Therapie Verwendung findet? Eine Möglichkeit, eine Grenze zu ziehen zwischen spezifischen und unspezifischen Faktoren, stellt der Bezug zur Theorie dar (vgl. Lohr et al., 2003). Demnach fallen all diejenigen Faktoren in die Kategorie der spezifischen Faktoren, die nach der theoretischen Konzeption der jeweiligen Therapie einen wichtigen, theoretisch begründeten Veränderungsmechanismus in Gang setzen. Die Schwierigkeit bei diesem Ansatz besteht allerdings darin, dass die Trennung zwischen spezifischen und unspezifischen Faktoren von der Therapieform bzw. -richtung abhängig ist, und somit unterschiedlich ausfällt. Demnach können Faktoren, die aus einer Therapioperspektive zu den unspezifischen Faktoren zählen, aus einer anderen Therapioperspektive zu den spezifischen gehören (vgl. Herbert & Gaudiano, 2005). Dies macht eine eindeutige Festlegung problematisch, was wiederum das Design einer „unspezifischen“ Kontrollgruppe erschwert, das über verschiedene Studien hinweg, der Vergleichbarkeit wegen, konstant gehalten werden könnte.

Aus diesen Erläuterungen wird deutlich, dass die Frage nach der inhaltlichen Festlegung der unspezifischen Faktoren komplex und schwierig ist, weshalb sie in den folgenden Abschnitten von verschiedenen Seiten beleuchtet und näher erläutert werden soll.

Generell bietet sich allerdings ein pragmatischer Ansatz für die Entscheidung an: Faktoren, die in allen Therapien eine Rolle spielen, egal ob sie von der Theorie her als spezifisch erachtet werden oder nicht (z.B. Rolle des Faktors „Empathie“ in der Gesprächspsychotherapie vs. in anderen Therapierichtungen) sollten als unspezifische Faktoren angesehen werden, denn wenn sie in allen Therapierichtungen eine Rolle spielen, stellen sie kein Entscheidungskriterium für mögliche Unterschiede dar.

---

2 Die Bezeichnungen „allgemeine“ und „unspezifische“ Wirkfaktoren sollen im Folgenden äquivalent verwendet werden, auch wenn diese nach Auffassung einiger Autoren zu unterscheiden sind (z.B. Castonguay, 1993). Lohr et al. (2003) sehen z.B. den Begriff „unspezifische Faktoren“ als einen Oberbegriff, dem der Term „allgemeine Faktoren“ als Teilmenge untergeordnet ist, doch erscheint diese Unterteilung nicht unbedingt notwendig und überzeugend. In der Literatur findet man meist eine Gleichstellung der beiden Begriffe und entsprechend ähnlich stellen sich auch die Definitionen der beiden Terme dar: „... nonspecific factors comprise those that are not unique to any given intervention, but that occur across a wide range of treatments. (Herbert & Gaudiano, 2005), S.897) vs. „Common factors are those dimensions of the treatment setting (therapist, client, therapy) that are not specific to any particular technique“ (Lambert. 2005).

## 2. Das Placebo-Konzept

---

Zu berücksichtigen ist in jedem Fall, dass das Konzept der unspezifischen bzw. allgemeinen Faktoren - trotz der genannten Umsetzungsprobleme - den Vorteil bietet, dass die darin zusammengefassten Faktoren nicht im Sinne des zugrunde liegenden Wirkmechanismus als unwirksam erachtet werden, wie es bei der Placebo-Konzeption im medizinischen Kontext der Fall ist. Das bedeutet, dass den unspezifischen bzw. allgemeinen Wirkfaktoren ein Anteil am Therapieerfolg zugesprochen wird. Der Beleg der spezifischen Wirksamkeit einer Therapie erfolgt somit über den Nachweis ihrer „inkrementellen Wirksamkeit“, die über die Effekte der allgemeinen Faktoren hinausgeht (vgl. Lohr et al., 2003).

Entscheidend ist damit in der Diskussion um die unspezifischen/allgemeinen Faktoren noch ein weiterer Gesichtspunkt: wenn diese Faktoren das Therapieergebnis positiv beeinflussen, dann müssen sie in jedem Fall überprüft werden, unabhängig davon, ob in einem Forschungsdesign der Wirksamkeitsnachweis einer Therapie durch den Vergleich mit einer unspezifisch behandelten Kontrollgruppe (z.B. in Form einer Unterstützung gebenden Therapie) oder einer Standardtherapie erfolgen soll. D.h. auch beim Vergleich zweier spezifischer Therapien erscheint es notwendig, das Vorhandensein bzw. das Ausmaß der unspezifischen Faktoren zu erfassen, um die Wirkung der spezifischen Faktoren miteinander vergleichen zu können. Ansonsten kann die Hypothese nicht ausgeschlossen werden, dass eine Ungleichheit zwischen den Therapieergebnissen auf das unterschiedliche Ausmaß bzw. die unterschiedliche Qualität der gemeinsamen Faktoren zurückzuführen ist (vgl. Borkovec & Sibrava, 2005; Jacobson, 1991). Eine Berücksichtigung der relevanten unspezifischen/allgemeinen Faktoren hat bisher allerdings auch in den so genannten RCTs (randomized controlled studies) nur ungenügend stattgefunden (Herbert & Gaudiano, 2005).

Es bleibt also festzuhalten, dass es prinzipiell sinnvoll und wichtig erscheint, unspezifische bzw. allgemeine Faktoren in Therapiestudien zu erfassen. Bevor man sich allerdings mit dem Problem auseinandersetzen kann, wie diese Variablen sinnvoll erhoben werden können, ist es unumgänglich, sich zunächst der schwierigen Frage zuzuwenden, welche Faktoren in eine solche Erfassung mit aufgenommen werden sollten.

### **3. Unspezifische/allgemeine Wirkfaktoren in der Psychotherapie**

Um herauszufinden, welche Faktoren in allen psychotherapeutischen Verfahren einen positiven Einfluss auf den Therapieerfolg haben und somit im Rahmen von Psychotherapiestudien als unspezifische bzw. allgemeine Faktoren erfasst werden sollten, bieten sich zwei mögliche Herangehensweisen: eine theoretische und eine empirische. In den folgenden Abschnitten sollen beide berücksichtigt werden.

#### **3.1. Theoretische Konzeptionen allgemeiner Wirkfaktoren**

Es gibt eine Vielzahl an theoretischen Konzeptionen, die sich im Rahmen der Outcome- sowie der Prozessforschung mit dem Thema „allgemeine Wirkfaktoren“ befasst haben, weshalb eine vollständige Auflistung den Rahmen dieser Arbeit sprengen würde. Daher sollen im Folgenden lediglich einige theoretische Konzepte vorgestellt werden, die nachfolgende Ansätze und Forschungsarbeiten besonders beeinflusst haben.

##### **3.1.1. Das Modell von Frank (1961)**

Als einer der ersten Forscher untersuchte Frank (Frank, 1961, zit. nach Grawe, 1998; Frank & Frank, 1991, zit. nach Wampold et al., 2005) den Einfluss unspezifischer Faktoren auf den therapeutischen Prozess. Er geht davon aus, dass Therapie in erster Linie durch Induktion von Hoffnung wirkt, d.h. durch die Weckung einer positiven Erwartung auf Besserung. Die Patienten kämen in einem Zustand der Hoffnungslosigkeit bzw. „Demoralisierung“ in die Therapie und dieser Zustand werde durch die Therapie verändert. Demnach sei der generelle Wirkfaktor die Veränderungserwartung. Laut Frank gibt es vier zentrale Faktoren, die dabei unabhängig von der therapeutischen Schule eine Rolle spielen:

- Die *Therapiebeziehung* zwischen Hilfeempfänger und sanktioniertem Hilfegeber
- Das *Bestehen eines formalisierten Behandlungsangebotes* im institutionellen Rahmen
- Die *Vermittlung eines Behandlungsrational*s
- Die Durchführung eines mit dem Rational konsistenten Behandlungsvorgehens („*therapeutische Rituale*“)

Nach der Theorie von Frank sei der „Inhalt“ der therapeutischen Rituale von untergeordneter Bedeutung für den Heilungsprozess, solange die „Interventionen“ glaubwürdig und plausibel in Be-



zug auf das Behandlungsrational seien. Er schließt dabei die Wirksamkeit spezifischer Interventionen nicht aus, betont aber die grundlegende Notwendigkeit der genannten allgemeinen Faktoren, damit die spezifischen Faktoren ihre Wirkung entfalten könnten (Frank, 1989).

#### **3.1.2. Das Modell von Garfield (1982)**

Garfield greift in seinem Buch „Psychotherapie – Ein eklektischer Ansatz“ das Konzept von Frank auf und stellt die Hypothese auf, „... dass trotz vieler offensichtlicher Differenzen hinsichtlich theoretischer Orientierungen und Verfahren viele voneinander unterscheidbare Therapieschulen auf im Grunde gemeinsamen Faktoren aufbauen, die sich zur Erklärung ziemlich vieler Änderungen heranziehen lassen, von denen man glaubte, sie würden nur aufgrund ihrer jeweiligen schulenspezifischen Vorgehensweisen eintreten“ (S.134).

Folgende allgemeine Faktoren werden von Garfield als therapeutische Wirkmechanismen angeführt:

- Das *Hilfe-Aufsuchen bei einem gesellschaftlich legitimierten Helfer*, was zu Entwicklung von Zuversicht beim Klienten führt
- Das Erleben von Interesse und verständnisvoller *Unterstützung durch den Therapeuten*
- Das Bestehen einer *Therapeut-Klient-Beziehung*, die durch gegenseitige Achtung und Wertschätzung bestimmt ist
- Die *Erzeugung von Einsicht und Verständnis* auf Seiten des Klienten durch die Einordnung der eigenen Erfahrungen in ein „Glaubenssystem“
- Das *Erleben von Katharsis und Release-Phänomenen* (Entlastung durch die Freisetzung von Emotionen)
- Die *Anregung zu neuen Verhaltensweisen* durch den Therapeuten
- Die *Veränderung der Selbstwahrnehmung*

Im Vergleich zu Frank führt Garfield in seiner Schrift die zur Veränderung führenden „therapeutischen Rituale“ näher aus und benennt einzelne Prozesse, die zur Veränderung führen (z.B. Katharsis). Zudem integriert er in seinen Ansatz auch Aspekte, die eher dem Therapieoutcome zuzuordnen sind als den Wirkfaktoren, wie z.B. die Veränderung der Selbstwahrnehmung.

#### **3.1.3. Das Modell von Orlinsky & Howard (1987)**

In ihrem „generic model of psychotherapy“ stellen Orlinsky und Howard, basierend auf einem Review der Forschungsliteratur, fünf Variablen zusammen, die ihrer Ansicht nach in allen Therapieformen wichtige Bestandteile des therapeutischen Prozesses darstellen:

---

### 3. Unspezifische/allgemeine Wirkfaktoren in der Psychotherapie

---

- Das *Bestehen eines Abkommens* bzgl. der Therapie bzw. das Vorhandensein eines therapeutischen Vertrages, der die Bedingungen und Ziele der Therapie beinhaltet
- Die *Anwendung therapeutischer Interventionen* (Diagnostik bzw. Einschätzung der Probleme des Patienten und Einsatz von Techniken zu deren Behebung)
- Die *Entwicklung eines therapeutischen Bündnisses*, das sich durch Rollen- Engagement, empathischer Resonanz und beidseitiger Zustimmung auszeichnet
- Der *Status der Selbstbezogenheit* (z.B. „psychologische Offenheit“ vs. „Verteidigungshaltung“), sowohl auf Patienten-, als auch auf Therapeutenseite, der den Grad der Einflussmöglichkeit durch therapeutische Interventionen bestimmt
- Das *Erreichen von unmittelbaren Zielen* innerhalb der Therapiesitzung

Ähnlich wie die oben genannten Ansätzen betonen Howard & Orlinsky die Bedeutung der Rahmenbedingungen und des therapeutischen Bündnisses. Zudem wird den therapeutischen Interventionen eine wichtige Rolle zuerkannt. Ähnlich wie Garfield (1982) führen die Autoren zudem das Erreichen von Teilzielen als bedeutsame Variable an. Im Unterschied zu Frank (1961) und Garfield berücksichtigen Orlinsky und Howard in ihrem Modell den Faktor „Selbstbezogenheit“ (self-relatedness), der ihrer Auffassung nach verschiedene Aspekte des Selbst-Managements widerspiegeln (z.B. Selbstaufmerksamkeit, Selbstregulation, Selbstwert). Das Ausmaß der Selbstbezogenheit bestimmt, inwieweit der Patient Informationen aufnehmen könne, die in diesem Moment nicht mit den eigenen Schemata in Einklang stünden. Idealerweise bestehe ein Zustand, den Howard und Orlinsky als „Offenheit gegenüber Erfahrung“ (openness to experience) bezeichneten.

#### **3.1.4. Das Modell von Grawe et al. (1994; Grawe, 2000, 2005)**

In ihrer viel zitierten Studie zur Wirksamkeit von Psychotherapie postulieren Grawe et al. (1994) als Fazit ihrer Untersuchung, dass es verschiedene allgemeine Wirkprinzipien gebe, die in allen Psychotherapien die positive Veränderung auf Seiten des Patienten bewirkten. Zunächst wurden von den Autoren drei Wirkprinzipien beschrieben, die in späteren Arbeiten von Grawe (z.B. 2005) zu folgenden fünf Wirkfaktoren ausdifferenziert wurden :

- *Ressourcenaktivierung*: die Einbeziehung individueller Fähigkeiten, Eigenschaften und Interessen des Patienten tragen zum Therapieerfolg bei, da der Patient in der Therapie

auch seine positiven Seiten und Stärken erfahren könne. Unter diesen Aspekt subsumieren Grawe et al. auch Förderung der Erwartungen beim Patienten hinsichtlich der Wirksamkeit der Therapie und der motivationalen Bereitschaft, sowie die Erstellung eines plausiblen Erklärungsmodells.

- *Therapiebeziehung*: Die therapeutische Beziehung wird von den Autoren ebenfalls als ein Faktor angesehen, der einen signifikanten Einfluss auf das Therapieergebnis hat.
- *Problemaktualisierung*: die Probleme, die Anlass für die Therapie sind, sollten in der Therapie direkt erfahrbar gemacht werden (z.B. durch Exposition).
- *Motivationale Klärung*: im Verlauf der Therapie sollte der Patient ein Verständnis für die prädisponierenden, auslösenden und aufrecht erhaltenden Faktoren seines problematischen Erlebens und Verhaltens entwickeln.
- *Problembewältigung*: der Patient soll vom Therapeuten bei der Bewältigung bisher nicht lösbarer Probleme unterstützt und angeleitet werden, so dass neue Kompetenzerfahrungen möglich werden.

Stärker noch als die Ansätze von Garfield und Orlinsky & Howard bildet Grawe in seinem Modell die Prozesse ab, die seinen Ergebnissen zufolge einen Therapieerfolg bewirken. Diese Konzeption einer „allgemeinen Psychotherapie“ hat die Therapieforschung sehr beeinflusst und den Bereich der „Prozessforschung“ ganz wesentlich mitbestimmt.

#### **3.1.5. Vergleich der Modelle**

Vergleicht man die vier dargestellten Modelle allgemeiner Wirkfaktoren, werden Gemeinsamkeiten und Unterschiede deutlich. Gemeinsam ist allen Konzeptionen, dass sie die Schaffung positiver Therapieerwartungen beim Patienten als einen zentralen Wirkmechanismus beinhalten. Ein wesentlicher Bestandteil für die Schaffung positiver Erwartungen stellen dabei die institutionellen Rahmenbedingungen („Aufsuchen eines legitimierte Helfers“; Therapievertrag) dar. Die Bedeutung eines Behandlungsrationalen für den Aufbau von Therapieerwartungen wird ebenfalls in fast allen Modellen berücksichtigt (Behandlungsrational, „Glaubenssystem“, plausibles Erklärungsmodell). Eine weitere Übereinstimmung zwischen den Ansätzen stellt die Betonung der therapeutischen Beziehung als allgemeines Wirkprinzip dar, auch wenn diese, je nach Autor, durch unterschiedliche Merkmale bestimmt zu werden scheint (z.B. Rollen-Engagement, empathische Resonanz und wechselseitige Zustimmung bei Orlinsky & Howard (1987) vs. Achtung und Wertschätzung bei Garfield (1982), der die „verständnisvolle Unterstützung“ als separaten

### 3. Unspezifische/allgemeine Wirkfaktoren in der Psychotherapie

---

Punkt anführt). Unterschiede zwischen den Modellen zeichnen sich v.a. in Hinblick auf die Berücksichtigung von Interventionen und therapeutischen Teilzielen ab. Während Frank (1961) noch sehr vage von „therapeutischen Ritualen“ spricht, die in erster Linie konsistent mit dem Behandlungsrational sein sollen, und keine Teilziele benennt, führt Garfield (1982) bestimmte Wirkprozesse an, die für den Therapieerfolg von Bedeutung sein sollen: er benennt hier „Katharsis und Release-Phänomene“ und die Anregung zu neuen Verhaltensweisen. Zudem listet er die „Veränderung der Selbstwahrnehmung“ auf, die als Teilziel bzw. „Stundenoutcome“ anzusehen ist. Orlinsky & Howard (1987) hingegen führen hier „therapeutische Interventionen“ im Sinne spezifischer Techniken und „Erreichen von unmittelbaren Zielen innerhalb der Therapiesitzung“ nur als Oberbegriffe an.

Das Modell von Orlinsky & Howard weist noch eine Besonderheit auf: es ist das einzige Modell, in dem mit dem Faktor „Status der Selbstbezogenheit“ explizit eine „Patientenvariable“ aufgeführt wird. Grawe (1994; 2000) fokussiert zum einen auf die Prozesse, die in den Therapiesitzungen angestoßen werden und ihrerseits zu Veränderung führen (Ressourcenaktivierung) und zum anderen auf Ergebnisse des therapeutischen Arbeitens (motivationale Klärung, Problembewältigung).

Prüft man die aufgeführten Konzepte hinsichtlich ihrer Nützlichkeit für die Abgrenzung zwischen spezifischen und unspezifischen/allgemeinen Faktoren, so muss man feststellen, dass es sich bei den Modellen – mit Ausnahme vielleicht des Modells von Frank – zumeist um übergreifende Konzeptionen handelt, die man vielleicht als „Metamodelle“ bezeichnen könnte. Insbesondere trifft dies auf den Ansatz von Grawe zu. Diese Modelle beinhalten zum einen Oberbegriffe, unter die sich verschiedene spezifische und unspezifische Interventionen subsumieren lassen, und zum anderen werden Prozesse aufgeführt, die durch diese Interventionen in Gang gesetzt werden (z.B. Katharsis-Phänomene; Problemaktualisierung). Das „Allgemeine“ der einzelnen Elemente der Konzeptionen besteht eher in den Prozessen (z.B. Klärung), die durch sie angestoßen werden, bzw. in deren Auswirkungen (z.B. Veränderung der Selbstwahrnehmung). Es handelt sich dabei weniger um „allgemeine Faktoren“, die spezifischen, also therapieschulen-typischen Interventionen gegenüber gestellt werden könnten.

Diese Herangehensweise und die daraus resultierenden Modelle sind von großem Wert für Forschung und Praxis, da sie übergeordnete Konzeptionen darstellen und v.a. auf den Ebenen von Prozess und Outcome Vergleiche zwischen verschiedenen Therapieformen zulassen. Für die Abgrenzung zwischen spezifischen und unspezifischen Variablen im Sinne des oben skizzierten Ansatzes eignen sie sich allerdings nur begrenzt. Hier erscheint es erforderlich, einen „additiven Ansatz“ zu verwenden, in dem unspezifische und spezifische Faktoren differenziert und parallel erfasst werden können.

Dies scheint mit dem „dualen Basismodell der Psychotherapie“ von Schulte (1996) besser möglich zu sein, weshalb dieses Modell im nächsten Abschnitt kurz vorgestellt werden soll.

#### **3.1.6. Das Modell von Schulte (1996)**

Dem „dualen Basismodell“ des therapeutischen Prozesses entsprechend gibt es zwei zentrale Aufgaben für den Therapeuten: zum einen die Förderung des so genannten „Basisverhaltens“ des Patienten, was nach Schulte die Grundlage für die „eigentliche“ Therapie darstellt, und zum anderen die Durchführung spezieller therapeutischer Methoden zur Modifikation der Probleme oder Verhaltensauffälligkeiten des Patienten. Beide Aufgaben seien „für die Durchführung einer erfolgreichen Therapie notwendig und insofern gleichberechtigt.“ (Schulte, 1996, S.64). Bestandteile dieses Modells seien drei psychotherapeutische Grundelemente, nämlich die therapeutischen Techniken, die therapeutische Beziehung, sowie Klienteneigenschaften, die sich als Motivation spezifizieren ließen. Unter den therapeutischen Techniken versteht Schulte die Anwendung spezieller Therapiemethoden, die Veränderung ermöglichen und Lernprozesse initiieren, wie z.B. die kognitive Umstrukturierung oder Veränderung emotionaler Schemata.

Um diese Techniken anwenden zu können, müssten allerdings gewisse Rahmenbedingungen gegeben sein, nämlich das Basisverhalten des Patienten. Dieses könne vom Therapeuten durch die Gestaltung der therapeutischen Beziehung und die Förderung der Motivation (d.h. in erster Linie der Erwartung) des Patienten beeinflusst werden. Diese therapeutischen Aufgaben zur Verbesserung des Basisverhaltens seien unspezifisch, da sie für alle Therapierichtungen gelten. Um die Voraussetzung für die Anwendung spezieller Therapiemethoden zu schaffen, müsse der Therapeut auf die Motivation des Patienten und damit auf dessen Basisverhalten Einfluss nehmen. Dies gelinge über die Gestaltung der therapeutischen Beziehung, die von Therapeutenseite acht Merkmale aufweisen müsse:

- *Expertenstatus*: der Therapeut soll für den Patienten ein kompetenter Experte sein.

### 3. Unspezifische/allgemeine Wirkfaktoren in der Psychotherapie

---

- *Unbedingte Wertschätzung*: der Therapeut soll dem Patienten gegenüber eine positive Haltung signalisieren.
- *Empathie*: der Therapeut soll sich verständnisvoll verhalten
- *Vertrauen*: der Therapeut soll sich dem Patienten gegenüber als vertrauenswürdig erweisen.
- *Autonomie*: der Therapeut soll dem Patient in der Therapie Freiheiten lassen.
- *Konsens*: der Therapeut soll bzgl. der Ziele und Vorgehensweisen in der Therapie mit dem Patienten einer Meinung sein bzw. zumindest das Gespräch darüber suchen.
- *Professionalität*: der Therapeut soll die therapeutische Distanz wahren.
- *Akzeptanz*: der Therapeut akzeptiert den Patienten in seinen interaktionellen Besonderheiten.

Nach dem Modell von Schulte kann man also zwischen spezifischen und unspezifischen Variablen unterscheiden. Durch die unspezifischen Faktoren wird die Basis für die Anwendung spezifischer Interventionen geschaffen, wobei nach Schulte in manchen Fällen auch das Vorhandensein unspezifischer Faktoren für die Verbesserung der Symptomatik ausreicht, was er mit der „Placebowirkung“ von Kontrollgruppen in Verbindung bringt (Schulte, 1996, S. 66). Als unspezifische Variablen benennt Schulte auf Seiten des Therapeuten die oben angeführten acht Merkmale, die eine gute therapeutische Beziehung ausmachen, und auf Seiten des Patienten die „Therapiemotivation“, die sich in erster Linie als positive Therapieerwartung manifestiert. Damit in direktem Zusammenhang sieht Schulte bestimmte Verhaltensweisen des Patienten, die das Basisverhalten ausmachen: neben der Therapienachfrage (= der Patient kommt in Therapie und bleibt) sind das die Mitarbeit des Patienten (= aktive Mitwirkung), seine Selbstöffnung (= Bereitschaft des Patienten über sehr persönliche Aspekte seines Lebens und seines Verhaltens zu berichten, auch wenn dies mit negativen Gefühlen verbunden ist) und die Bereitschaft zur Erprobung neuer Verhaltensweisen.

#### **3.1.7. Der Ansatz von Wampold (2005)**

Ähnlich wie Schulte geht Wampold in seiner Konzeption von einem additiven Modell der Wirkfaktoren aus: „However, there are convincing reasons to believe that specific treatment effects and placebo effects are additive, at least in well-designed studies.“ (S.840). Seiner Ansicht nach ist der Begriff „Placebo“ ein Symbol für den heilenden psychotherapeutischen Kontext und um-

fasst alle Aspekte, die für den Patienten bedeutsam sind. Folgende Faktoren sieht Wampold als „essential but not unique“ an:

- Die *Beziehung* zu einem ausgebildeten Therapeuten.
- Die *Unterstützung* des Therapeuten.
- Das *empathische Eingehen* des Therapeuten auf den Patienten.
- Die *Erwartung*, dass die Therapie wirksam sein wird.

Unterstützung für dieses „duale“ oder additive Konzept der Wirksamkeit von Psychotherapie liefert die Argumentation von Lambert (2005) in Bezug auf frühe Therapieerfolge. Im Unterschied zur Pharmakotherapie seien in der Psychotherapie zeitlich gesehen frühe Verbesserungen in der Therapie Hinweise auf einen langfristig positiven Therapieerfolg. Nach Lambert signalisiere eine frühzeitige Reaktion des Patienten auf die Therapie seine Veränderungsbereitschaft und eine Reaktion auf allgemeine Faktoren, nicht auf spezifische Interventionen, da diese zu Therapiebeginn noch gar nicht zum Einsatz kämen.

#### **3.1.8. Fazit**

Die oben aufgeführten Modelle wurden dargestellt, um auf theoretischer Basis eine Auswahl der relevanten Faktoren treffen zu können, die im Rahmen von Psychotherapiestudien als allgemeine Wirkfaktoren erfasst werden sollten.

Da das Ziel darin besteht, zwischen allgemeinen/unspezifischen und spezifischen Faktoren unterscheiden zu können, um deren jeweilige Auswirkung auf den Therapieerfolg überprüfen zu können, sollen nur solche Variablen ausgewählt werden, die keine Oberbegriffe für Interventionsstrategien darstellen.

Ebenso wenig erscheint es sinnvoll, formulierte Teilziele als Variablen zu erfassen, da zwischen Einflussvariable und Ergebnis unterschieden werden soll (unabhängige bzw. Covariable vs. abhängige Variable). Als letztes sollen diejenigen Variablen ausgeschlossen werden, die strukturelle Merkmale beschreiben (z.B. das formalisierte Behandlungsangebot; die „Legitimation“ des Heilers; der Schritt des Hilfe-Aufsuchens). Diese Faktoren sind zweifelsfrei von großer Bedeutung, doch sollten sie im Rahmen von Psychotherapie-Studien durch die Schaffung von Rahmenbedingungen kontrollierbar sein (z.B. Therapeuten mit gleichwertigem Ausbildungsstand in Therapie- und Kontrollgruppen; gleiche Räumlichkeiten; gleiche Akquise der Teilnehmer etc.) und somit keiner Überprüfung mittels eines Messinstrumentes bedürfen.

### 3. Unspezifische/allgemeine Wirkfaktoren in der Psychotherapie

---

Unter Berücksichtigung der oben genannten Ausschlusskriterien erscheinen für die Erfassung unspezifischer Faktoren zwei Faktoren von besonderer Bedeutung:

- Die (positive) Erwartung des Patienten hinsichtlich des Therapieverlaufes (Veränderungserwartung), sowie
- die therapeutische Beziehung, wobei deren „Bestandteile“ noch einer näheren Betrachtung unterzogen werden müssen.

In Kombination mit diesen beiden Hauptvariablen sind noch weitere Variablen von Interesse.

Dies sind zum einen auf Therapeutenseite:

- Die Glaubwürdigkeit des – wie auch immer gearteten – Behandlungsrationalen und
- der vom Patienten wahrgenommene „Expertenstatus“ des Therapeuten bzw. dessen professionelles Auftreten.

Und zum anderen auf Patientenseite:

- Die Selbstöffnung des Patienten<sup>3</sup>, sowie
- die Bereitschaft zur Mitarbeit (aktive Mitarbeit, Bereitschaft zur Erprobung neuer Verhaltensweisen i.S. Schultes).

## 3.2. Das Konzept der therapeutischen Beziehung

Wie im vorangegangenen Abschnitt bereits dargestellt, wird dem Konstrukt der therapeutischen Beziehung in nahezu allen Arbeiten, die sich mit dem Wirkmechanismus von Psychotherapie auseinandersetzen, eine große Bedeutung zugemessen. So bemerkte beispielsweise Horvath (1994) in der Einführung zu seinem Sammelband über die „working alliance“: „... It appears that the alliance is currently the best model of the in-therapy pantheoretical process variable“ (Horvath, 1994, S.2). Tatsächlich konnten Forscher unterschiedlicher theoretischer Orientierungen, die mit verschiedensten Messinstrumenten die Wirkung der therapeutischen Beziehung untersucht haben, durchgehend einen Zusammenhang zwischen der Beziehung zwischen Therapeut und Patient und dem Therapieerfolg feststellen (Horvath & Symonds, 1991, zit. nach Martin et al., 2000).

Nach Gaston (1990) hat die spezielle Beziehung zwischen Therapeut und Patient drei Funktionen:

---

3 Die Variable „Selbstöffnung“ im Sinne Schultes (1996) scheint in engem Zusammenhang zu stehen mit dem Faktor „self-relatedness“ bei Orlinsky & Howard (1987), da sie den für die Therapie geeigneten Zustand der „self-relatedness“ als Zustand der „psychologischen Offenheit“ bezeichnen.



### 3. Unspezifische/allgemeine Wirkfaktoren in der Psychotherapie

---

- Die therapeutische Beziehung hat selbst eine therapeutische Wirkung.
- Die therapeutische Beziehung stellt eine essentielle Grundbedingung für die Wirksamkeit therapeutischer Interventionen dar.
- Die therapeutische Beziehung interagiert mit verschiedenen Formen therapeutischer Interventionen.

Wenig Übereinstimmung findet man in der Literatur allerdings bezüglich der postulierten Konstituenten und auch bezüglich der Bezeichnungen für das Konstrukt. Verwendet werden die Begriffe „therapeutische Beziehung“, „therapeutisches Bündnis“, „therapeutische Allianz“ bzw. „therapeutic alliance“, „helping alliance“ und „working alliance“. Daher soll in diesem Abschnitt zunächst eine kurze Darstellung über die Entwicklung des Konzepts und ein Überblick erfolgen, bevor die Entscheidung bezüglich der im Rahmen dieses Ansatzes zu erfassenden Komponenten erläutert wird.

Es lassen sich prinzipiell drei Gruppen von Variablen unterscheiden, die in Zusammenhang mit der therapeutischen Beziehung bzw. der therapeutischen Allianz gebracht werden, nämlich Patientenvariablen, Therapeutenvariablen und Interaktionsvariablen, d.h. Variablen, die in irgendeiner Form die Beziehung und Interaktion zwischen Therapeut und Patient beschreiben. Unterschiede zwischen den verschiedenen Konzeptionen der therapeutischen Beziehung ergeben sich zum einen – wie oben angedeutet – hinsichtlich der Bezeichnungen und zum anderen hinsichtlich des Umfangs und der Art der integrierten Variablen.

Nach Sachse (2006) stellen die therapeutische Beziehung und die therapeutische Allianz zwei Grundkonzepte dar, die es zu differenzieren gilt: der Begriff der „therapeutischen Beziehung“ beschränkt sich seiner Auffassung nach auf diejenigen Beziehungsaspekte, die vom Therapeuten gestaltet werden (z.B. Empathie, Akzeptanz, Wertschätzung etc.), also Therapeutenvariablen. Der Begriff „therapeutische Allianz“ hingegen umfasst nach Sachse auch Aspekte, die sich auf beide Interaktionspartner beziehen, also nicht nur Therapeuten-, sondern auch Patienten- (z.B. Aufnahmebereitschaft des Patienten) und Interaktionsvariablen. Andere Forscher treffen diese Unterscheidung allerdings nicht, sondern verwenden die Begriffe „therapeutische Beziehung“ und

### 3. Unspezifische/allgemeine Wirkfaktoren in der Psychotherapie

---

„therapeutische Allianz“ äquivalent (vgl. z.B. Henry & Strupp, 1994; Hank & Krampen, 2006; Corbiere et al., 2006).

Möglicherweise spiegeln die verschiedenen Bezeichnungen für die besondere Beziehung zwischen Therapeut und Patient weniger inhaltliche Unterschiede, denn Versuche der Abgrenzung zwischen verschiedenen Forschergruppen wider, die relativ zeitgleich diese Thematik untersucht haben. Aus dieser Perspektive scheint es zulässig, die verschiedenen Termini als äquivalent zu betrachten und stellvertretend zu verwenden, wie es im weiteren Text auch gehandhabt werden soll.

Zum ersten Mal wurde das Konzept der therapeutischen Beziehung nach Hentschel (2005) unter der Bezeichnung „therapeutische Allianz“ in den Schriften zur psychoanalytischen Psychotherapie erwähnt. Schon von Freud (1912; S. 371, zit. nach Hentschel, 2005) wurde eine positive Form der Übertragung auf den Arzt in Abgrenzung zu unbewussten Übertragungsphänomenen (Widerstand bzw. verdrängte erotische Regungen) beschrieben: „Die andere bewussteinfähige und unanstößige Komponente bleibt bestehen und ist in der Psychoanalyse genau ebenso Trägerin des Erfolges wie bei anderen Behandlungsmethoden.“

Zu Beginn handelte es sich beim Konzept der therapeutischen Allianz also um ein psychodynamisches Paradigma, das zunächst in diesem Rahmen ausgebaut und überarbeitet wurde (Horvath, 1994). Später wurde dieses Konzept auf interpersonale und klient-zentrierte Ansätze ausgeweitet. In einem weiteren Schritt erfolgte dann in den 70er Jahren mit der Entwicklung von Verfahren, die auf die Erfassung allgemeiner therapeutischer Wirkfaktoren ausgerichtet waren, die Loslösung der Operationalisierungen des Konzeptes aus der Schulengebundenheit (Hank & Krampen, 2006).

Eine therapieschulenunabhängige Definition der therapeutischen Allianz (in diesem Fall: „therapeutic working alliance“) lieferte u.a. Bordin (1979), der folgende drei Aspekte als notwendige Bestimmungsstücke ansieht: die Übereinstimmung zwischen Therapeut und Patient bezüglich der Ziele, die Übereinstimmung bezüglich der Aufgaben der Therapie und die Entwicklung einer therapeutischen Bindung (Wir-Gefühl). Unter Aufgaben versteht Bordin dabei die spezifischen

### 3. Unspezifische/allgemeine Wirkfaktoren in der Psychotherapie

---

Aktivitäten von Patient und Therapeut, die zur Erreichung der zu benennenden Therapieziele führen, d.h. die Durchführung der spezifischen Interventionen. Die Konzeptualisierungen dieser Aufgaben unterscheiden sich zwischen den verschiedenen Therapierichtungen. Entscheidend für die therapeutische Beziehung ist nach Bordin, dass sich Therapeut und Patient einig sind darüber, welche Ziele erreicht werden sollen und auf welche Art und Weise. Die Bindung entsteht aus dem Erleben von Gemeinschaft in Bezug auf die zusammen zu bewältigenden Aufgaben und spiegelt sich wider in der Erfahrung des Sich-Mögens, Sich-Vertrauens, wechselseitigen Respektierens und des gemeinsamen Engagements (Bordin, 1994).

Der Ansatz von Bordin führte zu einer Reihe von Forschungsarbeiten und beeinflusste die Entwicklung zahlreicher Fragebögen und Inventare zur Erfassung der therapeutischen Allianz (z.B. WAI: Working Alliance Inventory, Horvath & Greenberg, 1986; CALPAS: California Psychotherapy Alliance Scale; Gaston, 1991). Andere Instrumente postulieren mehr Dimensionen als die von Bordin genannten, wie z.B. die ARM (Agnew Relationship Measure, Agnew-Davies et al., 1998) mit den Subskalen Bindung, Partnerschaft, Vertrauen, Offenheit und Patienteninitiative. Gemeinsam ist den meisten Ansätzen, dass sie von einer Mehrdimensionalität der therapeutischen Allianz ausgehen (so auch Hentschel, 2005; S.388: „Therapeutische Allianz ist eindeutig ein mehrdimensionales Konzept.“). Obwohl die verschiedenen Allianz-Skalen in unterschiedlichen Forschergruppen entwickelt wurden und verschiedene Subskalen enthalten, ergaben sich allerdings in vergleichenden Untersuchungen hohe Interkorrelationen zwischen den Skalen (z.B. Hatcher & Barends, 1996), was für eine hohe inhaltliche Überschneidung spricht.

Nach Luborsky (1994), der ähnlich wie Bordin auf jahrzentelange Forschungsarbeit zum Thema therapeutische Allianz zurückblicken kann, sind zwei Aspekte einer hilfreichen Beziehung zu berücksichtigen: zum einen die Erfahrung des Patienten, dass der Therapeut sich hilfreich und unterstützend verhält, und zum anderen das Gefühl, gemeinsam gegen das anzukämpfen, was den Patienten behindert. Ein Aspekt, den auch Horvath (1994) als „Herzstück“ der Allianz ansieht: „The concept focuses on the importance of the client and therapist forming a partnership against the common foe of the client's debilitating pain“ (Horvath, 1994; S.1). In ihrem Review der ver-

### 3. Unspezifische/allgemeine Wirkfaktoren in der Psychotherapie

---

schiedenen Allianz-Skalen kommen Horvath & Luborsky (1993) zu dem Schluss, dass es zwei Hauptaspekte der therapeutischen Allianz gibt, die sich in den meisten Konzeptionen wiederfinden: zum einen eine affektive Bindung zwischen Therapeut und Patient, und zum anderen die Zusammenarbeit bzw. die Bereitschaft gemeinsam in den therapeutischen Prozess zu investieren.

Obwohl sich die verschiedenen Ansätze zur Beschreibung und Erfassung der therapeutischen Allianz stark überschneiden, gibt es bislang keinen allgemein gültigen Konsens darüber, welche Bestimmungsstücke nun die therapeutische Beziehung im Kern ausmachen. Um diese Schwierigkeit zu verdeutlichen sind im Folgenden einige der wichtigsten Konzeptionen aufgelistet. Die jeweiligen Bestimmungsstücke wurden - soweit möglich - unterschieden nach Therapeuten-, Patienten- und Interaktionsvariablen. Zudem sind zugehörige Messinstrumente aufgeführt<sup>4</sup>.

Tabelle 1

#### Überblick: Komponenten der therapeutischen Beziehung/therapeutischen Allianz

	<b>Therapeutenvariablen</b>	<b>Patientenvariablen</b>	<b>Interaktionsvariablen</b>	<b>Messinstrumente</b>
<i>Schulte (1996)</i> <sup>5</sup>	- Expertenstatus - Wertschätzung - Empathie - Gewährung von Autonomie - Professionalität - Akzeptanz des Pat.	- Vertrauen	- Konsens bzgl. der Ziele und des Vorgehens	
<i>Bordin (1979)</i>			- Übereinstimmung bzgl. der Therapieziele - Übereinstimmung bzgl. der Aufgaben - Bindung/„Wir-Gefühl“	z.B. WAI (Working Alliance Inventory, Horvath & Greenberg, 1986)
<i>Hartley &amp; Strupp (1983)</i>	- Beitrag des Ther. zur Allianz (z.B. Vermittlung von Expertise; Verständnis; Respekt)	- Beitrag des Pat. zur Allianz (z.B. Widerstand; Motivation)	- Therapeut-Patient-Interaktionen	VTAS (Vanderbilt Therapeutische Beziehungsskala, Hartley & Strupp, 1983)

4 Ein ausführlicher Überblick der verschiedenen Messinstrumente findet sich bei Hank & Krampen (2006).

5 Von Schulte werden alle acht aufgeführten Komponenten als Therapeutenvariablen angesehen, da sie von diesem beeinflussbar sind. Die Komponenten „Vertrauen“ und „Konsens“ wurden in der tabellarischen Darstellung dennoch in die Spalten „Patientenvariablen“ bzw. „Interaktionsvariablen“ aufgeführt, da ihre Ausprägung deutlich vom Patienten abhängt.

### 3. Unspezifische/allgemeine Wirkfaktoren in der Psychotherapie

<i>Gaston (1991)</i>	- Verständnis und Engagement des Therapeuten	- Zusammenarbeit mit dem Therapeuten - Arbeitsfähigkeit des Patienten	- Übereinstimmung bzgl. Vorgehen u. Ziele	CALPAS (California Psychotherapy Alliance Scale; Gaston, 1991)
<i>Luborsky (1994)</i>	- Der Therapeut wird als hilfreich und unterstützend erlebt (durch Wärme; Unterstützung; Respekt; Wertschätzung u.a.)	- Glaube an den Wert der Behandlung	- Die Beziehung wird als hilfreich erlebt (durch Wir-Gefühl; Übereinstimmung bzgl. Störungskonzept; Fähigkeit zur Kooperation u.a..)	HAQ (Helping Alliance Questionnaire, Alexander & Luborsky, 1986)
<i>Agnew-Davies et al. (1998)</i>		- Vertrauen - Offenheit - Initiative	- Bündnis: Akzeptanz, Verständnis, Unterstützung - Partnerschaft: gemeinsame Arbeit an den therapeutische Aufgaben	ARM (Agnew Relationship Measure, Agnew-Davies et al., 1998)

Die tabellarische Auflistung verdeutlicht die Schwierigkeit der Festlegung bzw. der Definition des komplexen Konstruktes der „therapeutischen Beziehung“ bzw. der „therapeutischen Allianz“. Welche Variablen stellen separate Faktoren dar, die ihrerseits Einfluss auf die therapeutische Beziehung nehmen, und welche Variablen sind als Subskalen unter das Konstrukt zu subsumieren? Zum Beispiel die Variablen „Empathie, Akzeptanz und Wertschätzung des Therapeuten“<sup>6</sup> - Sind sie Bestandteile der therapeutischen Beziehung oder beeinflussen sie diese und sind eher dem Komplex „Kompetenz des Therapeuten“ zuzuordnen? Ähnlich verhält es sich auf Patientenseite z.B. mit der Variable „Offenheit“ des Patienten. Hinsichtlich der Interaktionsvariablen scheint es größere Übereinstimmung zu geben. Hier werden regelmäßig zwei Aspekte als relevante Komponenten aufgeführt, zum einen ein inhaltsbezogener Aspekt, nämlich die Übereinstimmung zwischen Therapeut und Patient hinsichtlich der therapeutischen Inhalte (Ziele, Vorgehen bzw. Aufgaben der Therapie, Störungskonzept etc.) und zum anderen ein gefühlsbezogener Aspekt, also die erlebte Beziehung zwischen Therapeut und Patient („Wir-Gefühl“, Erleben von Bindung und Partnerschaft). Dass diese beiden Aspekte in enger Verbindung zu den oben genannten Therapeuten- und Patientenvariablen stehen ist meines Erachtens unmittelbar nachvollziehbar, was allerdings die Frage nach deren Unabhängigkeit nicht klärt.

6 Die drei Basisvariablen „Empathie, Echtheit, Akzeptanz“ werden von C. Rogers als Hauptwirkfaktoren der Gesprächspsychotherapie angeführt (z.B. Kriz, 1991), weshalb sie aus dieser Perspektive auch nicht als allgemeine, sondern als (gesprächspsychotherapie-) spezifische Wirkfaktoren angesehen werden könnten. Da diese Faktoren aber in allen Therapieschulen als bedeutsam für den Therapieerfolg erachtet werden („In gewisser Weise sind alle Therapeuten Rogerianer“, Psychologie heute, 2009, 21, S. 46), sollen sie dennoch als allgemeine Wirkfaktoren behandelt werden.

Für die Erfassung allgemeiner Wirkfaktoren im Rahmen von Therapiestudien erscheint es in erster Linie wichtig, dass die relevanten Faktoren berücksichtigt werden, um Gruppenvergleiche zu ermöglichen. Die Zusammenhänge zwischen den Variablen sind für dieses Vorhaben von untergeordneter Relevanz bzw. können mithilfe ihrer Erfassung vielleicht besser untersucht werden. Die Trennung von Therapeutenvariablen, Patientenvariablen und Interaktionsvariablen erscheint aus dieser Sicht von Vorteil, da sie von weniger Vorannahmen bzgl. der Interdependenzen der einzelnen Variablen ausgeht. Daher sollen im Weiteren unter den Begriffen „therapeutische Beziehung“/„therapeutische Allianz“ nur diejenigen Variablen zusammengefasst werden, die die Interaktion zwischen Therapeut und Patient betreffen (Übereinstimmung bzgl. der Aufgaben und Ziele/Zusammenarbeit; emotionale Bindung), während therapeutenbezogene Variablen (v.a. Empathie Wertschätzung, Akzeptanz) und patientenbezogene Variablen (Mitarbeit, Selbstöffnung) als separate Faktoren neben der therapeutischen Beziehung behandelt werden sollen, wie es im Falle der letztgenannten Variablen ja im vorangegangenen Abschnitt bereits geschehen ist.

### **3.3. Empirische Untersuchungen zu allgemeinen Wirkfaktoren**

In diesem Kapitel soll zunächst insgesamt die Bedeutung allgemeiner Wirkfaktoren für den Behandlungserfolg anhand empirischer Studien dargestellt werden. Empirische Befunde sollen weiterhin zur Klärung der Frage herangezogen werden, welche unspezifischen bzw. allgemeinen Faktoren erfasst werden sollten.

Ähnlich wie im vorangegangenen Abschnitt müssen aufgrund des Umfangs auch hier Einschränkungen gemacht werden. In erster Linie soll daher auf metaanalytische Arbeiten Bezug genommen werden. Weiterhin soll unter empirischer Perspektive eine Betrachtung derjenigen Variablen erfolgen, die sich in den vorangegangenen Kapiteln als bedeutsam erwiesen haben.

#### **3.3.1. Zur Effektivität allgemeiner Wirkfaktoren: Metaanalysen**

Lambert (1992) stellt auf der Basis der Ergebnisse verschiedener Reviews in seinem Artikel zum Thema „Ergebnisforschung in der Psychotherapie“ ein Schaubild vor, das in etwa die prozentualen Anteile der verschiedenen therapeutischen Einflussfaktoren wiedergeben soll. Demnach gehen 40% der Verbesserung auf Faktoren zurück, die außerhalb des therapeutischen Einflusses lie-

### 3. Unspezifische/allgemeine Wirkfaktoren in der Psychotherapie

---

gen („extratherapeutic change“, z.B. durch zufällige Ereignisse oder soziale Unterstützung). 30% seien auf allgemeine Wirkfaktoren zurückzuführen („common factors“), wobei Lambert hierunter in erster Linie Therapeutenvariablen versteht (Empathie, Wärme, Akzeptanz, Unterstützung). 15% der Verbesserung kämen durch Erwartungseffekte zustande, die Lambert mit Placebo-Effekten gleichsetzt, und lediglich für 15% seien die spezifischen Techniken (wie z.B. Biofeedback, Hypnose oder systematische Desensibilisierung) verantwortlich. Der Autor betont, dass diese Angaben nicht mit statistischen Berechnungen gewonnen worden seien, so dass die Prozentangaben vielleicht eine Präzision vermittelten, die nicht gegeben sei. Dennoch gingen diese Angaben - so Lambert - auf die Ergebnisse empirischer Studien zurück, so dass eine solide Basis zur Untermauerung dieser Aussagen gegeben sei. Genauere Angaben zur Herkunft dieser Ergebnisse liefert der Autor allerdings nicht.

Fasst man die bei Lambert getrennten Variablengruppen der „common factors“ und der „expectancy“ zusammen, so ergibt sich ein Anteil von 45% am Behandlungserfolg, der nach Auffassung des Autors auf diese Faktoren zurückgeht, während nur 15% den spezifischen Techniken zuzurechnen ist. Demnach bestünde eine Wirksamkeitsrelation zwischen allgemeinen Faktoren (incl. der Erwartungseffekte) und spezifische Faktoren von 3:1.

In einer späteren Arbeit schätzen Lambert & Ogles (2004) das Verhältnis zwischen allgemeinen und spezifischen Faktoren allerdings etwas anders ein: demnach läge die Effektstärke (ES) für die Placebo-Therapie im Schnitt bei .42, während die durchschnittliche ES für die Psychotherapie insgesamt (allgemeine plus spezifische Faktoren) .82 erreiche.

Eine Metaanalyse von 69 Studien führten Bowers & Clum (1988) durch, wobei sie verhaltenstherapeutische Behandlungen mit Placebo-Kontrollbedingungen verglichen, um den zusätzlichen Beitrag am Behandlungserfolg der spezifischen Interventionen gegenüber den unspezifischen Wirkfaktoren (in dieser Studie „Placebo-Faktoren“) zu eruieren. Eine Besonderheit der Studie besteht darin, dass die Autoren eine Einschätzung der Glaubwürdigkeit der Placebo-Bedingungen vornahmen („Glaubwürdigkeits-Rating“), um den Einfluss der Glaubwürdigkeit auf die Effektstärken bestimmen zu können. Bowers & Clum gingen davon aus, dass die Erwartung bzw. der Glaube der Patienten an die Effektivität der vorgeschlagenen Interventionen einen basalen unspezifischen Wirkfaktor darstellt und dass dieser wiederum von der Glaubwürdigkeit der Placebo-Kontrollbedingungen abhängt. Diese wird ihrerseits durch verschiedene Aspekte beeinflusst, z.B. dadurch, dass ein Kontakt mit einem qualifizierten Therapeuten stattfindet, dass die

### 3. Unspezifische/allgemeine Wirkfaktoren in der Psychotherapie

---

Vermittlung eines Therapierationales erfolgt und eine den anderen Therapiegruppen entsprechende Häufigkeit und Dauer der Therapiesitzungen gewährleistet ist.

Einbezogen wurden alle Studien, die eine unbehandelte Kontrollgruppe (KG), eine „Placebo“-Gruppe und eine komplette Behandlungsgruppe untersucht hatten. Dabei sollten die Placebo-Gruppen mit den Behandlungsgruppen in allen Aspekten mit Ausnahme der spezifischen Interventionen übereinstimmen. Entsprechend bestand die Annahme, dass Unterschiede in den Behandlungsergebnissen zwischen den Placebo-Gruppen und den Behandlungsgruppen auf die Wirkung der spezifischen Faktoren zurückführbar sind, während Unterschiede zwischen den Placebo-Gruppen und den unbehandelten Kontrollgruppen der Wirkung der unspezifischen Faktoren zugeschrieben werden sollten.

Die mittlere Effektstärke (ES)<sup>7</sup> für die Therapie insgesamt (behandelte Gruppen vs. unbehandelte Gruppen) betrug  $d = .76$ . Die Effektstärke der spezifischen Faktoren, gemittelt über alle Studien, betrug  $d = .55$ ; für die Studien mit sehr glaubwürdigen Placebo-Gruppen betrug die ES für die spezifischen Faktoren  $.60$ , was den Erwartungen der Autoren widersprach. Die Schätzung für die Effekte der unspezifischen Faktoren lag bei  $d = .21$ . Die Berechnung ergab somit, dass die spezifischen Faktoren der Verhaltenstherapie einen mehr als doppelt so großen Anteil am Behandlungserfolg haben wie die unspezifischen Effekte. Entgegen der Annahme der Autoren ergaben sich keine Auswirkungen der Glaubwürdigkeit der Placebo-Kontrollgruppen auf die Effektstärken. Möglicherweise lag dies an den sehr weit gefassten Kriterien, die zu der Einschätzung hoher Glaubwürdigkeit führte (z.B. Ratings, Anhaltspunkte im Studiendesign, pharmakologische Placebos).

Bowers & Clum werteten das Ergebnis ihrer Metaanalyse insgesamt als Beleg dafür, dass es sich bei psychologischen Placebos (definiert als unspezifische Interventionen, die eine Interaktion zwischen Therapeut und Patient und ein Gespräch über die Probleme des Patienten beinhalten) um selbständige Formen der Therapie handelt, die einen eigenen Behandlungseffekt erzielen. Ebenfalls bestätigen ließ sich aber andererseits die Annahme, dass spezifische Techniken einen deutlichen Zugewinn der Behandlung, zusätzlich zu den unspezifischen Effekten, bewirken.

---

7 Die Effektstärken wurden nach der üblichen Formel, die von Smith & Glass (1977, zit. nach Bowers & Clum, 1988) empfohlen wurde, berechnet: der Mittelwert der Placebo- bzw. der nichtbehandelten KG nach der Behandlung wurde subtrahiert vom Mittelwert der Behandlungsgruppe nach Therapieende und geteilt durch die Standardabweichung der Placebo- bzw. unbehandelten KG.



### 3. Unspezifische/allgemeine Wirkfaktoren in der Psychotherapie

---

Barker et al. (1988) führten eine Metaanalyse mit ähnlicher Zielsetzung wie die vorgenannten Autoren durch (Vergleich der Wirksamkeit psychotherapeutischer Behandlung im Verhältnis zur Effektivität unspezifischer Faktoren einerseits und unbehandelter Kontrollgruppenbedingungen andererseits), allerdings mit strengeren Eingangskriterien. Es wurden hier nur Studien berücksichtigt, in denen die Erwartungen der Patienten hinsichtlich des Therapieerfolges überprüft wurden und sich diesbezüglich keine Unterschiede zwischen Behandlungs- und unspezifischen Kontrollgruppen (in dieser Analyse: „nonspecific factors control“: NSF) zeigten. Aufgrund dieser Kriterien konnten nur 17 von 200 Studien, die einem Review unterzogen worden waren, für die Untersuchung ausgewählt werden. Die Erwartungen der Patienten wurden meist mit Hilfe eines Fragebogens überprüft („expectancy check“), der Items über die Logik, die Verständlichkeit und Nützlichkeit des Vorgehens enthielt und die Frage, ob man die Behandlung auch einem Freund empfehlen würde.

Die mittlere Effektstärke für die spezifischen Interventionen (Behandlungsgruppe vs. NSF) lag zum Behandlungsende (Post-Messung) bei  $d = .55$ . Die ES der allgemeinen Faktoren (NSF vs. unbehandelte KG) erreichte einen Wert von  $d = .47$ . Ein Wert von  $d = 1.06$  wurde für die Effektivität der Gesamtbehandlung ermittelt (Behandlungsgruppe vs. unbehandelte KG). Die Ergebnisse bestätigten nach Aussagen der Autoren die Annahme, dass es sich bei der Gesamteffektstärke um eine additive Zusammensetzung der Effekte der allgemeinen und spezifischen Faktoren handelt. Die Ergebnisse dieser Metaanalyse sprechen dafür, dass etwa die Hälfte der Effekte auf die Wirkung der allgemeinen Faktoren zurückgeht. Zusätzlich zu den Effektstärken am Behandlungsende errechneten Barker et al. auch Effektstärken für die Follow-Up-Messung. Dabei zeigte sich die Gesamt-ES nahezu unverändert ( $d = .96$ ). Allerdings verringerte sich die ES der spezifischen Interventionen auf  $d = .17$ , während die ES der allgemeinen Faktoren auf  $d = .72$  anstieg. Somit liegt die Schlussfolgerung nahe, dass der Anteil der allgemeinen Faktoren am Gesamterfolg zwischen der Post-Messung und dem Zeitpunkt der Follow-Up-Messung zunimmt.

Auch Stevens et al. (2000) führten eine metaanalytische Überprüfung von 80 Outcome-Studien durch, um die Zusammenhänge zwischen allgemeinen Faktoren („common factors“) und spezifischen Interventionen auf der einen und Behandlungsergebnissen auf der anderen Seite zu überprüfen. Die Autoren definierten Psychotherapie ebenfalls als „the additive product of (a) com-

### 3. Unspezifische/allgemeine Wirkfaktoren in der Psychotherapie

---

mon (nonspecific, placebo) factors found in many different treatment approaches, plus (b) specific factors proper to particular treatment methods and theories“ (p. 273). Unter den „common factors“ fassten sie in erster Linie Erwartungseffekte und Beziehungsvariablen zusammen (Wärme, Aufmerksamkeit, Verständnis, Unterstützung).

Speziell interessierte sie der Einfluss allgemeiner vs. spezifischer Behandlungseffekte auf verschiedene Aspekte des Behandlungserfolgs, nämlich auf das subjektive Wohlbefinden, auf die Anzahl und Ausprägung der Symptome und auf das allgemeine Funktionsniveau. Entsprechend dem Phasen-Modell der Psychotherapie von Howard et al. (1993) gingen die Autoren davon aus, dass es drei Phasen von Verbesserungen beim Patienten gibt (zunächst gesteigertes Wohlbefinden, dann Symptomreduktion und schließlich eine Erhöhung des allgemeinen Funktionsniveaus). In Bezug auf dieses Modell formulierten Stevens et al. die Annahme, dass die allgemeinen Wirkfaktoren eher für die frühen Therapieerfolge verantwortlich seien, während die späteren Verbesserungen eher auf die spezifischen Faktoren zurückzuführen sein dürften.

Die mittleren gewichteten Effektstärken lagen bei  $r^8 = .11$  für die allgemeinen Behandlungseffekte (common factors), bei  $r = .19$  für die spezifischen Behandlungseffekte und bei  $r = .28$  für die Gesamt-Behandlungs-Effekte (entsprechende  $d_s = .22, .38, .56$ ).

Somit zeigte sich, dass die Gesamt-Effektstärke annähernd die Summe aus der common-factor-ES und der ES der spezifischen Effekte darstellte, wobei die ES der spezifischen Faktoren der Wirkung der allgemeinen Faktoren deutlich überlegen war. Die Annahme der Autoren über die differentielle Wirkung der allgemeinen Faktoren auf die verschiedenen Phasen der Psychotherapie und damit die verschiedene Aspekte ließ sich nicht bestätigen. Allerdings zeigten sich Unterschiede hinsichtlich des Schweregrades der Erkrankung: bei gravierenden Erkrankungen (z.B. schizophrene Erkrankungen) ergaben sich keine signifikanten Effekte der allgemeinen Faktoren.

Insgesamt fielen die Effektstärken in dieser Metaanalyse zwar relativ gering aus, was die Autoren v.a. hinsichtlich der Bedeutung der allgemeinen Faktoren anders erwartet hatten, aber es zeigten sich die erwarteten Relationen (die Wirksamkeit der Gesamtbehandlungen übertraf die Wirksamkeit der common-factor-Gruppen, die ihrerseits der Nicht-Behandlung überlegen war. Die Wirkung der spezifischen Faktoren scheint gemäß dieser Studie die Effektivität der allgemeinen Faktoren etwa im Verhältnis 2:1 zu übertreffen.

---

8 Das übliche Maß für die Angabe von ES (d) wurde in dieser Studie umgerechnet in das Maß r mit  $r=d/(d^2+4)^{1/2}$  da dieses Vorgehen nach Ansicht der Autoren ein geringeres Ausmaß an Korrekturformeln erforderlich macht.

Um das Problem des Designs nicht-äquivalenter Kontrollgruppen herauszustellen, gingen Baskin et al. (2003) in einer Analyse von 21 Primärstudien der Frage nach, inwiefern strukturelle Unterschiede zwischen unspezifischen Behandlungs-Gruppen und „bona fide“ Therapiegruppen in Psychotherapiestudien zur Verzerrung der Ergebnisse beitragen. Sie stellten die Annahme auf, dass Unterschiede wie z.B. eine ungleiche Anzahl und Dauer der Therapieeinheiten, unterschiedliches Training der Therapeuten oder die Einschränkung der Themen zu einer überhöhten Einschätzung der spezifischen Effekte gegenüber der Wirkung der unspezifischen Faktoren führten. Entsprechend sollten die ES der Studien mit nicht-adäquaten Kontrollgruppen (KG) für die spezifischen Faktoren höher ausfallen als die ES der Studien mit strukturell äquivalenten Kontrollbedingungen. Die Einschätzung, ob es sich bei den Placebo-KG um äquivalente oder nicht-äquivalente KG handelte, nahmen unabhängige „blinde“ Rater vor.

Für die Studien, die nicht-äquivalente KG untersucht hatten, ergab sich eine ES für die spezifischen Interventionen (Behandlungsgruppe vs. unspezifische/Placebo-KG) von  $d = .47$ . Die Studien mit strukturäquivalenten KG erbrachten hingegen nur eine ES für die spezifischen Faktoren von  $d = .15$ . Entsprechend der Annahme der Autoren zeigten sich die spezifischen Therapien dann den unspezifischen Kontrollgruppen überlegen, wenn sie strukturell nicht ebenbürtig waren. Bei Studien mit einem sorgfältigen Design der Kontrollgruppen verschwanden diese Unterschiede allerdings fast gänzlich, so dass die Autoren es als bestätigt ansahen, dass spezifische Faktoren keine Effekte aufweisen, die über die Wirksamkeit der allgemeinen Faktoren hinausgehen.

Entsprechend betonen die Autoren die Bedeutung des Designs der unspezifischen Kontrollgruppen für die Wirksamkeit der Therapie („Outcome“) und somit auch für die Aussage der jeweiligen Studien.

#### **3.3.1.1. Fazit**

Zur Veranschaulichung sind in Tabelle 2 die Ergebnisse der dargestellten Metaanalysen noch einmal zusammengefasst.

Tabelle 2

### 3. Unspezifische/allgemeine Wirkfaktoren in der Psychotherapie

#### Überblick: Effektstärken (ES) spezifischer und allgemeiner Wirkfaktoren

Autoren der Studie	Gesamt-ES (d)	ES der spezifischen Faktoren	ES der allgemeinen Faktoren	Verhältnis Anteil spez. vs. allg. Faktoren am Behandlungserfolg	Besonderheit
Lampert (1992)	-	-	-	1 : 3	Keine statistischen Verfahren zur Gewinnung der Ergebnisse
Lambert (2004)	.82	[.40]	.42	1 : 1	
Bowers & Clum (1988)	.76	.55	.21	2,5 : 1	Überprüfung der Glaubwürdigkeit (kein Effekt gefunden)
Barker, Funk & Houston (1988)	1.06 (.96)*	.55 (.17)*	.47 (.72)*	1 : 1 (1 : 4)*	Kontrolle der Erwartungen in beiden Gruppen; ES-Berechnung zum Follow-Up-Zeitpunkt
Stevens, Hynan & Allen (2000)	.56	.38	.22	2 : 1	
Baskin et al. (2003)	-	.47** .15 <sup>+</sup>	-	-	Vergleich der ES für strukturäquivalente u. Nicht-äquivalente KG

Gesamt-ES = Vergleich Therapie-Gruppe vs. Unbehandelte KG; ES der spez. Faktoren = Vergleich Therapiegruppe vs. Unspezifische KG; ES der allgemeinen Faktoren = Vergleich Unspezifische KG v. Unbehandelte KG.

\* Kursiv: Ergebnisse zum Zeitpunkt d. Follow-Up-Messung. \*\* ES für Studien mit nicht-struktur-äquivalenten KG; + Kursiv: ES für die Studien mit struktur-äquivalenten KG

Aus Tabelle 2 wird ersichtlich, dass sich die Ergebnisse der dargestellten Studien deutlich unterscheiden. Dies gilt sowohl für die Beurteilung der Gesamtwirksamkeit der Psychotherapie (Ränge von  $d = .56$  bis  $1.06$ ) und vor allem für die Einschätzung der Anteile von spezifischen und unspezifischen Faktoren am Behandlungserfolg (outcome). Einigkeit besteht letztlich nur dahingehend, dass die allgemeinen Faktoren ihren Beitrag zum Behandlungserfolg leisten, unklar bleibt allerdings welchen.

Als grundsätzliche Schwierigkeit wird beim Vergleich der Arbeiten deutlich, dass es über die verschiedenen Studien hinweg keine einheitliche Definition gibt, was als „common factor“ zu gelten hat - nach Weinberger (1995) das „common factor dilemma“. Zudem wurde nur partiell

kontrolliert (z.B. Barker et al., 1988, Baskin et al., 2003), ob die common-factor-Gruppen tatsächlich hinsichtlich der relevanten Faktoren (z.B. der Erwartung, der therapeutischen Beziehung, der Struktur) das gleiche Niveau erreichten wie die Behandlungsgruppen und somit valide Vergleichsgruppen darstellten. Entsprechend den Ergebnissen von Baskin et al. (2000) ist anzunehmen, dass bislang in manchen Studien Kontrollgruppen verwendet wurden, die diesen Namen eigentlich nicht verdienen, da sie die relevanten Variablen (v.a. die allgemeinen Faktoren) eben nicht durch Konstanthaltung im Vergleich zur Therapiegruppe kontrolliert haben, sondern eher zur Verzerrung und Überschätzung der Effektivität der spezifischen Faktoren beigetragen haben.

Deutlich zeigt sich bei der Betrachtung dieser Ergebnisse der Bedarf nach einer einheitlichen Konzeption der allgemeinen Faktoren auf der einen Seite und v.a. nach einer Überprüfung ihrer Ausprägung in den zu vergleichenden Gruppen auf der anderen Seite. Da, wie es Baskin et al. formulieren, gilt: „This assertion [that the purported active ingredients were responsible for the benefits of the treatment] could only be made as long as the placebo treatment was sufficiently well designed to rule out factors such as the therapeutic relationship, expectation for benefits, the healing context and so forth.“ (Baskin et al., 2003, S. 973ff.)

#### **3.3.2. Zur Effektivität allgemeiner Wirkfaktoren: Studien zu einzelnen Variablen**

Da im letzten Abschnitt Studien vorgestellt wurden, in denen die Gesamt-Wirksamkeit allgemeiner Faktoren auf dem Prüfstein stand, aber wenige Aussagen zu einzelnen Variablen getroffen wurden, soll es in diesem Kapitel um Studien gehen, die das Ziel verfolgt haben, die Zusammenhänge zwischen einzelnen Variablen aus dem Bereich der allgemeinen Wirkfaktoren und dem Therapieergebnis („Outcome“) näher zu untersuchen. Im Fokus des Interesses stehen dabei diejenigen Variablen, die sich in den vorangegangenen Kapiteln, ausgehend von theoretischen Konzeptionen, als bedeutsam erwiesen haben.

Das waren:

- Die (positive) Erwartung des Patienten hinsichtlich des Therapieverlaufes (Veränderungserwartung).
- Die therapeutische Beziehung mit den Bestandteilen „Übereinstimmung hinsichtlich des therapeutischen Vorgehens“ (Aufgaben und Ziele) und „emotionale Bindung“.
- Die Empathie, Wertschätzung und Akzeptanz des Patienten durch den Therapeuten.
- Die Glaubwürdigkeit der Behandlung.

### 3. Unspezifische/allgemeine Wirkfaktoren in der Psychotherapie

---

- Der „Expertenstatus“ des Therapeuten bzw. dessen professionelles Auftreten.
- Die Selbstöffnung des Patienten
- die Bereitschaft des Patienten zur Mitarbeit (aktive Mitarbeit, Bereitschaft zur Erprobung neuer Verhaltensweisen).

In ihrer groß angelegten Untersuchung versuchten Orlinsky, Grawe und Parks (1994) Ergebnisse der psychotherapeutischen Prozessforschung („Was ist Psychotherapie?“) und der Outcome-Forschung („Wirkt Psychotherapie?“) zusammen zu bringen, um die entscheidende Frage zu klären: „Was wirkt in der Psychotherapie?“ Dazu führten sie eine ausgiebige Literaturrecherche durch und analysierten etwa 2300 Studien hinsichtlich der Prozessvariablen. Als strukturierende Grundlage diente ihnen dabei das allgemeine Psychotherapiemodell („Generic Model of Psychotherapy“, Orlinsky & Howard, 1987, siehe Kapitel 3.1.4), das folgende Aspekte des psychotherapeutischen Prozesses unterscheidet: einen Aspekt, der sich auf das therapeutische Angebot bezieht (das Bestehen eines Abkommens bzgl. der Therapie, „therapeutic contract“), einen technischen Aspekt (die therapeutischen Interventionen), einen interpersonellen Aspekt (die therapeutische Beziehung, „therapeutic bond“), einen intrapersonellen Aspekt (die Selbstbezogenheit des Patienten, „self-relatedness“), einen klinischen Aspekt (der Auswirkungen der Therapiestunde) und schließlich einen zeitlichen Aspekt.

Im Folgenden soll nur auf diejenigen Bereiche Bezug genommen werden, die für die vorliegende Studie von Bedeutung sind, d.h. auf das therapeutische Angebot, den interpersonellen, sowie den intrapersonellen Aspekt<sup>9</sup>. Generell haben die Autoren eine von der vorliegenden Darstellung abweichende Klassifizierung der Variablen vorgenommen. Um einer Verwirrung vorzubeugen, erfolgt die Darstellung hier entsprechend dem in diesem Text bislang verfolgten Konzept, d.h. gemäß den Kategorien „Therapeutenvariablen“, „Patientenvariablen“ und „Interaktionsvariablen“, wobei die Variable „Glaubwürdigkeit der Behandlung“ hier den Therapeutenvariablen zugeordnet werden soll. Da in einem Großteil der Primärstudien keine Angaben zu den Effektstärken gemacht wurden, werden von Orlinsky et al. in erster Linie Angaben über signifikante positive Korrelationen zwischen der relevanten Variable und dem Outcome-Maß aufgelistet. Im vorliegenden Text werden entsprechend Prozentangaben hinsichtlich der Häufigkeit von signifikanten

---

<sup>9</sup> Der technische Aspekt bezieht sich auf die spezifischen Interventionen, die hier nicht im Mittelpunkt des Interesses stehen. Bei dem klinischen Aspekt handelt es sich gleichzeitig um ein Outcome-Maß, weshalb er in dieser Untersuchung außen vor gelassen werden soll, und der zeitliche Aspekt stellt eine von außen entweder festzulegende oder zumindest leicht messbare, und damit kontrollierbare Einheit dar.

positiven Korrelationen dargestellt, die sich auf die jeweilige Anzahl der überprüften Studien beziehen.

#### **3.3.2.1. Therapeutenvariablen**

Die Glaubwürdigkeit des Therapeuten („therapist credibility versus unsureness“) weist in 59% der Studien einen positiven Zusammenhang mit dem Therapieerfolg auf. Noch deutlichere Ergebnisse lassen sich für die Variable „therapist skillfulness“ finden: hier zeigt sich in knapp 70% der Studien ein positiver Zusammenhang mit dem Therapieerfolg. Die ES dieses Faktors lassen eine große Streubreite erkennen, viele liegen allerdings über .20. Die Beiträge des Therapeuten zur therapeutischen Beziehung<sup>10</sup> werden in 67% der Untersuchungen als bedeutsam für den Erfolg gewertet, v.a. wenn die Beurteilung aus Sicht des Patienten erfolgte.

Ein positiver Zusammenhang zwischen Therapeutenverhalten und Therapieerfolg konnte auch in einer Studie von Williams & Chambless (1990) bei agoraphobischen Patienten belegt werden, die mit einer in vivo Expositionstherapie behandelt wurden. Es zeigte sich in diesem Fall, dass sowohl das Selbstvertrauen des Therapeuten – nach Patienteneinschätzung – mit  $r = .39$ , als auch eine mitfühlende Haltung des Therapeuten und sein „Involviert-Sein“ mit  $r = .30$  mit dem Therapieerfolg korrelierten.

Als weiterer Beleg für den Einfluss der Therapeutenvariablen ist weiterhin eine Untersuchung von Burns & Nolen-Hoeksema (1992) anzuführen. Die Autoren konnten den kausalen Effekt der Empathie des Therapeuten auf den Erfolg einer kognitiv-behavioralen Therapie bei depressiven Patienten mit Hilfe von Strukturgleichungsmodellen nachweisen.

#### **3.3.2.2. Patientenvariablen**

Einen deutlich positiven Zusammenhang mit dem Therapieerfolg fanden Orlinsky et al. für die Variable „patient cooperation“, also für die Mitarbeit des Patienten (in 69% der Studien). Entsprechend zeigte sich ein negativer Zusammenhang zwischen Outcome und der Variable „Widerstand des Patienten“ (patient resistance). Ähnliches ergab sich für die Variable „patient role engagement“ (pos. Zusammenhänge in 65% der Studien). Als weitere Patientenvariable wurde der Aspekt „self-relatedness“ (meist in Form von „Offenheit“ vs. „Verteidigung“) untersucht. Hier war ein sehr deutlicher Zusammenhang mit dem Therapieerfolg feststellbar, der sich in ei-

---

<sup>10</sup> Leider fehlen konkretere Angaben in der Studie, welche Faktoren hier untersucht wurden. Anzunehmen ist, dass es sich dabei in erster Linie um die Faktoren Verständnis, Empathie, Akzeptanz, Wertschätzung und Unterstützung handelte.

nem Prozentsatz von 80% positiver Korrelationen und zum Teil recht hohen Effektstärken ausdrückt.

#### 3.3.2.3. Interaktionsvariablen

Orlinsky et al. fassen unter dem Begriff „therapeutic contract“, neben verschiedenen formalen Aspekten, auf die hier nicht näher eingegangen werden soll, auch die Punkte „expectational clarity“ und „goal consensus“<sup>11</sup> zusammen. Beide scheinen gemäß den untersuchten Studien eine Rolle für den Therapieerfolg zu spielen, allerdings nur, wenn die Patientenbeurteilung herangezogen wurde. Sofern Effektstärken (ES) berechnet wurden, lagen diese eher im unteren Bereich.

#### 3.3.2.4. Zur therapeutischen Beziehung

Generell zeigt ein Großteil der überprüften Studien bei Orlinsky et al. einen deutlich positiven Zusammengang zwischen der therapeutischen Beziehung und dem Outcome, wobei die ES bei etwa einem Viertel über .25 lagen.

Ähnliche Ergebnisse lieferte auch eine Metaanalyse von Martin et al. (2000), in der 79 Studien auf den Zusammenhang zwischen der therapeutischen Beziehung und dem Therapieerfolg hin untersucht wurden. In den Primärstudien wurden unterschiedliche Messinstrumente zur Erfassung der therapeutischen Allianz verwendet, wobei das WAI (Working Alliance Inventory, Horvath & Greenberg, 1986) am häufigsten verwendet wurde. In dieser Untersuchung wurde als gewichteter Schätzer für die Effektstärke das Maß  $r^{12}$  verwendet als eine Variation der Produkt-Moment-Korrelation.

Die gewichtete Gesamtkorrelation zwischen Allianz und Outcome lag bei Martin et al. bei .22, was einem moderaten Zusammenhang zwischen diesen Variablen entspricht. Das Ergebnis des durchgeführten Homogenitätstests sprach für eine homogene Population, was bedeutet, dass es keine Hinweise auf Moderatorvariablen gab, welche den Zusammenhang zwischen Allianz und Outcome und somit die Varianz beeinflusst hätten.

Martin et al. sehen in diesen Ergebnissen eine Bestätigung der Annahme, dass es sich bei der therapeutischen Allianz um einen eigenständigen Wirkfaktor handelt, unabhängig von den durchgeführten Interventionen, der sich von frühen Therapieerfolgen unterscheiden lässt.

---

11 Die Autoren zählen diesen Aspekt zum therapeutischen Kontrakt hinzu; andernorts – wie auch in der vorangehenden Darstellung – wird dieser Punkt der therapeutischen Beziehung zugerechnet.

12 
$$\bar{r} = \frac{\sum [N_i r_i]}{\sum [N_i]}$$
 mit  $r_i$  als der Korrelation in Studie  $i$  und  $N_i$  als der Anzahl an Individuen in Studien  $i$ .



### 3. Unspezifische/allgemeine Wirkfaktoren in der Psychotherapie

---

Unterstützt wird diese Aussage von einer Studie von Krupnick et al. (1996), die im Rahmen der NIMH Collaborative Depression Study nachweisen konnten, dass die therapeutische Allianz signifikante Zusammenhänge mit den Therapieergebnissen in allen vier Behandlungsbedingungen (IPT, kognitive VT, Pharmakotherapie + klinische Betreuung, Placebo + klinische Betreuung) aufwies und, dass durch die Allianz mehr Varianz aufgeklärt wurde als durch den spezifischen Therapieansatz.

Innerhalb einer großen Menge an Prozessvariablen die in der Studie von Orlinsky et al., und ergänzend in anderen Studien, untersucht worden sind, wurde im vorangehenden Abschnitt der Fokus auf diejenigen Aspekte gerichtet, die „interessante Kandidaten“ für die Auswahl relevanter und in Studien zu kontrollierender allgemeiner Wirkfaktoren darstellen. Dabei zeigte sich die Beteiligung des Patienten am therapeutischen Prozess – in erster Linie in Form von Offenheit bzw. im Sinne der Bereitschaft sich zu öffnen – als eine der wichtigsten Determinanten für den Therapieerfolg. Es erscheint plausibel, dass therapeutische Interventionen in ihrer Wirksamkeit abgeschwächt werden können durch eine Verteidigungshaltung bzw. den Widerstand von Patienten und dass auf der anderen Seite Patienten, die eine höhere Kooperationsbereitschaft zeigen und aktiv an der Therapie teilnehmen, mehr von effektiven Interventionen profitieren.

Die therapeutische Beziehung stellte sich als ein ebenso wichtiger Faktor für ein Gelingen des therapeutischen Prozesses heraus, und dies v.a. wenn sie aus der Perspektive des Patienten beurteilt wurde. Zum einen scheint die therapeutische Beziehung eine direkte Wirkung auf die Verfassung des Patienten zu haben und zum anderen scheint sie innerhalb des therapeutischen Systems Einfluss zu nehmen auf die Bedingungen, unter denen Therapie wirksam werden kann. Auf Therapeutenseite konnten den Variablen Empathie, Bestätigung, Zusammenarbeit und selbstkongruentes Engagement für die Therapie neben der erlebten Kompetenz in der Anwendung verschiedener spezifischer Interventionen als relevante Faktoren ausgemacht werden.

Aufgrund der großen Anzahl an gesichteten Studien und Befunden können diese Ergebnisse als gesichert und zuverlässig angesehen werden: „These consistent process-outcome relations, based on literally hundreds of empirical findings, can be considered *facts* established by 40-plus years of research on psychotherapy.“ (Orlinsky, et al., 1994, S. 361). Die Konsequenzen für das Design von Psychotherapiestudien erscheinen naheliegend: „In designing future studies, researchers in

particular should remember to control for the effects of process variables that have been found consistently related to outcome. These include, among others, the overall quality of the therapeutic relationship, therapist skill, patient cooperation versus resistance, patient openness versus defensiveness,[ and treatment duration]“(ebd., S.364).

Was bedeuten diese Ergebnisse nun für die Frage nach den zu erfassenden allgemeinen Wirkfaktoren? Die Studie von Orlinsky et al., die sicher als eine der umfassendsten in diesem Bereich angesehen werden kann, liefert solide empirische Belege für die Variablen „therapeutische Beziehung“, „Selbstöffnung des Patienten“, „Bereitschaft des Patienten zur Mitarbeit“, „Empathie, Wertschätzung und Akzeptanz des Patienten durch den Therapeuten“ und „Expertenstatus des Therapeuten“. Die in den vorangegangenen Kapiteln als zentraler Faktor angesehene Variable „Veränderungserwartung des Patienten“ wurde in der Untersuchung allerdings nicht berücksichtigt und auch die Bedeutung der Variablen „Glaubwürdigkeit der Behandlung“ blieb eher im Unklaren. Daher soll im Folgenden speziell auf diese Faktoren eingegangen werden.

Devilley und Borkovec (2000) untersuchten in einer Studie sowohl den Aspekt der Glaubwürdigkeit, als auch den Aspekt der Veränderungserwartung. Die Autoren entwickelten hierzu einen Fragebogen, der beide Faktoren erfassen sollte und als Kontrollinstrument für den Einsatz bei klinischen Outcome-Studien gedacht war („Unless comparison conditions are equivalent on these factors in the early stages of treatment, investigators cannot rule out differential levels of credibility and/or expectancy (...) as alternative explanations for any differences in observed outcome between the compared conditions.“ (Devilley & Borkovec ,2000, S.73f.). Der Faktor Glaubwürdigkeit wurde dabei von den Autoren eher als kognitionsbezogene, der Faktor Veränderungserwartung eher als emotionsbezogene Facette konzipiert. Die Variable Glaubwürdigkeit beschreibt, wie glaubhaft, überzeugend und folgerichtig die Behandlung dem Patienten erscheint. Der Begriff „Erwartung“ bezieht sich auf die antizipierten Verbesserungen, die nach Ansicht des Patienten durch die Therapie erreicht werden können.

Beide Konstrukte hatten sich in früheren Studien als signifikante Prädiktoren für den Therapieerfolg erwiesen (z.B. Borkovec & Costello, 1993, bzgl. der Erwartung und Kirsch & Henry, 1977, bzgl. der Glaubwürdigkeit, beide zitiert nach Devilly & Borkovec, 2000), wobei sich die Variable „Glaubwürdigkeit“ als schwächere Prädiktorvariable gezeigt hatte.

### 3. Unspezifische/allgemeine Wirkfaktoren in der Psychotherapie

---

Untersucht wurden die Zusammenhänge zwischen Glaubwürdigkeit/Erwartung und Outcome in drei unabhängigen Studien mit Stichproben aus unterschiedlichen Patientenpopulationen. Neben der Fähigkeiten der beiden Skalen, das Ergebnis der Therapie vorauszusagen, wurde die Faktorstruktur des Fragebogens, seine interne Konsistenz, sowie die Retest-Reliabilität überprüft.

Die psychometrischen Kennwerte des Fragebogens erwiesen sich durchgehend als gut und die Zwei-Faktoren-Lösung wurde durch die empirischen Daten unterstützt. Hinsichtlich der Zusammenhänge zwischen den Skalen und dem Outcome zeigte sich allerdings ein weniger klares Bild. Für die Variable „Erwartung“ ergaben sich signifikante Zusammenhänge, sofern zur Messung des Outcomes die Veränderung spezifischer Symptome und nicht das „allgemeine Funktionsniveau“ erfasst wurde. Für die Variable „Glaubwürdigkeit der Behandlung“ konnten nur deutlich schwächere, im Trend vorhandene Zusammenhänge gefunden werden, was den Ergebnissen früherer Untersuchungen (z.B. Borkovec & Costello, 1993; Bowers & Clum, 1988) entsprach.

Ähnlich wie Devilly und Borkovec (2000) entwickelte Schulte (2005) einen Fragebogen zur Erfassung von Patientenerwartungen und zur Beurteilung der Therapieglaubwürdigkeit, allerdings im Unterschied zu den erstgenannten Autoren für den deutschsprachigen Raum. Nach Schulte unterscheiden sich die beiden Konstrukte weniger hinsichtlich der Betonung rationaler (Glaubwürdigkeit) vs. emotionaler (Erwartung) Aspekte, als vielmehr hinsichtlich ihres Zeitbezuges: demnach handle es sich bei der Einschätzung der Glaubwürdigkeit vornehmlich um eine Beurteilung bzw. Wahrnehmung eines aktuell vorliegenden Sachverhaltes, während sich in der Erwartung die Vorhersage eines zukünftigen Sachverhaltes spiegele. Der Fragebogen von Schulte entstand auf der Grundlage zweier selbstkonstruierter Fragebögen zum Therapieprozess und besteht aus drei Skalen, die faktorenanalytisch gewonnen wurden: Hoffnung auf Besserung, Furcht vor Veränderung und Passung (was in etwa der Skala „Glaubwürdigkeit“ entspricht).

Vergleichbar mit den Ergebnissen der vorangegangenen Untersuchung ergaben sich nicht für alle Erfolgsmaße bedeutsame Zusammenhänge mit der Erfolgserwartung/Hoffnung auf Besserung bzw. Passung/Glaubwürdigkeit. Signifikante Ergebnisse waren nur dann zu finden, wenn der Therapieerfolg durch die Restsymptomatik (Postwerte der Symptomfragebogen) oder durch eine retrospektive Erfolgsbeurteilung zu Therapieende erhoben wurde. Hoffnung und Passung (erhoben nach der 4. Sitzung) erklären zusammen 16,2% der Varianz der retrospektiven Erfolgsbeur-

teilung, wobei die Spearman Rangkorrelation für die Variable Hoffnung .40 und für die Variable Passung .36 beträgt.

Einen weiteren Beleg für die Verbindung zwischen Erwartung und Therapieerfolg lieferte die Studie von Meyer et al. (2002), in der Daten aus dem National Institute of Mental Health Treatment of Depression Collaborative Research Program analysiert wurden. Die Autoren fanden einen Zusammenhang von .22 zwischen der Behandlungserwartung der Patienten und den Erfolgsmaßen (BDI, HAM-D). Der Zusammenhang zwischen therapeutischer Allianz (VTAS) und Therapieerfolg lag allerdings mit .49 fast doppelt so hoch und die Autoren konnten nachweisen, dass die therapeutische Allianz als Mediatorvariable für die Beziehung zwischen Erwartung und Erfolg fungiert.

Neben der indirekten Wirkung der Erwartung über die Allianz scheint diese Variable aber auch einen direkten Effekt auf das Ergebnis der Therapie zu haben, was eine Studie von Joyce et al. (2003) zeigen konnte.

Insgesamt kann festgehalten werden, dass sich für die Variablen „Glaubwürdigkeit des Therapieangebotes“ und „Veränderungserwartung des Patienten“ empirische Belege anführen lassen, die den Zusammenhang mit dem Therapieerfolg nachweisen. Allerdings scheint dieser Zusammenhang v.a. für das Konstrukt „Glaubwürdigkeit“ schwächer und weniger konsistent zu sein als für die vorangegangenen Variablen, wie z.B. die therapeutische Beziehung, und abhängig zu sein von der Operationalisierung des Therapieerfolges.

#### **3.3.3. Zusammenfassung**

Das Ziel der vorausgehenden Abschnitte war es, verschiedene Fragen hinsichtlich der Bedeutsamkeit allgemeiner Wirkfaktoren zu klären: Spielen allgemeine Wirkfaktoren eine – mit empirischen Methoden nachweisbare – Rolle für den Erfolg einer Therapie? Wenn ja, in welchem Umfang? Und *welche* Faktoren sind in diesem Zusammenhang von besonderer Bedeutung?

Die erste Frage lässt sich vom Standpunkt der vorgestellten Studien her meines Erachtens eindeutig mit ja beantworten. Bei der zweiten Frage treten allerdings Schwierigkeiten auf: die Forschungsergebnisse sind in diesem Fall keineswegs einheitlich, die ermittelten Effektstärken variieren stark und erlauben keine abschließende Einschätzung des Wirkungsumfangs allgemeiner Faktoren. Deutlich wird dabei das Kernproblem, dessentwegen die Messung und Kontrolle der allgemeinen Faktoren in klinischen Wirksamkeitsstudien unbedingt notwendig erscheint. Solan-

ge die Ausprägung der allgemeinen Faktoren in den Untersuchungsgruppen nicht mittels eines vergleichbaren und psychometrisch überprüften Messinstrumentes erhoben und somit eine Äquivalenz der verschiedenen Gruppen sicher gestellt wird, bleiben alle Aussagen über die Größe der Effekte letztlich Spekulation. Demnach liegt die Notwendigkeit für die Anwendung eines solchen Verfahrens auf der Hand.

Zur Klärung der Frage, welche Faktoren in diesem Zusammenhang erfasst werden sollten, wurden die Ergebnisse aus dem Abschnitt 3.1 über die theoretischen Konzeptionen herangezogen und die entsprechenden Variablen näher untersucht. Für alle acht Faktoren<sup>13</sup> („Glaubwürdigkeit/Nachvollziehbarkeit des Therapieangebotes“, „Empathie, Akzeptanz, Echtheit und Wertschätzung des Therapeuten“, „Kompetenz des Therapeuten“, „Veränderungserwartung des Patienten“, „Offenheit des Patienten“, „Bereitschaft des Patienten zur Mitarbeit“, „Emotionale Bindung zwischen Therapeut und Patient“ und „Zusammenarbeit zwischen Therapeut und Patient“) konnten empirische Belege gefunden werden. Allerdings fielen die Befunde nicht für alle Variablen mit gleicher Deutlichkeit aus: während z.B. für die Beziehungsvariablen eine Vielzahl an empirischen Nachweisen mit überzeugenden Korrelationen vorliegen, sind die Belege z.B. für den Faktor „Glaubwürdigkeit“ im Vergleich als eher inkonsistent zu bewerten.

Dennoch sind die empirischen Nachweise für alle genannten Variablen ausreichend und überzeugend genug, um eine regelmäßige Erfassung in Psychotherapiestudien zu fordern.

Entsprechend soll sich der nächste Abschnitt mit den Fragen auseinandersetzen, welche Instrumente derzeit zu diesem Zweck zur Verfügung stehen.

### **3.4. Messinstrumente zur Erfassung allgemeiner Wirkfaktoren**

Das umfangreichste Arsenal an Messinstrumenten ist sicherlich, zumindest im angloamerikanischen Raum, im Bereich der therapeutischen Allianz zu finden<sup>14</sup>. Bei diesen Verfahren handelt es sich entweder um Patientenbeurteilungen oder Therapeutenbeurteilungen in Form von Fragebö-

---

13 Die beiden Aspekte der therapeutischen Beziehung („Bindung zwischen Therapeut und Patient“ und „Zusammenarbeit zwischen Therapeut und Patient“) werden als einzelne Faktoren gezählt.

14 Einige der Instrumente wurden bereits im Kapitel 3.2 aufgeführt. Ein ausführlicher Überblick findet sich bei Hank & Krampen (2006).

gen oder um Verhaltensbeobachtungen von Fremdratern, wobei sich für die Patientenbeurteilungen nicht selten die höchsten Zusammenhänge mit dem Therapieergebnis zeigen (z.B. Horvath & Symonds, 1991). Die Vorteile dieser Verfahren liegen sicherlich in ihrer weiten Verbreitung, ihrer hohen Akzeptanz im englischsprachigen Raum, und der sich daraus ergebenden Dichte an Studien, die diese Verfahren verwendet und/oder überprüft haben. Insgesamt sprechen die Ergebnisse der Untersuchungen für eine hohe Messgenauigkeit der Verfahren, wobei aber etwas im Unklaren bleibt, *was* genau sie eigentlich messen (Hatcher & Barends, zit. nach Hank & Krampen, 2006). Die Schwierigkeiten liegen z.T. in der nicht einheitlichen Definition der therapeutischen Allianz/Beziehung und der sich daraus ergebenden sehr unterschiedlichen Skalenbildungen, wie in Kapitel 3.2 bereits ausgeführt wurde. Diese Unterschiedlichkeit ist teilweise auf die Schulenzugehörigkeit mancher, v.a. älterer Ansätze und der daraus entstandenen Skalen zurückzuführen, die für den hier skizzierten Ansatz ebenfalls als Nachteil zu werten ist. Als größtes Problem für die Verwendung dieser Verfahren zur Kontrolle allgemeiner Wirkfaktoren im Kontext psychotherapeutischer Wirksamkeitsstudien ist allerdings die häufig gegebene Vermischung von Beziehungselementen mit Therapieergebnissen (siehe auch: Hank & Krampen, 2006). Als ein Beispiel sei hier der HAQ (Helping Alliance Questionnaire, Alexander & Luborsky, 1986) genannt, einer der wenigen Fragebogenverfahren zur Allianz-Messung die auch ins Deutsche übersetzt und an einer deutschen Stichprobe überprüft wurden (Bassler et al., 1995). Der HAQ beinhaltet neben Items, die auf die Beziehung zum Therapeuten im engeren Sinn abzielen (z.B. „Ich habe das Gefühl, mich auf den Therapeuten verlassen zu können.“) und Items, die sich auf die Veränderungserwartung beziehen (z.B. „Ich glaube, dass mir die Behandlung hilft.“) auch Aussagen, die inhaltlich dem Ergebnis bzw. Erfolg der Therapie zuzurechnen sind (z.B. „Ich habe einige neue Einsichten gewonnen.“ oder „Seit kurzem fühle ich mich besser.“). Nicht verwunderlich erscheint es daher, dass eine explorative Überprüfung der Itemstruktur der deutschen Version (Bassler et al., 1995) neben dem Faktor „Beziehungszufriedenheit“ einen Faktor lieferte, der von den Autoren als „Erfolgswilligkeit“ bezeichnet wurde, was eindeutig ein Maß für den Therapieerfolg darstellt.

Es liegt auf der Hand, dass es für das Vorhaben, ein Messinstrument zur Erfassung und Kontrolle allgemeiner Wirkfaktoren zu entwickeln, wenig sinnvoll ist, Erfolgsmaße zu erfassen. Das Ziel besteht ja gerade darin, die allgemeinen Faktoren als unabhängige Variablen bzw. Covariablen in den verschiedenen Gruppen zu erfassen, um deren Einfluss auf den Therapieerfolg berechnen und kontrollieren zu können, und somit präzisere Aussagen über die Wirkung der spezifischen

### 3. Unspezifische/allgemeine Wirkfaktoren in der Psychotherapie

---

Interventionen zu ermöglichen. Bei allen Variablen, die Therapieerfolg erfassen, handelt es sich aber um abhängige Variablen, die unterscheidbar bleiben müssen von den allgemeinen Faktoren. Aus diesem Grund sind alle Instrumente, die Erfolgsmessungen beinhalten, für den genannten Zweck nicht geeignet.

Ähnliche Schwierigkeiten treten auch bei denjenigen Verfahren auf, die auf der Grundlage der in Kapitel 3.1 als „Metamodelle“ bezeichneten Konzepte entstanden sind, wie z.B. dem „generic model of psychotherapy“ von Orlinsky & Howard, 1987) oder dem Ansatz einer allgemeinen Psychotherapie von Grawe (2000). Als Beispiel sei hier der Stundenbogen für die Einzel-Psychotherapie und -Beratung (STEP) von Krampen und Wald (2001) angeführt, der sich auf die von Grawe postulierten Wirkprinzipien beruft. Dieser Fragebogen, der sowohl in einer Therapeuten-, als auch in einer Patientenform vorliegt, besteht aus drei Subskalen, die sich auf die Klärungs-, die Problembewältigungs- und die Beziehungsperspektive beziehen. Die Skala „motivationale Klärung und Entwicklung“ beinhaltet, ähnlich wie der HAQ, Items, die Auswirkungen der Therapie beschreiben und somit dem Therapieerfolg zuzurechnen sind (z.B. „In der heutigen Sitzung habe ich mehr Einsicht in meine Person und Schwierigkeiten gewonnen.“), was die oben genannten Probleme der Trennung zwischen unabhängigen und abhängigen Variablen mit sich bringt.

Als weitere Schwierigkeit kommt aber hinzu, dass die Items der Skala „Hilfe zur Problembewältigung“ keine Differenzierung von spezifischen Faktoren und allgemeinen Faktoren erlauben (z.B. „Heute habe ich praktische Hinweise zur Bewältigung meiner Schwierigkeiten erhalten.“). Wenn es aber gerade darum geht, die Auswirkungen allgemeiner Faktoren von denen der spezifischen Interventionen bzw. der speziellen therapeutischen Techniken zu trennen, macht eine Vermischung dieser beiden Aspekte auf Itemebene keinen Sinn. Aus diesem Grund sind entsprechende Verfahren – ohne ihren Wert für eine allgemeine Therapieevaluation in Frage stellen zu wollen – für Designs mit Kontrollgruppen in der Psychotherapieforschung ungeeignet.

Neben den genannten Verfahren gibt es evaluierte Fragebögen, die einzelne Konzepte aus dem Bereich der allgemeinen Faktoren erfassen, wie zum Beispiel das „Psychotherapy Expectancy Inventory – Revised“ (Bleyen et al., 2001) und den „Credibility/Expectancy Questionnaire“ (Devilly & Borkovec, 2000) oder den „Patientenfragebogen zur Therapieerwartung und Therapieevaluation (PATHEV)“ von Schulte (2005) und den auf der Bindungstheorie gründenden „Biele-

felder Fragebogen zu Klientenerwartungen“ (BFKE) von Höger (1999). Alle der genannten Fragebögen befassen sich mit wichtigen Einzelaspekten der allgemeinen Faktoren, greifen für das beschriebene Vorhaben aber zu kurz, da wichtige Aspekte außer acht gelassen werden.

#### **3.4.1. Fazit**

Betrachtet man die Schwierigkeiten der aktuell zur Verfügung stehenden Messinstrumente, so drängt sich die Schlussfolgerung auf, dass gegenwärtig kein Verfahren existiert, das die für bedeutsam erachteten Wirkfaktoren adäquat, d.h. unter Ausschluss sowohl von spezifischen Elementen, als auch von Erfolgsmaßen, erfassen kann. Gleichzeitig besteht ein dringender Bedarf für ein solches Instrument, um die allgemeinen Wirkfaktoren im Rahmen von Kontrollgruppenstudien im Bereich der Psychotherapieforschung routinemäßig erfassbar und damit kontrollierbar zu machen: Erst wenn die Anteile von allgemeinen und spezifischen Faktoren am Therapieergebnis abschätzbar werden, sind Gruppenvergleiche und Aussagen über den Erfolg einer spezifischen Therapie möglich.

Aus diesem Grund scheint es gerechtfertigt – trotz der insgesamt gesehen großen Fülle an bereits existierenden Messinstrumenten – ein neues Verfahren für den Einsatz in der Psychotherapieforschung zu entwickeln, das in der Lage ist, alle – oder zumindest einen Großteil – der relevanten allgemeinen Faktoren zu erfassen.

Im Folgenden soll also die Entwicklung und statistische Überprüfung eines Messinstrumentes zur Erfassung allgemeiner Wirkfaktoren in der Psychotherapie dargestellt werden.



## **III. Empirischer Teil**

### **1. Konzeption des Fragebogens zur Erfassung allgemeiner Wirkfaktoren**

#### **1.1. Vorgehen**

Um einen Überblick über die Vorgehensweise bei der Entwicklung des Fragebogens zu ermöglichen sind im Folgenden die einzelnen Konstruktionsschritte dargestellt:

- (1) Konzeptionelle Überlegungen
- (2) Theoretisch begründete Festlegung der Fragebogenskalen
- (3) Itemsammlung
- (4) Expertenbefragung
- (5) Erstellen der ersten Fragebogenversion
- (6) Durchführung der Fragebogenuntersuchung
- (7) Auswertung der Ergebnisse und faktorenanalytisch begründete Itemselektion
- (8) Revision der ersten Fragebogenversion

#### **1.2. Formale Aspekte**

##### **1.2.1. Wahl des Verfahrens**

Zur Erfassung von allgemeinen therapeutischen Wirkfaktoren eignen sich prinzipiell entweder die Verfahren der Verhaltensbeobachtung oder Fragebogenverfahren. Bei der Verhaltensbeobachtung codieren unabhängige, aber geschulte Beobachter Sequenzen aus Therapiegesprächen nach einem vorgegebenen Beurteilungssystem, Fragebögen werden in Zusammenhang mit Therapieeinschätzungen meist entweder von den Therapeuten oder den Patienten ausgefüllt.

Die verschiedenen Formen der Verhaltensbeobachtung bringen den Vorzug mit sich, dass Verzerrungen bzw. Verfälschungen durch die Versuchspersonen, d.h. in diesem Fall durch die Patienten oder die Therapeuten, im Unterschied zu Fragebogenverfahren relativ unwahrscheinlich sind.

Auf der anderen Seite muss man bei einer Verhaltensbeobachtung mit Verzerrungstendenzen auf Seiten der Beobachter rechnen, wie z.B. Kontrastfehler, Erwartungseffekte oder Fehler aufgrund der Überforderung der Differenzierungsfähigkeit (vgl. Fisseni, 1990). Weiterhin ist als Nachteil

der Verhaltensbeobachtung der enorme Material- und Zeitaufwand anzuführen: die Therapiestunden müssen gefilmt und anschließend von trainierten Ratern abschnittsweise eingeschätzt werden.

Zudem erscheinen nicht alle der in Frage stehenden Konstrukte gleichermaßen der Beobachtung zugänglich: während z.B. die therapeutische Beziehung, sofern die dafür repräsentativen Verhaltensweisen im Vorfeld klar definiert wurden, gut von externen Beobachtern einzuschätzen sein dürfte, so ist dies z.B. im Falle der Veränderungserwartung der Patienten deutlich schwieriger.

Aus diesen Gründen wurde als Verfahren zur Erfassung der allgemeinen Faktoren bzw. der relevanten Therapiebedingungen die Fragebogenmethode gewählt, auf deren Vor- und Nachteile im Folgenden noch näher eingegangen werden soll.

### **1.2.2. Vorteile der Fragebogenmethode**

Im Gegensatz zur Verhaltensbeobachtung stellt die Fragebogenmethode ein sehr ökonomisches, d.h. ein einfach durchzuführendes und leicht auszuwertendes Verfahren dar, was einen entscheidenden Vorzug darstellt. Auch der geringe Zeit- und Materialaufwand sprechen für den Einsatz dieser Methode. Weiterhin lassen sich mit einem Fragebogen, sofern es sich um ein Selbstbeurteilungsinstrument handelt, auch innere Zustände, Erlebnisweisen und Kognitionen erfragen, die sich von außen nur schwer beobachten lassen. Durch die Erfassung solcher von außen nur begrenzt wahrzunehmender Merkmale einerseits und die standardisierte Form der Itemdarbietung andererseits nimmt der Fragebogen eine Zwischenstellung zwischen einem „subjektiven“ und einem „objektiven“ Erhebungsverfahren ein (vgl. Mummendey, 1987). Und schließlich findet bei Fragebogenerhebungen durch die Abbildung der zu erfassenden Merkmale im Zahlenrelativ eine Quantifizierung der Merkmale statt, die Vergleiche in verschiedener Form ermöglichen, wie z.B. zwischen verschiedenen Aspekten der Zuverlässigkeit oder zwischen individuellen oder gruppenorientierten Scores (Fisseni, 1990).

### **1.2.3. Schwächen der Fragebogenmethode und Lösungsansätze**

Neben den entscheidenden Vorteilen der Fragebogenverfahren müssen allerdings auch die Schwächen dieser Methode berücksichtigt werden. Hier ist die hohe Sensibilität gegenüber sprachlichen Nuancen zu nennen: kleine sprachliche Veränderungen an der Item-Formulierung können zu großen Bedeutungsunterschieden führen (vgl. Mummendey, 1987). Daher sind eine

sehr sorgfältige Formulierung der Items, unter Berücksichtigung verschiedener Kriterien, und mehrere Durchgänge zu deren Überprüfung dringend erforderlich.

Eine weitere Schwäche stellt die Anfälligkeit von Fragebögen gegenüber Antwortstilen und Antworttendenzen dar. Zu den wichtigsten zählt die „Aquieszenz“ bzw. „Ja-Sage-Tendenz“, also die Einstellung der Probanden, den Antworten eher zuzustimmen. Des Weiteren muss damit gerechnet werden, dass die Reaktionen der Probanden durch die Tendenz beeinflusst werden, sozial erwünschte Antworten zu geben. Im vorliegenden Fall ist zudem auch der so genannte „leniency bias“ (Guilford, 1954; zit. nach Podsakoff, 2003) also die Tendenz zur Milde, von Bedeutung. Diese bezeichnet die Neigung der urteilenden Person, dann bessere Einschätzungen abzugeben, wenn es um die Beurteilung einer Person geht, die sie gut kennt oder zu der eine bedeutsame Beziehung besteht. Da die Beziehung zwischen Therapeut und Patient zweifelsohne eine bedeutsame Beziehung darstellt, ist immer dann mit einer Tendenz zur Milde bei den Patienten zu rechnen, wenn es um die Einschätzung der Therapeuten geht.

Eine völlige Vermeidung des Einflusses von Antwortstilen bzw. -tendenzen erscheint bei der Verwendung der Fragebogenmethode nicht erreichbar. Dies würde eine Verschleierung der eigentlichen Messintention erfordern, was weder durchführbar, noch sinnvoll wäre. Allerdings können bestimmte Strategien zur Abschwächung der Einflüsse der beiden genannten Antworttendenzen eingesetzt werden. Eine Möglichkeit zur Vermeidung von Aquieszenz stellt die Verwendung positiv und negativ formulierter Items dar, um Inhalt und Antwortstil voneinander zu trennen (vgl. DeVellis, 2003).

Der Tendenz zur sozialen Erwünschtheit kann durch die Formulierung der Instruktion entgegen gewirkt werden. Als Möglichkeiten bietet sich hier z.B. die Aufforderung an, „möglichst offen und ehrlich“ zu antworten. Zudem kann eine gewisse Reduktion der Sozialen-Erwünschtheits-Tendenz durch die Garantie der Anonymität erreicht werden (vgl. Mummendey, 1987). Ein Ansatz, die Tendenz zur Milde zu reduzieren, besteht darin, dem Urteilenden die Trennung zwischen Erhebung und Auswertung zu verdeutlichen, bzw. die fehlende Konsequenz des Urteils. Alle genannten Lösungsansätze zur Reduktion der Schwächen der Fragebogenmethode wurden bei der Konstruktion des Fragebogens berücksichtigt.

### **1.2.4. Wahl der Perspektive**

Neben der Entscheidung für eine Erhebungsmethode, musste noch die Frage geklärt werden, aus wessen Perspektive die infrage stehenden Merkmale erfasst werden sollten. Zur Wahl standen die Perspektive des Therapeuten auf der einen Seite, sowie die Perspektive des Patienten auf der anderen Seite. Für die Erfassung der relevanten Therapiebedingungen aus Sicht des Therapeuten sprach, dass diesem aufgrund seiner Fachkompetenz eine präzisere und klarere Bewertung möglich sein sollte. Auf der anderen Seite war zu bedenken, dass es auch um die Einschätzung therapeutischer Verhaltensweisen gehen sollte und dies aus Therapeutensicht vermutlich mit Verzerrungstendenzen einhergehen würde. Möglicherweise reagieren verschiedene Therapeuten in ihrem Urteil über die eigene Therapie idiosynkratisch, d.h. einige urteilen vielleicht optimistisch, andere unnachsichtig (vgl. Garfield, 1982). Zudem sollten auch Merkmale erhoben werden, die eher interne Zustände oder Vorgänge des Patienten darstellten, als beobachtbares Interaktionsverhalten, wie z.B. die Veränderungserwartung des Patienten.

Diese Merkmale lassen sich zwar von außen aufgrund von Äußerungen des Patienten und seinem Interaktionsverhalten erschließen, bleiben aber direkt schwer einschätzbar. Diese beiden Punkte sprachen für die Wahl der Patientenperspektive. Ein weiteres wichtiges Argument dafür stellten die Ergebnisse der Metaanalyse von Horvath & Symonds (1991) dar. In diese Analyse wurden 20 Studien einbezogen, in denen der Zusammenhang zwischen der therapeutischen Beziehung bzw. der therapeutischen Allianz und dem Therapieoutcome überprüft worden war. Die Autoren hatten dabei u.a. die Stärke der Zusammenhänge zwischen Allianzmessungen und Outcome aus drei Perspektiven verglichen, nämlich der Therapeuten-, der Patienten- und der Beobachterperspektive. Dabei hatte sich gezeigt, dass die Einschätzungen der therapeutischen Beziehung aus Patienten- und auch aus Beobachterperspektive stärker mit den verschiedenen Outcome-Maßen korreliert waren als die Einschätzungen aus Therapeutenperspektive. Zudem erwiesen sich in einer metaanalytischen Untersuchung von Martin et al. (2000) die Einschätzungen der therapeutischen Allianz aus Patientensicht über die Zeit hinweg stabiler als aus Sicht der Therapeuten oder der Beobachter.

Für die Konstruktion des Fragebogens fiel daher, unter Berücksichtigung der genannten Punkte, die Wahl auf die Perspektive des Patienten.

## **1.3. Inhaltliche Aspekte**

### **1.3.1. Festlegung der Fragebogenskalen**

Die Festlegung der a-priori-Skalen erfolgte entsprechend der im ersten Abschnitt dieses Textes in Kapitel 3 dargestellten theoretischen Konzeptionen und empirischen Befunde. Ziel war es, diejenigen Variablen in den Fragebogen mit aufzunehmen, die sowohl aus konzeptioneller Sicht, als auch in Hinblick auf die empirischen Daten (d.h. hinsichtlich ihres Zusammenhanges mit dem Therapieerfolg) für relevant erachtet werden.

Daraus ergaben sich folgende Ausgangsskalen

#### **1.3.1.1. Die Veränderungserwartung des Patienten**

Diese Skala soll die vom Patienten erwarteten Auswirkungen der Therapie erfassen. Damit ist sowohl der eher emotionale Aspekt der Hoffnung auf Hilfe und Linderung der Belastungen gemeint (Hoffnung auf Therapieerfolg), als auch die eher rational/kognitive Antizipation von positiven Veränderungen, die durch die Therapie erreicht werden können (z.B. das bessere Zurechtkommen im Alltag).

Das Konzept der Veränderungserwartung weist die größte Nähe zum Placebo-Konzept aus dem Bereich der Medizin bzw. Pharmakologie auf, da auch in diesem Fall der positiven Wirkungserwartung Veränderungspotential zugeschrieben wird.

#### **1.3.1.2. Die Offenheit des Patienten**

Mit der „Offenheit des Patienten“ ist seine Bereitschaft gemeint, sich zu öffnen und auf den therapeutischen Prozess einzulassen. In erster Linie bedeutet dies, offen zu berichten und mit möglichst wenig Verzerrungen zugunsten der interpersonellen Eindrucksbildung oder anderer Ziele von der eigenen Problematik zu berichten. Weiterhin ist mit Offenheit auch das Annehmen-Können von Informationen gemeint, die nicht mit dem gegenwärtigen Selbst-Schema übereinstimmen, sowie die Bereitschaft, im Rahmen des psychotherapeutischen Prozesses unangenehme Gefühle zuzulassen (z.B. Scham- und Schuldgefühle etc.).

#### **1.3.1.3. Die Bereitschaft des Patienten zur Mitarbeit**

Unter diesem Begriff sollen Aspekte zusammengefasst werden, die sich – in Abgrenzung zur Offenheit des Patienten – eher auf den aktiven Teil der Mitarbeitsbereitschaft beziehen. Damit ist v.a. das Engagement des Patienten gemeint und das Ausmaß an Mühe und Zeit, das er im Rah-

men der Therapie zu investieren bereit ist (z.B. in Hinblick auf Hausaufgaben, Termine...). Gleichzeitig soll mit dieser Skala auch die „Rollenerwartung“ erfasst werden, um eine aktive Mitarbeitsbereitschaft und eine eher passive Helferwartung unterscheiden zu können.

#### **1.3.1.4. Die therapeutische Beziehung**

Wie in Kapitel 3.2 des vorangehenden Abschnitts dieses Textes beschrieben werden unter dem Begriff „therapeutische Beziehung“ lediglich diejenigen Aspekte dieses Konstruktes subsumiert, die sich auf die Interaktion zwischen Therapeut und Patient beziehen. Diejenigen Bestandteile die entweder therapeutische Verhaltensweisen oder patientenbezogene Anteile beschreiben sollen separat erfasst werden. Es lassen sich zwei Interaktions-Aspekte unterscheiden:

##### **a) Die Zusammenarbeit zwischen Therapeut und Patienten („goal and task“)**

Diese Skala soll helfen die Frage zu klären, inwieweit sich der Patient in der Therapie, unter Berücksichtigung der jeweiligen Rollen, als gleichberechtigter Partner in der Therapie erlebt („Wir ziehen beide an einem Strang.“). Dies bezieht sich in erster Linie auf die Übereinstimmung hinsichtlich der therapeutischen Ziele und – in gewissen Grenzen – auch auf die Vorgehensweise und die Aufgaben in der Therapie.

##### **b) Die emotionale Bindung zwischen Therapeut und Patient („bond“)**

Dieser Aspekt der therapeutischen Beziehung umfasst die eher emotionale Komponente des gemeinsamen Bemühens um Veränderung, das „Wir-Gefühl“ im therapeutischen Prozess. Damit ist die Erfahrung des Patienten gemeint, sich in der Therapie aufgehoben, sicher und „am richtigen Platz“ zu fühlen, und das Erleben von Sympathie gegenüber dem Therapeuten.

#### **1.3.1.5. Die Empathie, Echtheit und Akzeptanz von Seiten des Therapeuten**

Diese Skala beinhaltet den Beitrag des Therapeuten zum Aufbau und Erhalt der therapeutischen Beziehung, der sich im Wesentlichen aus den „rogerianischen“ Variablen zusammensetzt. Das Sich-Einfühlen-Können in die Sichtweise und das Erleben des Patienten, das Verständnis für seine Verhaltensweisen und die unbedingte Akzeptanz sind hier als entscheidende Bestandteile zu nennen. Das Zeigen von Interesse an den Erfahrungen des Patienten und das Engagement hinsichtlich des therapeutischen Fortschrittes kommen hinzu. Und schließlich stellt die „Echtheit“, die Authentizität des Therapeuten ebenfalls einen wichtigen Teilaspekt dar.

### **1.3.1.6. Die Kompetenz des Therapeuten**

In Abgrenzung zu den therapeutenseitigen Beziehungsvariablen soll diese Skala die vom Patienten erlebte fachliche Kompetenz des Therapeuten beschreiben. Diese bezieht sich ebenso auf sein Fachwissen, also z.B. die bereit gestellten Informationen, als auch auf seine Überzeugungskraft und sein Auftreten. Erfasst werden soll, inwieweit der Therapeut als kompetenter und professioneller Helfer wahrgenommen wird, der weiß, was er tut und warum.

### **1.3.1.7. Die Glaubwürdigkeit und Nachvollziehbarkeit des Therapieangebotes**

Diese Skala bezieht sich auf den Inhalt des therapeutischen Angebotes. Hierbei geht es um die Frage, ob und inwieweit der Patient die Therapie als sinnvoll, passend und nachvollziehbar in Hinblick auf seine Schwierigkeiten erlebt. Insbesondere in Hinblick auf den Einsatz von „unspezifischen Kontrollgruppen“ im Rahmen von Psychotherapiestudien sollte mithilfe dieser Skala ein „Glaubwürdigkeits-Check“ möglich sein.

## **1.3.2. Itemsammlung**

Ausgehend von den im vorangehenden Kapitel beschriebenen a-priori-Skalen fand eine Sammlung und Generierung von Items statt. In diesem Stadium der Fragebogenentwicklung wurde v.a. Wert auf eine große inhaltliche Bandbreite gelegt, um möglichst viele Aspekte der jeweiligen Skala zu erfassen. Die einzelnen Items wurden unter Berücksichtigung bereits vorhandener, inhaltlich verwandter Instrumente (z.B. die Alliance-Fragebögen) und in Diskussionen mit Fachleuten (v.a. Psychologische Psychotherapeuten) entwickelt.

Die Items wurden in Aussageform formuliert, da der Fragebogen als Ratingskala konzipiert wurde. Formal wurde dabei auf die Einhaltung von Formulierungsregeln geachtet, um das Ausmaß von Missverständlichkeiten zu reduzieren (vgl. z.B. Mummendey, 1987). Dazu gehören z.B. die Vermeidung doppelter Verneinungen, die Verwendung kurzer Sätze und der Verzicht auf Fremdwörter.

Wie oben bereits angeführt wurden zu einer Verminderung möglicher Aquieszenz-Einflüsse sowohl positive, als auch negative Items formuliert, um Inhalt und Antwortstil besser voneinander trennen zu können (vgl. DeVellis, 2003), wobei maximal die Hälfte der Items pro Skala negativ gepolt wurden. Zunächst wurden insgesamt 121 Items entwickelt, die sich gleichmäßig auf die

einzelnen Skalen verteilt. D.h. es gab 15 Items pro Skala, mit Ausnahme der Skala „Offenheit“, die 16 Items umfasste.

## 2. Expertenbefragung

### 2.1. Ziel und Vorgehensweise

Nachdem eine erste Überprüfung der entwickelten Items hinsichtlich ihrer Formulierung und der genereller Passung durch zwei Fachleute (Diplom-Psychologinnen in Ausbildung zum Psych. Psychotherapeuten) erfolgt war, wurde zur weiteren Überprüfung der generierten Items und deren Zuordnung zu den vorgegebenen Skalen eine Befragung von Experten durchgeführt. Ziel war zum einen die Eliminierung derjenigen Items, die von den Experten nicht mit großer Übereinstimmung den entsprechenden Skalen zugeordnet werden konnten, und zum anderen auch eine Überprüfung der Diskriminierungsfähigkeit zwischen den Skalen.

Dieses Vorgehen sollte der Verbesserung der inhaltlichen Validität dienen, indem in einem ersten Schritt überprüft werden sollte, inwieweit die Items in angemessener Weise die entsprechenden Inhaltsbereiche (Skalen) repräsentieren (DeVellis, 2003; Fisseni, 1990). Daher musste es sich notwendigerweise bei den zur Einschätzung der Itempassung herangezogenen Personen um Sachverständige handeln, die über Fachwissen im Bereich der Skaleninhalte verfügten.

Entsprechend wurden als Experten Personen ausgewählt, die als Psychotherapeuten tätig waren. Die Experten bekamen eine Liste mit allen 121 Items präsentiert, die in zufälliger Reihenfolge angeordnet waren, mit der Angabe, dass es sich dabei um Aussagen über die Therapie, den Therapeuten, den Patienten und das Verhältnis zwischen beiden aus Sicht des Patienten handle. Die Items sollten nun von den Experten jeweils einer der acht vorgegebenen Skalen zugeordnet werden, deren Benennungen und Abkürzungen aufgeführt waren (1. „Glaubwürdigkeit/Nachvollziehbarkeit des Therapieangebotes“ (G), 2. „Glaube/Hoffnung/Veränderungserwartung des Patienten“ (VE), 3. „Empathie, Akzeptanz und Echtheit des Therapeuten (E), 4. „Offenheit des Patienten“ (O), 5. Kompetenz des Therapeuten (K), 6. „Bereitschaft des Patienten zur Mitarbeit“ (M), 7. „Emotionale Bindung zwischen Therapeut und Patient“ (B) und 8. „Zusammenarbeit zwischen Therapeut und Patient“ (Z). Zudem sollten sie einige persönliche Angaben machen zum Geschlecht, Alter, Dauer der Tätigkeit als Psychotherapeut/ Psychotherapeutin und Art der therapeutischen Ausbildung.



## 2.2. Stichprobenbeschreibung

An der Expertenbefragung nahmen 31 Personen teil, wobei es sich bei allen um Diplom-Psychologen handelte. Davon hatten 28 eine verhaltenstherapeutische Ausbildung begonnen oder abgeschlossen, jeweils eine Therapeutin hatte eine Systemische Therapie-Ausbildung bzw. eine Gesprächspsychotherapie-Ausbildung (eine Angabe fehlte). 7 Therapeuten/Therapeutinnen hatten mindestens eine Zusatzausbildung in Systemischer Therapie, Gesprächspsychotherapie, Hypnose, Gestalttherapie oder Psychoanalytischer Therapie. Der Altersdurchschnitt der Therapeuten lag bei 38,03 Jahren (SD=8,99) in einem Bereich zwischen 26 und 61 Jahren. Im Mittel lag die Berufserfahrung als Psychotherapeut bei 7,73 Jahren (SD= 9,16). Weitere Angaben zur Stichprobe finden sich in Tabelle 3.

Tabelle 3

Alter, Geschlecht und Berufserfahrung der Experten (N=31)

<b>Alter (Angabe in Jahren)</b>	<b>N</b>
26-35	15
36-45	11
46-55	2
56-65	2
Summe	30*
<b>Geschlecht</b>	
weiblich	22 (71%)
männlich	9 (29%)
Summe	31 (100%)
<b>Berufserfahrung als Therapeut (Angabe in Jahren; Monaten)</b>	<b>N</b>
Weniger als 1 Jahr	7
1 - 2;11	5
3 - 4;11	8
5-10	3
Über 10 Jahre	7
<b>Summe</b>	<b>30*</b>

\*Eine fehlende Angabe

### 2.3. Auswertung: Itemvorauswahl und Skalen-Neubestimmung

Als Kriterium für die Akzeptanz eines Items für die erste Version des Fragebogens wurde eine Übereinstimmungsrate von 75% angestrebt (vgl. Fuchs, 1998), d.h. mindestens 75% der Experten sollten das Item der entsprechenden Skala zuordnen. Es sollten dabei für jede Skala mindestens 10 Items ausgewählt werden. Dies war v.a. wichtig im Hinblick auf die später durchzuführende Faktorenanalyse, da man ab einer Anzahl von 10 oder mehr Items pro a priori-Faktor mit einer präzisen Faktorenlösung rechnen kann, sofern man eine Mindeststichprobengröße von N=150 erreicht (vgl. Guadagnoli & Velicer, 1988). In Tabelle 4 werden die Ergebnisse dargestellt, die sich gemäß dem 75%-Kriterium ergaben.

Tabelle 4  
Ergebnisse der Expertenbefragung

Skalenbezeichnung	Anzahl der Items gesamt	Anzahl der richtig zugeordneten Items
Glaubwürdigkeit/Nachvollziehbarkeit des Therapieangebotes (G)	15	8 (53%)
Veränderungserwartung des Patienten (VE)	15	12 (80%)
Empathie, Akzeptanz, Echtheit des Therapeuten (E)	15	4 (27%)
Offenheit des Patienten (O)	16	3 (19%)
Kompetenz des Therapeuten (K)	15	13 (87%)
Bereitschaft des Patienten zur Mitarbeit (M)	15	7 (47%)
Emotionale Bindung zwischen Therapeut und Patient (B)	15	4 (27%)
Zusammenarbeit zwischen Therapeut und Patient (Z)	15	6 (40%)
<b>Summe</b>	<b>121</b>	<b>57 (47%)</b>

Wie aus Tabelle 4 ersichtlich wird, erwies sich die Zuordnung der Items zu den entsprechenden Skalen als relativ schwierig - lediglich 57 der insgesamt 121 Items wurden entsprechend dem 75%-Kriterium richtig zugeordnet. Ein Blick auf die Einzelskalen zeigt, dass die Zuweisung der Items zu den Skalen „Veränderungserwartung des Patienten“ (VE) mit 12 Items über dem 75%-

Niveau und „Kompetenz des Therapeuten“ (K) mit 13 Items über dem 75%-Niveau wenig Schwierigkeiten bereitete.

Bei den anderen fünf Skalen konnten weniger als 10 Items von den Experten richtig zugewiesen werden, wobei sich offensichtlich besondere Probleme bei den Skalen „Empathie, Akzeptanz, Echtheit des Therapeuten“ (E), „Offenheit des Patienten“ (O) und „Emotionale Bindung zwischen Therapeut und Patient“ (B) zeigten, mit jeweils weniger als 5 richtig zugeordneten Items. Folgende Erklärungen kamen für dieses Ergebnis in Betracht: a) die Items waren nicht eindeutig formuliert, b) die Skalen waren in der Instruktion zu wenig klar definiert bzw. beschrieben und c) die Skalen selbst, d.h. die dahinterliegenden Konstrukte, haben untereinander z.T. große Überlappungsbereiche, die jeweils auf einen gemeinsamen Faktor hinweisen.

Gegen die Annahme, dass die Zuordnungsprobleme auf die mehrdeutige Formulierung zurückzuführen sind, sprach die Tatsache, dass vorab die Passung der Items von zwei Experten überprüft worden war, wobei allerdings die Zugehörigkeit der Items zu den Skalen angegeben war.

Die zweite Hypothese, wonach die Zuweisungsschwierigkeiten durch eine mangelnde Definition der Skalen bedingt waren, schien, zumindest für manche Skalen, eher in Betracht zu kommen. Während es bei Skalenbezeichnungen wie „Emotionale Bindung zwischen Therapeut und Patient“ oder „Empathie, Akzeptanz, Echtheit des Therapeuten“ vermutlich wenig Unklarheiten gegeben haben dürfte, könnten die Skalenbezeichnungen „Offenheit des Patienten“ oder „Bereitschaft zur Mitarbeit“ möglicherweise unterschiedlich ausgelegt werden. Die Überprüfung der einzelnen Bögen legte tatsächlich bei manchen Experten die Annahme nahe, dass sie unter „Offenheit“ eher „Ehrlichkeit“ i.S. eines offenen Ansprechens von Kritik verstanden hatten (z.B. bei dem Item „Manchmal erscheint mir mein Therapeut abweisend.“) als die Selbstöffnung des Patienten.

Die dritte Annahme postuliert, dass den Schwierigkeiten bei der Itemzuordnung die große Überlappung der Konstrukte selbst zugrunde lag. Einige Beispiele sollen diese Erklärungsmöglichkeit verdeutlichen: Das Item „Ich spüre, dass meinem Therapeuten etwas an mir liegt.“ gehört zur Skala „Emotionale Bindung zwischen Therapeut und Patient (B)“, da es ein Beziehungs-Item darstellt. Gleichzeitig ist darin indirekt aber auch die empathische, wertschätzende Haltung des Therapeuten angesprochen, ohne die tatsächlich eine emotionale Bindung nicht möglich ist.

Ähnlich verhält es sich mit dem Item „Manchmal fühle ich mich dem Therapeuten ausgeliefert“, das ebenfalls zur Skala „Emotionale Bindung zwischen Therapeut und Patient (B)“ gehört, aber auch einen Aspekt der Zusammenarbeit tangiert.

Möglicherweise spiegelte sich hier, zumindest in Bezug auf einen Teil der Skalen, eine Problematik wider, die aus der Forschung zur Dimensionalität der therapeutischen Allianz bzw. der Instrumente zur Erfassung der therapeutischen Allianz bekannt ist.

Im Folgenden soll dies kurz erläutert werden: Es wurden im Laufe der letzten Jahrzehnte verschiedene theoretische Konzeptualisierungen der therapeutischen Allianz entwickelt, die u.a. auch in Form von Messinstrumenten zur Erfassung der Allianz aus der Perspektive des Patienten operationalisiert wurden. Dazu gehören beispielsweise das „Working Alliance Inventory“ (WAI, Horvath & Greenberg, 1986) mit den drei Subskalen „Goals“, „Tasks“ and „Bond“, der „Helping Alliance Questionnaire“ (HAQ, Alexander & Luborsky, 1986) mit den zwei Subskalen „Helping Alliance“ und „Collaboration“ und die „California Psychotherapy Alliance Scale“ (CALPAS, Gaston, 1991) mit den vier Subskalen „Patient Working Capacity“, „Patient Commitment“, „Working Strategy Consensus“ und „Therapist Understanding and Involvement“. Obwohl die Subskalen der einzelnen Instrumente unterschiedliche theoretische Dimensionen abbilden sollten, ergaben einige empirische Überprüfungen hohe Korrelationen zwischen den Subskalen, was die Annahme der Mehrdimensionalität in Frage stellte und die Hypothese eines einzelnen „General-faktors“ der therapeutischen Allianz aufbrachte (Salvio et al., 1992).

Auch faktorenanalytische Überprüfungen lieferten wenig Evidenz für eine klare Subskalen-Diskrimination innerhalb einzelner Instrumente (z.B. Gaston & Marmar, 1994; Tracey & Kokotovic, 1989; Corbiere et al., 2006). Selbst eine faktorielle Überprüfung über drei Messinstrumente hinweg (WAI, HAQ, CALPAS) erbrachte einen starken Hauptfaktor für die therapeutische Allianz, der 37% der Varianz erklärte und daneben fünf weitere nachgeordnete Faktoren, die je maximal 5% zur Varianzaufklärung beitrugen (Hatcher & Barends, 1996). Obwohl die Existenz eines General-Faktors der therapeutischen Allianz nicht generell anerkannt ist (vgl. z.B. Andrusyna et al., 2001), wird doch die Problematik der Differenzierung zwischen den Subskalen aufgrund großer inhaltlicher und statistischer Überlappungsbereiche deutlich.

In Bezug auf die Expertenbefragung in der vorliegenden Untersuchung könnte das bedeuten, dass sich in den Schwierigkeiten der Experten, die Items den vorgegebenen Skalen zuzuordnen, eben diese Differenzierungsproblematik darstellte. Dies schien v.a. für die Skalen „Empathie, Akzeptanz und Echtheit des Therapeuten“ (E), „Emotionale Bindung zwischen Therapeut und Patient“ (B) und „Zusammenarbeit zwischen Therapeut und Patient“ (Z) zu gelten, da diese konzeptionell den „klassischen“ Allianzskalen „bond“ und „agreement on task and goals“ (vgl. Agnew-Davies, 1998) am nächsten stehen.

Aufgrund dieser Überlegungen in Zusammenhang mit den Ergebnissen der Expertenbefragung wurden diese drei Skalen („Empathie, Akzeptanz und Echtheit des Therapeuten“ (E), „Emotionale Bindung zwischen Therapeut und Patient“ (B) und „Zusammenarbeit zwischen Therapeut und Patient“ (Z)) für die weitere Untersuchung zu einer gemeinsamen Skala mit der Bezeichnung „Therapeutische Beziehung“ (TB) zusammengefasst, unter der Annahme, dass offenbar das Gemeinsame der therapeutischen Beziehung stärker ins Gewicht fällt als die Unterschiede, die sich durch die verschiedenen Facetten der ursprünglichen Subskalen ergeben („Sich-Vom-Therapeut-Verstanden-Fühlen“ vs. „Sich-Gegenseitig-Verstehen“ vs. „An-Gemeinsamen-Zielen-Arbeiten“).

Die Skala „Offenheit des Patienten“ (O), die als Fähigkeit des Patienten konzipiert worden war, in der Therapie persönlich Bedeutsames preiszugeben und auch Unangenehmes zulassen zu können, ist als separate Komponente im Vergleich zu den zentralen Komponenten der therapeutischen Beziehung („bond“ und „agreement on task and goals“) aufzufassen (vgl. Agnew-Davies, 1998). Dennoch zeigten sich auch hier, ebenso wie bei der Skala „Bereitschaft des Patienten zur Mitarbeit (M)“ Zuordnungsprobleme. Wie oben bereits erwähnt, scheinen diese, zumindest was die Skala „Offenheit des Patienten“ (O) betrifft, teilweise auf eine mangelnde Skalendefinition zurückzugehen.

Allerdings lässt sich auch hier, analog zu den Überlegungen zur therapeutischen Beziehung, die Annahme formulieren, dass den Skalen „Offenheit des Patienten“ (O) und „Bereitschaft des Patienten zur Mitarbeit (M)“ eine gemeinsame Grundstruktur zugrunde liegt, bzw. genauer, dass die Skala „Offenheit des Patienten“ (O) möglicherweise als ein Teilbereich des Konzeptes „Bereitschaft zur Mitarbeit“ anzusehen ist. Auch hier ein Beispiel: das Item „Ich gehe das Risiko ein, in der Therapie auch unangenehme Gefühle zu erleben“ ist als Item der Skala „Offenheit des Patienten“ (O) konzipiert, stellt gleichzeitig aber notwendigerweise auch einen Teilaspekt der Mitar-

beit dar. Ausgehend von diesem Gedanken wurden diese beiden Skalen ebenfalls für die weitere Untersuchung zusammengefasst unter der Bezeichnung „Bereitschaft des Patienten sich auf die Therapie einzulassen und mitzuarbeiten“ (MO).

Die Skala „Glaubwürdigkeit/Nachvollziehbarkeit des Therapieangebotes“ (G) wurde trotz unbefriedigender Zuordnungsergebnisse in ihrer ursprünglichen Form beibehalten, da in diesem Fall inhaltliche Überlappungen am ehesten mit der Skala „Kompetenz des Therapeuten“ (K) bestehen, die selbst aber gute Zuordnungsergebnisse aufwies. Weitere Analysen sollten daher zeigen, ob sich eine Beibehaltung der Skala „Glaubwürdigkeit/Nachvollziehbarkeit des Therapieangebotes“ (G) als sinnvoll erweisen würde.

Auf eine Neuformulierung der Items zur Erhöhung der Prägnanz bzw. Eindeutigkeit der Skalenzugehörigkeit, wie es die erstgenannte Erklärungsmöglichkeit nahe legen würde, wurde verzichtet. Ein Grund dafür war die Tatsache, dass den oben genannten Überlegungen bezüglich der Überlappungen der Konstrukte mehr Bedeutung zugeschrieben wurde für die Erklärung der schlechten Zuordnungsergebnisse der Expertenbefragung. Zudem hätte eine „Prägnanzerhöhung“ notwendigerweise zu einer Reduktion der inhaltlichen Bandbreite geführt, da die Items sich inhaltlich ähnlicher geworden wären, womit auch immer eine Verringerung der Varianz eingetreten wäre.

Um die Annahme der Überlappungsbereiche zwischen den Skalen „Empathie/ Akzeptanz/ Echtheit des Therapeuten“ (E), „Emotionale Bindung zwischen Therapeut und Patient“ (B) und „Zusammenarbeit zwischen Therapeut und Patient“ (Z), sowie den beiden Skalen „Offenheit des Patienten“ (O) und „Bereitschaft des Patienten zur Mitarbeit (M)“ zu überprüfen, wurde eine weitere Analyse der Daten aus der Expertenbefragung vorgenommen. Dabei wurden die Häufigkeiten der richtigen Zuordnungen errechnet für die zusammengefassten Skalen „Therapeutische Beziehung“ (TB) bzw. „Bereitschaft des Patienten, sich auf die Therapie einzulassen und mitzuarbeiten“ (MO), d.h. eine Zuordnung wurde als zutreffend gewertet, wenn das betreffende Item einer der Skalen zugeschrieben wurde, die unter zur Gesamtskala „TB“ respektive „MO“ zusammengefasst worden waren.

## 2. Expertenbefragung

---

Sofern die Annahme der Überlappungsbereiche zwischen den Einzelskalen zutreffen sollte, sollte die Anzahl der richtig zugeordneten Items (gemäß dem 75%-Kriterium) für die zusammengefassten Skalen „TB“ und „MO“ jeweils deutlich über der Summe der richtig zugeordneten Items aus den Einzelskalen der ersten Auswertung liegen. Dies nämlich würde nahe legen, dass die Experten die entsprechenden Items zwar unterschiedlichen Skalen zugeordnet haben (Ergebnis der 1. Auswertung, vgl. Tabelle 4), diese Zuordnungen aber nicht über alle Skalen hinweg variierten, sondern im Wesentlichen innerhalb derjenigen Skalen erfolgten, zwischen denen, wie oben erläutert, große Überlappungsbereiche angenommen wurden.

In Tabelle 5 werden die Ergebnisse dargestellt, die sich für die Zuordnungen nach der Zusammenfassung der Skalen „Empathie, Akzeptanz und Echtheit des Therapeuten“ (E), „Emotionale Bindung zwischen Therapeut und Patient“ (B) und „Zusammenarbeit zwischen Therapeut und Patient“ (Z) zur Gesamtskala „Therapeutische Beziehung“ (TB) einerseits und der Skalen „Offenheit des Patienten“ (O) und „Bereitschaft des Patienten zur Mitarbeit“ (M) zur Gesamtskala „Bereitschaft des Patienten, sich auf die Therapie einzulassen und mitzuarbeiten“ (MO) andererseits gemäß dem 75%-Kriteriums ergaben.

Tabelle 5

Ergebnisse der Expertenbefragung nach Zusammenfassung der Skalen E, B und Z zur Gesamtskala „Therapeutische Beziehung“ (TB) und der Skalen O und M zur Skala „Bereitschaft des Patienten, sich auf die Therapie einzulassen und mitzuarbeiten“ (MO)

Skalenbezeichnung	Anzahl der Items gesamt	Anzahl der richtig zugeordneten Items
I. Glaubwürdigkeit/Nachvollziehbarkeit des Therapieangebotes (G)	15	8 (53%)
II. Veränderungserwartung des Patienten (VE)	15	12 (80%)
III. Therapeutische Beziehung (TB) (aus den Skalen Empathie, Akzeptanz und Echtheit des Therapeuten (E), Emotionale Bindung zwischen Therapeut und Patient (B) und Zusammenarbeit zwischen Therapeut und Patient (Z))	45	35 (78%)
IV. Bereitschaft des Patienten, sich auf die Therapie einzulassen und mitzuarbeiten (MO) (aus den Skalen Offenheit des Patienten (O) und Bereitschaft des Patienten zur Mitarbeit (M))	31	15 (48%)
V. Kompetenz des Therapeuten (K)	15	13 (87%)
<b>Summe</b>	<b>121</b>	<b>83 (67%)</b>

Wie aus Tabelle 5 ersichtlich, ergab sich für die neue Skala „Therapeutische Beziehung“ (TB) die Anzahl von 35 richtig zugeordneten Items (gemäß dem 75%-Kriterium) aus einem Pool von 45 Items, was einem Prozentsatz von 78% entspricht. Vergleicht man dieses Ergebnis mit den Werten für die Einzelskalen aus der 1. Auswertung (vgl. Tabelle 4) wird deutlich, dass die Zusammenlegung der Skalen einen deutlichen Zugewinn erbracht hat: in der 1. Auswertung konnten von den Einzelskalen „Empathie, Akzeptanz und Echtheit des Therapeuten“ (E), „Emotionale Bindung zwischen Therapeut und Patient“ (B) und „Zusammenarbeit zwischen Therapeut und Patient“ (Z) zusammengenommen nur 14 der 45 Items (31%) richtig zugeordnet werden (gemäß dem 75%-Kriterium). Dementsprechend konnte durch die Zusammenlegung der Skalen ein deutlicher Zugewinn von 21 Items erzielt werden (35 richtig zugeordnete Items in der 2. Auswertung gegenüber 14 Items in der 1. Auswertung). Dieses Ergebnis spricht für das Vorhandensein von Überlappungen zwischen den ursprünglichen Einzelskalen, weswegen die Zusammenlegung der Skalen zur Gesamtskala „Therapeutische Beziehung“ (TB) beibehalten wurde.



Hinsichtlich der neuen Skala „Bereitschaft des Patienten, sich auf die Therapie einzulassen und mitzuarbeiten“ (MO) erbrachte die 2. Auswertung die Anzahl von 15 richtig zugeordneten Items (gemäß dem 75%-Kriterium) aus einem Pool von 31 Items, was einem Prozentsatz von 48% entspricht. Ein Vergleich mit den Ergebnissen der 1. Auswertung für die Einzelskalen (vgl. Tabelle 4) zeigt, dass durch die Zusammenlegung der Skalen „Offenheit des Patienten“ (O) und „Bereitschaft des Patienten zur Mitarbeit“ (M) 5 Items hinzugewonnen werden konnten (15 richtig zugeordnete Items in der 2. Auswertung gegenüber 10 Items in der 1. Auswertung).

Im Vergleich zur neuen Skala „Therapeutische Beziehung“ (TB) fiel der Zugewinn durch die Zusammenlegung deutlich geringer aus. Ausgehend von der oben skizzierten Annahme, dass die ursprüngliche Skala „Offenheit des Patienten“ (O) als Teilkonzept der Skala „Bereitschaft des Patienten zur Mitarbeit“ (M) verstanden werden könnte, bedeutet dies, dass vermutlich ein Teil der ursprünglich der Skala „Offenheit des Patienten“ (O) zugehörigen Items von den Experten der Skala „Bereitschaft des Patienten zur Mitarbeit“ (M) zugerechnet wurde, aber nicht umgekehrt. Dafür spricht auch ein Vergleich der Prozentsätze der richtig zugeordneten Items zwischen den ursprünglichen Skalen und der Gesamtskala: während die Zusammenlegung im Vergleich mit der Einzelskala für die Skala „Bereitschaft des Patienten zur Mitarbeit“ (M) nur eine Veränderung um einen Prozentpunkt brachte (48% vs. 47%, vgl. Tabellen 5 und 4), ergab sich für die Gesamtskala im Vergleich zur Skala „Offenheit des Patienten“ (O) ein Zuwachs um 29% (48% vs. 19%, vgl. Tabellen 5 und 4).

Da die zusammgelegten Skalen „Therapeutische Beziehung“ (TB) und die Skala „Bereitschaft des Patienten, sich auf die Therapie einzulassen und mitzuarbeiten“ (MO) gegenüber den ursprünglichen Skalen zufriedenstellende bis deutliche Verbesserungen in der Itemzuordnung aufwiesen und die Zusammenlegungen inhaltlich plausibel waren, wurden die neuen Skalen für die weitere Untersuchung in dieser Form beibehalten.

Um diese Überlegungen auch statistisch mithilfe von Faktorenanalysen überprüfen zu können und eine voreilige Itemselektion zu vermeiden, wurden für die zusammgelegten Skalen alle Items aus den ursprünglichen Skalen beibehalten, die gemäß dem 75%-Kriterium ausgewählt worden waren.

## 2. Expertenbefragung

---

Zudem wurde das Einschlusskriterium auf eine 70%-Übereinstimmung zwischen den Experten gesenkt um für die nachfolgende Untersuchung zumindest annähernd die 10 gewünschten Items pro a priori-Faktor zu gewinnen. Dafür erschien es geeigneter, diejenigen Items mit relativ hoher Übereinstimmung in die Fragebogen-Vorversion aufzunehmen, als neue Items zu generieren. Entsprechend dieser Vorgehensweise ergaben sich für die Fragebogen-Vorversion folgende Skalen und Itemzahlen (siehe Tabelle 6)

Tabelle 6

Skalenbezeichnungen und Itemzahlen für die Fragebogen-Vorversion nach Auswertung der Expertenbefragung (70%-Übereinstimmung)

<b>Skalenbezeichnung</b>	<b>Anzahl der Items gesamt</b>	<b>Anzahl der richtig zugeordneten Items</b>
I. Glaubwürdigkeit/ Nachvollziehbarkeit des Therapieangebotes (G)	15	9 (60%)
II. Veränderungserwartung des Patienten (VE)	15	13 (87%)
III. Therapeutische Beziehung (TB) (aus den Skalen Empathie, Akzeptanz und Echtheit des Therapeuten (E), Emotionale Bindung zwischen Therapeut und Patient (B) und Zusammenarbeit zwischen Therapeut und Patient (Z))	45	35 (78%)
IV. Bereitschaft des Patienten, sich auf die Therapie einzulassen und mitzuarbeiten (MO) (aus den Skalen Offenheit des Patienten (O) und Bereitschaft des Patienten zur Mitarbeit (M))	31	17 (55%)
V. Kompetenz des Therapeuten (K)	15	14 (93%)
<b>Summe</b>	<b>121</b>	<b>88 (73%)</b>

Eine Liste der ausgewählten Items befindet sich unter Angabe der Expertenübereinstimmung (in Prozent) im Anhang.

Da die Polung der Items als Auswahlkriterium gegenüber der Übereinstimmungsrate der Experten nachgestellt wurde, ergaben sich für die Skalen unterschiedliche Anteile negativ gepolter Items. Von den 88 Items waren insgesamt 30 Items (34%) negativ gepolt, wobei sich diese ungleich über die Skalen verteilten. Der Anteil negativ gepolter Items variierte zwischen 21% für

die Skala V. „Kompetenz des Therapeuten“ (K) und 44% für die Skala I. „Glaubwürdigkeit/Nachvollziehbarkeit des Therapieangebotes“ (G).

### **3. Überarbeitungen und Erstellung der ersten Version des Fragebogens**

#### **3.1. Überarbeitungen der Formulierungen**

Die 88 Items des Fragebogens wurden im Anschluss an die Expertenbefragung erneut hinsichtlich ihrer Formulierungen überprüft. Zunächst wurden von einer Expertin (Psychotherapeutin in Ausbildung) die Items noch einmal auf Doppeldeutigkeit und Missverständlichkeit hin durchgesehen. Danach schätzten vier Experten (Psychotherapeutinnen in Ausbildung) für jedes Item die Wahrscheinlichkeit einer positiven Extremantwort, d.h. maximale Zustimmung bei positiv formulierten Items und maximale Ablehnung bei negativ formulierten Items, ein. Diejenigen Items, bei denen die Experten im Vorfeld von einer sehr hohen Zustimmungsrates ausgingen, wurden umformuliert. Dabei wurde darauf geachtet, dass der Inhalt unverändert blieb, aber die Intensität der Aussagen abgeschwächt wurde.

#### **3.2. Erstellen der 1. Version des Fragebogens**

Nachdem die als Aussagen formulierten Items noch einmal überprüft und überarbeitet worden waren, konnte nun die 1. Fragebogenversion erstellt werden. Dafür mussten verschiedene Entscheidungen hinsichtlich des Antwortformats, des Bezugszeitraumes der Aussagen und der Itemreihung getroffen werden, auf die im folgenden kurz eingegangen werden soll.

##### **3.2.1. Antwortformat**

Der Fragebogen wurde als Ratingskala konzipiert. Bei diesem Antwortformat, das häufig im Rahmen von Persönlichkeits- oder Einstellungstests verwendet wird, wird der Grad der Zustimmung- bzw. der Ablehnung zu einer vorgelegten Aussage als Indikator für die Ausprägung des betreffenden Merkmals herangezogen (vgl. Moosbrugger & Kelava, 2007). Bei der Wahl des Antwortformats kann entweder eine bipolare oder eine unipolare Skala gewählt werden. Bipolare Skalen haben einen Zustimmungsbereich von einem positiven Pol (=Zustimmung) über einen indifferenten Bereich bis zu einem negativen Pol. Bei unipolaren Skalen gibt es einen Punkt, der das geringste Ausmaß an Zustimmung bzw. Ablehnung markiert, sowie einen

Pol, der die stärkste Zustimmung bzw. Ablehnung angibt. Zudem ist eine Entscheidung hinsichtlich der Anzahl und Bezeichnung der Skalenstufen zu treffen.

Für den konzipierten Fragebogen wurde eine siebenstufige, bipolare und numerische Ratingskala gewählt. Auf jeder Seite des Fragebogens wurden zudem die verbalen Entsprechungen der numerischen Skalenstufen in Form von Beurteilungen (=Abstufungen des Zutreffens) angegeben (① = „stimmt überhaupt nicht“; ② = „stimmt nicht“; ③ = „stimmt eher nicht“ ; ④ = „stimmt teils-teils“; ⑤ = „stimmt eher“; ⑥ = „stimmt“; ⑦ = „stimmt voll und ganz“). Die Skala wurde bipolar konstruiert, da es inhaltlich sinnvoll erschien, sowohl zustimmende als auch ablehnende Einschätzungen zuzulassen.

Bei der Bezeichnung der Skalenstufen wurde eine Kombination aus numerischer und verbaler Skala gewählt, um die Vorteile beider Optionen zu nutzen: die Zahlenzuordnung soll den Eindruck der Gleichheit zwischen den Stufenabständen vermitteln (vgl. Moosbrugger & Kelava, 2007). Zudem ist dadurch die Übersichtlichkeit der Darstellung gewährleistet. Die verbalen Bezeichnungen wiederum haben den Vorzug, dass die Interpretation der Skalenpunkte weniger Spielraum zulässt. Zudem sind die Testpersonen zufriedener, wenn nicht nur die Extremwerte verbale Beschreibungen aufweisen (Dickinson & Zellinger, 1980, zit. nach Moosbrugger & Kelava, 2007). Es wurde die Anzahl von sieben, anstatt der häufig verwendeten fünf Stufen gewählt, da generell – bis zu einer gewissen Obergrenze – Validität und Reliabilität steigen, wenn man mehr Antwortkategorien zulässt (Matell & Jacoby, 1971, zit. nach Bühner, 2006). Eine noch differenziertere Abstufung erschien inhaltlich nicht sinnvoll. Zudem würde dies vermutlich die Probanden in ihrer Unterscheidungsfähigkeit überfordern.

Ein schwierigeres Problem stellte die Frage dar, ob eine Mittelkategorie verwendet werden sollte. Mehrfach konnte nämlich empirisch nachgewiesen werden, dass diese Kategorie häufig als Ausweichoption von Probanden genutzt wird, die sich nicht festlegen können oder wollen. Gleichzeitig scheinen motivierte Probanden diese Kategorie zu vermeiden (Bortz & Döring, 2002). Auf der anderen Seite ist es schwierig, den Probanden die mittlere Kategorie vorzuenthalten, da sie so zu einer Entscheidung gezwungen werden (Bühner, 2006). Dieser zweite Aspekt wurde schließlich als bedeutsamer angesehen, weswegen die mittlere Kategorie („stimmt teils-teils“) verwendet wurde.

### 3.2.2. Festlegung des Bezugszeitraumes

Bei der Wahl des Zeitraumes, auf den sich die Urteile der Patienten beziehen sollten, standen im Wesentlichen zwei Optionen zur Wahl. Eine Möglichkeit bestand darin, die Patienten um ein Gesamturteil über alle bis zum Erhebungsdatum durchgeführten Therapiesitzungen zu bitten. Die andere Möglichkeit stellte die Beurteilung der letzten, also der am kürzesten zurückliegenden Therapiestunde dar, womit die wiederholte Darbietung des Fragebogens verbunden war. Der Vorteil einer Gesamtbeurteilung war in erster Linie in der Ökonomie und Einfachheit der Anwendung zu sehen.

Zudem wird dabei generell der Therapieprozess nicht durch wiederholte Messungen beeinflusst (vgl. Bühner, 1998). Auf der anderen Seite bietet der auf eine Therapiestunde begrenzte Bezugszeitraum den Vorzug, dass eine Erfassung der möglichen Veränderungen in den Merkmalsausprägungen über den Therapieverlauf möglich ist. Dies erschien auf jeden Fall für die Skala „Therapeutische Beziehung“ von Bedeutung, da verschiedene Studien für diesen Bereich Schwankungen über den Zeitraum der Therapie nachweisen konnten, weswegen eine häufigere Messung empfohlen wird (Stiles et al., 1998; Agnew-Davies, 1998).

Auch bei den anderen Skalen erschienen Schwankungen über den Therapieverlauf plausibel, weshalb die Möglichkeit wiederholter Messungen als Vorteil erachtet wurde. Ein weiteres Argument für die Verwendung der letzten Therapiestunde bestand darin, dass damit die Gefahr von Verzerrungen der Urteile durch Gedächtniseffekte als geringer erachtet wurden, da die Zeitabstände zwischen Ereignis und Beurteilung in diesem Fall kürzer wären und nicht viele Erinnerungen von den Probanden „gemittelt“ werden müssten.

Als Bezugszeitraum wurde daher die letzte Therapiestunde gewählt. Um allerdings mit den Items der a-priori-Skala „Veränderungserwartung des Patienten“ möglichst nicht die direkte Nachwirkung der letzten Therapiestunde zu erfassen (=Stunden-Outcome), wurde für diese Items der Bezugszeitraum auf die Woche vor der letzten Therapiestunde festgelegt. Damit ergab sich für den Fragebogen eine Unterteilung in zwei Abschnitte: die Items des ersten Teils erforderten eine Einschätzung der letzten Therapiestunde, während sich die Items im zweiten Teil auf die Woche vor dieser Sitzung bezogen.

### **3.2.3. Itemreihung und Ergänzung**

Die Reihenfolge der Items wurde zufällig gewählt (vgl. Bühner, 2006). Dazu wurden alle Itemnummern auf kleine Zettel geschrieben und verdeckt gemischt. Die Abfolge der Items wurde dann durch wiederholtes Ziehen eines Zettels aus dem Pool festgelegt. Allerdings ergab sich durch die Einteilung in die beiden Abschnitte eine Einschränkung der Zufälligkeit. Zudem wurde darauf geachtet, dass der Fragebogen nicht mit einem umgepolten, also negativ formulierten Item beginnen sollte. Und schließlich wurde die Reihenfolge umgestellt, sofern zwei Items mit sehr ähnlichem Inhalt aufeinander folgten.

Der Fragebogen wurde schließlich noch um zwei Zusatzitems ergänzt. Das erste Item erfragte die Zufriedenheit mit der letzten Therapiestunde, während das zweite Item zum Thema hatte, wie typisch die letzte Stunde im Vergleich mit den anderen bisherigen Therapiesitzungen war. Der vollständige Fragebogen in seiner 1. Version befindet sich im Anhang (incl. Instruktionen).

## **4. Fragebogenuntersuchung**

Zur statistischen bzw. testtheoretischen Überprüfung des Fragebogens wurden Daten einer Analysestichprobe ausgewertet. Diese wurden durch eine Erhebung an Psychotherapiepatienten gewonnen.

Im Folgenden werden zunächst das Vorgehen bei der Untersuchung und die Stichprobenmerkmale beschrieben. Im zweiten Abschnitt werden die Auswertungsschritte und die Ergebnisse dargestellt.

### **4.1. Stichprobe**

#### **4.1.1. Kriterien für die Stichprobengewinnung**

Für die Wahl der Analysestichprobe sind bei der Entwicklung eines Fragebogens zwei Aspekte von besonderer Bedeutung, die miteinander in engem Zusammenhang stehen: die Größe und die Repräsentativität der Stichprobe.

Die Stichprobe muss ausreichend groß sein, um die Varianz zu eliminieren, die nicht auf die Items, sondern auf die Versuchspersonen zurückgeht. Bei einer zu geringen Anzahl an Personen besteht die Gefahr, dass die Korrelationen zwischen den Items zu einem nicht unerheblichen Teil durch zufällige Faktoren beeinflusst werden (DeVellis, 2003). Hinsichtlich der Frage, ab wie vie-

len Versuchspersonen eine Stichprobe groß genug ist, gibt es allerdings keinen Konsens. Nunnally (1978, zitiert nach DeVellis, 2003) gibt als Richtwert eine Anzahl von 300 Personen an. DeVellis (2003) bemerkt hierzu allerdings, dass die Erfahrung gezeigt habe, dass auch mit kleineren Stichproben erfolgreich Skalen entwickelt worden seien. Gemäß der Studie von Guadagnoli & Velicer (1988) genügt für den Erhalt einer stabilen Faktorenlösung mit guter Annäherung an die Populations-Parameter eine Stichprobe von 150 Beobachtungen, sofern ausreichend Variablen ( $\geq 10$ ) einen a priori Faktor determinieren.

Ausgehend von diesen Angaben wurde für die vorliegende Untersuchung als Untergrenze ein N von 200 angesetzt, mit dem Zielwert von 250 Versuchspersonen.

Auch hinsichtlich der Repräsentativität ist die Stichprobengröße von Bedeutung, da die Wahrscheinlichkeit der Nicht-Repräsentativität bei einer kleinen Stichprobe deutlich höher liegt. Man kann zwei Arten von Nicht-Repräsentativität unterscheiden (DeVellis, 2003): die erste bezieht sich auf mögliche Unterschiede in der Ausprägungsstärke der zu erfassenden Merkmale in der Untersuchungsstichprobe im Vergleich zur Population. Z.B. könnte eine stichprobenbedingte Einschränkung des Wertebereiches dazu führen, dass sich die Mittelwerte der Skalen, die man für die Untersuchungsstichprobe erhält, bedeutsam von denen unterscheiden, die man für die Population erwarten würde. Allerdings disqualifizieren nicht-repräsentative Mittelwerte die Stichprobe nicht unbedingt für den Zweck der Skalenentwicklung, da die Zusammenhänge zwischen den Items davon unbeeinflusst bleiben.

Die zweite Art der Nicht-Repräsentativität ist problematischer, da sie sich auf die Beziehung der Items untereinander auswirkt. Dieser Typus von Nicht-Repräsentativität liegt vor, wenn sich die Stichprobe qualitativ von der Population unterscheidet. Wenn sich nämlich die zu erfassenden Merkmale in ihrer Relation zueinander in der Stichprobe anders verhalten als in der Population, sind die Ergebnisse der statistischen Auswertungen (z.B. die Faktorenstruktur) nicht auf die Population zu übertragen. Konkret könnte dies z.B. bedeuten, dass die Items für die Probanden der Stichprobe eine andere Bedeutung haben, als für die Population.

Für die vorliegende Untersuchung bedeutete dies in Hinblick auf den ersten Typus von Nicht-Repräsentativität, dass sich die Personen, die den Fragebogen auszufüllen hatten, möglichst untereinander im Ausprägungsgrad der Merkmale unterscheiden sollten. In Anbetracht des Geltungs-

#### 4. Fragebogenuntersuchung

---

bereiches des Fragebogens erschien es daher sinnvoll, den Fragebogen an Psychotherapiepatienten auszugeben, die sich zu dieser Zeit in Therapie befanden. Da bei einer solchen Fragebogenuntersuchung die Rekrutierung der Patienten zur Wahrung der Anonymität durch den betreffenden Therapeuten erfolgen musste, konnte eine bewusste oder unbewusste Selektion durch den Therapeuten nicht ganz ausgeschlossen werden. Zur Erhöhung der Merkmalsvarianz sollte der Fragebogen an alle Patienten unabhängig von der Sitzungsanzahl ausgegeben werden, sofern sie den Therapeuten mindesten dreimal gesehen hatten. Hintergrund war der Gedanke, dass Patienten, die eine hohe Sitzungsanzahl aufweisen, im Schnitt vermutlich eine höhere Ausprägung der in Frage stehenden Merkmale haben, da diejenigen mit geringerer Merkmalsausprägung mit einer höheren Wahrscheinlichkeit die Therapie vorzeitig abbrechen.

Einen schwer zu kontrollierenden Faktor stellte allerdings die Selbstselektion der Patienten dar, da die Teilnahme an der Untersuchung selbstverständlich freiwillig war. Eine Einschränkung der Streuungsbreite der Merkmalsausprägung stellte die Möglichkeit dar, dass diejenigen Patienten, die eine höhere Ausprägung der Merkmale aufwiesen, auch eher zur Teilnahme bereit sein könnten. Um diesen Effekt abzumildern, wurde in der Instruktion auf die Anonymität verwiesen und auf die Tatsache, dass der Therapeut den Bogen nicht zu lesen bekäme, was auch der Tendenz zur Abgabe sozial erwünschter Antworten entgegenwirken sollte.

Die zweite Art von Nicht-Repräsentativität bezieht sich, wie oben erläutert, auf die qualitative Repräsentativität. Da der Fragebogen künftig in der Psychotherapieforschung, unabhängig vom Störungsbild der Patienten und der angewandten Therapiemethode, eingesetzt werden soll, sollte die Untersuchungsstichprobe möglichst ein kleines Abbild der Population der sich in Psychotherapie befindenden Patienten darstellen, wie man sie in der Praxis findet. Berücksichtigt wurden dabei soziodemographische Variablen, die Häufigkeit der Störungsbilder und die Therapiemethoden bzw. -schulen. Da ein Großteil der Patienten gesetzlich krankenversichert ist und meist eine von den Krankenkassen finanzierte Therapie in Anspruch genommen wird, beschränkte sich die Rekrutierung der Stichprobe auf Patienten, die nach den Richtlinienverfahren behandelt wurden (Verhaltenstherapie, tiefenpsychologisch fundierte Psychotherapie, analytische Psychotherapie).



### 4.1.2. Vorgehen

Die Rekrutierung der Stichprobe erfolgte in zwei Phasen. Im Juli 2005 wurden Anfragen an Ambulanzen und therapeutische Ausbildungsinstitute verschiedener Therapierichtungen (Verhaltenstherapie, analytische Therapie, tiefenpsychologisch fundierte Therapie) bezüglich einer möglichen Teilnahme an der Untersuchung versandt, mit dem Angebot, das Projekt in einer der Ambulanzsitzungen vorzustellen. Von den Instituten und Ambulanzen, zu denen Kontakt aufgenommen worden war, erklärten sich 8 zur Teilnahme bereit (6 Institute der Christoph-Dornier-Stiftung, die Hochschulambulanz der Universität Tübingen und die Private Tübinger Akademie für Verhaltenstherapie), alle mit verhaltenstherapeutischer Orientierung. An den beteiligten Instituten wurde jeweils eine Liste ausgeteilt, in die sich die Therapeuten unter Angabe der Anzahl der von ihnen behandelten Patienten eintragen konnten, sofern sie bereit waren an der Studie teilzunehmen.

In der zweiten Phase wurden im September 2005 Briefe an alle niedergelassenen Psychologischen Psychotherapeutinnen und -therapeuten im Landkreis Tübingen verschickt (N=88; davon 37 Therapeuten/Therapeutinnen mit VT-Ausbildung, 37 mit analytisch/ tiefenpsychologischer Ausbildung, 10 mit tiefenpsychologischer Ausbildung und 4 mit tiefenpsychologischer und verhaltenstherapeutischer Ausbildung). Dem Anschreiben der Untersuchungsleiterin lag ein Schreiben des Institutsleiters mit der Bitte um Unterstützung des Projekts, sowie eine frankierte und adressierte Postkarte bei, welche die Therapeuten unter Angabe der Patientenanzahl zurücksenden sollten, wenn sie sich an der Untersuchung beteiligen wollten. Zudem wurden zwei verhaltenstherapeutisch orientierte Therapeuten in Würzburg, sowie je ein psychoanalytisch orientierter Therapeut und ein tiefenpsychologisch arbeitender Therapeut in Berlin angeschrieben, zu denen institutsintern Kontakte bestanden. Von den niedergelassenen Therapeuten erklärten sich 14 zur Teilnahme an der Untersuchung bereit, davon 7 mit verhaltenstherapeutischer Orientierung, vier mit tiefenpsychologischer Ausrichtung, zwei mit tiefenpsychologisch/analytischer Ausbildung und ein Psychoanalytiker.

Entsprechend den Angaben der Therapeuten hinsichtlich der Anzahl der von ihnen behandelten Patienten wurde für jeden teilnehmenden Therapeuten ein „Paket“ zusammengestellt (siehe Anhang), bestehend aus einem Anschreiben, einem Bogen „Angaben des Therapeuten“, der entsprechenden Anzahl an Fragebögen mit Umschlag und einem an die Untersuchungsleiterin adressierten und frankierten Rückumschlag. Das Anschreiben enthielt eine kurze Darstellung des Inhalts

#### 4. Fragebogenuntersuchung

---

und der Zielsetzung der Studie, sowie Hinweise zum Ablauf der Fragebogenuntersuchung. Auf dem Bogen „Angaben des Therapeuten“ befanden sich, neben der Instruktion für den Therapeuten, Felder, in denen er Angaben zur eigenen Person machen sollte (Alter, Geschlecht, Berufs- und Therapieausbildung, Berufserfahrung und Angaben zum Ort, an dem die Therapien durchgeführt wurden). Die Therapeuten wurden gebeten, die Patienten kurz über den Inhalt der Studie zu informieren und dabei auf die Anonymität der Datenerhebung hinzuweisen.

Wenn sich die Patienten zur Teilnahme bereit erklärten, bekamen sie den Fragebogen mit nach Hause, mit der Bitte, ihn zur nächsten Sitzung ausgefüllt im verschlossenen Umschlag wieder mitzubringen. Auf der ersten Seite des insgesamt vierseitigen Fragebogens, der den Arbeitstitel „Fragebogen zur Erfassung des Therapieerlebens“ trug, befand sich eine Instruktion für die Patienten, in der auch noch einmal auf die Wahrung der Anonymität hingewiesen wurde, und auf die Tatsache, dass der Therapeut den Fragebogen nicht zu sehen bekäme. Ebenfalls auf der ersten Seite befanden sich die Felder für die soziodemographischen Angaben des Patienten (Alter, Geschlecht, Schulabschluss, Partnerschaft) und die Angaben zur Störungsdauer, Belastungsstärke, Vorbehandlung und Zeitdauer seit der letzten Therapiestunde.

Auf den Umschlag sollten vom Therapeuten die Diagnosen (nach ICD-10) eingetragen und Angaben zur Sitzungsfrequenz, der Anzahl der bisher durchgeführten Sitzungen und der Anzahl der geplanten Sitzungen gemacht werden, sowie die Angabe, ob ein Kassenantrag gestellt wurde. Zudem sollte vom Therapeuten eine Code-Nr. vergeben werden.

Ursprünglich war es vorgesehen, dass die Therapeuten warten sollten, bis alle Patienten die Fragebögen zurückgebracht hätten, um diese dann zusammen in einem großen Umschlag mit den Angaben des Therapeuten zurückzuschicken. Bei der praktischen Umsetzung zeigte sich aber, dass viele Fragebögen einzeln zurückkamen und dann nur sehr schwer die zugehörigen Therapeuten ermittelt werden konnten. Daher wurden der Bogen „Angaben des Therapeuten“ und der Umschlag für die Fragebögen in der Folge um einen Therapeutencode ergänzt (Anfangsbuchstabe des Vornamens der Mutter und deren Geburtstag), um die oben beschriebenen Schwierigkeiten zu vermeiden.

Die Datenerhebung erfolgte zwischen Juli 2005 und November 2006, wobei der Großteil der Fragebögen im Zeitraum zwischen August 2005 und Januar 2006 ausgefüllt wurde. Unter den

Therapeuten, die sich mit den meisten Patienten für die Untersuchung angemeldet hatten, wurden am Ende der Datenerhebung Büchergutscheine im Wert von insgesamt 500 Euro verteilt. Für die Dateneingabe und die deskriptiven Analysen wurde das Programm „SPSS 14.0 for Windows“ verwendet.

### 4.1.3. Beschreibung der Stichprobe

#### 4.1.3.1. Therapeuten

Es nahmen insgesamt 36 Therapeuten an der Fragebogenuntersuchung teil. Von 35 Therapeuten wurde das Formblatt „Angaben des Therapeuten“ (siehe Anhang) ausgefüllt und zurückgeschickt. Von einem Therapeuten wurde der Bogen nicht zurückgesendet und konnte aufgrund der anonymisierten Erhebung auch nicht mehr nachgefordert werden. Die folgenden Angaben beziehen sich daher auf N=35 Therapeuten.

Alle Therapeuten waren Diplom-Psychologen. Der Altersdurchschnitt der Therapeuten lag bei 38,29 Jahren (SD=9,98), in einem Bereich zwischen 27 und 62 Jahren. 31 Therapeuten hatten eine verhaltenstherapeutische Ausbildung begonnen oder abgeschlossen, 3 Therapeuten hatten eine psychoanalytische und/oder tiefenpsychologische Ausbildung (eine Angabe fehlte). Eine oder mehrere zusätzliche Therapieausbildungen hatten 6 Therapeuten (z.B. Gesprächspsychotherapie, Systemische Therapie, Hypnotherapie) und 3 Therapeuten arbeiteten in mehr als einer der unten aufgeführten Ambulanzen. Im Mittel lag die Therapieerfahrung bei 8,5 Jahren (SD= 8,25). Nähere Angaben zu Alter, Geschlecht, Therapieerfahrung und Ort der Therapiedurchführung finden sich in Tabelle 7.

Tabelle 7

Angaben zu Alter, Geschlecht und Berufserfahrung der Therapeuten und zum Ort der Therapiedurchführung (N=35)

<b>Alter (in Jahren)</b>	<b>N</b>
26 - 35	19
36 - 45	8
46 - 55	4
56 - 65	4
<b>Summe</b>	<b>35</b>

#### 4. Fragebogenuntersuchung

---

<b>Geschlecht</b>	<b>N</b>
weiblich	31 (88,6%)
männlich	4 (11,4%)
Summe	35 (100%)
<b>Therapieerfahrung (Angabe in Jahren; Monaten)</b>	<b>N</b>
< 1 Jahr	1
1 - 2;11	6
3 - 4;11	8
5 - 10;0	12
>10;0	8
<b>Summe</b>	<b>35</b>

<b>Ort der Therapiedurchführung</b>	
Ambulanzen der Christoph-Dornier-Stiftung (Institute Tübingen und Berlin)	6
Hochschulambulanz der Universität Tübingen	6
Ambulanz der Tübinger Akademie für Verhaltenstherapie (TAVT)	14
Eigene Praxis	9
<b>Summe</b>	<b>35</b>

Gemäß den Angaben der Kassenärztlichen Bundesvereinigung (2006) zur vertragsärztlichen Versorgung in Deutschland liegt der Anteil weiblicher psychologischer Psychotherapeuten bei 67%, womit der Frauenanteil in der Untersuchungsstichprobe deutlich über dem Durchschnitt lag. Auffällig unterrepräsentiert waren in der Untersuchung mit 8,8% die Therapeuten und Therapeutinnen mit tiefenpsychologischer und/oder analytischer Ausbildung, die in der vertragsärztlichen Versorgung zusammengenommen einen Anteil von 55% ausmachen (KBV, 2006).

Von den 36 Therapeuten und Therapeutinnen wurden insgesamt 222 Patienten-Fragebögen zurückgeschickt. Die Anzahl der Patienten pro Therapeut schwankte zwischen 1 und 36 ( $M=20,86$ ;  $SD= 11,44$ ). Nähere Angaben zur Häufigkeit finden sich in Tabelle 8.

Tabelle 8

Angaben zur Anzahl der behandelten Patienten pro Therapeut

<b>Anzahl zurückgesendeter Patientenfragebögen pro Therapeut</b>	<b>N</b>
1-5	16
6-10	17
11-15	2
>15	1
<b>Summe</b>	<b>36</b>

#### 4.1.3.2. Patienten

An diejenigen Therapeuten, die sich prinzipiell zur Teilnahme an der Untersuchung bereit erklärt hatten, wurden 659 Fragebögen ausgegeben bzw. verschickt. Von 222 Patienten wurden Fragebögen ausgefüllt und zurückgegeben. Dies entspricht einer Rücklaufquote von 33,7 %.

In Hinblick auf fehlende Werte („missing values“) wurde ein Cut-Off-Kriterium von mehr als 20% fehlender Antworten (18 Items) für den Ausschluss eines kompletten Datensatzes festgelegt. Diesem Kriterium folgend wurden 6 Datensätze von der weiteren Analyse ausgeschlossen. Bei drei der sechs Fälle, die ausgeschlossen wurden, handelte es sich um zufallsbedingt fehlende Werte, da die betreffenden Patienten jeweils vergessen hatten, eine Seite auszufüllen. In den drei anderen Fällen hatten die Patienten Schwierigkeiten im Umgang mit dem Fragebogen (Instruktion wurde falsch verstanden; Items erschienen zu schwierig; häufige Mehrfachantworten pro Item). Es ist anzunehmen, dass diese Schwierigkeiten primär personenbedingt und nicht auf das Messinstrument zurückzuführen waren, da sonst ein größerer Teil der Befragten entsprechend viele fehlende Werte hätte aufweisen müssen. Insgesamt gingen 216 Datensätze in die Auswertung ein.

Im Folgenden werden die wichtigsten Daten zur Beschreibung der Stichprobe tabellarisch aufgeführt. Um die Repräsentativität der Stichprobe abschätzen zu können wurde eine Untersuchung als Referenz herangezogen, deren Ziel es war, Basisdaten zur Qualitätssicherung in der ambulan-

ten Psychotherapie zu erheben, und die an einer Stichprobe von N=804 Psychotherapiepatienten durchgeführt wurde (Scheidt et al., 1998).

### a) Soziodemographische Angaben

Tabelle 9  
Angaben zum Geschlecht

<b>Geschlecht</b>	<b>N</b>	<b>Prozent</b>
weiblich	150	69,4
männlich	66	30,6
<b>Summe</b>	<b>216</b>	<b>100</b>

Es haben deutlich mehr Frauen an der Untersuchung teilgenommen als Männer. Das Verhältnis von ca. 70:30 entspricht allerdings ungefähr der Geschlechtsverteilung der Patienten, die in psychotherapeutischen Praxen anzutreffen sind (vgl. Scheidt et al. (1998): 74,0% weibliche Patienten, 26,0% männliche Patienten).

Tabelle 10  
Angaben zum Alter

<b>Alter (Angaben in Jahren)</b>	<b>N</b>	<b>Prozent</b>
18-25	42	19,4
26-35	54	25,0
36-45	67	31,0
46-55	37	17,1
56-65	10	4,6
>65	6	2,8
<b>Summe</b>	<b>216</b>	<b>98,9</b>

Tabelle 11  
Angaben zum Schulabschluss

<b>Schulabschluss</b>	<b>N</b>	<b>Prozent</b>
Kein Schulabschluss	2	0,9
Hauptschulabschluss	27	12,5
Realschulabschluss	64	29,6
Abitur	67	31,0
Abgeschlossenes Studium (Universität; FH)	55	25,5
<b>Summe</b>	<b>215</b>	<b>99,5</b>

Im Mittel waren die Patienten 37,7 Jahre alt (SD=11,8). Der Range reichte von 18 bis 74 Jahre. Die Altersverteilung der Patienten und die Angaben zum höchsten Schulabschluss weichen insgesamt nicht erheblich von denen der Referenzstichprobe ab (Scheidt et al., 1998). Allerdings finden sich in der vorliegenden Stichprobe mehr Patienten in der Altersgruppe 18-25 Jahre als in der zitierten Vergleichsstichprobe (Scheidt et al. (1998): 9,8%). In der Referenzstichprobe wurde als höchster Schulabschluss nicht das abgeschlossene Studium aufgeführt, sondern das Abitur bzw. die Fachhochschulreife (Scheidt et al. (1998): 46%). Legt man die Kategorien „Abitur“ und „Abgeschlossenes Studium“ aus der vorliegenden Untersuchung zusammen, ergibt sich ein Prozentsatz von 56,5%, was für ein überdurchschnittlich hohes Bildungsniveau der Untersuchungsstichprobe spricht. Die Kombination aus hohem Bildungsniveau und einer Häufung an Patienten in der Altersklasse 18-25J. spricht für einen hohen Anteil an Studenten in der untersuchten Stichprobe, was vermutlich an der räumlichen und institutionellen Nähe der beteiligten Praxen und Ambulanzen zur Universität liegt.

Auf Angaben zur Berufstätigkeit der Patienten wurde verzichtet, um die Anonymität zu wahren, da das Wissen um den Beruf die Rekonstruktion der Identität des Befragten deutlich erleichtern würde (vgl. Rost, 2004).

Tabelle 12  
Angaben zur Partnerschaft

<b>Partnerschaft</b>	<b>N</b>	<b>Prozent</b>
Verheiratet/ feste Partnerschaft	138	63,9

#### 4. Fragebogenuntersuchung

---

Keine feste Partnerschaft	76	35,2
<b>Summe</b>	<b>214</b>	<b>99,1</b>

Leider fehlen in der zitierten Referenzuntersuchung Angaben zur Partnerschaft der untersuchten Patienten, so dass hier keine Vergleichsangaben möglich sind.

### **b) Störungsbezogene Angaben**

#### Diagnosen

Die Angaben zu den Diagnosen der Patienten wurden von den Therapeuten nach den Vorgaben der ICD-10, Kapitel F (Dilling et al., 2000) codiert. Zur besseren Veranschaulichung wurden die Erstdiagnosen der Patienten unter der jeweiligen Kapitelbezeichnung bzw. –nummer der ICD-10 zusammengefasst. Diagnosen aus den Kapiteln F7-F9/F99 wurden nicht vergeben, so dass auf eine Auflistung in der Tabelle verzichtet wurde. An N=101 (46,8%) der Patienten wurde mehr als eine Diagnose vergeben, was hier aber nicht separat aufgelistet werden soll. Die Häufigkeit von zweit und Drittdiagnosen unterscheidet sich nicht wesentlich von der Referenzstichprobe (vgl. Scheidt et al., (1998): 54% Zweitdiagnosen).

Tabelle 13

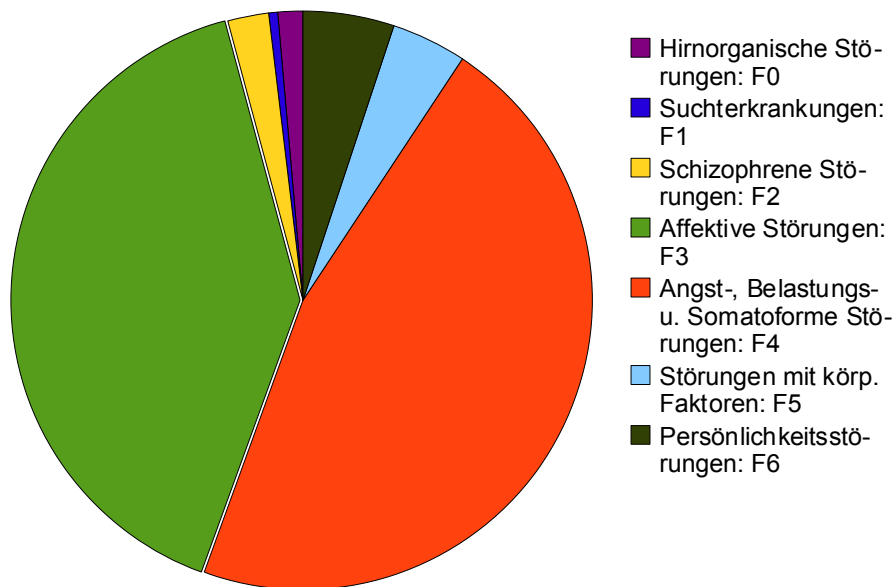
#### Erstdiagnose nach ICD-10

<b>Erstdiagnose</b>	<b>N</b>	<b>Prozent</b>
F0: Organische, einschließlich symptomatischer psychische Störungen	3	1,4
F1: Psychische und Verhaltensstörungen durch psychotrope Substanzen	1	0,5
F2: Schizophrenie, schizotype und wahnhafte Störungen	5	2,3
F3: Affektive Störungen	87	40,3
F4: Neurotische-, Belastungs- und somatoforme Störungen	100	46,3
F5: Verhaltensauffälligkeiten mit körp. Störungen /Faktoren	9	4,2
F6: Persönlichkeits- und Verhaltensstörungen	11	5,1
<b>Summe</b>	<b>216</b>	<b>100</b>



In Abbildung 1 ist die Häufigkeitsverteilung der Erstdiagnosen als Tortendiagramm dargestellt.

Abbildung 1: Verteilung der Erstdiagnosen nach ICD-10



Die Häufigkeitsverteilung der Erstdiagnosen in der vorliegenden Stichprobe kann als typisch für die psychotherapeutische Praxis angesehen werden (vgl. Scheidt et al., 1998).

### Dauer der Problematik

Die Angabe zur Dauer der Problematik wurde vom Patienten gemacht. Da ein freies Antwortformat vorgegeben war, fanden sich häufig Aussagen, die keinen klaren Bezug zu messbaren zeitlichen Einheiten hatten (z.B. „schon ewig“). Sofern es Anhaltspunkte für den subjektiv erlebten Beginn der Problematik gab und die Angabe plausibel erschien (z.B. „seit der Pubertät“), wurde sie unter Berücksichtigung des Alters des Patienten in der entsprechenden Kategorie vermerkt. Ansonsten wurde sie unter „fehlende Angabe“ vermerkt.

Tabelle 14  
Angaben zur Dauer der Problematik

<b>Dauer der Problematik</b> (Angaben in Jahren; Monaten)	<b>N</b>	<b>Prozent</b>
< 1	16	7,4
1-2;11	39	18,1
3-4;11	41	19,0
5-10	39	18,1
>10	62	28,7
<b>Summe</b>	<b>197</b>	<b>91,3</b>

Die mittlere Störungsdauer lag bei 8,0 Jahren (SD=8,58). Leider fehlen in der zitierten Referenzuntersuchung Angaben zur Störungsdauer, so dass hier keine Vergleichsangaben möglich sind.

### Belastungsstärke

Die Einschätzung der aktuellen Belastungsstärke wurde ebenfalls vom Patienten getroffen, wobei hier eine fünfstufige Skala von „gar nicht“ bis „sehr stark“ vorgegeben war.

Tabelle 15  
Angaben zur Belastungsstärke

<b>Aktuelle Belastungsstärke</b>	<b>N</b>	<b>Prozent</b>
gar nicht	4	1,9
wenig	24	11,1
mittelmäßig	89	41,2
stark	51	23,6
sehr stark	44	20,4
<b>Summe</b>	<b>212</b>	<b>98,2</b>

### Vorbehandlung

Auch die Angabe zur Vorbehandlung wurde vom Patienten vorgenommen.

Tabelle 16  
Angaben zur Vorbehandlung

Vorbehandlung	N	Prozent
ja	113	52,3
nein	103	47,7
<b>Summe</b>	<b>216</b>	<b>100</b>

Es wird deutlich, dass knapp über die Hälfte der Patienten bereits vorbehandelt worden ist und somit über eine gewisse „Therapieerfahrung“ verfügte. Gleichzeitig spricht dies aber auch für eine gewisse Schwere bzw. Chronifizierung der Störung, was den Ergebnissen der Tabelle „Dauer der Problematik“ entspricht. Der Prozentsatz der vorbehandelten Patienten lag über dem Referenz-Wert von 42,3% aus der Vergleichsstichprobe (Scheidt et al., 1998).

Ursprünglich war geplant gewesen, die Zeitdauer zwischen der letzten Therapiestunde und dem Ausfüllen des Fragebogens mitzuerfassen. Hier zeigte sich aber ein ungünstiger Reihungseffekt bzw. eine nicht ausreichend genaue Formulierung als Problem: da die Frage „Wie lange liegt die letzte Therapiestunde zurück“ unmittelbar nach der Angabe zur Vorbehandlung folgte, bezogen die meisten Patienten die Frage offensichtlich auf die vorhergehende Aussage. Aus diesem Grund wurde diese Frage von der weiteren Auswertung ausgeschlossen.

### c) Therapiebezogene Angaben

In Tabelle 17 werden die tiefenpsychologisch und psychoanalytisch behandelten Patienten in einer Gruppe zusammengefasst. Dies stellt zwar inhaltlich eine Pauschalisierung dar, erscheint aber sinnvoll im Hinblick auf die geringe Anzahl nicht verhaltenstherapeutisch arbeitender Therapeuten, die in dieser Stichprobe vertreten war.

Tabelle 17  
Angaben zur Therapierichtung

Therapierichtung	N	Prozent
Verhaltenstherapie	191	88,4
Psychoanalyse/ Tiefenpsychologie	24	11,1
<b>Summe</b>	<b>215</b>	<b>99,5</b>

#### 4. Fragebogenuntersuchung

---

Tabelle 17 zeigt, dass in dieser Stichprobe ein deutliches Ungleichgewicht zwischen den Fallzahlen verhaltenstherapeutisch behandelter Patienten und psychoanalytisch bzw. tiefenpsychologisch behandelter Patienten bestand.

Tabelle 18  
Angaben zur Therapiefrequenz

<b>Therapiefrequenz</b>	<b>N</b>	<b>Prozent</b>
< 2 Sitzungen/Monat	12	5,6
2-3 Sitzungen/Monat	31	14,4
1 Sitzung/Woche	151	69,9
2-3 Sitzungen/Woche	20	9,3
> 5 Sitzungen/Woche	1	0,5
<b>Summe</b>	<b>215</b>	<b>99,5</b>

Der Großteil der Therapien wurde im Rahmen der üblichen, von den Krankenkassen empfohlenen Frequenz von einer Therapiestunde pro Woche durchgeführt. 20% der Therapien fanden niedrigerfrequenter statt. In knapp 10% der Therapien wurden mehrmals wöchentlich Sitzungen durchgeführt.

Tabelle 19  
Angaben zur Sitzungsanzahl (incl. der probatorischen Sitzungen)

<b>Sitzungsanzahl</b> (bis zum Zeitpunkt der Datenerhebung)	<b>N</b>	<b>Prozent</b>
0-10	34	15,7
11-30	111	51,4
31-50	42	19,4
51-80	19	8,8
81-160	4	1,9
161-240	3	1,4
>240	2	0,9
<b>Summe</b>	<b>215</b>	<b>99,5</b>

Aus Tabelle 19 wird ersichtlich, dass die Anzahl der bisher durchgeführten Therapiestunden für einen Großteil der Patienten (86,5%) im Bereich der kassenüblichen Kurz- bzw. Langzeittherapi-

en (25+5 bzw. 45+5 Therapiestunden) lag. Über 50% der Therapiestunden lagen im Bereich zwischen 11 und 30 Sitzungen.

Tabelle 20

Angaben zur Anzahl geplanter Sitzungen (incl. der probatorischen Sitzungen)

Geplante Sitzungsanzahl	N	Prozent
0-30	38	17,6
31-50	100	46,3
51-65	38	17,6
66-80	22	10,2
81-160	6	2,8
161-240	2	0,9
>240	5	2,3
<b>Summe</b>	<b>211</b>	<b>97,7</b>

### 4.1.3.3. Zusammenfassung

Die Stichprobengröße lag mit der Anzahl von 216 Datensätzen zwischen der festgelegten Untergrenze von N=200 und der Zielgröße von N=250.

Der Vergleich der vorliegenden Patientenstichprobe mit der Referenzstichprobe aus der Untersuchung von Scheidt et al. (1998) zeigt, dass in Hinblick auf die soziodemographischen Variablen keine wesentlichen Unterschiede zwischen den Stichproben festzustellen sind, mit Ausnahme des höheren Bildungsniveaus bzw. der stärkeren Besetzung der Altersklasse 18-25J (höherer Studentenanteil). Die Häufigkeiten der Erstdiagnosen nach ICD-10-Kategorien sind in den beiden Stichproben ebenso vergleichbar wie die Anzahl der Mehrfachdiagnosen. Eine Abweichung zeigt sich lediglich in der Anzahl der Vorbehandlungen. Insgesamt sprechen die Vergleiche dafür, dass die untersuchte Patientenstichprobe in zufriedenstellendem Maß die Population der Psychotherapiepatienten abbildet und in dieser Hinsicht eine qualitative Repräsentativität gegeben ist. Hinsichtlich der Verteilung der Patienten auf die verschiedenen Therapieschulen bzw. -methoden muss allerdings von einer Nicht-Repräsentativität der Untersuchungsstichprobe gesprochen werden, da hier ein deutliches Ungleichgewicht zugunsten der Verhaltenstherapie besteht.

Die andere Form der möglichen Nicht-Repräsentativität einer Stichprobe bezieht sich, wie bereits angeführt, auf die stichprobenbedingte Einschränkung der Unterschiede im Ausprägungs-

grad der zu erfassenden Merkmale. Um dies zu überprüfen, wurde eine mögliche Selektion der Patienten durch die Therapeuten stichprobenartig abgefragt, was keine diesbezüglichen Hinweise erbrachte. Wie Tabelle 19 zeigt, gelang es auch, Patienten aus verschiedenen Phasen der Therapie in die Untersuchung aufzunehmen. Allerdings befanden sich nur 15,7% der Patienten im ersten Therapieabschnitt (0-10. Sitzung), womit diese Gruppe – mit dem vermeintlich größten Anteil an Patienten mit geringerer Merkmalsausprägung – deutlich unterrepräsentiert war. Die Selbst-Selektion der Patienten wurde nicht überprüft.

## 5. Auswertung und Ergebnisse

Das Ziel der Datenauswertung stellt die statistische Überprüfung der theoretisch hergeleiteten Struktur des Fragebogens dar. Zudem soll mithilfe einer primär faktorenanalytischen Auswertung eine fundierte Grundlage für die Itemselektion geschaffen werden. Da die Wahl der Verfahren für die Überprüfung der Fragebogenstruktur und die Auswahl und Analyse der Items eine wesentliche Rolle spielt, werden diesbezügliche Entscheidungen genauer erläutert. In den folgenden Kapiteln sollen zunächst jeweils die Auswertungsschritte beschrieben werden, bevor die Darstellung der Ergebnisse erfolgt.

### 5.1. Deskriptive Auswertung und Berechnung der Itemschwierigkeiten

#### 5.1.1. Vorgehen

Zu Beginn wird eine Aufbereitung der Rohdaten durchgeführt, die auch eine Umpolung der negativ formulierten Items beinhaltet. Anhand dieses Datensatzes erfolgt dann eine erste Überprüfung hinsichtlich der deskriptiven Eigenschaften (Angaben von Mittelwert mit Standardabweichung, Median, Kurtosis, Schiefe, Minimal- und Maximalwert pro Item). Zur statistischen Überprüfung der Itemverteilungen soll dann der Kolmogorov-Smirnov Test auf Normalverteilung durchgeführt werden (vgl. Bühner, 2006)<sup>15</sup>.

In einem nächsten Schritt werden die Itemschwierigkeiten berechnet, um eine Vorselektion der Items zu ermöglichen. Der Schwierigkeitsindex eines Items ist ein Maß dafür, wieviele Proban-

---

<sup>15</sup> Die Normalverteilung von Testitems ist aus statistischer Sicht vorteilhaft, da sie bestimmte statistische Auswertungsmethoden ermöglicht. Allerdings sind im klinisch/therapeutischen Bereich normalverteilte Fragebogenitems eher als Ausnahme anzusehen (vgl. Bühner, 2006).

den das Item in Schlüsselrichtung beantwortet haben. Da Items mit einer sehr geringen Schwierigkeit (=hoher „Schwierigkeits“-Index)<sup>16</sup> nicht zwischen Personen mit verschiedenen Merkmalsausprägungen ausreichend differenzieren (Deckeneffekt)<sup>17</sup>, ist es sinnvoll, extrem leichte Items als eine erste Maßnahme der Itemselektion vor der Durchführung weiterer Auswertungsschritte zu eliminieren. Zu diesem Zweck soll für alle Items die Schwierigkeit berechnet werden nach der Formel:

$$P_1 = \frac{\sum_{v=1}^n [x_{vi} - \min(x_i)]}{n[\max(x_i) - \min(x_i)]} * 100 \quad (\text{Moosbrugger \& Kelava, 2007, S. 75})$$

Der Schwierigkeitsindex ergibt sich demnach als Quotient aus der bei dem betreffenden Item erreichten Punktschme aller n Probanden und der maximal erreichbaren Punktschme aller n Probanden, multipliziert mit 100. Da der Wertebereich der Itemantworten nicht bei 0 beginnt, sondern bei 1 (s.u. 3.2.1), müssen zudem das potentiell erreichbare Minimum einer Itemantwort ( $\min(x_i)$ ) von jeder realisierten Itemantwort ( $x_{vi}$ ) im Zähler und die minimal erreichbare Punktschme der n Probanden im Nenner abgezogen werden (vgl. Moosbrugger & Kelava, 2007).

Für die Selektion von Items aufgrund ihres Schwierigkeitsindex gibt es keine allgemein gültigen Regeln. Häufig werden Items mit einem Schwierigkeitsindex > 85.00 eliminiert (z.B. Fuchs, 1998), Moosbrugger und Kelava (2007) hingegen geben einen Wert von 95.00 als Cut-Off-Kriterium an. Für die vorliegende Untersuchung wird für die Eliminierung von Items ein Schwierigkeitsindex >= 90.00 festgelegt, d.h., dass alle Items mit einem Wert >= 90,00 ausgeschlossen werden sollen. Nach der Eliminierung der leichtesten Items sollen die faktorenanalytischen Auswertungsschritte erfolgen.

## 5.1.2. Ergebnisse

### 5.1.2.1. Deskriptive Ergebnisse

Die Aufbereitung der Rohdaten und die ersten Auswertungsschritte erfolgen mit dem Softwarepaket „SPSS 14.0 for Windows“ (2005). Nach der Eliminierung der sechs Datensätze mit mehr als 20% missing values gehen in die weitere Auswertung 216 Datensätzen ein. Den Stufen der Likertskala werden, den sieben verbalen Abstufungen entsprechend, Zahlenwerte zugeordnet,

---

16 Der Begriff „Leichtigkeitsindex“ wäre vermutlich hilfreicher, da hohe Indexwerte – im testtheoretischen Sinn – für leichte Aufgaben stehen.

17 Das Gleiche gilt für Items mit zu hoher Schwierigkeit, d.h. geringem Schwierigkeitsindex (Bodeneffekt), womit aber in der vorliegenden Untersuchung nicht zu rechnen war.

## 5. Auswertung und Ergebnisse

von „stimmt überhaupt nicht“ = 1 bis „stimmt voll und ganz“ = 7. Für die weiteren Auswertungsschritte werden die 30 negativ gepolten Items umgepolt, d.h. in diesen Fällen wird die Einschätzung „stimmt überhaupt nicht“ mit 7 codiert, die Einschätzung „stimmt nicht“ mit 6 etc. In Tabelle 21 sind die Ergebnisse der deskriptiven Auswertung für die 88 Items aufgeführt, wobei die Items zwar mit Nummer, aber nicht in der Reihenfolge gemäß der Darbietung im Fragebogen aufgelistet sind, sondern entsprechend ihrer postulierten Zugehörigkeit zu den a-priori-Faktoren<sup>18</sup>. Aus Platzgründen wird auf eine Darstellung der Formulierungen verzichtet. Eine vollständige Ansicht der Itemformulierungen befindet sich im Anhang.

Tabelle 21

Deskriptive Ergebnisse (Angaben von Mittelwert mit Standardabweichung, Median, Kurtosis, Schiefe, Minimal- und Maximalwert pro Item)

<b>Item-Nr. (Skalen-Nr.)</b>	<b>M (SD)</b>	<b>Md</b>	<b>Schiefe</b>	<b>Kurtosis</b>	<b>Min</b>	<b>Max</b>
19 (I)	5.58 (1.237)	6.00	-1.569	2.762	1	7
45 (I)	5.80 (1.006)	6.00	- .466	- .834	4	7
47 (I)	5.86 (1.011)	6.00	- .584	- .711	4	7
48 (I)	5.89 (1.046)	6.00	- .612	- .798	4	7
56 (I)	5.62 (1.139)	6.00	-1.064	1.375	1	7
62 (I)	6.02 ( .767)	6.00	- .721	.606	4	7
83 (I)	5.58 ( .959)	6.00	- .242	- .874	4	7
86 (I)	4.44 (1.399)	4.00	- .300	- .519	1	7
89 (I)	4.21 (1.686)	4.00	- .258	- .998	1	7
68 (II)	5.18 (1.357)	6.00	- .832	.456	1	7
71 (II)	5.40 (1.272)	6.00	- .755	.155	1	7
73 (II)	5.36 (1.293)	6.00	- .820	.316	1	7
74 (II)	5.07 (1.348)	5.00	- .400	- .406	1	7
75 (II)	5.54 (1.095)	6.00	- .752	.568	1	7
76 (II)	5.40 (1.125)	6.00	- .606	.032	1	7
78 (II)	6.06 ( .894)	6.00	- .631	- .425	4	7
80 (II)	5.72 (1.076)	6.00	- .298	-1.176	4	7
82 (II)	5.07 (1.323)	5.00	- .846	.541	1	7
84 (II)	5.39 (1.070)	6.00	- .887	1.115	1	7
85 (II)	5.67 (1.054)	6.00	- .274	- 1.119	4	7
87 (II)	5.77 ( .894)	6.00	- .317	- .620	4	7
90 (II)	5.00 (1.367)	5.00	- .657	.213	1	7
1 (III)	3.18 (2.132)	2.00	.549	-1.235	1	7
4 (III)	5.36 (1.707)	6.00	-1.142	.282	1	7

18 Skalenbezeichnung: I.= Glaubwürdigkeit/Nachvollziehbarkeit des Therapieangebotes, II.= Veränderungserwartung des Patienten, III.= Therapeutische Beziehung, IV.= Bereitschaft des Patienten, sich auf die Therapie einzulassen und mitzuarbeiten und V.= Kompetenz des Therapeuten (K)



## 5. Auswertung und Ergebnisse

5 (III)	6.21 (1.330)	7.00	-2.364	5.443	1	7
6 (III)	6.62 ( .613)	7.00	-1.996	6.175	3	7
8 (III)	6.63 ( .611)	7.00	-1.679	2.801	4	7
10 (III)	6.63 ( .697)	7.00	-2.180	4.819	4	7
12 (III)	6.25 ( .900)	6.50	-1.028	.162	4	7
13 (III)	5.80 (1.085)	6.00	- .424	-1.111	4	7
15 (III)	6.39 ( .822)	7.00	-1.388	1.445	4	7
16 (III)	6.63 ( .619)	6.00	-1.812	3.624	4	7
17 (III)	5.90 (1.034)	6.00	- .610	- .767	4	7
18 (III)	6.66 ( .565)	7.00	-1.601	2.443	4	7
21 (III)	5.70 (1.146)	6.00	-1.503	3.008	1	7
27 (III)	5.65 (1.267)	6.00	-1.083	.985	1	7
29 (III)	6.22 ( .686)	6.00	- .837	1.377	4	7
30 (III)	5.76 (1.080)	6.00	-1.148	1.743	1	7
31 (III)	5.83 (1.059)	6.00	- .927	.321	2	7
33 (III)	5.74 ( .974)	6.00	- .779	.490	2	7
35 (III)	6.13 ( .890)	6.00	-1.212	1.609	3	7
38 (III)	6.42 ( .719)	7.00	-1.515	3.393	3	7
39 (III)	6.36 ( .771)	6.00	-1.461	3.051	3	7
40 (III)	6.15 ( .826)	6.00	-1.346	3.420	2	7
43 (III)	6.24 ( .819)	6.00	-1.093	1.193	3	7
44 (III)	6.30 ( .775)	6.00	- .932	.407	4	7
46 (III)	6.00 (1.059)	6.00	- .799	- .593	4	7
49 (III)	6.23 ( .931)	6.00	-1.106	.326	4	7
51 (III)	5.86 (1.045)	6.00	-1.601	4.286	1	7
53 (III)	6.14 ( .954)	6.00	- .940	- .058	4	7
55 (III)	5.90 (1.063)	6.00	-1.788	4.851	1	7
57 (III)	5.60 (1.248)	6.00	-1.398	2.473	1	7
58 (III)	5.85 ( .957)	6.00	- .431	- .749	4	7
60 (III)	5.93 ( .846)	6.00	- .704	.365	3	7
61 (III)	6.00 ( .930)	6.00	-1.312	2.881	2	7
65 (III)	6.01 ( .928)	6.00	-1.510	3.828	2	7
70 (III)	5.79 (1.057)	6.00	- .924	.537	2	7
7 (IV)	3.64 (1.584)	4.00	.203	- .699	1	7
11 (IV)	5.52 (1.244)	6.00	- .956	.872	1	7
20 (IV)	5.13 (1.640)	6.00	-1.057	.258	1	7
22 (IV)	6.12 ( .925)	6.00	- .882	- .037	4	7
24 (IV)	6.38 ( .706)	6.00	-1.024	.915	4	7
26 (IV)	5.89 ( .987)	6.00	- .625	- .582	4	7
36 (IV)	6.15 ( .913)	6.00	- .8160	- .244	4	7
42 (IV)	5.32 (1.570)	6.00	-1.080	.342	1	7
52 (IV)	5.03 (1.458)	5.00	- .761	.173	1	7
59 (IV)	5.77 (1.199)	6.00	-1.356	2.277	1	7

## 5. Auswertung und Ergebnisse

63 (IV)	5.98 (.814)	6.00	-.647	.150	4	7
69 (IV)	4.95 (1.397)	5.00	-.571	-.433	1	7
72 (IV)	5.29 (1.390)	6.00	-.861	.205	1	7
77 (IV)	5.55 (1.001)	6.00	-.142	-1.037	4	7
79 (IV)	6.46 (.721)	7.00	-1.336	1.596	4	7
81 (IV)	4.94 (1.690)	6.00	-.803	-.338	1	7
88 (IV)	5.26 (1.390)	6.00	-.950	.628	1	7
2 (V)	6.13 (.916)	6.00	-1.724	5.602	1	7
3 (V)	5.69 (1.372)	6.00	-1.514	1.993	1	7
9 (V)	5.53 (1.197)	6.00	-1.105	1.014	1	7
14 (V)	6.53 (.708)	7.00	-1.584	2.351	4	7
23 (V)	5.29 (1.338)	6.00	-.966	.587	1	7
25 (V)	5.73 (1.088)	6.00	-1.624	3.802	1	7
28 (V)	5.63 (1.122)	6.00	-.755	.156	2	7
32 (V)	5.21 (1.254)	5.00	-.783	.580	1	7
34 (V)	5.94 (.972)	6.00	-1.189	2.010	2	7
37 (V)	5.81 (1.003)	6.00	-1.068	1.072	2	7
41 (V)	6.20 (.748)	6.00	-1.957	10.516	1	7
50 (V)	6.34 (.723)	6.00	-.761	-.169	4	7
54 (V)	6.44 (.764)	7.00	-1.450	1.899	4	7
64 (V)	5.93 (.942)	6.00	-1.078	1.569	2	7

Aus Tabelle 21 wird ersichtlich, dass die Items zum größten Teil einen Mittelwert  $>5.00$  bzw. einen Median von  $\geq 5$  haben. Dies bedeutet, dass die meisten Items - im testtheoretischen Sinn - leicht sind, da die meisten Patienten ein zustimmendes Urteil („stimmt eher“, „stimmt“, „stimmt voll und ganz“) abgegeben haben, während ablehnende („stimmt eher nicht“, „stimmt nicht“, „stimmt überhaupt nicht“) oder unentschiedene („stimmt teils-teils“) Urteile deutlich seltener sind. Betrachtet man den Range der Items, so wird deutlich, dass bei einer erheblichen Anzahl an Items gar nicht die ganze Breite der Antwortkategorien ausgenutzt wurde (Range  $<6$ ). Zudem weisen nahezu alle Items eine linksschiefe Verteilung (Sch  $<0$ ) auf, was bedeutet, dass mehr Urteile oberhalb des Mittelwertes liegen als darunter. Hinsichtlich der Kurtosis ist das Bild weniger einheitlich; hier zeigen sich sowohl steilgipfelige (Kur  $>0$ ), als auch flachgipflige (Kur  $<0$ ) Verteilungen.

Zur statistischen Überprüfung der Itemverteilungen wurde der Kolmogorov-Smirnov Test auf Normalverteilung durchgeführt (vgl. Bühner, 2006). Dieser Test lieferte in allen Fällen ein signifikantes Ergebnis ( $p = 0.000$ ), was bedeutet, dass alle Items signifikant von einer Normalverteilung abweichen (Darstellung der Ergebnisse: siehe Anhang).

Inhaltlich bedeutet dies, dass ein Großteil der Patienten zustimmend auf die Fragen nach Glaubwürdigkeit des Therapieangebotes, Veränderungserwartung, Therapeutischer Beziehung, Mitarbeit und Kompetenz des Therapeuten (als a-priori-Skalen) geantwortet haben. Bei den weiteren Auswertungsschritten müssen die Abweichungen von der Normalverteilung berücksichtigt werden.

### 5.1.2.2. Itemschwierigkeiten

Die Itemschwierigkeiten sind in Tabelle 22 aufgelistet.

Tabelle 22  
Schwierigkeitsindices (P)

Item-Nr. (Skala)	P	Item-Nr. (Skala)	P	Item-Nr. (Skala)	P
19 (I)	76.40	13 (III)	80.08	11 (IV)	75.39
45 (I)	80.03	15 (III)	89.81	20 (IV)	68.76
47 (I)	80.93	16 (III)	93.83	22 (IV)	85.28
48 (I)	81.48	17 (III)	81.64	24 (IV)	89.74
56 (I)	77.04	18 (III)	94.34	26 (IV)	81.48
62 (I)	83.64	21 (III)	78.37	36 (IV)	85.8
83 (I)	76.40	27 (III)	77.55	42 (IV)	71.99
86 (I)	57.32	29 (III)	86.98	52 (IV)	67.22
89 (I)	53.57	30 (III)	79.40	59 (IV)	79.53
68 (II)	69.69	31 (III)	80.42	63 (IV)	83.02
71 (II)	73.30	33 (III)	78.95	69 (IV)	65.89
73 (II)	72.71	35 (III)	85.43	72 (IV)	71.42
74 (II)	67.90	38 (III)	90.34	77 (IV)	75.77
75 (II)	75.69	39 (III)	89.35	79 (IV)	91.05
76 (II)	73.38	40 (III)	85.83	81 (IV)	65.74
78 (II)	84.27	43 (III)	87.36	88 (IV)	71.01
80 (II)	78.74	44 (III)	88.27	2 (V)	85.43
82 (II)	67.91	46 (III)	83.41	3 (V)	78.25
84 (II)	73.24	49 (III)	87.19	9 (V)	75.55
85 (II)	77.83	51 (III)	80.92	14 (V)	92.21
87 (II)	79.44	53 (III)	85.67	23 (V)	71.44
90 (II)	66.67	55 (III)	81.61	25 (V)	78.81
1 (III)	36.31	57 (III)	76.74	28 (V)	77.08
4 (III)	72.71	58 (III)	80.78	32 (V)	70.19
5 (III)	86.76	60 (III)	82.18	34 (V)	82.25
6 (III)	93.67	61 (III)	83.41	37 (V)	80.16
8 (III)	93.83	65 (III)	83.56	41 (V)	86.59
10 (III)	93.83	70 (III)	79.78	50 (V)	88.97

## 5. Auswertung und Ergebnisse

---

12 (III)	87.54	7 (IV)	44.02	54 (V)	90.70
				64 (V)	82.17

\*Grau unterlegt sind alle Items mit einem Schwierigkeitsindex  $\geq 90.00$

Wie aus Tabelle 22 ersichtlich, weisen die Items zum größten Teil eine geringe bis sehr geringe Schwierigkeit auf. Es zeigt sich eine geringe Schwankungsbreite hinsichtlich der Schwierigkeitsindizes: mit Ausnahme des Items 1 liegen alle Werte zwischen 72.71 und 93.83. Über dem Cut-Off-Wert liegen die Items Nr. 6, Nr. 8, Nr. 10, Nr. 16, Nr. 18, Nr. 38, Nr. 79, Nr. 14 und Nr. 54 ( $\Sigma = 9$ ) (vgl. Tabelle 22), wobei Item 38 genau am Grenzbereich liegt. Da es sich dabei um ein inhaltlich wichtiges Item handelt („In der letzten Therapiestunde ... fühlte ich mich von meinem Therapeut als Mensch akzeptiert“), das einen zentralen Aspekt der therapeutischen Beziehung repräsentiert, wird dieses Item zu Gunsten der Inhaltsvalidität beibehalten. Daher werden nur die 8 anderen Items ausgeschlossen.

Somit verbleiben 80 Items für die weitere Auswertung.

## 5.2. Faktorenanalysen

### 5.2.1. Vorgehen

#### 5.2.1.1. Allgemeine Überlegungen

Nach der ersten Vorselektion der Items auf der Grundlage der Schwierigkeitsindizes wird für die weitere Itemauswahl und die Überprüfung der Struktur des Fragebogens ein faktorenanalytisches Vorgehen in mehreren Stufen gewählt. Der große Vorteil dieses Vorgehens besteht darin, dass vor einer Eliminierung von Items zunächst eine empirische Überprüfung des theoretisch hergeleiteten Gesamtkonzepts, d.h. der Anzahl der Faktoren und der Zuordnung der Indikatoren erfolgt. Ausgehend von dieser empirischen Fundierung kann dann eine schrittweise, statistisch begründete Eliminierung von Items erfolgen. Im Folgenden soll das Vorgehen erläutert und beschrieben werden.

Die Grundannahme der Faktorenanalyse ist, dass es latente, d.h. nicht direkt beobachtbare Variablen gibt (=Faktoren), welche die gemessenen Werte der beobachtbaren Variablen (=Indikatoren) beeinflussen und die Korrelationen zwischen diesen gemessenen Werten bedingen („common factor model“). Es wird also angenommen, dass die Indikatoren deshalb korrelieren, weil sie eine gemeinsame Ursache haben, nämlich das zugrundeliegende Konstrukt. Ziel der Faktorenanalyse

ist es, empirisch die Anzahl und die Eigenschaften der latenten Variablen zu bestimmen, die den beobachteten Varianzen und Kovarianzen bzw. Korrelationen der Indikatoren zugrundeliegen (vgl. DeVellis, 2003). Dabei wird die Varianz eines jeden Indikators in zwei Anteile zerlegt, nämlich in diejenige Varianz, die durch den latenten Faktor erklärt wird („common variance“) und in diejenige, die für diesen Indikator spezifisch ist („unique variance“: bestehend aus individueller systematischer Varianz und Fehlervarianz) (vgl. Brown, 2006). Man unterscheidet zwei Hauptarten von Faktorenanalysen, die beide auf dem „common factor model“ basieren, nämlich die explorative Faktorenanalyse (EFA) und die konfirmatorische Faktorenanalyse (CFA), wobei letztere eine Komponente so genannter Strukturgleichungsmodelle (SEM) darstellt. Im Gegensatz zur EFA erfordert die konfirmatorische Faktorenanalyse vorab eine genaue Spezifizierung bezüglich der Anzahl der Faktoren, dem Muster der Faktorladungen und anderen Parametern. Ziel der CFA ist zu überprüfen, wie gut die empirische Korrelations (bzw. Kovarianz)-Matrix durch die vorher festgelegte Faktorenlösung reproduziert wird (=modellprüfendes Verfahren). Da die CFA im Gegensatz zur EFA eine gute empirische und theoretische Fundierung benötigt, auf deren Grundlage die Spezifizierung des Faktorenmodells vorgenommen werden kann, ist es sinnvoll und üblich, unter Berücksichtigung der theoretischen Grundlage und der daraus resultierenden Vorannahmen, zunächst eine explorative Faktorenanalyse durchzuführen. Ausgehend von diesen Ergebnissen kann dann eine konfirmatorische Faktorenanalyse angeschlossen werden (vgl. Brown, 2006).

Bevor eine Faktorenanalyse durchgeführt werden kann, sind im Vorfeld, ausgehend von den Stichprobendaten und theoretischen Überlegungen, verschiedene Entscheidungen hinsichtlich der Wahl der passenden Auswertungsstrategie bzw. -methode zu treffen. Dabei sind auch die Verteilungseigenschaften der Indikatoren und die Annahmen bezüglich des Skalenniveaus zu berücksichtigen.

### **5.2.1.2. Verteilungscharakteristika**

Wie bereits aufgeführt weisen alle 80 Indikatoren deutliche Abweichungen von der Normalverteilung auf. Bei einem Großteil der Items ist dabei eine deutliche Linksschiefe zu beobachten, was sich auch in den hohen Schwierigkeitsindices (=leichte Items) widerspiegelt. Viele Schätzmethoden erfordern eine multivariate Normalverteilung der gemessenen Variablen. Andernfalls droht eine deutliche Verzerrung der Ergebnisse (vgl. Brown, 2006). Da die multivariate Normalverteilung die univariate Normalverteilung der Indikatoren voraussetzt, kann im Falle der vorliegenden Daten nicht von einer multivariaten Normalverteilung ausgegangen werden. Daher ist es

notwendig, einen Schätzalgorithmus auszuwählen, der gegenüber Verletzungen der Normalverteilung robust ist bzw. diese nicht notwendigerweise voraussetzt.

### **5.2.1.3. Skalenniveau**

Häufig wird bei der Verwendung von Rating-Skalen „per fiat“ davon ausgegangen, dass diese Intervallskalenniveau besitzen, ohne dass dies näher überprüft wird (vgl. Bühner, 2006). Dies setzt aber voraus, dass die Abstände zwischen den Skalenpunkten in gleicher Weise den Abständen im Urteil der Probanden entsprechen. Dies ist jedoch weder für numerische, noch für verbale Ratingskalen sicherzustellen (vgl. Moosbrugger & Kelava, 2007). Daher ist es empfehlenswert, nur von einem Ordinalskalenniveau auszugehen und die Indikatoren als kategoriale Variablen zu behandeln.

### **5.2.1.4. Auswahl eines geeigneten Messmodells**

Zur Berechnung von Faktorenanalysen auf der Grundlage nicht-normalverteilter, ordinaler Indikatorvariablen sind die traditionellen linearen Modelle der Faktorenanalyse nicht geeignet, da sie von kontinuierlichen beobachteten Variablen ausgehen. Die Voraussetzung einer multivariaten Normalverteilung gilt auch für die Anwendung konventioneller Strukturgleichungsmodelle (SEM), zu denen die konfirmatorische Faktorenanalyse (CFA) als ein Bestandteil zählt. Strukturgleichungsmodelle bestehen aus einer messtheoretischen Komponente und einem Strukturmodell. Im Messmodell werden die Zusammenhänge zwischen den beobachteten Variablen (=Indikatoren) und den latenten Variablen spezifiziert und überprüft. Die konfirmatorische Faktorenanalyse stellt die messtheoretische Komponente im SEM-Ansatz dar. Im Strukturmodell werden dagegen die Beziehungen der latenten Variablen zueinander spezifiziert (Bühner, 2006). Der Umgang mit nichtkontinuierlichen Variablen stellt für den SEM-Ansatz – und damit für die Durchführung einer CFA - ein Problem dar, für das keine klare Lösung zur Verfügung steht (Glöckner-Rist & Hoijtink, 2003).

Im Gegensatz dazu wurden die Modelle der so genannten Item-Response-Theorie (IRT) speziell für die Verwendung bei binären oder ordinalen Indikatorvariablen entwickelt, wobei von probabilistischen, nicht-linearen Beziehungen zwischen den beobachteten Indikatoren und den latenten Variablen ausgegangen wird (Glöckner-Rist & Hoijtink, 2003). Im Gegensatz zur CFA im SEM-Ansatz besteht das Hauptanliegen der IRT-Modelle darin zu spezifizieren, wie der Ausprägungsgrad einer latenten Eigenschaft auf der einen Seite und die Eigenschaften eines Items auf der anderen Seite zur Itemantwort einer Person beitragen. Die Wahrscheinlichkeit für die Wahl

einer bestimmten Antwortkategorie wird graphisch dargestellt mit der Item-Response-Funktion (IRF) bzw. Item-Charakteristik-Kurve (ICC), welche die nicht-lineare Regression einer Antwortwahrscheinlichkeit auf die latente Eigenschaft (trait) wiedergibt.

Viele Ziele der IRT sind denen der CFA ähnlich: so können IRT-Modelle z.B. genutzt werden, um die latenten Dimensionen ordinaler Item-Antworten zu explorieren und die psychometrischen Eigenschaften eines Tests zu überprüfen (Brown, 2006). Die Verwendung von IRT-Modellen bietet durch deren Fähigkeit im Umgang mit binären bzw. polytomen Items einen großen Vorteil gegenüber dem linearen SEM -Ansatz. Auf der anderen Seite bringt der SEM-Ansatz u.a. den Vorzug mit sich, die Beziehungen zwischen manifesten und latenten Variablen modellieren zu können (CFA). Eine Kombination beider Ansätze stellt somit eine wünschenswerte Lösung dar, um eine konfirmatorische Faktorenanalyse bei ordinalen Daten durchführen zu können. Dass dies möglich ist, konnte in verschiedenen Arbeiten belegt werden, in denen die formale Äquivalenz von IRT- und SEM-Messmodellen für dichotome und polytome Itemantworten bewiesen wurde (Glöckner-Rist & Hoijtink, 2003; Takane & de Leeuw, 1987).

Verwendet werden kann dabei als IRT-Modell das multidimensionale zwei-Parameter „normal ogive model“ (McDonald, 1981). Die Wahrscheinlichkeit für eine bestimmte Itemantwort hängt von der zu erfassenden Merkmalsausprägung der Person und den Eigenschaften des jeweiligen Items ab, wobei zwei Item-Parameter berücksichtigt werden: die Diskriminationsfähigkeit des Items und ein Schwierigkeitsparameter (=Lage des Items auf der latenten Variable: threshold). Dabei erhält eine Person  $i$  einen Wert  $\theta_i$  (=Faktorwert) auf jeder latenten Dimension (=Faktor, latente Variable) und für jedes Item  $j$  gibt es einen Diskriminationswert  $\alpha_j$  für jede latente Dimension (entspricht der Trennschärfe des Items in der klassischen Testtheorie). Der Itemdiskriminationsparameter  $\alpha_j$  entspricht der Faktorladung im CFA- bzw. EFA-Modell, da durch den entsprechenden Wert in beiden Fällen die Beziehung zwischen der latenten Variable und der Indikatorvariable repräsentiert wird (Brown, 2006). Graphisch wird der Diskriminationswert durch die Steigung der IRF-Kurve dargestellt. Für jedes (dichotome) Item wird weiterhin ein Schwellenwert (=threshold)  $\tau_i / \mu_i$  angegeben, welcher in Beziehung steht zu dem aus der klassischen Testtheorie stammenden Begriff der Item-Schwierigkeit, d.h beide Werte können durch eine Transformation jeweils ineinander überführt werden (Glöckner-Rist & Hoijtink, 2003). Bei Items mit mehreren Antwortmöglichkeiten (z.B. Items einer Likert-Skala) werden entsprechend mehrere

Schwellenwerte berechnet (Anzahl der Schwellen/thresholds = Anzahl der Antwortmöglichkeiten - 1), die jeweils den Punkt auf der horizontalen Merkmalsachse (=latente Variable) angeben, an dem die Wahrscheinlichkeit  $p$  dem Wert 0.5 entspricht, dass die nächst höhere Antwortkategorie gewählt wird.

Die Parameter eines multidimensionalen Zwei-Parameter-Normalogivenmodells sind also interpretierbar wie die Parameter linearer Modelle der Faktorenanalyse.

Durch die Integration von IRT-Modellen in einen erweiterten Strukturgleichungs-Modellierungsansatz können die Vorzüge beider Ansätze nutzbar gemacht werden, z.B. für die Überprüfung von Messinstrumenten. Zur Verfügung gestellt wird ein solch erweiterter SEM-Ansatz durch das Programm Mplus (Muthén & Muthén, 1998-2006), das für alle folgenden Berechnungen verwendet wird. Somit können in Mplus mit Hilfe des multidimensionalen Zwei-Parameter-Normalogivenmodells sowohl explorative (EFA), als auch, durch die Einbettung in ein erweitertes Strukturgleichungsmodell, konfirmatorische Faktorenanalysen (CFA) mit ordinalen Daten durchgeführt werden. Entscheidend dabei ist die Annahme, dass den nicht-kontinuierlich verteilten manifesten Variablen kontinuierlich verteilte latente Variablen zugrunde liegen<sup>19</sup>.

### 5.2.1.5. Umgang mit missing data

Ein weiteres wichtiges Thema stellt die Frage nach dem richtigen Umgang mit lückenhaften Datensätzen („missing data“) dar. Das Fehlen einzelner Itemantworten ist ein Sachverhalt, der die Auswertung von Datensätzen erschwert, aber in der Praxis kaum zu vermeiden ist. Häufig verwendete Strategien, wie z.B. „listwise deletion“, haben sich als problematisch erwiesen, da sie mit einem Verlust statistischer power und einer Verzerrung in der Schätzung der Parameter, der Standardfehler und der Teststatistik einhergehen (Allison, 2003). In den letzten Jahren wurden daher neuere Ansätze entwickelt, die in der Lage sind, auch die Datensätze mit fehlenden Werten zu berücksichtigen<sup>20</sup>. Auch das Programm Mplus ermöglicht durch die Verwendung solch neuer Schätzverfahren, lückenhafte Datensätze ohne Informationsverlust in die Auswertung mit einzu-

---

19 Die Beziehungen zwischen den beobachteten, ordinalen Variablen und den kontinuierlichen latenten Variablen werden dabei durch eine Reihe von Probit-Regressionsgleichungen dargestellt (Muthén & Muthén, 1998-2006).

20 Voraussetzung dafür ist, dass das Zustandekommen der fehlenden Werte nicht von den beobachtbaren Variablen abhängt und damit nicht systematisch ist. Eine der Bedingungen „missing completely at random (MCAR)“ oder „missing at random“ (MAR) muss erfüllt sein.



beziehen (Muthén & Muthén, 1998-2006), wobei sich eine Überprüfung anhand der Verteilungsmuster der missing data empfiehlt<sup>21</sup>.

### 5.2.2. Explorative Faktorenanalyse

Entsprechend der bereits kurz skizzierten Vorgehensweise, soll der nach der ersten Vor-Selektion aus 80 Items bestehende Datensatz in einem ersten faktorenanalytischen Auswertungsschritt einer explorativen Faktorenanalyse unterzogen werden. Im Vorfeld einer explorativen Faktorenanalyse bedarf es keiner Vorab-Spezifizierung hinsichtlich der Anzahl der latenten Variablen und der Art der Beziehungen zwischen den Indikatoren und der latenten Variablen, d.h. der Faktorladungen (Brown, 2006).

Die Durchführung einer EFA mit den vorliegenden Daten erscheint im ersten Moment im Widerspruch zum bisherigen Vorgehen zu stehen, da ja durchaus Annahmen über die Zusammenhänge zwischen den Indikatoren (= Items) und den latenten Variablen (= a-priori-Skalen) bestanden - die Items wurden ja in der Absicht generiert, Indikatoren für die latenten Variablen darzustellen. Dem ist allerdings entgegen zu halten, dass diese Vorannahmen über die Zusammenhänge zwischen den Indikatorvariablen zunächst auf theoretischen Überlegungen basierten und noch einer empirischen Bestätigung entbehren. Aus diesem Grund erscheint es angebracht, zunächst eine Faktorenanalyse durchzuführen, die keine Vorannahmen beinhaltet (EFA) und somit auch der Möglichkeit Raum lässt, dass sich eine andere Konstellation latenter Dimensionen zeigt als a-priori angenommen. Dennoch stellen natürlich die a-priori-Skalen die Grundlage dar, auf deren Basis die gewonnenen Faktorenlösungen inhaltlich interpretiert werden sollen.

Die Analyseeinheiten für die EFA stellen die Korrelationen (bzw. Kovarianzen) zwischen den Items dar<sup>22</sup>. Mit Hilfe der EFA wird daraus eine Matrix aus Faktorladungen ( $\Lambda$ ) generiert, die am besten die Korrelationen zwischen den Items erklärt (vgl. Brown, 2006). Um diese zu erhalten, muss zunächst eine geeignete Methode zur Schätzung des Faktorenmodells gewählt werden,

---

21 In Mplus wird dafür in der Sektion „covariance coverage“ der Outputdatei der Anteil der Daten angegeben, der für jeden Indikator und für jedes Indikatorenpaar zur Verfügung stand (=Kovarianz-Deckung). Solange dieser Wert 50% nicht unterschreitet, gibt es keine Probleme bei der Parameterschätzung (vgl. Brown, 2006). Zudem lässt sich über die angegebenen Häufigkeitsverteilungen der missing data (=“missing data pattern frequencies“) ein Eindruck der Verteilungsmuster gewinnen.

22 Im Falle ordinaler Daten handelt es sich dabei um polychorische Korrelationen. Dabei werden die Zusammenhänge zwischen je zwei unbeobachteten, kontinuierlichen Variablen auf der Grundlage beobachteter ordinaler Variablen geschätzt.

dann muss die angemessene Anzahl der Faktoren bestimmt werden, sowie eine geeignete Rotationsstechnik, sofern man nicht von unkorrelierten Faktoren ausgeht.

### **5.2.2.1. Wahl einer geeigneten Schätzmethode zur Faktorenextraktion**

Wie bereits angemerkt spielt das Skalenniveau und die Verteilungsform der Daten eine entscheidende Rolle bei der Auswahl des richtigen Schätzverfahrens, da viele Verfahren eine multivariate Normalverteilung der Daten erfordern. Ein Schätzalgorithmus, der für die Anwendung bei kategorialen Indikatoren geeignet ist, stellt der „weighted least squares mean and variance adjusted“-Schätzer dar (WLSMV, Muthén & Muthén, 1998-2006). Der WLSMV-Schätzer liefert robuste Schätzungen der Standardfehler, wobei eine robuste asymptotische Kovarianz-Matrix verwendet wird, sowie eine Mittelwert- und Varianz angepasste  $\chi^2$ -Teststatistik. Im Gegensatz zu anderen WLS-Schätzern führt der WLSMV-Schätzer auch bei kleinen Stichproben zu stabilen Parameterschätzungen (Flora & Curran, 2004).

### **5.2.2.2. Bestimmung der Faktorenanzahl**

Die Bestimmung der Anzahl der relevanten latenten Dimensionen/Faktoren (Faktorenextraktion) stellt den kritischsten Schritt innerhalb der explorativen Faktorenanalyse dar, da sowohl eine „Unterfaktorisierung“, als auch eine „Überfaktorisierung“ die Gültigkeit des resultierenden Faktorenmodells und seiner Parameterschätzungen beeinträchtigt (Brown, 2006). Die Konsequenzen der „Überfaktorisierung“ werden allerdings als weniger schwerwiegend angesehen als die Folgen der „Unterfaktorisierung“ (z.B. Fabrigar et al., 1999). Die Entscheidung für die Anzahl der Faktoren sollte in erster Linie auf der Basis gründlicher Überlegungen hinsichtlich der Interpretierbarkeit/Plausibilität der Faktoren und unter Berücksichtigung des theoretischen Hintergrundes, also der theoretisch begründeten Faktorenanzahl, erfolgen (vgl. Bühner, 2006).

Zu bedenken ist dabei auch, dass sich so genannte „triviale Faktoren“ bilden können, die auf Methodeneffekte zurückzuführen sind, z.B. durch die Gruppierung sehr ähnlich oder negativ formulierter Items (Brown, 2006; Podsakoff et al., 2003; Horan et al., 2003). In verschiedenen Studien konnte gezeigt werden, dass die Ladungen auf solchen Konstrukt-irrelevanten Methodenfaktoren mit den Ladungen auf den substantiellen Faktoren vergleichbar oder sogar höher als diese sind (z.B. Harvey et al., 1985) und mit der Re-Formulierung der Items in positive Richtung verschwinden (Idaszak & Drasgow, 1987). Weiterhin ist darauf zu achten, dass die Faktoren „gut definiert“ sind, d.h. dass sie durch drei bzw. vier oder mehr Indikatoren repräsentiert werden, da sie sich sonst als instabil erweisen können (Fabrigar et al., 1999).

Neben inhaltlichen Überlegungen stehen statistische Regeln und Kennwerte zur Verfügung, die bei der Faktorenauswahl hilfreich sein können. Zu nennen ist hier z.B. die Regel „Eigenwert  $>1$ “, auch Kaiser-Guttman-Regel genannt<sup>23</sup>. Obwohl diese Regel häufig Verwendung findet, ist ihre Anwendung mit einigen Problemen verbunden und führt oft zu Über- und Unterschätzung der Faktorenanzahl in Abhängigkeit von der Variablenanzahl, der Reliabilität der Faktoren, des Verhältnisses von Indikatoren zu Faktoren u.a., weswegen einige Forscher von der Anwendung dieser Regel abraten (vgl. Preacher & McCallum, 2003).

Einen anderen populären Ansatz stellt der Scree-Test dar<sup>24</sup>. Der Scree-Test hat sich in der Praxis bewährt, wurde aber aufgrund seiner Subjektivität ebenfalls oft kritisiert (Bühner, 2006).

Verwendet man für die EFA den ML- (maximum likelihood) oder WLSMV-Schätzer, so hat man den großen Vorteil, dass zusätzlich zu den oben genannten Verfahren auch „goodness-of-fit“-Indices zur Verfügung stehen, die zur Bestimmung der angemessenen Faktorenanzahl herangezogen werden können (vgl. Preacher & McCallum, 2003). Die Goodness-of-fit-Kennwerte liefern Informationen darüber, wie gut die Parameter des Faktorenmodells in der Lage sind, die Stichprobenkorrelationen zu reproduzieren. Unter dem WLSMV-Schätzer werden im Rahmen der EFA folgende Fit-Indices angegeben:  $\chi^2$ , RMSEA (=root mean square error of approximation) und RMR (= root mean square residual; durchschnittlicher Residualwert). Der  $\chi^2$ -Wert sollte nicht signifikant werden ( $p > .05$ ,  $.01$ ), da ein nicht-signifikanter Wert für eine geringe Diskrepanz zwischen der beobachteten Kovarianzmatrix  $\Sigma$  und der implizierten Kovarianzmatrix  $\Sigma(\theta)$  spricht. Da der  $\chi^2$ -Test allerdings nur eingeschränkt interpretierbar ist, da er abhängig ist von der Stichprobengröße, sollten in erster Linie die anderen Fit-Indices berücksichtigt werden (Brown, 2006)<sup>25</sup>.

---

23 Eigenwerte der Faktoren = Varianz der Indikatoren, die sukzessive durch die Faktoren erklärt wird, d.h. ein Eigenwert wird gebildet aus der Summe der quadrierten Ladungen über alle Items auf einem Faktor. Somit ist der Eigenwert ein Maß für die „Wichtigkeit“ eines Faktors. Gemäß der Kaiser-Guttman-Regel sollte man für die Extraktion der Faktoren die Anzahl der Eigenwerte  $>1$  auf Grundlage der nicht-reduzierten Korrelationsmatrix bestimmen und ausgehend von dieser Menge die Anzahl nichttrivialer latenter Dimensionen festlegen (Brown, 2006).

24 Beim Scree-Test handelt es sich nicht um einen Test im eigentlichen Sinne, sondern eher um eine Daumenregel. Dabei sucht man nach einem bedeutsamen Eigenwertabfall (Knick) im Screeplot (graphische Darstellung Höhe der Eigenwerte) und betrachtet in der weiteren Auswertung nur diejenigen Faktoren, deren Eigenwerte vor dem Knick liegen.

25 Häufig wird zusätzlich der  $\chi^2$ -Wert durch die Freiheitsgrade geteilt, wobei ein Wert  $<2$  als gut angesehen wird. Für den RMR gibt es keinen allgemein anerkannten Cut-Off-Wert, wobei u.a. aber ein Wert  $< .05$  vorgeschlagen wird (z.B. Byrne, 1989). Für den RMSEA gilt ein Wert  $< .06$  als gut (Hu & Bentler, 1999).

Ziel ist es, bei der Verwendung der Goodness-of-Fit-Ansätze im Rahmen einer EFA, diejenige Faktorenlösung zu finden, die sowohl die beobachteten Korrelationen besser reproduzieren kann als sparsamere Lösungen mit weniger Faktoren, als auch vergleichbare Werte zu komplexeren Lösungen mit mehr Faktoren aufweist (Brown, 2006). Daher ist es notwendig verschiedene Lösungen mit unterschiedlicher Faktorenanzahl zu vergleichen, wobei immer die Plausibilität und Interpretierbarkeit der Faktoren als entscheidende Auswahlkriterien mitberücksichtigt werden müssen.

Neben den bisher genannten Kriterien ist es zudem für die Bestimmung der Faktorenzahl sinnvoll, die Ladungsverteilungen der Items zu berücksichtigen.

Sofern verschiedene Methoden der Faktorenextraktion zu inkonsistenten Ergebnissen führen, kann es auch sinnvoll sein, zunächst die Rotation mit verschiedenen Lösungen, d.h. mit unterschiedlicher Faktorenzahl, durchzuführen und die Entscheidung für die angemessene Anzahl der Faktoren in erster Linie aufgrund der Interpretierbarkeit der Lösung zu treffen (Preacher & McCallum, 2003).

### 5.2.2.3. Wahl der Rotationstechnik

Rotationsmethoden werden eingesetzt um eine gut zu interpretierende Lösung aus einer unendlich großen Menge von Alternativlösungen zu finden, wobei das Kriterium der Einfachstruktur nach Thurstone (1947) angewendet wird<sup>26</sup>. Die Rotation verändert nicht den Model-Fit einer Lösung (z.B. den  $\chi^2$ -Wert), da es sich nur um eine mathematische Transformation, nämlich die Rotation im multidimensionalen Raum handelt.

Unterschieden werden dabei orthogonale und oblique Rotationstechniken, wobei im ersten Fall die Einschränkung getroffen wird, dass die Faktoren unkorreliert sein sollen (=90°-Winkel zwischen den Faktor-Achsen im multidimensionalen Raum), während im zweiten Fall die Faktoren miteinander korrelieren dürfen<sup>27</sup>. Eine oblique Rotation ist zu empfehlen, da sie eine realistischere Repräsentation der Zusammenhänge zwischen den Faktoren liefert. Zudem würde sich, im

---

26 Demnach sollte zum einen jeder Faktor definiert sein durch eine Teilmenge von Indikatoren, die auf diesem Faktor hoch laden, und zum anderen sollte jeder Indikator idealerweise hoch auf einem Faktor („hohe Hauptladung“) und nur geringfügig, bzw. nicht-bedeutsam auf anderen Faktoren laden („geringe Nebenladungen“). Als bedeutsam werden dabei meist Ladungen  $\geq .30$  bis  $.40$  angesehen (Brown, 2006).

27 Die Annahme der Unkorreliertheit der Faktoren und die damit verbundene Verwendung orthogonaler Rotationsmethoden sind weit verbreitet und populär (z.B. auch durch die Vorgabe in vielen Statistikprogrammen), aber in den meisten Fällen schwer zu rechtfertigen: solange es kein Wissen darüber gibt, wie die Faktoren zueinander in Beziehung stehen, gibt es keinen Grund zu der Annahme, dass sie vollständig unabhängig sind (Preacher & McCallum, 2003).

Falle tatsächlich unkorrelierter Faktoren, die Lösung einer orthogonalen Rotation nicht von derjenigen einer obliquen Rotation unterscheiden (Floyd & Widaman, 1995). Ein weiteres Argument für die oblique Rotation stellt die Beobachtung dar, dass oblique Lösungen einer EFA sicher besser in eine CFA verallgemeinern lassen als orthogonale Lösungen, die üblicherweise zu einem schlechten Model-Fit führen (Brown, 2006). Empfohlen wird für die oblique Rotation häufig die Promax-Rotation als Methode der Wahl (z.B. Bühner, 2006).

### 5.2.2.4. Evaluation und Optimierung der Faktorenlösung

Zur Sicherung der Qualität der Faktorenlösung sollten die Faktoren abschließend hinsichtlich ihrer inhaltlichen Bedeutung bzw. ihrer empirischen Relevanz überprüft werden. Neben der inhaltlichen Interpretierbarkeit sollte dabei v.a. darauf geachtet werden, ob sich Methodenfaktoren gebildet haben, z.B. durch ähnlich formulierte oder negativ formulierte Items (vgl. Brown, 2006). Im Anschluss an diese Überprüfung sollte die Faktorenlösung optimiert werden. Dafür kommen zwei Prozesse in Frage:

a) die Eliminierung schlecht definierter Faktoren: dies betrifft sowohl Faktoren, die auf Methodeneffekte zurückzuführen sind, als auch Faktoren, auf denen nur zwei oder drei Items bedeutsam laden bzw. Faktoren, die nur durch Items mit geringer Ladung definiert sind.

b) die Eliminierung problematischer Indikatoren: dies gilt für Items mit nicht-bedeutsamen Ladungen auf allen Faktoren und Items mit bedeutsamen Ladungen auf mehr als einem Faktor.

Es gibt keine allgemein gültige Konvention, ab wann eine Ladung als „nicht-bedeutsam“ anzusehen ist, wobei häufig als Kriterium Werte  $<.30-.40$  angegeben werden (vgl. Brown, 2006). Für diese Untersuchung wurde der Cut-Off-Wert auf  $<.40$  festgelegt. Hinsichtlich der Mehrfach-Ladungen kann zur Itemselektion das so genannte Fürntratt-Kriterium herangezogen werden<sup>28</sup>. Gemäß dem Fürntratt-Kriterium liegt der Cutoffwert für die Eliminierung von Items bei  $a^2/h^2 <.50$ . Zusätzlich sollte bei der Itemselektion auf der Basis statistischer Kriterien unbedingt die Inhaltsvalidität mit berücksichtigt werden: wenn das Item für das Konstrukt wichtig erscheint, sollte es trotz geringer Ladung beibehalten werden (vgl. Bühner, 2006).

### 5.2.2.5. Zusammenfassung

Aus den vorangegangenen Ausführungen lässt sich folgende Vorgehensweise ableiten:

- **Zusammenfassung der Antwortkategorien:** die Daten sollen als ordinale, kategoriale Daten behandelt werden. Da, wie in Tab. 21 aufgeführt, nicht bei allen Indikatoren alle

---

<sup>28</sup> Demnach sollte die quadrierte Hauptladung ( $a^2$ ) eines Items mindestens 50% seiner Kommunalität ( $h^2 = \sum a^2$ , Quadratsumme aus allen Faktorladungen eines Items) ausmachen.

sieben Antwortkategorien von den Probanden genutzt wurden, für die Berechnung von  $\chi^2$ -Tests aber eine Kategorienhäufigkeit von mind. 5% gefordert wird (vgl. Bortz, 2005), sollen zunächst Antwortkategorien so zusammengefasst werden, dass für jedes Item alle Antwortkategorien mind. 5% der Antworten beinhalten.

- **Faktorextraktion:** die Extraktion der Faktoren soll mit Hilfe des Programmes M-plus (Version 4.2) unter Verwendung des WLSMV-Schätzers erfolgen.
- **Überprüfung der missing data-Verteilung:** die Überprüfung soll anhand der „covariance coverage“-Werte (Kovarianz-Deckung) durchgeführt werden.
- **Faktoreselektion:** die Auswahl der Faktoren soll mithilfe verschiedener Kriterien (Kaiser-Guttman-Regel, Scree-Test, Goodness-of-model-fit), unter Berücksichtigung der theoretischen Annahmen (5 a-priori-Faktoren) erfolgen. Sollte sich keine eindeutige Lösung abzeichnen, soll zunächst eine Faktorrotation durchgeführt und im Anschluss die Entscheidung für die passende Faktorenlösung aufgrund der Interpretierbarkeit der Faktoren getroffen werden.
- **Faktorrotation:** die Rotation soll oblique mit der Promax-Methode erfolgen.
- **Evaluation der Faktorenlösung:** es soll eine Prüfung durchgeführt werden, ob alle Faktoren eine substantielle Bedeutung aufweisen und ob sich keine trivialen Faktoren (z.B. durch Methodeneffekte) gebildet haben.
- **Optimierung der Faktorenlösung:**
  - a) Eliminierung schlecht definierter Faktoren, die nur durch zwei oder drei Items mit bedeutsamen Ladungen oder nur durch Items mit geringer Ladung definiert sind.
  - b) Eliminierung problematischer Indikatoren (Items mit hohen Ladungen auf mehr als einem Faktor: Fürntratt-Kriterium ( $a^2/h^2 < .50$ .)); Items mit nicht-bedeutsamen Ladungen ( $< .40$ ) auf allen Faktoren), unter Berücksichtigung der Inhaltsvalidität.

### 5.2.3. Konfirmatorische Faktorenanalyse

Wie bereits kurz dargelegt, ist es das Ziel einer konfirmatorischen Faktorenanalyse (CFA), ein vorab bestimmtes Modell zu testen, das die Zusammenhänge zwischen latenten Variablen (Faktoren) und Indikatorvariablen (Items) beschreibt. Die CFA ist damit in erster Linie ein strukturprüfendes Verfahren. Das spezifizierte Modell wird dabei in Form einer Hypothesentestung auf





seine Übereinstimmung mit der empirischen Kovarianz-/bzw. Korrelationsmatrix hin überprüft (vgl. Moosbrugger & Kelava, 2007)<sup>29</sup>.

Neben der implizierten, d.h. der aus dem spezifizierten Modell errechneten, Kovarianzmatrix werden zudem folgende Modellparameter über wiederholte Rechenschritte geschätzt: nicht-standardisierte und standardisierte Regressionsgewichte (= Faktorladungen der Indikatoren auf den latenten Variablen), Kovarianzen/Korrelationen der latenten Variablen (=Maß für die Zusammenhänge zwischen den Faktoren) und Fehlervarianzen (=Varianz, die nicht durch die Prädiktoren aufgeklärt werden kann). Um diese so genannten „freien Parameter“ schätzen zu können, wird das spezifizierte Modell in Grundgleichungen zerlegt. Die Schätzung beginnt mit so genannten Startwerten, die als vorläufige Werte eingesetzt werden. Diese Werte werden entweder vom Programm automatisch vorgegeben oder vom Anwender eingesetzt (vgl. Bühner, 2006).

Grundlegend wird davon ausgegangen, dass die Indikatoren nur auf einer latenten Variablen laden und mögliche Gemeinsamkeiten mit anderen Indikatoren auf die Korrelationen der Konstrukte (=der latenten Variablen) zurückführbar sind. Daher werden in der Regel die Nebenladungen der Indikatoren, ebenso wie die Fehler-Kovarianzen auf null fixiert, in der Annahme, dass alle Messfehler zufallsbedingt auftreten.

### Graphische Darstellung

Für die graphische Darstellung der Modelle (in Form von Pfadmodellen) wird normalerweise folgende Notation verwendet (vgl. Bühner, 2006):

	= latente Variable/Faktor ( $\xi$ )
	= beobachtete Variable/Indikator/Item ( $x_1, x_2, \dots$ )
	= gerichtete Beziehung: Regressionsgewicht (in der CFA: Faktorladung)
	= ungerichtete Beziehung: Korrelation oder Kovarianz
$\delta_1, \delta_2, \delta_3 \dots$	= Fehlerterme

---

<sup>29</sup> Der Vergleich zwischen implizierter Kovarianz-/Korrelationsmatrix ( $\Sigma$ ) und beobachteter Kovarianz-/ Korrelationsmatrix ( $S$ ) geschieht anhand der gewichteten Differenz beider Matrizen (=Residualmatrix), die in iterativen Rechenschritten ermittelt wird. Wenn die gewichtete Differenz zwischen beiden Matrizen nicht mehr verringert werden kann, stoppt der Iterationsprozess und die Schätzung hat konvergiert (vgl. Bühner, 2006). Die Bestimmung und Gewichtung der Differenz erfolgt in Abhängigkeit von der gewählten Schätzmethode.

### 5.2.3.1. Voraussetzungen

Um eine CFA durchführen zu können, müssen zwei Bedingungen erfüllt sein. Die erste betrifft die so genannte Modell-Identifikation. Dies bedeutet, dass genügend Informationen vorliegen müssen, um alle Parameter schätzen und das Modell eindeutig berechnen zu können<sup>30</sup>.

Die zweite Bedingung betrifft die Metrik der latenten Variablen. Da diese nicht „von sich aus“ über eine Metrik verfügen, muss diese im Vorfeld festgelegt werden. Dies kann z.B. dadurch erfolgen, dass für jede latente Variable die unstandardisierte Ladung eines Indikators auf 1 festgesetzt wird. Dieser Indikator wird somit zum „Referenzindikator“, weshalb dafür auch ein für das latente Konstrukt bedeutsamer Indikator ausgewählt werden sollte (vgl. Bühner, 2006). Eine andere Möglichkeit besteht darin, die Varianz der latenten Variablen auf 1 festzulegen. Dieses Vorgehen bringt den Vorzug mit sich, dass damit die Kovarianzen zwischen den latenten Variablen den Korrelationen entsprechen.

### 5.2.3.2. Festlegung der Schätzmethode und des Rahmenmodells

Die verschiedenen Schätzmethode unterscheiden sich dahingehend, mit welcher Funktion die Abweichung zwischen implizierter und beobachteter Kovarianzmatrix berechnet wird (Diskrepanzfunktion, „fitting function“). Entscheidend für die Wahl des Schätzverfahrens ist, entsprechend dem Vorgehen bei der explorativen Faktorenanalyse, die Skalenqualität und die Verteilungsformen der Indikatoren. Bei ordinalen/kategorialen Daten wird üblicherweise („per default“) der WLSMV-Schätzer („weighted least squares mean and variance adjusted“, Muthén & Muthén, 1998-2006) verwendet.

In M-plus wird ein Rahmenmodell für die CFA bei ordinalen, nicht normalverteilten Indikatorvariablen bereitgestellt, das die Verwendung latenter kontinuierlicher und normal verteilter Variablen ( $y^*$ ) vorsieht, die „hinter“ den beobachteten Variablen stehen (vgl. Muthén & Asparouhov, 2002). Für die Korrelationsmatrix  $S$  werden anstatt der Korrelationen der beobachteten Variablen daher auch die Korrelationen der zugrunde liegenden  $y^*$ -Variablen verwendet, die über die Mittelwerte der Schwellen (threshold)-Parameter berechnet werden<sup>31</sup>.

---

30 Die Anzahl der bekannten Parameter, d.h. in erster Linie der empirischen Varianzen und Kovarianzen, muss größer sein als die Anzahl der zu schätzenden freien Parameter, was bedeutet, dass das Modell eine positive Anzahl von Freiheitsgraden (df) aufweisen muss. In diesem Fall spricht man von einem „überidentifizierten Modell“ (vgl. Brown, 2006). Bei nicht-identifizierten Modellen (Anzahl der freien Parameter > Anzahl der bekannten Parameter) kann eine Identifikation auch durch weitere Parameterrestriktionen erzielt werden, d.h., dass weitere Parameter auf einen bestimmten Wert festgelegt oder bestimmte Parameter gleichgesetzt werden (z.B. Faktorladungen).

31 Im Gegensatz zur CFA mit kontinuierlichen Indikatoren werden die Residual-Varianzen der kategorialen Indikatoren nicht geschätzt, weshalb die Messfehler des CFA-Modells ( $\theta$ ) keine frei geschätzten Parameter darstellen,



### 5.2.3.3. Modell-Evaluation

Durch den Modelltest wird geprüft, ob die modelltheoretische Kovarianz- bzw. Korrelationsmatrix signifikant von der beobachteten Kovarianz- bzw. Korrelationsmatrix abweicht. Als Prüfgröße wird dabei der „value of the fitting function“ (F) verwendet: je kleiner dieser Wert wird, umso besser gibt die implizierte Kovarianz-/Korrelationsmatrix die beobachtete wieder. Dieser Wert geht ein in die Berechnung des  $\chi^2$ -Wertes:  $\chi^2 = N * F$  (vgl. Brown, 2006). Dieser  $\chi^2$ -Wert wird dann mit Hilfe der  $\chi^2$ -Verteilung auf Signifikanz geprüft, wobei die Anzahl der Freiheitsgrade (df) aus der Anzahl der beobachteten Parameter (b) und der Anzahl der freien Parameter (f) errechnet wird<sup>32</sup>. In diesem  $\chi^2$ -Test werden die Hypothesen geprüft, ob das Modell zu den empirischen Daten passt (H0) oder ob das Modell von der Datenstruktur abweicht (H1). Die Nicht-Signifikanz des  $\chi^2$ -Wertes ist damit ein Beleg für einen exakten Modell-Fit. Der  $\chi^2$ -Wert wurde als erster Fit-Index für die Modellpassung entwickelt, geriet allerdings als Prüfverfahren bald in die Kritik, da er einige Nachteile aufweist<sup>33</sup> (vgl. Brown, 2006).

Daher wurde eine Vielzahl anderer Kennwerte für die Modellpassung, so genannte „Fit-Indizes“, entwickelt, die weniger anfällig und mit weniger strikten Annahmen verbunden sind. Der  $\chi^2$ -Wert und der dazu gehörige p-Wert werden zwar berechnet, u.a. auch weil der  $\chi^2$ -Wert in die Berechnung vieler Fit-Indizes mit eingeht, und sie müssen immer in der Ergebnisdarstellung mit angegeben werden. Die Entscheidung für die Annahme bzw. Zurückweisung eines Modells wird aber meist aufgrund der anderen Fit-Indizes getroffen (vgl. Brown, 2006).

---

sondern den Wert der Differenz von 1 minus dem Produkt aus quadrierter Faktorladung und Faktorvarianz annehmen ( $\theta = 1 - \lambda^2 \phi$ , vgl. Brown, 2006). Die quadrierten komplett standardisierten Faktorladungen geben jeweils den Varianzanteil von  $y^*$  an, der durch den latenten Faktor erklärt wird, und nicht den aufgeklärten Varianzanteil der beobachteten kategorialen Variable.

32  $df = \{(b) * (b+1) / 2\} - f$  (vgl. Bühner, 2006).

33 In vielen Fällen (z.B. bei einer geringen Stichprobengröße oder nicht-normalverteilten Daten) kommt es vor, dass die zugrunde liegende Verteilung gar nicht  $\chi^2$ -verteilt ist. Weiterhin ist der  $\chi^2$ -Test von der Stichprobengröße abhängig, was bei größer werdendem N zu einer vermehrten Zurückweisung eigentlich passender Modelle führt, und schließlich basiert der  $\chi^2$ -Test auf der äußerst strikten Annahme von  $S = \Sigma$ .

Bei den Fit-Indizes existieren verschiedene Untergruppen<sup>34</sup>. Jede Unterkategorie von Fit-Indizes stellt unterschiedliche Informationen über den Modell-Fit bereit, weshalb es ratsam ist, mindestens einen Index aus jeder Untergruppe zu berücksichtigen und anzugeben (vgl. Brown, 2006). Folgende Fit-Indices sollen verwendet werden<sup>35</sup>:

### *Komparative Fit-Indizes:*

**CFI** (Comparative-Fit-Index; Bentler, 1990)/**TLI** (Tucker-Lewis-Index; Tucker & Lewis, 1973): Für diese beiden populären und zuverlässigen Indizes liegt der Wertebereich in der Regel zwischen null und eins, wobei der TLI nicht normiert ist und daher auch Werte  $>1.0$  annehmen kann. Generell gilt, dass höhere Werte für einen besseren Modell-Fit sprechen. Als Cut-off-Wert wird für beide Indizes ein Wert  $>.95$  angegeben (Hu & Bentler, 1999).

### *Fit-Index mit Sparsamkeits-Korrektur:*

**RMSEA** (Root-Mean-Square-Error of Approximation; Steiger & Lind, 1980): hier steht ein hoher Wert für einen schlechten Modell-Fit. Je komplexer das Modell (=geringere Anzahl an df), umso höher der RMSEA und umso schlechter der Modell-Fit. Das bedeutet, dass beim RMSEA die Sparsamkeit eines Modells (=wenig Restriktionen, viele frei zu schätzende Parameter) belohnt wird. Als Cut-off-Wert wird für den RMSEA meist ein Wert  $<.06$  angegeben (vgl. Hu & Bentler, 1999), wobei ein Wert  $\leq .05$  als „close fit“ bezeichnet wird, also als fast perfekter Modell-Fit (vgl. Brown, 2006). Für Stichproben  $< 250$  wird allerdings auch ein Wert  $<.80$  als akzeptabel angesehen (vgl. Bühner, 2006).

### *Absoluter Fit-Index:*

**WRMR** (Weighted Root Mean Square Residual; Muthén and Muthén, 1998-2006)

---

34 Diese prüfen in unterschiedlicher Form entweder die Abweichung von einem „Nullmodell“ (=Independence-Modell), in dem alle Parameter mit Ausnahme der Varianzen der beobachteten Variablen auf Null fixiert sind, oder die Abweichung von einem Modell, das die Stichprobenkovarianz exakt repliziert (=Saturated-Modell) (vgl. Bühner, 2006). Diejenigen Fit-Indizes, die den Unterschied in der Daten-Reproduktion zwischen dem spezifizierten Modell und einem Null-Modell angeben, werden „komparative Fit-Indizes“ genannt. Fit-Indizes, die zum Vergleich ein saturiertes Modell heranziehen, werden als „absolute Fit-Indizes“ bezeichnet. Weiterhin gibt es noch Fit-Indizes, welche die Komplexität eines Modells mit berücksichtigen (Fit-Indizes mit „Sparsamkeits-Korrektur“).

35 Es werden dabei auch Cut-off-Werte angegeben, v.a. in Anlehnung an die Empfehlungen von Hu & Bentler (1999).

Dieser Index eignet sich besonders dann, wenn sich die Varianzen in der Stichprobe stark unterscheiden, wenn Modelle mit Mittelwerten und Schwellen (thresholds) verwendet werden, sowie bei fehlender Normalverteilung. Als Cut-off-Wert wird ein Wert  $<1.0$  empfohlen (Yu, 2002).

Neben den Fit-Indizes sollten bei der Prüfung der Modell-Güte allerdings noch weitere Aspekte berücksichtigt werden, welche die Interpretierbarkeit und Güte der Parameter-Schätzungen betreffen, wie z.B. das Fehlen von „out-of-range“-Parameter-Werten (wie z.B. einer negativen Fehler-Varianz oder standardisierte Faktorladungen  $>1$ ), sowie die statistische Signifikanz<sup>36</sup>, die Richtung und die Größe der Parameterschätzungen (Brown, 2006).

Neben den Faktorladungen der Indikatoren sollten zudem die Faktor-Interkorrelationen betrachtet werden. Unter Berücksichtigung des theoretischen Rahmens und der damit verbundenen Erwartungen bzgl. der Zusammenhänge zwischen den Faktoren, sollten die Faktor-Interkorrelationen auch in Hinblick auf die diskriminative Validität der latenten Konstrukte interpretiert werden. Nähern sich die Korrelationen dem Wert 1.0 an, so liegt die Vermutung nahe, dass die entsprechenden Faktoren keine distinkten Konstrukte repräsentieren. Als Cut-off-Kriterium für eine schlechte diskriminative Validität wird häufig ein Wert  $>.80$  bzw.  $.85$  angegeben (vgl. Cohen et al., 2003).

### 5.2.3.4. Modell-Modifikation („Respezifizierung“)

Die verschiedenen Fit-Indizes geben zwar Auskunft über die globale Passung des Modells, können aber keine genaueren Angaben darüber machen, an welchen Stellen die Spezifizierung des Modells ungenügend ist. Diese Überprüfung wird durch die so genannten „Modifikations-Indizes“ (M.I.) bzw. EPC-Werte möglich (Sörbom, 1989)<sup>37</sup>.

Beide Informationen (Modifikationsindex und EPC) sollte man heranziehen, um zu entscheiden, ob eine Neu-Spezifizierung tatsächlich eine bedeutsame Verbesserung erbringen könnte. Aller-

---

36 Die Signifikanz der Parameterschätzungen (z.B. der Faktorladungen) kann mit Hilfe der Standardfehler berechnet werden: der unstandardisierte Parameterwert („Estimate“/Est.) wird dafür durch den Standardfehler (SE) geteilt. Dieser Quotient (Est./SE) kann als z-Wert aufgefasst werden und somit ab einem Wert von  $\pm 1.96$  ( $\alpha = .05$ ) bzw.  $\pm 2.58$  ( $\alpha = .01$ ) oder mehr als statistisch signifikant gelten.

37 Modifikations-Indizes können für alle fixierten Parameter berechnet werden (z.B. für die auf null fixierten Nebenladungen der Indikatoren) und geben annäherungsweise an, um wie viel der f-Wert kleiner würde, wenn der entsprechende Parameter frei geschätzt würde. Ein Problem besteht allerdings darin, dass die Modifikations-Indizes, ähnlich wie die  $\chi^2$ -Werte, von der Stichprobengröße abhängig sind. Daher werden zusätzlich so genannte EPC-Werte (expected parameter change) für alle Modifikations-Indizes berechnet, die unabhängig von der Stichprobengröße in standardisierter oder unstandardisierter Form angeben, wie stark sich der entsprechende Wert des Parameters annähernd verändern würde, sofern der Parameter frei geschätzt würde.

dings sollte man nur dann eine solche Modell-Veränderung vornehmen, wenn dieser Schritt empirisch bzw. konzeptuell begründbar ist. Sofern die Re-Spezifizierung eines Modells lediglich auf Grundlage der Modifikation-Indizes vorgenommen wird, besteht die Gefahr, dass fehlerhafte Modelle spezifiziert werden (vgl. MacCallum, 1986). Jöreskog (1993) empfiehlt daher, die Modifizierung eines Modells mit der Freisetzung desjenigen Parameters zu beginnen, der den größten Modifikations-Index aufweist, unter der Voraussetzung, dass dieser Schritt inhaltlich interpretierbar ist. Sollte dies nicht der Fall sein, soll man sich den zweitgrößten Modifikations-Index vornehmen usw.<sup>38</sup>

Eine andere Form der Modell-Revision stellt die Eliminierung „schlechter“ Parameter dar. Dieses Vorgehen erscheint besonders im Kontext der Fragebogenkonstruktion von Bedeutung, da sich hiermit die Möglichkeit einer empirisch fundierten Item-Selektion bietet. Mit diesem Vorgehen werden zweierlei Ziele verfolgt: zum einen eine Verbesserung der Modellpassung, zum anderen aber auch eine statistisch begründete Kürzung des Fragebogens. Zu berücksichtigen sind hier als Selektionskriterien zum einen die statistische Signifikanz der Parameter (v.a. der Faktorladungen), zum anderen aber auch die Modifikationsindizes. Zeigen sich hier bei einem Item z.B. hohe Ladungen auf anderen Faktoren, sollte eher noch als eine Freisetzung des entsprechenden Parameters die Eliminierung des Indikators in Erwägung gezogen werden, wobei zur Entscheidung wiederum auch inhaltliche Überlegungen herangezogen werden sollten. Neben der fehlenden statistischen Signifikanz stellt auch die Höhe der Parameterwerte, d.h. in erster Linie der Faktorladungen, ein Selektionskriterium dar. Brown (2006) gibt als üblichen Richtwert für bedeutsame Faktor-Ladungen einen Wert von  $.30 / .40$  an, wobei er anmerkt, dass diese Richtlinie für die meisten Anwendungsgebiete der CFA wohl zu liberal ist. Ein Richtwert von  $\geq .50$  erschien für das vorliegende Projekt angemessen.

### **5.2.3.5. EFA im Rahmen einer CFA**

Die Durchführung einer explorativen Faktorenanalyse (EFA) im Vorfeld einer CFA dient zur Überprüfung der Anzahl der theoretisch hergeleiteten Faktoren bzw. Konstrukte und der Ladungsverteilungen der Indikatoren. Das im Rahmen der CFA vorab zu bestimmende Modell wird

---

<sup>38</sup> Sobald man beginnt, das Ausgangsmodell aufgrund der Modifikationsindizes neu zu spezifizieren, verlässt man streng genommen den Bereich konfirmatorischer Analysen, da man dann eigentlich explorativ und modellgenerierend vorgeht. Daher sollten alle auf diese Weise gewonnenen „neuen“ Modelle wiederum konfirmatorisch mit Hilfe einer unabhängigen Stichprobe überprüft werden (vgl. Brown, 2006). Auf diese Thematik soll im Diskussionsenteil Bezug genommen werden.

mithilfe der durch die EFA ermittelten Struktur spezifiziert. Die Faktoren der EFA werden dann als latente Variablen in das Modell für die CFA übernommen. Die Ladungen der Indikatoren auf den latenten Variablen können entweder, ausgehend von den Ergebnissen der EFA, vorgegeben oder als freie Parameter geschätzt werden. Damit stellen die Ergebnisse der EFA eine solide empirische Grundlage für die Spezifizierung des Modells im Rahmen der CFA dar. Dennoch kommt es häufig bei dieser Vorgehensweise zu schlecht passenden CFA-Lösungen, da es viele Möglichkeiten der „Miss-Spezifikation“ gibt, auf die in der EFA keine Hinweise zu finden sind (Brown, 2006).

Eine mögliche Fehl-Spezifizierung kann u.a. dadurch entstehen, dass die Gemeinsamkeiten von Indikatoren in Wirklichkeit auf korrelierte Fehler („correlated uniqueness“) statt auf einen gemeinsamen Faktor zurückgehen. Da es im Rahmen einer EFA nicht möglich ist, korrelierte Fehler (v.a. Methodeneffekte) der Indikatorvariablen zu erkennen und zu berücksichtigen, können die Ergebnisse der EFA in diesem Fall einen zusätzlichen Faktor nahelegen, obwohl es sich dabei eigentlich um ein Methodenartefakt handelt (vgl. Brown, 2006). Solch ein „künstlicher Faktor“ kann z.B. durch die Verwendung unterschiedlicher Erhebungsinstrumente, durch die unterschiedliche „Anfälligkeit“ von Items gegenüber Antworttendenzen oder durch sehr ähnlich oder negativ formulierte Items entstehen (vgl. Podsakoff et al., 2003).

Eine Möglichkeit, um die Struktur des Messmodells besser im Vorfeld der eigentlichen CFA explorieren zu können, bietet die so genannte E/CFA-Methode („exploratory factor analysis within the CFA framework“, Jöreskog, 1969; Jöreskog & Sörbom, 1979; Muthén and Muthén, 1998). Mithilfe dieser Methode können zusätzliche Informationen gewonnen werden, die für die Spezifizierung passender CFA-Lösungen hilfreich sind. Dabei werden innerhalb eines CFA-Modells dieselben Restriktionen vorgenommen wie im EFA-Modell, d.h. die Faktorvarianzen werden auf 1 fixiert (als „Metrik-Geber“), die Faktor-Kovarianzen werden frei geschätzt, ebenso wie die Faktorladungen der Indikatoren. Eine Ausnahme bilden dabei die Nebenladungen der so genannten „Anker-Items“ (=Items mit der höchsten Ladung pro Faktor und den niedrigsten Nebenladungen in der EFA), die auf null fixiert werden, um die Identifikation des Modells zu gewährleisten. Die Anzahl der Restriktionen entspricht dabei der Anzahl der latenten Variablen/Faktoren zum Quadrat ( $m^2$ ). Der entscheidende Unterschied zum „richtigen“ CFA-Modell besteht darin, dass weder die Kovarianzen der Fehler, noch die Nebenladungen der Indikatoren auf null fixiert werden, weshalb dieses Vorgehen Informationen über diese Parameter liefert, wie z.B. Hinweise auf

korrelierte Fehler (über die Modifikationsindizes/M.I.) und signifikante Nebenladungen ( $z$ -Test:  $\text{Est./SE} > \pm 1.96$  ( $\alpha = .05$ ) bzw.  $\pm 2.58$  ( $\alpha = .01$ )). Der Modell-Fit entspricht dem der EFA, sofern alle Indikatoren der ursprünglichen EFA in diese Berechnung mit eingehen.

### 5.2.3.6. Überprüfung von Methodeneffekten

Es gibt verschiedene Verfahren, um im Rahmen einer CFA das Vorliegen von Methodeneffekten, z.B. aufgrund negativer Itemformulierungen, zu überprüfen (vgl. Podsakoff et al., 2003; Horan et al., 2003; Hildebrandt & Temme, 2006). Dazu gehört die „correlated traits, correlated uniqueness“-Methode (CTCU-Methode, Marsh & Grayson, 1995). Der Vorteil dieser Methode liegt darin, dass sie die Überprüfung einer spezifisch formulierten Hypothese bezüglich des Methoden-Artefakts möglich macht, ohne dass dieser Effekt direkt messbar gemacht werden müsste (vgl. Podsakoff et al., 2003). Bei der CTCU-Methode wird ein Modell spezifiziert, in dem nicht nur die Ladungen der negativ formulierten Indikatoren auf ihrem jeweiligen Faktor („trait“), sondern auch Korrelationen zwischen den negativ formulierten Indikatoren zugelassen werden („correlated uniqueness“). Damit wird das Modell um die Korrelationen der Fehlerterme der negativ formulierten Items erweitert<sup>39</sup>.

Dieses gemäß der CTCU-Methode spezifizierte Modell kann ebenso wie ein Modell, das keine Methodeneffekte berücksichtigt, auf seinen Modell-Fit hin überprüft werden.

### 5.2.3.7. Modellvergleich

Um zu entscheiden, ob ein Modell mit korrelierten Fehlertermen tatsächlich besser die Zusammenhänge widerspiegelt als ein Modell mit unkorrelierten Fehlern, kann die Differenz der  $\chi^2$ -Werte auf Signifikanz geprüft werden. Der direkte Vergleich zweier Modelle ist dann möglich, wenn es sich um so genannte „nested models“ handelt<sup>40</sup>. Im vorliegenden Fall würde es sich bei dem Modell ohne korrelierte Fehlerterme um das genestete Modell handeln, da die Festlegung der Nullkorrelationen der Fehlerterme eine Reduzierung der frei geschätzten Parameter darstellte. Das Modell mit frei zu schätzenden Fehlerkorrelationen würde das entsprechende „parent model“ darstellen. Die Signifikanzprüfung der  $\chi^2$ -Werte der beiden Modelle ist allerdings bei der Verwendung des WLSMV-Schätzers etwas komplizierter als bei der Verwendung der auf normal-

---

39 Vorsicht ist allerdings bei der Verwendung der Standardfehler der Parameterschätzungen für die Signifikanzprüfung geboten, da diese verzerrt werden können (vgl. Podsakoff et al., 2003).

40 Ein „nested model“ beinhaltet eine Teilmenge der frei geschätzten Parameter eines anderen Modells (= „parent model“) bei ansonsten gleich bleibender Struktur (gleiche Anzahl an Faktoren und Indikatoren, vgl. Brown, 2006). Die beiden Modelle unterscheiden sich nur hinsichtlich der Anzahl an fixierten und frei geschätzten Parametern.

verteilten Daten basierenden Schätzer (z.B. ML), da die berechneten Differenzwerte nicht  $\chi^2$ -verteilt sind und zudem die Freiheitsgrade (df) anders berechnet werden. In Mplus ist der statistische Modellvergleich aber mit Hilfe eines zweistufigen Verfahrens durchführbar (vgl. Muthén and Muthén, 1998-2006). Interpretiert wird das Ergebnis aber wie ein üblicher  $\chi^2$ -Differenztest in Abhängigkeit von den Freiheitsgraden

### 5.2.3.8. Zusammenfassung

Aus den vorangegangenen Ausführungen lässt sich folgende Vorgehensweise für die weitere Auswertung der vorliegenden Daten ableiten.

- **E/CFA:** Durchführung einer EFA im CFA-Modell. Dabei sollen die Faktoren aus der ersten EFA mit den dazu gehörigen Indikatoren in das Modell übernommen werden. Zur Identifikation sollen die Varianzen der Faktoren auf eins festgesetzt und pro Faktor soll ein Item als Anker-Item ausgewählt werden (hohe Hauptladung, geringe Nebenladungen), dessen Nebenladungen für alle anderen Faktoren auf null gesetzt werden sollen. Die Nebenladungen für alle anderen Items werden dabei, ebenso wie die Kovarianzen der Fehler, frei geschätzt. Die Modifikationsindizes (bzw. EPC-Werte) sollen auf Hinweise bezüglich korrelierter Fehler überprüft und die Parameterwerte (v.a. die Nebenladungen) auf Signifikanz getestet werden. Dabei soll der Wert von  $\pm 2.58$  ( $\alpha = .01$ ) für Est./SE als Kriterium für die weitere Item-Eliminierung herangezogen werden.
- **Modell-Spezifizierung:** Ausgehend von der E/CFA-Lösung soll das Modell für die CFA spezifiziert werden<sup>41</sup>. Zudem muss die Metrik des Faktors definiert werden (Anker-Items als Markervariable oder Faktorvarianz auf eins setzen). Weiterhin soll das Modell als Pfaddiagramm dargestellt werden. Neben einem Modell ohne Berücksichtigung möglicher Methodeneffekte, soll ggf. zusätzlich ein Modell mit korrelierten Fehlern (CTCU-Methode) spezifiziert werden.
- **Modell-Schätzung:** Die Schätzung soll mit Hilfe des Programmes M-plus (Version 4.2) unter Verwendung des WLSMV-Schätzers erfolgen.
- **Modell-Evaluation:** Überprüfung der Passung über die Goodness-of-fit-Indizes. Zudem soll eine Überprüfung der Parameterwerte erfolgen („Out-of-range“-Schätzungen; Signifikanz der Parameterschätzungen, Höhe und Richtung der Faktorladungen; Faktorinterkorrelationen). Weiterhin sollen die Modifikationsindizes und EPC-Werte herangezogen

---

41 Beschreibung der frei geschätzten und fixierten Parameter: Liste der Indikatoren für jeden Faktor; ggf. Angaben zu Nebenladungen und korrelierten Fehlern.

werden, um mögliche Schwachstellen des Modells ausfindig zu machen. Das Ergebnis soll als Pfaddiagramm mit komplett standardisierten Parametern dargestellt werden. Konkurrierende Modelle (mit und ohne korrelierte Fehlerterme) sollen hinsichtlich ihrer Passung mit Hilfe eines  $\chi^2$ -Differenztests verglichen werden. Entsprechend des Ergebnisses soll die Entscheidung für eines der Modelle erfolgen.

- **Respezifizierung des Modells:** Ausgehend von den Ergebnissen der CFA soll eine Respezifizierung des ausgewählten Modells über weitere Item-Eliminierungen erfolgen. Kriterien sind neben einem guten Modell-Fit, auch die Kürze des Fragebogens und seine Fähigkeit, die konstruktrelevanten Inhalte abzudecken.

Kriterien für die Eliminierung von Items sind a) die Modifikationsindizes (M.I.>5.00) bzw. EPC-Werte und b) die Faktorladungen ( $a < .50$ ). Daneben sollen die Faktorinterkorrelationen ( $< .80$ ) und die Anzahl der Indikatoren (=Fragebogenlänge zwischen 30 und 40 Items) berücksichtigt werden.

### 5.2.4. Ergebnisse

#### 5.2.4.1. Ergebnisse der Explorativen Faktorenanalyse

##### **Zusammenfassung der Antwortkategorien**

Durch die Vorgabe, dass jede Antwortkategorie mind. 5% der Antworten enthalten muss, ergibt sich eine zwischen den Items variierende Anzahl von Antwortkategorien, was aber für die weitere Auswertung unproblematisch ist. Die Zusammenlegung konnte in der Outputdatei kontrolliert werden (SUMMARY OF CATEGORICAL DATA PROPORTIONS pro Kategorie  $\geq 0.05$ ) und erwies sich als fehlerfrei.

##### **Überprüfung der missing data-Verteilung**

Die Kovarianz-Deckung liegt in allen Fällen über 0.950 (95%). Die Verteilung der Häufigkeiten zeigt in 162 Fällen komplette Datensätze und eine unsystematische Verteilung bei den lückenhaften Datensätzen, d.h. in max. drei Fällen das gleiche „Lückenmuster“. Daher kann man davon ausgehen, dass die Annahme „missing at random“ erfüllt und die Ergänzung der lückenhaften Datensätze zulässig ist.

##### **Faktorenselktion**

###### a) Eigenwerte:

Die Eigenwerte der extrahierten Faktoren sind in Abbildung 2 dargestellt.



Abbildung 2: Eigenwerte

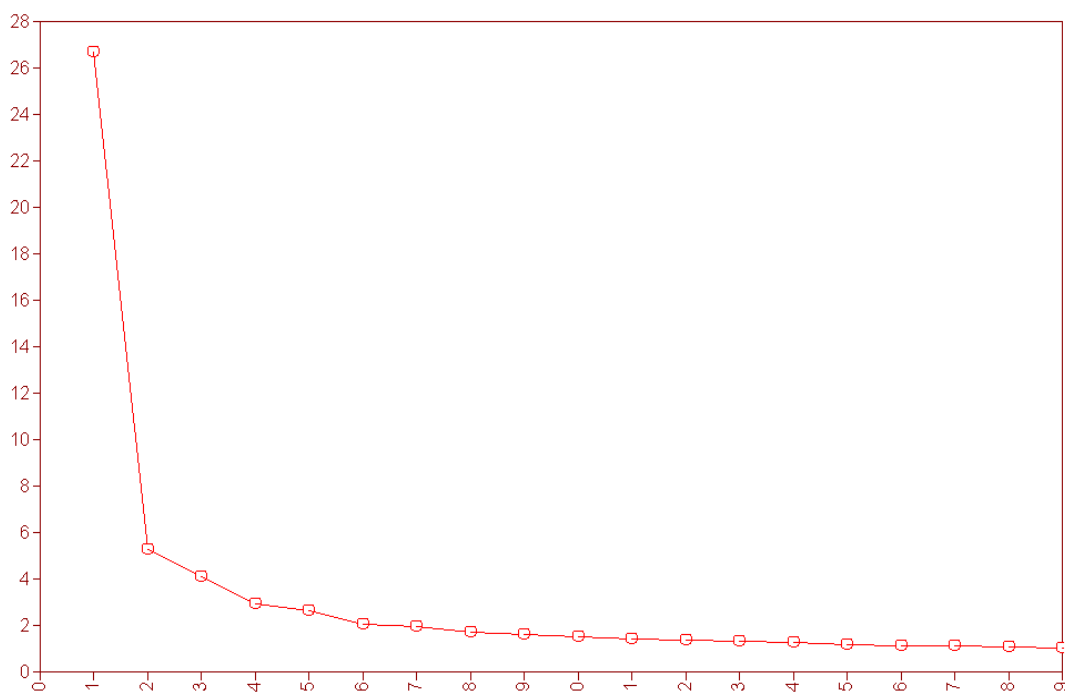
EIGENVALUES FOR SAMPLE CORRELATION MATRIX									
1	2	3	4	5	41	42	43	44	45
26.366	5.234	4.102	2.878	2.601	0.402	0.395	0.353	0.349	0.328
6	7	8	9	10	46	47	48	49	50
2.015	1.767	1.728	1.599	1.504	0.304	0.283	0.265	0.256	0.245
11	12	13	14	15	51	52	53	54	55
1.432	1.389	1.313	1.257	1.181	0.241	0.218	0.206	0.203	0.188
16	17	18	19	20	56	57	58	59	60
1.127	1.100	1.043	0.991	0.966	0.172	0.156	0.154	0.135	0.121
21	22	23	24	25	61	62	63	64	65
0.938	0.921	0.887	0.856	0.841	0.111	0.102	0.098	0.085	0.071
26	27	28	29	30	66	67	68	69	70
0.778	0.742	0.714	0.689	0.646	0.057	0.047	0.037	0.030	0.018
31	32	33	34	35	71	72	73	74	75
0.626	0.587	0.539	0.529	0.523	0.012	0.005	-0.011	-0.022	-0.035
36	37	38	39	40	76	77	78	79	80
0.508	0.497	0.473	0.434	0.414	-0.038	-0.050	-0.063	-0.072	-0.088

Wie aus Abbildung 2 ersichtlich, wurden 18 Faktoren mit einem Eigenwert >1,0 extrahiert. Gemäß der Kaiser-Gutman-Regel müssten diese 18 Faktoren in die weitere Analyse eingehen.

**b) Scree-Test:**

In Abbildung 3 ist der Scree-Plot des Mplus-Outputs (EFA) abgebildet. Auf der x-Achse ist die Anzahl der Faktoren abgetragen und auf der Y-Achse die dazugehörigen Eigenwerte.

Abbildung 3: Scree-Plot



Anmerkung: Der Übersicht wegen sind nur die Eigenwerte >1 im Scree-Plot dargestellt.

Entsprechend des Selektionskriteriums (Anzahl der Faktoren bis zum „Knick“ im Eigenwerteverlauf) könnten entweder 2 oder 5 Faktoren für die weitere Auswertung berücksichtigt werden.

### c) Goodness-of-fit-Indices:

Ausgehend von den 5-a-priori-Faktoren werden zunächst die 5-Faktoren-Lösung, sowie zusätzlich die „nächstliegenden“, d.h. die 2-, 3-, 4-, sowie die 6-, 7- und 8-Faktoren-Lösungen, hinsichtlich ihrer Goodness-of-fit-Indices betrachtet. Die Ergebnisse sind in Tabelle 23 aufgeführt.

Tabelle 23:

#### Goodness-of-fit-Indices (EFA)

	$\chi^2$	df	p	$\chi^2/\text{df}$	RMSEA	RMR
2-Faktoren-Lösung	422.89	121	0.000	3.49	0.107	0.083
3-Faktoren-Lösung	330.83	128	0.000	2.58	0.086	0.067
4-Faktoren-Lösung	299.94	132	0.000	2.27	0.077	0.061
5-Faktoren-Lösung	272.58	134	0.000	2.03	0.069	0.056

## 5. Auswertung und Ergebnisse

6-Faktoren-Lösung	253.93	133	0.000	1.91	0.065	0.051
7-Faktoren-Lösung	249.88	138	0.000	1.81	0.061	0.048
8-Faktoren-Lösung	241.67	139	0.000	1.74	0.058	0.046

Anmerkungen: RMSEA= root mean square of approximation (Cut-Off<0.06); RMR= root mean square residual (Cut-Off <0.05).

Berücksichtigt werden die Werte  $\chi^2/df$ , RMSEA und RMR als Indikatoren für einen guten Model-Fit<sup>42</sup>. Wie aus Tabelle 23 ersichtlich ergibt sich für die 6-Faktoren-Lösung ein  $\chi^2/df$ -Verhältnis < 2, wobei die anderen Werte noch über den Cut-Off-Werten liegen. Für die 7-Faktoren-Lösung zeigt sich ein  $\chi^2/df$ -Verhältnis < 2, ein RMR < 0.05 und ein RMSEA nahe dem Cutoff-Wert < 0.06. Bei der 8-Faktoren-Lösung liegen alle drei Werte unterhalb des Cut-Off-Wertes. Demnach würden sich für die weitere Auswertung die 7- oder 8-Faktorenlösung anbieten.

Die Anzahl von 18 Faktoren, die nach der Kaiser-Gutman-Regel in die weitere Analyse eingehen müsste, erscheint angesichts der 5 a-priori-Faktoren schwer nachvollziehbar. Da sich die Anwendung dieser Regel ohnehin als eher problematisch erwiesen hat (vgl. Preacher & McCallum, 2003), wird zur Faktorenauswahl den anderen Auswahlkriterien mehr Gewicht zugesprochen. Doch auch die Lösungen, die sich aus der Betrachtung des Scree-Plot einerseits und der Berücksichtigung der Fit-Indices andererseits anbieten, weichen deutlich voneinander ab: der Scree-Plot spricht entweder für eine 2- oder eine 5-Faktoren-Lösung, die Fit-Indices hingegen für eine 7- oder 8-Faktoren-Lösung. Aufgrund der Uneindeutigkeit werden daher zunächst eine Faktorrotation durchgeführt, um aufgrund inhaltlicher Gesichtspunkte eine Entscheidung treffen zu können.

### Faktorenselktion nach obliquer Rotation

Verglichen werden sollen nun die 2-, 5-, 7- und 8-Faktorenlösungen hinsichtlich der Interpretierbarkeit und Plausibilität der Faktoren. Als zusätzliches Kriterium werden die Ladungsverteilungen mitberücksichtigt. Den Ausgangspunkt für die Beurteilung der Interpretierbarkeit stellen die 5 a-priori-Skalen dar, die im folgenden mit den zugehörigen Item-Nummern aufgelistet sind:

*I. Glaubwürdigkeit/Nachvollziehbarkeit des Therapieangebotes (G)*

[19, 45, 47, 48, 56, 62, 83, 86, 89]

*II. Glaube/Hoffnung/Erwartung von Veränderung (VE)*

---

42 Die signifikanten Ergebnisse des  $\chi^2$ -Tests werden aufgrund der Stichprobenabhängigkeit des Verfahrens nicht berücksichtigt.

[68, 71, 73, 74, 75, 76, 78, 80, 82, 84, 85, 87, 90]

III. *Therapeutische Beziehung (TB)* (beeinhaltet Aspekte aus folgenden Ursprungsskalen: „Empathie, Akzeptanz und Echtheit des Therapeuten“, „Emotionale Bindung zwischen Therapeut und Patient“ und „Zusammenarbeit zwischen Therapeut und Patient“ )

[1, 4, 5, 12, 13, 15, 17, 21, 27, 29, 30, 31, 33, 35, 38, 39, 40, 43, 44, 46, 49, 51, 53, 55, 57, 58, 60, 61, 65, 70]

IV. *Bereitschaft des Patienten, sich einzulassen und mitzuarbeiten (MO)* (beeinhaltet Aspekte aus folgenden Ursprungsskalen: „Offenheit des Patienten“ und „Bereitschaft des Patienten zur Mitarbeit“)

[7, 11, 20, 22, 24, 26, 36, 42, 52, 59, 63, 69, 72, 77, 81, 88]

V. *Kompetenz des Therapeuten (K)*

[2, 3, 9, 23, 25, 28, 32, 34, 37, 41, 50, 64]

In den Tabellen 24 bis 27 sind die Lösungen für 2, 5, 7 und 8 Faktoren angegeben. Um die Darstellungen übersichtlicher zu halten, sind pro Faktorenlösung für jeden Faktor nur diejenigen Items angegeben, die auf keinem anderen Faktor höher laden und deren Ladung über .40 liegen (Angabe der Item-Nr. und Angabe der Hauptladung in Klammern). Zu jedem Faktor ist zudem eine mögliche Skalenbezeichnung angegeben, ebenso wie je drei Beispiel-Items, um die Faktorenterpretation nachvollziehbar zu machen.

Tabelle 24:  
Faktorenlösung mit 2 Faktoren

<b>Faktorenlösung mit 2 Faktoren</b>		
<b>Faktor 1</b>	<b>Faktor 2</b>	
Item 45 (.41)	Item 01 (.40)	Item 51 (.41)
Item 47 (.63)	Item 20 (.41)	Item 55 (.58)
Item 48 (.65)	Item 63 (.48)	Item 57 (.45)
Item 12 (.56)	Item 56 (.63)	Item 60 (.72)
Item 13 (.58)	Item 62 (.68)	Item 61 (.66)
Item 15 (.63)	Item 68 (.42)	Item 65 (.56)
Item 17 (.55)	Item 71 (.51)	Item 70 (.45)
Item 46 (.66)	Item 73 (.64)	Item 11 (.54)
Item 49 (.71)	Item 74 (.51)	Item 72 (.41)
Item 53 (.70)	Item 75 (.65)	Item 02 (.72)

## 5. Auswertung und Ergebnisse

Item 58 (.46) Item 22 (.60) Item 26 (.44) Item 36 (.66) Item 78 (.57) Item 80 (.72) Item 85 (.72) Item 87 (.66) Item 27 (.50) Item 44 (.50) Item 24 (.43) Item 59 (.45) Item 81 (-.54) Item 50 (.49)	Item 76 (.61) Item 82 (.45) Item 84 (.51) Item 21 (.51) Item 28 (.84) Item 29 (.60) Item 30 (.48) Item 31 (.67) Item 33 (.53) Item 35 (.52) Item 38 (.63) Item 39 (.55) Item 40 (.63) Item 43 (.55) Item 03 (.70) Item 09 (.88) Item 23 (.57) Item 25 (.66) Item 88 (.41) Item 32 (.73) Item 34 (.70) Item 37 (.83) Item 41 (.55) Item 64 (.87) Item 42 (.40)
<p><b>Mögliche Bezeichnung:</b>  <i>„Passung zwischen Therapeut und Patient“</i></p> <p><b>Beispiel-Items:</b>            Item 53: „In der letzten Therapiestunde ... bin ich mir vorgekommen wie ein Schüler“            Item 49: „...hatte ich das Gefühl, dass ich mich verteidigen muss“            Item 15: „...hat mich das Verhalten meines Therapeuten irritiert.“</p>	<p><b>Mögliche Bezeichnung:</b>  <i>„Positives Therapieerleben“</i></p> <p><b>Beispiel-Items :</b>            Item 09: „In der letzten Therapiestunde ... habe ich gemerkt, dass mein Therapeut über ein großes Fachwissen verfügt.“            Item 61: „... hat sich mein Therapeut sehr einfühlsam verhalten.“            Item 75: „In der Woche vor der letzten Therapiesitzung ... hatte ich das Gefühl, dass die Therapie mir helfen wird.“            Item 23: In der letzten Therapiestunde ... ging mein Therapeut planvoll und zielorientiert vor.“</p>

Anmerkungen: Faktorenladungen: angegeben sind nur die Hauptladungen der Items (in Klammern)  $\geq .40$ .

Tabelle 25:  
Faktorenlösung mit 5 Faktoren

<b>Faktorenlösung mit 5 Faktoren</b>				
Faktor 1	Faktor 2	Faktor 3	Faktor 4	Faktor 5
Item 86 (.45)	Item 45 (.40)	Item 62 (.47)	Item 47 (.54)	Item 20 (.45)
Item 68 (.74)	Item 89 (.47)	Item 21 (.58)	Item 48 (.50)	Item 11 (.65)
Item 71 (.75)	Item 01 (.42)	Item 29 (.71)	Item 12 (.65)	Item 42 (.43)
Item 73 (.80)	Item 56 (.40)	Item 30 (.55)	Item 13 (.56)	Item 81 (.75)
Item 74 (.79)	Item 23 (.45)	Item 31 (.44)	Item 15 (.67)	Item 88 (.63)
Item 75 (.81)		Item 33 (.69)	Item 17 (.52)	
Item 76 (.84)		Item 35 (.56)	Item 46 (.62)	
Item 78 (.50)		Item 38 (.75)	Item 49 (.74)	

## 5. Auswertung und Ergebnisse

Item 80 (.60)		Item 39 (.75)	Item 53 (.73)	
Item 82 (.60)		Item 40 (.70)	Item 58 (.55)	
Item 84 (.59)		Item 43 (.43)	Item 22 (.65)	
Item 85 (.51)		Item 57 (.43)	Item 26 (.42)	
Item 87 (.57)		Item 60 (.54)	Item 36 (.67)	
Item 90 (.67)		Item 61 (.73)	Item 27 (.43)	
Item 69 (.58)		Item 65 (.75)	Item 44 (.51)	
		Item 70 (.40)	Item 51 (.41)	
		Item 02 (.41)	Item 24 (.51)	
		Item 09 (.68)	Item 59 (.45)	
		Item 25 (.40)	Item 50 (.57)	
		Item 28 (.70)		
		Item 32 (.61)		
		Item 34 (.51)		
		Item 37 (.63)		
		Item 41 (.46)		
		Item 64 (.75)		
Mögliche Bezeichnung: <i>„Veränderungserwartung“</i>	Mögliche Bezeichnung: <i>„Nachvollziehbarkeit und Zielorientierung der Therapie“</i>	Mögliche Bezeichnung: <i>„Therapeutische Beziehung und Kompetenzwirkung des Therapeuten“</i>	Mögliche Bezeichnung: <i>„Passung zwischen Therapeut und Patient“</i>	Mögliche Bezeichnung: <i>„Bereitschaft des Patienten, sich einzulassen und mitzuarbeiten“</i>

Anmerkungen: Faktorenladungen: angegeben sind nur die Hauptladungen der Items  $\geq .40$ .

### Beispiel-Items für Faktor 1:

Item 75: „In der Woche vor der letzten Therapiesitzung ... hatte ich das Gefühl, dass die Therapie mir helfen wird.“

Item 73: „In der Woche vor der letzten Therapiesitzung ... habe ich mir gedacht, dass ich am Ende der Therapie mit meinem Alltag besser zurechtkommen werde.“

Item 76: „In der Woche vor der letzten Therapiesitzung ... habe ich geglaubt, dass die Therapie zu einer Verbesserung der Symptome führen wird.“

### Beispiel-Items für Faktor 2:

Item 23: „In der letzten Therapiestunde ... ging mein Therapeut planvoll und zielorientiert vor.“

Item 89: „In der Woche vor der letzten Therapiesitzung ... wusste ich, was mich in der nächsten Therapiesitzung erwartet.“

### Beispiel-Items für Faktor 3:

Item 09: „In der letzten Therapiestunde ... habe ich gemerkt, dass mein Therapeut über ein großes Fachwissen verfügt.“

Item 61: „In der letzten Therapiestunde ... hat sich mein Therapeut sehr einfühlsam verhalten.“

Item 39: „In der letzten Therapiestunde ... habe ich gespürt, dass ich meinem Therapeuten vertrauen kann.“

### **Beispiel-Items für Faktor 4:**

Item 53: „In der letzten Therapiestunde ... bin ich mir vorgekommen wie ein Schüler“

Item 49: „In der letzten Therapiestunde ... hatte ich das Gefühl, dass ich mich verteidigen muss“

Item 15: „In der letzten Therapiestunde ... hat mich das Verhalten meines Therapeuten irritiert.“

### **Beispiel-Items für Faktor 5:**

Item 81: „In der Woche vor der letzten Therapiesitzung ... habe ich gemerkt, dass ich mich sehr bemühen muss, um eine Veränderung zu erreichen.“

Item 11: „In der letzten Therapiestunde ... habe ich gemerkt, dass ich ernsthaft an mir arbeiten muss.“

Item 20: „In der letzten Therapiestunde ... habe ich mich getraut, Dinge zu sagen, die mir peinlich sind.“

Tabelle 26:  
Faktorenlösung mit 7 Faktoren

Faktorenlösung mit 7 Faktoren						
Faktor 1	Faktor 2	Faktor 3	Faktor 4	Faktor 5	Faktor 6	Faktor 7
Item 45 (.40)	Item 11 (.51)	Item 62 (.47)	Item 47 (.54)	Item 09 (.60)	Item 20 (.74)	Item 68 (.64)
Item 89 (.55)	Item 81 (.81)	Item 21 (.40)	Item 48 (.57)	Item 28 (.69)	Item 42 (.67)	Item 71 (.80)
Item 01 (.51)	Item 88 (.72)	Item 29 (.71)	Item 12 (.70)	Item 32 (.49)	Item 52 (.60)	Item 73 (.82)
Item 56 (.64)		Item 33 (.66)	Item 13 (.53)	Item 37 (.55)	Item 59 (.40)	Item 74 (.77)
Item 31 (.49)		Item 35 (.50)	Item 15 (.59)	Item 64 (.58)		Item 75 (.81)
Item 55 (.55)		Item 38 (.73)	Item 17 (.46)			Item 76 (.80)
Item 02 (.49)		Item 39 (.78)	Item 46 (.55)			Item 78 (.56)
Item 03 (.56)		Item 40 (.74)	Item 49 (.73)			Item 80 (.53)
Item 23 (.57)		Item 43 (.42)	Item 53 (.64)			Item 82 (.63)
Item 25 (.44)		Item 44 (.45)	Item 58 (.47)			Item 84 (.54)
		Item 60 (.51)	Item 22 (.67)			Item 87 (.53)
		Item 61 (.74)	Item 26 (.46)			Item 90 (.63)
		Item 65 (.61)	Item 36 (.62)			Item 69 (.47)
		Item 70 (.50)	Item 85 (.46)			
		Item 41 (.40)	Item 27 (.40)			
			Item 50 (.60)			
Mögliche Bezeichnung: <i>„Nachvollziehbarkeit und Zielorientierung der Therapie“</i>	Mögliche Bezeichnung: <i>„Bereitschaft zur Mitarbeit“</i>	Mögliche Bezeichnung: <i>„Therapeutische Beziehung“</i>	Mögliche Bezeichnung: <i>„Passung zwischen Therapeut und Patient“</i>	Mögliche Bezeichnung: <i>„Kompetenzwirkung des Therapeuten“</i>	Mögliche Bezeichnung: <i>„Bereitschaft des Patienten, sich einzulassen“/„Offenheit“</i>	Mögliche Bezeichnung: <i>„Veränderungserwartung“</i>

Anmerkungen: Faktorenladungen: angegeben sind nur die Hauptladungen der Items  $\geq .40$ .

### Beispiel-Items für Faktor 1:

Item 23: „In der letzten Therapiestunde ... ging mein Therapeut planvoll und zielorientiert vor.“

Item 56: „In der letzten Therapiestunde ... erschien mir das Therapieangebot einleuchtend.“

Item 89: „In der Woche vor der letzten Therapiesitzung ... wusste ich, was mich in der nächsten Therapiesitzung erwartet.“

### Beispiel-Items für Faktor 2:

Item 81: „In der Woche vor der letzten Therapiesitzung ... habe ich gemerkt, dass ich mich sehr bemühen muss, um eine Veränderung zu erreichen.“

Item 88: „... habe ich mir gedacht, dass ich für eine Verbesserung noch viel tun muss.“



Item 11: „In der letzten Therapiestunde. ... habe ich gemerkt, dass ich ernsthaft an mir arbeiten muss.“

**Beispiel-Items für Faktor 3:**

Item 61: „ In der letzten Therapiestunde ... hat sich mein Therapeut sehr einfühlsam verhalten.“

Item 39: „ ... habe ich gespürt, dass ich meinem Therapeut vertrauen kann.“

Item 38: „ ... fühlte ich mich von meinem Therapeut als Mensch akzeptiert.“

**Beispiel-Items für Faktor 4:**

Item 53: „In der letzten Therapiestunde ... bin ich mir vorgekommen wie ein Schüler“

Item 49: „ ... hatte ich das Gefühl, dass ich mich verteidigen muss“

Item 15: „ ... hat mich das Verhalten meines Therapeuten irritiert.“

**Beispiel-Items für Faktor 5:**

Item 28: „ In der letzten Therapiestunde ... habe ich gemerkt, dass mein Therapeut viel Erfahrung hat.“

Item 09: „ ... habe ich gemerkt, dass mein Therapeut über ein großes Fachwissen verfügt.“

Item 64: „ ... habe ich gemerkt, dass mein Therapeut gut ausgebildet ist.“

**Beispiel-Items für Faktor 6:**

Item 20: „ In der letzten Therapiestunde ... habe ich mich getraut, Dinge zu sagen, die mir peinlich sind.“

Item 42: „ ... bin ich das Risiko eingegangen, auch unangenehme Gefühle zu erleben.“

Item 52: „ ... konnte ich über meinen Schatten springen.“

**Beispiel-Items für Faktor 7:**

Item 75: „In der Woche vor der letzten Therapiesitzung ... hatte ich das Gefühl, dass die Therapie mir helfen wird.“

Item 73: „In der Woche vor der letzten Therapiesitzung ... habe ich mir gedacht, dass ich am Ende der Therapie mit meinem Alltag besser zurechtkommen werde.“

Item 76: „In der Woche vor der letzten Therapiesitzung ... habe ich geglaubt, dass die Therapie zu einer Verbesserung der Symptome führen wird.“

Tabelle 27:  
Faktorenlösung mit 8 Faktoren

Faktorenlösung mit 8 Faktoren							
Faktor 1	Faktor 2	Faktor 3	Faktor 4	Faktor 5	Faktor 6	Faktor 7	Faktor 8
Item 89 (.56)	Item 47 (.52)	Item 62 (.45)	Item 19 (.40)	Item 11 (-.55)	Item 20 (.75)	Item 68 (.63)	Item 09 (.61)

## 5. Auswertung und Ergebnisse

Item 01 (.52)	Item 48 (.57)	Item 21 (.40)	Item 51 (.49)	Item 81 (-.78)	Item 42 (.66)	Item 71 (.79)	Item 28 (.71)
Item 56 (.63)	Item 12 (.70)	Item 29 (.71)		Item 88 (-.73)	Item 52 (.59)	Item 73 (.81)	Item 32 (.55)
Item 31 (.47)	Item 13 (.53)	Item 33 (.67)			Item 59 (.40)	Item 74 (.78)	Item 34 (.41)
Item 55 (.62)	Item 15 (.60)	Item 35 (.49)				Item 75 (.81)	Item 37 (.58)
Item 03 (.49)	Item 17 (.47)	Item 38 (.73)				Item 76 (.81)	Item 64 (.60)
Item 23 (.54)	Item 46 (.55)	Item 39 (.78)				Item 78 (.58)	
	Item 49 (.73)	Item 40 (.73)				Item 80 (.54)	
	Item 53 (.65)	Item 43 (.42)				Item 82 (.62)	
	Item 58 (.47)	Item 44 (.45)				Item 84 (.55)	
	Item 22 (.67)	Item 60 (.51)				Item 87 (.55)	
	Item 26 (.47)	Item 61 (.74)				Item 90 (.63)	
	Item 36 (.62)	Item 65 (.61)				Item 69 (.47)	
	Item 85 (.47)	Item 70 (.50)					
	Item 27 (.40)						
	Item 50 (.60)						
Mögliche Bezeichnung: <i>„Nachvollziehbarkeit und Zielorientierung der Therapie“</i>	Mögliche Bezeichnung: <i>„Passung zwischen Therapeut und Patient“</i>	Mögliche Bezeichnung: <i>„Therapeutische Beziehung“</i>	Mögliche Bezeichnung: ?	Mögliche Bezeichnung: <i>„Bereitschaft zur Mitarbeit“</i>	Mögliche Bezeichnung: <i>„Bereitschaft sich einzulassen“/“Offenheit“</i>	Mögliche Bezeichnung: <i>„Veränderungserwartung“</i>	Mögliche Bezeichnung: <i>„Kompetenzwirkung des Therapeuten“</i>

Anmerkungen: Faktorenladungen: angegeben sind nur die Hauptladungen der Items  $\geq$ /.40.

### Beispiel-Items für Faktor 1:

Item 56: „In der letzten Therapiestunde ... erschien mir das Therapieangebot einleuchtend.“

Item 89: „In der Woche vor der letzten Therapiesitzung ... wusste ich, was mich in der nächsten Therapiesitzung erwartet.“

Item 23: „In der letzten Therapiestunde ... ging mein Therapeut planvoll und zielorientiert vor.“

### Beispiel-Items für Faktor 2:

Item 53: „In der letzten Therapiestunde ... bin ich mir vorgekommen wie ein Schüler“

Item 49: „... hatte ich das Gefühl, dass ich mich verteidigen muss“

Item 15: „... hat mich das Verhalten meines Therapeuten irritiert.“

### Beispiel-Items für Faktor 3:

Item 61: „In der letzten Therapiestunde ... hat sich mein Therapeut sehr einfühlsam verhalten.“

Item 39: „... habe ich gespürt, dass ich meinem Therapeut vertrauen kann.“

Item 38: „... fühlte ich mich von meinem Therapeut als Mensch akzeptiert.“

### Beispiel-Items für Faktor 4:

Item 19: „In der letzten Therapiestunde ... war das Vorgehen für mich nachvollziehbar.“

Item 51: „... hatte ich das Gefühl, dass wir beide an einem Strang ziehen.“

### **Beispiel-Items für Faktor 5:**

Item 81: „In der Woche vor der letzten Therapiesitzung ... habe ich gemerkt, dass ich mich sehr bemühen muss, um eine Veränderung zu erreichen.“

Item 88: „... habe ich mir gedacht, dass ich für eine Verbesserung noch viel tun muss.“

Item 11: „In der letzten Therapiestunde ... habe ich gemerkt, dass ich ernsthaft an mir arbeiten muss.“

### **Beispiel-Items für Faktor 6:**

Item 20: „In der letzten Therapiestunde ... habe ich mich getraut, Dinge zu sagen, die mir peinlich sind.“

Item 42: „... bin ich das Risiko eingegangen, auch unangenehme Gefühle zu erleben.“

Item 52: „... konnte ich über meinen Schatten springen.“

### **Beispiel-Items für Faktor 7:**

Item 75: „In der Woche vor der letzten Therapiesitzung ... hatte ich das Gefühl, dass die Therapie mir helfen wird.“

Item 73: „In der Woche vor der letzten Therapiesitzung ... habe ich mir gedacht, dass ich am Ende der Therapie mit meinem Alltag besser zurecht kommen werde.“

Item 76: In der Woche vor der letzten Therapiesitzung ... habe ich geglaubt, dass die Therapie zu einer Verbesserung der Symptome führen wird.“

### **Beispiel-Items für Faktor 8:**

Item 28: „In der letzten Therapiestunde ... habe ich gemerkt, dass mein Therapeut viel Erfahrung hat.“

Item 09: „... habe ich gemerkt, dass mein Therapeut über ein großes Fachwissen verfügt.“

Item 64: „... habe ich gemerkt, dass mein Therapeut gut ausgebildet ist.“

Vergleicht man die vier Faktoren-Lösungen, so fällt auf, dass ein Faktor über alle Lösungen stabil bleibt, der nicht in der ursprünglichen Konzeption vorgesehen war: der Faktor mit der Bezeichnung „Passung zwischen Therapeut und Patient“, auf dem Items aus allen a-priori-Skalen laden. Abgesehen von diesem Faktor stellen die vier Faktorenlösungen unterschiedliche „Ausdifferenzierungsstufen“ der a-priori-Konzepte dar:

In der 2-Faktorenlösung mit 69 Indikatoren  $\geq .40$  laden Items aller a-priori-Skalen auf einem Faktor (abgesehen von denjenigen Items, die auf dem oben genannten stabilen Faktor „Passung

zwischen Therapeut und Patient“ laden), wobei die Items aus dem Bereich „Kompetenz des Therapeuten“ (Item 64, Item 09, Item 28) die höchsten Werte aufweisen.

Die 5-Faktoren-Lösung mit 69 Indikatoren  $\geq .40$  weist im Vergleich zu den a-priori-Skalen folgende Aufspaltung auf (abgesehen von dem stabilen Faktor „Passung zwischen Therapeut und Patient“): gemäß der ursprünglichen Konzeption bilden sich ein Faktor aus Items der a-priori-Skala „Bereitschaft des Patienten, sich einzulassen und mitzuarbeiten“ und ein Faktor aus Items der a-priori-Skala „Erwartung von Veränderung“. Ein Faktor ergibt sich aus einer Kombination von Items der Skalen „Therapeutische Beziehung“ und „Kompetenz des Therapeuten“ und der letzte Faktor („Nachvollziehbarkeit und Zielorientierung der Therapie“) setzt sich aus drei Items der a-priori-Skala „Nachvollziehbarkeit des Therapieangebotes“ und einem Item der Skala „Kompetenz des Therapeuten“ zusammen.

In der 7-Faktorenlösung mit 66 Indikatoren  $\geq .40$  teilen sich die Items des Faktors „Therapeutische Beziehung und Kompetenzwirkung des Therapeuten“ aus der 5-Faktoren-Lösung entsprechend der ursprünglichen Konzeption auf zwei Faktoren auf. Ebenso teilen sich die Items des Faktors „Bereitschaft des Patienten sich einzulassen und mitzuarbeiten“ auf zwei neue Faktoren auf („Bereitschaft sich einzulassen/Offenheit“ und „Bereitschaft zur Mitarbeit“). Der Faktor „Nachvollziehbarkeit und Zielorientierung der Therapie“ konnte im Vergleich zur 5-Faktorenlösung 5 Items dazu gewinnen. Die Faktoren „Passung zwischen Therapeut und Patient“ und „Veränderungserwartung“ bleiben weitgehend unverändert.

In der 8-Faktoren-Lösung mit 65 Indikatoren  $\geq .40$  ergab sich zusätzlich zu den Faktoren der 7-Faktorenlösung ein Faktor aus zwei Einzelfaktoren (Item 19 und Item 51), für den keine passende Bezeichnung gefunden werden konnte.

Ausgehend von der Interpretierbarkeit der Faktoren erscheinen die 5- und die 7-Faktorenlösung am plausibelsten. Für die 5-Faktorenlösung spricht, dass sich der Faktor „Bereitschaft des Patienten, sich einzulassen und mitzuarbeiten“ entsprechend der a-priori-Konzeption gebildet hat, und dass – im Vergleich zur 7-Faktorenlösung – mehr Items der Skala „Therapeutische Beziehung und Kompetenzwirkung des Therapeuten“ eine Ladung  $\geq .40$  aufweisen. Für die 7-Fakto-

ren-Lösung spricht hingegen, dass sich die Skalen „Therapeutische Beziehung“ und „Kompetenzwirkung des Therapeuten“ entsprechend der Ursprungskonzeption bildeten. Zudem weist der Faktor „Nachvollziehbarkeit und Zielorientierung der Therapie“ deutlich mehr Items und zudem höhere Ladungen auf. Auch die Items der Faktoren „Bereitschaft sich einzulassen/Offenheit“ und „Bereitschaft zur Offenheit“ haben höhere Ladungen.

In Abwägung der Vor- und Nachteile der beiden Lösungen wird die 7-Faktoren-Lösung insgesamt als geeigneter angesehen. Am meisten ins Gewicht fällt dabei die Trennung der Faktoren „Therapeutische Beziehung“ und „Kompetenzwirkung des Therapeuten“, da diese konzeptionell als zwei getrennte Einheiten angesehen werden, auch wenn eine inhaltliche Nähe plausibel ist. Zudem ist nur in der 7-Faktoren-Lösung der Faktor „Nachvollziehbarkeit und Zielorientierung der Therapie“ gut definiert, da er in der 5-Faktoren-Lösung lediglich durch vier Items mit Ladungen im Bereich von .40-.47 bestimmt wird. Für die Wahl der 7-Faktoren-Lösung spricht schließlich auch die Beobachtung von Fabrigar et al. (1999), dass generell eine Unterfaktorisierung größere Probleme bereitet als eine Überfaktorisierung, was in Zweifelsfällen eher die Wahl der Lösung mit mehr Faktoren nahelegt. Daher wurde die 7-Faktorenlösung mit folgenden Faktorenbezeichnungen als passende Lösung ausgewählt<sup>43</sup>:

1. *„Nachvollziehbarkeit und Zielorientierung der Therapie“*
2. *„Bereitschaft zur Mitarbeit“*
3. *„Therapeutische Beziehung“*
4. *„Passung zwischen Therapeut und Patient“*
5. *„Kompetenzwirkung des Therapeuten“*
6. *„Bereitschaft, sich einzulassen/Offenheit“*
7. *„Veränderungserwartung“*

### **Evaluation der Faktorenlösung:**

Hinsichtlich der inhaltlichen Bedeutung bzw. empirischen Relevanz der Skalen 2, 3, 5, 6 und 7 besteht wenig Unsicherheit, da diese Faktoren größtenteils den ursprünglichen Skalen entsprechen – abgesehen von der Aufspaltung der Skala *„Bereitschaft des Patienten, sich einzulassen und mitzuarbeiten“* in die zwei Einzelskalen *„Bereitschaft zur Mitarbeit“* und *„Bereitschaft, sich einzulassen/Offenheit“*. Der Faktor *„Nachvollziehbarkeit und Zielorientierung der Therapie“* besteht zwar aus Items dreier Ursprungsskalen (*„Glaubwürdigkeit/Nachvollziehbarkeit des*

---

43 Eine Darstellung der kompletten Faktorenlösung befindet sich im Anhang.

*Therapieangebotes*“, „*Therapeutische Beziehung*“ und „*Kompetenz des Therapeuten*“), ist aber in seiner Zusammensetzung stimmig, da alle Items den Aufbau, die Planung und die Nachvollziehbarkeit der Therapie zum Thema haben, wenn auch in unterschiedlichen Facetten.

Als problematisch hingegen erweist sich der Faktor „*Passung zwischen Therapeut und Patient*“. Es fällt zum einen auf, dass auf diesem Faktor Items aus allen a-priori-Skalen laden, was eine inhaltliche Interpretation erschwert. Zum anderen wird bei genauerer Betrachtung deutlich, dass nahezu alle Items, die auf diesem Faktor laden, negativ formuliert sind (Item47, Item48, Item12, Item13, Item15, Item17, Item46, Item49, Item53, Item58, Item22, Item26, Item36, Item85, Item24, Item50). Lediglich ein Item (Item27) ist positiv formuliert. Diese Beobachtung legt die Vermutung nahe, dass es sich bei diesem Faktor um einen so genannten „trivialen Faktor“ handelt, genauer gesagt, um einen Methodenfaktor, der auf die negative Formulierung zurückzuführen ist. Dies würde bedeuten, dass die Gemeinsamkeit, die sich in den Korrelationen zwischen den Items ausdrückt, nicht auf eine inhaltliche Überschneidung zurückgeht sondern auf die Ähnlichkeit hinsichtlich der Formulierungen, was die Patienten bei den entsprechenden Items zu einem ähnlichen Antwortverhalten motiviert haben könnte (Brown, 2006; Podsakoff et al., 2003, Horan et al., 2003). Dafür spricht weiterhin die Beobachtung, dass auch diejenigen negativ formulierten Items, die nicht primär auf dem Faktor „*Passung zwischen Therapeut und Patient*“ laden, substantielle Nebenladungen (zwischen .35 und .44) auf diesem Faktor haben. Weiterhin weisen viele Items des Faktors „*Passung zwischen Therapeut und Patient*“ ihrerseits substantielle Nebenladungen auf denjenigen Faktoren auf, für die sie ursprünglich konzipiert worden sind (z.B. Item47, Item85).

Gegen die Annahme eines Methodenfaktors spricht allerdings, dass es auch ein positiv formuliertes Item gibt, das auf diesem Faktor lädt. Zudem gibt es auch andere positiv formulierte Items, die substantielle Nebenladungen auf diesem Faktor haben (z.B. Item38, Item59, Item41). Und schließlich haben einige der hoch auf dem Faktor „*Passung zwischen Therapeut und Patient*“ ladenden Items keine relevanten Nebenladungen (z.B. Item12, Item22, Item50). Denkbar ist, dass bei diesem Faktor inhaltliche und methodische Gemeinsamkeiten konfundiert sind. Da im Rahmen einer EFA keine weiteren Schritte zur Klärung dieser Frage möglich sind, wird der Faktor „*Passung zwischen Therapeut und Patient*“ vorerst beibehalten. Beabsichtigt wird damit, eine vorschnelle Eliminierung von Items zu verhindern. Nach Durchführung der konfirmatori-

schen Faktorenanalyse sollte sich durch den Vergleich verschiedener Modelle eine klarere Aussage treffen lassen, ob es sich dabei um einen Methodenfaktor handelte oder nicht.

### **Optimierung der Faktorenlösung**

#### Eliminierung schlecht definierter Faktoren

Faktor 2 („*Bereitschaft zur Mitarbeit*“) ist nur durch drei Indikatoren bestimmt, was im Rahmen der EFA als problematisch anzusehen ist. Zwei der Indikatoren (Item81, Item88) laden hoch auf dem Faktor (.81; .70), was aber auch auf die sehr ähnliche Formulierung der Items zurückzuführen sein könnte. Faktor 2 wird dennoch vorerst beibehalten, da in weiteren Analysen (CFA) die Stabilität des Faktors überprüft werden sollte.

Ähnliches gilt für Faktor 6 („*Bereitschaft, sich einzulassen/Offenheit*“), der durch vier Indikatoren bestimmt ist, von denen einer lediglich eine Ladung von .40 aufweist (Item 59). Auch dieser Faktor wird für die weiteren Auswertungsschritte übernommen. Alle anderen Faktoren erscheinen sowohl hinsichtlich der Itemanzahl, als auch hinsichtlich der Höhe der Ladungen ausreichend definiert.

#### Eliminierung problematischer Indikatoren

Das Kriterium einer bedeutsamen Hauptladung ( $\geq .40$ ) erfüllen folgende 14 Items nicht: Item 83, Item 63, Item 86, Item 04, Item 51, Item 57, Item 72, Item 24, Item 34, Item 07, Item 19, Item 05, Item 30, Item 77.

Gemäß dem Fürntratt-Kriterium sollten folgende 15 Items ausgeschlossen werden: Item 83, Item 63, Item 86, Item 04, Item 51, Item 57, Item 72, Item 24, Item 34, Item 41, Item 47, Item 62, Item 85, Item 80, Item59 (Auflistung: siehe Anhang).

Aus der Auflistung wird ersichtlich, dass sich ein Teil der Items hinsichtlich beider Kriterien nicht für eine Weiterverwendung eignet, während andere sich nur in Bezug auf ein Kriterium als unzureichend erweisen. Insgesamt zeigen 20 Items problematische Werte auf.

Hinsichtlich der Inhaltsvalidität erscheint eine Eliminierung der meisten Items akzeptabel, da deren inhaltliche Domänen noch durch andere Items abgedeckt werden. Problematisch ist dies bei Item 59, da der zugehörige Faktor „*Bereitschaft, sich einzulassen/Offenheit*“ (Faktor 6) ohne dieses Item nur noch aus 3 Indikatoren gebildet würde. Daher wird dieses Item trotz eines Verhältnisses  $a^2/h^2$  von 0,48 beibehalten.

Entsprechend werden insgesamt 19 Items von der weiteren Auswertung ausgeschlossen, womit sich folgende Faktorenlösung mit insgesamt 61 Items ergibt:

Tabelle 28:  
Faktorenlösung mit 7 Faktoren nach Itemeliminierung

<b>Faktorenlösung mit 7 Faktoren</b>						
<i>„Nachvollziehbarkeit und Zielorientierung der Therapie“</i>	<i>„Bereitschaft zur Mitarbeit“</i>	<i>„Therapeutische Beziehung“</i>	<i>„Passung zwischen Therapeut und Patient“</i>	<i>„Kompetenzwirkung des Therapeuten“</i>	<i>„Bereitschaft des Patienten, sich einzulassen“/„Offenheit“</i>	<i>„Veränderungserwartung“</i>
Item 45 (.40)	Item 11 (.51)	Item 21 (.40)	Item 48 (.57)	Item 09 (.60)	Item 20 (.74)	Item 68 (.64)
Item 89 (.55)	Item 81 (.81)	Item 29 (.71)	Item 12 (.70)	Item 28 (.69)	Item 42 (.67)	Item 71 (.80)
Item 01 (.51)	Item 88 (.72)	Item 33 (.66)	Item 13 (.53)	Item 32 (.49)	Item 52 (.60)	Item 73 (.82)
Item 56 (.64)		Item 35 (.50)	Item 15 (.59)	Item 37 (.55)	Item 59 (.40)	Item 74 (.77)
Item 31 (.49)		Item 38 (.73)	Item 17 (.46)	Item 64 (.58)		Item 75 (.81)
Item 55 (.55)		Item 39 (.78)	Item 46 (.55)			Item 76 (.80)
Item 02 (.49)		Item 40 (.74)	Item 49 (.73)			Item 78 (.56)
Item 03 (.56)		Item 43 (.42)	Item 53 (.64)			Item 82 (.63)
Item 23 (.57)		Item 44 (.45)	Item 58 (.47)			Item 84 (.54)
Item 25 (.44)		Item 60 (.51)	Item 22 (.67)			Item 87 (.53)
		Item 61 (.74)	Item 26 (.46)			Item 90 (.63)
		Item 65 (.61)	Item 36 (.62)			Item 69 (.47)
		Item 70 (.50)	Item 27 (.40)			
			Item 50 (.60)			
$\Sigma= 10$	$\Sigma= 3$	$\Sigma= 13$	$\Sigma= 14$	$\Sigma= 5$	$\Sigma= 4$	$\Sigma= 12$

### 5.2.4.2. Ergebnisse der Konfirmatorischen Faktorenanalyse (CFA)

#### EFA im Rahmen einer CFA

In die E/CFA-Auswertung gehen die sieben Faktoren der ausgewählten EFA-Lösung mit 61 Indikatoren ein. Folgende Items werden als Ankeritems (hohe Haupt- und geringe Nebenladungen in der EFA) ausgewählt:

Faktor **„Nachvollziehbarkeit/Zielorientierung der Therapie“** (Abkürzung: „nach“): Item 23  
(„In der letzten Therapiestunde ... ging mein Therapeut planvoll und zielorientiert vor.“)

Faktor **„Bereitschaft zur Mitarbeit“** (Abkürzung: „mitar“): Item 81  
(„In der Woche vor ... habe ich gemerkt, dass ich mich sehr bemühen muss, um eine Veränderung zu erreichen.“)



Faktor „ <b>Therapeutische Beziehung</b> “ (Abkürzung: „bez“): („In der letzten Therapiestunde ... habe ich gespürt, dass mein Therapeut mich schätzt.“)	Item 33
Faktor „ <b>Passung zwischen Therapeut und Patient</b> “ (Abkürzung: „pas“): („In der letzten Therapiestunde ... hatte ich das Gefühl, dass ich mich verteidigen muss.“)	Item 49
Faktor „ <b>Kompetenzwirkung des Therapeuten</b> “ (Abkürzung: „komp“): („In der letzten Therapiestunde ... habe ich gemerkt, dass mein Therapeut viel Erfahrung hat.“)	Item 28
Faktor „ <b>Bereitschaft des Pat. sich einzulassen/sich zu öffnen</b> “ (Abkürzung: „offen“): („In der letzten Therapiestunde ... bin ich das Risiko eingegangen, auch unangenehme Gefühle zu erleben.“)	Item 42
Faktor „ <b>Veränderungserwartung</b> “ (Abkürzung: „ver“): („In der Woche vor ... habe ich geglaubt, dass die Therapie zu einer Verbesserung der Symptome führen wird.“)	Item 76

Für diese Items werden die Ladungen auf den jeweils übrigen sechs Faktoren auf null fixiert. Weiterhin werden die Faktorvarianzen auf eins festgesetzt. Damit ergibt sich mit  $7 \cdot 6 + 7 = 49$  die geforderte Anzahl an Restriktionen ( $m^2 = 72 = 49$ ). Als Startwerte werden zudem die Hauptladungen der Indikatoren eingesetzt.

Das Ergebnis zeigt mit  $\chi^2 = 225.32$  (df: 130) und einem RMSEA = 0.058 (weitere Ergebnisse: CFI = 0.96, TLI = 0.98, WRMR = 0.638) bessere Werte auf als das EFA-Modell ( $\chi^2 = 249.88$  (df: 138), RMSEA = 0.061), was an der Eliminierung der aufgrund der EFA-Ergebnisse ausgewählten Items liegt. Die Parameterschätzungen für die ausgewählten Indikatoren erweisen sich für alle Faktoren als signifikant auf beiden Signifikanzniveaus (Est./S.E.  $> \pm 1.96$  ( $\alpha = .05$ ) bzw.  $\pm 2.58$  ( $\alpha = .01$ )).

Für den nächsten Schritt der Itemselektion sind v.a. die an den Standardfehlern (S.E.) standardisierten Parameterschätzungen (Est.) der Nebenladungen interessant. Als Kriterium für die Eliminierung wurde ein Est./S.E.-Wert  $\geq \pm 2.58$  ( $\alpha = .01$ ) festgesetzt. Diesen Cut-Off-Wert überschreiten folgende Items (mit \* versehene Items weisen auf zwei Faktoren Doppelladungen auf): Item 17\*, Item 60, Item 09, Item 11, Item 65, Item 36\*, Item 56, Item 78\*, Item 21, Item 55, Item 70, Item 59, Item 45, Item 44, Item 87, Item 13, Item 27, Item 38, Item 69 ( $\Sigma = 19$  Items).

Es werden zwei Modifikationsindizes angegeben: eine Residuenkorrelation zwischen Item 55 und Item 56 (M.I. = 5.079; StdYX E.P.C<sup>44</sup> = 0.284), sowie eine Residuenkorrelation zwischen

---

<sup>44</sup> StdYX E.P.C. = vollstandardisierter „expected parameter of change“-Wert.

Item 65 und Item 21 (M.I. = 5.447, StdYX E.P.C. = 0.218). Bei der erstgenannten Korrelation könnte es sich um einen Reihenfolgeeffekt handeln, obwohl beide Items auch auf einem Faktor laden und möglicherweise auch inhaltlich mehr miteinander gemeinsam haben als mit den anderen Indikatoren. Bei den Items 65 und 21 handelt es sich um „Empathie-Items“, was die gemeinsame Restvarianz erklärt. Es ergaben sich keine Residuenkorrelationen für die negativ formulierten Items, wie z.B. Item 78 und Item 87, allerdings signifikante Nebenladungen auf dem Faktor „Passung zwischen Therapeut und Patient“. Folgende Überlegungen sind für die Auswahl der im nächsten Schritt zu eliminierenden Items ausschlaggebend:

Unter den Items mit auf dem .01-Level signifikanten Nebenladungen befinden sich vier negativ formulierte Items (Item 45, Item 78, Item 87, Item 44), die auf dem Faktor „Passung zwischen Therapeut und Patient“ Nebenladungen aufweisen. Hierbei könnte es sich um inhaltliche Nebenladungen handeln oder aber um methodenbedingte, was bedeuten würde, dass die Gemeinsamkeiten der genannten Items mit den Items des Faktors „Passung zwischen Therapeut und Patient“ primär auf die negative Formulierung der Items zurückzuführen sind. Da auch positiv formulierte Items signifikante Nebenladungen auf diesem Faktor aufweisen (z.B. Item 59) erscheint eine Trennung zwischen Methoden- und „Inhalts“-Effekt an dieser Stelle nicht möglich. Daher werden die Items 45, 78, 87 und 44 noch vor der Überprüfung von Methodeneffekten im Rahmen der CFA-Berechnung aus dem Itempool entfernt.

Unter den Items mit signifikanten Nebenladungen befindet sich auch ein Item des Faktors „Bereitschaft zur Mitarbeit“ (Item 11). Da dieser Faktor ohnehin nur aus drei Items besteht (Item 81, Item 88, Item 11) und somit bereits an der Grenze zur empirischen Unteridentifizierung liegt, bedeutet die Entscheidung für die Eliminierung des Items 11 aufgrund der signifikanten Nebenladung gleichzeitig auch die Entscheidung für die Eliminierung des Faktors „Bereitschaft zur Mitarbeit“.

Somit ergibt sich für diesen Schritt der Itemselektion folgende Auswahl:

Item 17, Item 60, Item 09, Item 65, Item 36, Item 56, Item 21, Item 55, Item 45, Item 78, Item 87, Item 44, Item 70, Item 59, Item 11, Item 81, Item 88 (aufgrund der Faktoreliminierung). Diejenigen Items, die nur Nebenladungen auf dem Faktor „Bereitschaft zur Mitarbeit“ hatten, können aufgrund der Faktoreliminierung beibehalten werden (Item 13, Item 36, Item 27, Item 38, Item 69).

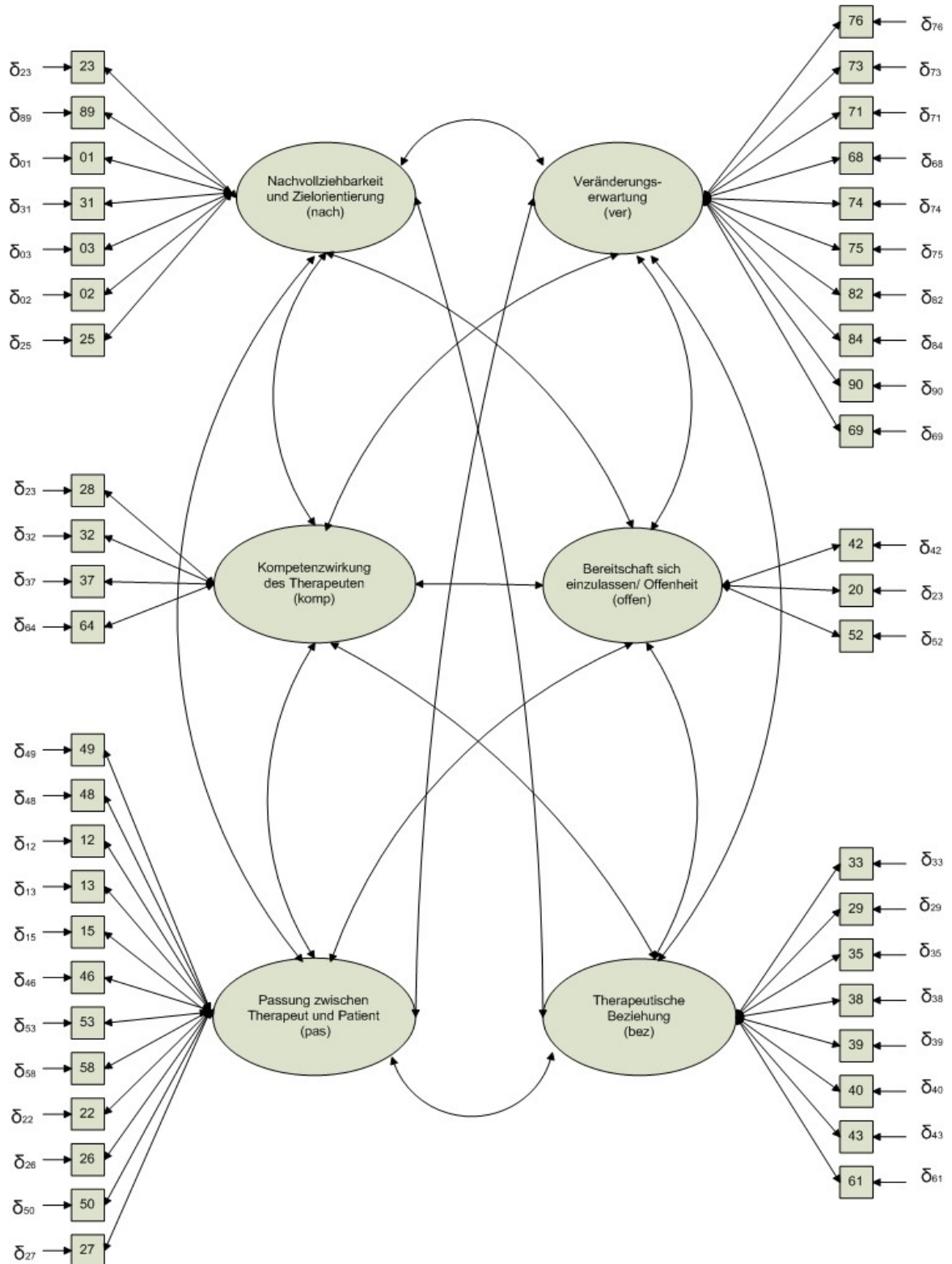
Somit ergibt sich eine Summe von 17 eliminierten Items und eine Summe von 44 Items und sechs Faktoren (Abkürzungen: „nach“, „pas“, „bez“, „ver“, „offen“ und „komp“) für die weitere Auswertung.

### **Modellspezifizierung**

#### Modell 1

Zunächst wird das Modell ohne korrelierte Fehlerterme spezifiziert (Modell 1). Den Ergebnissen der E/CFA-Analyse entsprechend besteht dieses Modell aus sechs Faktoren mit insgesamt 44 Items. Zur Bestimmung der Metrik der latenten Variablen (Faktoren) werden die Ankeritems aus der E/CFA-Untersuchung als Referenzindikatoren verwendet, was bedeutet, dass deren unstandardisierte Faktorladungen auf eins festgesetzt werden. Die Interkorrelationen zwischen den Faktoren werden ebenso wie die Faktorladungen der Indikatoren frei geschätzt, während die Interkorrelationen der Items (=Fehlerterme) auf null fixiert werden. In Abbildung 4 ist das Pfaddiagramm des Modells 1 dargestellt.

Abbildung 4:  
Pfaddiagramm Modell 1

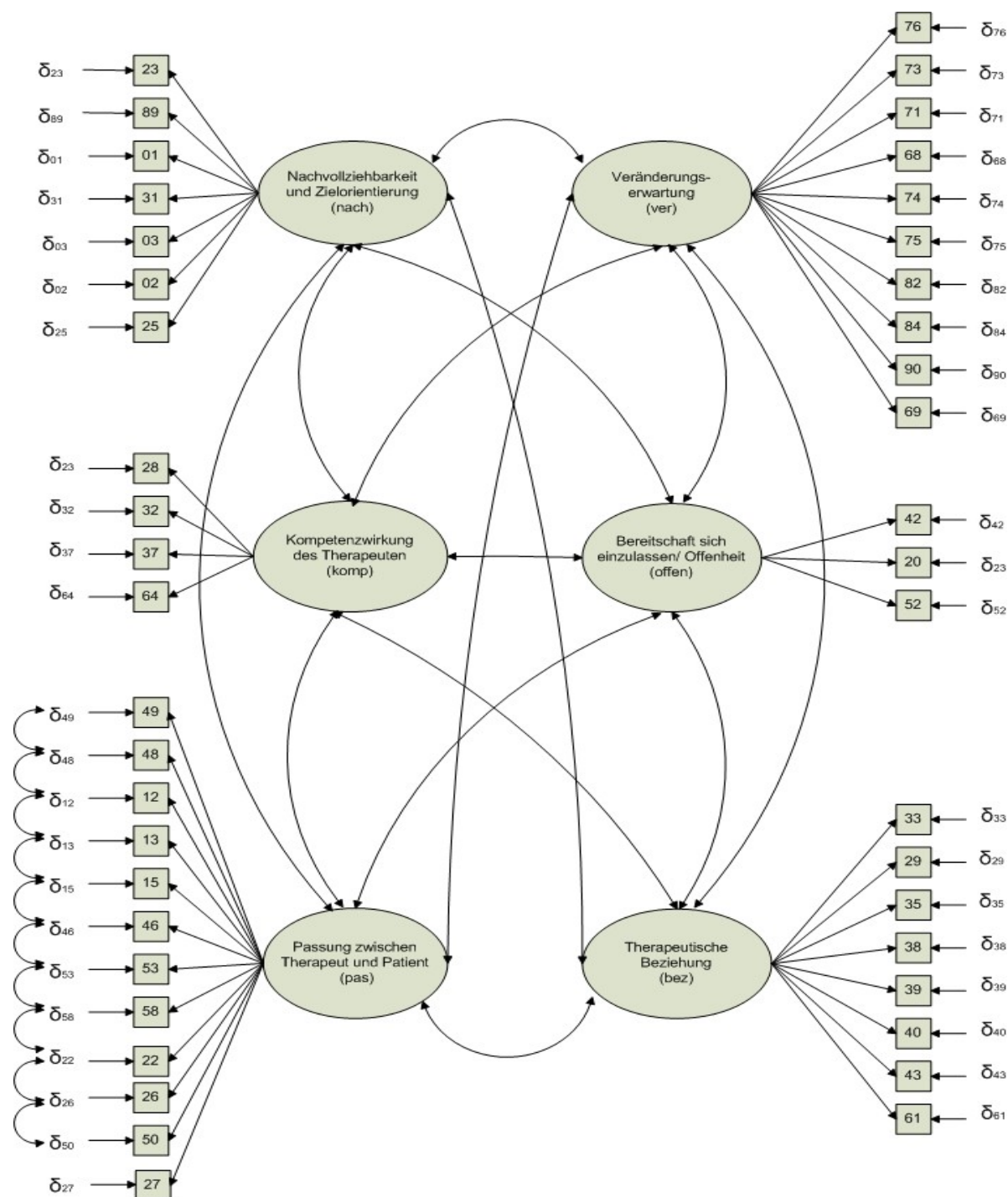


## Modell 2

Modell 2 setzt sich aus der gleichen Anzahl an Indikatoren und Faktoren mit derselben Zuordnung zusammen. Der Unterschied besteht lediglich darin, dass die Korrelationen der negativ formulierten Items nicht auf null fixiert, sondern frei geschätzt werden. Dadurch soll das Vorliegen eines Methodeneffektes aufgrund der negativen Itemformulierungen überprüft werden.

In Abbildung 5 ist das Pfaddiagramm des Modells 2 dargestellt.

Abbildung 5 : Pfaddiagramm Modell 2



Um das Pfaddiagramm übersichtlich zu halten, ist pro negativ formuliertem Item lediglich eine Korrelation zwischen den Fehlertermen dargestellt. Eine vollständige Auflistung würde 10 Doppelpfeile pro Fehlerterm enthalten (= alle Fehlerterme der neg. formulierten Items korrelieren miteinander), was zu einer unübersichtlichen graphischen Darstellung geführt hätte.

## Modellevaluation

### Modell 1: Modell mit unkorrelierten Fehlertermen

Die Ergebnisse der Modellschätzung ( $\chi^2$ -Wert und Fit-Indices) für Modell 1 sind in Tabelle 29 aufgeführt.

Tabelle 29:  
Goodness-of-fit-Indices (CFA) für das Modell 1

	$\chi^2$	df	p	CFI	TLI	RMSEA	WRMR
Modell 1	227.085	106	0.000	0.943	<b>0.975</b>	0.073	1.005

CFI = Comparative-Fit-Index (Cut-Off > .95); TLI = Tucker-Lewis-Index (Cut-Off > .95); RMSEA = root mean square of approximation (Cut-Off < 0.06); WRMR = Weighted Root Mean Square Residual (Cut-Off < 1.0). Werte, die für einen guten Modellfit sprechen sind fett gedruckt.

Wie aus Tabelle 29 ersichtlich, weist nur der CFI auf einen akzeptablen Fit hin. Alle anderen Werte sprechen für eine schlechte Modellpassung. Alle Parameterschätzungen weisen auf beiden Signifikanzniveaus (Est./SE >  $\pm 1.96$  ( $\alpha = .05$ ) bzw.  $\pm 2.58$  ( $\alpha = .01$ )) signifikante und hinsichtlich der Richtung erwartungsgemäße Werte auf und es zeigen sich keine „Out-of-range“-Parameter. Es wurden insgesamt 13 Modificationsindizes > 5.00 angegeben, die bedeutsame Nebenladungen für die Items 48, 27, 12 und 58 aufzeigen.

### Modell 2: Modell mit korrelierten Fehlertermen für die negativ formulierten Items

Die Ergebnisse der Modellschätzung ( $\chi^2$ -Wert und Fit-Indices) für Modell 2 sind in Tabelle 30 aufgeführt.

Tabelle 30:  
Goodness-of-fit-Indices (CFA) für das Modell 2

	$\chi^2$	df	p	CFI	TLI	RMSEA	WRMR
Modell 2	200.47	105	0.000	<b>0.955</b>	<b>0.980</b>	0.065	<b>0.899</b>

CFI = Comparative-Fit-Index (Cut-Off > .95); TLI = Tucker-Lewis-Index (Cut-Off > .95); RMSEA = root mean square of approximation (Cut-Off < 0.06); WRMR = Weighted Root Mean Square Residual (Cut-Off < 1.0). Werte, die für einen guten Modellfit sprechen sind fett gedruckt.

Aus Tabelle 30 wird ersichtlich, dass drei der vier Fitindizes für einen guten Modellfit sprechen. Die Parameterschätzungen der Faktorladungen und Faktorinterkorrelationen weisen auf beiden Signifikanzniveaus (Est./S.E. > ±1.96 ( $\alpha = .05$ ) bzw. ±2.58 ( $\alpha = .01$ )) signifikante und hinsichtlich der Richtung erwartungsgemäße Werte auf und es zeigen sich keine „Out-of-range“-Parameter. Die Höhe der Faktorladungen und der Faktorinterkorrelationen ist der Abbildung 5 zu entnehmen. Für Item 48 ergaben sich allerdings negative, und somit nicht erwartungsgemäße Korrelationen mit den anderen negativ formulierten Items. Dies weist auf Spezifizierungsprobleme bezüglich des Items 48 hin.

Es wurde ein Modifikationsindex > 5.00 für Item 48 angegeben. Zudem wurden zwei On/by-Statements (Hinweise auf funktionale Abhängigkeiten zwischen den latenten Variablen) ausgegeben, die aber – möglicherweise aufgrund einer Unteridentifizierung- nicht berechnet werden können.

### Vergleich Modell 1 mit Modell 2

Um eine Entscheidung treffen zu können, ob die augenscheinlich vorhandene bessere Passung von Modell 2 gegenüber Modell 1 auch statistisch nachzuweisen ist, wird ein  $\chi^2$ -Differenztest durchgeführt. Das Ergebnis ist in Tabelle 31 aufgeführt.

Tabelle 31:

#### Ergebnis des $\chi^2$ -Differenztests zwischen Modell 1 und Modell 2

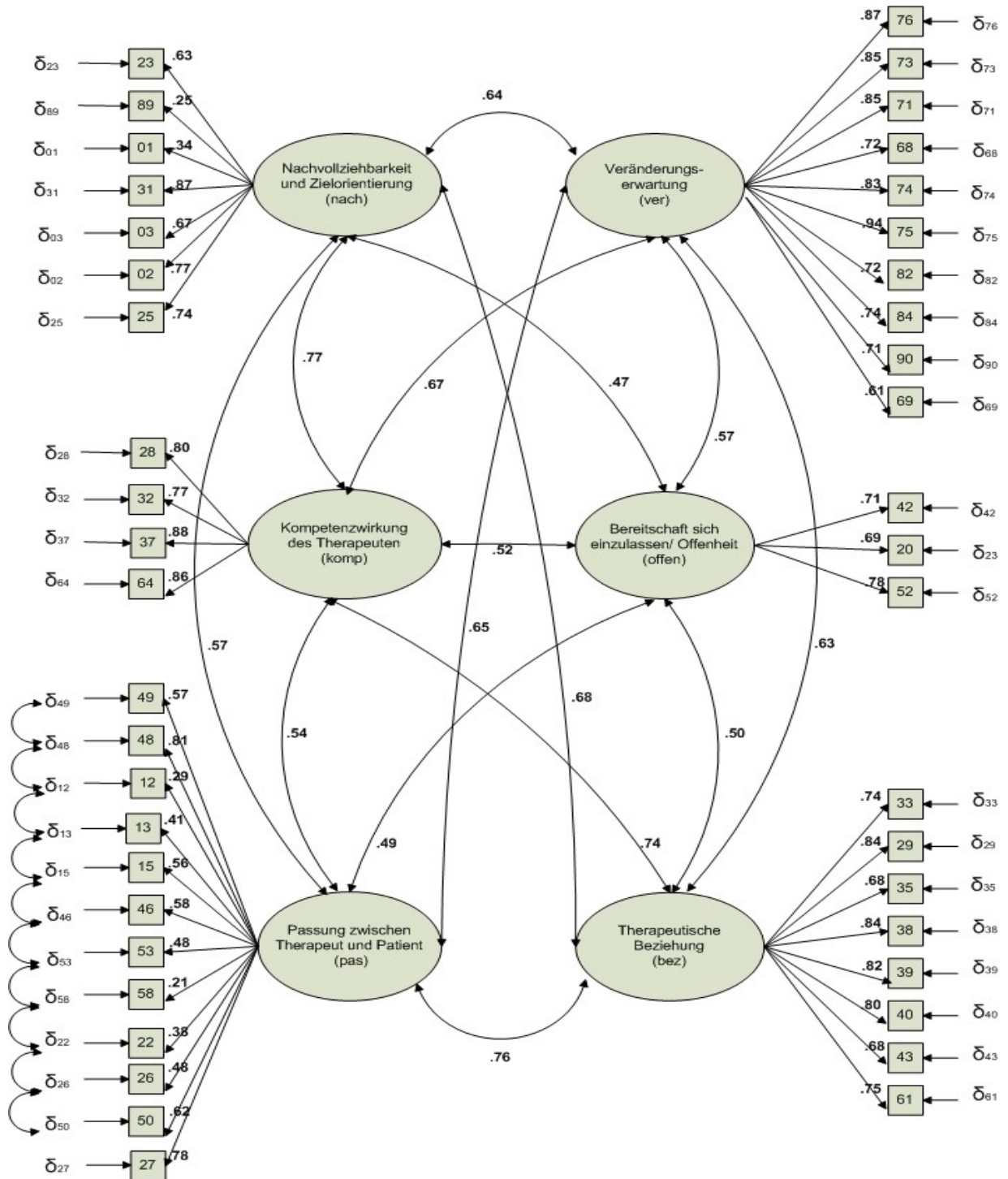
$\chi^2_{DIFF}$	df	p
86.74	21	< 0.001

Das Ergebnis des Differenztests bedeutet, dass die Freisetzung der fixierten Parameter (d.h. das Zulassen von Korrelationen zwischen den Fehlertermen der negativ formulierten Items) zu einer signifikant besseren Modellpassung führte ( $\alpha = 0.001$ ). Inhaltlich ist dies als Beleg für das Vorliegen eines Methodeneffektes zu sehen.

Aufgrund dieses Ergebnisses wird das Modell 2, das Korrelationen zwischen den negativ formulierten Items zulässt, als das passende Modell für die weitere Auswertung gewählt.

In Abbildung 6 ist das Pfaddiagramm des Modells 2 mit komplett standardisierten Parameterschätzungen (Faktorladungen und Faktorinterkorrelationen) dargestellt.

Abbildung 6 : Pfaddiagramm Modell 2 mit Parameterschätzungen



Um das Pfaddiagramm übersichtlich zu halten, ist pro negativ formuliertem Item lediglich eine Korrelation zwischen den Fehlertermen dargestellt. Eine vollständige Auflistung würde 10 Doppelpfeile pro Fehlerterm enthalten (= alle Fehlerterme der neg. formulierten Items korrelieren miteinander), was zu einer unübersichtlichen graphischen Darstellung geführt hätte.



### Respezifizierung des Modells

Die Respezifizierung des Modells und die damit verbundene Itemselektion bzw. Itemeliminierung erfolgt in mehreren Schritten. Als Kriterien für die Eliminierung werden an erster Stelle die Modifikationsindizes ( $> 5.00$ ) und dann die Faktorladungen ( $< .50$ ) herangezogen, ebenso wie die Signifikanz der Parameterschätzungen. Als weitere Kriterien werden die Gesamtanzahl der Indikatoren und die „Skalenlänge“ berücksichtigt. Zur abschließenden Beurteilung des Gesamtmodells werden außerdem die Fit-Indizes und die Höhe der Faktorinterkorrelationen betrachtet. Gemäß dem beschriebenen Vorgehen werden sukzessive 13 Items eliminiert. Dies sind in der entsprechenden Reihenfolge: Item 48, Item 58, Item 89, Item 12, Item 01, Item 22, Item 13, Item 26, Item 53, Item 69, Item 84, Item 31 und Item 50.

### Evaluation des respezifizierten Modells

Die Ergebnisse der Modellschätzung ( $\chi^2$ -Wert und Fit-Indices) für das respezifizierte Modell sind in Tabelle 32 aufgeführt.

Tabelle 32:

Goodness-of-fit-Indices (CFA) für das respezifizierte Modell (2R)

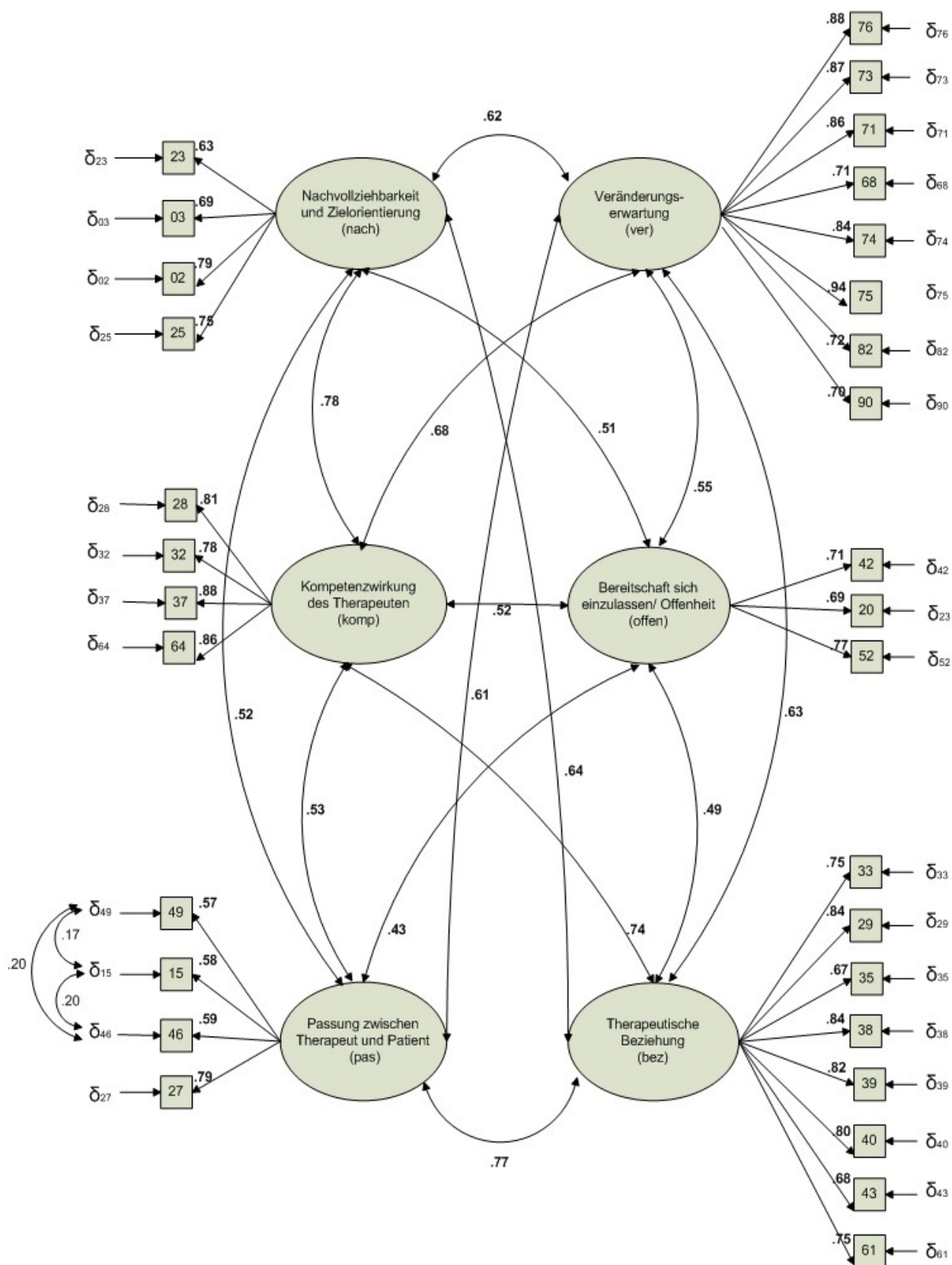
	$\chi^2$	df	p	CFI	TLI	RMSEA	WRMR
Modell 2R	141.18	81	0.000	<b>0.969</b>	<b>0.989</b>	<b>0.059</b>	<b>0.801</b>

CFI = Comparative-Fit-Index (Cut-Off  $> .95$ ); TLI = Tucker-Lewis-Index (Cut-Off  $> .95$ ); RMSEA = root mean square of approximation (Cut-Off  $< 0.06$ ); WRMR = Weighted Root Mean Square Residual (Cut-Off  $< 1.0$ ). Werte, die für einen guten Modellfit sprechen sind fett gedruckt.

Wie Tabelle 32 zu entnehmen ist, sprechen alle vier Fit-Indizes für eine gute Übereinstimmung des Modells mit den Daten.

Alle Parameterschätzungen erweisen sich als signifikant (z-Werte zwischen 4.10 und 9.34 für die Parameterschätzungen der Faktorladungen) und zeigen hinsichtlich Höhe und Richtung erwartungsgemäße Werte. Alle Faktorladungen liegen über  $.50$ . Die Faktorinterkorrelationen liegen zwischen  $.43$  und  $.78$ . Die Faktorladungen und die Faktorinterkorrelationen sind in Abbildung 7 dargestellt.

Abbildung 7 : Pfaddiagramm Modell 2R mit Parameterschätzungen



In Tabelle 33 sind die Skalen- bzw. Faktoren-Interkorrelationen der Übersichtlichkeit halber noch einmal separat aufgeführt.

Tabelle 33:  
Interkorrelationen der Skalen/Faktoren

	nach	ver	pas	bez	komp	offen
nach	-					
ver	.62	-				
pas	.52	.61	-			
bez	.64	.63	.77	-		
komp	.78	.68	.53	.74	-	
offen	.51	.55	.43	.49	.53	-

**Abkürzungen:** nach = „Nachvollziehbarkeit/Zielloerientierung der Therapie“; ver = „Veränderungserwartung“; pas = „Passung zwischen Therapeut und Patient“; bez = „Therapeutische Beziehung“; komp = „Kompetenzwirkung des Therapeuten“; offen = „Bereitschaft des Pat. sich einzulassen/ sich zu öffnen“.

Es ergeben sich keine Modifikationsindizes  $> 5.00$ . Es werden allerdings noch vier Nebenladungen und eine Residuenkorrelation ausgegeben, die aber hinsichtlich ihrer Höhe als nicht bedeutsam eingeschätzt werden (StdYX E.P.C.: 0.15 für die Residuenkorrelation und .19 -.28 für die Nebenladungen).

Das respezifizierte Modell enthält 31 Indikatoren und 6 Faktoren.

### Abschließende Bewertung des Modells

Insgesamt liefert die abschließende Modellevaluation zufriedenstellende Ergebnisse. Die Fit-Indices liegen alle innerhalb der vorgegebenen Wertebereiche, die Faktorladungen weisen alle Beiträge über dem festgesetzten Mindestwert auf, es zeigen sich durchweg signifikante Parameterschätzungen und es treten keine „Out-of-range-Parameter“ auf. Der  $\chi^2$ -Wert erweist sich allerdings als signifikant, was bedeutet, dass kein exakter Modell-Fit erzielt werden konnte. Wie im vorangegangenen Kapitel bereits diskutiert wurde, kann ein signifikanter  $\chi^2$ -Wert bei einem relativ komplexen Modell, wie dem vorliegenden, aber toleriert werden, sofern die Fit-Indizes für ein gut passendes Modell sprechen (vgl. Brown, 2006).

Die Faktorinterkorrelationen liegen mit .43-.78 im moderaten bis hohen Bereich, wobei der kritische Wert von .80 nicht überschritten wird. Die höchsten Korrelationen finden sich zwischen den Faktoren „Therapeutische Beziehung“ (bez) und „Passung zwischen Therapeut und Patient“

(pas), „Therapeutische Beziehung“ (bez) und „Kompetenzwirkung des Therapeuten“ (komp), sowie „Kompetenzwirkung des Therapeuten“ (komp) und „Nachvollziehbarkeit/ Ziellorientierung der Therapie“ (nach). Inhaltlich ist die Nähe zwischen diesen Faktoren nachvollziehbar und plausibel. Da die Interkorrelationen aber unterhalb des kritischen Wertes von .80 liegen, kann dennoch davon ausgegangen werden, dass die betreffenden Faktoren trotz ihrer inhaltlichen Nähe ausreichend unterscheidbare Konstrukte darstellen.

Es ergeben sich im finalen Modell keine Modifikationsindizes  $> 5.00$ , was bedeutet, dass es keine Hinweise auf das Vorliegen größerer Fehlspezifikationen gibt. Es werden noch fünf Modifikationsindizes mit Indexwerten  $< 5.00$  angezeigt (vier Nebenladungen und eine Residuenkorrelation), die hinsichtlich ihrer E.P.C.-Werte als nicht salient betrachtet werden (Nebenladungen  $< .30$ ). Daher wird auf eine weitergehende Modellveränderung ausgehend von diesen Modifikationsindizes verzichtet. Dies erscheint auch aufgrund der Gefahr einer Überbewertung stichprobenbedingter Besonderheiten sinnvoll („capitalization on change“, vgl. MacCallum et al., 1992).

Die Anzahl von 31 Indikatoren für die sechs Faktoren/Skalen liegt im angestrebten Bereich von 30-40 Items. Die Anzahl der Indikatoren pro Faktor schwankt zwischen 3 und 8. Eine gleichmäßige Verteilung von 5 (bzw. 6) Items pro Faktor wäre in Hinblick auf die Skalenstabilität wünschenswerter. Den anderen, bereits aufgeführten, Selektionskriterien wurde aber bei der Itemeliminierung mehr Gewicht zugesprochen, weshalb die unterschiedliche Anzahl von Items pro Faktor akzeptiert wird. Zu bemerken ist aber, dass der Faktor „Bereitschaft des Pat. sich einzulassen/sich zu öffnen“ (offen) mit der Menge von drei Indikatoren an der Grenze zur empirischen Untereidentifizierung liegt, wobei die Ladungen zwischen .69 und .77 für seine Stabilität sprechen. Im Vergleich mit dem Modell ohne korrelierte Fehlerterme bzw. Residuen (Modell 1) hat sich das Modell mit korrelierten Fehlern (Modell 2) als das passendere erwiesen. Dies bedeutet, dass es sinnvoll ist, die auf die negative Itemformulierung zurückzuführende gemeinsame Varianz der betreffenden Items im Modell zu berücksichtigen (Methodeneffekt).

## **6. Darstellung der revidierten Form des Fragebogens**

### **6.1. Darstellung der Skalen und Items**

Die in den vorangegangenen Kapiteln beschriebenen Schritte der Auswertung und Itemselektion liefern im Endergebnis sechs Skalen (Faktoren) mit insgesamt 31 Items (Indikatoren). Abschlie-

## 6. Darstellung der revidierten Form des Fragebogens

End sollen die Skalenbezeichnungen auf ihre Passung und ihre „Handhabbarkeit“ hin überprüft und gegebenenfalls angepasst werden.

Im Folgenden sind die sechs Skalen und die dazu gehörigen Items mit Item-Nummern, Formulierungen und Faktorladungen (a) aufgeführt. Sofern eine neue Skalenbezeichnung gewählt wird, ist die alte Bezeichnung in Klammern aufgeführt.

Tabelle 34:

### Darstellung der Skalen und Items des revidierten Fragebogens

#### **1. Faktor: „Veränderungserwartung des Patienten“ : 8 Items**

<b>Item-Nr.</b>	<b>Itemformulierung</b>	<b>a</b>
Item 76	„In der Woche vor der letzten Therapiesitzung ... habe ich geglaubt, dass die Therapie zu einer Verbesserung der Symptome führen wird.“	.88
Item 73	„In der Woche vor der letzten Therapiesitzung ... habe ich mir gedacht, dass ich am Ende der Therapie mit meinem Alltag besser zurecht kommen werde.“	.87
Item 71	„In der Woche vor der letzten Therapiesitzung ... war ich überzeugt, dass es mir am Ende der Therapie deutlich besser gehen wird.“	.86
Item 68	„In der Woche vor der letzten Therapiesitzung ... habe ich in der Therapie einen Weg gesehen, um mit meinen Schwierigkeiten fertig zu werden.“	.71
Item 74	„In der Woche vor der letzten Therapiesitzung ... war ich überzeugt, dass sich durch die Therapie bald Verbesserungen einstellen werden.“	.84
Item 75	„In der Woche vor der letzten Therapiesitzung ... hatte ich das Gefühl, dass die Therapie mir helfen wird.“	.94
Item 82	„In der Woche vor der letzten Therapiesitzung ... habe ich gedacht, dass ich mich am Ende der Therapie viel wohler fühlen werde.“	.72
Item 90	„In der Woche vor der letzten Therapiesitzung ... war ich überzeugt, dass sich mein Leben durch die Therapie nachhaltig verändern wird.“	.70

#### **2. Faktor: „Passung zwischen Therapeut und Patient“: 4 Items**

<b>Item-Nr.</b>	<b>Itemformulierung</b>	<b>a</b>
Item 49	„In der letzten Therapiestunde ... hatte ich das Gefühl, dass ich mich verteidigen muss.“	.57
Item 15	„In der letzten Therapiestunde ... hat mich das Verhalten meines Therapeuten irritiert.“	.58
Item 46	„In der letzten Therapiestunde ... habe ich mich unwohl gefühlt.“	.59
Item 27	„In der letzten Therapiestunde ... habe ich mich sicher gefühlt.“	.79

#### **3. Faktor: „Therapeutische Beziehung“: 8 Items**

<b>Item-Nr.</b>	<b>Itemformulierung</b>	<b>a</b>
-----------------	-------------------------	----------

## 6. Darstellung der revidierten Form des Fragebogens

Item 33	„In der letzten Therapiestunde ... habe ich gespürt, dass mein Therapeut mich schätzt.“	.75
Item 29	„In der letzten Therapiestunde ... habe ich mich gut mit meinem Therapeuten verstanden.“	.84
Item 35	„In der letzten Therapiestunde ... hatte ich das Gefühl, dass mein Therapeut mir seine ehrliche Meinung sagt.“	.67
Item 38	„In der letzten Therapiestunde ... fühlte ich mich von meinem Therapeuten als Mensch akzeptiert.“	.84
Item 39	„In der letzten Therapiestunde ... habe ich gespürt, dass ich meinem Therapeuten vertrauen kann.“	.82
Item 40	„In der letzten Therapiestunde ... haben wir gut zusammengearbeitet.“	.80
Item 43	„In der letzten Therapiestunde ... habe ich gemerkt, dass meinem Therapeut positive Veränderungen genauso wichtig sind wie mir.“	.68
Item 61	„In der letzten Therapiestunde ... hat sich mein Therapeut sehr einfühlsam verhalten.“	.75

### 4. Faktor: „Fachkompetenz des Therapeuten“ (Kompetenzwirkung d.T.): 4 Items

Item-Nr.	Itemformulierung	a
Item 28	„In der letzten Therapiestunde ... habe ich gemerkt, dass mein Therapeut viel Erfahrung hat.“	.81
Item 32	„In der letzten Therapiestunde ... hatte ich den Eindruck, dass mein Therapeut ein Spezialist für meine Beschwerden ist.“	.78
Item 37	„In der letzten Therapiestunde ... habe ich gemerkt, dass mein Therapeut viel über psychische Probleme weiß.“	.88
Item 64	„In der letzten Therapiestunde ... habe ich gemerkt, dass mein Therapeut gut ausgebildet ist.“	.86

### 5. Faktor: „Überzeugungskraft des Therapeuten“ (Nachvollziehbarkeit/Zielorientierung der Therapie): 4 Items

Item-Nr.	Itemformulierung	a
Item 03	„In der letzten Therapiestunde ... hat mein Therapeut schwierige Zusammenhänge verständlich erklärt.“	.70
Item 02	„In der letzten Therapiestunde ... wirkte das Auftreten meines Therapeuten professionell.“	.80
Item 23	„In der letzten Therapiestunde ... ging mein Therapeut planvoll und zielorientiert vor.“	.63
Item 25	„in der letzten Therapiestunde ... wirkte mein Therapeut überzeugt von seinem Vorgehen.“	.75

### 6. Faktor: „Bereitschaft des Patienten sich einzulassen“ (Bereitschaft des Pat. sich einzulassen/ sich zu öffnen“): 3 Items

Item-Nr.	Itemformulierung	a
Item 20	„In der letzten Therapiestunde... habe ich mich getraut, Dinge zu sagen, die mir peinlich	.69

	<i>sind.</i> “	
Item 42	„In der letzten Therapiestunde... bin ich das Risiko eingegangen, auch unangenehme Gefühle zu erleben.“	.71
Item 52	„In der letzten Therapiestunde... konnte ich über meinen Schatten springen.“	.77

## 6.2. Beurteilung des endgültigen Fragebogens

### 6.2.1. Gütekriterien

#### 6.2.1.1. Objektivität

Die Objektivität eines Tests oder Fragebogens gibt an, inwieweit die Ergebnisse unabhängig von der Person des Untersuchers sind (vgl. Bühner, 2006). Dabei wird zwischen Durchführungs-, Auswertungs- und Interpretationsobjektivität unterschieden.

#### **Durchführungsobjektivität**

Die Durchführungsobjektivität ist dann gewährleistet, wenn die Bedingungen, unter denen ein Test oder Fragebogen durchgeführt wird, für alle Probanden gleich sind. Dies ist bei der Durchführung von Leistungstests von größerer Bedeutung als für Fragebögen, die Eigenschaften oder Einstellungen erfassen, da im ersten Fall z.B. Zeitbegrenzungen und Hilfestellungen einen Einfluss auf das Testergebnis haben. Im Falle von Fragebögen ist in erster Linie die Instruktion von Bedeutung.

Da die erste Version des vorliegenden Fragebogens von den betreffenden Therapeuten an die Patienten verteilt wurde, konnte die Instruktion nur in Textform vorgegeben werden. Auf der ersten Seite des Fragebogens befand sich die Instruktion für die Patienten (siehe Anhang B). Den Patienten wurde das Vorgehen erläutert, sie wurden gebeten, spontan und aufrichtig zu antworten, mit dem Hinweis, dass es keine richtigen und falschen Antworten gibt. Diese Ausführungen sollten die Ausgangsbedingungen für die Patienten vor dem Ausfüllen möglichst ähnlich machen. Die Bedingungen, unter denen die Patienten den Fragebogen dann tatsächlich ausfüllten, konnten nicht überprüft werden.

Die Therapeuten wurden separat über das Vorgehen informiert (siehe Anhang B). Ihnen wurde inhaltlich, aber nicht wortwörtlich vorgegeben, was sie den Patienten mitteilen sollten. Mit Hilfe dieses Vorgehens sollte, soweit möglich, die Durchführungsobjektivität gewährleistet werden.

Dennoch ist festzuhalten, dass an diesem Punkt ein gewisses „Standardisierungsdefizit“ blieb, da die Art und Weise, wie die Therapeuten das Vorgehen den Patienten vermittelt haben, nicht kontrolliert werden konnte. Eine standardisierte Form der Vorgabe durch eine andere Person als den Therapeuten wäre sowohl hinsichtlich des Datenschutzes, als auch hinsichtlich der Organisation schwierig gewesen, so dass in diesem Fall keine andere Lösung praktikabel erschien.

### **Auswertungsobjektivität**

Ein Test bzw. Fragebogen ist dann hinsichtlich seiner Auswertung objektiv, wenn unterschiedliche Auswerter zum gleichen formalen Ergebnis (z.B. Punktwert) gelangen.

Im Falle eines Fragebogens bedeutet dies, dass gleichen Itemantworten gleiche numerische Werte zugeordnet werden. Dies ist bei dem vorliegenden Fragebogen durch die Wahl des numerischen Antwortformats (7-stufige-Likertskala) gegeben, so dass von einer hohen Auswertungsobjektivität ausgegangen werden kann.

### **Interpretationsobjektivität**

Von Interpretationsobjektivität spricht man, wenn verschiedene Beurteiler aus den Ergebnissen die gleichen Schlüsse ziehen. Eine hohe Interpretationsobjektivität setzt eine Normierung und ausreichend geprüfte Gütekriterien, sowie standardisierte Interpretationsmöglichkeiten voraus (vgl. Bühner 2006).

Ein wichtiger Schritt in Richtung Sicherung der Gütekriterien wurde bei dem vorliegenden Fragebogen mit der Konstruktion auf der Basis der faktorenanalytisch gewonnenen Ergebnisse getan. Den nächsten Schritt stellen eine Validierungsstudie und schließlich eine Normierung anhand einer großen Normstichprobe dar.

Im Kontext der Darstellung der Objektivitätsaspekte soll abschließend noch die Beteiligung der behandelnden Therapeuten angesprochen werden. Im optimalen Fall sollten die Therapeuten in keiner Weise an der Datenerhebung und -auswertung beteiligt sein und die Patienten sollten sich sicher sein können, dass die Daten anonym bleiben. Durch die Trennung zwischen Datenerhebung bzw. -auswertung und Therapie sollte eine wie auch immer geartete Beeinflussung durch den behandelnden Therapeuten ausgeschlossen werden. Die Zusicherung der Anonymität sollte den Einfluss der Tendenz zur Abgabe sozial erwünschter Antworten reduzieren.

In der vorliegenden Untersuchung wurde versucht, dies soweit als möglich zu berücksichtigen. Die Patienten wurden in der Instruktion darüber informiert, dass aufgrund der anonymen Form der Datenerhebung keine Rückschlüsse auf einzelne Personen möglich seien und, dass der Therapeut den Fragebogen nicht zu lesen bekäme, da dieser in einem verschlossenen Umschlag zu-



rückgegeben werde. Zudem wurde die Dateneingabe nicht vom behandelnden Therapeuten durchgeführt.

Allerdings blieb, wie oben beschrieben, aus organisatorischen und datenschutzrechtlichen Gründen keine andere Möglichkeit, als dass die Patienten von ihren Therapeuten über die Studie informiert und instruiert wurden.

### 6.2.1.2. Reliabilität

Als Maß der klassischen Testtheorie für die Messgenauigkeit eines Verfahrens gibt die Reliabilität den Anteil der Varianz der wahren Werte an der Varianz der beobachteten Werte an. Die meisten Methoden zur Reliabilitätsschätzung beruhen auf Korrelationen. Dafür wird ein Merkmal mehrfach gemessen, um so – in der Annahme, dass sich das Merkmal selbst nicht verändert – eine Schätzung des wahren Wertes und des Fehlers zu erhalten. Zu berücksichtigen ist bei der Durchführung einer Reliabilitätsschätzung immer die Äquivalenz, d.h. die Gleichwertigkeit der Messungen, da nur bei dem Vorhandensein gewisser Äquivalenzeigenschaften eine präzise Schätzung möglich ist (vgl. Bühner, 2006).

Man unterscheidet hinsichtlich der Äquivalenz von Messungen fünf Stufen, nämlich die streng parallele, die essenziell parallele, die tau-äquivalente, die essenziell tau-äquivalente und die tau-kongenerische Messung<sup>45</sup>. In Bezug auf die Ergebnisse einer konfirmatorischen Faktorenanalyse bedeutet die essenzielle Parallelität, dass alle Testteile, d.h. Items eines Faktors gleich hohe Faktorladungen und gleich hohe Fehlervarianzen aufweisen. Eine essenziell tau-äquivalente Messung liegt dann vor, wenn die Items eines Faktors bei unterschiedlich hohen Fehlervarianzen gleich hohe Ladungen aufweisen. Und tau-kongenerische Items hingegen müssen eindimensional sein d.h. nur auf einem Faktor laden, wobei sich aber die Faktorladungen und Fehlervarianzen unterscheiden können (Osburn, 2000, S.345, zit. nach Bühner, 2006).

Bei der Entwicklung eines Fragebogens wird meist eine Reliabilitätsschätzung durch die Berechnung eines bzw. mehrerer Konsistenzkoeffizienten (im Falle mehrerer Skalen) vorgenommen.

---

45 Bei der streng parallelen Messung wird davon ausgegangen, dass die betreffenden Tests bzw. Items „dieselbe Eigenschaft mit der derselben Skala und für alle Personen gleich gut messen“ (Lord & Novick, 1968, S.48, zit. nach Bühner, 2006). Entsprechend sollten die Mittelwerte und die Varianzen der beiden Tests bzw. der Items gleich sein. Im Fall der essenziell parallelen Messung kann sich der wahre Wert im Vergleich zwischen den Messungen um eine additive Konstante verschieben. Liegen tau-äquivalente Messungen vor, so bedeutet dies, dass die Messfehler zwischen den Messungen variieren, wobei aber die Mittelwerte der Tests bzw. Items gleich sind. Bei essenziell tau-äquivalenten Messungen liegt wiederum im Vergleich zu tau-äquivalenten Messungen eine Verschiebung um eine additive Konstante vor. Bei tau-kongenerischen Messungen dürfen die Tests bzw. Items unterschiedliche Maßeinheiten, Mittelwerte und Fehlervarianzen aufweisen, wobei davon ausgegangen wird, dass die wahren Werte perfekt miteinander korrelieren.

Dabei wird in den meisten Fällen die Formel für Cronbachs- $\alpha$  verwendet (Cronbach, 1951). Cronbachs- $\alpha$  gibt an, wie eng die Items eines Faktors miteinander zusammenhängen, ist aber kein Maß für die Eindimensionalität oder Homogenität eines Faktors bzw. einer Skala, was oft fälschlicherweise angenommen wird (Cortina, 1993). Weiterhin wird oft bei der Verwendung dieses Reliabilitätsmaßes übersehen, dass Cronbachs- $\alpha$  nur dann eine Schätzung der wahren Reliabilität liefert, wenn die Items (d.h. die Messungen) essenziell tau-äquivalent sind (Osburn, 2000, S.345, zit. nach Bühner, 2006). Ist diese Voraussetzung nicht erfüllt, so stellt Cronbachs- $\alpha$  nur einen unteren Grenzwert der Reliabilität dar, d.h. die wahre Reliabilität fällt mindestens so hoch aus oder höher (McDonald, 1981). Zu bedenken ist weiterhin, dass es bei der Berechnung von Cronbachs- $\alpha$  zu Über- oder Unterschätzung der wahren Reliabilität kommen kann, wenn die Annahme der Unkorreliertheit der Messfehler verletzt wird (Shevlin et al., 2000).

Um diesen Problemen aus dem Weg zu gehen, wurden für die Reliabilitätsschätzung der Faktoren eines konfirmatorischen Faktorenmodells andere Maße entwickelt, vor allem da die Voraussetzung der essenziellen Tau-Äquivalenz nur selten erfüllt ist, was im faktorenanalytischen Sinne die Annahme gleicher Faktorladungen für die Indikatoren einer latenten Variablen bedeutet (vgl. Hildebrandt & Temme, 2006).

Zu diesen alternativen Maßen zählt die Schätzung der Faktorreliabilität nach folgender Formel::

$$\rho_c = \frac{\left(\sum_{i=1}^q \lambda_{ij}\right)^2 \phi_{jj}}{\left(\sum_{i=1}^q \lambda_{ij}\right)^2 \phi_{jj} + \sum_{i=1}^q \theta_{ii}} \quad (\text{McDonald, 1970, zit. nach Hildebrandt \& Temme, 2006})$$

Dabei stellt  $\lambda_{ij}$  die Faktorladung dar,  $\Phi_{jj}$  die Varianz der latenten Variable und  $\theta_{ii}$  die Residualvarianz. Analog zu Cronbachs- $\alpha$  stellt die Faktorreliabilität ein Maß für die Reliabilität über alle Indikatoren eines Faktors dar und kann dabei Werte zwischen 0 und 1 annehmen. Als Anforderungskriterium hinsichtlich der Höhe der Reliabilität eines Faktors wird ein Schwellenwert von  $\geq .60$  angegeben (Bagozzi & Yi, 1988). Zu berücksichtigen sind allerdings bei der Reliabilitätsbeurteilung verschiedene Einflussgrößen, z.B. die Bandbreite des Konstrukts, die Redundanz der Formulierungen und die Anzahl der Indikatoren. Die Faktorreliabilität steigt, ebenso wie Cronbachs- $\alpha$ , mit zunehmender Anzahl an Indikatoren, weswegen man bei längeren Skalen strengere Kriterien ansetzen sollte als bei kurzen (vgl. Hildebrandt & Temme, 2006).

Da die Indikatoren der sechs Faktoren in der vorliegenden Untersuchung innerhalb jeder Skala unterschiedlich hohe Faktorladungen aufweisen (vgl. Tabelle 34), und somit nicht von essenziell

tau-äquivalenten Messungen ausgegangen werden kann, wird zur Reliabilitätsschätzung die oben dargestellte Formel zur Berechnung der Faktorreliabilität  $\rho_c$  statt der üblicherweise verwendeten Formel für Cronbachs- $\alpha$  verwendet. Die Ergebnisse sind in Tabelle 35 dargestellt.

Tabelle 35:  
Faktorreliabilitäten  $\rho_c$

1. Faktor: „Veränderungserwartung des Patienten“	(8 Items)	<b>.94</b>
2. Faktor: „Passung zwischen Therapeut und Patient“	(4 Items)	<b>.73</b>
3. Faktor: „Therapeutische Beziehung“	(8 Items)	<b>.92</b>
4. Faktor: „Fachkompetenz des Therapeuten“	(4 Items)	<b>.90</b>
5. Faktor: „Überzeugungskraft des Therapeuten“	(4 Items)	<b>.80</b>
6. Faktor: „Bereitschaft des Patienten sich einzulassen“:	(3 Items)	<b>.77</b>

Tabelle 35 macht deutlich, dass alle Werte über dem Schwellenwert von .60 liegen, so dass man bei allen Faktoren/Skalen von einer ausreichenden bis guten Höhe der Reliabilität sprechen kann.

### **6.2.1.3. Validität**

Grundsätzlich werden drei Arten von Validität unterschieden: Inhaltsvalidität, Konstruktvalidität und Kriteriumsvalidität (Bryant, 2000). Im Folgenden soll auf diese drei Validitätsformen in Hinblick auf den endgültigen Fragebogen kurz Bezug genommen werden.

#### **Inhaltsvalidität**

Unter Inhaltsvalidität wird die Fähigkeit eines Messinstruments verstanden, das in Frage stehende Merkmal hinreichend genau zu erfassen. Die ausgewählten Items sollen eine repräsentative Teilmenge des Universums aller denkbaren Items darstellen, die das Merkmal abbilden. Die Inhaltsvalidität lässt sich schwer in Form eines numerischen Wertes messen, sondern wird eher „aufgrund logischer und fachlicher Überlegungen“ bestimmt (nach Michel & Conrad, 1982, zitiert nach Bühner, 2006).

Um ein hohes Maß an Inhaltsvalidität zu gewährleisten, wurden die Items des Fragebogens unter Berücksichtigung verschiedener Informationsquellen generiert: Literatur in Form von theoretischen Abhandlungen und empirischen Studien, Messinstrumente mit ähnlichem Geltungsbereich und Erfahrungswerte aus der therapeutischen Praxis. Weiterhin wurden die entwickelten Items

mehrfach verschiedenen Psychologischen Psychotherapeuten<sup>46</sup> zur Einschätzung vorgelegt und entsprechend den Rückmeldungen verändert. Und schließlich stellte die durchgeführte Expertenbefragung ebenfalls eine Methode zur Sicherung der Inhaltsvalidität dar.

Im Prozess der Itemselektion, der primär auf der Basis faktorenanalytischer Verfahren, und somit in erster Linie hinsichtlich statistischer Kriterien erfolgte, wurden die Entscheidungen für oder gegen die Eliminierung von Items immer auch unter Berücksichtigung der Inhaltsvalidität getroffen.

### **Konstruktvalidität**

Der Begriff „Konstruktvalidität“ kann unterschiedlich weit gefasst werden: im weitesten Sinne wird er als Oberbegriff für alle Validitätsarten verwendet. Im engeren Sinne, auf den hier Bezug genommen werden soll, umfasst die Konstruktvalidität die konvergente, diskriminante und faktorielle Validität (vgl. Bühner, 2006). Ein Ziel bei der Untersuchung der Konstruktvalidität besteht in der Überprüfung theoretischer Annahmen über Zusammenhangsstrukturen latenter Konstrukte mit Hilfe empirischer Daten. Hierbei können unterschiedliche empirische Methoden eingesetzt werden, die ein hypothetisch-deduktives Vorgehen erlauben (vgl. Moosbrugger & Kelava, 2007). Ist die konvergente Validität eines Tests bzw. einer Skala hoch, so finden sich hohe Zusammenhänge mit ähnlichen Konstrukten. Entsprechend werden bei der diskriminanten Validität geringe Zusammenhänge mit fremden Konstrukten erwartet. Von faktorieller Validität spricht man dann, wenn eine aus der theoretischen Definition eines Konstrukts (bzw. mehrerer Konstrukte) abgeleitete Struktur durch eine Faktorenanalyse an empirischen Daten nachgewiesen werden kann (vgl. Moosbrugger & Kelava, 2007).

Der vorliegende Fragebogen wurde mit der Zielsetzung einer hohen faktoriellen Validität entwickelt. Das Ergebnis des abschließenden Modellvergleichs zwischen dem spezifizierten, also vorgegebenen Modell und den empirischen Daten in der konfirmatorischen Faktorenanalyse (vgl. Tabelle 32 ) spricht für eine gute Übereinstimmung und somit für das Vorhandensein einer faktoriellen Validität. Allerdings ist zu bedenken, dass im Rahmen der Itemselektion schrittweise Modellveränderungen vorgenommen wurden und es sich somit nicht mehr um ein konfirmatorisches Vorgehen im engeren Sinne handelt.

---

<sup>46</sup> Die beteiligten Psychotherapeuten waren zum Teil approbiert und zum Teil befanden sie sich noch in der Ausbildung.

Die Höhen der Korrelationen zwischen den Faktoren sprechen erwartungsgemäß für eine inhaltliche Nähe der Skalen (vgl. Tabelle 33), wobei eine ausreichende Unterscheidbarkeit der Konstrukte besteht ( $r < .80$ ), was als Hinweis auf das Vorliegen einer diskriminativen Validität zu werten ist. In der vorliegenden Untersuchung wurden parallel keine Daten erhoben, die eine Berechnung der konvergenten Validität zulassen würden (z.B. Fragebogen aus dem Bereich der „Alliance Scales“).

### **Kriteriumsvalidität**

Die Kriteriumsvalidität ist ein Maß für den Zusammenhang zwischen der durch den Test bzw. Fragebogen erfassten Merkmalsausprägung und externen Kriterien, die mit dem Merkmal in Zusammenhang stehen (vgl. Bühner, 2006). Diese Übereinstimmung wird in Form von Korrelationen erfasst, wobei die Messung des Kriteriums entweder im Vorfeld der Merkmalerfassung (retrospektive Validität), zeitgleich mit dieser (konkurrente Validität) oder später (prognostische Validität) erfolgen kann.

Die Überprüfung der Kriteriumsvalidität steht im Falle des vorliegenden Fragebogens noch aus. Von besonderem Interesse wäre hierbei eine Überprüfung des Zusammenhangs zwischen den erfassten Merkmalsbereichen (Skalen) und dem Therapieerfolg, was im weitesten Sinne eine Überprüfung der prognostischen Validität darstellen würde.

### **6.2.2. Inhaltliche Darstellung**

Im Folgenden sollen die einzelnen Skalen des Fragebogens inhaltlich beschrieben werden.

Die Skala „**Veränderungserwartung des Patienten**“ bezieht sich auf die Annahme des Patienten, dass die Therapie zu einer Verbesserung seiner Beschwerden führen wird. Es werden dabei verschiedene Aspekte berücksichtigt, z.B. die Erwartung von Symptomreduktion, die Annahme einer verbesserten Fähigkeit im Alltag zurechtzukommen oder eine generelle Hoffnung auf Hilfe und Erfolg durch die Therapie. Trotz dieser unterschiedlichen Facetten ist die Skala als relativ homogen anzusehen, was sich auch in der hohen Reliabilitätsschätzung niederschlägt.

Die Items der Skala „**Passung zwischen Therapeut und Patient**“ beschreiben das eher gefühlsbezogene als rational begründbare Erleben, in der Therapie am „richtigen Ort“ bzw. beim richtigen Therapeuten zu sein. Die Ursachen für dieses Gefühl von Stimmigkeit bzw. Nicht-Stimmigkeit werden dabei nicht angesprochen, sondern lediglich deren Folge. Die Gründe können sowohl auf Patientenseite, als auch auf Therapeutenseite zu finden sein und hängen inhaltlich mit

den anderen Skalen zusammen (z.B. therapeutische Beziehung). Denkbar ist allerdings auch, dass das Erleben von Nicht-Passung auf Ursachen zurückgeht, die nicht im Rahmen des Fragebogens erhoben werden (z.B. externe Faktoren), oder die Ursachen vom Patienten nicht benennbar sind, so dass ein Zugewinn an Informationen durch diese Items möglich ist.

Inhaltlich eng verbunden mit dem Faktor „Passung zwischen Therapeut und Patient“ ist die Skala „**Therapeutische Beziehung**“<sup>47</sup>. Darin finden sich acht Items, die beziehungsfördernde Verhaltensweisen des Therapeuten (z.B. Empathie) und deren Wahrnehmung (z.B. Wertschätzung, Akzeptanz, Echtheit und Vertrauen) beim Patienten beschreiben und die sich auf die Zusammenarbeit zwischen Therapeut und Patient beziehen („goal and task/collaboration“). Diese Facetten der Beziehung zwischen Therapeut und Patient finden sich in den meisten Konzeptionen zur therapeutischen Beziehung bzw. Allianz (z.B. Bordin, 1979).

Die Skala „**Fachkompetenz des Therapeuten**“ beschreibt mit vier Items die – vom Patienten wahrgenommene – fachliche Eignung des Therapeuten. Diese bezieht sich auf die Darstellung von Fachwissen, eine fundierte Ausbildung und die Demonstration von Erfahrung in der Behandlung psychischer Beschwerden. Eingeschätzt werden soll, inwieweit der Therapeut die Person des „legitimierten Heilers“ (vgl. Frank, 1961) repräsentiert.

Mit der Skala „**Überzeugungskraft des Therapeuten**“ hingegen soll das konkrete Vorgehen in der Therapie und die Glaubwürdigkeit des Therapeuten hinsichtlich dieses Vorgehens erfasst werden. Die Items beziehen sich sowohl auf die Nachvollziehbarkeit des therapeutischen Angebotes, als auch auf das Verhalten des Therapeuten: Kann der Therapeut den Eindruck vermitteln, dass er hinter seinem Konzept steht und selbst von der Wirksamkeit überzeugt ist? Die beiden Skalen „Fachkompetenz des Therapeuten“ und „Professionelles Auftreten des Therapeuten“ geben zwei Facetten des Therapeutenverhaltens bzw. der Therapeutenwirkung wieder, was eine gewisse inhaltliche Nähe mit sich bringt<sup>48</sup>.

---

47 Diese inhaltliche Nähe spiegelt sich auch in der Faktoreninterkorrelation von .77 wieder.

48 Ähnlich wie bei den Skalen „Passung zwischen Therapeut und Patient“ und „Therapeutische Beziehung“ spiegelt sich diese Nähe in der Skaleninterkorrelation wieder (.78).

Die Items der Skala „**Bereitschaft des Patienten sich einzulassen**“ beschreiben das Engagement des Patienten und seine Bereitschaft, auch für anstrengende Prozesse in der Therapie offen zu sein. Mit drei Items stellt diese die kürzeste Skala dar.

### 6.2.3. Vergleich mit den A-priori-Skalen

Im Verlauf der Entwicklung des Fragebogens sind nicht nur hinsichtlich der Itemanzahl, sondern auch hinsichtlich der Anzahl und Bezeichnungen der Skalen Veränderungen notwendig geworden. In Tabelle 36 sind die drei Entwicklungsstufen dargestellt:

Tabelle 36:  
Bezeichnungen der Fragebogenskalen

Theoretisch hergeleitete A-priori-Skalen	<ul style="list-style-type: none"> <li>• Veränderungserwartung</li> <li>• Offenheit des Patienten</li> <li>• Bereitschaft des Patienten zur Mitarbeit</li> <li>• Therapeutische Beziehung: a) Zusammenarbeit b) emotionale Bindung</li> <li>• Empathie, Echtheit und Akzeptanz von Seiten des Therapeuten</li> <li>• Kompetenzwirkung des Therapeuten</li> <li>• Glaubwürdigkeit und Nachvollziehbarkeit der Therapie</li> </ul>
Skalenbezeichnung der 1. Fragebogenversion	<ul style="list-style-type: none"> <li>• Veränderungserwartung</li> <li>• Bereitschaft des Patienten, sich einzulassen und mitzuarbeiten</li> <li>• Therapeutische Beziehung</li> <li>• Kompetenzwirkung des Therapeuten</li> <li>• Glaubwürdigkeit und Nachvollziehbarkeit der Therapie</li> </ul>
Skalenbezeichnung der endgültigen Fragebogenversion	<ul style="list-style-type: none"> <li>• Veränderungserwartung</li> <li>• Bereitschaft des Patienten sich einzulassen</li> <li>• Therapeutische Beziehung</li> <li>• Passung zwischen Therapeut und Patient</li> <li>• Fachkompetenz des Therapeuten</li> <li>• Überzeugungskraft des Therapeuten</li> </ul>

Über den Prozess der theoretischen und faktorenanalytischen Fragebogenentwicklung und Itemanalyse hinweg gab es bei einer Skala keinerlei Veränderungen – sieht man von der beabsichtigten Reduktion der Itemanzahl ab, und zwar bei der Skala „Veränderungserwartung“. Die beiden ursprünglich getrennten Patientenvariablen „Offenheit des Patienten“ und „Bereitschaft zur Mit-

arbeit“ wurden zunächst zusammengefasst. In den Faktorenanalysen zeigten sich wiederum getrennte Strukturen, wobei nur die Skala „Bereitschaft des Patienten sich einzulassen“, bestehend aus „Ofenheits-Items“, eine ausreichende Stabilität aufwies<sup>49</sup>. Die ursprünglich nach Therapeuten- und zwei Interaktionsaspekten („goal and task“; „bond“) der therapeutischen Beziehung getrennten Skalen „Empathie, Echtheit und Akzeptanz von Seiten des Therapeuten“, „Zusammenarbeit“ und „emotionale Bindung zwischen Therapeut und Patient“ waren infolge der Expertenbefragung zusammengefasst worden zur Skala „therapeutische Beziehung“, was sich in den Faktorenanalysen als stabile Lösung erwies. Bei der faktorenanalytischen Überprüfung wurde ein zusätzlicher Faktor ermittelt, die Skala „Passung zwischen Therapeut und Patient“<sup>50</sup>.

Die Skalen „Fachkompetenz des Therapeuten“ und „Überzeugungskraft des Therapeuten“ sind beide aus der a-priori-Skala „Kompetenzwirkung des Therapeuten“ hervorgegangen und repräsentieren offenbar zwei Unteraspekte von Therapeutenkompetenz. Die a-priori Skala „Glaubwürdigkeit und Nachvollziehbarkeit der Therapie“ ließ sich hingegen faktorenanalytisch nicht abbilden. Die Items dieser Ausgangsskala laden auf demselben Faktor wie die Items der Skala „Überzeugungskraft des Therapeuten“, wurden aber infolge der angewendeten Selektionskriterien eliminiert<sup>51</sup>.

### 6.3. Bezeichnung des Fragebogens

Von Beginn an war der beschriebene Fragebogen als ein Instrument zur Erfassung allgemeiner bzw. unspezifischer Wirkfaktoren in der Psychotherapie konzipiert worden. Daher liegt die Überlegung nahe, ihn ebenso zu benennen (z.B. „Fragebogen zur Erfassung allgemeiner Wirkfaktoren“) und entsprechend abzukürzen. Dagegen spricht allerdings die Tatsache, dass der Begriff „allgemeine Wirkfaktoren“ – zumindest im deutschen Sprachraum – sehr eng mit dem Konzept der allgemeinen Psychotherapie von Klaus Grawe (z.B. 2000) verbunden ist. Da der vorliegende Fragebogen aufgrund seiner Bestimmung als Kontrollinstrument in der Psychotherapieforschung in mancher Hinsicht von dem Grawe'schen Konzept der allgemeinen Wirkfaktoren abweicht, erscheint auch eine abweichende Bezeichnung angemessen. Entsprechend der Annahme des additiven Zusammenwirkens von allgemeinen und spezifischen Faktoren, derzufolge die allgemeinen

---

49 Die Eliminierung des Faktors „Mitarbeit des Patienten“ soll in der Diskussion noch einmal kritisch beleuchtet werden.

50 Die methodischen Schwierigkeiten mit diesem Faktor wurden andernorts bereits dargelegt und werden in der Diskussion erneut diskutiert.

51 Zur inhaltlichen Bedeutung: siehe Diskussion.



## 6. Darstellung der revidierten Form des Fragebogens

---

Faktoren gewissermaßen die Basis für die Wirkung der spezifischen Elemente darstellen, wird folgende Bezeichnung gewählt: „Fragebogen zur Erfassung relevanter Therapiebedingungen“, entsprechend abgekürzt mit „FERT“.

Eine Darstellung der überarbeiteten Fragebogenversion befindet sich in Anhang B.

## IV. Diskussion und Ausblick

### 1. Diskussion der Vorgehensweise

#### 1.1. Konzeption

Zwei Annahmen waren für die Entwicklung des „Fragebogens zur Erfassung relevanter Therapiebedingungen“ (FERT) grundlegend: erstens die Annahme, dass sowohl spezifische als auch unspezifische bzw. allgemeine Faktoren am Erfolg von Psychotherapie beteiligt sind, und zweitens die Annahme, dass sich ihre Wirkungsanteile additiv zusammensetzen bzw. sich klar voneinander unterscheiden lassen. Dass es allgemeine, schulenunspezifische Wirkfaktoren gibt, dürfte heute im Angesicht der vorhandenen Forschungsergebnisse kaum in Frage gestellt werden (vgl. z.B. Orlinsky et al., 1994). Streitbarer dürfte hingegen die Annahme der additiven Wirkungsweise sein: hier könnte eingewendet werden, dass es sich bei dem Zusammenwirken allgemeiner und spezifischer Faktoren um einen interaktiven Prozess handelt, dessen Anteile kaum voneinander zu trennen sind. Diesem Einwand lässt sich entgegen halten, dass einige Studien, die allgemeine und spezifische Therapiegruppen miteinander verglichen haben, Ergebnisse vorgelegt haben, welche die „additive These“ unterstützen (Lambert, 2004; Stevens et al., 2000; Bowers & Clum, 1988; Barker et al., 1988).

Der kritische Leser könnte an dieser Stelle allerdings anmerken, dass die Ergebnisse der Studien keineswegs einheitlich sind und sich hinsichtlich der Anteile von spezifischen und allgemeinen Faktoren am Therapieerfolg deutlich unterscheiden. Und dabei scheinen – wie es die Studie von Baskin et al. (2003) nahe legt – die Differenzen zwischen den Therapiegruppen mit zunehmender Qualität der unspezifischen Kontrollgruppe geringer zu werden, was eine allgemeingültige Aussage hinsichtlich der additiven Wirkung der Faktoren erschwert. Hier lässt sich erwidern, dass das Problem der nicht-äquivalenten Kontrollgruppen gerade ein Argument für die Entwicklung eines Fragebogens darstellt, der allgemeine Faktoren zu erfassen in der Lage ist. Denn wie sollte sonst die Äquivalenz zwischen Therapie- und Kontrollgruppe hinsichtlich der allgemeinen Bedingungen überprüft werden wenn nicht mit einem solchen Instrument? Damit ist dann auch die Möglichkeit gegeben, die Anteile der unspezifischen Faktoren näher zu untersuchen.

In Hinblick auf die Konzeption stellt die Unterscheidung zwischen allgemeinen und unspezifischen Faktoren einen zentralen Aspekt da. Wie andernorts bereits angesprochen, werden einige Variablen in manchen Theorien als spezifisch erachtet und in anderen nicht<sup>52</sup>. Dies erschwert die Zuordnung der in Frage kommenden Variablen zu den allgemeinen vs. den spezifischen Faktoren. Hier wurde im Kontext der Fragebogenkonzeption ein pragmatischer Ansatz gewählt: wenn ein Faktor in allen Therapierichtungen eine Rolle spielt, stellt er kein Entscheidungskriterium für Unterschiede und somit keinen spezifischen Faktor dar, unabhängig von der Theorie. Oder wie es Wampold (2005) formulierte: „What we call common factors can be thought of as a set of aspects that are incidental to the major therapies but that are common to all (or most) therapies“ (S. 837).

Die Entscheidung darüber, welche Faktoren mit dem Fragebogen erfasst werden sollten (a-priori-Faktoren), wurde aufgrund einer ausführlichen Literaturrecherche, die sich sowohl auf theoretische Abhandlungen, als auch auf empirische Studien bezog, getroffen. Es bleibt aber einzuwenden, dass es aufgrund des Fehlens einer allgemeingültigen Theorie bzw. Modellvorstellung bezüglich der allgemeinen Wirkfaktoren nicht eindeutig feststellbar ist, ob alle relevanten Faktoren bei der Entwicklung des Fragebogens berücksichtigt wurden. Zu bedenken ist dabei auch die Wahl der Patientenperspektive, durch welche die Erfassung bestimmter Faktoren kaum möglich ist: der Faktor „adherence“ beispielsweise, also die „Schulen- bzw. Manualtreue“ des Therapeuten (vgl. Wampold, 2001) – sofern man diesen Faktor zu den unspezifischen Variablen hinzurechnen mag –, wurde bei der Konzeption des Fragebogens nicht berücksichtigt. Die Einschätzung dieser Variable wäre nur im Rahmen von Verhaltensbeobachtungen durch unabhängige Rater möglich gewesen - eine Methode, die allerdings mit einem hohen Zeit- und Materialaufwand verbunden ist.

Einen unspezifischen Faktor, der ursprünglich nicht explizit, d.h. in Form einer eigenen Skala berücksichtigt worden war, stellt die Variable „allegiance“ dar, also die Überzeugung des Therapeuten, dass die von ihm durchgeführte Therapie erfolgreich sein wird bzw. kann (vgl. Wampold, 2001). Bei der Konzeption des Instrumentes wurden allerdings Items zur Erfassung der Therapeutenüberzeugung – entsprechend natürlich aus der Patientenperspektive – generiert, die Eingang in die a-priori-Skala „Kompetenzwirkung des Therapeuten“ fanden (z.B. „In der letzten

---

<sup>52</sup> Dies gilt z.B. für den Faktor „Empathie“ im Kontext der Gesprächstherapie im Vergleich zu anderen Therapierichtungen.

Therapiestunde ... wirkte mein Therapeut überzeugt von seinem Vorgehen.“), womit dem Aspekt der „allegiance“ auch in der Konzeption Rechnung getragen wurde.

Zu den bedeutsamen unspezifischen Faktoren gehört z.B. auch die Variable Therapiedauer (vgl. Orlinsky et al., 1994), die ebenso wie andere Faktoren nicht bei der Konzeption des Fragebogens berücksichtigt wurde. Dabei handelt es sich aber um Faktoren, die mittels des Studiendesigns kontrolliert werden können und daher keiner separaten Überprüfung durch einen Patientenfragebogen bedürfen.

Da es das Ziel des Fragebogens ist, Variablen zu erfassen, die einen Einfluss auf das Therapieergebnis haben, erschien es von besonderer Bedeutung, auf die Trennung zwischen allgemeinen (Einfluss-)Faktoren und Ergebnisvariablen zu achten, um unabhängige und abhängige Variablen nicht zu vermengen. Da es bei sich bei einer Therapie notwendigerweise um einen über eine gewissen Zeitdauer und über mehrere abgegrenzte Zeiteinheiten verlaufenden Prozess handelt, könnte kritisch angemerkt werden, dass es sich z.B. bei der Variable „Therapieerwartung“ auch um ein frühes Therapieergebnis handeln könnte, in dem Sinne, dass sich der anfängliche Zustand der „Demoralisierung“ des Patienten ändert. Dagegen lässt sich ins Feld führen, dass der Aspekt der Erwartung, der von den allgemeinen Faktoren die größte Nähe zum ursprünglichen Placebo-Konzept aufweist, der eigentlichen Veränderung zeitlich vorausgeht. Im Sinne des Modells von Schulte (1996) kann die Erwartung, ebenso wie die anderen allgemeinen Faktoren, als notwendige Grundbedingung für den Erfolg (spezifischer) therapeutischer Interventionen gesehen werden, die somit eine Voraussetzung für Therapieerfolg und kein Therapieergebnis darstellt.

### **1.2. Expertenbefragung**

Die Befragung der Experten stellte eine Überprüfung der Eignung der generierten Items zur Sicherung der Inhaltsvalidität dar. Die dabei aufgetretenen Zuordnungsschwierigkeiten hinsichtlich der Ausgangsskalen „Empathie, Akzeptanz und Echtheit des Therapeuten“, „Offenheit des Patienten“ und „Emotionale Bindung zwischen Therapeut und Patient“ wurden andernorts bereits ausführlich dargestellt und diskutiert, ebenso wie die sich daraus ergebenden Veränderungen hinsichtlich der Anzahl und Bezeichnung der Skalen.

Als Problem lässt sich hierzu sicherlich anführen, dass es anhand der Ergebnisse schwer nachzuziehen ist, inwieweit die Zuordnungsschwierigkeiten möglicherweise auch auf die fehlenden Skalenbeschreibungen zurückgehen. Hinweise hierfür gab es zumindest bei der Skala „Offenheit

des Patienten“, die offenbar zum Teil als „Ehrlichkeit“ missverstanden wurde (d.h. der Patient gibt offen zu, was ihm missfällt). Eine kurze Erläuterung zu den Skalenbezeichnungen wäre sicherlich hilfreich im Sinne einer klareren Zuordnung gewesen bzw. hätte zumindest Missverständnisse bezüglich der Skalenbedeutung als Ursache für Zuordnungsschwierigkeiten unwahrscheinlich gemacht. Ergänzend wäre es zudem sinnvoll gewesen, neben der schlichten Zuordnung von den Experten die Formulierungen „qualitativ“, d.h. hinsichtlich Verständlichkeit und Eindeutigkeit, überprüfen zu lassen – das hätte die Entscheidung bzgl. Neuformulierung/Neuauswahl von Items vs. Neudefinition der Skalen deutlich erleichtert.

### **1.3. Fragebogenuntersuchung**

#### **1.3.1. Durchführung**

Um Daten für die Überprüfung der Fragebogenstruktur und der Itemanalyse zu gewinnen, wurde die erste Version des Fragebogens an eine Stichprobe von Patienten ausgegeben, die sich zum Zeitpunkt der Datenerhebung in ambulanter Psychotherapie befanden. Die Patienten sollten mind. 18 Jahre alt sein und sich in Einzelpsychotherapie befinden; außer diesen Einschränkungen bezüglich des Alters und der Settingbedingung wurden keine Einschlusskriterien formuliert, so dass es sich um eine wenig selektierte, „natürliche“ Patientenstichprobe handelte, die sich unter „Realbedingungen“ in Therapie befand. Um dem Ansatz Rechnung zu tragen, dass der zu entwickelnde Fragebogen allgemeine, d.h. schulenunabhängige Faktoren erfassen sollte, wurde angestrebt, Therapeuten unterschiedlicher Therapierichtungen (Richtlinienverfahren) für die Untersuchung zu gewinnen. Das Ziel bestand darin, jeweils eine in etwa gleich große Anzahl an Probanden, die sich in psychoanalytischer, tiefenpsychologischer und verhaltenstherapeutischer Behandlung befanden, in die Studie mit einbeziehen zu können.

Es zeigte sich allerdings bei den Therapeuten eine sehr unterschiedlich ausgeprägte Bereitschaft zur Teilnahme an der Studie in Abhängigkeit von der Schulenzugehörigkeit. Dies führte zu einer deutlichen Überrepräsentativität der Verhaltenstherapie in der untersuchten Stichprobe.

Ein Grund für die zögerliche Beteiligung v.a. der psychoanalytisch arbeitenden Therapeuten mag in der Besorgnis gelegen haben, dass die Durchführung einer Fragebogenerhebung den Therapieverlauf bzw. die Übertragungsprozesse negativ beeinflussen könnte, was folgendes Zitat aus

einer E-mail eines psychoanalytisch arbeitenden Therapeuten verdeutlicht, der seine Kollegen einer Psychoanalytischen Arbeitsgemeinschaft um Mitarbeit an der Studie gebeten hatte:

*„Bei den Kollegen kamen grundsätzliche Bedenken über die Beforschung unserer Behandlungen mit Methoden, die der Psychoanalyse nicht angemessen sind, auf. (...) Unsere Patienten sind nicht auf das Ausfüllen von Fragebögen eingestellt, eine solche Einführung würde die Behandlungen verändern (...)“*

Auch andere Gründe mögen dazu geführt haben, dass einige der angeschriebenen und angesprochenen Kollegen nicht zur Teilnahme an der Studie bereit waren, wobei über diese nur spekuliert werden kann. Es ist allerdings nicht von der Hand zu weisen, dass durch diese „Selbstselektion“ der Therapeuten eine gewisse Stichprobenverzerrung stattfand.

Angesichts der Tatsache, dass der Fragebogen nur an diejenigen Therapeuten verschickt bzw. ausgegeben wurde, die sich vorab zur Teilnahme bereit erklärt hatten, erscheint die Rücklaufquote von 33,7% relativ gering. Dieses Ergebnis ist besonders auf den geringen Rücklauf der Fragebögen aus den Ambulanzen der Christoph-Dornier-Stiftung zurückzuführen (5,2%). Möglicherweise war dies durch die Tatsache bedingt, dass die Therapien in den CDS-Instituten aufgrund des Behandlungskonzepts unregelmäßiger stattfanden als in anderen Ambulanzen, so dass sich eine Vereinbarung über die Rückgabe des Fragebogens als schwieriger erwies. Ein anderer Grund mag darin gelegen haben, dass die anderen Therapeuten direkt angesprochen bzw. angeschrieben worden waren, während in den CDS-Ambulanzen die Akquise über die jeweiligen Institutsteiler erfolgte. Dies mag einen Einfluss auf die Motivation bzw. den empfundenen „Verpflichtungsgrad“ der beteiligten Therapeuten gehabt haben.

Als einen anderen kritischen Aspekt hinsichtlich der Repräsentativität der Stichprobe kann man die mögliche Selektion der teilnehmenden Patienten von Seiten der Therapeuten durch die Ausgabe der Fragebögen anführen. Nicht ganz auszuschließen ist, dass manche Therapeuten – in Anbetracht der Tatsache, dass in den Fragebögen Aussagen über die Therapie und z.B. über die therapeutische Beziehung getroffen werden – nicht allen Patienten den Fragebogen ausgeteilt haben. Und als letzter Punkt ist schließlich zu diesem Thema anzumerken, dass sich die Patienten natürlich selbst entschieden, ob sie an der Studie teilnehmen wollten oder nicht, was wiederum einen Selektionsprozess beinhaltete.

Als Verbesserungsidee könnte man z.B. den Anreiz für die Patienten zur Teilnahme erhöhen durch die Auszahlung eines „Versuchspersonengeldes“ o.ä., wodurch ein vom Inhalt der Befragung, also der Therapie, unabhängiger Motivationsreiz gesetzt werden könnte. Zudem könnte man den Rücklauf der Patienten überprüfen über eine Angabe der Therapeuten, wie viele Fragebögen sie ausgegeben und wie viele sie ausgefüllt zurückbekommen haben, um die Rücklaufquote berechnen zu können. Die Frage, wie man eine größere Zahl nicht-verhaltenstherapeutisch-orientierter Therapeuten zur Teilnahme motivieren könnte, wäre zu überlegen. Möglicherweise ist hier generell ein engerer Austausch und ein Annäherungsprozess nötig, um die Gründe für die Zurückhaltung besser verstehen und entsprechend darauf reagieren zu können.

Hinsichtlich der Durchführungsobjektivität sind – trotz standardisierter Instruktionen für die Therapeuten und Patienten – zwei Aspekte kritisch anzumerken: erstens könnte die Art und Weise, wie die beteiligten Therapeuten die Patienten um die Mitarbeit gebeten haben, die „Voreinstellung“ der Patienten hinsichtlich des Fragebogens und damit schließlich auch das Antwortverhalten beeinflusst haben. Im Rahmen dieser Studie war, wie andernorts bereits diskutiert, kein anderes Vorgehen möglich. Von Vorteil wäre aber prinzipiell gewesen, wenn die Informationsvermittlung über die Fragebogenuntersuchung unabhängig vom Therapeuten stattgefunden hätte. Und schließlich bleibt noch darauf hinzuweisen, dass nicht kontrollierbar war, wann und unter welchen Bedingungen die Patienten den Fragebogen ausgefüllt haben, was prinzipiell ebenfalls einen Einfluss gehabt haben könnte auf das Antwortverhalten.

### **1.3.2. Auswertung**

Zwei Aspekte der Vorgehensweise sind in Bezug auf die Auswertung der Fragebogendaten hervorzuheben. Zum einen ist dies die Verknüpfung der Itemanalyse, mit dem Ziel der Entdeckung und Eliminierung ungeeigneter Items, mit der strukturellen Überprüfung des Konstruktes im Rahmen der Faktorenanalysen. Zum anderen ist dies die strenge Berücksichtigung des Skalenniveaus und der Verteilungscharakteristika der Items durch die kategoriale Auswertung. Den großen Vorteil dabei stellt die parallele Berücksichtigung von Itemkennwerten und gesamtstrukturellen Merkmalen des Konstrukts bei der Itemanalyse dar, ohne dass ungerechtfertigte Annahmen hinsichtlich des Skalenniveaus getroffen werden.

Als kritischen Punkt kann man an dieser Stelle allerdings das Problem der Stichprobengröße anführen. Generell erfordert die Überprüfung von Modellen im Rahmen einer konfirmatorischen Faktorenanalyse (CFA) eine größere Stichprobe, wenn es sich um kategoriale Indikatoren handelt, da in diesem Fall eine polychorische Korrelationsmatrix verwendet wird (Brown, 2006), weshalb die Stichprobe von  $N=216$  in Anbetracht der Modellkomplexität (80 Indikatoren zu Beginn der faktorenanalytischen Auswertung) als zu klein angesehen werden kann. Allerdings lässt sich entgegen halten, dass Flora und Curran (2004) zeigen konnten, dass der WLSMV-Schätzer – im Gegensatz zu anderen Schätzern – auch bei kleinen Stichproben, verschiedenen Stufen an Abweichungen von der Normalverteilung und unterschiedlichen Komplexitätsgraden der Modelle angemessene Teststatistiken liefert, so dass bei Verwendung dieses Schätzers keine allzu gravierenden Abweichungen aufgetreten sein dürften.

Zur Vorgehensweise bei der Itemselektion lässt sich anmerken, dass es sich primär um eine Selektion anhand statistischer Werte handelte. Dies ist per se noch nicht ungewöhnlich, da auch die klassische Itemselektion anhand von Itemkennwerten in diesem Sinne „statistisch“ ist. Als eher ungewöhnlich erscheint möglicherweise, dass bei der dargestellten Vorgehensweise auch die Skalen bzw. Faktoren des Fragebogens auf dem Prüfstand waren. Dies führte z.B. zur Eliminierung des Faktors „Bereitschaft zur Mitarbeit“, da sich dieser Faktor nach der Durchführung der ersten faktorenanalytischen Auswertungsschritte als „empirisch unteridentifiziert“ erwiesen hatte. Man könnte nun vielleicht kritisch anmerken, dass diese Selektion anhand statistischer Kennwerte zu einem Verlust eines inhaltlich höchst relevanten Aspektes geführt haben könnte und die Kennwerte eventuell nur auf eine mangelnde Qualität der Itemformulierungen zurückzuführen seien. Entgegenen könnte man hier, dass sich in den statistischen Werten vermutlich ein inhaltlicher Effekt widerspiegelt. Möglicherweise handelt es sich bei dem Aspekt der „aktiven Mitarbeit“ des Patienten – in der Differenzierung zur Selbstöffnung – doch eher um ein verhaltenstherapeutisch geprägtes Konzept (Beispiel: „Hausaufgaben“), womit dieses eher einen spezifischen, denn einen allgemeinen Wirkfaktor darstellen würde, der sich entsprechend nicht in allen Patientensubgruppen gleichermaßen abbildet. Aus dieser Perspektive erscheint die Eliminierung dieses Faktors auch inhaltlich plausibel und sinnvoll.

Ebenfalls kurz noch einmal angeführt werden soll die – andernorts ausführlich diskutierte – Problematik des Methodenfaktors, also die Frage, ob durch die Gemeinsamkeit der negativen For-



mulierungen ein künstlicher, inhaltlich trivialer Faktor geschaffen wurde. Diese Überlegungen betreffen den im Rahmen der Faktorenanalysen neu gewonnenen Faktor „Passung zwischen Therapeut und Patient“. Nach Abwägen der unterschiedlichen Argumente pro und contra Methodenartefakt wurde die Entscheidung getroffen, diesem Faktor inhaltliche Relevanz zuzugestehen und in Bezug auf das CFA-Modell Residuenkorrelationen zwischen den negativ formulierten Items zuzulassen. Dieser Schritt kann sicherlich kritisch bewertet werden. Aber an dieser Stelle sind die inhaltliche Plausibilität und Bedeutsamkeit dieser Lösung mehr ins Gewicht gefallen als der Nachteil des statistisch nicht optimalen Zulassens von Residuenkorrelationen.

Ein letzter, sicherlich aber nicht unbedeutender Punkt ist an dieser Stelle noch anzuführen. Im Rahmen der konfirmatorischen Faktorenanalyse wurde durch die Eliminierung von Items eine Modell-Respezifizierung durchgeführt, was bedeutet, dass das ursprünglich getestete Ausgangsmodell verändert wurde – im Rahmen der Itemauswahl. Damit handelte es sich aber streng genommen nun nicht mehr um ein konfirmatorisches, sondern vielmehr um ein exploratives bzw. modellgenerierendes Vorgehen. Um sichergehen zu können, dass die postulierte Fragebogenstruktur nicht nur auf stichprobenbedingten Zufalls-Charakteristika zurückgeht, ist unbedingt eine Replizierung an einer unabhängigen Studie vonnöten. Im Rahmen dieser Studie könnte dann eine im eigentlichen Sinne konfirmatorische, d.h. das Konstrukt bestätigende, Faktorenanalyse durchgeführt werden.

## **2. Diskussion der Ergebnisse und Ausblick**

Der Fragebogen zur Erfassung allgemeiner Therapiebedingungen (FERT) besteht in seiner revidierten Form aus sechs Skalen mit 31 Items. Die Modevaluation im Rahmen der konfirmatorischen Faktorenanalyse erbrachte gute Werte der Fit-Indices, was als Beleg für die Modellgüte angesehen werden kann. Die Höhe der Faktorladungen ist zufriedenstellend, ebenso wie die Höhe der Kennwerte für die Faktorreliabilitäten. Insgesamt sprechen die Ergebnisse der Untersuchung für die psychometrische Qualität des FERT, der geeignet erscheint, in ökonomischer Weise die Vergleichbarkeit zwischen Therapiegruppen hinsichtlich der Ausprägung unspezifischer Wirkfaktoren zu überprüfen und somit einen Beitrag zur Verbesserung der Psychotherapieforschung zu liefern.

Dennoch sollten ein paar kritische Anmerkungen nicht fehlen. Zu erwähnen sind hier die links-schiefen Verteilungen der Items, sowie deren hohe Mittelwerte bzw. Mediane. Auch nach der Eliminierung von Items aufgrund der Schwierigkeit handelt es sich bei den verbleibenden Indikatoren nahezu durchgängig um leichte bis sehr leichte Items. Zu diskutieren sind hier als mögliche Erklärungen einerseits Aspekte, die mit der Fragebogenkonstruktion zusammenhängen (Itemformulierungen; unpassende Antwortstufen) und andererseits inhaltliche bzw. stichprobenbedingte Aspekte. Konkret stellt sich die Frage: Spiegelt sich in den positiven Urteilen der Patienten ihre tatsächliche Einschätzung wider oder wird aufgrund einer ungeeigneten Konstruktion ein Deckeneffekt provoziert? Einiges spricht dafür, dass es sich dabei um einen „realen Effekt“ handelt, d.h. dass die Urteile der Patienten tatsächlich positiv ausfallen. Es liegt schließlich in der Intention des Therapeuten, positive Therapiebedingungen herzustellen, die wiederum den Grundstock für die Wirksamkeit spezifischer Interventionen darstellen. Insofern wäre eine Normalverteilung der Items kein gutes Zeugnis für die therapeutischen Bemühungen bzw. Bedingungen<sup>53</sup>. Hinsichtlich des Antwortformates ist es dennoch von zentraler Bedeutung, negative Urteile zuzulassen, da diese natürlich vorkommen – wenn auch nicht so häufig wie positive, weshalb auf die Möglichkeit negativer Urteile nicht verzichtet werden sollte. Für die inhaltlich bedingte Verteilungsschiefe spricht weiterhin, dass sich dieser Effekt auch in anderen Untersuchungen zu Fragebögen mit „verwandtem Inhaltsbereich“ wiederfindet. Zu nennen sind hier beispielsweise Studien zu bzw. mit verschiedenen Allianz-Skalen (z.B. Gaston, 1991; Agnew-Davies, 1998; Hilsenroth et al., 2004), die Ergebnisse der Überprüfung der psychometrischen Eigenschaften des „Bonner Fragebogen für Therapie und Beratung“ (BFTB; Fuchs, 1998) oder die Untersuchung zum „Patientenfragebogen zur Therapieerwartung und Therapieevaluation“ (PATHEV, Schulte, 2005).

Ebenfalls kritisch anzumerken sind die relativ hohen Interkorrelationen zwischen den Skalen bzw. Faktoren des FERT. Diese liegen zwar unterhalb des festgelegten Cut-Off-Wertes, aber immer noch im mittleren bis hohen Bereich. Auch hier könnte die Frage aufgeworfen werden, ob dies an den ungenügend differenzierenden Itemformulierungen liegt – und somit ein Konstruktionsproblem darstellt – oder, ob dieser Effekt ein Abbild der „realen“ Zusammenhänge wiedergibt. Hier lässt sich eine ähnliche Argumentation führen wie im Fall der Nicht-Normalverteilung der Indikatoren. Für die Annahme, dass es sich bei den gefundenen Korrelationen um ein Abbild

---

<sup>53</sup> Selektionsbedingte Stichprobenverzerrungen, die zu einer Homogenisierung der Urteile und damit zu einem Varianzverlust beigetragen haben könnten, wurden als mögliche Einflussfaktoren andernorts bereits diskutiert.

realer Zusammenhänge handelt, spricht die inhaltliche Plausibilität. Die höchsten Interkorrelationen weisen diejenigen Faktoren auf, die auch inhaltlich die größte Nähe aufweisen, obwohl sie dennoch jeweils unterscheidbare Aspekte erfassen. Ein weiteres Argument für die inhaltliche Bedeutsamkeit der Faktor-Interkorrelationen liefern wiederum Studien zu „verwandten“ Instrumenten (z.B. Marmar et al., 1989; Gaston, 1991; Agnew-Davies, 1998; Fuchs, 1998; Krampen & Wald, 2001), die ähnliche hohe Zusammenhänge gefunden haben.

Bezüglich des Inhalts erscheint es lohnenswert noch einmal die Veränderungen der Skalen über den Untersuchungsprozess hinweg zu betrachten. Die Abwägungen bezüglich der Beibehaltung des neuen Faktors „Passung zwischen Therapeut und Patient“ (pro/contra Methodenfaktor) wurden an anderer Stelle bereits dargestellt. Inhaltlich geht es in dieser Skala um die Erfahrung von Stimmigkeit und Passung, bezogen auf das emotionale Erleben des Patienten. Im Berner Stundenbogen 2000 (Stubo, Grawe & Regli, 2000), der entsprechend dem Graweschen Wirkfaktorenmodell konzipiert wurde, findet sich die Skala „Positive Bindungserfahrungen: Aufgehobensein in der Therapie“, die zumindest der Skalenbezeichnung nach eine gewisse Parallele zur Skala „Passung zwischen Therapeut und Patient“ aufweist. Bei näherer Betrachtung der einzelnen Items werden allerdings inhaltliche Unterschiede deutlich (Beispiele: „Ich finde, der Therapeut müsste meinen Gefühlen mehr Beachtung schenken“; „Ich finde die Sichtweise des Therapeuten von meinen Problemen zu einfach.“ vs. „In der letzten Therapiestunde hatte ich das Gefühl, dass ich mich verteidigen muss.“ im FERT). Auffällig ist aber in jedem Fall, dass alle vier Items dieser Skala negativ formuliert sind, wohingegen sich sonst keine negativ formulierten Items im Berner Stundenbogen finden. Die Skala „Passung zwischen Therapeut und Patient“ des FERT beinhaltet zwar auch ein positiv formuliertes Item, alle anderen aber sind invers gepolt, während sich in den anderen Skalen lediglich positiv formulierte Items finden. Als Fazit dieser Parallele lässt sich folgern, dass die Frage nach dem Einfluss eines Methodeneffektes in nachfolgenden Studien weiter gestellt werden sollte und noch nicht als abschließend geklärt angesehen werden kann.

Verändert hat sich im Entwicklungsprozess auch die a-priori-Skala „Glaubwürdigkeit und Nachvollziehbarkeit der Therapie“. Ursprünglich war diese Skala in Anlehnung an die Überlegungen von Frank (1961) und anderen konzipiert worden, denen zufolge nur eine Therapie wirksam sein könne, die dem Patienten durch die Vermittlung eines Behandlungsrationalis plausibel und glaub-

würdig erscheine. Die Items dieser Ausgangsskala beinhalteten also Aussagen über die Therapieinhalte, wie z.B. „In der letzten Therapiestunde .... erschien mir das Therapieangebot einleuchtend“. Im Verlauf der faktorenanalytischen Skalenüberprüfungen und Item-Eliminierung wurde allerdings deutlich, dass diese Skala so nicht beizubehalten war. Der Fokus bewegte sich weg vom Inhalt der Therapiesitzung hin zum Auftreten bzw. Verhalten des Therapeuten (z.B. „In der letzten Therapiestunde ... wirkte mein Therapeut überzeugt von seinem Vorgehen“), weshalb diese Skala letztendlich auch umbenannt wurde in „Überzeugungskraft des Therapeuten“. Ausgehend von der Annahme, dass es sich dabei nicht um einen zufälligen Stichproben-Effekt handelt, und dass dieses Ergebnis auch nicht auf eine mangelhafte Fragebogenkonstruktion zurückgeht, kann man dieses Ergebnis als Hinweis darauf deuten, dass die Glaubwürdigkeit des Therapeuten einen bedeutsameren allgemeinen Wirkfaktor darstellt als die Glaubwürdigkeit des Behandlungsrationalen. Oder noch etwas pointierter formuliert: »Wenn der Therapeut überzeugt ist, ist dem Patienten die Plausibilität des Rationalen egal«. Eine mögliche Erklärung hierfür wäre, dass der Patient auf das Therapeutenverhalten als Informationsquelle hinsichtlich der Glaubwürdigkeit zurückgreift, da er selbst mit der Einschätzung der inhaltlichen Plausibilität überfordert ist. Dies könnte auch ein Erklärungsansatz für die Ergebnisse von Devilly & Borkovec (2000) sein, die im Rahmen ihrer Studien nur sehr schwache Zusammenhänge zwischen den Glaubwürdigkeitseinschätzungen der Patienten und dem Therapieerfolg finden konnten.

Es scheint in jedem Fall lohnend, diese etwas provokative These in weiteren Untersuchungen zu prüfen.

Auch andere Fragestellungen sind augenblicklich noch nicht zufriedenstellend geklärt. Zum Abschluss erscheint es sinnvoll, diese noch einmal in den Fokus zu setzen und davon ausgehend ein Konzept für weiterführende Untersuchungen zu erstellen.

An erster Stelle sei hier die Notwendigkeit einer Validierungsstudie genannt. Durch das modellgenerierende Vorgehen im Rahmen der Itemselektion wurde in der vorliegenden Untersuchung streng genommen keine konfirmatorische Faktorenanalyse durchgeführt. Daher ist die Überprüfung des Modells an einer zweiten, unabhängigen Stichprobe notwendig, um die faktorielle Validität zu bestätigen. In diesem Zusammenhang sollte zudem durch einen Vergleich mit ähnlichen Instrumenten (z.B. Allianz-Skalen) die konvergente Validität untersucht werden. Die Durchführung einer Validierungsstudie böte zudem die Möglichkeit, die postulierten Zusammenhänge zwischen den allgemeinen Faktoren und dem Therapieerfolg zu überprüfen und in ihrer Größe

abzuschätzen, was gewissermaßen eine Bestimmung der prognostischen Validität darstellen würde.

Auch andere Fragen sind in Bezug auf den Fragebogen zur Erfassung allgemeiner Therapiebedingungen noch offen und verdienen eine nähere Beleuchtung. Dies betrifft u.a. die Frage nach dem Einfluss der einzelnen Faktoren untereinander. Stellt beispielsweise die therapeutische Beziehung eine Moderatorvariable für die Erfolgserwartung dar? Denkbar wäre in diesem Zusammenhang etwa eine Erweiterung des im Rahmen der konfirmatorischen Faktorenanalyse getesteten Modells zu einem Strukturgleichungsmodell, in dem auch Beziehungen der Faktoren untereinander modelliert werden könnten. Weiterhin erscheint es nötig, die postulierte Passung des Fragebogens auch für andere Therapierichtungen zu untersuchen, um somit die Generalisierbarkeit zu bestätigen. Und schließlich wäre es sicher sinnvoll im Rahmen einer Therapiestudie die Veränderung der Variablenausprägung über den Therapieverlauf zu erheben, wie es zumindest für die therapeutische Allianz auch in anderen Studien bereits geschehen ist (z.B. Stiles et al., 2004).

Trotz dieser offenen Punkte lässt sich abschließend folgendes Fazit ziehen:

Auch wenn sich der Fragebogen zur Erfassung allgemeiner Therapiebedingungen (FERT) noch weiteren Überprüfungen stellen muss, so scheint es doch sinnvoll, dieses Instrument auch jetzt schon im Rahmen der Psychotherapieforschung zur Überprüfung der Kontrollgruppenäquivalenz einzusetzen. Denn neben einer Kontrolle der strukturellen Äquivalenz der Gruppen (z.B. in puncto Therapiedauer und Sitzungsanzahl) stellt die Vergleichbarkeit der allgemeinen Faktoren die entscheidende Voraussetzung dafür dar, dass Unterschiede im Therapieergebnis tatsächlich auf die durchgeführten Interventionen zurückgeführt werden können.

### **3. Zusammenfassung**

Das Ziel der hier vorgestellten Untersuchung war die Konstruktion und Überprüfung eines Fragebogens zur Erfassung allgemeiner Wirkfaktoren in der Psychotherapie, der dazu beitragen soll, die Aussagekraft und damit die Qualität von vergleichenden Psychotherapiestudien zu erhöhen. Grundlegend dafür ist die Annahme, dass sich der Erfolg einer Psychotherapie sowohl auf therapieschulen-spezifische, als auch auf unspezifische bzw. allgemeine Faktoren zurückführen lässt und dass sich die Wirkung dieser Faktoren additiv zusammensetzt (z.B. Stevens et al., 2000). Ausgehend von dieser Grundannahme lässt sich folgern, dass Aussagen über die Auswirkung

spezifischer Interventionen auf den Therapieerfolg nur dann getroffen werden können, wenn die Auswirkungen der unspezifischen Faktoren bekannt sind bzw. kontrolliert werden. Da bislang kein adäquates Messinstrument zur Verfügung stand, das zu diesem Zweck hätte eingesetzt werden können, war es das Ziel des dargestellten Projektes, ein neues Instrument zu entwickeln. Ausgehend von Theorien und Modellen zur Wirkungsweise von Psychotherapie wurden, unter Berücksichtigung von empirischen Befunden, die a-priori-Skalen des Fragebogens festgelegt. Im Anschluss daran erfolgte die Generierung von Items, wobei diese in mehreren Schritten sowohl hinsichtlich ihrer Passung, als auch hinsichtlich ihrer Formulierung überarbeitet wurden. Eine weiterführende Überprüfung der inhaltlichen Passung der Items, sowie der Fragebogenskalen erfolgte mithilfe einer Expertenbefragung, in welcher die Teilnehmer eine Zuordnung der Ausgangsitems zu den a-priori-Skalen vornahmen. Aufgrund der Ergebnisse dieser Befragung fand die teilweise Zusammenlegung von a-priori-Skalen auf nunmehr fünf Ausgangsskalen, sowie eine Itemvorauswahl statt.

Mit einem Itempool von 88 Items wurde die erste Version des Fragebogens zur Erfassung allgemeiner Wirkfaktoren erstellt. Zur Gewinnung von Daten für eine fundierte Itemanalyse und -selektion wurde eine Fragebogenuntersuchung zu einem Messzeitpunkt durchgeführt, an der 220 Patienten teilnahmen, die sich zum Zeitpunkt der Untersuchung in ambulanter Einzeltherapie bei 36 Therapeuten unterschiedlicher Schulenzugehörigkeit befanden. Zur Überprüfung der Fragebogenstruktur und zur Selektion der Items wurde ein mehrstufiges faktorenanalytisches Vorgehen gewählt. Zunächst wurde eine zweistufige explorative Faktorenanalyse (EFA) und im Anschluss eine konfirmatorische Faktorenanalyse (CFA) mit Modellvergleich durchgeführt, in deren Rahmen auch die schrittweise Itemselektion erfolgte. Das abschließende CFA-Modell zeigte einen guten Modellfit und bedeutsame Faktorladungen der Indikatoren, mit hohen, aber akzeptablen Faktorinterkorrelationen.

Der endgültige Fragebogen, mit der Bezeichnung „Fragebogen zur Erfassung relevanter Therapiebedingungen“ (FERT), besteht aus insgesamt 31 Items und sechs Skalen, die jeweils eine gute Faktorreliabilität aufweisen.

## V. Literaturangaben

- Agnew-Davies, R., Stiles, W.B., Hardy, G.E., Barkham, M. & Shapiro, D.A. (1998). Alliance structure assessed by the Agnew Relationship Measure (ARM). *British Journal of Clinical Psychology*, 37, 155-172.
- Alexander, L.B. & Luborsky, L. (1986). The Penn Helping Alliance Scales. In L.S. Greenberg & W.M. Pinsoff (Eds.). *The psychotherapeutic process: A research Handbook* (pp.325-366). New York: Guilford Press.
- Allison, P.D. (2003). Missing data techniques for structural equation modeling. *Journal of Abnormal Psychology*, 112, 545-557.
- Andrusyna, T.P., Tang, T.Z., DeRubeis, R.J. & Luborsky, L. (2001). The factor structure of the working alliance inventory in cognitive-behavioral therapy. *The Journal of Psychotherapy Practice and Research*, 10, 173-178.
- Barker, S.L., Funk, S.C. & Houston, B.K. (1988). Psychological treatment versus nonspecific factors: A meta-analysis of conditions that engender comparable expectations for improvement. *Clinical Psychology Review*, 8, 579-594.
- Baskin, T.W., Tierney, S.C., Minami, T. & Wampold, B.E. (2003). Establishing specificity in psychotherapy: A meta-analysis of structural equivalence of placebo controls. *Journal of Consulting and Clinical Psychology*, 71 (6), 673-679.
- Bassler, M., Potratz, B. & Krauthauser, H. (1995). Der Helping Alliance Questionnaire (HAQ) von Luborsky. *Psychotherapeut*, 40, 23-32.
- Beauducel, A. & Wittmann, W.W. (2005). Simulation study on fit indices in confirmatory factor analysis based on data with slightly distorted simple structure. *Structural Equation Modeling*, 12, 41-75.
- Bentler, P.M. (1990). Comparative fit indices in structural models. *Psychological Bulletin*, 107, 238-246.
- Bleyen, K., Vertommen, H., Vander Steene, G. & Van Audenhove C. (2001). Psychometric properties of the psychotherapy expectancy inventory -revised (PEI-R). *Psychotherapy Research*, 11 (1), 69-83.
- Bordin, E.S. (1979). The generalizability of the psychoanalytic concept of the working alliance. *Psychotherapy: Theory, research and practice*, 16, 252-260.
- Bordin, E.S. (1994). Theory and research on the therapeutic working alliance: New directions. In: A.O. Horvath & L.S. Greenberg (Eds.): *The working alliance: Theory, research and practice* (pp. 13-37). New York: Wiley.

- Borkovec, T.D. & Costello, E. (1993). Efficacy of applied relaxation and cognitive-behavioral therapy in the treatment of generalized anxiety disorder. *Journal of Consulting and Clinical Psychology*, 61 (4), 611-619.
- Borkovec, T.D. & Sibrava, N.J. (2005). Problems with the use of placebo conditions in psychotherapy research, suggested alternatives, and some strategies for the pursuit of the placebo phenomenon. *Journal of clinical psychology*, 61 (7), 805-818.
- Bortz, J. (2005). *Statistik für Human- und Sozialwissenschaftler* (6. Aufl.). Heidelberg: Springer.
- Bortz, J. & Döring, N. (2002). *Forschungsmethoden und Evaluation: Für Human- und Sozialwissenschaftler* (3.Aufl.). Berlin: Springer.
- Bowers, T.G. & Clum, G.A. (1988). Relative contribution of specific and nonspecific treatment effects: meta-analysis of placebo-controlled behavior therapy research. *Psychological Bulletin*, 103 (3), 315-323.
- Brown, T.A. (2006). *Confirmatory factor analysis for applied research (Methodology in the social sciences)*. New York: Guilford Press.
- Bryant, F.B. (2000). Assessing the Validity of Measurement. In L.G. Grimm & P.R. Yarnold (Eds.): *Reading and understanding MORE multivariate statistics*. Washington, DC: American Psychological Association.
- Bühner, M. (2006). *Einführung in die Test- und Fragebogenkonstruktion* (2. Auflage). München: Pearson.
- Burlingame, G., Lambert, M., Reisinger, C., Neff, W., Mosier, J. (1995). Pragmatics of tracking mental health outcome in managed care settings. *Journal of Mental Health Administration*, 22(3), 226-235.
- Byrne, B.M. (1989). *A primer of LISREL: basic applications and programming for confirmatory factor analytic models*. New York: Springer.
- Burns, D.D. & Nolen-Hoeksema, S. (1992). Therapeutic empathy and recovery from depression in cognitive-behavioral therapy: A structural equation model. *Journal of Consulting and Clinical Psychology* 60 (3), 441-449.
- Castonguay, L.G. (1993). „Common factors“ and „nonspecific variables“: Clarification of the two concepts and recommendations for research. *Journal of Psychotherapy Integration*, 3, 267-268.
- Cohen, J., Cohen, P., West, S.G. & Aiken, L.S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: Erlbaum.
- Corbiere, M., Bisson, J., Lauzon, S. & Ricard, N. (2006). Factorial validation of a french short-form of the Working Alliance Inventory. *International Journal of Methods in Psychiatric Research*, 15(1), 36-45.



- Cortina, J. (1993). What is Coefficient alpha? An Examination of theory and application. *Journal of Applied Psychology*, 78 (1),98-104.
- Cronbach, L.J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297-334.
- DeVellis, R.F. (2003). *Scale development: theory and applications* (2nd ed.) Applied social research methods series; v.26. Thousand Oaks: Sage.
- Devilley, G.J. & Borkovec, T.D. (2000). Psychometric properties of the credibility/expectancy questionnaire. *Journal of Behavior Therapy and Experimental Psychiatry*, 31, 73-86.
- Dilling, H., Mombour, W. & Schmidt M.H. (2004). *Internationale Klassifikation psychischer Störungen, ICD-10 Kapitel V (F)., Klinisch-diagnostische Leitlinien* (4. Auflage). Bern: Hans Huber.
- Enck, P., Benedetti, F. & Schedlowski, M. (2008). New insights into the placebo and nocebo responses. *Neuron*, 59, 195-206.
- Fabrigar, L.R., Wegener, D.T., MacCallum, R.C. & Strahan, E.J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 4, 272-299.
- Fisseni, H.J. (1990). *Lehrbuch der psychologischen Diagnostik*. Göttingen: Hogrefe.
- Flora, D.B. & Curran, P.J. (2004). An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological Methods*, 9 (4), 466-491.
- Floyd, F.J. & Widaman, K.F. (1995). Factor analysis in the development and refinement of clinical assessmentinstruments. *Psychological Assessment*, 7, 286-299.
- Frank, J.D. (1961). *Persuasion and healing: A comparative study of psychotherapy*. New York: Schocken Books.
- Frank, J.D. & Frank, L.B. (1991). *Persuasion and healing: A comparative study of psychotherapy* (3rd. ed.). Baltimore: Johns Hopkins University Press.
- Freud, S. (1912). Zur Dynamik der Übertragung. *GW*, Bd VIII. *Imago*, London (1955), 364-374.
- Fuchs, T. (1998). *Entwicklung und Validierung eines katamnestischen Fragebogens: der Bonner Fragebogen für Therapie und Beratung*. Witterschlick/Bonn: Wehle.
- Garfield, S.L. (1982). *Psychotherapie. Ein eklettischer Ansatz*. Weinheim: Beltz.
- Gaston L. (1990). The concept of the alliance and its role in psychotherapy: Theoretical and empirical considerations. *Psychotherapy*, 27 (2), 143-153.

- Gaston, L. (1991). Reliability and criterion-related validity of the California Psychotherapy Alliance Scales-Patient version. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*, 3, 68-74.
- Gaston, L. & Marmar, C.R. (1994). The California psychotherapy alliance scales. In: A.O. Horvath & L.S. Greenberg (Eds.): *The working alliance: Theory, research and practice* (pp. 85-108). New York: Wiley.
- Glöckner-Rist, A. & Hoijtink, H. (2003). The best of both worlds: factor analysis of dichotomous data using Item Response Theory and Structural Equation Modeling. *Structural Equation Modeling*, 10 (4), 544-565.
- Grawe, K., R. Donati & F. Bernauer (1994). *Psychotherapie im Wandel. Von der Konfession zur Profession*. Göttingen: Hogrefe.
- Grawe, K. (2000): *Psychologische Therapie*. Hogrefe, 2. korrig. Auflage.
- Grawe, K. & Regli, D. (2000). *Berner Stundenbogen 2000*. Unveröffentlichtes Instrument.
- Grawe, K. (2005). (Wie) kann Psychotherapie durch empirische Validierung wirksamer werden? *Psychotherapeutenjournal*, 1/2005, 4-11. Psychotherapeutenverlag.
- Guadagnoli, E. & Velicer, W.F. (1988). Relation of sample size to the stability of component patterns. *Psychological Bulletin*, 103 (2), 265-275.
- Hank, P. & Krampen, G. (2006). Diagnostik der therapeutischen Beziehung. In M. Hermer & B. Röhrle (Hrsg.), *Handbuch der therapeutischen Beziehung*. Tübingen: DGVT-Verlag.
- Hartley, D.E. & Strupp, H.H. (1983). The therapeutic alliance: Its relationship to outcome in brief psychotherapy. In J. Masling (Ed.). *Empirica studies of psychoanalytic theories*, Vol. 1; pp. 1-38. Hillsdale, NJ: Analytical Press.
- Harvey, R.J., Billings, R.S. & Nilan, K.J. (1985). Confirmatory factor analysis of the job diagnostic survey: Good news and bad news. *Journal of Applied Psychology*. 70 (3), 461-468.
- Hatcher, R.L. & Barends, A.W. (1996). Patient's view of the alliance in psychotherapy: Exploratory factor analysis of three alliance measures. *Journal of Consulting and Clinical Psychology*, 64 (6), 1326-1336.
- Henry, W. & Strupp, H. (1994). The therapeutic alliance as interpersonal process. In: A.O. Horvath & L.S. Greenberg (Eds.): *The working alliance: Theory, research and practice* (pp. 51-84). New York: Wiley.
- Hentschel, U. (2005). Die therapeutische Allianz. Teil 1: Die Entwicklungsgeschichte des Konzepts und moderne Forschungsansätze. *Psychotherapeut*, 50, 305-317.

- Herbert, J. & Gaudiano B. (2005). Moving from empirically supported treatment lists to practice guidelines in psychotherapy: The role of the placebo concept. *Journal of Clinical Psychology*, 61 (7), 893-908.
- Hildebrandt, L. & Temme, D. (2006). Probleme der Validierung mit Strukturgleichungsmodellen. SFB 649 Discussion Paper 2006-082. Sonderforschungsbereich 649: Economic Risk. ISSN 1860-5664. <http://sfb649.wiwi.hu-berlin.de>.
- Hilsenroth, M.J., Peters, E.J. & Ackerman, S.T. (2004). The development of therapeutic alliance: Patient and therapist perspectives across treatment. *Journal of Personality Assessment*, 83 (3), 332-344.
- Höger, D. (1999). Der Bielefelder Fragebogen zu Klientenerwartungen (BFKE). *Psychotherapeut*, 44, 159-166.
- Horan, P.M., DiStefano C. & Motl, R.W. (2003). Wording effects in self-esteem scales: Methodological artifact or response style? *Structural Equation Modeling*, 10(3), 435-455.
- Horvath, A.O. & Greenberg, L.S. (1986). The development of the Working Alliance Inventory. In: L.S. Greenberg & W.M. Pinsoff (Eds.): *The psychotherapeutic process: A research handbook* (pp. 529-556). New York: Guilford Press.
- Horvath, A.O. & Greenberg, L.S. (1994). Introduction. In: A.O. Horvath & L.S. Greenberg (Eds.): *The working alliance: Theory, research and practice* (pp. 1-12). New York: Wiley.
- Horvath, A.O. & Luborsky, L. (1993). The role of the therapeutic alliance in psychotherapy. *Journal of Consulting and Clinical Psychology*, 61, 561-573.
- Horvath, A.O. & Symonds, B.D. (1991). Relation between working alliance and outcome in psychotherapy: A meta-analysis. *Journal of Counseling Psychology*, 38, 139-149.
- Horvath, A.O. (1996). Research on the Alliance, in A.O. Horvath & L.S. Greenberg (Eds.). In: A.O. Horvath & L.S. Greenberg (Eds.): *The working alliance: Theory, research and practice* (pp. 259-80). New York: Wiley.
- Horvath, A.O. (2005). The therapeutic relationship: Research and theory. An introduction to the Special Issue. *Psychotherapy Research*, 15, 3-7.
- Howard, K., Lueger, R., Maling, M. & Martinovich, Z. (1993). A phase model of psychotherapy outcome: Causal mediation of change. *Journal of Consulting and Clinical Psychology*, 61, 678-685.
- Hu, L. & Bentler, P.M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling*, 6 (1), 1-55.
- Idaszak, J. & Drasgow, F. (1987). A revision of the job diagnostic survey: Elimination of a measurement artefact. *Journal of Applied Psychology*, 72, 69-74.

- Jaccard, J. & Wan, C.K. (1996). LISREL approaches to interaction effects in multiple regression. Thousand Oaks, CA: Sage.
- Jacobson, N. S. (1991). Behavioral versus insight-oriented marital therapy: Labels can be misleading. *Journal of Consulting and Clinical Psychology*, 59, 142-145.
- Jöreskog, K.G. (1969). A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika*, 34, 183-202.
- Jöreskog, K.G. (1993). Testing structural equation models. In K.A. Bollen & J.S. Long (Eds.): *Testing structural equation models* (pp. 294-316). Newbury Park, CA: Sage.
- Jöreskog, K.G. & Sörbom, D.(1979). *Advances in factor analysis and structural equation models.* (J. Magidson, Ed.). Cambridge, MA: Abt Books.
- Joyce, A.S., Ogradniczuk, J.S., Piper, W.E. & McCallum, M. (2003). The alliance as mediator of expectancy effects in short-term individual therapy. *Journal of Consulting and Clinical Psychology*, 71 (4), 672-679.
- Kassenärztliche Bundesvereinigung (2006). *Grunddaten zur vertragsärztlichen Versorgung in Deutschland 2006.*
- Kendall, P.C., Holmbeck, G. & Verduin, T. (2004). Methodology, design, and evaluation in psychotherapy research. In M. J. Lambert (Ed.): *Bergin and Garfield's handbook of psychotherapy and behavior change*, 5th edition (pp.16-43). New York: Wiley.
- Kirsch, I. (2005). Placebo psychotherapy: Synonym or oxymoron? *Journal of Clinical Psychology*, 61 (7), 791-803.
- Kirsch I. & Henry, D. (1977). Extinction versus credibility in the desensitization of speech anxiety. *Journal of Consulting and Clinical Psychology*, 45, 1052-1059.
- Krampen, G. & Wald, B. (2001). Kurzinstrumente für die Prozessevaluation und adaptive Indikation in der Allgemeinen und Differentiellen Psychotherapie und Beratung. *Stundenbogen für die Einzel-Psychotherapie und -Beratung (STEP)*. *Diagnostica*, 47 (1), 43-50.
- Kriz, J. (1991). *Grundkonzepte der Psychotherapie*. Weinheim: Psychologie Verlags Union.
- Krupnick, J.L., Sotsky, S.M., Simmens, S., Moyer, J., Elkin, I., Watkins, J. & Pilkonis, P.A. (1996). The role of the therapeutic alliance in psychotherapy and pharmacotherapy outcome: Findings in the National Institute of mental Health treatment of Depression Collaborative Research Program. *Journal of Consulting and Clinical Psychology*, 64 (3): 532-539.
- Lambert, M.J. (1992). Psychotherapy outcome research: Implications for integrative and eclectic therapists. In M.R. Goldfried, & J.C. Norcross (Eds.): *Handbook of psychotherapy integration* (pp. 94-129). New York: Basic Books.

- Lambert, M.J. (2005). Early response in psychotherapy: Further evidence for the importance of common factors rather than „placebo effects“. *Journal of Clinical Psychology*, 61 (7), 855-869.
- Lambert M.J. & Ogles, B.M. (2004). The efficacy and effectiveness of psychotherapy. In: M.J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (5th ed., pp.139-193). New York: Wiley.
- Leber, P. (2000). The use of placebo control groups in the assessment of psychiatric drugs: An historical context. *Biological Psychiatry*, 47, 699-706.
- Lohr, J.M., DeMaio, C. & McGlynn, F.D. (2003). Specific and nonspecific treatment factors in the experimental analysis of behavioral treatment efficacy. *Behavior Modification*, 27 (3), 322-368.
- Luborsky, L. (1994). Therapeutic alliances as predictors of psychotherapy outcomes: Factors explaining the predictive success. In: A.O. Horvath & L.S. Greenberg (Eds.): *The working alliance: Theory, research and practice* (pp. 38-48). New York: Wiley.
- MacCallum, R.C. (1986). Specification searches in covariance structure modeling. *Psychological Bulletin*, 100, 107-120.
- MacCallum, R.C., Roznowski, M. & Necowitz, L.B. (1992). Model modifications in covariance structure analysis: The problem of capitalization on change. *Psychological bulletin*, 111, 490-504.
- Margraf, J. (2008). *Kosten und Nutzen der Psychotherapie. Eine kritische Literaturlauswertung*. Heidelberg: Springer.
- Marmar, C.R., Weiss, D.S. & Gaston, L. (1989). Toward the validation of the California Therapeutic Alliance Rating System. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*, 1(1), 46-52.
- Marsh, H.W. & Grayson, D. (1995). Latent variable models of multitrait-multimethod data. In: R.H. Hoyle (Ed.), *Structural equation modeling: Concept, issues, and applications* (pp. 177-198). Thousand Oaks, CA: Sage.
- Martin D.J., Garske J.P. & Davis M.K. (2000). Relation of the therapeutic alliance: A meta-analytic review. *Journal of Consulting and Clinical Psychology*. 68 (3), 438-450.
- McDonald, R.P. (1970). Theoretical foundations of principal factor analysis, canonical factor analysis and alpha factor analysis. *British Journal of Mathematical and Statistical Psychology*, 23 (1), 1-21.
- McDonald, R.P. (1981). The dimensionality of tests and items. *British Journal of Mathematical and Statistical Psychology*, 34, 100-117.
- Meyer, B., Pilkonis, P.A., Krupnick, J.L., Egan, M.K., Simmens, S.J. & Sotsky, S.M. (2002). Treatment expectancies, patient alliance, and outcome: Further analyses from the National In-

- stitute of Mental Health treatment of Depression collaborative Research Program. *Journal of Consulting and Clinical Psychology*, 70 (4), 1051-1055.
- Moosbrugger, H. & Kelava, A. (Hrsg.) (2007). *Testtheorie und Fragebogenkonstruktion*. Heidelberg: Springer.
- Mummendey, Hans Dieter (1987). *Die Fragebogen-Methode*. Göttingen: Hogrefe.
- Muthén, B. O. & Asparouhov, T. (2002). Latent variable analysis with categorical outcomes: Multiple-group and growth modeling in Mplus. Unpublished manuscript. (Available at [www.statmodel.com/mplus/examples/webnote.html#web4](http://www.statmodel.com/mplus/examples/webnote.html#web4)).
- Muthén, B. O. & Muthén, L. (1998-2006). *MPlus (Version 4.2) [Software]*: Muthén & Muthén.
- Muthén, L.K. and Muthén, B.O. (1998-2006). *Mplus User's Guide*. Fourth Edition. Los Angeles, CA: Muthén & Muthén.
- Oken, B. (2008). Placebo effects: clinical aspects and neurobiology. *Brain*, 131, 2812-2823.
- Orlinsky, D.E. & Howard, K. I. (1987). A generic model of psychotherapy. *Journal of Integrative and Eclectic Psychotherapy*, 6 (1), 6-27.
- Orlinsky, D.E., Grawe, K. & Parks B.K. (1994). Process and outcome in psychotherapy – Noch einmal. In: A.E. Bergin & S.L. Garfield (Eds.), *Handbook of psychotherapy and behavior change*. New York: Wiley.
- Podsakoff, P.M., MacKenzie, S.B., Lee, J. & Podsakoff, N.P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. 88 (5), 879-903.
- Preacher, K.J. & MacCallum, R.C. (2003). Repairing Tom Swift's electric factor analysis machine. *Understanding Statistics*, 2 (1), 13-43.
- Richter, R. & Harfst, T. (2006). *Anpassung der Psychotherapierichtlinien und evidenzbasierte Psychotherapie*. *Psychotherapeutenjournal* 2/2006. Psychotherapeutenverlag.
- Rost, J. (2004). *Lehrbuch Testtheorie-Testkonstruktion (2. Aufl.)*. Bern: Hans Huber.
- Sachse, R. (2006). *Therapeutische Beziehungsgestaltung*. Göttingen: Hogrefe.
- Salvio, A.-M., Beutler, L.E., Wood, J.M. & Engle, D. (1992). The strength of the therapeutic alliance in three treatments for depression. *Psychotherapy Research*, 2, 31-36.
- Scheidt, C.E., Seidenglanz K., Dieterle, W., Hartmann, A., Bower, N., Hillenbrand, D., Sczudlek, G., Strassée, F. & Wirsching M. (1998). Basisdaten zur Qualitätssicherung in der ambulanten Psychotherapie. *Psychotherapeut*, 43, 92-101.
- Schulte, D. (1996). *Therapieplanung*. Göttingen: Hogrefe.

- Schulte, D. (2005). Messung der Therapieerwartung und Therapieevaluation von Patienten (PATHEV). *Zeitschrift für Klinische Psychologie und Psychotherapie*, 34 (3), 176-187.
- Shevlin, M., Miles, J., Davies, M. & Walkers, S. (2000). Coefficient alpha: A useful indicator of reliability? *Personality and Individual Differences*, 28 (2), 229-237.
- Sörbom, D. (1989). Model modification. *Psychometrika*, 54, 371-384.
- Steiger, J.H. & Lind, J.M. (1980). Statistically based tests for the number of common factors. Paper presented at the meeting of the Psychometric Society, Iowa City, IA.
- Studel, J. (1956). Woher stammt das Wort Placebo? *Deutsche Medizinische Wochenschrift* 81(1956), S. 1482.
- Stevens, S.E., Hynan M.T. & Allen, M. (2000). A meta-analysis of common factor and specific treatment effects across the outcome domains of the phase model of psychotherapy. *Clinical Psychology: Science and Practice*, 7 (3), 2000.
- Stiles, W.B., Glick, M.J., Osatuke, K., Hardy, G.E., Shapiro, D.A., Agnew-Davies, R., Rees, A. & Barkham, M. (2004). Patterns of Alliance Development and the Rupture-Repair Hypothesis: Are Productive Relationships U-Shaped or V-Shaped?. *Journal of Counseling Psychology* 51: 81-92.
- Suh, C.S., Strupp, H.H. & O'Malley, S.S. (1986). The Vanderbilt Process Measures: The Psychotherapy Process Scale (VPPS) and the Negative Indicators Scale (VNIS). In L.S. Greenberg & W.M. Pinsof (Eds.), *The psychotherapeutic process: A research handbook* (pp. 285-324). New York: Guilford.
- Takane, Y. & De Leeuw, J. (1987). On the relationship between item response theory and factor analysis of discretized variables. *Psychometrika*, 52 (3), 393-408.
- Thurstone, L.L. (1947). *Multiple-factor analysis*. Chicago: University of Chicago Press.
- Tracey, T.J. & Kokotovic, A.M. (1989). Factor structure of the Working Alliance Inventory. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*, 1, 207-210.
- Tucker, L.R & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, 38, 1-10. Mahwah, NJ: Erlbaum.
- Wampold, E. (2001). *The great psychotherapy debate*. Mahwah, NJ: Erlbaum.
- Wampold, B., Minami, T., Callen Tierney, S., Baskin T. & Bhati, K. (2005). The placebo is powerful: estimating placebo effects in medicine and psychotherapy from randomized clinical trials. *Journal of Clinical Psychology*, 61(7), 835-854.
- Weinberger, J. (1995). Common factors aren't so common: The common factors dilemma. *Clinical Psychology: Science and Practice*, 2, 45-69.

- Williams, K.E. & Chambless, D.L. (1990). The relationship between therapist characteristics and outcome of in vivo exposure treatment for agoraphobia. *Behavior Therapy*, 21 (1), 111-116.
- Yu, C. (2002). Evaluating cutoff criteria of model fit indices for latent variable models with binary and continuous outcomes. Doctoral dissertation, University of California, Los Angeles.
- Yuan, K.H. & Bentler, P.M. (2000). Three likelihood-based methods for mean and covariance structure analysis with nonnormal missing data. In M.E. Sobel & M.P. Becker (Eds.): *Sociological methodology 2000* (pp. 165-200). Washington, DC: American Sociological Association.



# VI. Anhang

## Anhang A: Expertenbefragung

### Instruktion für die Therapeuten und Itemliste

#### Instruktion

Auf den folgenden Seiten finden Sie 121 Items. Diese beinhalten Aussagen über die Therapie, den Therapeuten, den Patienten selbst und das Verhältnis zwischen beiden aus der Sicht des Patienten.

Bitte ordnen Sie jedes Item jeweils einer der unten aufgeführten acht Skalen zu indem Sie in jeder Itemzeile ein Kreuz in der der entsprechenden Skala zugeordneten Zelle machen. Jedes Item soll nur einer Skala zugeordnet werden, auch wenn verschiedene Zuordnungen denkbar wären.

Außerdem möchte ich Sie vorab noch bitten, ein paar wenige persönliche Angaben zu machen, die eventuell für die Auswertung eine Rolle spielen.

**Vielen herzlichen Dank für die Mithilfe!!!**

#### Persönliche Angaben

Geschlecht: w o m o

Alter: .....

Als PsychotherapeutIn tätig seit .....

Art der therapeutischen Ausbildung:.....

---

#### Skalenbezeichnungen und Abkürzungen:

- |   |      |
|---|------|
| 1. Glaubwürdigkeit / Nachvollziehbarkeit des Therapieangebotes: | „G“  |
| 2. Glaube/Hoffnung/Veränderungserwartung des Patienten:         | „VE“ |
| 3. Empathie/Akzeptanz/Echtheit des Therapeuten:                 | „E“  |
| 4. Offenheit des Patienten:                                     | „O“  |
| 5. Kompetenz des Therapeuten:                                   | „K“  |
| 6. Bereitschaft des Patienten zur Mitarbeit:                    | „M“  |
| 7. Emotionale Bindung zwischen Therapeut und Patient:           | „B“  |
| 8. Zusammenarbeit zwischen Therapeut und Patient:               | „Z“  |

## Anhang A: Expertenbefragung

### Itemliste

	G = Glaubwürdigkeit VE = Veränderungserwartung..... E = Empathie, Echtheit & Akzeptanz O = Offenheit	K = Kompetenz M = Mitarbeit B = Emotionale Bindung Z = Zusammenarbeit	G	V E	E	O	K	M	B	Z
1	„Mein Therapeut ist sehr einfühlsam.“									
2	„Oft sehe ich meinen Therapeuten als Gegner an.“									
3	„Die Erläuterungen des Therapeuten helfen mir, die Wirkungsweise der Therapie zu verstehen.“									
4	„Ich bin bereit, viel Anstrengung in die Therapie zu investieren.“									
5	„Ich sehe in der Therapie einen Weg, um mit meinen Schwierigkeiten fertig zu werden.“									
6	„Oft habe ich den Eindruck, dass die Therapie eigentlich nichts mit meinen Problemen zu tun hat.“									
7	„Ich kann es meinem Therapeuten sagen, wenn ich anderer Meinung bin als er.“									
8	„In den Therapiestunden haben wir auch immer mal wieder etwas zu lachen.“									
9	„Es lohnt sich, wenn man sich als Patient in der Therapie anstrengt“									
10	„Ich habe wenig Hoffnung, dass sich durch die Therapie etwas an meinen Problemen ändert.“									
11	„Ich bin bereit für eine Verbesserung der Symptome viel Mühe in Kauf zu nehmen.“									
12	„Ich habe keine Angst, in der Therapie über schwierige Themen zu sprechen.“									
13	„Der Erfolg der Therapie hängt auch davon ab, wie viel ich dafür tue.“									
14	„In den Therapiesitzungen habe ich oft das Gefühl, dass ich mich verteidigen muss.“									
15	„Meistens freue ich mich darauf, meinen Therapeuten zu sehen.“									
16	„Ich möchte mich in meiner Freizeit nicht mit den Themen aus der Therapie beschäftigen.“									
17	„Am Ende der Therapie wird es mir deutlich besser gehen.“									
18	„Mein Therapeut hilft mir dabei, mich besser zu fühlen.“									
19	„Mein Therapeut verlangt oft zuviel von mir.“									
20	„Ich bin mir unsicher, wo die Therapie hinführen soll.“									
21	„Mein Therapeut erkennt meine Bedürfnisse und geht auf sie ein.“									
22	„Von der Therapie erwarte ich mir wenig.“									
23	„Mein Therapeut würde mir nie ehrlich seine Meinung über mich sagen.“									
24	„Ich traue mich nachzufragen, wenn ich etwas nicht verstanden habe.“									
25	„Die Ziele und Inhalte der Therapie legt mein Therapeut fest.“									
26	„Nach der Therapie werde ich in meinem Alltag besser zurechtkommen.“									
27	„Mein Therapeut kann auch schwierige Zusammenhänge verständlich erklären.“									
28	„Das Auftreten meines Therapeuten ist sehr professionell.“									
29	„Mein Therapeut verfügt über sehr viel Fachwissen.“									
30	„In der Therapie sprechen wir häufig über das weitere Vorgehen.“									
31	„Das Vorgehen in der Therapie ist auf meine Schwierigkeiten abgestimmt.“									
32	„Es werden sich durch die Therapie bald Verbesserungen einstellen.“									
33	„Meinen Therapeuten würde ich jederzeit an einen guten Freund weiterempfehlen.“									
34	„Ich habe den Eindruck, dass ich in der Therapie nicht ernst genommen werde.“									
35	„Mir ist es sehr wichtig, dass mein Therapeut einen guten Eindruck von mir hat.“									

## Anhang A: Expertenbefragung

	G = Glaubwürdigkeit VE = Veränderungserwartung..... E = Empathie, Echtheit & Akzeptanz O = Offenheit	K = Kompetenz M = Mitarbeit B = Emotionale Bindung Z = Zusammenarbeit	G	V E	E	O	K	M	B	Z
36	„Es ist wichtig, dass man in einer Therapie ernsthaft an sich arbeitet.“									
37	„Die Therapie hilft mir dabei, meine Probleme zu lösen.“									
38	„Ich erhalte in der Therapie viel Unterstützung.“									
39	„Ich fühle mich sicher in der Therapie.“									
40	„Die Therapie wird zu einer Verbesserung der Symptome führen.“									
41	„Mein Therapeut interessiert sich für das, was ich ihm erzähle.“									
42	„Manchmal fühle ich mich dem Therapeuten ausgeliefert.“									
43	„Es gibt vieles, was mir wichtiger ist als die Therapie.“									
44	„Häufig wirkt mein Therapeut unsicher.“									
45	„Das Vorgehen in der Therapie ist für mich nachvollziehbar.“									
46	„Mein Leben wird sich durch die Therapie nachhaltig verändern.“									
47	„In der Therapie kann ich auch Dinge sagen, die mir peinlich sind.“									
48	„Oft fällt es mir schwer, den Aufforderungen meines Therapeuten nachzukommen.“									
49	„Es fällt mir schwer, die Therapietermine einzuhalten.“									
50	„Für meinen Therapeuten ist mein Verhalten nachvollziehbar.“									
51	„Oft scheint mir mein Therapeut gar nicht richtig zuzuhören.“									
52	„Ich verstehe mich gut mit meinem Therapeuten.“									
53	„Manchmal erscheint mir mein Therapeut abweisend.“									
54	„Mein Therapeut weiß, was er tut.“									
55	„Mein Therapeut geht planvoll und zielorientiert vor.“									
56	„Ich erzähle das, was mein Therapeut hören will.“									
57	„Mein Therapeut hilft mir dabei, Lösungen für meine Probleme zu finden.“									
58	„Es ist vor allem die Aufgabe des Therapeuten, eine Verbesserung der Symptome herbeizuführen.“									
59	„Eigentlich ist mir mein Therapeut unsympathisch.“									
60	„Meine Mitarbeit in der Therapie erscheint mir eher nebensächlich.“									
61	„Ich frage mich oft, was mein Therapeut wirklich von mir denkt.“									
62	„Ich weiß, was mich in den Therapiesitzungen erwartet.“									
63	„Es gibt einige Dinge, von denen mein Therapeut nichts wissen soll.“									
64	„Den Ausführungen des Therapeuten kann ich gut folgen.“									
65	„Ich habe Angst, dass mein Therapeut sieht, wie es wirklich in mir aussieht.“									
66	„Mein Therapeut kann sich in meine Lage versetzen.“									
67	„Ich kann den Inhalt der Therapiestunden wesentlich mitgestalten.“									
68	„Ich glaube, dass mein Therapeut ein Spezialist für meine Beschwerden ist.“									
69	„Ich spüre, dass meinem Therapeuten etwas an mir liegt.“									
70	„Ich halte meinen Therapeuten für sehr kompetent.“									
71	„Mein Therapeut meint auch das, was er sagt.“									
72	„Vor meinem Therapeuten kann ich meine Gefühle offen zeigen.“									
73	„Ich bin bei meinem Therapeuten in guten Händen.“									
74	„Manchmal überlege ich mir, zu einem anderen Therapeuten zu gehen.“									
75	„Ich habe das Gefühl, dass mein Therapeut nicht sehr erfahren ist.“									
76	„Mein Therapeut weiß viel über psychische Probleme.“									
77	„Wenn ich völlig verzweifelt wäre, würde ich mich an meinen Therapeuten wenden.“									
78	„Ich versuche mich so zu verhalten, wie es mein Therapeut von mir erwartet.“									

## Anhang A: Expertenbefragung

	G = Glaubwürdigkeit VE = Veränderungserwartung..... E = Empathie, Echtheit & Akzeptanz O = Offenheit	K = Kompetenz M = Mitarbeit B = Emotionale Bindung Z = Zusammenarbeit	G	V E	E	O	K	M	B	Z
79	„Ich werde mich am Ende der Therapie viel wohler fühlen.“									
80	„In der Therapie kann ich mich wirklich fallen lassen.“									
81	„Ich fühle mich von meinem Therapeuten als Mensch akzeptiert.“									
82	„Die Therapie wird mir nur etwas bringen, wenn ich mich selbst um Veränderung bemühe.“									
83	„Bei meinen Problemen kann mir niemand helfen.“									
84	„Mit meinem Therapeuten könnte ich auch gut in einem Team zusammenarbeiten.“									
85	„In der Therapie arbeiten wir gemeinsam an der Erreichung der Ziele.“									
86	„Mein Therapeut wirkt sehr selbstsicher.“									
87	„Ich gehe das Risiko ein, in der Therapie auch unangenehme Gefühle zu erleben.“									
88	„Oft erlebe ich die Therapie als unnötige Quälerei.“									
89	„Positive Veränderungen sind meinem Therapeuten genauso wichtig wie mir.“									
90	„Im Grunde ist es meinem Therapeuten egal, wie es mir geht.“									
91	„Es fällt mir schwer, meinem Therapeuten zu widersprechen.“									
92	„Meistens fühle ich mich von meinem Therapeuten verstanden.“									
93	„Ich fühle mich oft unwohl, wenn ich mit meinem Therapeuten zusammen bin.“									
94	„In der Therapie rede ich offen über alle Probleme.“									
95	„In den Therapiesitzungen frage ich mich oft, was das eigentlich soll.“									
96	„Wenn ich ein konkretes Anliegen habe, nimmt sich mein Therapeut Zeit dafür.“									
97	„Manchmal denke ich, dass sich bei mir nie etwas ändern wird.“									
98	„Häufig habe ich das Gefühl, dass mir mein Therapeut etwas aufzwingen will.“									
99	„Mein Therapeut weiß oft nicht, was er sagen soll.“									
100	„In der Therapie ziehen wir beide an einem Strang.“									
101	„Man muss in der Therapie auch über seinen Schatten springen.“									
102	„Häufig ist mir unklar, was der Therapeut vorhat.“									
103	„Oft fürchte ich, dass ich vor meinem Therapeuten die Fassung verliere.“									
104	„Die Ziele der Therapie habe ich gemeinsam mit meinem Therapeuten ausgearbeitet.“									
105	„Das Therapieangebot erscheint mir einleuchtend.“									
106	„Es fällt mir schwer zu erklären, was in der Therapie gemacht wird.“									
107	„Mein Therapeut hat Verständnis für meine Probleme.“									
108	„Oft fürchte ich, in der Therapie etwas Falsches zu sagen.“									
109	„Ich setze viel Hoffnung in diese Therapie.“									
110	„Häufig will mich mein Therapeut zu etwas überreden.“									
111	„Ich verstehe nicht, wie die Therapie zu einer Verbesserung meiner Symptome führen soll.“									
112	„Ich zweifle häufig daran, dass mir die Therapie etwas bringt.“									
113	„Ich habe Vertrauen in die Fähigkeiten meines Therapeuten.“									
114	„Bei der Planung der Therapie kann ich mitreden.“									
115	„Ich habe eine Vorstellung davon, wie die Therapieziele zu erreichen sind.“									
116	„Das Vorgehen in der Therapie erscheint mir sinnvoll.“									
117	„Ich kann mich auf die Therapie einlassen.“									
118	„Mein Therapeut hat mir erklärt, wie mir die Therapie helfen kann.“									
119	„Ich fürchte, dass mir die Therapie auch nicht weiterhelfen kann.“									
120	„Ich glaube, dass mein Therapeut gut ausgebildet ist.“									
121	„In der Therapie komme ich mir oft vor wie ein Schüler.“									

**Übereinstimmung der Experten bei der Zuordnung der Items zu den Skalen**

Skalenbenennung und Itemauswahl (70%-Kriterium)

I. Glaubwürdigkeit / Nachvollziehbarkeit des Therapieangebotes (G)

Item-Nr. *		% der Experten- übereinstimmung bei der Zuordnung
45.	„Das Vorgehen in der Therapie ist für mich nachvollziehbar.“	83,9
62.	„Ich weiß, was mich in den Therapiesitzungen erwartet.“	87,1
95.	„In den Therapiesitzungen frage ich mich oft, was das eigentlich soll.“	80,6
102.	„Häufig ist mir unklar, was der Therapeut vorhat.“	77,4
105.	„Das Therapieangebot erscheint mir einleuchtend.“	93,5
106.	„Es fällt mir schwer zu erklären, was in der Therapie gemacht wird.“	90,3
111.	„Ich verstehe nicht, wie die Therapie zu einer Verbesserung meiner Symptome führen soll.“	77,4
116.	„Das Vorgehen in der Therapie erscheint mir sinnvoll.“	87,1
115.	„Ich habe eine Vorstellung davon, wie die Therapieziele zu erreichen sind.“	71,0
<b>N = 9 (davon 4 negativ gepolte Items = 44%)</b>		

\* Die Item-Nummerierung bezieht sich auf die Reihenfolge der Items aus der Expertenbefragung; negativ gepolte Items sind grau unterlegt)

II. Glaube/ Hoffnung/ Veränderungserwartung des Patienten (VE)

Item-Nr. *		% der Experten- übereinstimmung bei der Zuordnung
5.	„Ich sehe in der Therapie einen Weg, um mit meinen Schwierigkeiten fertig zu werden.“	90,3
10.	„Ich habe wenig Hoffnung, dass sich durch die Therapie etwas an meinen Problemen ändert.“	83,9
17.	„Am Ende der Therapie wird es mir deutlich besser gehen.“	100
22.	„Von der Therapie erwarte ich mir wenig.“	96,8
26.	„Nach der Therapie werde ich in meinem Alltag besser zurechtkommen.“	100
32.	„Es werden sich durch die Therapie bald Verbesserungen einstellen.“	93,5
37.	„Die Therapie hilft mir dabei, meine Probleme zu lösen.“	74,2
40.	„Die Therapie wird zu einer Verbesserung der Symptome führen.“	87,1
46.	„Mein Leben wird sich durch die Therapie nachhaltig verändern.“	100
79.	„Ich werde mich am Ende der Therapie viel wohler fühlen.“	100
97.	„Manchmal denke ich, dass sich bei mir nie etwas ändern wird.“	100
109.	„Ich setze viel Hoffnung in diese Therapie.“	100
119.	„Ich fürchte, dass mir die Therapie auch nicht weiterhelfen kann.“	77,4
<b>N = 13 (davon 4 negativ gepolte Items = 31%)</b>		

Die Item-Nummerierung bezieht sich auf die Reihenfolge der Items aus der Expertenbefragung; negativ gepolte Items sind grau unterlegt)

III. Therapeutische Beziehung (TB)

Item-Nr. *		% der Experten-Übereinstimmung bei der Zuordnung
1.	„Mein Therapeut ist sehr einfühlbar.“	93,5
2.	„Oft sehe ich meinen Therapeuten als Gegner an.“	80,6
8.	„In den Therapiestunden haben wir auch immer mal wieder etwas zu lachen.“	90,3
14.	„In den Therapiesitzungen habe ich oft das Gefühl, dass ich mich verteidigen muss.“	80,6
15.	„Meistens freue ich mich darauf, meinen Therapeuten zu sehen.“	100
23.	„Mein Therapeut würde mir nie ehrlich seine Meinung über mich sagen.“	83,3
34.	„Ich habe den Eindruck, dass ich in der Therapie nicht ernst genommen werde.“	83,9
38.	„Ich erhalte in der Therapie viel Unterstützung.“	87,1
39.	„Ich fühle mich sicher in der Therapie.“	100
41.	„Mein Therapeut interessiert sich für das, was ich ihm erzähle.“	77,4
42.	„Manchmal fühle ich mich dem Therapeuten ausgeliefert.“	87,1
50.	„Für meinen Therapeuten ist mein Verhalten nachvollziehbar.“	93,5
51.	„Oft scheint mir mein Therapeut gar nicht richtig zuzuhören.“	83,3
52.	„Ich verstehe mich gut mit meinem Therapeuten.“	100
53.	„Manchmal erscheint mir mein Therapeut abweisend.“	83,9
59.	„Eigentlich ist mir mein Therapeut unsympathisch.“	100
61.	„Ich frage mich oft, was mein Therapeut wirklich von mir denkt.“	93,5
66.	„Mein Therapeut kann sich in meine Lage versetzen.“	80,0
67.	„Ich kann den Inhalt der Therapiestunden wesentlich mitgestalten.“	80,6
69.	„Ich spüre, dass meinem Therapeuten etwas an mir liegt.“	100
71.	„Mein Therapeut meint auch das, was er sagt.“	80,6
81.	„Ich fühle mich von meinem Therapeuten als Mensch akzeptiert.“	93,5
84.	„Mit meinem Therapeuten könnte ich auch gut in einem Team zusammenarbeiten.“	93,5
85.	„In der Therapie arbeiten wir gemeinsam an der Erreichung der Ziele.“	80,6
89.	„Positive Veränderungen sind meinem Therapeuten genauso wichtig wie mir.“	87,1
90.	„Im Grunde ist es meinem Therapeuten egal, wie es mir geht.“	96,8
92.	„Meistens fühle ich mich von meinem Therapeuten verstanden.“	93,1
93.	„Ich fühle mich oft unwohl, wenn ich mit meinem Therapeuten zusammen bin.“	93,5
96.	„Wenn ich ein konkretes Anliegen habe, nimmt sich mein Therapeut Zeit dafür.“	87,1
100.	„In der Therapie ziehen wir beide an einem Strang.“	90,3
104.	„Die Ziele der Therapie habe ich gemeinsam mit meinem Therapeuten ausgearbeitet.“	80,6
107.	„Mein Therapeut hat Verständnis für meine Probleme.“	93,5
110.	„Häufig will mich mein Therapeut zu etwas überreden.“	80,6
114.	„Bei der Planung der Therapie kann ich mitreden.“	90,3
121.	„In der Therapie komme ich mir oft vor wie ein Schüler.“	96,8
<b>N = 35 (davon 12 negativ gepolte Items = 34%)</b>		

Die Item-Nummerierung bezieht sich auf die Reihenfolge der Items aus der Expertenbefragung; negativ gepolte Items sind grau unterlegt)

## Anhang A: Expertenbefragung

### IV. Bereitschaft des Patienten, sich auf die Therapie einzulassen und mitzuarbeiten (MO)

Item-Nr. *		% der Experten- übereinstimmung bei der Zuordnung
4.	„Ich bin bereit, viel Anstrengung in die Therapie zu investieren.“	93,5
11.	„Ich bin bereit für eine Verbesserung der Symptome viel Mühe in Kauf zu nehmen.“	93,3
13.	„Der Erfolg der Therapie hängt auch davon ab, wie viel ich dafür tue.“	77,4
16.	„Ich möchte mich in meiner Freizeit nicht mit den Themen aus der Therapie beschäftigen.“	83,9
36.	„Es ist wichtig, dass man in einer Therapie ernsthaft an sich arbeitet.“	93,5
43.	„Es gibt vieles, was mir wichtiger ist als die Therapie.“	83,9
47.	„In der Therapie kann ich auch Dinge sagen, die mir peinlich sind.“	80,6
49.	„Es fällt mir schwer, die Therapietermine einzuhalten.“	73,3
56.	„Ich erzähle das, was mein Therapeut hören will.“	74,2
60.	„Meine Mitarbeit in der Therapie erscheint mir eher nebensächlich.“	83,9
63.	„Es gibt einige Dinge, von denen mein Therapeut nichts wissen soll.“	83,9
65.	„Ich habe Angst, dass mein Therapeut sieht, wie es wirklich in mir aussieht.“	77,4
82.	„Die Therapie wird mir nur etwas bringen, wenn ich mich selbst um Veränderung bemühe.“	87,1
87.	„Ich gehe das Risiko ein, in der Therapie auch unangenehme Gefühle zu erleben.“	83,9
94.	„In der Therapie rede ich offen über alle Probleme.“	100
101.	„Man muss in der Therapie auch über seinen Schatten springen.“	90,0
117.	„Ich kann mich auf die Therapie einlassen.“	86,7
<b>N = 17 (davon 7 negativ gepolte Items = 41%)</b>		

\* Die Item-Nummerierung bezieht sich auf die Reihenfolge der Items aus der Expertenbefragung; negativ gepolte Items sind grau unterlegt)

### V. Kompetenz des Therapeuten (K)

Item-Nr. *		% der Experten- übereinstimmung bei der Zuordnung
27.	„Mein Therapeut kann auch schwierige Zusammenhänge verständlich erklären.“	83,9
28.	„Das Auftreten meines Therapeuten ist sehr professionell.“	96,8
29.	„Mein Therapeut verfügt über sehr viel Fachwissen.“	100
44.	„Häufig wirkt mein Therapeut unsicher.“	96,8
54.	„Mein Therapeut weiß, was er tut.“	83,9
55.	„Mein Therapeut geht planvoll und zielorientiert vor.“	71,0
68.	„Ich glaube, dass mein Therapeut ein Spezialist für meine Beschwerden ist.“	90,3
70.	„Ich halte meinen Therapeuten für sehr kompetent.“	100
75.	„Ich habe das Gefühl, dass mein Therapeut nicht sehr erfahren ist.“	100
76.	„Mein Therapeut weiß viel über psychische Probleme.“	100
86.	„Mein Therapeut wirkt sehr selbstsicher.“	86,7
99.	„Mein Therapeut weiß oft nicht, was er sagen soll.“	96,7
113.	„Ich habe Vertrauen in die Fähigkeiten meines Therapeuten.“	83,9
120.	„Ich glaube, dass mein Therapeut gut ausgebildet ist.“	90,3
<b>N = 14 (davon 3 negativ gepolte Items = 21%)</b>		

\* Die Item-Nummerierung bezieht sich auf die Reihenfolge der Items aus der Expertenbefragung; negativ gepolte Items sind grau unterlegt

**Gesamtsumme: 88 Items (davon 30 negativ gepolte Items = 34%)**

## Anhang B: Fragebogenuntersuchung

### Anschreiben an die Therapeuten



EBERHARD KARLS  
UNIVERSITÄT  
TÜBINGEN



Tübingen, den 17.08.2005

Sehr geehrte Frau Kollegin, sehr geehrter Herr Kollege,

ich wende mich an Sie mit der Bitte um Unterstützung bei meiner Doktorarbeit.

Als Diplom-Psychologin und Doktorandin der Christoph-Dornier-Stiftung am Universitätsklinikum in Tübingen (wissenschaftlicher Leiter: Prof. Dr. N. Birbaumer) und der Abteilung „Klinische und Entwicklungspsychologie“ des Psychologischen Instituts (Lehrstuhl Prof. Dr. M. Hautzinger) möchte ich allgemeine Wirkfaktoren in Psychotherapien mit Hilfe eines Patienten-Fragebogens erfassen.

Mich interessieren Faktoren, die einen positiven Einfluss auf das Therapieergebnis haben, aber nicht spezifisch für eine bestimmte Therapieschule bzw. -richtung sind. Der Fragebogen soll in der Psychotherapieforschung eingesetzt werden, um die Aussagekraft von Therapievergleichsstudien zu erhöhen und die Fairness beim Vergleich zwischen verschiedenen Therapien zu gewährleisten.

Im Moment benötige ich für die erste Erhebungsphase möglichst viele Patienten aller großen Therapieschulen. Ich würde mich sehr freuen, wenn ich dabei auch Ihre Patienten einbeziehen könnte.

Nähere Informationen zum Konzept und zum Ablauf der Studie habe ich als Anlage beigelegt. Um zu klären, ob Sie zu einer Zusammenarbeit bereit wären, werde ich Sie in den nächsten Tagen anrufen. Bei dieser Gelegenheit können wir auch gern weitergehende Fragen klären.

Mit freundlichen Grüßen

Dipl.-Psych. Konstanze Vollmann



## Informationsblatt für die Therapeuten

### Fragebogen zur Erfassung des Therapieerlebens



Sehr geehrte Frau Kollegin, sehr geehrter Herr Kollege,

zunächst einmal vielen herzlichen Dank für Ihr Interesse an diesem Projekt.

#### Worum es geht:

Das Ziel meines Forschungsvorhabens ist es, einen Patientenfragebogen zu entwickeln und zu evaluieren. Dieser Fragebogen soll in der Lage sein, allgemeine Bedingungen und Variablen zu erfassen, die einen Einfluss auf den Therapieverlauf haben. Eingesetzt werden soll dieser Fragebogen in der Therapieforschung, d.h. im Rahmen von Therapievergleichsstudien und Placebo-Kontrollgruppenstudien.

#### Was der Fragebogen bringen soll:

In der Therapieforschung werden vor allem Studien durchgeführt, um die Wirksamkeit spezifischer Behandlungstechniken zu überprüfen. Verschiedenste Forschungsergebnisse belegen aber, dass für einen erfolgreichen Therapieverlauf nicht nur diese spezifischen Interventionen relevant sind. Es scheint gewisse allgemeine Bedingungen oder Faktoren zu geben, die einerseits direkt einen Einfluss auf das Therapieergebnis haben und andererseits Voraussetzungen dafür sind, dass spezifische Interventionen wirken können (z.B. Orlinsky und Howard, 1986; Schulte, 1996). Die Frage nach der Wirksamkeit spezifischer Techniken kann demnach eigentlich erst dann geklärt werden, wenn festgestellt werden kann, ob und wie diese allgemeinen Bedingungen realisiert wurden (Fuchs, 1998).

Im Rahmen meines Projektes soll also ein Instrument entwickelt werden, mit dessen Hilfe diese allgemeinen Bedingungen erfasst werden können.

#### Weshalb ich Sie um Ihre Mithilfe bitte:

Der Fragebogen mit seinen Skalen und Items wurde auf der Grundlage verschiedener Forschungsergebnisse und theoretischer Konzepte zur Wirkungsweise von Psychotherapie entwickelt. In einem nächsten Schritt muss eine sorgfältige Itemanalyse durchgeführt werden. Hierfür ist es nötig, den Fragebogen einer großen Stichprobe an Patienten vorzulegen und anhand der dadurch gewonnenen Daten die Qualität der Items zu überprüfen. Um eine genügend große Stichprobe zu bekommen, bin ich auf die Unterstützung von praktizierenden Kolleginnen und Kollegen angewiesen, und deshalb bitte ich auch Sie um Ihre Mithilfe.

#### Was Ihre Aufgabe wäre:

- Ihre Patienten über das Forschungsprojekt informieren und um ihre Teilnahme bitten (Einschlusskriterien: mind. 18 Jahre alt, Einzeltherapie, mind. 3 Sitzungen bisher).
- Die Patienten instruieren, dass sie den Fragebogen zuhause (oder nach der Therapiestunde) ausfüllen (Zeitaufwand: ca. 10-20 min.), ihn in den beiliegenden Umschlag stecken und diesen verschlossen zur nächsten Therapiestunde wieder mitbringen sollen.
- Auf den Umschlägen den Code des Patienten, die ICD-10-Diagnose(n) und die Angaben über die Sitzungen vermerken und den Bogen mit den Angaben zum Therapeuten ausfüllen.
- Die Umschläge und den Bogen in den beiliegenden frankierten und adressierten Umschlag stecken und abschicken.

Leider ist es mir nicht möglich, die Teilnahme der Patienten zu vergüten. Ich hoffe aber, im Rahmen dieses Projektes ein nützliches Instrument zu entwickeln, von dessen Einsatz langfristig auch die Patienten profitieren können. Unter den teilnehmenden Therapeuten werden als kleines Dankeschön Büchergutscheine verlost.

Ich wäre Ihnen sehr dankbar, wenn Sie sich zur Unterstützung dieses Projektes bereit erklären würden!

Dipl.-Psych. Konstanze Vollmann

## 1. Fragebogenversion

### Fragebogen zur Erfassung des Therapieerlebens Code-Nr. ....

Im folgenden Fragebogen geht es um Ihre persönlichen Erfahrungen in der Therapie. Uns interessiert zum einen, wie Sie die letzte (gemeint ist die am kürzesten zurückliegende) Therapiestunde erlebt haben. Zum anderen möchten wir erfahren, wie Sie in der Woche vor dieser letzten Sitzung über die Therapie gedacht haben.

Dazu werden Ihnen im *ersten Abschnitt* Aussagen vorgegeben, die beschreiben, wie manche Patienten die Therapiestunden erleben. Bitte schätzen Sie für jede Aussage ein, wie gut sie auf Ihre *letzte Therapiestunde* zutrifft.

Im *zweiten Abschnitt* werden Ihnen Aussagen vorgegeben, die ausdrücken, wie manche Patienten über die Therapie denken. Überlegen Sie bei jeder Aussage, inwieweit diese auf Sie zutrifft, und kreuzen Sie die Antwortmöglichkeit an, die am besten passt. Beziehen Sie sich dabei bitte auf *die Woche vor der letzten Therapiestunde*.

Bitte versuchen Sie, möglichst spontan und aufrichtig zu antworten. Es gibt keine richtigen oder falschen Antworten. Uns interessiert, was für Sie persönlich zutrifft. Lassen Sie bitte keine Antworten aus.

Aus Gründen der Vereinfachung wird im Fragebogen nur die männliche Form „der Therapeut“ verwendet, anstatt immer beide Formen „der Therapeut/die Therapeutin“ aufzuführen.

Diese Untersuchung dient allein wissenschaftlichen Forschungszwecken und wird anonym durchgeführt und ausgewertet. Das heißt, dass keine Rückschlüsse auf Ihre Person möglich sind, und auch, dass Ihr Therapeut den Fragebogen nicht zu sehen bekommt.

Geben Sie bitte Ihren Namen nicht an, sondern machen Sie lediglich folgende Angaben zu Ihrer Person:

Alter: .....

Geschlecht:  m.  w.

Schulabschluss: kein Schulabschluss .....

Hauptschule.....

Realschule.....

Abitur.....

Abgeschlossenes Studium (Universität; FH).....

Partnerschaft: verheiratet/feste Partnerschaft.....

keine feste Partnerschaft.....

Wie lange leiden Sie unter der Problematik, derentwegen sie die Therapie begonnen haben?  
.....

Wie stark fühlen Sie sich zurzeit dadurch belastet?

sehr stark.....

stark.....

mittelmäßig.....

wenig.....

gar nicht.....

Waren Sie früher schon in psychologischer Behandlung?

ja  nein

Wie lange liegt die letzte Therapiestunde zurück ?

.....

Bitte wenden



Fragebogen zur Erfassung des Therapieerlebens		Code-Nr. ....
32.	hatte ich den Eindruck, dass mein Therapeut ein Spezialist für meine Beschwerden ist.	① ② ③ ④ ⑤ ⑥ ⑦
33.	habe ich gespürt, dass mein Therapeut mich schätzt.	① ② ③ ④ ⑤ ⑥ ⑦
34.	hatte ich Vertrauen in die Fähigkeit meines Therapeuten, mir zu helfen.	① ② ③ ④ ⑤ ⑥ ⑦
35.	hatte ich das Gefühl, dass mein Therapeut mir seine ehrliche Meinung sagt.	① ② ③ ④ ⑤ ⑥ ⑦
36.	hatte ich Angst, dass mein Therapeut sehen könnte, wie es wirklich in mir aussieht.	① ② ③ ④ ⑤ ⑥ ⑦
37.	habe ich gemerkt, dass mein Therapeut viel über psychische Probleme weiß.	① ② ③ ④ ⑤ ⑥ ⑦
38.	fühlte ich mich von meinem Therapeuten als Mensch akzeptiert.	① ② ③ ④ ⑤ ⑥ ⑦
39.	habe ich gespürt, dass ich meinem Therapeuten vertrauen kann.	① ② ③ ④ ⑤ ⑥ ⑦
40.	haben wir gut zusammengearbeitet.	① ② ③ ④ ⑤ ⑥ ⑦
41.	wirkte mein Therapeut selbstsicher.	① ② ③ ④ ⑤ ⑥ ⑦
42.	bin ich das Risiko eingegangen, auch unangenehme Gefühle zu erleben.	① ② ③ ④ ⑤ ⑥ ⑦
43.	habe ich gemerkt, dass meinem Therapeuten positive Veränderungen genauso wichtig sind wie mir.	① ② ③ ④ ⑤ ⑥ ⑦
44.	hatte ich den Eindruck, dass mein Therapeut nicht wahrnimmt, wie ich mich fühle.	① ② ③ ④ ⑤ ⑥ ⑦
45.	war mir unklar, was der Therapeut vorhatte.	① ② ③ ④ ⑤ ⑥ ⑦
46.	habe ich mich unwohl gefühlt.	① ② ③ ④ ⑤ ⑥ ⑦
47.	habe ich nicht verstanden, wie die Therapie zu einer Verbesserung meiner Symptome führen soll.	① ② ③ ④ ⑤ ⑥ ⑦
48.	habe ich mich gefragt, was das bringen soll.	① ② ③ ④ ⑤ ⑥ ⑦
49.	hatte ich das Gefühl, dass ich mich verteidigen muss.	① ② ③ ④ ⑤ ⑥ ⑦
50.	musste mein Therapeut lange überlegen, bevor er mir eine Antwort geben konnte.	① ② ③ ④ ⑤ ⑥ ⑦
51.	hatte ich das Gefühl, dass wir beide an einem Strang ziehen.	① ② ③ ④ ⑤ ⑥ ⑦
52.	konnte ich über meinen Schatten springen.	① ② ③ ④ ⑤ ⑥ ⑦
53.	bin ich mir vorgekommen wie ein Schüler.	① ② ③ ④ ⑤ ⑥ ⑦
54.	wirkte mein Therapeut unerfahren.	① ② ③ ④ ⑤ ⑥ ⑦
55.	stimmten wir darin überein, an welchen Problemen wir arbeiten müssen.	① ② ③ ④ ⑤ ⑥ ⑦
56.	erschien mir das Therapieangebot einleuchtend.	① ② ③ ④ ⑤ ⑥ ⑦
57.	hatte ich das Gefühl, dass ich mit meinem Therapeuten auch gut in einem Team zusammenarbeiten könnte.	① ② ③ ④ ⑤ ⑥ ⑦
58.	wollte mich mein Therapeut von etwas überzeugen.	① ② ③ ④ ⑤ ⑥ ⑦
59.	habe ich offen über alle Probleme geredet.	① ② ③ ④ ⑤ ⑥ ⑦
60.	habe ich viel Unterstützung erhalten.	① ② ③ ④ ⑤ ⑥ ⑦
61.	hat sich mein Therapeut sehr einfühlsam verhalten.	① ② ③ ④ ⑤ ⑥ ⑦
62.	erschien mir das Vorgehen angemessen.	① ② ③ ④ ⑤ ⑥ ⑦
63.	konnte ich mich auf die Therapie einlassen.	① ② ③ ④ ⑤ ⑥ ⑦

Bitte wenden →

## Anhang B: Fragebogenuntersuchung

### Fragebogen zur Erfassung des Therapieerlebens Code-Nr. ....

Die Zahlen bedeuten folgendes:	① = „stimmt überhaupt nicht“	④ = „stimmt teils-teils“	⑦ = „stimmt eher“
	② = „stimmt nicht“		⑥ = „stimmt“
	③ = „stimmt eher nicht“		⑤ = „stimmt voll und ganz“

64. habe ich gemerkt, dass mein Therapeut gut ausgebildet ist.	① ② ③ ④ ⑤ ⑥ ⑦
65. hatte ich das Gefühl, dass mein Therapeut mich versteht.	① ② ③ ④ ⑤ ⑥ ⑦
66. Mit der letzten Therapiestunde war ich zufrieden.	① ② ③ ④ ⑤ ⑥ ⑦
67. Die letzte Therapiestunde war irgendwie anders als die bisherigen Stunden.	① ② ③ ④ ⑤ ⑥ ⑦

#### II. Abschnitt

In der Woche vor der letzten Therapiesitzung...

68. habe ich in der Therapie einen Weg gesehen, um mit meinen Schwierigkeiten fertig zu werden.	① ② ③ ④ ⑤ ⑥ ⑦
69. habe ich mich in der Lage gefühlt, aktiv etwas gegen meine Beschwerden zu tun.	① ② ③ ④ ⑤ ⑥ ⑦
70. habe ich mich darauf gefreut, meinen Therapeuten wiederzusehen.	① ② ③ ④ ⑤ ⑥ ⑦
71. war ich überzeugt, dass es mir am Ende der Therapie deutlich besser gehen wird.	① ② ③ ④ ⑤ ⑥ ⑦
72. habe ich mich häufig mit den Themen aus der Therapie beschäftigt.	① ② ③ ④ ⑤ ⑥ ⑦
73. habe ich mir gedacht, dass ich am Ende der Therapie mit meinem Alltag besser zurechtkommen werde.	① ② ③ ④ ⑤ ⑥ ⑦
74. war ich überzeugt, dass sich durch die Therapie bald Verbesserungen einstellen werden.	① ② ③ ④ ⑤ ⑥ ⑦
75. hatte ich Gefühl, dass die Therapie mir helfen wird.	① ② ③ ④ ⑤ ⑥ ⑦
76. habe ich geglaubt, dass die Therapie zu einer Verbesserung der Symptome führen wird.	① ② ③ ④ ⑤ ⑥ ⑦
77. habe ich gemerkt, dass mir vieles wichtiger ist als die Therapie.	① ② ③ ④ ⑤ ⑥ ⑦
78. hatte ich das Gefühl, dass die Therapie mir wenig bringen wird.	① ② ③ ④ ⑤ ⑥ ⑦
79. fiel es mir schwer, den nächsten Therapietermin einzuhalten.	① ② ③ ④ ⑤ ⑥ ⑦
80. habe ich geglaubt, dass sich bei mir nie etwas ändern wird.	① ② ③ ④ ⑤ ⑥ ⑦
81. habe ich gemerkt, dass ich mich sehr bemühen muss, um eine Veränderung zu erreichen.	① ② ③ ④ ⑤ ⑥ ⑦
82. habe ich gedacht, dass ich mich am Ende der Therapie viel wohler fühlen werde.	① ② ③ ④ ⑤ ⑥ ⑦
83. wäre es mir schwer gefallen zu erklären, was in den Sitzungen gemacht wird.	① ② ③ ④ ⑤ ⑥ ⑦
84. hatte ich Hoffnung auf Verbesserung.	① ② ③ ④ ⑤ ⑥ ⑦
85. habe ich befürchtet, dass mir die Therapie auch nicht weiterhelfen kann.	① ② ③ ④ ⑤ ⑥ ⑦
86. hatte ich eine Vorstellung davon, wie ich die Therapieziele erreichen kann.	① ② ③ ④ ⑤ ⑥ ⑦
87. habe ich gedacht, dass sich durch die Therapie wenig an meinen Problemen ändern wird.	① ② ③ ④ ⑤ ⑥ ⑦
88. habe ich mir gedacht, dass ich für eine Verbesserung noch viel tun muss.	① ② ③ ④ ⑤ ⑥ ⑦
89. wusste ich, was mich in der nächsten Therapiesitzung erwartet.	① ② ③ ④ ⑤ ⑥ ⑦
90. war ich überzeugt, dass sich mein Leben durch die Therapie nachhaltig verändern wird.	① ② ③ ④ ⑤ ⑥ ⑦

Herzlichen Dank für Ihre Mitarbeit!

## Angaben des Therapeuten

### Fragebogen zur Erfassung des Therapieerlebens

#### Angaben des Therapeuten/der Therapeutin

Sehr geehrte Frau Kollegin, sehr geehrter Herr Kollege,

zunächst einmal vielen herzlichen Dank für Ihre Bereitschaft, dieses Projekt zu unterstützen.

Für die Untersuchung kommen Patienten infrage, die mind. 18 Jahre alt sind, sich in Einzel-Psychotherapie befinden und bisher mindestens drei Sitzungen bei Ihnen hatten (incl. probatorische Sitzungen).

Bitte informieren Sie Ihre Patienten, dass es sich um eine Datenerhebung zu Forschungszwecken handelt und die Daten in anonymer Form erhoben werden. Wenn die Patienten zur Teilnahme bereit sind, sollen sie den Bogen zuhause ausfüllen, ihn in den beiliegenden Umschlag stecken und diesen verschlossen zur nächsten Therapiesitzung wieder mitbringen. Bitte weisen Sie darauf hin, dass kein Name auf den Fragebogen geschrieben werden soll.

Bitte vermerken Sie selbst dann auf den Umschlägen mit den ausgefüllten Fragebögen noch den Code des Patienten, Ihre eigene Code-Nummer (siehe unten) und die Diagnose(n) des Patienten (nach ICD-10), sowie die Angaben über die Sitzungen (Frequenz, bisherige Anzahl, geplante Anzahl).

Für die Rücksendung der Umschläge mit den Fragebögen und dieses Bogens liegt ein weiterer frankierter und adressierter Umschlag bei.

Es wäre nett, wenn Sie folgende, für die Auswertung relevante Angaben machen, und diesen Bogen zusammen mit dem Fragebogen zurücksenden könnten:

#### **Code-Nr. Therapeut .....**

Tragen Sie bitte hier, ebenso wie auf den Umschlägen mit den Fragebögen, Ihren persönlichen Code ein, bestehend aus:

Anfangsbuchstabe des Vornamens Ihrer Mutter und deren Geburtstag (Bsp.: K120843)

#### **Angaben zu Ihrer Person:**

Alter: ..... Geschlecht:  m.  w.

Berufsausbildung: .....

Therapieausbildung: .....

Berufserfahrung (in Jahren): .....

#### **Angaben zur Therapie:**

Die Therapie wird durchgeführt in....

einer Ambulanz der Christoph-Dornier-Stiftung.....

einer Hochschulambulanz.....

der Ambulanz eines Ausbildungsinstitutes.....

einer Psychotherapie-Praxis.....

.....

**Vielen herzlichen Dank!**



### Test der Items auf Normalverteilung

	Kolmogorov-Smirnov(a)		
	Statistic	df	Sig.
Item 19, Skala I	,328	214	,000
Item 45, Skala I	,248	212	,000
Item 47, Skala I	,264	215	,000
Item48, Skala I	,251	216	,000
Item 56, Skala I	,285	212	,000
Item 62, Skala I	,309	215	,000
Item 83, Skala I	,252	214	,000
Item 86, Skala I	,146	214	,000
Item 89, Skala I	,158	215	,000
Item 68, Skala II	,257	215	,000
Item 71, Skala II	,242	216	,000
Item 73, Skala II	,247	215	,000
Item 74, Skala II	,203	216	,000
Item 75, Skala II	,264	216	,000
Item 76, Skala II	,272	216	,000
Item 78, Skala II	,226	213	,000
Item 80, Skala II	,204	214	,000
Item 82, Skala II	,223	215	,000
Item 84, Skala II	,268	213	,000
Item 85, Skala II	,223	215	,000
Item 87, Skala II	,248	214	,000
Item 90, Skala II	,198	214	,000
Item 1, Skala III	,231	213	,000
Item 4, Skala III	,301	215	,000
Item 5, Skala III	,308	214	,000
Item 6, Skala III	,403	216	,000
Item 8, Skala III	,417	216	,000
Item 10, Skala III	,425	216	,000
Item12, Skala III	,297	214	,000
Item 13, Skala III	,213	215	,000
Item 15, Skala III	,327	216	,000
Item 16, Skala III	,415	216	,000
Item17, Skala III	,248	216	,000



## Anhang B: Fragebogenuntersuchung

Item 18, Skala III	,429	215	,000
Item 21, Skala III	,305	215	,000
Item 27, Skala III	,265	216	,000
Item 29, Skala III	,285	215	,000
Item 30, Skala III	,295	216	,000
Item 31, Skala III	,311	212	,000
Item 33, Skala III	,292	213	,000
Item 35, Skala III	,286	215	,000
Item 38, Skala III	,309	214	,000
Item 39, Skala III	,292	216	,000
Item 40, Skala III	,283	214	,000
Item 43, Skala III	,255	215	,000
Item 44, Skala III	,281	216	,000
Item 46, Skala III	,251	214	,000
Item 49, Skala III	,286	216	,000
Item 51, Skala III	,303	214	,000
Item 53, Skala III	,256	214	,000
Item 55, Skala III	,318	213	,000
Item 57, Skala III	,285	215	,000
Item 58, Skala III	,233	215	,000
Item 60, Skala III	,292	216	,000
Item 61, Skala III	,284	215	,000
Item 65, Skala III	,309	216	,000
Item 70, Skala III	,288	216	,000
Item 7, Skala IV	,131	209	,000
Item 11, Skala IV	,275	214	,000
Item 20, Skala IV	,284	215	,000
Item 22, Skala IV	,244	214	,000
Item 24, Skala IV	,304	216	,000
Item 26, Skala IV	,267	216	,000
Item 36, Skala IV	,260	216	,000
Item 42, Skala IV	,287	213	,000
Item 52, Skala IV	,232	212	,000
Item 59, Skala IV	,278	215	,000
Item 63, Skala IV	,291	215	,000
Item 69, Skala IV	,234	215	,000

## Anhang B: Fragebogenuntersuchung

Item 72, Skala IV	,253	214	,000
Item 77, Skala IV	,226	216	,000
Item 79, Skala IV	,346	216	,000
Item 81, Skala IV	,241	215	,000
Item 88, Skala IV	,224	215	,000
Item 2, Skala V	,283	215	,000
Item 3, Skala V	,325	213	,000
Item 9, Skala V	,306	214	,000
Item 14, Skala V	,380	216	,000
Item 23, Skala V	,266	213	,000
Item 25, Skala V	,322	210	,000
Item 28, Skala V	,261	216	,000
Item 32, Skala V	,224	213	,000
Item 34, Skala V	,290	216	,000
Item 37, Skala V	,324	215	,000
Item 41, Skala V	,299	215	,000
Item 50, Skala V	,297	216	,000
Item 54, Skala V	,339	215	,000
Item 64, Skala V	,292	215	,000
Item 66, Zufriedenheit	,267	216	,000
Item 67, Stunde anders	,235	216	,000

a Lilliefors Significance Correction

**Ergebnis der Explorativen Faktorenanalyse (EFA)**

Faktorlösung mit 7-Faktoren

	ziel	pas	mit	bez	komp	offen	ver	h <sup>2</sup>	a <sup>2</sup> /h <sup>2</sup>
ITEM45	0,401	0,354	-0,058	0,030	-0,051	-0,133	0,090	0,319	0,50
ITEM47	0,481	0,537	0,027	-0,283	-0,043	0,100	0,179	0,644	0,45
ITEM48	0,311	0,565	0,145	-0,153	0,009	-0,062	0,326	0,571	0,56
ITEM83	0,304	0,354	-0,132	-0,065	0,085	0,140	-0,153	0,290	0,43
ITEM89	0,550	0,083	-0,125	-0,027	0,020	-0,143	-0,137	0,365	0,83
ITEM01	0,506	-0,179	0,022	0,112	-0,040	-0,129	0,037	0,321	0,80
ITEM12	-0,104	0,696	0,211	-0,058	0,031	0,016	-0,074	0,550	0,88
ITEM13	0,108	0,532	-0,180	0,012	0,022	-0,007	-0,062	0,332	0,85
ITEM15	0,198	0,592	-0,122	0,271	-0,093	-0,170	-0,086	0,523	0,67
ITEM17	-0,176	0,460	-0,132	0,081	0,005	0,232	-0,037	0,322	0,66
ITEM46	0,170	0,553	-0,224	0,208	-0,044	-0,145	-0,001	0,451	0,68
ITEM49	-0,026	0,734	0,132	0,081	-0,060	-0,022	0,068	0,572	0,94
ITEM53	-0,037	0,644	-0,124	0,244	-0,088	-0,007	-0,111	0,511	0,81
ITEM58	-0,075	0,467	-0,012	0,158	-0,130	0,089	-0,176	0,305	0,72
ITEM07	-0,263	-0,026	0,078	0,031	-0,078	0,141	0,049	0,105	0,66
ITEM20	0,024	-0,048	0,080	-0,131	0,135	0,742	0,079	0,601	0,92
ITEM22	-0,093	0,670	0,106	0,009	0,016	0,026	-0,032	0,471	0,95
ITEM26	-0,169	0,460	0,034	-0,025	0,212	0,218	-0,009	0,334	0,63
ITEM36	-0,028	0,621	-0,114	0,013	0,009	0,174	-0,100	0,440	0,88
ITEM63	0,177	0,139	-0,063	0,388	-0,141	0,301	0,186	0,350	0,43
ITEM77	0,145	0,383	0,171	-0,011	-0,088	0,007	0,055	0,208	0,71
ITEMR19	0,388	0,090	0,183	-0,093	-0,086	0,117	0,228	0,274	0,55
ITEMR56	0,643	0,082	0,051	-0,026	0,014	0,280	0,124	0,517	0,80
ITEMR62	0,465	0,091	0,035	0,468	-0,045	0,078	0,016	0,453	0,48
ITEMR86	0,117	0,172	-0,120	-0,264	0,110	0,169	0,290	0,252	0,33
ITEMR68	0,130	-0,048	-0,063	-0,069	0,003	0,184	0,644	0,477	0,87
ITEMR71	-0,007	0,006	-0,028	0,264	-0,102	-0,037	0,800	0,722	0,89
ITEMR73	-0,095	-0,252	0,019	0,286	-0,037	0,087	0,820	0,836	0,80
ITEMR74	-0,070	-0,077	-0,152	0,219	0,043	0,015	0,774	0,683	0,88
ITEMR75	0,041	-0,043	0,028	0,194	0,071	-0,010	0,813	0,708	0,93
ITEMR76	-0,036	-0,086	0,006	0,087	0,152	0,048	0,803	0,687	0,94
ITEMR78	0,047	0,380	0,047	0,099	0,014	-0,271	0,585	0,574	0,60
ITEMR80	-0,048	0,457	-0,181	-0,231	-0,010	-0,028	0,525	0,574	0,48
ITEMR82	0,022	0,027	0,078	0,138	0,000	-0,031	0,634	0,429	0,94
ITEMR84	0,052	0,023	0,114	0,043	0,044	0,234	0,540	0,366	0,80
ITEMR85	0,140	0,466	-0,234	-0,118	0,073	-0,086	0,420	0,495	0,44
ITEMR87	0,060	0,445	-0,045	-0,171	0,020	-0,077	0,534	0,524	0,54
ITEMR90	-0,032	0,091	-0,010	-0,004	0,095	0,073	0,627	0,417	0,94
ITEMR04	0,071	0,074	-0,173	0,168	0,188	-0,089	0,079	0,118	0,30
ITEMR05	0,153	0,044	-0,010	0,191	0,000	-0,043	0,007	0,064	0,57
ITEMR21	-0,004	0,088	-0,160	0,397	0,271	0,196	0,035	0,304	0,52
ITEMR27	0,090	0,395	-0,141	0,204	0,142	0,067	0,067	0,255	0,61
ITEMR29	0,074	0,086	-0,057	0,707	0,063	-0,096	0,110	0,541	0,92
ITEMR30	-0,075	0,141	-0,103	0,376	0,282	0,111	0,105	0,280	0,50
ITEMR31	0,485	0,038	-0,040	0,297	0,144	-0,038	0,098	0,358	0,66
ITEMR33	-0,071	0,050	-0,068	0,664	0,081	0,023	0,096	0,469	0,94
ITEMR35	-0,115	0,225	0,228	0,497	0,190	-0,087	0,093	0,415	0,59
ITEMR38	-0,030	0,237	0,187	0,732	0,108	-0,131	0,013	0,657	0,82
ITEMR39	-0,138	0,233	0,123	0,779	0,077	-0,195	0,067	0,744	0,82
ITEMR40	0,222	0,049	-0,081	0,736	-0,045	-0,037	0,021	0,604	0,90
ITEMR43	0,133	0,036	0,113	0,421	0,014	0,052	0,148	0,234	0,76
ITEMR44	0,000	0,382	-0,069	0,449	-0,056	0,156	0,008	0,380	0,53
ITEMR51	0,273	0,287	0,068	0,295	-0,131	0,152	0,089	0,214	0,41
ITEMR55	0,554	0,063	0,077	0,127	-0,038	0,251	-0,025	0,398	0,77
ITEMR57	0,209	0,147	-0,002	0,348	0,100	0,187	-0,173	0,261	0,46
ITEMR60	0,287	-0,111	-0,030	0,509	-0,025	0,162	0,204	0,423	0,61
ITEMR61	0,191	-0,125	-0,118	0,741	-0,035	0,158	-0,045	0,643	0,85
ITEMR65	-0,130	0,080	-0,126	0,611	0,250	0,092	0,037	0,485	0,77
ITEMR70	0,008	0,128	0,027	0,496	-0,082	-0,091	0,370	0,415	0,59
ITEMR11	0,280	-0,092	0,513	-0,118	0,117	0,225	-0,012	0,428	0,61
ITEMR24	-0,064	0,393	0,254	0,244	-0,258	0,246	0,120	0,424	0,36
ITEMR42	0,087	0,017	0,049	0,075	-0,013	0,673	-0,021	0,469	0,96
ITEMR52	-0,015	0,111	-0,002	0,010	0,058	0,597	0,110	0,385	0,93
ITEMR59	-0,110	0,339	-0,085	0,185	-0,010	0,398	-0,042	0,329	0,48
ITEMR69	0,118	0,055	-0,189	-0,025	-0,056	0,187	0,473	0,315	0,71
ITEMR72	0,113	-0,003	0,111	0,152	-0,054	0,207	0,260	0,162	0,42
ITEMR81	-0,092	-0,084	0,805	0,027	0,041	0,092	-0,085	0,682	0,95
ITEMR88	0,049	0,062	0,718	0,010	0,049	-0,092	-0,011	0,533	0,97
ITEMR02	0,486	0,034	0,131	0,159	0,275	0,086	-0,075	0,368	0,64
ITEMR03	0,555	-0,179	-0,008	0,192	0,166	0,145	-0,052	0,428	0,72
ITEMR09	0,296	-0,097	0,048	0,186	0,587	0,102	-0,038	0,490	0,70
ITEMR23	0,566	-0,148	-0,089	0,095	0,126	-0,076	0,148	0,403	0,80
ITEMR25	0,442	0,003	0,025	0,120	0,323	0,023	0,001	0,315	0,62
ITEMR28	0,061	-0,091	0,065	0,162	0,686	0,052	0,122	0,531	0,89
ITEMR32	0,077	-0,011	0,099	0,254	0,487	-0,049	0,173	0,350	0,68
ITEMR34	0,114	0,091	0,115	0,248	0,358	-0,040	0,301	0,316	0,41
ITEMR37	0,192	-0,013	0,064	0,176	0,554	0,126	0,082	0,402	0,76
ITEMR41	0,213	0,215	0,257	0,404	0,100	0,041	-0,161	0,358	0,46
ITEMR50	0,155	0,600	0,179	0,111	0,080	-0,114	-0,050	0,450	0,80
ITEMR64	0,276	-0,061	-0,027	0,290	0,576	-0,019	0,022	0,497	0,67
ITEMR64	0,234	-0,081	-0,060	0,251	0,613	0,004	0,027	0,504	0,745

## Fragebogen zur Erfassung relevanter Therapiebedingungen (FERT)

Datum: .....

Code-Nr.: .....

### FERT

Auf der folgenden Seite finden Sie eine Reihe von Aussagen, die sich auf die Therapie beziehen. Im *ersten Abschnitt* geht es um Einschätzungen der letzten, d.h. der am kürzesten zurückliegenden Therapiesitzung. Die Aussagen im *zweiten Abschnitt* beziehen sich auf die Woche vor der letzten Therapiesitzung.

Wählen Sie für jede Aussage diejenige Antwort, die für Sie am besten zutrifft. Bitte versuchen Sie, möglichst spontan und aufrichtig zu antworten. Es gibt keine richtigen oder falschen Antworten. Uns interessiert Ihre persönliche Meinung.

Lesen Sie bitte alle Sätze durch und lassen Sie keine Antwort aus. Wenn Sie sich unsicher sind, wählen Sie die am ehesten zutreffende Antwort.

Aus Gründen der Vereinfachung wird im Fragebogen nur die männliche Form „der Therapeut“ verwendet, anstatt immer beide Formen „der Therapeut/die Therapeutin“ aufzuführen.

Diese Untersuchung dient allein wissenschaftlichen Forschungszwecken und wird anonym durchgeführt und ausgewertet.

Geben Sie bitte Ihren Namen nicht an, sondern machen Sie lediglich folgende **Angaben zu Ihrer Person:**

**Alter:** ..... **Geschlecht:**  m.  w.

**Schulabschluss:** kein Schulabschluss .....   
Hauptschule.....   
Realschule/mittlere Reife.....   
Abitur.....   
Abgeschlossenes Studium (Universität; FH).....

**Partnerschaft:** verheiratet o. feste Partnerschaft.....   
keine feste Partnerschaft .....

**Wie lange leiden Sie unter der Problematik, derentwegen sie die Therapie begonnen haben (Angabe in Monaten/Jahren)?**

.....

**Wie stark fühlen Sie sich zurzeit dadurch belastet?**

sehr stark.....   
stark.....   
mittelmäßig.....   
wenig.....   
gar nicht.....

Datum: .....

Code-Nr.: .....

**FERT**

Die Zahlen bedeuten folgendes:

- ① = „stimmt überhaupt nicht“    ② = „stimmt nicht“    ③ = „stimmt eher nicht“    ④ = „stimmt teils-teils“  
 ⑤ = „stimmt eher“    ⑥ = „stimmt“    ⑦ = „stimmt voll und ganz“

**1. Abschnitt**

In der letzten Therapiesitzung...

1. wirkte das Auftreten meines Therapeuten professionell.	①	②	③	④	⑤	⑥	⑦
2. hat mein Therapeut schwierige Zusammenhänge verständlich erklärt.	①	②	③	④	⑤	⑥	⑦
3. hat mich das Verhalten meines Therapeuten irritiert.	①	②	③	④	⑤	⑥	⑦
4. habe ich mich getraut, Dinge zu sagen, die mir peinlich sind.	①	②	③	④	⑤	⑥	⑦
5. ging mein Therapeut planvoll und zielorientiert vor.	①	②	③	④	⑤	⑥	⑦
6. wirkte mein Therapeut überzeugt von seinem Vorgehen.	①	②	③	④	⑤	⑥	⑦
7. habe ich mich sicher gefühlt.	①	②	③	④	⑤	⑥	⑦
8. habe ich gemerkt, dass mein Therapeut viel Erfahrung hat.	①	②	③	④	⑤	⑥	⑦
9. habe ich mich gut mit meinem Therapeuten verstanden.	①	②	③	④	⑤	⑥	⑦
10. hatte ich den Eindruck, dass mein Therapeut ein Spezialist für meine Beschwerden ist.	①	②	③	④	⑤	⑥	⑦
11. habe ich gespürt, dass mein Therapeut mich schätzt.	①	②	③	④	⑤	⑥	⑦
12. hatte ich das Gefühl, dass mein Therapeut mir seine ehrliche Meinung sagt.	①	②	③	④	⑤	⑥	⑦
13. habe ich gemerkt, dass mein Therapeut viel über psychische Probleme weiß.	①	②	③	④	⑤	⑥	⑦
14. fühlte ich mich von meinem Therapeuten als Mensch akzeptiert.	①	②	③	④	⑤	⑥	⑦
15. habe ich gespürt, dass ich meinem Therapeuten vertrauen kann.	①	②	③	④	⑤	⑥	⑦
16. haben wir gut zusammengearbeitet.	①	②	③	④	⑤	⑥	⑦
17. bin ich das Risiko eingegangen, auch unangenehme Gefühle zu erleben.	①	②	③	④	⑤	⑥	⑦
18. habe ich gemerkt, dass meinem Therapeut positive Veränderungen genauso wichtig sind wie mir.	①	②	③	④	⑤	⑥	⑦
19. habe ich mich unwohl gefühlt.	①	②	③	④	⑤	⑥	⑦
20. hatte ich das Gefühl, dass ich mich verteidigen muss.	①	②	③	④	⑤	⑥	⑦
21. konnte ich über meinen Schatten springen.	①	②	③	④	⑤	⑥	⑦
22. hat sich mein Therapeut sehr einfühlsam verhalten.	①	②	③	④	⑤	⑥	⑦
23. habe ich gemerkt, dass mein Therapeut gut ausgebildet ist.	①	②	③	④	⑤	⑥	⑦

Datum: .....

Code-Nr.: .....

**FERT**

Die Zahlen bedeuten folgendes:

- ① = „stimmt überhaupt nicht“    ② = „stimmt nicht“    ③ = „stimmt eher nicht“    ④ = „stimmt teils-teils“  
⑤ = „stimmt eher“    ⑥ = „stimmt“    ⑦ = „stimmt voll und ganz“

**2. Abschnitt**

In der Woche vor der letzten Therapiesitzung....

24. habe ich in der Therapie einen Weg gesehen, um mit meinen Schwierigkeiten fertig zu werden.	①	②	③	④	⑤	⑥	⑦
25. war ich überzeugt, dass es mir am Ende der Therapie deutlich besser gehen wird.	①	②	③	④	⑤	⑥	⑦
26. habe ich mir gedacht, dass ich am Ende der Therapie mit meinem Alltag besser zurecht kommen werde.	①	②	③	④	⑤	⑥	⑦
27. war ich überzeugt, dass sich durch die Therapie bald Verbesserungen einstellen werden.	①	②	③	④	⑤	⑥	⑦
28. hatte ich das Gefühl, dass die Therapie mir helfen wird.	①	②	③	④	⑤	⑥	⑦
29. habe ich geglaubt, dass die Therapie zu einer Verbesserung der Symptome führen wird.	①	②	③	④	⑤	⑥	⑦
30. habe ich gedacht, dass ich mich am Ende der Therapie viel wohler fühlen werde.	①	②	③	④	⑤	⑥	⑦
31. war ich überzeugt, dass sich mein Leben durch die Therapie nachhaltig verändern wird.	①	②	③	④	⑤	⑥	⑦

Vielen Dank!